The Box–Cox transformation technique: a review

R. M. SAKIA

Sokoine University of Agriculture, Department of Crop Science and Production, Box 3005, Morogoro, Tanzania

Abstract. Box & Cox (1964) proposed a parametric power transformation technique in order to reduce anomalies such as non-additivity, non-normality and heteroscedasticity. Although the transformation has been extensively studied, no bibliography of the published research exists at present. An attempt is made here to review the work relating to this transformation.

1 Introduction

Many important results in statistical analysis follow from the assumption that the population being sampled or investigated is normally distributed with a common variance and additive errror structure. When the relevant theoretical assumptions relating to a selected method of analysis are approximately satisfied, the usual procedures can be applied in order to make inferences about unknown parameters of interest. In situations where the assumptions are seriously violated several options are available (see Graybill, 1976, p. 213).

(i) Ignore the violation of the assumptions and proceed with the analysis as if all assumptions are satisfied.
(ii) Decide what is the correct assumption in place of the one that is violated and use a valid procedure that takes into account the new assumption.
(iii) Design a new model that has important aspects of the original model and satisfies all the assumptions, e.g. by applying a proper transformation to the data or filtering out some suspect data point which may be considered outlying.
(iv) Use a distribution-free procedure that is valid even if various assumptions are violated.

Most researchers, however, have opted for (iii) which has attracted much attention as documented by Thoeni (1967) and Hoyle (1973) among others. In this paper the parametric power transformation proposed by Box & Cox (1964) is reviewed in the context of model simplification as well as that of finding a metric in which the theoretical assumptions made in an analysis are more nearly satisfied.

2 The Box–Cox transformation and some alternative versions

Tukey (1957) introduced a family of power transformations such that the transformed values are a monotonic function of the observations over some admissible range and indexed by

$$y_i^{(\lambda)} = \begin{cases} y_i^{\lambda}; & \lambda \neq 0 \\ \log y_i; & \lambda = 0 \end{cases} \tag{1}$$

for $y_i > 0$. However, this family has been modified by Box & Cox (1964) to take account of the discontinuity at $\lambda = 0$, such that

$$y_i^\lambda = \begin{cases} (y_i^\lambda - 1)/\lambda; & \lambda \neq 0 \\ \log y_i; & \lambda = 0 \end{cases} \tag{2}$$

and that for unknown λ

$$\mathbf{y}^{(\lambda)} = (y_1^{(\lambda)}, y_2^{(\lambda)}, \ldots, y_n^{(\lambda)})' = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

where $\mathbf{X}$ is a matrix of known constants, $\boldsymbol{\theta}$ is a vector of unknown parameters associated with the transformed values and $\boldsymbol{\varepsilon} \sim$ MVN $(0, \sigma^2 I_n)$ is a vector of random errors. The transformation in equation (2) is valid only for $y_i > 0$ and, therefore, modifications have had to be made for negative observations. Box & Cox proposed the shifted power transformation with the form

$$y_i^{(\lambda)} = \begin{cases} \{(y_i + \lambda_2)^{\lambda_1} - 1\}/\lambda_1; & \lambda_1 \neq 0 \\ \log(y_i + \lambda_2); & \lambda_1 = 0 \end{cases} \tag{3}$$

where λ_1 is the transformation parameter and λ_2 is chosen such that $y_i > -\lambda_2$.

Manly (1976) suggested another alternative which can be used with negative observations and which is claimed to be effective at turning skew unimodal distributions into nearly symmetric normal-like distributions and is of the form

$$y_i^{(\lambda)} = \begin{cases} (\exp(\lambda y_i) - 1)/\lambda; & \lambda \neq 0 \\ y_i; & \lambda = 0 \end{cases} \tag{4}$$

John & Draper (1980) introduced the so-called modulus transformation which is considered to normalize distributions already possessing some measure of approximate symmetry and carries the form

$$y_i^{(\lambda)} = \begin{cases} \operatorname{sign}(y_i)\{(|y_i| + 1)^\lambda - 1\}/\lambda; & \lambda \neq 0 \\ \operatorname{sign}(y_i)\{\log(|y_i| + 1)\}; & \lambda = 0 \end{cases} \tag{5}$$

Bickel & Doksum (1981) suggested another modification so that distributions of $y_i^{(\lambda)}$ with unbounded support such as the normal distribution can be included.

For $\lambda > 0$, the extension is

$$y_i^{(\lambda)} = \{|y_i|^\lambda \operatorname{sign}(y_i) - 1\}/\lambda \tag{6}$$

It is important to note that the range of $y^{(\lambda)}$ in equations (1)–(3) and (5) is restricted according to whether λ is positive or negative. This implies that the transformed values do not cover the entire range $(-\infty, +\infty)$ and, hence, their distributions are of bounded support. Consequently, only approximate normality is to be expected.

