A Look at the Box-Cox Transformation: The Use of Linear Versus Semilogarithmic Functional Forms for the Econometric Model of the Central Bank of Barbados

by

Peter Welch

June 1986

Nonlinearities in econometric and statistical models are playing an increasingly important role; this has not been the case in the past. This is due to the fact that advanced desk-top software is becoming easily available. It is also a fact that the statistical techniques used in both estimation and specification tests are biased towards linearity (Salmon and Wallis (1982)).

"The fact that the best estimator of a parameter is nonlinear should no longer be regarded as a deterrent to employing that estimator" (Harvey 1985). Accordingly, an examination of nonlinear estimators and their properties is important for econometricians, policymakers and all those involved in the difficult task of building structural models.

One of the best ways of exploring this avenue of econometric research is by an analysis of some of the techniques used to provide nonlinear estimates. This paper will examine the various approaches to nonlinear estimation. In particular the Gauss-Newton method is applied to the Box-Cox transformation. This transformation is used to determine the relationship of explanatory variables (relative prices, incomes and lagged nontraded output) to the assumed dependent variable, output of nontraded goods in the Barbadian economy. It is demonstrated that on the assumption that we have the ideal properties for the disturbances, the proper choice of functional form for the explanatory variables is their logarithms.

The linear pseudomodel is then applied to the model in an attempt to illustrate the direction that future research could take when the covariance structure of the disturbance is under question.

SECTION 1

NONLINEAR ESTIMATION

Estimation in nonlinear models is made difficult whether one adopts the least squares or maximum likelihood principle because the first order conditions of the maximization or minimization problem usually cannot be easily solved analytically. As a result least squares or maximum likelihood estimates have to be provided by iterative numerical techniques.

An iterative technique can be described as follows: an initial estimate is obtained and a new estimate which is hoped to be an improvement on the original is computed by a given rule. This process is then repeated until convergence occurs. If the procedure is successful, the final estimate should satisfy all the properties required of that particular estimation principle. The rules governing the iterative procedure provide the basis of a particular optimization algorithm.

There is a wide range of algorithms available. They differ in the extent to which they employ partial derivatives. Apart from this the choice of particular algorithm will depend to some extent on the type of function to be maximised or minimised.

Judge et al (1980) provide an adequate appraisal of these algorithms.

In discussing the general problem of nonlinear estimation it will be assumed that a criterion function f(p) is to be minimised with respect to the n parameters in the vector p. The n x 1 vector of first partial derivatives will be denoted by g(p) or simply g, while the n x n matrix of second partial derivatives the hessian will be written

$$\frac{d^2 f(p)}{dp\, dp'} = G(p)$$

where the operator d^2 refers to the second partial differentiation unless otherwise indicate.

A general recursion relation considered by most algorithms is the following

$$p^* = \hat{p} + l\ d\ (\hat{p})$$

where $\hat{p}$ is the current approximation to the minimum, p* is the revised estimate, l is a positive scalar known as the step-length and d(P) is the direction vector.

Typically the direction vector is the matrix product of the gradient vector and a positive definite matrix:

$$d\ (\hat{p}) = \hat{c}\ g\ (\hat{p}) \qquad (3)$$

where $\hat{c}$ is a positive definite matrix which leads to an 'acceptable step' (Fomby, Hill and Johnson (1984)).

The general form of the recursion relation is

$$p^* = \hat{p} + L\ \hat{c}\ g\ (\hat{p}) \qquad (4)$$

Methods involving such iterations are called gradient methods, since the direction vector is a function of the gradient vector g. Some of these gradient methods are presented below.

Methods of Steepest Descent

The choice of $\hat{c}$ as the identity matrix is motivated by the fact that g points in the direction of the most rapid decrease of the function f (.) at $p = \hat{p}$

The method of steepest descent is not without its problems. The method may converge to a saddle point rather than a maximum; if the maximum lies on a narrow ridge then there is a tendency for successive steps to oscillate back and forth across the ridge so that conversnce is slow.

