T032
2/2/84

(84)

RESTRICTED COLLECTION

SOME PROBLEMS ON ECONOMETRIC REGRESSION ANALYSIS

MAITREYI CHAUDHURI

INDIAN STATISTICAL INSTITUTE LIBRARY
No. T032
Date 2 FEB 1984
CALCUTTA

A thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Calcutta

1979

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

RESTRICTED COLLECTIO

PREFACE AND ACKNOWLEDGEMENTS

This thesis presents studies on some of the problems of single equation econometric regression analysis. The problems are

1. Omission of regressors from a single equation regression model,

and

2. Handling of errors in variable models with trending or autocorrelated errors.

The structure of the thesis is as follows :

Chapter 1 attempts a brief survey of existing literature on the problems of

(a) Omission of relevant regressors from a regression equation and misspecification of algebraic forms.

(b) Autocorrelation of disturbances.

and

(c) Errors in variables.

Chapter 2 considers the problem of omission of regressors from a single equation regression model with nonstochastic regressors and spherical disturbances. The results of this chapter have been published in the Journal of Econometrics 1977, vol.5, pp. 301-313.

Chapter 3 considers the problem of omission of regressors from a regression equation having stochastic regressors and/or autocorrelated disturbances.

STATISTICAL INSTITUTE LIBRARY
No. 1022
12 FEB 1984

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In Chapter 4 we have compared the asymptotic MSEs of the OLS estimators of the regression coefficients in the fully specified model with the MSEs of the OLS estimators of regression coefficients when one of the variables has been omitted from the true model.

Chapter 5 considers the problems of estimation relating to errors in variable models where the errors of observations either have trend components or the errors are autocorrelated.

In Chapter 6 we take an overview of the entire investigation and point to the directions to which further researches can be made.

Appendices 1,2 and 3 contain the derivations of some of the results of Chapters 2,3 and 4 respectively.

Grateful acknowledgements are due to Professor N. Bhattacharya who guided me through every step in the preparation of this thesis. I am thankful to him for the benefit of his time and attention that I could draw upon so liberally. I also thank Dr. Dipankar Coondoo, Dr. Kajal Lahiri, Dr. Robin Mukherjee and Mr. Nityananda Sarkar who made many useful suggestions in various contexts. I have also been benefited from the invaluable suggestions and comments by Professor A. Zellner and Professor J.B. Ramsey on various parts of the thesis. To them I extend my gratefulness. I acknowledge my indebtedness to the Indian Statistical Institute for the facilities

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

granted to me. I owe my sincere thanks to all my teachers and colleagues at Economic Research Unit of the Indian Statistical Institute for their helpful cooperation. Finally I thank Mr. Arun Das for excellent typing and Mr. Harish Prasad for rolling the stencils.

Maitreyi Chauduri

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CONTENTS

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 0

ABSTRACT

Very often in econometric analysis one adopts the classical linear regression model. The classical linear regression model is given by

$$Y = X\beta + \varepsilon \tag{0.1}$$

where $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix}$,

$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$ and $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

The underlying assumptions are :

$$E(\varepsilon) = \underset{\sim}{0} \text{ (null vector)}$$

$$\text{and} \quad E(\varepsilon\varepsilon') = \sigma^2 I_n \tag{0.2}$$

X is nonstochastic and rank of X is $k \leq n$.

If, in addition, ε is assumed to be normally distributed, the model is called classical normal linear regression model. Ordinary least squares (OLS) methods of estimation and hypothesis testing are based on this model.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

But the assumptions on ε's and x's may not be fulfilled in reality; or, in other words, the model may not be correctly specified. One class of problems arises when some of the regressors are omitted from the equation and/or some additional regressors are wrongly included in the model; or when the algebraic form of the regression equation is misspecified. In such cases OLS method would fail to give satisfactory estimates of the regression coefficients.

Another class of problems is created when $E(\varepsilon\varepsilon') \neq \sigma^2 I_n$. Generalised least squares techniques are called for in such situations.

Problems also arise when the regressors (X) are stochastic. There is little trouble if X is stochastic but fully independent of ε. However, if the regressors and disturbances are correlated, OLS estimates cease to be unbiased. The danger is particularly great if the regressor values and the disturbances in the same observational equation are correlated. In this case, OLS estimates of β's are not even asymptotically unbiased. This kind of complication arises in two important situations :

(a) where the regressors are observed with errors

and (b) where the equation is embedded in simultaneous equation models where several current endogenous

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

variables are determined through the simultaneous interactions of the structural relationships in the model.

This study, is largely concerned with

1. Problems of omission of regressors from a single equation regression model leading to autocorrelation among the disturbances

2. MSE criterion in the context of specification error analysis with stochastic regressors

and

3. Handling of errors in variable models with trending or autocorrelated errors.

Below we give a summary of different chapters in the thesis.

Chapter 1. A survey of previous researches.

This chapter gives a brief survey of existing literature on three main problems of econometrics (single equation methods) to provide a background to the investigations reported in this thesis. The problems are those arising due to

(a) Omission of relevant regressors from a regression equation and misspecification of algebraic forms

(b) Autocorrelation of disturbances

(c) Errors in variables.

(a) Omission of relevant regressors from a regression equation and misspecification of algebraic forms : The survey has been organised under the following heads :

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(i) The consequences of using OLS procedures for estimating the regression coefficients of a misspecified model.

(ii) Applications of specification analysis.

(iii) Different tests of misspecification and their applications.

(iv) The residual variance criterion.

(v) The method of using least squares to approximate unknown regression functions.

and (vi) Consequences of misspecification in simultaneous equation systems.

(b) Autocorrelation of disturbances : Here we have not considered the distributed lag models or simultaneous equation models although a few references have been cited. Different methods of testing the randomness of disturbances in a regression model have been discussed. A brief survey of different methods of estimating the regression coefficients in a regression model with autocorrelated disturbances has been given.

(c) Errors in variables : Single equation errors in variable models have been considered. We have discussed.

(i) Effects of errors of observations on OLS estimation of regression coefficients.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(ii) Effects of errors on the regression line.

(iii) Methods of estimation.

and (iv) Other relevant problems.

Chapter 2. Autocorrelated disturbances in the light of specification analysis.— Part 1

This chapter examine the consequences of omission of relevant regressors from a regression equation with nonstochastic regressors. In the literature on econometric methods it is wellknown that one of the important causes of autocorrelation among disturbances is omission of relevant regressors from a regression equation. When the disturbances are autocorrelated, they are generally assumed to follow the Markov scheme

$$\varepsilon_t = \rho \, \varepsilon_{t-1} + u_t$$

where, $|\rho| < 1$, $E(u_t) = 0 \; \forall \, t$ and $cov(u_t, u_s) = \sigma_u^2 \, \delta_{ts}$, δ_{ts} being the Kronecker delta.

In this case, the OLS formulae for estimating the sampling variances of the estimated regression coefficients tend to give serious underestimates in some important situations.

If, however, the effect of omission of regressors be examined following the approach of specification analysis due to Theil (1957), the usual formulae appear to overestimate the sampling

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

variances of the estimated regression coefficients of the misspecified equation. So we arrive at something like a contradiction. This point has been examined carefully in this chapter from the point of view of specification analysis.

Suppose, from the equation (0.1), let the last (k-m) regressors have been omitted. So, the misspecified model is

$$y = X^{+}\beta^{+} + \varepsilon^{+} \qquad (0.3)$$

where X^{+} contains the first m colums of X. $\beta^{+} = (\beta_1^{+}, \beta_2^{+}, \ldots, \beta_m^{+})'$ and $\varepsilon^{+} = (\varepsilon_1^{+}, \varepsilon_2^{+}, \ldots, \varepsilon_n^{+})'$. β^{+} is different from β. β^{+} should be defined in such a way that it may capture to the extent possible the partial influence of the omitted regressors on the regressand; or, in other words, β_i^{+}'s should be defined in such a way that they may enable the regression function to approximate as closely as possible the systematic component of Y, i.e., $\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$.

The OLS estimates of β_i^{+}'s will, in general, be biased estimates of β_i's and may not be so meaningful. If, however, the interest lies mainly in predicting y from the truncated set of regressors and not in estimating individual structural parameters, the equation may still be quite useful.

β^{+} has been defined so that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E\{(y-X^{+}\beta^{+})'(y-X^{+}\beta^{+})\} \text{ is minimum} \quad (0.4)$$

(0.4) is minimised for

$$\beta^{+} = (X^{+\prime}X^{+})^{-1}X^{+\prime}X\beta = P\beta \quad (0.5)$$

and $$\varepsilon^{+} = y - X^{+}\beta^{+} \quad (0.6)$$

The newly defined ε^{+} has some peculiar properties.

$$E(\varepsilon^{+}) = (X-X^{+}P)\beta = z = (z_1, z_2, \ldots, z_n)'$$

$$E(\varepsilon^{+}\varepsilon^{+\prime}) = zz' + \sigma^2 I_n \neq \sigma^2 I_n \quad (0.7)$$

and $$D(\varepsilon^{+}) = E\{\varepsilon^{+} - E(\varepsilon^{+})\}\{\varepsilon^{+} - E(\varepsilon^{+})\}' = \sigma^2 I_n$$

So, ε^{+}'s do not follow a Markov scheme

$$\varepsilon_t^{+} = \rho\,\varepsilon_{t-1}^{+} + u_t, \quad |\rho| < 1 \text{ and } u_t \text{ spherical} \quad (0.8)$$

It has been observed that the OLS estimator gives unbiased estimates of β_i^{+}'s, $i = 1, 2, \ldots, m$.

$$D(\hat{\beta}^{+}) = E(\beta^{+}-\hat{\beta}^{+})(\beta^{+}-\hat{\beta}^{+})' = \sigma^2(X^{+\prime}X^{+})^{-1} \quad (0.9)$$

σ^2 is estimated by $\dfrac{e^{+\prime}e^{+}}{n-m}$

where e^{+} is the OLS residual given by

$$e^{+} = y - X^{+}\hat{\beta}^{+} \quad (0.10)$$

where $\hat{\beta}^{+}$ is the OLS estimator of β^{+} in (0.3).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

It can be shown that

$$E(e^{+\prime}e^{+}) = \beta'(X - X^{+}P)'(X - X^{+}P)\beta + (n-m)\sigma^2 \tag{0.11}$$

Thus, $$\frac{E(e^{+\prime}e^{+})}{n-m} > \sigma^2 \quad \text{(in general)} \tag{0.12}$$

So, the usual OLS formula $\frac{e^{+\prime}e^{+}}{n-m}(X^{+\prime}X^{+})^{-1}$ gives an overestimate of $D(\hat{\beta}^{+})$. This, however, is a familiar result proved by Theil (1957).

The Durbin-Watson (D-W) (1950,1951) test statistic for testing the randomness of disturbances of the misspecified model is given by

$$d = \frac{\sum_{t=2}^{n}(e_t^{+} - e_{t-1}^{+})^2}{\sum_{t=1}^{n} e_t^{+2}} \tag{0.13}$$

$$\operatorname*{plim}_{n\to\infty} d = 2(1 - \rho_o) \tag{0.14}$$

where $$\rho_o = \frac{\tilde{\rho}}{1 + \frac{\sigma^2}{\sigma_o^2}}$$

and $$\tilde{\rho} = \frac{\lim_{n\to\infty} \frac{1}{n} \sum_{t=2}^{n} z_{\infty,t}\, z_{\infty,t-1}}{\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} z_{\infty,t}^2}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\sigma_o^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} z_{\infty,t}^2$

where $z_{\infty,t}$ is the t-th element in the vector $(X - X^+ P_\infty)\beta$ and $P_\infty = \lim_{n\to\infty} P$.

If $\tilde{\rho} = 0$, then $\rho_o = 0$. If, however, $\tilde{\rho}$ is positive and σ^2/σ_o^2 is such that ρ_o is appreciably greater than zero, d would often come out to be significantly smaller than 2 in large samples.