3 Estimation of the transformation parameter

Box & Cox (1964) proposed maximum likelihood as well as Bayesian methods for the estimation of the parameter λ. Further Bayesian considerations have also been outlined by Pericchi (1981) and Sweeting (1984). Some robust adaptations of the estimation procedures have been studied by Carroll (1980; 1982 a), Bickel & Doksum (1981), Carroll and Ruppert (1985), Taylor (1983, 1985 a, b, 1987) and Carroll & Ruppert (1987). The extension of the Box–Cox procedure to multivariate data has been proposed in Andrew *et al.* (1971, 1973), Dunn & Tubbs (1980) and Beauchamp & Robson (1986). An approximation for the precision of the maximum likelihood estimate has been presented by Draper & Cox (1969) with a subsequent correction by Hinkley (1975). Cressie (1978) suggested a simple graphical procedure to estimate the transformation parameter by utilizing the principle of one degree or freedom for non-additivity in a two-way table without replication while Hernandes & Johnson (1980) proposed to estimate the transformation by minimizing the Kullback–Leibler information in examining the large sample behaviour of transformation to normality. A more analytical procedure to estimate the transformation parameter which simultaneously corrects for the non-additivity and heterogeneity of residuals has been proposed by Hinkley (1985). The approach is based on likelihood analysis for local deviations from a normal theory linear model. A non-parametric estimator of the transformation based on Kendall's rank correlation was proposed by Han (1987), and is found to be more consistent and efficient than the maximum likelihood estimator. The application of the Box–Cox technique to simple random effects models was suggested by Solomon (1985) and has been extended to cover all mixed models by Sakia (1988). Computer programs for estimating λ are given by Chang (1977 a) with a modification by Huang *et al.* (1978).

4 Hypothesis tests and other inferences on the transformation parameter

It is frequently of interest to test whether the estimate of the Box–Cox transformation parameter conforms to a hypothesized value. Box & Cox (1964) used the asymptotic distribution of the likelihood ratio to test some hypotheses about the parameter. Andrews (1971) further proposed a test for the value of the parameter whose null distribution is known and which is easier to calculate. This test was developed by ignoring the Jacobian of transformation whose omission was later investigated and, hence, incorporated in another test by Atkinson (1973) in which the power of the three tests was also compared. Tests derived from the likelihood were found to be more powerful. A robust competitor to these tests has been proposed by Carroll (1980). Atkinson's score statistic has been further standardized by Lawrance (1987 a) and a simulated comparison of the two tests indicate that the standardized statistic has improved standard normal behaviour relative to Atkinson's test. Lawrance's statistic has been extended to models where response and mean may both be transformed by Hinkley (1988). Furthermore, as an alternative to the likelihood ratio confidence interval for testing hypotheses about the transformation parameter, Lawrance (1987 b) has given an asymptotically justifiable expression for the estimated variance of λ which leads to more efficient hypothesis tests on λ than Atkinson's (1985, p. 100) which is based on a regression analogy for constructed variables. A more recent improvement of both Atkinson's test and Lawrance's standardized score statistic has been proposed by Wang (1987) and is said to give a more accurate approximation to the standard normal. A simulated comparison by Atkinson and Lawrance (1989) of the test statistics concluded that, in general, Atkinson's test is very similar to Lawrance's test. They suggested, however, that the small samples used in the study could have masked the superiority of Lawrance's test as claimed earlier. While Draper & Cox (1969) have shown that the estimation of λ is fairly robust to non-normality as long as the variable has a reasonably symmetric distribution, this may not be the case when skewness is encountered. Thus, Poirier (1978) investigated the effect of the transformation in limited dependent variables, i.e. variables which have been possibly censored or truncated thus introducing some skewness. The procedure maximizes the likelihood function of the truncated normal distribution. Although the asymptotic properties of the maximum likelihood estimators are known, little is known about their small sample properties. Spitzer (1978) examined the small sample properties of the parameter estimates employing a Box–Cox transformation. The procedure was found to be usefully implemented under the assumption of approximate normality. For forecasting purposes, the forecasts were unbiased and their variances were remarkably low. Bickel & Doksum (1981) studied consistency properties of the Box–Cox estimates of the transformation parameter in the linear model as well as the asymptotic variances of these estimates. They found that in linear regression models with small to moderate error variances, the asymptotic variances of these parameters are much larger when the transformation parameter is unknown than when it is known and in some unstructured models the cost of not knowing λ was found to be moderate to small. Moreover, they concluded that the performance of all Box–Cox type procedures is unstable and highly dependent on the parameters of the model in structured models with small to moderate error variances. This statement has been refuted by way of clarification by Box & Cox (1982). Also in response to the work of Bickel & Doksum (1981), further discussions have been presented by Carroll & Ruppert (1981), Carroll (1982 a) and Hinkley & Runger (1984). Doksum & Wong (1983) have used asymptotic and Monte Carlo methods to study the effect of estimation of parameters on tests of hypotheses and concluded by asymptotic efficiency results that when the Box–Cox transformation is used, tests used on transformed data have good power properties. It is generally accepted therefore, that the standard methods for the normal theory linear model are justifiable when applied to the transformed variable as if the transformation parameter was known beforehand, i.e. not making an allowance for its estimation from the data. An incorporation of the Box–Cox transformation in situations where the theoretical considerations already provide a regression function has been examined by Wood (1984) and Carroll & Ruppert (1984, 1988) by transformating simultaneously the response and the theoretical model and, by a Monte Carlo study, concluded that for estimating the model parameters there is little cost for not knowing the correct transformation *a priori.* More recently, the subject of transforming theoretical or empirical models has been examined by Ruppert *et al.* (1989) in fitting the Michaelis–Menten model, as well as its error structure. Rudemo *et al.* (1989) applied the power transformation for the logistic model in bioassay where it was found to perform well on the basis of the data set used. A theoretical as well as a simulation comparison of the conditional and unconditional tests of hypothesis after a Box–Cox power transformation in linear models with a single error vector has been conducted by Wixley (1986). Unconditional likelihood ratio tests are shown to have the more correct level.