The Method of Newton

The problems inherent in the method of steepest descent led to the consideration of new methods of optimization

(Harvey[1]). It was soon realized that approximating the function by a quadratic expansion could form the basis of relatively efficient computational schemes. Such schemes employ second, as well as as first derivatives. The basic procedure obtained from this approach is known as Newton-Raphson or simply Newton's Method.

The criterion on objective function expanded about $\bar{p}$ (the minimum) follows

$$f(p) = f(\bar{p}) + (p - \bar{p})\, g(\bar{p}) + \tfrac{1}{2}\,(p-\bar{p})\, G(\bar{p}) \qquad (5)$$

Differentiating with respect to p yields.

$$g(p) = g(\bar{p}) + (p - \bar{p})\, G(\bar{p}) \qquad (6)$$

Since $g(\bar{p}) = 0$ this implies that

$$\bar{p} \doteq p - G^{-1}(\bar{p})\; g\;(p) \qquad (7)$$

where $\doteq$ denotes approximately equal to:

$\tilde{p}$ is usually unknown. The solution adopted in Newton-Raphson is to evaluate the Hessian at the current estimate. $\hat{p}$ on th grounds that this will yield an acceptable approximation to $G(\bar{p})$ if $\hat{p}$ is reasonably close to $\bar{p}$. This iterative scheme becomes:

---

(1) Ibid

$$p^* = \hat{p} - G^{-1}(\hat{p})\; g\;(\hat{p})$$

which is sometimes written as:

$$p^* = \hat{p} - G^{-1}g \qquad (9)$$

where it is understood that G and g are evaluated at the current estimate.

Modifications

The Newton-Raphson scheme will only progress towards a minimum if $G(\hat{p})$ is positive definite. While it is true that the Hessian is always positive it is still possible for $f(p^*)$ to exceed $f(\hat{p})$. However, this represents a case of 'overshooting' and such an occurrence can always be avoided by introducing a variable step length, L' into the scheme. Hence

$$p^* = \hat{p} - L'\; G^{-1}g \qquad (10)$$

A number of techniques have been devised for modifying the basic Newton-Ralphson method so as to ensure that the gradient is always premultiplied by a positive definite matrix.

The iterative scheme for such methods can be written as:

$$p^* = \hat{p} - L'\; H\; g \qquad (11)$$

where $H = H(\hat{p})$ is an n x n positive definite matrix. In the method of quadratic Mill climbing proposed by Goldfeld, Quandt and Trotter (1966). H is set equal to

$$(G + mI)^{-1} \qquad (12)$$

The positive scaler m varies as the iterations proceed its value being chosen as to ensure that H is positive. When G is negative definite, this may be guaranteed by setting m to a value greater than the modulus of the largest eigen-value of G. When G is singular any positive value of m ensures that H is positive definite. If m is large, the iteration will be similar to the method of steepest descent. However, as the minimum is approached G will tend to become positive definite, and in this case m may be set eual to zero so that the method collapses to Newton-Rophson. Although quadratic hill climbing is a useful technique in many problems, its efficiency is likely to fall markedly as n increases, since the matrix inversions and eigenvalue evaluations will impose a heavy computational burden.

QUASI-NEWTON

The main advantage of Quasi-Newton on variable metric method is that they do not require the Hessian to be explicitly evaluated. The iterative scheme is given in equation 11. At each iteration H is updated in such a way as to yield a series of positive definite matrices that eventually converge to the inverse of the Hessian. A common choice for the starting matrix is the identity matrix. The first iteration is therefore carried out bythe methods of steepest descent, but the procedure gradually tends towards Newton-Raphson as the minimum is approached. An example of a Quasi-Newton algorithm is the Davidson-Fletcher Powell algorithm (Ammemiya (1983)).

The explicit form of the criterion function has not yet been specified. In the literature on nonlinear functions, two criterion have been usually specified. The log-likelihood function and the sums of squares.

