

Division of Research
Graduate School of Business Administration
The University of Michigan

June 1979

COMPARING REGRESSION MODELS
INVOLVING RATIO VARIABLES

Working Paper No. 180

Roger L. Wright

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

None of this material is to be quoted or
reproduced without the express permission
of the Division of Research.

COMPARING REGRESSION MODELS INVOLVING RATIO VARIABLES

Roger L. Wright
The University of Michigan

1. Introduction

Regression models are often specified to include ratio variables such as household consumption per member, regional income per capita, savings as a fraction of income, and profit of a firm per unit of its assets or sales. In cross sectional studies, the motivation for utilizing ratio variables is usually to eliminate the influence of size to better isolate non-size effects, (Kuh and Meyer, 1955). In the analysis of time series data, ratio variables arise in adjusting for seasonality and inflation.

Ratio variables are also used to implement weighted least squares estimation for heteroscedastic regression models, (Johnston, 1972, pp. 214-221.) As Kuh and Meyer emphasized, the Gauss-Markov Theorem for heteroscedastic linear models provides a systematic and rigorous basis for introducing ratio variables in certain circumstances. Various authors, including Fisher (1957), Glejser (1969), Harvey (1976), Park (1966), Rutenmiller and Bowers (1968), and Taylor (1978), have discussed the estimation of heteroscedastic models, but they have not been developed with sufficient generality to encompass the common and useful practice of utilizing both deflated and unde-flated independent variables in the same model.

In this paper we will examine a very broad family of nonlinear

regression models which generalize Harvey's multiplicative model for heteroscedasticity to allow deflation (or inflation) of independent variables. Rao's scoring method (Rao, 1973, pp. 367-70) will be used to develop a convenient procedure for calculating maximum likelihood estimators.

There has always seemed to have been confusion about the relationship of correlation to the use of ratio variables. Several authors, including Pearson (1897), Kuh and Meyer (1955) and Madansky (1964), have discussed the spurious correlation that may arise from using a common variable to deflate both dependent and independent variables. Other writers have warned against using multiple correlation coefficients to compare the goodness of fit of models with alternative transformations of the dependent variable. (On the other hand Granger and Newbold [1976] find that maximizing R^2 can be an appropriate strategy for analyzing a family of transformations of both dependent and independent variables.) We will develop here a simple way of using the standard errors of the regressions to compare the likelihood of alternative specifications involving heteroscedasticity or deflated dependent variables.

Our methods of model specification and estimation will be illustrated by reanalyzing Mooz's (1978) capital cost data for nuclear power plants. Some of the issues of concern here were raised in a rather similar setting by Griliches (1972).

2. Model and Notation

Throughout this paper we will examine the following nonlinear, heteroscedastic regression model:

$$Y_i = \sum_{j=1}^m e^{Z'_{ji}\alpha_j} X'_{ji}\beta_j + u_i \quad (1 \leq i \leq n) \quad (1a)$$

Here $\{Z'_{ji}, X'_{ji}, Y_i: 1 \leq i \leq n, 1 \leq j \leq m\}$ are observed data, and $\{\alpha_j, \beta_j: 1 \leq j \leq m\}$ are parameters. In general we consider Z_{ji} and α_j to be vectors of size $p_j \times 1$, and X_{ji} and β_j to be of size $q_j \times 1$.

The following assumptions are maintained:

- (i) The disturbances $\{u_i: 1 \leq i \leq n\}$ are mutually independent and normally distributed with mean zero and variance

$$E(u_i^2) = \sigma_e^2 e^{2Z'_{1i}\alpha_1} \quad (1b)$$

with Z_{1i} and α_1 as in the initial term of the summation in (1a), and $\sigma_e^2 > 0$ an additional parameter.

- (ii) The parameters $\{\alpha_j: 1 \leq j \leq \ell\}$, $\{\beta_j: 1 \leq j \leq m\}$, and σ_e^2 are to be estimated, while $\{\alpha_j: \ell + 1 \leq j \leq m\}$ are prespecified.