5 Empirical determination of the functional form

It has been acknowledged in econometric studies that the determination of the functional relationship that may exist between some variables of interest need not be based on *a priori* economic rationale. The simplest procedure that has been accepted and successfully applied is the Box–Cox transformation technique. In much of the research that has been undertaken, the following functional form has been accepted as standard:

$$y_i^{(\lambda_0)} = \beta_0 + \sum_{j=1}^{q} \beta_j x_{ji}^{(\lambda_j)} + \varepsilon_i \tag{7}$$

where $y_i^{(\lambda_0)}$ and $x_{ji}^{(\lambda_j)}$ are the transformed regressand and regressants, respectively, and ε_i represents the random errors. The presence of a model constant is a prerequisite to preserving the scale invariance as indicated by Schlesselman (1971). Zarembka (1968) considered the functional form in the demand for money as related to both real income and interest rate with a common λ while White (1972) estimated the liquidity trap with the same restrictions on λ. Heckman & Polacheck (1974) studied the relationship between earning, schooling and experience with a generalized λ whereas Kau & Lee (1976) considered the functional form between population density and the distance from the central business district with $\lambda_1 = 1$. Khan & Ross (1977) determined the aggregate form of the import demand equation with a common λ and Mills (1978) considered the functional form for the UK for money. (for further comments, see also Oxley, 1982). Chang (1977 b, 1980) estimated the functional relationship for demand of meat in the USA with $\lambda_0 \neq \lambda_1$ which has also been discussed by Gemmil (1980). Spitzer (1976) considered the relationship between the demand for money and the liquidity trap with a generalized Box–Cox parameter. An examination of the aggregate import demand equation by Boylan *et al.* (1980) was constrained to a common λ and a further examination by Boylan & O'Muircheartaigh (1981) constrained $\lambda_1 = \lambda_3, \lambda_1 \neq \lambda_2$. This was further generalized by Boylan *et al.* (1982). Lin & Huang (1983) estimated the generalized functional form for the yield trend of wheat, corn and soybean. Newman (1977) estimated the relationship between the incidence of malaria and the mortality rate and concluded that the functional specificiation obtained by using the Box–Cox procedure was superior to earlier specifications. Some different procedures for estimating the transformation parameter in normal error models have been examined by Spitzer (1982 a, b) which, although leading to essentially the same estimates, differ in terms of computational time. Poirier (1978) studied some estimation methodology when the error terms are truncated normal. For some discussion of the interpretation of estimated coefficients in Box–Cox models, see Poirier & Melino (1978), Huang & Kelingos (1979), Mallela (1980) and Huang & Grawe (1980). The generalized Box–Cox transformation has also been applied to model price changes (e.g. Milon *et al.*, 1984) and demand and supply elasticities (Bessler *et al.*, 1984). Soybean yield functions have been examined by Miner (1982) and Davison *et al.* (1989) have modelled US soybean export. They concluded that the transformation provides approximately normally distributed error terms, a condition which is important for hypothesis testing and the construction of confidence intervals. It is important however, to point out that when certain *a priori* restrictions are placed on the transformation parameter, some behavioural properties are also unnecessarily forced upon the function. Since the Box–Cox transformation procedure calls for the resulting functional form to be entirely an outcome of the estimation process, any form of restrictions to be imposed on an *a priori* basis should be avoided as much as possible.