The solution of the normal equations in the case of the sum of squared errors as criterion function in general yield estimations which are not a linear function of the sample y. This however is a crucial condition used to show unbiasedness of linear least-squared estimator. Consequently the nonlinear least square estimator is neither linear nor unbiased in general and hence it is not BLUE (best linear unbiased); in fact,it is difficult to derive its small sample properties. However, provided that the errors have zero mean and are indepedently identically distributed with variance s^2, it can be shown that even if the error distribution is non normal the nonlinear least-square estimator of the parameter vector is consistent and asymptotically normally distributed, (Judge et al 1982). The consistency of the least-square estimator implies that, given a large sample, the estimated vector is likely to be close to the true parameter. As a result, the Newton-Raphson procedures will give the same results (provided the sample is large enough and certain conditions on the disturbance vector hold) whether the criterion function is the log-likelihood or the sums of squared errors.

GAUSS NEWTON

The Gauss-Newton method adopts the sum of squared errors as its criterion

The derivation follows:

If $f(p) = s(p) = \Sigma e_t^2$ (13)

then

$$g(p) = \frac{ds(p)}{dp} = 2 \Sigma \frac{de_t}{dp} e_t \qquad (14)$$

while the Hessian is

$$G(p) = \frac{d^2 s}{dp dp'} (p)$$

$$= 2 \Sigma \left\{ \frac{de_t}{dp} \frac{de_t}{dp'} + \frac{d^2 e_t}{dp} e_t \right\} \qquad (15)$$

using (9) we have

$$p^* = \hat{p} \left[\Sigma \frac{de_t}{dp} \cdot \frac{de_t}{dp'} + \frac{d^2 e_t}{dp dp'} e_t \right]^{-1} \frac{de_t}{dp} \cdot e_t \Big\} \qquad (16)$$

Given that the terms involving the second derivatives will be small, this results.

$$p^* = \hat{p} - \left[\Sigma \frac{de_t}{dp} \cdot \frac{de_t}{dp'} \right]^{-1} \Sigma \frac{de_t}{dp} e_t \qquad (17)$$

This is the Gauss Newton

If we write

$$Z_t = - \frac{de_t}{dp} \qquad (18)$$

this implies that

$$p^* = \hat{p} + (\Sigma Z_t Z_t')^{-1} \Sigma Z_t e_t \qquad (19)$$

A slightly different scheme is due to Marquadt (Ammemiya)[2] is

$$p^* = \hat{p} + (\Sigma Z_t Z_t' + mI)^{-1} \Sigma Z_t e_t \qquad (20)$$

This study employs the Gauss-Newton method. After converting the Box-Cox problem to a nonlinear least-square problem the Gauss-Newton technique is applied to estimate the parameters.

SECTION 2

The Box-Cox Transformation

The Box-Cox transformation (Box-Cox (1964) provides a method of choosing among a family of competing models not necessarily having the same dependent variables.

The Box-Cox procedure utilizes the following definitions.

(2) Ibid

The power transformation of the randem variable z

$$Z^{(L)} = \frac{Z^L - 1}{L}, \quad L \neq 0 \qquad (21)$$

$$\mathrm{Ln}Z, \quad L = 0$$

$$Z^{(o)} = \mathrm{Ln}Z \text{ as } \lim_{L \to o} [(Z^L - 1)/L] = \mathrm{Ln}Z$$

by L'Hopital's rule.