When the D-W statistic comes out to be significantly less than 2, one generally tries to re-estimate β^+ by fitting a Markov scheme to ε_t^+'s, i.e.,

$$\varepsilon_t^+ = \rho_o \, \varepsilon_{t-1}^+, \; |\rho_o| < 1 \text{ and } u_t \text{ is the spherical disturbance term} \qquad (0.15)$$

Here the symbol ρ_o has been used in anticipation of subsequent results. The methods of re-estimation discussed are

(a) Cochrane-Orcutt two-step method (1949).

(b) Prais-Winsten method (vide Rao, 1968).

(c) Durbin two-step procedure (1960).

It has been observed that all the above two-step procedures give

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

inconsistent estimates of β^+. The probability limit of the Cochrane-Orcutt two-step estimator is approximately equal to that of Prais-Winsten estimator ; whereas, the probability limit of the Durbin two-step estimator is different from those of Cochrane-Orcutt two-step and Prais-Winsten estimators.

Chapter 3. Autocorrelated disturbances in the light of specification analysis - Part II.

In this chapter we have mainly extended the results in Chapter 2 to the case where x's are stochastic and ε_t's in the true model are themselves autocorrelated. In this chapter we have also considered the subcases where

(i) x's are stochastic and ε_t's is the true model are spherical.

(ii) x's are nonstochastic and ε_t's in the true model follow a Markov scheme.

The model considered is

$$y = X\beta + \varepsilon \qquad (0.16)$$

x's are stochastic and

$\varepsilon_t = \rho\, \varepsilon_{t-1} + u_t$, $|\rho| < 1$ and u_t spherical with mean 0 and variance $\sigma_u^2 \; \forall\, t$. So, in this case,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(\varepsilon) = 0$$

and $$E(\varepsilon\varepsilon') = \sigma^2 V$$

where $$\sigma^2 = \frac{\sigma_u^2}{1-\rho^2}$$

and

$$V = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}$$

From the model (0.16), the last (k-m) regressors have been omitted. So, the misspecified model is of the form

$$y = X^+\beta^+ + \varepsilon^+ \qquad (0.17)$$

where β^+ is obtained by minimising

$$E(y - X^+\beta^+)'(y - X^+\beta^+)$$

It has been observed that the above expression is minimum when

$$\beta^+ = \{E(X^{+\prime} X^+)\}^{-1} E(X^{+\prime} X)\beta$$

$$= \bar{P}\beta \quad . \qquad (0.18)$$

~~the expression in (0.17) is minimum~~

INDIAN STATISTICAL INSTITUTE LIBRARY
No. T032
Date 2 FEB 1984
CALCUTTA

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\varepsilon^{+} = y - X^{+}\beta^{+} \tag{0.19}$$

Here ε^{+} has the following properties :

$E(\varepsilon^{+}) = E(X - X^{+}\bar{P})\beta = \tilde{z} = (\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_n)' \neq \underset{\sim}{0}$ (in general)

$E(\varepsilon^{+}\varepsilon^{+\prime}) = E(\tilde{z}\,\tilde{z}') + \sigma^2 V$ and

$D(\varepsilon^{+}) = E\{\varepsilon^{+} - E(\varepsilon^{+})\}\{\varepsilon^{+} - E(\varepsilon^{+})\}' = \sigma^2 V$

(0.20)

From the above results, it is obvious that ε_t^{+}'s do not follow a Markov scheme given by

$$\varepsilon_t^{+} = \rho\ \varepsilon_{t-1}^{+} + u_t\ ;\ |\rho| < 1 \text{ and } u_t \text{ spherical.}$$

In the special case where $E(\tilde{z}\,|X^{+}) = \underset{\sim}{0}$ (i.e. where the regressions of X^{-} on X^{+} are strictly linear),

$$E(\varepsilon^{+}) = \underset{\sim}{0} \tag{0.21}$$

It has been observed that the OLS estimator fails to give unbiased estimates of β^{+}. But, under the assumption that

$$\underset{n \to \infty}{\text{plim}} \left(\frac{1}{n} X'X\right) \text{ exists,} \tag{0.22}$$

OLS give consistent estimate of β^{+}.

In the special case, where $E(\tilde{z}\,|X^{+}) = \underset{\sim}{0}$, the OLS give unbiased estimate of β^{+}.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In the general case, where $E(\tilde{z} \mid X^+) \neq \underset{\sim}{0}$, let us define

$$X_o^+ = \begin{pmatrix} 1 & x_{11} - \bar{x}_1 & \cdots & x_{m1} - \bar{x}_m \\ 1 & x_{12} - \bar{x}_1 & \cdots & x_{m2} - \bar{x}_m \\ \vdots & \vdots & & \vdots \\ 1 & x_{1n} - \bar{x}_1 & \cdots & x_{mn} - \bar{x}_m \end{pmatrix}$$

$$\bar{x}_i = \frac{\sum_{t=1}^{n} x_{it}}{n},$$

$$\beta_o^+ = [(\beta_1^+ + \sum_{i=2}^{m} \beta_i^+ \bar{x}_i), \beta_2^+, \ldots, \beta_m^+]'$$

$$= (\tilde{\beta}_1^+, \beta_2^+, \ldots, \beta_m^+)'$$

$$(X_o^{+\prime} X_o^+)^{-1} = \frac{1}{\Delta} (a_{ij}), \quad i = 1,2,\ldots,m, \quad j = 1,2,\ldots,m,$$

where $a_{1j} = a_{j1} = 0$ for $j = 2,3,\ldots,m$

and $\Delta = \mathrm{Det}(X_o^{+\prime} X_o^+)$.

Let, $\hat{\beta}_o^+$ be the OLS estimator of β_o^+.

$$V(\hat{\tilde{\beta}}_1^+) = [\text{the } (1,1) \text{ element in the matrix } V^+] + \mu_1 \sigma^2$$

$$+ 2\sigma^2 E \left\{ \frac{a_{11}^2}{2} \sum_{j=1}^{n-j} \rho^j (n-j) \right\} \qquad (0.23)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where V^+ is the dispersion matrix of $(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} \tilde{z}$.

$$\mu_1 = E(\frac{a_{11}}{\Delta}).$$

$$V(\hat{\beta}_i) = [\text{the } (i,i) \text{ element in the matrix } V^+] + \mu_i \sigma^2$$
$$+ 2\sigma^2(\rho q_1 + \rho^2 q_2 + \dots + \rho^{n-1} q_{n-1}), \quad i = 2,3,\dots,m. \tag{0.24}$$

where $\mu_i = E(\frac{a_{ii}}{\Delta})$

$$q_t = E \frac{\sum_{j=1}^{n-t} a_j^{(i)} a_{j+t}^{(i)}}{\Delta^2}, \quad t = 1,2,\dots,n-1$$

$$a_j^{(i)} = a_{i2}(x_{2j} - \bar{x}_2) + a_{i3}(x_{3j} - \bar{x}_3) + \dots + a_{im}(x_{mj} - \bar{x}_m)$$

σ^2 is estimated by $\frac{e^{+\prime} e^+}{n - m}$ in the OLS procedure.

$$e^+ = y - X_o^+ \hat{\beta}_o^+ \tag{0.25}$$

Under the assumption that $\tilde{z}$'s are homoscedastic with variance σ_z^2,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(e^{+\prime}e^{+}) = n(\sigma^2_{\tilde{z}} + \sigma^2) - \text{trace } W - m\sigma^2 + \frac{2\sigma^2}{}[a_{11} \sum_{j=1}^{n-1} \rho^j (n-j) +$$

$$+ a_{22} \sum_{j=1}^{n-1} \rho^j \sum_{i=1}^{n-j} (x_{2i} - \bar{x}_2)(x_{2,i+j} - \bar{x}_2)$$

$$+ a_{23} \sum_{j=1}^{n-1} \rho^j \{ \sum_{i=1}^{n-j} (x_{2i} - \bar{x}_2)(x_{3,i+j} - \bar{x}_3) + \sum_{i=1}^{n-j} (x_{3i} - \bar{x}_3)(x_{2,i+j} - \bar{x}_2) \} + \cdots$$

$$+ a_{mm} \sum_{j=1}^{n-1} \rho^j \sum_{i=1}^{n-j} (x_{mi} - \bar{x}_m)(x_{m,i+j} - \bar{x}_m)]$$

$$= (n-m)\sigma^2 + n\sigma^2_z - E(S_o) - \text{trace } W. \tag{0.26}$$

where $W = E[(X_o^{+\prime} X_o^{+})^{-1} X_o^{+\prime} \tilde{z}\tilde{z}' X_o^{+}]$.

If $\rho = 0$, $E(S_o)$ vanishes. The term $n\sigma^2_{\tilde{z}}$ - trace W may be taken as the effect of omission on the residual sum of squares.

Whether $\frac{e^{+\prime}e^{+}}{n-m}$ is an overestimate, or underestimate or unbiased estimate of σ^2 will depend on the relative magnitudes of $n\sigma^2_{\tilde{z}}$ - trace(W) and $E(S_o)$.

The standard OLS formulae for estimating the variance of $\hat{\tilde{\beta}}_1^{+}$ and $\hat{\beta}_i^{+}$, $i = 2,3,\ldots,m$ are given by

$$\frac{e^{+\prime}e^{+}}{n-m} \frac{a_{11}}{\Delta} \tag{0.27}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\frac{e^{+\prime}_{n} e^{+}}{n-m} \frac{a_{i1}}{\Delta}$ $i = 2,3,\ldots,m$ respectively (0.28)

From (0.23), (0.24), (0.26), (0.27) and (0.28), we find that the standard OLS formulae for estimating $V(\hat{\beta}_1^+)$ and $V(\hat{\beta}_i^+)$, $i = 2,3,\ldots,m$ can be in error in two respects. The estimate of σ^2 may not be unbiased and at the same time many terms in the expression for $V(\hat{\beta}_1^+)$ or $V(\hat{\beta}_i^+)$, $i = 2,3,\ldots,m$ may be neglected. This result is, however not surprising. A similar result holds for the OLS estimation of variances of the estimated regression coefficients when the disturbances in the correctly specified model are autocorrelated and there is no omission of regressors.

Next, the speical case $E(\tilde{z} \mid X^+) = \underset{\sim}{0}$ has been considered. Let X^- be the $(n \times (k-m)$ matrix which contains the last $(k-m)$ columns of the matrix X. When the regressions of X^- on X^+ are perfectly linear,

$$E(\tilde{z} \mid X^+) = \underset{\sim}{0} \quad [\text{since } \tilde{z} = (X - X^+ \bar{P})\beta = (X^- - X^+ \bar{P}^*)\beta^{**}]$$

where $\bar{P}^* = \{E(X^{+\prime} X^+)\}^{-1} E(X^{+\prime} X^-)$

and $\beta^{**} = (\beta_{m+1}, \beta_{m+2}, \ldots, \beta_k)'$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The following set of regression equations have been considered.

$$X^- = X^+ \delta + \nu \tag{0.29}$$

We have further assumed that ν is independent of X^+. So,

$$E(\nu | X^+) = E(\nu | X_0^+) = \underset{\sim}{0}$$

$$\delta = \bar{p}^*$$

and $E(\tilde{z} | X_0^+) = E(\nu | X_0^+) \beta^{**} = \underset{\sim}{0}$

$$E(\tilde{z} \tilde{z}' | X_0^+) = E(\nu \beta^{**} \beta^{**\prime} \nu' | X_0^+) = E(\nu \beta^{**}\beta^{**\prime} \nu')$$

(0.30)

We can write

$$\nu = \begin{pmatrix} \nu_{m+1,1} & \nu_{m+2,1} & \cdots & \nu_{k,1} \\ \nu_{m+1,2} & \nu_{m+2,2} & \cdots & \nu_{k,2} \\ \vdots & \vdots & & \vdots \\ \nu_{m+1,n} & \nu_{m+2,n} & & \nu_{kn} \end{pmatrix}$$

We assume that these (k-m) column vectors in (0.30) are mutually independent. This means that any two regressors in X^- have zero partial correlation if the influence of X^+ has been eliminated. Each vector of disturbances in (0.30) has been assumed to be homoscedastic and the i-th vector follows the Markov scheme given by

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\nu_{m+i,t} = \rho_i \, \nu_{m+i,t-1} + \eta_{m+i,t} \; ; \; |\rho_i| < 1, \; i = 1,2,\ldots,(k-m)$$

$$t = 2,3,\ldots, n \qquad (0.31)$$

$\eta_{m+i,t}$ is the spherical disturbance term with

$E(\eta_{m+i,t}) = 0 \;\forall\; t$ and $E(\eta_{m+i,t}, \eta_{m+i,s}) = \sigma^2_{m+i,\eta} \, \delta_{ts}$, δ_{ts} being the Kronecker delta.