6 Variance heteroscedasticity and autocorrelation of the error structure

Although the Box–Cox procedure and, in particular, the maximum likelihood method, has been shown to be robust to non-normality so long as there is reasonable symmetry in the disturbances (Draper & Cox, 1969), Zarembka (1974) has indicated that the procedure is not robust with respect to heteroscedasticity. There is a bias in estimating the transformation parameter towards that transformation of the dependent variable which leads to the stabilization of the error variance. This problem has prompted some modification of the Box–Cox procedure to take into account the estimation of λ in models with heteroscedastic error. Much of the work is based on assuming (or empirically estimating) the relationship between the variance and the mean. For example, Zarembka (1974) assumed a relationship of the form

$$V(y_i) = \sigma^2[E(y_i)]^\delta \qquad E(y_i) > 0 \tag{8}$$

where δ was assumed to be known.

More recently, Lahiri & Egy (1981) assumed the variance of the transformed response to be of the form

$$V(y_i^{(\lambda)}) = \sigma^2 z_i^\delta \tag{9}$$

for exogenously given z_i and both σ^2 and δ to be unknown. However, both the above forms have been modified by Sarkar (1985) to take into effect the heteroscedasticity of the transformed response. It is based on assuming that the variance of a Box–Cox transformed variable can be approximated by Bartlett's (1947) variance stabilization procedure as

$$V(y_i^{(\lambda)}) \approx V(y_i)[E(y_i)]^{2\lambda - 2} \tag{10}$$

Then by further assuming equation (8) but with δ unknown, we obtain

$$V(y_i^{(\lambda)}) \simeq \sigma^2[E(y_i)]^{2\lambda-2+\delta} \quad (11)$$

Maximum likelihood estimation is used to estimate the transformation parameter λ and the form of heteroscedasticity δ simultaneously. Note that for homoscedasticity to be achieved $\delta = 2 - 2\lambda$, a case which has been considered earlier by Box & Hill (1974), Pritchard *et al.* (1977), Hinz & Eagles (1976) and in a slightly different context by Pritchard & Bacon (1977). Further consideration of variance stabilization was undertaken by Dunn & Tubbs (1980) in the context of several multivariate populations with possibly unequal covariance matrices. The transformation was used to enhance homogeneity of the covariance, leading to reduction of the heterogeneity effect upon the analysis of the linear model. Seaks & Layson (1982) extended the Box–Cox transformation to correct for heteroscedasticity and autocorrelation in econometric models in which they used some analogous variation of the standard model to estimate the transformation parameter. In an earlier paper, Savin & White (1978) stressed that failure to correct for autocorrelation in a Box–Cox transformation can cause misleading results in the sense that an indication of autocorrelation might really be a problem of functional form in disguise. A more generalized procedure of the Box–Cox transformation in models with a heteroscedastic error has been considered by Blayblock *et al.* (1980) and Blayblock & Smallwood (1985) by estimating the analytic form of heteroscedasticity simultaneously with the non-stochastic part of the model in empirically determining the functional form of the import demand equation. Due to the interdependent relationship between heteroscedasticity in the error term and the transformation parameter on the dependent variable, a failure to incorporate a variance stabilization modification in Box–Cox transformation may adversely affect the conclusions.

7 Effect of outliers and influential cases

The selection of a transformation may be properly viewed as model selection and, in this initial phase of analysis, influential cases can have particularly important and lasting effects that are difficult to uncover in the subsequent analysis. Thus, an outlying observation in the original scale may conform in the transformed scale. It is therefore necessary to find if the evidence for the particular transformation is spread evenly throughout the data or just within a few cases. Atkinson (1982) introduced some diagnostic displays of outlying and influential observations in multiple regression and their possible reduction by a transformation. Some criteria for estimating the transformation through the use of constructed variables are suggested by Atkinson (1983, 1985) and this has been extended to the power transformation after a shift in location. Carroll (1980, 1982 b) proposed robust estimators of the transformation parameter by replacing the normal likelihood with an objective function that is less sensitive to outlying responses. Cook & Wang (1983) described a method to measure the influence of cases on the Box–Cox likelihood estimate of the response transformation parameter in linear regression and this is indicated to be superior to Atkinson's for the purpose of detecting influential cases for a transformation. Atkinson (1986) further extended his work by deriving some expressions for estimating the effect of deletion of observations on the estimate of the transformation parameter and some subsequent tests of hypotheses.