In general the model is

$$Y_t^{(L_1)} = B_1 + B_2 X_{t2}^{(L_2)} + \ldots\ldots\ldots t B_k X_{tk}^{(L_k)} + e_t$$

$$= f(x_t, B, L_{(2)})$$

denote $L_{(2)} = \{L_2, \ldots\ldots\ldots\ldots L_k\}$ $\quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_k \end{bmatrix}$ (22)

and $L = \{L_1, \ldots\ldots L_k\}$

The maximum likelihood estimates for B and s^2 (the variance) can be solved for conditional on values for $L_1, L_2 \ldots\ldots\ldots L_k$ (Fomby Hill and Johnson, 1984)

$$\hat{B}(L) = \left(X^{(L_{(2)})'} X^{(\hat{L}_{(2)})}\right)^{-1} X^{(\hat{L}_{(2)})'} Y^{(L_1)} \qquad (23)$$

$$\hat{s}^2 = \left(Y^{(L_1)} - X^{(L_{(2)})} \hat{B}(L)\right) \left((Y^{(L_1)} - X^{(L_{(2)})} \hat{B}(\_)\right)/T \qquad (24)$$

where $Y^{(L_1)} = \left(Y_2^{(L_1)}, \ldots\ldots\ldots\ldots, Y_T^{(L_1)}\right)'$

and $X^{(L_{(2)})} = \left(i, X_2^{(L_2)}, X_3^{(L_3)}, \ldots\ldots\ldots, X_k^{(L_k)}\right)$

where i is a TX1 vector of ones and $X_2^{(L_2)}, \ldots\ldots\ldots X_k^{(L_k)}$ are each TX1 vectors of observations on $X_{t2}^{(L_1)}, \ldots\ldots\ldots X_{tk}^{(L_k)}$ respectively.

The error vector e is assumed to have the following properties.

$$E = (e) \qquad E(ee') = \hat{s}I$$

The concentrated log likelihood junction is then

$$L(L_1, L_2, \ldots\ldots\ldots\ldots L_k / Y, X)$$

$$= \text{Const} + (L_1 - 1) \sum_{L=1} \mathrm{Ln}Y_t - \frac{1}{2} \mathrm{Ln}\acute{s}^2(L) \qquad (25)$$

On the assumption that $L_1 = 1$ the Jacobian term falls out of the analysis, and the implication here is that we can use non linear least squares.

This is similar to the approach used by Spitzer (1982). Spitzer considers models with $L_2 = L_3 = \ldots\ldots L_k = 0$ and L_1 to be determined by the data. By using a scaling method due to Zarembka (1968) he reduces the estimation problem to one of nonlinear least square estimation. The Jacobian term vanishes. Spitzer assures us that if $\{l_2\ l_3 \ldots\ldots l_4\} \neq$ zero vector, we can still use this scaling technique with models more complicated than the model he used for expository purposes. This appears feasible given that the Jacobian term does not involve the parameters in the design matrix. The estimated parameters however would need to be transformed back to their original parameters (parameters before scaling technique is applied).

Care must be taken when interpreting results. Spitzer (1984) demonstrates the effect of scaling on hypothesis testing . Scaling can effect the significance or insignificance of parameter estimates. Spitzer argues that if one uses the Box-Cox scaling technique to compare two models the same scaling factor for the dependent should be used. Merely shifting the decimal point of the dependent variable affects hypothesis testing.

The approach used here is based on considerations of simplicity. It is assumed that $L_1 = 1$ and $L_2 = L_3 = L_4 = \ldots L_k = L$ This type of approach is hinted at in Judge [(3)]. He says "it may not be clear whether there is a relationship between a dependent variable y and an independent variable x or the logarithm of x. This uncertainty can be accounted for by using a Box-Cox transformation of x. The kind of model referred to by Judge follows:

$$y_t = B_1 + B_2 \frac{(X_t - 1)}{L} + e_t \qquad (26)$$

SECTION 3

Application

The model to be examined for nonlinearities is the non-tradable product market. Worrell (1984) describes output in this sector as being affected both by demand and supply; a problem of simultaneity therefore arises. McClean (1979) sees this sector however, as having considerable excess capacity and hence being demand side determined. The demand side is therefore the short side of the market. This latter assumption is adopted here. However, the partial adjustment mechanism utilized by Worrell is employed. The model can be

---

(3) Judge, Hill, Griffiths, Lütkempoml, Lee

written as

$$BQN = f(BYR, BPN/BPT, BQN(-1))$$

In the original model Worrell included the loan interest rates. The variables in the relationship were all expressed on their logarithms. His analysis concludes that insensitivity to relative prices limits the scope for expenditure switching policies. The study by Worrell also concludes that interest rates had no effect.