- (iii) The variables Z_{ji} and X_{ji} are either nonstochastic or stochastic but distributed independently of the u_i with a distribution not involving the parameters α_j , β_j , and σ_e^2 .

- (iv) $\sum_{i=1}^n Z_{1i} = 0$. The importance of this assumption, which imposes no real lack of generality, will become clear in Section 3.

- (v) The parameters to be estimated are identifiable.

Some additional notation will simplify subsequent analysis:

$$X_{ji}^* = e^{Z_{ji}'\alpha_j - Z_{1i}'\alpha_1} X_{ji} \quad (1 \leq j \leq m) \quad (2a)$$

$$Z_{ji}^* = (X_{ji}^{*\prime} \beta_j) Z_{ji} \quad (1 \leq j \leq \ell) \quad (2b)$$

$$Y_i^* = e^{-Z_{1i}'\alpha_1} Y_i \quad (2c)$$

$$\epsilon_i = e^{-Z_{1i}'\alpha_1} u_i \quad (2d)$$

In addition we will use the following partitioned vectors and matrices:

$$\underline{X} = \begin{bmatrix} Z_{11}^{*\prime} & Z_{21}^{*\prime} & Z_{\ell 1}^{*\prime} & X_{11}^{*\prime} & X_{m1}^{*\prime} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Z_{1n}^{*\prime} & Z_{2n}^{*\prime} & Z_{\ell n}^{*\prime} & X_{1n}^{*\prime} & X_{mn}^{*\prime} \\ \hline \sqrt{2}\alpha_{11}' & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sqrt{2}\alpha_{1n}' & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3a)$$

$$\underline{\beta}' = [\alpha_1' \mid \alpha_2' \cdots \alpha_\ell' \mid \beta_1' \cdots \beta_n'] \quad (3b)$$

$$\underline{\epsilon}' = \left[\epsilon_1 \cdots \epsilon_n \mid \frac{\epsilon_1^2}{\sqrt{2}\sigma} \cdots \frac{\epsilon_n^2}{\sqrt{2}\sigma} \right] \quad (3c)$$

The dimension of \underline{X} , $\underline{\beta}'$, and $\underline{\epsilon}'$ are respectively $(2n \times r+s)$, $(1 \times r+s)$, and $(1 \times 2n)$. Here $r = \sum_{j=1}^{\ell} p_j$ and $s = \sum_{j=1}^m q_j$. In section 3 we will often write ϵ_i , $\underline{\epsilon}$ and \underline{X} as $\epsilon_i(\underline{\beta})$, $\underline{\epsilon}(\underline{\beta}, \sigma^2)$, and $\underline{X}(\underline{\beta}, \sigma^2)$ to emphasize their functional dependence on the parameters.

Three special cases of this model are of special interest:

(i) If $\alpha_1 = 0$ and $\ell = 1$ so that all of the deflation coeffi-

coefficients $\alpha_2, \dots, \alpha_m$ are prespecified, we obtain a homoscedastic linear regression model involving q_1 independent variables represented by the vector X_{1i} and prespecified deflated independent variables given by $X_{2i}^*, \dots, X_{mi}^*$. In the most common applications, $p_j = 1$, Z_j is the logarithm of a measure of size, and $\alpha_j = -1$ ($2 \leq j \leq m$). Alternatively if $\alpha_j = 1$ we obtain an interaction variable.

- (ii) If $\alpha_1 = 0$ and $\ell \geq 2$, we have a nonlinear, homoscedastic regression model which generalizes an example discussed by Draper and Smith (1966, pp. 266-284). In this case the degree of deflation of X_2, \dots, X_ℓ is to be estimated from the data.
- (iii) If $X_{1i} = 0$ and $\ell = 1$, we have the linear model with multiplicative heteroscedasticity. Harvey (1976) discusses estimation of α_1 and β_j ($2 \leq j \leq m$) in this case. (β_1 is obviously unidentifiable and is taken to be zero.) Kmenta (1971, pp. 256-264) also discusses this case but with the added restriction that $p_1 = 1$.