8 Prediction in the original scale

One of the most controversial arguments in data transformation has been about which final scale to use in making some inferences, i.e. whether one should make unconditional inferences about the regression parameters in the correct but unknown scale as in Bickel &

Doksum (1981) or a conditional inference for an appropriately defined regression parameter in an estimated scale as advocated in a rebuttal by Box & Cox (1982) and Hinkley & Runger (1984). To avoid this problem, Carroll & Ruppert (1981) studied the cost of estimating the transformation parameter when inferences are to be made in the original scale of the observations. They concluded that for prediction as well as other problems in the original scale, there is a cost due to estimating the transformation parameter which is not generally severe. In the same manner Carroll (1982 c) focused attention on estimating the median of response in the original scale for a given regressant specifically when the transformation parameter is restricted to a finite set and their analysis shows that restricted estimation of the transformation parameter can possibly lead to inferences different from maximum likelihood estimation of the median response. Rather than estimating the median, Taylor (1986) has given an approximate method for estimating the mean of the dependent variable after a Box–Cox transformation. The approximation is shown to be inaccurate when the transformation parameter is close to zero and when the error variance is large. It is also shown that there is some cost in estimating the transformation parameter, although not a severe one. Sakia (1988, 1990) has extended the approximation to cover all balanced mixed models. In addition, expressions for the variance of the estimated mean and the bias due to retransformation are provided. In a Monte Carlo experiment Smallwood & Blayblock (1986) examined the forecasting performance of the power transformation and showed that the sign and magnitude of the transformation parameter influence the forecasting performance.

9 Conclusions

The Box–Cox transformation has been widely used since it was first proposed. It has inspired a large amount of research on its applicability as well as on the drawbacks arising from its use. However, one thing is clear; that seldom does this transformation fulfil the basic assumptions of linearity, normality and homoscedasticity simultaneously as originally suggested by Box & Cox (1964). The Box–Cox transformation has found more practical utility in the empirical determination of functional relationships in a variety of fields, especially in econometrics.

Acknowledgements

I am very grateful to Professor Dr H. Thoeni for his guidance and for a referee for pointing out some anomalies in the original draft. The German Academic Exchange Service (DAAD) is kindly acknowledged for financial support.

References

ANDREWS, D. F. (1971) A note on the selection of data transformation, *Biometrika*, 58, 249–254.

ANDREWS, D. F., GNANADESIKAN, R. & WARNER, J. L. (1971) Transformations of multivariate data, *Biometrics*, 27, 825–840.

ANDREWS, D. F., GNANADESIKAN, R. & WARNER, J. L. (1973) Method for assessing multivariate normality, in: P. R. KRISHNAIAH (Ed.) *Multivariate Analysis III*, pp. 95–115 (New York, Academic Press).

ATKINSON, A. C. (1973) Testing transformations to normality, *Journal of the Royal Statistical Society, Series B*, 35, 473–479.

ATKINSON, A. C. (1982) Regression diagnostics, transformation and constructed variables, *Journal of the Royal Statistical Society, Series B*, 44, 1–36.

ATKINSON, A. C. (1983) Diagnostic regression analysis and shifted power transformation, *Technometrics*, 25, 23–33.