This paper is concerned with the validity of alternative functional forms for the data. It questions the validity of claims based on incorrect functional specifications.

RESULTS OF ESTIMATION

TABLE ONE

Dependent Variable BQN

From 59 - 1		UNTIL 82 - 1	
obs 24		Deg. of freedom	19
R-Squared	.9895	$\bar{R}$ - Squared	.9873
Durbin Watson 1.5734			
SSR	5600.0861	SEE	

Parameters	Values	T-Statistics
B1	-1233.9390	- 2.4133
B2	19.8410	.8546
B3 or L	.3949	1.7619
B4	28.8709	.7903
B5	-257.0300	-1.0755

NB t_{19} (1%) = 2.861 t_{19} (5%) = 2.093

TABLE 2

Linear Pseudo Model

$$xi_t = \left. \frac{df(x_t, B1, B2, B3, B4)}{dB_i} \right|_{\hat{B}}$$

where

$$x_t = (1, BQN(-1_t), BYR_t \ (BPN/BPT)_t)$$

and

Bi refers to the parameters in Table one

f is the right hand side of 22 with $L = L_2 = L_3 = L_k$

L_1 is set equal to one

BQN(-1)	is lagged nontradable output sometimes written BQNL
BYR	is real output at time t
BPN/BPT	The opportunity cost of nontradable output in terms of tradable output sometimes written temp.
BPN	Price index of nontradable goods
BPT	Price index on tradable goods
B	is estimated parameter vector from the non linear least-square estimation.

$$Y_t = BQN_t + \sum xi_t \, Bi/_{\hat{B}} - f(x_t, \hat{B})$$

Dependent variable	Y		
Obs	24	Degrees of Freedom	19
R-Squared	.9914	R-squared	.9869
SSR	5690.7241	SEE	17.3064
Durbin Watson	1.5502		

Parameters of	Value	T-Statistics
Constant	-2011.786	-.3545
x2	21.32403	1.8341
x3	.3675	2,1844
x4	30.55088	2.0552
x5	-149.2229	- .5025

N.B. t_{19} (1 %) = 2.861 and t_{19} (5%) = 2.093

TABLE 3

LBQNL, LBYR, L(BPN/BPT) denote the logarithms of the respective quantities.

From	59 - 1	until 82 - 1	
obs	24	Degrees of Freedom	20
R - squared	.9876	R - squared	.9857
SSR	6612.1161	SEE	18.1826
Durbin Watson	1.3315		

Parameters of	Value	T-statistics
Constant	-2449.471	- 10.5805
LBQNL	140.8392	2.9429
LBYR	263.6263	4.4941
* LTEMP	- 78.2895	-1.9278

Calculated F-Statistic		Tabulated
		5% Level
F^1_{19} = 11.19740	L = 1	F^1_{19} = 4.38
F^1_{19} = .6278	L = 0.5	1% Level
		F^1_{19} = 8.18

The above results show that the parameter L is significantly different from one at the five percent level. The hypothesis of equality of the parameter value with 0.5 was not rejected at the five percent level. The hypothesis of equality with zero was not rejected as well. This statistic is not reported here however a brief comparison of the reported F's (in an asymptotic sense of course) and the statistics calculated on the basis of the square of the student t test using information in Table one corroborates this. For example:

$$\frac{(.3449 - 1)^2}{0.1975} = 11.0022$$

and

$$\frac{(0.3449)}{(0.1975)^2} = 3.0497$$

and

$$\frac{(0.3449 - 0.5)^2}{0.1975} = .6167$$

The F-test reported is more reliable than the t statistic since hyothesis testing using this particular software (RATS) is based upon a quadratic approximation to the likelihood surface.Curvature information is employed and problems of under or over estimation of variances because of these first derivative methods are not as harsh.