Furthermore if α_1 is known, it is easy to see that this case can be reformulated as a homoscedastic linear model in the deflated variables Y^*, X_2^*, \dots, X_m^* . This is the familiar case of weighted least squares.

3. Estimation

To study estimation, we reformulate the model (1a-b) using the definitions (2a-d) as

$$Y_i^* = \sum_{j=1}^m X_{ji}^* \beta_j + \epsilon_i$$

Here the $\{\epsilon_i : 1 \leq i \leq n\}$ are mutually independent and normally distributed with mean zero and variance σ^2 .

The transformation from Y_i to Y_i^* involves parameters $\underline{\beta}$ to be estimated. In such a case the likelihood function of the sample can be written as the product of the probability density of the $\{Y_i^* : 1 \leq i \leq n\}$ and the Jacobian J of the transformation from $Y_1 \cdots Y_n$ to $Y_1^* \cdots Y_n^*$, (Box and Cox, 1964, p. 215). From (2c) and assumption (iv) it is easy to see that

$$\begin{aligned} J &= \prod_{i=1}^n \left| \frac{dY_i^*}{dY_i} \right| \\ &= e^{-\sum_{i=1}^n Z_{1i}' \alpha_1} \\ &= 1 \end{aligned}$$

This will provide a very convenient simplification to our analysis.

The likelihood functions $L(\underline{\beta}, \sigma^2)$ can be expressed as

$$L(\underline{\beta}, \sigma^2) = (\sqrt{2\pi} \sigma)^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n \epsilon_i(\underline{\beta})^2}$$

so that

$$\log L(\underline{\beta}, \sigma^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \epsilon_i(\underline{\beta})^2 \quad (4)$$

where, using (1a) and (2d) we write

$$\epsilon_i(\underline{\beta}) = e^{-Z'_{li}\alpha} (Y_i - \sum_{j=1}^m e^{Z'_{ji}\alpha} X'_{ji}\beta_j) \quad (5)$$

From (4) it is clear that the maximum likelihood estimators of $\underline{\beta}$ and σ^2 , say $\hat{\underline{\beta}}$ and $\hat{\sigma}^2$, have some of the properties of ordinary least squares estimators:

$$(i) \quad \hat{\underline{\beta}} \text{ must minimize } \sum_{i=1}^n \epsilon_i(\underline{\beta})^2 \quad (6a)$$

$$(ii) \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_i(\hat{\underline{\beta}})^2, \text{ and} \quad (6b)$$

$$(iii) \quad \log L(\hat{\underline{\beta}}, \hat{\sigma}^2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\hat{\sigma}^2) - \frac{n}{2} \quad (6c)$$

Rather than attempting to directly solve the nonlinear normal equations associated with (6a), we develop an iterative estimation procedure utilizing Rao's method of scoring (Rao, 1973, pp. 367-70). For this we consider a trial value $\underline{\beta}_0$ and let σ_0^2 be defined from (6b):

$$\sigma_0^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_i(\underline{\beta}_0)^2.$$

We define the efficient score $S(\underline{\beta}_0)$ and the information $I(\underline{\beta}_0)$ to be:

$$S(\underline{\beta}_0) = \frac{\partial}{\partial \underline{\beta}'} \log L(\underline{\beta}_0, \sigma_0^2), \text{ and} \quad (7a)$$

$$I(\underline{\beta}_0) = E\left[-\frac{\partial^2}{\partial \underline{\beta} \partial \underline{\beta}'} \log L(\underline{\beta}_0, \sigma_0^2)\right] \quad (7b)$$

Then the revised estimate $\underline{\beta}_1$ is calculated as

$$\underline{\beta}_1 = \underline{\beta}_0 + I(\underline{\beta}_0)^{-1} s(\underline{\beta}_0) \quad (7c)$$

This procedure is continued until convergence of σ_0^2 . The resulting maximum likelihood estimator $\hat{\underline{\beta}}$ is asymptotically normal with mean $\underline{\beta}$ and covariance matrix $I(\underline{\beta})^{-1} \sigma^2$.