ATKINSON, A. C. (1985) *Plots, Transformation and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis* (Oxford, Clarendon Press).
ATKINSON, A. C. (1986) Diagnostic tests for transformation, *Technometrics*, 28, 29–37.
ATKINSON, A. C. and LAWRANCE, A. J. (1989) A comparison of asymptotic equivalent test statistics for regression transformation, *Biometrika*, 76, 223–229.
BARTLETT, M. S. (1947) The use of transformations, *Biometrics*, 3, 39–52.
BESSLER, D. A., WEAVER, R. D., WOHLGENANT, M. K., CHALLANT, J. A., KING, G. A. & GALLANT, A. R. (1984) Estimating functional forms with special reference to agriculture, *American Journal of Agricultural Economics*, 66, 204–224.
BEAUCHAMP, J. J. & ROBSON, D. S. (1986) Transformation considerations in discriminant analysis, *Communications in Statistics—Simulation and Computation*, 15(1), 147–179.
BICKEL, P. J. and DOKSUM, K. A. (1981) An analysis of transformations revisited, *Journal of the American Statistical Association*, 76, 296–311.
BLAYBLOCK, J. R. & SMALLWOOD, D. M. (1985) Box–Cox transformation and heteroscedastic error variance: Import demand equation revisited, *The International Statistical Review*, 53, 91–97.
BLAYBLOCK, J. R. & SMALLWOOD, D. M. (1986) Forecasting performance of models using the Box–Cox transformation, *Agricultural Economics Research* (*USDA*), 38(4), 14–24.
BLAYBLOCK, J., SLATHE, L. & GREEN, R. (1980) A note on the Box–Cox transformation under heteroscedasticity, *West Journal of Agricultural Economics*, 5(2), 129–135.
BOX, G. E. P. & COX, D. R. (1964) An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, 26, 211–252.
BOX, G. E. P. & COX, D. R. (1982) An analysis of transformation revisited, rebutted, *Journal of the American Statistical Association*, 77, 209–210.
BOX, G. E. P. & HILL, W. J. (1974) Correcting inhomogeneity of variance with power transformations weighting. *Technometrics*, 16, 385–389.
BOYLAN, T. A., CUDDY, M. P. & O'MUIRCHEARTAIGH, I. (1980) The functional form of the aggregate import demand equation, *Journal of International Economics*, 10, 561–566.
BOYLAN, T. A., CUDDY, M. P. & O'MUIRCHEARTAIGH, I. (1982) Import demand equation: an application of generalized Box–Cox methodology, *The International Statistical Review*, 50, 103–112.
BOYLAN, T. A. & O'MUIRCHEARTAIGH, I. G. (1981) The functional form of the U.K. demand for money: a critique of a paper by Mills, *Applied Statistics*, 30, 296–299.
CARROLL, R. J. (1980) A robust method for testing transformation to achieve approximate normality, *Journal of the Royal Statistical Society, Series B*, 42, 71–78.
CARROLL, R. J. (1982 a) Tests for regression parameters in power transformation models, *Scandinavian Journal of Statistics*, 9, 217–222.
CARROLL, R. J. (1982 b) Two examples of transformation when there are possible outliers, *Applied Statistics*, 31, 149–152.
CARROLL, R. J. (1982 c) Prediction and power transformations when the choice of power is restricted to a finite set, *Journal of the American Statistical Association*, 77, 908–915.
CARROLL, R. J. & RUPPERT, D. (1981) On prediction and the power transformation family, *Biometrika*, 79, 321–328.
CARROLL, R. J. & RUPPERT, D. (1984) Power transformations when fitting theoretical models to data, *Journal of the American Statistical Association*, 79, 321–615.
CARROLL, R. J. & RUPPERT, D. (1987) Diagnostics and robust estimation when transforming the regression model and the response, *Technometrics*, 29, 287–299.
CARROLL, R. J. & RUPPERT, D. (1988) *Transformations and Weighting in Regression* (London, Chapman and Hall).
CHANG, H. S. (1977 a) A computer program for Box–Cox transformation and estimation technique, *Econometrica*, 45, 1741.
CHANG, H. S. (1977 b) Functional form and the demand for meat in the United States, *Review of Economics and Statistics*, 59, 355–359.
CHANG, H. S. (1980) Functional form and the demand for meat in the United States: a reply, *Review of Economics and Statistics*, 62, 148–150.
COCHRAN, W. G. (1947) Some consequences when the assumptions for the analysis of variance are not satisfied, *Biometrics*, 3, 22–38.
COOK, R. D. & WANG, P. C. (1983) Transformation and influential cases in regression, *Technometrics*, 25, 337–343.
CRESSIE, N. A. C. (1978) Removing nonadditivity from two-way tables with one observation per cell, *Biometrics*, 34, 505–513.
DAVISON, C. W., ARNADE, C. A. & HALLAHAN, C. B. (1989). Box–Cox estimation of U.S. soyabean exports, *Journal of Agricutural Economics Research*, 41(3), 8–15.
DOKSUM, K. A. & WONG, C. W. (1983) Statistical tests based on transformed data, *Journal of the American Statistical Association*, 78, 411–417.
DRAPER, N. R. & COX, D. R. (1969) On distributions and their transformations to normality, *Journal of the Royal Statistical Society, Series B*, 31, 472–476.

DUNN, J. E. & TUBBS, J. D. (1980) VARSTAB: A procedure for determining homoscedastic transformations of multivariate normal populations, *Communications in Statistics—Simulation and Computation B*, 9(6), 589–598.

EISENHART, C. (1947) The assumption underlying the analysis of variance, *Biometrics*, 3, 1–21.