Defining $\sigma^{*2}_i = \sigma^2_{m+i,} \quad \beta^{**2}_{m+i}$, $i = 1,2,\ldots,(k-m)$,

$$\sigma^{*2} = \sum_{i=1}^{k-m} \sigma^{*2}_i \qquad (0.32)$$

and $\sigma^2_{m+i,} = \dfrac{\sigma^2_{m+i,\eta}}{1 - \rho_i^2}$, $i = 1,2,\ldots, (k-m)$

We have

$$V(\hat{\tilde{\beta}}_1^{+}) = \mu_1(\sigma^{*2}+\sigma^2) + 2E\Big[\frac{a_{11}^2}{2} \sum_{j=1}^{n-j} \Big\{ \sum_{i=1}^{k-m} (\rho_i)^j \, \sigma^{*2}_i \Big\} (n-j)\Big]$$

$$+ \, 2E\Big[\frac{a_{11}^2}{2} \sum_{j=1}^{n-j} \rho^j \, (n-j)\Big] \, \sigma^2 \qquad (0.33)$$

and

$$V(\hat{\beta}_i^{+}) = \mu_i(\sigma^2+\sigma^{*2}) + 2\Big[\Big(\sum_{j=1}^{k-m} \rho_j \, \sigma^{*2}_j + \rho \, \sigma^2\Big) q_1$$

$$+ \Big(\sum_{j=1}^{k-m} \rho_j^2 \, \sigma^{*2}_j + \rho^2 \sigma^2\Big) q_2 + \ldots + \Big(\sum_{j=1}^{k-m} \rho_j^{n-1} \, \sigma^{*2}_j + \rho^{n-1} \sigma^2\Big) q_{n-1}\Big]$$

$$i = 2,3,\ldots,m \qquad (0.34)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In the expression for ($E(e^{+'}e^{+})$ in (0.24), we can now obtain a simplified expression for trace W.

$$\text{trace } W = m\sigma^{*2} + \frac{2}{\Delta} E[a_{11}\{\sum_{j=1}^{-1}(\sum_{i=1}^{\kappa-m}\rho_i^j\sigma_i^{*2})(n-j)\}$$

$$+ a_{22}\sum_{j=1}^{n-1}(\sum_{i=1}^{\kappa-m}\rho_i^j\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{2t}-\bar{x}_2)(x_{2,t+j}-\bar{x}_2)\}$$

$$+ a_{23}\sum_{j=1}^{n-1}(\sum_{i=1}^{\kappa-m}\rho_i^j\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{2t}-\bar{x}_2)(x_{3,t+j}-\bar{x}_3)$$

$$+ \sum_{t=1}^{n-j}(x_{3t}-\bar{x}_3)(x_{2,t+j}-\bar{x}_2)\}$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ a_{mm}\sum_{j=1}^{n-1}(\sum_{i=1}^{\kappa-m}\rho_i^j\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{mt}-\bar{x}_m)(x_{m,t+j}-\bar{x}_m)\}]$$

$$= m\sigma^{*2} + E(S_o^*) \quad \text{(say)} \qquad (0.35)$$

The conclusions regarding the bias in estimating the sampling variances of $\hat{\tilde{\beta}}_1^+$ and $\hat{\beta}_i^+$'s, $i = 2,\ldots,m$ by the OLS method are similar to those in the general case.

In large samples when the $\tilde{z}_t$'s are positively autocorrelated, under certain assumptions the D-W statistic for testing the randomness of disturbances comes out as

$$\underset{n\to\infty}{\text{plim}}\, d = 2(1 - \underset{n\to\infty}{\text{plim}} \frac{\sum_{t=2}^{n} e_t^+ e_{t-1}^+}{\sum_{t=1}^{n} e_t^{+2}})$$

$$= 2(1 - \rho^*) \qquad (0.36)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\rho^* \neq \rho$.

If, $E(\tilde{z}/X^+) = \underset{\sim}{0}$,

$$\rho^* = \frac{\rho_{\tilde{\varepsilon}}\, \sigma^2_{\tilde{z}} + \rho\, \sigma^2}{\sigma^2_{\tilde{z}} + \sigma^2} \tag{0.37}$$

where $\rho = \underset{n\to\infty}{\text{plim}} \left(\frac{\sum_{t=2}^{n} \epsilon_t \epsilon_{t-1}}{\sum_{t=1}^{n} \epsilon_t^2} \right)$

$$\rho_{\tilde{z}} = \underset{n\to\infty}{\text{plim}} \left(\frac{\sum_{t=2}^{n} \tilde{z}_t \tilde{z}_{t-1}}{\sum_{t=1}^{n} \tilde{z}_t^2} \right)$$

and $\sigma^2_{\tilde{z}} = \underset{n\to\infty}{\text{plim}} \frac{1}{n} \sum_{t=1}^{n} \tilde{z}_t^2$

For $\rho > 0$, $\underset{n\to\infty}{\text{plim}}\ d$ is likely to be less than 2 since, sample $\rho_{\tilde{z}}$ is assumed to be positive.

When D-W statistic comes out to be significantly less than 2, one generally fits a Markov scheme

$\varepsilon^+_t = \rho^* \varepsilon^+_{t-1} + u_t$ to ε^+_t and tries to re-estimate β^+ by using the three methods of re-estimation discussed in Chapter 2.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Here also, in the general case, these three methods of re-estimation give inconsistent estimates of β^+. The probability limit of the Cochrane-Orcutt (1949) two-step estimator is approximately same as that of the Prais-Winstein (vide Rao, 1968) estimator. But the probability limit of the Durbin (1960) two-step estimator is different from the probability limits of Cochrane-Orcutt two-step and Prais-Winsten estimators.

In the special case, where the regressions of X^T on X^+ are strictly linear, i.e., where $E(\tilde{z}/X^+) = \underset{\sim}{0}$, under a set of fairly general conditions, all the three methods of re-estimation give consistent estimates of the β^+.

Next we consider the subcases (i) and (ii).

<u>Subcase (i)</u> x's are stochastic, ε's spherical.
All the results can be obtained by putting I_n for V in the previous results.

Here it has been observed that OLS give biased but consistent estimates of β^+. Only in the special case, where $E(\tilde{z}/X^+) = \underset{\sim}{0}$, OLS give unbiased estimates of β^+.

Let $\hat{\tilde{\beta}}_1^+$ and $\hat{\beta}_i^+$ be the OLS estimates of $\tilde{\beta}_1^+$ and β_i^+,

$i = 2, 3, \ldots, m.$

$$V(\hat{\tilde{\beta}}_1^+) = [\text{the } (1,1) \text{ element of } V^+ + \mu_1\sigma^2] \geq \mu_1\sigma^2 = V_{OLS}(\hat{\tilde{\beta}}_1^+) \quad (0.38)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and in general,

$$V(\hat{\beta}_i^+) = [\text{the } (i,i) \text{ element of } V^+ + \mu_i \sigma^2] \geq \mu_i \sigma^2 = V_{OLS}(\hat{\beta}_i^+)$$

$$i = 2,3,\ldots,m \qquad (0.39)$$

Here $E(e^{+\prime} e^+) = (n-m)\sigma^2$

$$+ E[\{I - X_o^+(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime}\}\tilde{z}]'[\{I - X_o^+(X_o^{+\prime}X_o^+)^{-1}X_o^+\}\tilde{z}]$$

$$\geq (n-m)\sigma^2 \qquad (0.40)$$

So, $\frac{e^{+\prime} e^+}{n-m}(X_o^{+\prime} X_o^+)^{-1}$ will give biased estimates of the variances of $V(\hat{\tilde{\beta}}_1^+)$, $V(\hat{\beta}_i^+)$ $i = 2,\ldots,m$ depending on the relative effects of the first term in (0.38) or (0.39) and the second term in (0.40).

Large sample properties of the usual OLS formulae for estimating the sampling variances of β^+ have been investigated under the assumption that $E(\tilde{z}/X^+) = \underset{\sim}{0}$.

All the three methods of re-estimation (mentioned before) give inconsistent estimates of β^+. But in the special case, where $E(\tilde{z}|X^+) = \underset{\sim}{0}$, the methods of re-estimation give consistent estimates of β^+.

<u>Subcase (ii)</u> The regressors (X's) are nonstochastic and ε's follow a Markov scheme

$$\varepsilon_t = \rho\, \varepsilon_{t-1} + u_t \quad |\rho| < 1 \text{ and } u_t \text{ spherical.}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Most of the conclusions for stochastic X and ε's following a Markov scheme, remain valid here also.

Chapter 4. MSE criterion in the context of specification error analysis with stochastic regressors.

Here we have considered a two variable stochastic regressor model from which one regressor has been omitted. The true model is

$$y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon \tag{0.41}$$

x_1 and x_2 are stochastic and ε is the spherical disturbance term. x_1, x_2, y and ε are $n \times 1$ vectors. Each of the variables has been taken as a deviation from its respective mean. ε is independent of x_1 and x_2.

From the above model, the regressor x_2 has been omitted. So, the misspecified model is

$$y = \beta_1^+ x_1 + \varepsilon^+ \tag{0.42}$$

Let $\hat{\beta}_1$ and $\hat{\beta}_1^+$ be the OLS estimators of β_1 and β_1^+ respectively. We have obtained conditions under which $\bar{V}(\hat{\beta}_1)$ the asymptotic variance of $\hat{\beta}_1$ is greater than $\bar{V}(\hat{\beta}_1^+)$ the asymptotic variance of $\hat{\beta}_1^+$.

$\bar{V}(\hat{\beta}_1^+) \leq \bar{V}(\hat{\beta}_1)$ if and only if

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\frac{\beta_2^2}{\bar{V}(\hat{\beta}_2)} \leq \frac{\delta_{21}^2}{\bar{V}(\hat{\delta}_{21})} \qquad (0.43)$$

where $\hat{\beta}_2$ is the OLS estimate of β_2 in (0.41),

$$\delta_{21} = \frac{\bar{E}(\frac{1}{n} \sum_1^n x_{2i} x_{1i})}{\bar{E}(\frac{1}{n} \sum_1^n x_{2i}^2)} \quad 1/$$

and $$\hat{\delta}_{21} = \frac{\sum_1^n x_{2i} x_{1i}}{\sum_1^n x_{2i}^2} .$$

Also,

$\lim_{n \to \infty} MSE(\hat{\beta}_1^+) \leq \lim_{n \to \infty} MSE(\hat{\beta}_1)$ if and only if

$$\frac{\beta_2^2}{\bar{V}(\hat{\beta}_2)} \leq \frac{1}{1 + \frac{\bar{V}(\hat{\delta}_{21})}{\delta_{21}^2}} \qquad (0.44)$$

When x_1 and x_2 jointly follow a bivariate normal distribution,

$$\frac{\bar{V}(\hat{\delta}_{21})}{\delta_{21}^2} = \frac{1 - \rho_{12}^2}{n \rho_{12}^2} \qquad (0.45)$$

1/ For some sequence of random variables

$$g_1, g_2, \ldots, g_n, \quad \bar{E}(g_n) = \lim_{n \to \infty} E(g_n)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where ρ_{12} is the population correlation coefficient between x_1 and x_2.

So, the condition in (0.44) becomes

$$\frac{\beta_2^2}{\bar{V}(\hat{\beta}_2)} \leq \frac{n\,\rho_{12}^2}{1 + (n-1)\,\rho_{12}^2} \tag{0.46}$$

The above result has been extended to the case where the true regression model consists of k regressors from which one regressor has been omitted.