Table 3 shows the semilog model. Choice of this model was informed by the hypothesis test. The model has the right signs and relative proces seem to have more importance. Moreover, a cox test applied to the semilog and linear, shows that the semilogarithmic form is better.

Results of Cox Test (R.E. Quandt, 1974)

First Test

$H_1: \; Y = XB_1 + U_1$

vs

$H_2: \; Y = ZB_2 + U_2$

$X_t = (1, \; BQN\;(-1)_t, \; BYR_t \; (BNP/BPT)_t$

$Z_t = \{1, LBQN\;(-1)_t, \; LBYR_t, \; L(BPN/BPT)_t$

$T_1 = -4.6340 \quad V_1 = 5.8311$

$d_{1x} = \dfrac{T_1}{V_1^{\frac{1}{2}}} = -1.9190 \quad N_{o.o5} = 1.96$

Second Test

$H_1: y = ZB_1 + U_1$

$H_2: Y = B_2 + U_2$

$T_1 = -2.3485 \quad V_1 = 5.8774$

$d_{1z} = -.9687$

This represents what one may indulge in describing as almost unqualified acceptance of the semilog formal Fisher and McAleer (1980). It should be noted that if one adheres however, to a strict interpretation of the statistical test of hypothesis, one would not reject the null hypothesis on the strength of the d1x statistic. On the face of the following results, some degree of subjective preference is however reasonable.

TABLE 4

Dependent variable BQN

From	59 - 1	Until 82 - 1	
Observations	24	Degrees of Freedom	20
R - squared	.9834	R - squared	.9809
SSR	8842.4991	SEE	21.0268
Durbin Watson	1.2129		

Parameters	Value	T. Statistics
Constant	-26.72155	-.6165
BQNL	.4394	3.322441

BYR	.4293	4.2348
BPN/BPT	826.7799	.1841

SECTION 4

The literature does not record that substantial an attempt to model the convariance structure of simple non-linear models.

One possible scheme is the use of locally equivalent alternative models. These functional forms are described by Godfrey and Wickens (1982).

If we write

$$Y_t = X_t' B + a_1 \hat{u}_{y-1} + \ldots a_G \hat{u}_{t-G} + e_t$$

we can test variance convariance structures using an LM statistic. In order to test for the null hypothesis of serial independence against that of autoregressive or moving average disturbance structure one evaluates the test statistic $K^2(G)$ which is G times the F statistic for the joint significance of $a_1\ a_2\ a_3, \ldots, a_G$. This test statistic is distributed as a chi-square with G degrees of freedom, with $\hat{u}_t$ being the residuals obtained from the maximum likelihood estimation of

$$Y_t = x_t' B + U_t \tag{27}$$

This technique is applied to the linear pseudo model on the assumption that the covariance structure for the disturbance in the linear pseudomodel (Malinvaud 1977) is no different to that in the actual model. An examination of how the linear pseudomodel is calculated implies that this assumption is not likely to be unreasonable. The linear pseudo model is already an approximation in itself.

Values of the Test statistic for the first second and third order statistics under the linear pseudo-model is

G X F (G, D) where D is the number of degrees of freedom (T-K) where T is the number of observations and K the number of data points. F is the calculated F statistic.

(a) 1 x (1.8952) = 1.8952

2 x .8624 = 1.7248

3 x .7572 = 2.2716.

Now the Chi-square statistic at the 1% and 5% level with G degrees of freedom is

DeG of Freedom	Tabulated	Value of Chi-square
1 F(1,17)=1.8951	5% 3.8415	1% 6.6350
2 F(2,15)= .8624	5% 5.99146	1% 9.21034
3.F(3,B) = .7572	5% 7.81473	1%11.3449

The null hypothesis of serial independence is accpeted at the 1 and 5 per cent levels.