GEMILL, G. (1980) Using the Box–Cox form for estimating demand. a comment, *Review of Economics and Statistics*, 62, 147–148.

GRAYBILL, F. A. (1976) *The Theory and Applications of the Linear Model* (London, Duxbury Press).

HAN, A. K. (1987) A non-parametric analysis of transformation, *Journal of Econometrics*, 35, 191–209.

HECKMAN, J. & POLSCHEK, S. (1974) Empirical evidence of the functional form of the earning–schooling relationship, *Journal of the American Statistical Association*, 69, 350–354.

HERNANDES, F. & JOHNSON, R. A. (1980) The large sample behaviour of transformations to normality, *Journal of the American Statistical Association*, 75, 855–861.

HINKLEY, D. V. (1975) On power transformation to symmetry, *Biometrika*, 62, 101–111.

HINKLEY, D. V. (1985) Transformation diagnostics for linear models, *Biometrika*, 72, 487–496.

HINKLEY, D. V. (1988) More on score tests for transformation in regression, *Biometrika*, 75, 366–369.

HINKLEY, D. V. & RUNGER, G. (1984) The analysis of transformed data, *Journal of the American Statistical Association*, 79, 302–320.

HINZ, P. & EAGLES, H. A. (1976) Estimation of a transformation for the analysis of some agronomic and genetic experiments, *Crop Science*, 16, 280–283.

HOYLE, M. H. (1973) Transformations: An introduction and a bibliography, *The International Statistical Review*, 41, 203–223.

HUANG, C. L. & GRAWE, O. R. (1980) Functional forms and the demand for meat in the United States: a comment, *Review of Economics and Statistics*, 62, 144–146.

HUANG, C. L. & KELINGOS, J. A. (1970) Conditional mean function and general specification of the disturbances in regression analysis, *South Economics Journal*, *45*, 710–717.

HUANG, C. L., MOON, L. C. & CHANG, H. S. (1978) A computer program using the Box–Cox transformation technique for the specification of functional form, *The American Statistician*, 32, 144.

JOHN, J. A. & DRAPER, N. R. (1980) An alternative family of transformations, *Applied Statistics*, 29, 190–197.

KAU, J. B. & LEE, C. F. (1976) The functional form in estimating the density gradient: an empirical investigation, *Journal of the American Statistical Association*, 71, 326–327.

KHAN, M. S. & ROSS, K. Z. (1977) The functional form of the aggregate import demand equation, *Journal of International Economics*, 7, 149–160.

LAHIRI, K. & EGY, D. (1981) Joint estimation and testing for functional form and heteroscedasticity, *Journal of Econometrics*, 15, 299–307.

LAWRANCE, A. J. (1987 a) The score statistic for regression transformation, *Biometrika*, 74, 275–279.

LAWRANCE, A. J. (1987 b) A note on the variance of the Box–Cox regression transformation estimate, *Applied Statistics*, 36, 221–223.

LIN, T. K. H. and HUANG, C. L. (1983) Use of Box–Cox transformation technique for fitting crop yield trends, *Agronomy Journal*, 75, 310–314.

MALLELA, P. (1980) Discrimination between linear and logarithmic forms: a note. *Review of Economics and Statistics*, 62, 142–144.

MANLY, B. F. (1976) Exponential data transformation, *The Statistician*, 25, 37–42.

MILLS, T. C. (1978) The functional form of the U.K. demand for money, *Applied Statistics*, 27, 52–57.

MILLON, J. W., GRESSEL, J. and MULKEY, D. (1984) Hedonic amenity valuation and functional form specification, *Land Economics*, 60, 378–387.

MINER, A. G. (1982) The contribution of weather and technology to U.S. soybeam yield. *Unpublished Dissertation*, University of Minnesota.

NEWMAN, P. (1977) Malaria and mortality, *Journal of the American Statistical Association*, 72, 257–263.

OXLEY, L. T. (1982) Box–Cox transformation and the demand for money, *Applied Statistics*, 31, 304.

PERICCHI, L. R. (1981) A Bayesian approach to transformations to normality, *Biometrika*, 68, 35–43.

POIRIER, D. J. (1978) The use of the Box–Cox transformation in limited dependent variable models, *Journal of the American Statistical Association*, 73, 285–287.

POIRIER, D. J. and MELINO, A. (1978) A note on the interpretation of regression coefficient within a class of truncated distributions, *Econometrica*, 46, 1207–1209.

PRITCHARD, D. J., DOWNIE, J. D. and BACON, D. W. (1977) Further considerations of heteroscedasticity in fitting kinetic models, *Technometrics*, 19, 227–236.