The true regression equation is

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \tag{0.47}$$

$y, x_1, x_2, \dots, x_k$ and ε are $(n \times 1)$ vectors. Each of the variables has been taken as a deviation from its mean. ε is spherical with the variance covariance matrix $\sigma^2 I_n$. ε is also independent of x's.

From the model (0.47) x_k has been omitted. So, the misspecified model is

$$y = \beta_1^+ x_1 + \beta_2^+ x_2 + \dots + \beta_{k-1}^+ x_{k-1} + \varepsilon^+ \tag{0.48}$$

Here $MSE(\hat{\beta}_1^+) \leq MSE(\hat{\beta}_1)$ if and only if

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\frac{\sigma^2}{n\,\beta_k^2\,\sigma^2_{k.123...k-1}} \geq 1 + \frac{1}{\theta^2_{\delta_{k1.23...k-1}}} \tag{0.49}$$

where $\sigma^2_{k.123...k-1} = \bar{E}(S^2_{k.123...k-1})$,

$S^2_{k.123...k-1}$ is the residual sum of squares in the linear regression of x_k on $x_1, x_2, \ldots, x_{k-1}$.

$$\theta^2_{\delta_{k.123...k-1}} = \frac{\rho^2_{1k.23...k-1}\,\dfrac{\sigma^2_{k.23\,...\,k-1}}{\sigma^2_{1.23\,...\,k-1}}}{\bar{V}(\hat{\delta}_{k1.23\,...\,k-1})}$$

$$\sigma^2_{1.23\,...\,k-1} = \bar{E}(S^2_{1.23\,...\,k-1})$$

where $S^2_{1.23\,...\,k-1}$ is the residual sum of squares in the linear regression of x_1 on $x_2, x_3, \ldots, x_{k-1}$.

$\rho_{1k.23...k-1}$ is the partial correlation coefficient between x_1 and x_k conditional on $x_2, x_3 \ldots x_{k-1}$

$$\hat{\delta}_{k1.23...k-1} = \text{the 1st element in } \{(X^{+\prime}, X^{+})^{-1} X^{+\prime}\, x_k\}.$$

$$\text{where} \quad X^{+} = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{(k-1),1} \\ x_{12} & x_{22} & \cdots & x_{(k-1),2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & & x_{(k-1),n} \end{pmatrix}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Under the assumption that $x_1, x_2, \ldots, x_k$ jointly follow a k variate normal distribution, the condition in (0.49) becomes

$$\theta^2_{\beta_k} = \beta^2_k \Big/ \left(\frac{\sigma^2}{n\, \sigma^2_{k.123\ldots k-1}} \right) \leq \frac{n\, \rho^2_{1k.23\ldots k-1}}{1 + (n-1)\, \rho^2_{1k.23\ldots k-1}} \qquad (0.50)$$

Mc Callum (1972) and Wickens (1972) showed that the asymptotic bias in estimating a regression coefficient by using a proxy for an otherwise relevant independent variable in a multiple regression model is smaller than that obtained by discarding the proxy altogether. Aigner (1974) expanded this analysis to consider the variance in addition to bias in the criterion function.

Aigner considered the model

$$y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad (0.51)$$

where x_1, x_2 and y jointly follow a trivariate normal distribution.

$\varepsilon \sim N(0, \sigma)$ and $(x_1, x_2) \sim N_2(\underset{\sim}{0}, \Sigma)$

$$\text{where} \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

A simple random sample of observations are available on y, x_1 and x^*_2 where

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$x_2^* = x_2 + u \qquad (0.52)$$

where $u \sim N(0, \sigma_u)$. u is assumed to be independent of all other basic variables.

Let $\hat{\beta}_{1P}$ be the OLS estimate of β_1 in (0.51). Obtaining sufficient condition under which $MSE(\hat{\beta}_{1P})$ is greater that $MSE(\hat{\beta}_1^+)$,[2/] Aigner observed that although the inclusion of proxy is not a superior strategy unequivocally, it can be recommended over a broad range of empirical situations.

However, the formula for $MSE(\hat{\beta}_1^+)$ obtained by Aigner was wrong. In fact, the true result for $MSE(\hat{\beta}_1^+)$ will have some more terms than those considered by Aigner. In this chapter we have rectified this mistake.

We have obtained the sufficient condition under which $MSE(\hat{\beta}_{1P}) < MSE(\hat{\beta}^+)$. Our conclusion is to include proxy in most of the cases. Our examination covers a wider range of values of ρ_{12} (the correlation coefficient between x_1 and x_2), n(the sample size) and

$$\lambda = \frac{\sigma_u^2}{\sigma_{22} + \sigma^2}.$$

~~than those considered by Aigner.~~

Considering the general case of k regressors we have also derived a necessary and sufficient condition under which

---

2/ β_1^+ is same as in (0.42).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\mathrm{MSE}(\hat{\beta}_{1P}) < \mathrm{MSE}(\hat{\beta}_1^+)$.

Let the true regression model be

$$y = x_1\beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \varepsilon \tag{0.53}$$

where $y_1, x_1, x_2, \dots, x_k$ jointly follow a (k+1) variate normal distribution. $\varepsilon \sim N(0,\sigma)$. $(x_1, x_2, \dots, x_k) \sim N_k(\underset{\sim}{0}, \Sigma)$, where

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

Suppose x_k has been observed with error. So,

$$x_k^* = x_k + u \tag{0.54}$$

where $u \sim N(0, \sigma_u)$ and u is independent of all other basic variables.

$\lim_{n\to\infty} \mathrm{MSE}(\hat{\beta}_{1P}) \le \lim_{n\to\infty} \mathrm{MSE}(\hat{\beta}_1^+)$ [3/] if and only if

$$\theta^2_{\beta_k} = \beta_k^2 \Big/ \left\{ \frac{\sigma^2}{n\,\sigma^2_{k.23\ldots k-1}(1-\rho^2_{1k.23\ldots k-1})} \right\}$$

$$\le \frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{\theta^2_{\delta_{k1.23\ldots k-1}} \left\{ 1 + \frac{\sigma_u^2}{\sigma^2_{k-1.23\ldots k-1}(1+\rho^2_{1k.23\ldots k-1}) + \sigma_u^2} \right\} + 1} \tag{0.55}$$

3/ β_1^+ is same as in (0.48).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and ~~for large n,~~

$$\theta^2_{\delta_{k1.23...k-1}} = \frac{n.\rho^2_{1k.23...k-1}}{1-\rho^2_{1k.23...k-1}}$$

where $\sigma^2_{k.23...k-1}$, $\rho^2_{1k.23...k-1}$ and $\delta_{k1.23...k-1}$ are same as before.

If sufficient data are available as in the book by Griliches and Ringstad (1971), (<u>vide</u> also Aigner 1974), the condition (0.55) can be verified easily.

<u>Chapter 5.</u> <u>Handling of errors in variable models with trending or autocorrelated errors.</u>

In this chapter we have considered the problems of estimation relating to (E-V) models where the errors of observations have either trend components or the errors are autocorrelated. Both linear and exponential trends have been considered.

(i) <u>Errors in variable models with linear trend in errors.</u>

The true model is

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1,2,...,n \qquad (0.56)$$

where ε_t is the spherical disturbance term with variance σ^2. x is stochastic and ε is independent of x.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{Let} \quad x_t^* = x_t + u_t$$

$$\text{and} \quad y_t^* = y_t + v_t \tag{0.57}$$

For linear trend,

$$u_t = c_1 + c_2 t + \tilde{u}_t$$

$$v_t = d_1 + d_2 t + \tilde{v}_t \tag{0.58}$$

$\tilde{u}_t$ and $\tilde{v}_t$ are spherical with mean 0 and variances $\sigma^2_{\tilde{u}}$ and $\sigma^2_{\tilde{v}}$ respectively. $\tilde{u}_t$'s and $\tilde{v}_t$'s are serially and mutually independent and independent of x and y. These are also independent of ε_t's. The E-V model is

$$y_t^* = \alpha + \beta x_t^* + \gamma t + \xi_t \tag{0.59}$$

where $\alpha = (\alpha + d_1 - \beta c_1)$

$$\gamma = (\alpha_2 - \beta c_2)$$

$$\xi_t = \varepsilon_t + \tilde{v}_t - \beta \tilde{u}_t$$

For exponential trend,

$$u_t = A_u e^{-d_u t} + \tilde{u}_t$$

$$\text{and} \quad v_t = A_v e^{-d_v t} + \tilde{v}_t \tag{0.60}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Assumptions about $\tilde{u}_t$ and $\tilde{v}_t$ are same as before.

So, the E-V model is

$$y_t^* = \alpha + \beta x_t^* - \gamma e^{-d_u t} + \delta e^{-d_v t} + \xi_t$$

$$= \alpha + \beta x_t^* + (\delta - \gamma) e^{-dt} + \xi_t$$

$$= \alpha + \beta x_t^* + \tilde{\gamma} e^{-dt} + \xi_t \qquad (0.61)$$

where $d_u = d_v = d$

$\gamma = \beta A_u$, $\delta = A_v$ and $\tilde{\gamma} = (\delta - \gamma)$.

For estimating such E-V models, instrumental variable technique has been adopted. Assuming $\{x_t\}$ series to be serially correlated, the matrix of instrumental variables can be taken as

$$Z_{(1)} = \begin{pmatrix} 1 & x_1^* & 2 \\ 1 & x_2^* & 3 \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}^* & n \end{pmatrix} \qquad (0.62)$$

For the case where u_t and v_t have exponential trends, we have assumed d to be known. A search procedure for estimating d could be suggested. But it is not easy to verify whether such estimates are reasonale.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In both the cases of linear trend and exponential trend in errors, the asymptotic efficiency of the I-V estimator of β with respect to that of the OLS estimator of β (in the case where $\tilde{u}_t = 0 \ \forall\ t$ and OLS is consistent.) is given by

$$E_1 = \rho_1^2 \quad \text{(provided } \{x_t\} \text{ is stationary)} \tag{0.63}$$

where $$\rho_1 = \bar{E} \frac{\sum_{t=1}^{n-1} (x_t - \bar{x}')(x_{t+1} - \bar{x}')}{\sum_{t=1}^{n-1} (x_t - \bar{x}')^2}, \qquad \bar{x} = \frac{\sum_1^n x_t}{n-1}$$

Following Karni and Weissman (1974) alternative I-V estimators of α, β, γ in (0.59) or α, β, $\tilde{\gamma}$ in (0.61) can be obtained by considering the matrix of instrumental variables

$$Z_{(2)} = \begin{pmatrix} 1 & x_3^* + x_1^* & 2 \\ 1 & x_4^* + x_2^* & 3 \\ \vdots & \vdots & \vdots \\ 1 & x_n^* + x_{n-2}^* & n-1 \end{pmatrix} \tag{0.64}$$

Here, the asymptotic efficiency of the I-V estimator of β with respect to that of the OLS estimator of β (in both the cases of linear trend in errors and exponential trend in errors) will be given by

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E_2 = \frac{2\rho_1^2}{(1+\rho_2)} > E_1 = \rho_1^2 \qquad (0.65)$$

where $\rho_2 = \bar{E}\left[\frac{\sum_{t=2}^{n-1}(x_{t+1}-\bar{x}_+)(x_{t-1}-\bar{x}_-)}{\sum_{t=1}^{n-1}(x_t-\bar{x}')^2}\right]$

and $\bar{x}_+ = \frac{\sum_{t=2}^{n-1} x_{t+1}}{(n-2)}$, ~~$\bar{x}'$ is same as before~~.

$\bar{x}_- = \left(\sum_{t=2}^{n-1} x_{t-1}\right)/(n-2)$

However, the methods of estimation suggested here can be used rather routinely even when the errors do not have a trend component.

Under the assumption that $\tilde{u}_t$'s are independently, symmetrically and identically distributed we have obtained up to order $\frac{1}{n}$, the small sample bias of I-V estimator of β in both the cases of linear trend and exponential trend in errors.