Because of the special form of the model that we are considering, the method of scoring is computationally quite convenient. From (5) it is easily seen that

$$\frac{\partial}{\partial \alpha_j'} \varepsilon_i(\underline{\beta}) = -Z_{ji}^* - \delta_{j1} \varepsilon_i Z_{1i}' \quad (1 \leq j \leq \ell)$$

$$\frac{\partial}{\partial \beta_j} \varepsilon_i(\underline{\beta}) = -X_{ji}^* \quad (1 \leq j \leq m)$$

so that (4) implies

$$\frac{\partial}{\partial \beta_j'} \log L(\underline{\beta}, \sigma^2) = \frac{1}{\sigma^2} \left(\sum_{i=1}^n \varepsilon_i Z_{ji}^* + \delta_{j1} \sum_{i=1}^n \varepsilon_i^2 Z_{1i}' \right) \quad (1 \leq j \leq \ell)$$

$$\frac{\partial}{\partial \beta_j} \log L(\underline{\beta}, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^n \varepsilon_i X_{ji}^* \quad (1 \leq j \leq m)$$

Here δ_{j1} is defined to be one if $j = 1$ and zero for other values of j .

From (3a-c) we find that

$$s(\underline{\beta}_0) = \underline{X}(\underline{\beta}_0, \sigma_0^2)' \underline{\varepsilon}(\underline{\beta}_0, \sigma_0^2) \quad (8a)$$

In a similar fashion one can confirm that

$$E\left[-\frac{\partial^2}{\partial \alpha_j' \partial \alpha_k'} \log L(\underline{\beta}, \sigma^2)\right] = \frac{1}{\sigma^2} \left(\sum_{i=1}^n Z_{ji}^* Z_{ki}' + \delta_{j1} \delta_{k1} \sum_{i=1}^n Z_{1i} Z_{1i}' \right)$$

$$(1 \leq j \leq k \leq \ell)$$

$$E\left(-\frac{\partial^2}{\partial \alpha_j \partial \beta_k'} \log L(\underline{\beta}, \sigma^2)\right) = \frac{1}{\sigma^2} \sum_{i=1}^n z_{ji}^* x_{ki}^{*'} \quad (1 \leq j \leq \ell, 1 \leq k \leq m)$$

$$E\left(-\frac{\partial^2}{\partial \beta_j \partial \beta_k'} \log L(\underline{\beta}, \sigma^2)\right) = \frac{1}{\sigma^2} \sum_{i=1}^n x_{ji}^* x_{ki}^{*'} \quad (1 \leq j \leq k \leq m)$$

which implies that

$$I(\underline{\beta}_0) = \underline{X}(\underline{\beta}_0, \sigma_0^2)' \underline{X}(\underline{\beta}_0, \sigma_0^2) \quad (8b)$$

Thus the revised estimate $\underline{\beta}_1$ is given by

$$\underline{\beta}_1 - \underline{\beta}_0 = [\underline{X}(\underline{\beta}_0, \sigma_0^2)' \underline{X}(\underline{\beta}_0, \sigma_0^2)]^{-1} \underline{X}(\underline{\beta}_0, \sigma_0^2)' \underline{\varepsilon}(\underline{\beta}_0, \sigma_0^2)$$

and the correction for $\underline{\beta}_0$ takes the form of an ordinary least squares estimator with $\underline{\varepsilon}$ as the dependent variable and \underline{X} as the data matrix. Moreover the ordinary standard errors of the final correction serve as standard errors of the final estimate $\hat{\underline{\beta}}$.

In the particular case that $X_{1j} = 0$, $\underline{X}'\underline{X}$ is block diagonal so that the scoring correction takes the form of one ordinary least squares estimator associated with α_1 and a second ordinary least squares estimator for $\alpha_2, \dots, \alpha_\ell, \beta_2, \dots, \beta_m$. The first of these two estimations is equivalent to Harvey's scoring procedure.