If we write the problem in nonlinear form as in equation 22 then the null hypothesis of serial independence against that of an autoregressive or moving average disturbance process can be performed again by calculating the lagrange multiplier statistic as T times the R-squared of the regression of $\hat{e}_t$ an

$$H_t, \quad \hat{e}_{t-1}, \ \ldots\ldots\ldots, \ \hat{e}_{t-G}$$

where $H_t = \dfrac{df\ (x_t, B, L)}{dA}$

where $A = \{B, L\}$

The evaluated statistic was

R - squared		T	TR^2
AR(1)	.0348	23	.8010
AR(2)	.0342	22	.8010
AR(30	.0512	21	1.0762

A comparison with the tabulated chi-square values shows that the hypothesis of serial independence is accepted at the one and five percent level.

CONCLUSION

We set out to show the importance of the correct functional forms in the design matrix. We demonstrated this with the Box-Cox transformation shown.

Although the Cox test does not reject the linear functional specification, there is some indication -- in the light of our serial independence hypothesis - that the semilog is preferable. Standard errors are lower and R-squares higher with this model. The descriptive statistics on the whole give better performance with the semilog.

Much of the analyses has been couched in asymptotic results. While there is a better chance of bringing down the sparrow with a cannon than with a pistol, we don't want to spend that long rolling the cannon into place to find that the bird is gone. However exact distribution theory in the context of nonlinear theory does not seem to be very practical at this time.

Any performance of nonlineal estimators should involve stochastic simulation. We should determine whether L =0 is therefore a better fit than L = 0.5.

PRINCIPAL DATA USED IN STUDY

APPENDIX

X2

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	14.038494	15.069897	14.667168	14.575254
63-	1	14.886664	15.856491	16.458094	17.085111
67-	1	18.054817	18.670730	19.407399	20.713387
61-	1	21.710302	20.987853	22.284613	22.590883
75-	1	22.211903	21.949144	22.128790	22.190344
79-	1	22.266473	22.838236		

X4

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	18.399904	18.230999	18.433833	18.485659
63-	1	19.623283	19.791194	20.798766	21.327607
67-	1	21.999760	22.307665	23.099782	24.100551
61-	1	23.870567	25.063824	25.299680	25.084073
75-	1	24.881686	25.252387	25.640098	26.097046
79-	1	26.871609	27.361633		

X5

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	-23.340914	-2.334432	-2.319699	-2.315004
63-	1	-2.330397	-2.337629	-2.299837	-2.308583
67-	1	-2.334637	-2.311159	-2.312469	-2.299931
61-	1	-2.269742	-2.280950	-2.277670	-2.291423
75-	1	-2.307047	-2.280711	-2.274456	-2.271969
79-	1	-2.295557	-2.326734	-2.328620	2.328588

X3

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	-10174.118720	-9874.761740	-9583.410642	-9495.783450
63-	1	-9900.675192	-10115.366362	-9530.703828	-9762.861826
67-	1	-10366.353651	-9982.114003	-10114.400208	10053.735633
61-	1	-9612.012781	-9840.949280	-9884.520508	-10102.214179
75-	1	-10328.451464	-9910.846300	-9854.987766	-9859.282862
79-	1	-10301.481826	-10934.969637	-10981.533681	-10953.133031

ZBBAR

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	-4537.119000	-4414.640000	-4312.494000	-4281.390000
63-	1	-4385.980000	-4437.801000	-4185.398000	-4240.014000
67-	1	-4416.223000	-4256.546000	-4265.026000	-4186.074000
61-	1	-4012.814000	-4074.544000	-4056.192000	-4134.962000
75-	1	-4230.375000	-4073.209000	-4038.443000	-4026.871000
79-	1	-4159.586000	-4360.612000	-4373.718000	-4373.498000

FOFB

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	180.067228	191.418585	185.497137	183.963091
63-	1	226.942366	252.891105	284.203246	314.159977
67-	1	359.502017	374.577391	412.399359	463.981648
61-	1	469.362255	492.359255	524.054511	527.441171
75-	1	518.094750	517.680836	529.964978	543.739372
79-	1	573.674971	607.180054	611.104170	601.511651