Following Gurian and Halperin (1971), under the assumptions that $\tilde{u}_t$'s are independently, identically and normally distributed, we have obtained the exact small sample biases of OLS estimators of β both including and excluding the time variable from the models (0.59) and (0.61).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In this chapter we have also considered the case where the measurement errors u_t's and v_t's are correlated. The true model is given by (0.56). The E-V model is

$$y_t^* = \alpha + \beta x_t^* + \xi_t \tag{0.66}$$

where

$$y_t^* = y_t + v_t$$
$$x_t^* = x_t + u_t \tag{0.67}$$

u_t and v_t have mean 0. u_t and v_t are mutually independent and independent of true values u_t and v_t are also independent of ε_t. But both the series $\{u_t\}$ and $\{v_t\}$ are serially correlated.

Let

$$\varepsilon_t = \rho_o \varepsilon_{t-1} + w_{ot}, \quad |\rho_o| < 1$$
$$u_t = \rho_1 u_{t-1} + w_{1t}, \quad |\rho_1| < 1 \tag{0.68}$$
$$v_t = \rho_2 v_{t-1} + w_{2t}, \quad |\rho_2| < 1$$

w_{ot}, w_{1t} and w_{2t} are spherical with mean 0 and variances $\sigma^2_{w_o}$, $\sigma^2_{w_1}$, and $\sigma^2_{w_2}$. For estimating α and β, I-V technique has been adopted. Under the assumption that u_t's are so small that rank ordering of the observed x^*'s give the rank ordering of true x's, the matrix of instrumental variables is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$Z = \begin{pmatrix} 1 & r_1 \\ 1 & r_2 \\ \vdots & \vdots \\ 1 & r_n \end{pmatrix}^{4/} \qquad (0.69)$$

where r_i is the rank of x_i^*.

We have also discussed the methods of estimating the variance-covariance matrix(V) of ξ consistently. Moreover, the following special cases

(a) $\rho_o = 0$

(b) $\rho_o = \rho_2 = 0$

(c) $\frac{\sigma_u^2}{\sigma_v^2} = \lambda$ (known)

and) $\rho_o = 0$ and $\rho_1 = \rho_2 = \rho^*$

have been discussed in details.

However, the case of autocorrelated errors is much more complicated than the case of trending errors. The workability of the methods of estimation suggested (specially for estimating V) in this case remains to be experimentally verified.

Chapter 6. Concluding Observations.

In this chapter we take an overview of the entire investigation and stressed the results reached and their significances. We have also pointed to the directions to which further researches can be made.

4/ This is in fact, an extension of Durbin's (1954) method of using the ranks of the observed x_t^*'s as instruments.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 1

A SURVEY OF PREVIOUS RESEARCHES

1.1 Introduction

Most of the econometric theory is mainly concerned with the estimation and testing of relationships among economic variables. For this, the first step is to specify the relationship or the model in mathematical form. The next step is to collect data relating to the economy or the sector to which the model refers. Finally, we use different statistical methods for estimating the parameters of the model and judge by suitable tests whether the model provides a realistic picture of the phenomenon being studied.

Although economic theory often specifies the exact functional relationships among its variables, a careful examination of economic data shows that no such exact relationships exist in reality. So, the task of econometric theory is to provide a link between the exact relationships of theory and the observed relationships of reality. So, need is felt for the specification of a stochastic error term in each relation. Hence, the probabilistic version of econometrics is that it specifies a functional relationship in which the observed independent variables and unobserved disturbances determine an observed dependent variable. The statistical properties of the disturbances are also specified.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Very often in econometric practice one adopts the classical linear regression model. The classical linear regression model is given by

$$y_i = x_{1i}\,\beta_1 + x_{2i}\,\beta_2 + \dots + x_{ki}\,\beta_k + \varepsilon_i, \quad i=1,2,\dots,n \tag{1.1.1}$$

or, in matrix notation,

$$y = X\beta + \varepsilon \tag{1.1.2}$$

where

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$X = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & & \\ x_{1n} & x_{2n} & \dots & x_{kn} \end{pmatrix}$$

$$= (x_1, \; x_2 \; \dots \; x_k)$$

very often the first column of X is $(1,1, \dots, 1)'$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \text{and} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The underlying assumptions are

$$E(\varepsilon) = \underset{\sim}{0} \tag{1.1.3}$$

$$E(\varepsilon\varepsilon') = \sigma^2 I_n \tag{1.1.4}$$

X is nonstochastic and fixed in repeated samples. Rank of $X = k \leq n$.

In such a case, the ordinary least squares(OLS)method give best lineare unbiased estimates of the regression coefficients β's in (1.1.1). If in addition ε is assumed to be normally distributed, the model is called classical normal linear regression model. Standard methods of interval estimation and hypothesis testing are based on this model.

But the assumptions on ε's and X's may not be fulfilled in reality, or, in other words,the model may not be correctly specified. The literature on econometric methods [vide Johnston (1972), Goldberger(1964), Theil (1971) etc.] is largely concerned with the statistical problems that arise when one or more of the assumptions of the classical model is found to be unrealistic.

One class of problem arises when some of the regressors are omitted from the equation and/or some additional regressors are wrongly included in the model, or when the algebraic form of the regression equation is misspecified. In such cases, OLS method

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

would fail to give satisfactory estimates of the regression coefficients.

Another class of problems is created by nonspherical disturbances. Here $E(\varepsilon\varepsilon') \neq \sigma^2 I_n$. When the diagonal elements of the dispersion matrix are unequal, but the off-diagonal elements are zero, we have the problem of heteroscedasticity. If, on the other hand, the diagonal elements are equal but the off-diagonal elements are non-zero, we have the problem of intercorrelated disturbances. Both the above two complications may also be present in a given problem. In such cases, OLS procedures of point estimation may not be optimal and interval estimation and hypothesis testing may be seriously wrong. Generalised least squares techniques are called for in such situations.

Problems also arise when the regressors (X) are stochastic. There is little trouble if X is stochastic but fully independent of ε. However, if the regressors and disturbances are correlated, OLS estimates cease to be unbiased. This is the case where one of the regressors is a lagged value of the Y variable. The danger is particularly great if the regressor values and the disturbances in the t-th observational equation in (1.1.1) are correlated. In this case, OLS estimates of β's are not even asymptotically unbiased. In this case, the least squares

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

estimates will be inconsistent. This kind of complication arises in two important situations :

(a) where the regressors are observed with errors

and (b) where the equation is embedded in simultaneous equation models where several current endogenous variables are determined through the simultaneous interactions of the structural relationships in the model.

This chapter is intended to give a brief survey of the existing literature on the following problems of econometric methodology.

1. Omission of regressors from econometric relationships, or more generally, the misspecification of algebraic forms of these relationships.

2. Autocorrelation of disturbances.

3. Presence of errors in observations on the regressors.

This survey is mostly confined to the discussion of single equation econometric models.

1.2 Omission of regressors and misspecification of algebraic forms.

1.2.1 Consequences of omission of regressors and misspecification of algebraic forms. A regressor may be omitted from an economic

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

relationship for three main reasons :

(a) The importance of the particular regressor may be unknown to the analyst.

(b) Reliable data on the regressor may not be available and the regressors may have to be replaced by a substitute, and

(c) Inclusion of the regressor may increase the risk of multicollinearity.

When a regressor is omitted or the algebraic form of the equation is wrongly specified, estimation of the regression coefficients in the misspecified equation poses some serious problems. This was first studied by Theil (1957). He considered the following regression model

$$y = X\beta + \varepsilon \qquad (1.2.1)$$

where X is a matrix of order $n \times k$ with rank k. The element in the i-th row and the j-th column of X is denoted by x_{ji}, $i = 1,2, \ldots, n$; $j = 1,2, \ldots, k$, It is further specified that the elements of X are real and nonstochastic ; y is a column vector of n real elements having y_i, $i = 1,2, \ldots, n$ as the i-th element and β is a $k \times 1$ vector of real parameters; ε is an $n \times n$ vector of disturbances and $E(\varepsilon) = 0$. Given these, the OLS estimate of β is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\hat{\beta} = (X'X)^{-1} X'y \qquad (1.2.2)$$

We know that $\hat{\beta}$ is an unbiased estimator of β and this $\hat{\beta}$ is BLUE if it is further assumed that

$$E(\varepsilon\varepsilon') = \sigma^2 I_n .$$

Now suppose that instead of using X as the matrix of observations, one uses some other real nonstochastic matrix X^+ of order $n \times k^+$ with rank k^+. So, the misspecified model may be written as

$$y = X^+ \beta^+ + \varepsilon^+ \qquad (1.2.3)$$

$$\therefore \quad E(\varepsilon^+) = X\beta - X^+ \beta^+$$

provided X^+ is also nonstochastic, and this implies $E(\varepsilon^+) \neq 0$ in general, for whatever β. If however, the columns of X and those of X^+ are linearly dependent, then we may have $E(\varepsilon^+) = 0$ for some choice of β^+. Now from (1.2.3)

$$\hat{\beta}^+ = (X^{+'} X^+)^{-1} (X^{+'} y)$$

$$\therefore \quad E(\hat{\beta}^+) = (X^{+'} X^+)^{-1} X^{+'} X\beta$$

$$= P\beta \qquad (1.2.4)$$

P is given by the auxilary regression equations

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$X = X^{+}P + \text{matrix of residuals} \qquad (1.2.5)$$

In the special case, suppose $x_1, x_2, \ldots, x_{k-1}$ are correctly specified and included in the regression equation and x_k is replaced by x_k^{+}. Thus $X = (x_1, x_2, \ldots, x_{k-1}, x_k)$ and $X^{+} = (x_1, x_2, \ldots, x_{k-1}\ x_k^{+})$.

Here, $$E(\beta_h) = \beta_h + P_{hk}\,\beta_k, \quad h = 1,2,\ldots,k-1, \qquad (1.2.6)$$

but $$E(\beta_k^{+}) = P_{kk}\,\beta_k \qquad (1.2.7)$$

where P's are given by the regression equation

$$x_{ki} = \sum_{h=1}^{k-1} P_{hk}\,x_{hi} + P_{kk}\,x_{ki}^{+} \qquad (1.2.8)$$

where x_{hi} and x_{ki}^{+} are the elements in the i-th row and h-th column, h=1,2,...k of X and i-th row and k-th column of X^{+} respectively.
The specification bias of the estimator $\hat{\beta}_h^{+}$ is given by

$$E(\hat{\beta}_h^{+}) - \beta_h = P_{hk}\,\beta_k, \quad h = 1,2,\ldots,k-1 \qquad (1.2.9)$$

The estimators $\hat{\beta}_1^{+}, \hat{\beta}_2^{+}, \ldots, \hat{\beta}_{k-1}^{+}$ have no specification bias if the corresponding regressors are uncorrelated with the incorrectly specified X_k.

Theil (1957) also considered the following special cases :

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(a) The true model is

$$y_i = \beta_o + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i \ , \quad i = 1,2,\ldots,n \qquad (1.2.10)$$

But the analyst fits a linear relationship of the form

$$y_i = \beta_o^+ + \beta_1^+ x_i + \varepsilon_i^+ \ , \quad i = 1,2,\ldots,n \qquad (1.2.11)$$

It is assumed that x_i is measured as a deviation from its mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ .$$

Here $E(\hat{\beta}_1^+) = \beta_1 + \beta_2 r$ (1.2.12)

where $r = \frac{m_3}{m_2} = \frac{\sum_{i=1}^{n} x_i^3}{\sum_{i=1}^{n} x_i^2}$.

Since the slope of (1.2.10) at the centre of gravity ($\bar{x} = 0$), is equal to β_1, we may conclude that $\beta_2 \frac{m_3}{m_2}$ is the specification bias.