4. An Application

Policy decisions about alternative sources of energy require projections of future capital costs based on a thorough understanding of past experience. Recently W.E. Mooz (1978) published a de-

tailed statistical analysis of the capital cost of nuclear power plants constructed in the U.S. We will use data collected and published by Mooz to illustrate the methods of analysis developed in Section 3.

In the preparation of his data, Mooz first carries out a rather complex interpolation to adjust annually reported capital costs for the inflation of construction costs to obtain an estimate of the total capital cost of each plant measured in millions of 1973 dollars. We label this variable as COST. Mooz also records the SIZE of each plant, measured in installed megawatts of capacity. The dependent variable selected by Mooz is the ratio $1000 \text{ COST}/\text{SIZE}$, which measures costs in 1973 dollars per kilowatt of installed capacity.

The following independent variables are included in Mooz's final model [p. 32]:

- SIZE - installed capacity in megawatts.
- CPIS - date of issuance of the construction permit, represented as a decimal equivalent of year and month, e.g. 67.17 for February, 1967.
- LOC1 - a dummy variable indicating the FPC Region I, the Northeastern U.S.
- EXP - the natural logarithm of the cumulative number of plants built by the architect-engineer.
- TOWER - a dummy variable indicating the inclusion of a cooling tower in the plant.

The data used by Mooz and also here is for 37 nuclear power plants having cost data available and considered by Mooz to be reliable. These plants were started between 1967 and 1972 and range

in capacity from 457 megawatts to 1130 megawatts. Their costs range from 208 to 881 million 1973 dollars.

In our re-examination of these data we will investigate the use of capacity to deflate both COST and the independent variables. For our purpose we introduce the variable $S=SIZE/GM(SIZE)$. Here $GM(SIZE)$ is the geometric mean of $SIZE$ within the 39 plant database, 813.788 megawatts.

We consider models of the form

$$\begin{aligned} \text{COST} = & (\beta_1 + \beta_2 \text{ SIZE}) S^{\alpha_2} + \beta_3 \text{ CPIS } S^{\alpha_3} + \beta_4 \text{ LOCI } S^{\alpha_4} \\ & + \beta_5 \text{ EXP } S^{\alpha_5} + \beta_6 \text{ TOWER } S^{\alpha_6} + u \end{aligned} \quad (9)$$

with $E(u) = 0$ and $E(u^2) = \alpha^2 S^{2\alpha_1}$ and sometimes with the constraint $\alpha_1 = \alpha_2$. (For our computations we have also scaled COST by a factor $1/379.745$)

It is easy to verify that this specification is within the family of ratio models (1a), with Z_j defined to be $\log(S)$. The table summarizes our numerical analysis, first considering (9) to be unconstrained (full), and then under various constraints. In each case the table shows the maximum likelihood estimates of the coefficients $\hat{\alpha}_j$ and $\hat{\beta}_j$ and their respective standard errors. The table also gives several measures of the goodness of fit of each specification, ($k=I, \dots, V$), namely the standard error $\hat{\sigma}_k$ determined by (6b), the likelihood ratio of each constrained model

Table. Summary of Illustrative Analyses

	j						(7) $\hat{\sigma}$	(8) L^*	(9) X^2	(10) df	(11) $X^2_{.05}$
	1	2	3	4	5	6					
I											
Full	$\hat{\alpha}_j$	1.72	1.65	5.92	3.26	.84	.19442	1.0	-		
	$S\alpha$	1.07	.83	2.74	1.99	1.96					
	$\hat{\beta}_j$.25-3	.31	.20	-.15	.19					
	$S\beta$	4.22-3	.04	.12	.07	.09					
II											
$\alpha_1 = \alpha_2$	$\hat{\alpha}_j$	1.07	1.09	4.51	2.30	1.10	.19624	.695	.73	1	3.81
	$S\alpha$.52	.46	2.40	1.51	1.87					
	$\hat{\beta}_j$	-1.21-3	.31	.22	-.16	.19					
	$S\beta$	2.79-3	.04	.11	.06	.09					
III											
α_j fixed	$\hat{\alpha}_j$	1.0	1.0	1.0	1.0	1.0	.20642	.097	4.67	6	12.59
	$\hat{\beta}_j$	-19.0	.30	.27	-.16	.20					
	$S\beta$	2.2	.03	.09	.04	.07					
IV											
α_j fixed	$\hat{\alpha}_j$	1.0	0.	0.	0.	0.	.20974	.052	5.92	6	12.59
	$\hat{\beta}_j$	-21.0	.31	.26	-.15	.16					
	$S\beta$	2.3	.03	.08	.04	.07					
V											
α_j fixed	$\hat{\alpha}_j$	0.	0.	0.	0.	0.	.22037	.007	9.77	6	12.59
	$\hat{\beta}_j$	-21.0	.31	.29	-.16	.17					
	$S\beta$	2.4	.04	.09	.05	.08					

