573

Economics Letters 30 (1989) 27–29
North-Holland

THE BOX–COX TRANSFORMATION *
Power Invariance and a New Interpretation

Marc GAUDRY

Université de Montréal, Montréal, Quebec, Canada H3C 3J7

Richard LAFERRIÈRE

Université de Montréal, Montreal, Quebec, Canada H3C 3J7, and Concordia University, Montreal, Quebec, Canada H3C 3J7

Received 25 October 1988
Accepted 8 January 1989

We show that a Box–Cox transformation on the dependent variable of linear regression models is invariant to power transformations of that variable even without the presence of a regression constant and, consequently, can sometimes be interpreted as a simple power transformation the estimation of which will then admit of a non-degenerate solution.

1. Introduction

Concern about the appropriate use of the Box–Cox transformation (henceforth BCT) applied to the dependent variable of models has centered on the influence of *scalar* transformations of that variable: Schlesselmann (1971) has shown that invariance of the Box–Cox parameter and its variance estimates required the presence of a constant term among the regressors; later, Spitzer (1984) pointed out that asymptotic *t*-statistics of the linear coefficients of the regression problem depend upon the units of measurement of the dependent variable, a predicament that was implicit in Burguette et al. (1982) and may well find its solution along the lines suggested by Dagenais and Dufour (1987). Here we want to focus on the influence of *power* transformations of the dependent variable and, secondarily, point out that the BCT can then sometimes be interpreted as a simple power transformation. This means that, when the BCT [1] is used as an approximation of the power transformation, the estimation of the latter now admits of a non-degenerate solution.

2. Power invariance of the BCT transformation

Consider the following model, written without observation subscripts

$$y^{(\lambda)} = \beta_0 + \sum_{k=1}^{K} \beta_k X_k + \epsilon, \tag{1}$$

* The research on which this paper is based was supported at various stages by Transport Canada, the F.C.A.R. program of the Québec goverment, the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Coucil of Canada.

[1] It was introduced as an improvement of Tukey's (1957) original power transformation [y^{λ}, for $\lambda \neq 0$ and $\log(y)$ for $\lambda = 0$] in order to avoid the discontinuity at $\lambda = 0$.

where $y^{(\lambda)}$ equals $(y^{\lambda}-1)/\lambda$ if $\lambda \neq 0$ and $\log(y)$ otherwise. We first note that (1) is a particular case of (2) with $s=1$,

$$y^{s^{(\lambda)}} = \beta_0 + \sum_{k=1}^{K} \beta_k X_k + \epsilon. \tag{2}$$

Now, since the left side of (2) is

$$y^{s^{(\lambda)}} = \frac{y^{s\lambda}-1}{\lambda} = s\left[\frac{y^{s\lambda}-1}{s\lambda}\right] = s\left[y^{(s\lambda)}\right], \tag{3}$$

eq. (2) can be rewritten as

$$s\left[y^{(\lambda^*)}\right] = \beta_0 + \sum_{k=1}^{K} \beta_k X_k + \epsilon, \tag{4}$$

where $\lambda^* = s\lambda$. Eq. (4) shows that the power transformation parameter s is equivalent to scaling the transformed dependent variable.

Moreover, it is also the case that

$$(sy)^{(\lambda)} = s^{\lambda} y^{(\lambda)} + s^{(\lambda)}, \tag{5}$$

for which Schlesselman (1971) demonstrated that the presence of the second term on the right side required a constant among the regressors in order to obtain the invariance of the λ parameter to the scalar s. It is also clear from his demonstration that, were the term $s^{(\lambda)}$ absent, the constant would not be required. This means that, in our problem, power invariance is insured even in the absence of the constant term: this is not surprising as s in (4) is applied to the transformed dependent variable. In effect then, the β parameters and their variances change exactly by s and the t-statistics are invariant, as in the usual linear regression case.

3. A new interpretation and an old problem

Since λ^* is unique, it can be interpreted as the power transformation s if (but only if) λ is assumed equal to 1. Clearly, one can only identify the product $s\lambda$ but, because of its unicity, setting λ equal to 1 means that λ^* is equivalent to the power transformation parameter s.

Now remember that the following regression problem:

$$y^{s} = \beta_0 + \sum_{k=1}^{K} \beta_k X_k + \epsilon \tag{6}$$

looks reasonable but has the degenerate solution $\beta_k = 0$, $\beta_0 = 1$ and $s=0$ if one maximizes the corresponding likelihood function. But our result above suggests a more satisfactory possibility: as long as $y^{s^{(\lambda)}}$ is a good representation of y^{s}, as it should be since they both have the same variance analysis [Box and Cox (1964)], the λ^* parameter in (4) can be used as a non-degenerate solution to this problem because it can be interpreted as a way of finding a value for s.

4. Conclusion and significance

It is clear from (3) that the invariance of the lambda parameter to a simple power transformation of the dependent variable also holds for the independent variables. Furthermore, the t-statistics of the β coefficients are unaffected by such transformations since a simple scaling of the explanatory variables is involved. In practice, the unicity of λ^* in (4) insures the appropriateness of many interpretations of the BCT, for instance as Arrow–Pratt's relative risk aversion measure [Montmarquette and Blais (1987)], as degree of inequality aversion in Atkinson's index [Guerrero (1987)], or as order of the general mean [Norris (1976)].

References

Box, G.E.P. and D.R. Cox, 1964, An analysis of transformations, Journal of the Royal Statistical Society, Series B, 211–243.

Burguette, J.S., A.R. Gallant and G. Souza, 1982, On unification of the asymptotic theory of nonlinear econometric models, Econometric Reviews 1, no. 2.

Dagenais, M.G. and J.-M. Dufour, 1987, Invariance, nonlinear models and asymptotic tests, Cahier no. 3287 (Centre de Recherche et de Développement en Économique, Université de Montréal, Montréal).

Guerrero, V.M., 1987, A note on the estimation of Atkinson's index of inequality, Economics Letters 25, 379–384.

Montmarquette, C. and A. Blais, 1987, A survey measure of risk aversion, Economics Letters 25, 27–30.

Norris, N., 1976, General means and statistical theory, The American Statistician 30, 8–12.

Schlesselmann, J., 1971, Power families: A note on the Box and Cox transformation, Journal of the Royal Statistical Society, Series B, 33, 307–311.

Spitzer, J.J., 1984, Variance estimates in models with the Box–Cox transformation: Implications for estimation and hypothesis testing, The Review of Economics and Statistics 66, 645–652.

Tukey, J.W., 1957, On the comparative anatomy of transformations.