Theil (1957) also gave some interesting results for the case where the correct specification for a bivariate relationship is linear in logarithms and the incorrect specification is linear in the variables themselves. Correct specification is

$$y_i = A \, x_i^{\beta} \, e^{\varepsilon_i} \qquad i = 1,2,\ldots,n \qquad (1.2.13)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where one fits

$$y_i = a + b\,x_i \quad , \quad i = 1,2,\ldots,n \qquad (1.2.14)$$

ε's are normally and independently distributed with common variance σ^2. Then expectation of elasticity (η) obtained from the fitted line at the centre of gravity is

$$E(\eta) = e^{\frac{1}{2}\sigma^2}\,\beta\left\{1 + \tfrac{1}{2}(\beta - 1)\,\gamma_1 \nu + \ldots\right\} \qquad (1.2.15)$$

where γ_1 and ν are the coefficients of skewness and of variation respectively of the independent variable x. The result (1.2.15) shows that here two types of errors have been made. One is the error due to nonlinearity of the true model. This is given by the expression between the curled brackets. If $\beta = 1$, there is no specification bias resulting from this source. Because then the relationship between x and y is linear. The second source of error is concerned with the treatment of the disturbances. log y contains the disturbance term ε which is normally distributed with zero mean and variance σ^2. But the specification adopted deals with y (not with log y). This leads to the factor $e^{\frac{1}{2}\sigma^2}$ which is always > 1, so that, for this reason, η is biased away from zero.

<u>1.2.2 Application of specification analysis</u>. Application of specification analysis can be found in the theory of aggregation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

of microeconomic relations (vide Theil (1954), Theil (1971). The matrix representation of this theory is due to Kloek (1961). Let us consider that there are N economic agents, each of which is characterized by a behavioral equation of the type

$$y_i = X_i \beta_i + \varepsilon_i , \quad i = 1,2,\ldots,N \tag{1.2.16}$$

where y_i and ε_i are column vectors each containing n elements. X_i is a matrix of order $n \times k$. The assumption is that for each pair of agents (i,j) any column of (y_i, X_i) has the interpretation as the corresponding column of (y_j, X_j). Only these two columns refer to two different agents. Moreover, in each X_i, let the elements in the first column be unity. The parameter vectors $\beta_1, \beta_2, \ldots, \beta_N$ of N agents can be arranged in a $k \times N$ matrix :

$$(\beta_1,\beta_2,\ldots,\beta_N) = \begin{pmatrix} \beta_{11} & \beta_{12} \cdots & \beta_{1N} \\ \beta_{21} & \beta_{22} \cdots & \beta_{2N} \\ \vdots & \vdots & \\ \beta_{k1} & \beta_{k2} \cdots & \beta_{kN} \end{pmatrix} \tag{1.2.17}$$

Now, from (1.2.16) we have

$$\bar{y} = \frac{1}{N} \sum_{i=1}^{N} X_i \beta_i + \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i \tag{1.2.18}$$

where $\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

But, no researcher will like to work with the above model because the labour would be prohibitive. He will perfer to estimate a macro relation of the form

$$\bar{y} = \bar{X}\beta + \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \qquad (1.2.19)$$

where $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$.

Here the OLS estimator of β is

$$\hat{\beta} = (\bar{X}'\bar{X})^{-1}\bar{X}'\bar{y} . \qquad (1.2.20)$$

Making usual assumptions,

$$E(\hat{\beta}) = \sum_{i=1}^{N} B_i \beta_i$$

where $B_i = (\bar{X}'\bar{X})^{-1}\bar{X}'\frac{1}{N}X_i \qquad i = 1,2,\ldots,N.$

B_i is the coefficient matrix of the auxiliary regression where $\frac{1}{N}X_i$ is regressed on $\bar{X}$. Writing these regressions as

$$\frac{1}{N}x_{\alpha hi} = \sum_{j=1}^{N} b_{jhi}\,\bar{x}_{\alpha j}$$

where $x_{\alpha hi}$ and $\bar{x}_{\alpha h}$ are respectively the (α, h)-th elements of X_i and $\bar{X}$ and b_{jhi} is (j,h)-th element of B_i.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let $\hat{\beta}_h$ be the h-th element of the vector β

$$\therefore E(\hat{\beta}_h) = \sum_{i=1}^{N} b_{hhi}\, \beta_{hi} + \sum_{j\neq h} \left(\sum_{i=1}^{N} b_{jhi}\, \beta_{ji} \right) \qquad h = 1,2,\ldots,k$$

$$= \frac{1}{N} \sum_{i=1}^{N} \beta_{hi} + \sum_{i=1}^{N} (b_{hhi} - \frac{1}{N})\, \beta_{hi} + \sum_{j\neq h} \left(\sum_{i=1}^{N} b_{jhi}\, \beta_{ji} \right) \qquad (1.2.21)$$

Since $\sum_{i=1}^{N} B_i = I_k$, $\sum_{i=1}^{N} b_{jhi} = 1$ if $h = j$

$= 0$ if $h \neq j$

The second and the third terms in (1.2.21) together give the aggregation bias of the macro coefficient β_h .

Another application of the specification analysis can be found in the theory of grouping of observations (Haitovsky(1973)). This can be illustrated in the case of a two regressor model

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad (1.2.22)$$

where y_1, x_1 and x_2 are $n \times 1$ column vectors.

This can be written in terms of deviations from overall means as

$$y = \beta_1 x_{1.0} + \beta_2 x_{2.0} + \varepsilon_o \qquad (1.2.23)$$

where y_o, $x_{1.0}$, $x_{2.0}$ and ε_o denote deviations from their respective means.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Suppose we have two tables, one showing averages of all values for each of a number of intervals of x_1 and the other giving averages of all the values for each of a number of intervals of x_2. Haitovsky proposed to obtain an estimate $\tilde{\beta}_1$ of β_1 by regressing y_o on $x_{1,0}$ using the table based on x_1 classification, and an estimate $\tilde{\beta}_2$ of β_2 by the simple regression of y_o on $x_{2,0}$ from the tables based on x_2 classification. Both the estimates will be biased, because in each case we have left out a relevant regressor from the regressor equation

$$E(\tilde{\beta}_1) = \beta_1 + \beta_2 \frac{\sum_{i=1}^{n(2)} x_{1.0,i}\, x_{2.0,i}}{\sum_{i=1}^{n} x^2_{1.0,i}} \qquad (1.2.24)$$

$$E(\tilde{\beta}_2) = \beta_1 \frac{\sum_{i=1}^{(2)} x_{2.0,i}\, x_{1.0,i}}{\sum_{i=1} x^2_{2.0,i}} + \beta_2 \qquad (1.2.25)$$

where the superscripts ((1) and (2)) on the summation sign indicates the table used for computation. Haitovsky's technique consists in solving for β_1 and β_2 from the equation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\beta_1 = \hat{\beta}_1 + \hat{\beta}_2 \frac{\sum_{i=1}^{(1)} x_{1.0,i}\, x_{2.0,i}}{\sum_{i=1} x_{1.0,i}^2} \tag{1.2.26}$$

$$\beta_2 = \tilde{\beta}_1 \frac{\sum_{i=1}^{(2)} x_{2.0,i}\, x_{1.0,i}}{\sum_{i=1}^{(2)} x_{2,0,i}^2} + \beta_2 \tag{1.2.27}$$

Obviously these solutions are unbiased estimators of β_1 and β_2. Finally,

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 \tag{1.2.28}$$

where $\bar{y}$, $\bar{x}_1$ and $\bar{x}_2$ are the respective means of y, x_1 and x_2.

An early application of specification analysis was made by Griliches (1957) in estimating the returns to scale in production when an input variable (e.g. managerial services) has been omitted from the Cobb-Douglas production function. Under usual assumptions, the omission of managerial inputs biases the estimate of the elasticity of output with respect to capital inputs upwards and the estimate of returns to scale downward. If the quality differences are disregarded in the measure of labour input, elasticity of labour inputs becomes downwardly biased, that of capital inputs upwardly biased and estimates of returns to scale

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

become downwardly biased. If geometric averages and sums are used when aggregating the inputs, there will be no bias unless there is a strong association between the individual elasticities and the levels of the corresponding inputs. But the estimates are not good. Even when we use the arithmetic sums, the elasticities may not lead to bias (except in the constant term) if the inputs being aggregated are used in approximately fixed proportions.

Mundlak (1961) (See also Hoch 1955), suggested a method of estimating the parameters of production functions free of management bias. Management input is generally omitted because of lack of units for its direct measurement. The underlying assumption is that the management does not change considerably over time, and for short periods, say, a few years, it remains constant. For the purpose of analysis, it is sufficient that the management remains constant for a period of two years, since at least two observations on each firm are required. The procedure deals with the linear form of a production function. So, this includes the Cobb-Douglas function where the variables are written in Logarithms. Analysis of covariance has been used to obtain unbiased estimates of the coefficient of the linear form of the production function. The bias of the estimates due to omission of management input has been evaluated and then

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

management has been estimated up to a multiplier. The method, however, does not solve the problem of estimating the conditional expectation of output for a given bundle of resources since the value to be estimated depends on the value of management, and therefore varies from firm to firm. As long as management is not measurable, it is not possible to solve such problems. One possibility is to ignore the management variable. In this case, the inference is directed to the average firm under the assumption that the management variable in the sample is measured from its average. Another possibility is to use the interfirm regression. Since it is subject to management bias, the correlation between management and other inputs will be reflected here. So, this can be used for the purpose of inferring about the productivity of a firm whose level of management is known.

1.2.3 Tests of specification errors. Ramsey (1969) (see also Ramsey 1974) suggested four different procedures of testing for the presence of specification errors in classical linear regression model. The specification errors considered are (a) omitted variables, (b) incorrect functional form, (c) simultaneous equation problems and (d) heteroscedasticity.

The model considered for regression is

$$y = X\beta + \varepsilon^{+} \tag{1.2.29}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where y is an $n \times 1$ vector. X is an $n \times k$ matrix of non-stochastic regressors of rank k, β is a $k \times 1$ vector of regression coefficients and ε^{+} is an $n \times 1$ vector of independent disturbances terms normally and identically distributed with mean zero and variance σ^2. Equation (1.2.29) specifies the null hypothesis. For each specification error, the alternative hypothesis is defined by specifying that the true model has some other specification.

Let the specification of the true models be

$$y = X\beta + z\gamma + \varepsilon \quad (1.2.30)$$

$$y = Z\gamma + \varepsilon \quad (1.2.31)$$

$$y = Z\gamma + W\delta + \varepsilon \quad (1.2.32)$$

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(\underset{\sim}{0}, \sigma^2 \quad) \quad (1.2.33)$$

In model (1.2.30), the specification of y, X and β are same as in (1.2.29). z is an $n \times 1$ nonstochastic regressor vector and γ is the corresponding coefficient, $\varepsilon \sim N(\underset{\sim}{0}, \sigma^2 I_n)$. If one uses (1.2.29) to estimate β, one has a misspecified model where the misspecification is that of omitting the regressor z. In model (1.2.31), Z is an $n \times k$ matrix of nonstochastic regressors and $\varepsilon \sim N(\underset{\sim}{0}, \sigma^2 I_n)$. If one considers a matrix X whose elements are obtained from the elements of the matrix Z by nonlinear transformations, and one uses (1.1.29) as the regression

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

model instead of (1.2.31), one would have a misspecified model, the misspecification being that of incorrect functional form of regressors. In (1.2.32), let y be an $n \times 1$ regressand vector, Z a $n \times k_1$ matrix of nonstochastic regressors of full rank with γ the corresponding coefficient vector and W an $n \times k_2$ matrix of stochastic regressors of rank k_2 in observed samples with δ the corresponding coefficient vector and $\varepsilon \sim N(0, \sigma^2 I_n)$, where the elements in each row of W and the corresponding ε_i, $i = 1, 2, \ldots, n$ are statistically dependent. Defining $X = (Z \vdots W)$ an $n \times k$ $(k = k_1 + k_2)$, matrix and $\beta' = (\gamma', \delta')$, a $k \times 1$ vector, if one were to assume that the model $y = X\beta + \varepsilon$ so defined satisfies the conditions given in (1.2.29), one has a misspecified model in which the misspecification is denoted by simultaneous equation problem. In (1.2.33) all the specifications in model (1.2.29) are correct except that $\sigma^2\Omega$ is diagonal and has unequal elements on the diagonal. Here, if one uses model (1.1.29), the specification error of heteroscedasticity is said to have been made.