$$L_k^* = \frac{L(\hat{\beta}_k, \hat{\sigma}_k^2)}{L(\hat{\beta}_I, \hat{\sigma}_I^2)} = \left[\frac{\hat{\sigma}_I}{\hat{\sigma}_k} \right]^n$$

and the chi-square statistic for testing the null hypothesis associated with the constrained model,

$$\chi^2 = -2 \log L_k^*$$

Under the null hypothesis that the constrained model is a correct specification, χ^2 asymptotically has the chi-square distribution with degrees of freedom (df) equal to the number of constraints, so that a critical value (e.g. $\chi^2_{.05}$) can be easily selected for testing the constrained model against the full model.

A careful examination of the table provides a number of helpful insights although our findings must be considered as rather tentative due to the relatively small sample available. In the full model, $\hat{\alpha}_1 = .6$ which suggests some heteroscedasticity in the relationship determining COST, with the standard deviation of the disturbance approximately in proportion to the square root of plant size. The high values of $\hat{\alpha}_2, \dots, \hat{\alpha}_5$ suggest a rather strong relationship between COST and SIZE, resembling a quadratic function of SIZE with strong interaction between SIZE and LOCI and EXP. The indication is that some deflation of COST is called for to adjust for heteroscedasticity but the independent variables should not be deflated as in weighted least squares but actually inflated for SIZE.

The second analysis (II: $\alpha_1 = \alpha_2$) is representative of several attempts to simplify the full model without undue loss of goodness of fit. The success of the simplification should be subjectively evaluated in terms of the plausibility of the specification and of the corresponding estimates, and the sacrifice of fit as measured by $\hat{\sigma}$, χ^2 , or (perhaps most suitably) by L^* . Under the constraint that $\hat{\alpha}_1 = \hat{\alpha}_2$, the standard deviation of the disturbance is estimated to be almost directly proportional to SIZE, and the first two terms of the predictive equation are almost quadratic in SIZE with the remaining terms similar to the full model. The likelihood ratio of this specification compared to the full model is about 7 to 10 and certainly "acceptable" in the technical sense of the asymptotic χ^2 test.

We might interject a note about the computation of the estimates. The scoring procedure was used with the initial values of the deflation coefficients $\alpha_1, \dots, \alpha_6$ taken to be one and the β_1, \dots, β_6 conditionally determined from ordinary least squares using the deflated variables X^*, Y^* . Then a second data set was set up corresponding to $\underline{X}, \underline{\varepsilon}$ and the corrections to the deflation coefficients were determined from a second ordinary least squares regression. This procedure was iterated until convergence of $\hat{\sigma}$, the (unadjusted) standard error of the conditional regression. In our experience, satisfactory convergence occurred in two to six cycles

although convergence was not uniformly good for all deflation coefficients. In fact $\hat{\alpha}_6$ usually oscillated rather widely, indicating poor identifiability of this parameter.

In analyses III-V of the table, the linear coefficients β_1, \dots, β_6 are estimated conditional on specified values of the deflation coefficients. Specification III in its deflated form is equivalent to the final model reported by Mooz. (Mooz uses 1000 COST/SIZE rather than COST/S. Moreover our COST is multiplicatively rescaled as already noted. Mooz's estimates can be obtained by multiplying all statistics of III except L^* and χ^2 by the factor 466.64.) Specification III achieves a considerable simplification at a rather high but not conventionally significant cost in goodness of fit. A comparison of III with II suggests that Mooz might find it desirable to add interaction terms of SIZE with both LOCI and EXP.