Y

ANNUAL DATA FROM 59 1 TO 82 1

59-	1	-4519.086228	-4420.558585	-4315.291137	-4273.053091
63-	1	-4388.622366	-4444.892105	-4200.001246	-4244.873977
67-	1	-4439.325017	-4260.323391	-4240.125359	-4157.055648
61-	1	-4029.976255	-4039.803255	-4034.346511	-4139.703171
75-	1	-4241.469750	-4073.189836	-4047.007978	-4044.610372
79-	1	-4171.860971	-4369.292054	-4371.922170	-4362.609651

ANNUAL DATA FROM 58 1 TO 82 1

BYR

58-	1	280.300000	324,300000	316·900000	325·800000
62-	1	328.100000	381·300000	389·600000	441,900000
66-	1	471.00000	510·000000	528·500000	578.100000
70-	1	645.000000	629·200000	714·000000	731.600000
74-	1	715.500000	700·600000	730·300000	757.500000
78-	1	793.200000	856.200000	897·700000	874.200000
82-		836.700000			

ANNUAL DATA FROM 58 1 TO 82 1

BPN

58-	1	.255000	.320000	.333000	.325000
62-	1	.350000	.388000	.343000	.377000
66-	1	.397000	.364000	.393000	.407000
70-	1	.470000	.499000	.536000	.649000
74-	1	.880000	1.000000	1.139000	1.272000
78-	1	1.347000	1.347000	1,437000	1,523000
82-	1	1.543000			

ANNUAL DATA FROM 58 1 TO 82 1

BPT

58-	1	.40000000	.40000000	.38200000	.34600000
62-	1	.36400000	.43600000	.40000000	.36400000
66-	1	.40000000	.41800000	.40000000	.41800000
70-	1	.45400000	.41800000	.47300000	.56400000
74-	1	.81600000	1.00000000	1.00400000	1.08900000
78-	1	1.14000000	1.27400000	1.58500000	1.69600000
82-	1	1.71800000			

ANNUAL DATA FROM 58 1 to 82 1

BQN

58-	1	166 900000	198.100000	185 500000	182.700000
62-	1	192 300000	224 300000	245.800000	269.600000
66-	1	309 300000	336 400000	370,800000	437,300000
70-	1	493 000000	452.200000	527.100000	545,900000
74-	1	522 700000	507 000000	517.500000	609.900000
82-	1	612 400000			

BIBLIOGRAPHY

AMMEMIYA T. (1983) "Handbook in Econometrics" Edited by Zvi Griliches and Michael D. Intriligator, North Holland Publishing Company

BOX G.E.P. and COX D.R. (1964) An Analysis of Transformations Journal of the Royal Statistical Society

CHOW, Gregory, CORSI Paolo (1982) "Evaluating the Reliability of Macroeconomic Models" John Wiley and Sons

FISHER, GORDAN, McALEER (1979) "On the Interpretation of the Cox Test inEconometrics." Economic Letters

FOMBY, Thomas B., Carter-Jill, Johnson, Stanley R. "Advanced Econometric Methods" Springer-Verlag

GOLDFELD, S.M. Quandt, R.E., Trotter H.F. (1966) "Maximization by Quadratic Hill-climbing" Econometrica

JUDGE G, Griffiths William E, Carter Mill, R. Tsoung-Cha Lee, John Wiley and Sons

PESARAN, M.M. and Deaton, A.S. (1978) "Testing Non-Nested Nonlinear Regression Models," Econometrica

SPITZER (1982) "A Primer on Box-Cox Estimation" Review of Economic Studies

SPITZER, John J. (1984) "Variance Estimates in Models with the Box-Cox Transformation. Implications for Estimation and Hypothesis Testing.

ZAREMBKA, P. (1968) "Functional Form in the Demand for Money" Journal of the American Statistical Assoc.

ZAMEMBKA P. (1974) Transformation of Variables in Econometrics. Edited by P. Sarambka