Next, Ramsey obtained Theil's (1965) BLUS residual vector $\breve{e}^+ = (\breve{e}_1, \breve{e}_2 \ldots \breve{e}_{n-k})'$ defind by

$$\breve{e}^+ = A'y \tag{1.2.34}$$

where the $(n - k) \times n$ matrix A satisfies the conditions

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$A'X = 0$$

and $$A'A = I_{n-k} \qquad (1.2.35)$$

Ramsey showed that for model specified in (1.2.30), (1.2.31) and (1.2.32), the use of the equation (1.2.29) as the true regression model leads to a BLUS residual vector $\tilde{e}^{+}$ whose distribution is $N(A'\xi, \bar{\sigma}^2 I_{n-k})$ where ξ is a $(n - k) \times 1$ nonstochastic vector whose precise definition depends on the particular misspecification and $\bar{\sigma}^2$ is the variance of $\tilde{e}_i^{+}$; $i = 1,2,\ldots,n$. Under quite general conditions, Ramsey showed that $A'\xi$ can be approximate by

$$A'\xi \simeq \alpha_1 q_1 + \alpha_2 q_2 + \ldots\ldots \quad \underline{1}/ \qquad (1.2.36)$$

where $(n - k) \times 1$ dimensional vectors q_j's, $j = 1,2,\ldots$ are defined by

$$q_j = A' \hat{y}^{(j+1)}, \quad j = 1,2, \ldots. \qquad (1.2.37)$$

and $$\hat{y}^{(j+1)} = (\hat{y}_1^{(j+1)}, \hat{y}_2^{(j+1)} \ldots\ldots \hat{y}_n^{(j+1)})$$ where

---

1/ If the regression (1.2.29) is not in deviation term and a constant term is not included in X, then

$$A'\xi \simeq \alpha_0 A'i + \alpha_1 q_1 + \alpha_2 q_2 + \ldots.$$

where $i = (1,1,\ldots,1)'$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\hat{y}_i^{\,j}$ is the j-th power of the least squares estimator of mean $E(y_i)$. 2/

In the case where the model (1.2.33) is true, the use of the model (1.2.29) implies that the distribution of $\tilde{e}^{+}$ (conditional on X) is $N(\underset{\sim}{0}, A'\Omega A)$.

Next we shall discuss briefly the tests developed by Ramsey. For each of the tests considered by Ramsey, the null hypothesis is

H_0 : Distribution of $\tilde{e}^{+}$ is $N(\underset{\sim}{0}, \sigma^2 I_{n-k})$ (1.2.38)

The alternative hypotheses H_i $i = 1,2$ are

H_1 : Distribution of $\tilde{e}^{+}$ is $N(A'\xi, \sigma^2 I_{n-k})$ (1.2.39)

and

H_2 : Distribution of $\tilde{e}^{+}$ is $N(\underset{\sim}{0}, \theta)$ (1.2.40)

where θ is a diagonal positive definite matrix.

1. Regression Specification Error Test (RESET)

Under the alternative hypothesis H_1, Ramsey considered the regression equation

---

2/ If however, the implicit relationship between z and y is not analytic, it may not be possible to express z in terms of a polynomial in $\hat{y}$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\tilde{e}^{+} = \alpha_o + \alpha_1 q_1 + \alpha_2 q_2 + \ldots + \alpha_p q_p + u \quad (1.2.41)$$ [3]

where u is a vector of independent disturbance term distributed as normal with zero mean and constant variance. H_o is tested by the usual F test for testing the joint significance of $\alpha_o, \alpha_1, \ldots, \alpha_p$. Under H_o, F statistic is distributed as a central F with (p+1) and (n-k-p-1) degrees of freedom.

Ramsey and Schmidt (1976) described PESET procedure using a regression equation of the form

$$\tilde{e}^{+} = \alpha_1 q_1 + \alpha_2 q_2 + \ldots + \alpha_p q_p + u \quad (1.2.42)$$

(i.e., there is no intercept term).

Here also a central F statistic with p and (n-k-p) degrees of freedom is used to test the joint significance of $\alpha_1, \alpha_2, \ldots, \alpha_p$. Ramsey and Schmidt (1976) denoted this as BLUS variant A test and the RESET procedure developed by Ramsey (1969) as BLUS variant B test. Whether BLUS variant A or BLUS variant B is more powerful will generally depend on the alternative hypothesis. However, the two tests appear to give almost identical results in some limited Monte Carlo experiments that have been done.

3/ How many q_j's are needed will depend on the particular problem. Ramsey found that using q_1, q_2 and q_3 was sufficient.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Ramsey and Schmidt (1976) also showed that BLUS variant A test is equivalent to a test obtained by regressing the OLS residuals $\tilde{e}^+$ on $M_X Q$ [where $Q = (q_1, q_2, \ldots, q_n)'$ and $M_X = [I - X(X'X)^{-1} X']$ and using the usual F test for testing the joint significance of the associated coefficients. This procedure is also identical to regressing y on (X, Q) and using the usual F statistic for testing the hypothesis that the coefficients associated with Q are zero.

2. Rank Specification Error Test (RASET)

Here Ramsey (1969) assumed that the second moment of $\tilde{e}_i^+$ varies monotonically with $q_{i1} = a_i' y^{(2)}$, $i = 1, 2, \ldots, n-k$, where a_i' is the i-th row of A. Since $\tilde{e}_i^{+2}$ is an estimate of the second moment of $\tilde{e}^+$, Ramsey considered the relationship between the rankings of $\tilde{e}_i^2$ and of q_{i1}. Rearranging q_{i1} into ascending order and permuting the elements of the vector e^+ conformably, one transforms $q_{i1} \rightarrow j$, $j = 1, 2, \ldots, n-k$ where j is the index of j-th largest q_{i1}. Then one assigns to each $\tilde{e}_i^{+2}$ an integer r_i such that r_i indicates the r-th largest value of $\tilde{e}_i^{+2}$. So n-k numbers $\{r_i\}$ is a permutation of integers $1, 2, 3, \ldots, n-k$. For testing H_o against H_1, Spearman's rank correlation coefficient given by

$$\tilde{R}_s = 1 - \frac{6}{(n-k)\,[(n-k)^2 - 1]} \sum_{i=1}^{n-k} (r_i - i)^2 \qquad (1.2.43)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

For large values one uses the asymptotic result that

$$t_R = \left\{ \frac{(n-k-2)\, R_S^2}{(1 - R_S^2)} \right\}^{\frac{1}{2}} \qquad (1.2.44)$$

is distributed as Student's t with (n-k-2) degrees of freedom under the null hypothesis.

3. Kolmogorov's Specification Error Test (KOMSET).

The distribution of the squared residuals e_i^{+2} depends upon the unknown scale factor $\bar{\sigma}^2$. RASET is invariance of this scale factor. Another way of avoiding this scale factor is to consider the distribution of

$$\eta_i = \frac{e_j^{+2}}{e_k^{+2}} \qquad j \neq k,\ j,k = 1,2, \ldots, (n-k) \qquad (1.2.45)$$

η_i's are so chosen that they are independent under H_o. Ramsey studied the distribution of η_i's under H_o and H_1. Then, for testing H_o against H_1, he applied Kolmogorov's test on these η_i's.

4. Bartlett's M specification Error Test (BAMSET).

Following Bartlett's M test for equality of variances, Ramsey (1969) also suggested a procedure for testing H_o against H_2.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In most of the models examined in Ramsey and Gilbert (1969, 1972), RASET has less power than either RESET or BAMSET and sometimes less than both.

Applications of Ramsey's tests. Ramsey (1968 a) first reported the use of his specification error tests, later fully described in Ramsey (1969). The tests were used to discriminate between eight different models showing the relationship between value added per unit of labour and wage rate. Despite the small sample sizes, the tests were fairly sensitive to various alternatives and enabled one to select a model in preference to the others. Gilbert (1969) used the tests to discriminate between alternative formulations of demand for money. Lee (1972) used them for discriminating between four alternative production functions. Ramsey and Zarembka (1971) used the tests to discriminate between the Cobb-Douglas, CES, VES, GPF and quadratic production functions using the data relating to U.S. manufacturing industries. Although all the coefficients estimates were statistically significant at 5 per cent level and all the R^2 values were greater than 0.99, CES came out as the best form of the production function and although the quadratic form had the highest value of R^2 it was most strongly rejected by specification error tests. Ramsey (1971) tested the adequacy of a market demand model fomulated in Ramsey (1972). Loeb (1976) applied the tests to compare

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

three quarterly investment models formulated by Anderson (1964, 1967), Eisner (1962) and Meyer-Glanber (1974) with data relating to thirteen manufacturing industries. The models were ranked in order of the number of times they failed to be rejected by the four Ramsey tests. This rank scheme has been compared to that found in a previous study by Jergenson, Hunter and Nadiri (JHN) (1970). In both the ranking schemes, Eisner's model proved to be the best one. In JHN scheme Meyer-Glanber model was ranked second and Anderson model last. But, according to Ramsey tests Anderson model was ranked second and Meyer-Glanber model last.

1.2.4 Residual variance criterion. Suppose the true model is given by (1.2.1) and the misspecified model is given by (1.2.2). The OLS residual for the model (1.2.2) is

$$e^{+} = y - X^{+}\hat{\beta}^{+} \tag{1.2.46}$$

Theil (1957) showed that

$$\frac{E(e^{+\prime}e^{+})}{n - k^{+}} \geq \sigma^{2} \tag{1.2.47}$$

Since $E(y - X\beta)'(y - X\beta) = \sigma^{2}$ (1.2.48)

(1.2.47) means that the residual variance estimator overestimates σ^{2} if the specification is incorrect. This provides a basis for selecting a specification with the smallest residual variance.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

But, this criterion does not work when neither specification is correct. Again, it can be applied in a straightforward manner only if the different specifications have the same regressand.

Koerts and Abrahamse (1970) pointed out that when the sample size is small, the analyst may make a wrong decision with a large probability by choosing a model with smaller residual variance.

Kloek (1975) showed that the probability of adopting the wrong model on the basis of the residual variance criterion converges to zero as the sample size increases. The true model is

$$y = X\beta + \varepsilon \tag{1.2.49}$$

where X is an $n \times k$ matrix of nonstochastic regressors. The disturbances (ε) are independently and identically distributed with zero mean.

The misspecified model is

$$y = Z\gamma + u \tag{1.2.50}$$

where Z is an $h \times h$ nonstochastic regressor matrix and u's are the disturbances. It is assumed that the space spanned by the columns of z does not contain $X\beta$. Hence, a specification which contains all the variables in the correct specification plus some irrelevant variables, is not incorrect in this set up.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let s_n^2 be the residual variance of the correctly specified model and t_n^2 be the residual variance of the incorrectly specified model. Then Kloek proved that

$$\lim_{n\to\infty} P[t_n^2 < s_n^2 + \lambda g^2] = 0 \quad \text{for every } \lambda < 1 \text{ and for a certain number } g > 0 \qquad (1.2.51)$$

Schmidt (1974) showed that the Theil's (1957) residual variance criterion still holds asymptotically when the disturbances (ε's) are autocorrelated. He considered the classical regression model

$$y = X\beta + \varepsilon \qquad (1.2.52)$$

(y is a vector of order $n \times 1$ and X is a matrix of nonstochastic regressors of order $n \times k$) with the difference that the disturbances ε's follow a Markov scheme

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t. \quad |\rho| < 1 \text{ and } u_t\text{'s are i.i.d. as } N(0,\sigma_u^2) \qquad (1.2.53)$$

Let the misspecified model be

$$y = Z\gamma + w \qquad (1.2.54)$$

(Z is of order $n \times h$).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Associated with this model, there is an estimate of ρ, say, $\hat{\rho}$ with probability limit ρ_o 4/.