Analyses IV and V explore other plausible, simple specifications. In IV, COST is linearly related to undeflated independent variables, but with the standard deviation of the disturbances proportional to SIZE, as in the heteroscedastic specification. V is a homoscedastic specification relating SIZE to the undeflated independent variables. The likelihood ratios of IV and V relative to III are about .54 and .07 respectively so III seems to be clearly superior among these three simple conditional models.

5. Concluding Comments

Evaluating the adequacy of alternative model specifications is an essential aspect of most data analysis. The coefficient of determination R^2 has seemingly become the statistic of choice for such comparisons when the dependent variable is invariant. Of course the standard error of the regression (unadjusted or adjusted) could serve equally well in this case.

When the comparisons involve transformations of the dependent variable, it is widely recognized that the coefficient of determination is generally misleading because of changes in the total sum of squares. Fortunately Box and Cox (1964) have given us a simple method of comparing models involving power and logarithmic transformations of the dependent variable. In the Box-Cox case, the dependent variable is multiplicatively standardized to have unitary geometric mean, and the standard error of the regression or the associated likelihood ratio is used for comparisons. We have shown in this paper that an analogous standardization of the deflation variables yields a similar approach to the comparison of specifications involving ratio transformations or heteroscedasticity. In fact a wide family of specifications involving both Box-Cox transformations and our deflations can be easily investigated provided that the dependent variable and the deflation variables are suitably standardized.

6. References

- Box, G.E.P. and Cox, D.R. (1964), "An Analysis of Transformations," Journal of the Royal Statistics Society, B, 26, 211-243.
- Draper, N.R. and Smith, H. (1966), Applied Regression Analysis, New York: Wiley.
- Fisher, G.R. (1957), "Maximum Likelihood Estimators with Heteroscedastic Errors," Review of the International Statistical Institute, 25, 52-55.
- Glejser, H. (1969), "A New Test for Heteroscedasticity," Journal of the American Statistical Association, 64, 316-323.
- Granger, C.W. and Newbold, P. (1976), "The Use of R^2 to Determine the Appropriate Transformation of Regression Variables," Journal of Econometrics, 4, 205-210.
- Griliches, Zvi, (1972), "Cost Allocation in Railroad Regulation," The Bell Journal of Economics and Management Science, 26-41.
- Harvey, A.C. (1976), "Estimating Regression Models with Multiplicative Heteroscedasticity," Econometrica, 44, 461-464.
- Johnston, J. (1972), Econometric Methods, Second Edition, New York: McGraw-Hill.
- Kmenta, J. (1971), Elements of Econometrics, New York: Macmillan.
- Kuh, E. and Meyer, J.R. (1955), "Correlation and Regression Estimates when the Data are Ratios," Econometrica, 23, 400-416.
- Madansky, A. (1964), "Spurious Correlation Due to Deflating Variables," Econometrica, 32, 652-655.
- Mooz, W.E. (1978), Cost Analysis of Light Water Reactor Power Plants (Report), Santa Monica: Rand, R-2304-DOE.
- Park, R.E. (1966), "Estimation with Heteroscedastic Error Terms," Econometrica, 34, 888.
- Pearson, K. (1897), "On a Form of Spurious Correlation which may Arise when Indices are used in the Measurement of Organs," Proceedings of the Royal Society of London, 60.

Rao, C.R. (1973), Linear Statistical Inference and its Applications,
Second Edition, New York: Wiley.

Rutemiller, H.C. and Bowers, D.A. (1968), "Estimation in a
Heteroscedastic Regression Model," *Journal of the American
Statistical Association*, 63, 552-557.

Taylor, W.E. (1978), "The Heteroscedastic Linear Model: Exact
Finite Sample Results," *Econometrica*, 46, 663-676.