PRITCHARD, D. J. & BACON, D. W. (1977) Accounting for heteroscedasticity in experimental designs, *Technometrics*, 19, 109–115.

RUDEMO, M., RUPPERT, D. & STREIBIG, J. C. (1989) Random effects models in nonlinear regression with application to bioassay, *Biometrics*, 45, 349–362.

RUPPERT, D., CRESSIE, N. & CARROLL, R. J. (1989) A transformation/weighting model for estimating Michaelis–Menten parameters, *Biometrics*, 45, 637–656.

SAKIA, R. M. (1988) Application of the Box–Cox transformation technique to linear balanced mixed analysis of variance models with a multi-error structure, *Unpublished PhD Thesis*, Universitaet Hohenheim, FRG.

SAKIA, R. M. (1990) Retransformation bias: a look at the Box–Cox transformation to linear balanced mixed ANOVA models, *Metrika*, 37, 345–351.

SARKAR, N. (1985) Box–Cox transformation and the problem of heteroscedasticity, *Communications in Statistics—Theory and Methods*, 14(2), 363–379.

SAVIN, N. E. & WHITE, K. J. (1978) Estimating and testing for functional form and autocorrelation, *Journal of Econometrics*, 8, 1–12.

SEAKS, T. G. & LAYSON, (1982) Box–Cox estimation with standard econometric problems, *Review of Economics and Statistics*, 64, 160–164.

SCHLESSELMAN, J. (1971) Power families. A note on the Box and Cox transformation, *Journal of the Royal Statistical Society, Series B*, 33, 307–371.

SMALLWOOD, D. M. & BLAYBLOCK, J. R. (1986) Forecasting performance of models using the Box–Cox transformation, *Agricultural Economics Research*, 38, 14–24.

SPITZER, J. J. (1976) The demand for money, the liquidity trap and functional forms, *The International Economics Review*, 17, 220–227.

SPITZER, J. J. (1978) A Monte Carlo investigation of the Box–Cox transformation in small samples, *Journal of the American Statistical Association*, 73, 488–495.

SPITZER, J. J. (1982 a) A primer on Box–Cox estimation, *Review of Economics and Statistics*, 64, 307–313.

SPITZER, J. J. (1982b) A fast and efficient algorithm for the estimation of parameters in models with the Box and Cox transformation, *Journal of the American Statistical Association*, 77, 760–766.

SOLOMON, P. J. (1985) Transformations for components of variance and covariance, *Biometrika*, 72, 233–239.

SWEETING, T. J. (1984) On the choice of the prior distribution for the Box–Cox transformed linear model, *Biometrika*, 71, 127–134.

TAYLOR, M. J. G. (1983) Power transformation to symmetry. *Unpublished Dissertation*, Department of Statistics, University of California, Berkeley.

TAYLOR, M. J. G. (1985 a) Power transformation to symmetry, *Biometrika*, 72, 145–152.

TAYLOR, M. J. G. (1985 b) Measure of location of skew distributions obtained through Box–Cox transformations, *Journal of the American Statistical Association*, 80, 427–432.

TAYLOR, M. J. G. (1986) The retransformed mean after a fitted power transformation, *Journal of the American Statistical Association*, 81, 114–118.

TAYLOR, M. J. G. (1987) Using a generalized mean as a measure of location, *Biometrical Journal*, 6, 731–738.

THOENI, H. (1967) Transformation of variables used in the analysis of experimental and observational data: a review. *Technical Report No.* 7, Iowa State University, Ames.

THOENI, H. (1969) A table for estimating the mean of a lognormal distribution, *Journal of the American Statistical Association*, 64, 632–636.

TUKEY, J. W. (1957) The comparative anatomy of transformations, *Annals of Mathematical Statistics*, 28, 602–632.

WANG, S. (1987) Improved approximation for transformation diagnostics, *Communications in Statistics—Theory and Methods*, 16(6), 1797–1819.

WHITE, K. J. (1972) Estimation of the liquidity trap with a generalized functional form, *Econometrica*, 40, 193–199.

WIXLEY, R. A. J. (1986) Unconditional likelihood tests for the linear model following Box–Cox transformation, *South African Statistical Journal*, 20, 67–83.

WOOD, J. T. (1974) An extension of transformations of Box and Cox, *Applied Statistics*, 23, 278–283.

ZAREMBKA, P. (1968) Functional form in the demand for money, *Journal of the American Statistical Association*, 63, 502–511.

ZAREMBKA, P. (1974) Transformation of variables in Econometrics, in: P. ZAREMBKA (ed.) *Frontiers of Econometrics*, pp. 81–104 (New York, Academic Press).