So, $y_{*,t} = y_t - \hat{\rho}\, y_{t-1}$

$$= x_{*,t}\beta + u_t + (\rho - \hat{\rho})\, \varepsilon_{t-1}, \quad t = 2,2,\ldots,n \tag{1.2.55}$$

where $x_{*,t} = x_t - \hat{\rho} x_{t-1}, \quad t = 2,3,\ldots,n$

and $x_t = (x_{1t}, x_{2t}, \ldots, x_{kt})$.

Let

$$M_{Z*} = I - Z_* (Z_*' Z_*)^{-1} Z_*' \tag{1.2.56}$$

where $Z_* = (z_{*2}, \ldots, z_{*n})'$

and $z_{*t} = z_t - \hat{\rho} z_{t-1}, \quad t = 2,3,\ldots,n$

where $z_t = (z_{1t}, z_{2t}, \ldots, z_{kt})$.

Then $\sigma_u^2 = \frac{1}{n} y_*' M_{Z_*} y_*$. $\quad$ (1.2.57)

Under the assumption that as $n \to \infty$, $\frac{1}{n} Z_*' Z_*$ converges to a positive definite matrix

$$\operatorname*{plim}_{n\to\infty} \sigma_u^2 \geq \sigma_u^2 \tag{1.2.58}$$

4/ If ρ is not estimated, but chosen on some aprioригgrounds, then clearly, $\hat{\rho} = \rho_o$. Also, if the model $y = Z\gamma + w$ is not to be transforms, $\rho = \rho_o = 0$. Thus the model $y = Z\gamma + w$ may be regarded as very general.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

It should also be noted that when the model $y = Z\gamma + \omega$ contains the same regressors as the true model(i.e. $Z = X$), the probability limit of the estimated error variance is smaller when the true value of ρ is used to perform the transformation (1.2.55) than when any other value of ρ is used.

1.2.5 Using least squares to approximate unknown regression functions. White (1976) obtained some results which may be interesting for one concerned with least squares approximation for unknown regression equations, but whose relevance for econometric practice is not quite clear. He considered models like

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1,2,\ldots,n \tag{1.2.59}$$

x_i are independent, identically distributed $1 \times r$ vectors with generalised density function $dF(x)$, x being a real valued $1 \times r$ vector and $g(x_i)$ is an unknown Borel function. Moreover, x_i and ε_i are independent, $E(\varepsilon_i x_i) = 0$, $E(\varepsilon_i) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2 I_n$.

The approximation model is

$$y_i = x_i\beta + u_i, \quad i = 1,2,\ldots,n \tag{1.2.60}$$

where x_i is a $1 \times k$ vector

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\therefore \quad u_i = g(\tilde{x}_i) - x_i\beta + \varepsilon_i, \quad i = 1,2,\ldots,n \tag{1.2.61}$$

It can be shown that the OLS estimator of β is

$$\hat{\beta} = (X'X)^{-1} X'y \tag{1.2.62}$$

(where X is an $n \times k$ matrix with typical row x_i) is a constant estimator of β. So, $x_i \hat{\beta}$ is a consistent estimate of a least squares approximation to $g(\tilde{x}_i)$ with weighting function $dF(\tilde{x})$.

White also considered the case of nonspherical disturbances, i.e., the case where

$$E(\varepsilon\varepsilon') = \Omega \neq \sigma^2 I_n \tag{1.2.63}$$

and Ω is positive definite. The generalised least squares estimate of β is

$$\hat{\hat{\beta}} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1} y \tag{1.2.64}$$

is a consistent estimator of β, and so $x_i \hat{\hat{\beta}}$ is a consistent estimate of a least squares approximation to $g(\tilde{x}_i)$ with weighting function $dF(\tilde{x})$.

<u>1.2.6 Consequences of misspecification in simultaneous equation systems.</u> Fisher (1966) considered the effect of omission of variables in the context of simultaneous equation systems. Let the particular equation considered in the system be

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y = Y\gamma + X_1\beta + W\eta + \varepsilon \qquad (1.265)$$

y is an n component vector of observations on an endogenous variable, Y is an $n \times m$ matrix of observations on m othe[r] endogenous variables; X_1 is an $n \times \ell$ matrix of observations on ℓ included exogenous variables, W is an $n \times q$ matrix of observations on q additional variables. ε is a vector of disturbances, γ, β and η are vectors of coefficient parameter. The analyst commits a specification error by assuming that $\eta = 0$. The general k-class estimator of $\alpha = (\gamma_1\ \beta)'$ is given by

$$\hat{\alpha}(k) = \begin{pmatrix} Y'Y - kv'v & Y'X_1 \\ X_1'Y & X_1'X_1 \end{pmatrix}^{-1} \begin{pmatrix} Y' - kv' \\ X_1' \end{pmatrix} y \qquad (1.2.66)$$

[w]here v is an $n \times m$ matrix of values of residuals from the [r]educed form regression of Y on X where $X = [X_1 \,\vdots\, X_2]$ is [a]n $n \times k$ matrix of observations on all exogenous variables in [t]he complete system, X_2 being the matrix of observations on [t]hose exogenous variables which are excluded from the equation [u]nder study. k is a scalar with $\operatorname{plim}_{n\to\infty} k = 1$ if $\eta = 0$. [Wh]en $k = 1$, $\hat{\alpha}(k)$ becomes the two-stage least squares estimator.

For limited information maximum likelihood estimator, k [i]s given by the smallest root of the determinental equation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\mathrm{Det}(Q_1 - \lambda Q) = 0 \qquad (1.2.67)$$

Q_1 and Q_2 are the variance-covariance matrices of the residuals from the regressions of the (m+1) endogenous variables corresponding to y and Y on the variables corresponding to X_1 and X respectively.

Now,

$$d(k) = \operatorname*{plim}_{n\to\infty} (\hat{\alpha}(k) - \alpha)$$

$$= \operatorname*{plim}_{n\to\infty} \begin{pmatrix} Y'Y-kv'v & Y'X_1 \\ X_1'Y & X_1'X_1 \end{pmatrix}^{-1} \begin{pmatrix} (Y'-kv')(W\eta+u) \\ X_1'(W\eta + u) \end{pmatrix}$$

$$\neq 0 \text{ (in general)} \qquad (1.2.68)$$

which shows the inconsistency.

Fisher (1966) considered in particular the problem whether the two stage least squares is affected more (or less) by specification errors than the limited information maximum likelihood estimator is. It was found that although two stage least squares and limited information maximum likelihood estimators (and other members of K class) have different sensitivities to specification error, neither is uniformly more robust than the other. Which estimator is less sensitive to the type of specification error depends on the unobservable nature of error committed and on the unobservable disturbances.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Summers' (1965) conducted Monte Carlo experiments to examine the small sample properties of two stage least squares (2SLS), limited information (least variance ratio) maximum likelihood (LISE) estimator, full information maximum likelihood estimator (FIML) estimator and ordinarly least squares (OLS) estimator when one variable was incorrectly excluded from one of the two equations in the model. When exogenous variables were independent, 2SLS was ranked first, LISE second, OLS third and FIML last. For correlated exogenous variables, 2SLS was again ranked first OLS second, FIML third and LISE last.

Cragg (1968) conducted an extensive series of Monte-Carlo experiments to study the effects of incorrect specification on small sample properties of several simultaneous equation estimators. For most purposes, the effect of misspecification was probably enough to render the estimates of the structural coefficients useless. Failing to specify as zero all coefficients for which this was correct had serious effects on the central tendencies and dispersions of the estimators when knowledge of these coefficients was important for the identification of an equation. FIML seemed the estimator most sensitive to this danger. These findings suggest that in the absence of a fairly complete and conformably held knowledge about which coefficients in a structure are very small or zero, successful econometric model building becomes a

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

very difficult task.

For more details on the effect of misspecification on small sample properties of different simultaneous equation estimators, see also Waud (1966), Glahe and Hunt (1970) and Byron (1972).

1.3 Autocorrelated disturbances. Next we give a brief review of considerable literature on the problems created by autocorrelated disturbances in single equation regression analysis. These problems may less exactly be called the time series complication.

In a pioneering paper Yule (1926) drew attention to the difficulties of interpreting the correlation between the two autocorrelated series. But in the subsequent literature among those authors who were concerned with the measurement of functional relationship between autocorrelated series, few made it clear that the significant factor in the analysis is the autocorrelation of the error term and not the autocorrelations of the time series themselves. This point was clearly brought out by Aitken (1934/1935) and later by Champernowne (1948) who carried the problem into the field of statistical estimation and sampling theory.

Cochrane and Orcutt (1949) pointed out that the error terms in many, if not most, economic relationships are highly positively autocorrelated. They claimed that such autocorrelations arise from three main sources :

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

1. Omission of important economic and non-economic regressors from the relationship and the autocorrelated nature of these regressors.

2. Faulty choice of the algebraic form of the relationship.

3. Autocorrelated errors of observations.

<u>1.3.1 Tests of randomness of disturbances.</u> Anderson (1941), Bartlett (1935), Dixon (1944), Durbin and Watson (1950, 1951), Geisser (1956), Neumann (1941,1942), Hart (1942(a), 1942(b)), Rubin (1945), Shewon and Johnson (1965) and White (1957,1961) among others, investigated the problem of testing for randomness of the disburbances in a regression equation. Most of these studies are relevant to using the null hypothesis of zero autocorrelation among disturbances. 5/

Two of the most important contributions are due to Von-Neumann (1941, 1942) and of Durbin and Watson (1950, 1951). The Von-Neumann test of randomness is as follows :

5/ For the cases where equation contain lagged values of explanatory variables (X's) and/or the lagged values of the dependent variable (y), considerable work has been done on the problems of testing of randomness of disturbances and estimation of regression coefficients. For details, references are made, among others, to Koyck (1954), Marriot and Pope (1954), Griliches (1961,1967), White (1961), Leeuw (1962), Tinsely (1962), Liviatan (1963), Morton (1964), Taylor and Wilson(1964), Almon (1965,1968), Copas (1966), Wallis (1967), Zellner and Geisel (1968), Lund and Holden (1968), Orcutt and Winokur (1969).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let $u_1, u_2, \ldots, u_n$ be a series of observations. Von-Neumann ratio for testing H_o : u's are i.i.d. $N(\mu, \sigma^2)$, is

$$Q = \frac{\frac{1}{n-1} \sum_{t=2}^{n} (u_t - u_{t-1})^2}{\frac{1}{n} \sum_{t=1}^{n} (u_t - \bar{u})^2} \tag{1.3.1}$$

where $\bar{u} = \frac{1}{n} \sum_{t=1}^{n} u_t$.

When Q is sufficiently small, it indicates positive autocorrelation, and when Q is large enough, it points towards negative autocorrelation. Under H_o, the significance points of Q have been computed by Hart(1942(a),1942(b)) for $n \leq 60$. For $n > 60$, the distribution of Q may be approximated by a normal distribution with mean $\frac{2n}{n-1}$ and variance $\frac{4}{n}$.

For a time series regression,

$$y_t = \beta_o + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + \varepsilon_t, \quad t = 1, 2, \ldots, n \tag{1.3.2}$$

where ε_t's are normally distributed with mean 0 and a common variance σ^2, Durbin and Watson (1950, 1951) proposed a small sample test of

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

H_0 : ε_t's are independently distributed

against H_1 : ε_t's are positively autocorrelated (1.3.3)

The test statistic is

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \quad (1.3.4)$$ 6/

$$\simeq 2(1 - r_1) \quad (1.3.5)$$

where $r_1 = \frac{\sum_{t=2}^{n} e_t \, e_{t-1}}{\sum_{t=1}^{n} e_1^2}$

and $e_t = y_t - (\hat{\beta}_o + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} + \ldots + \hat{\beta}_k x_{kt})$, $t = 1, 2, \ldots, k$ (1.3.6)

where $\hat{\beta}_o$, $\hat{\beta}_1 \ldots \hat{\beta}_k$ are the OLS estimates of $\beta_o, \beta_1 \ldots \beta_k$ in (1.3.1).

6/ In deriving the D-W test statistic, the underlying assumption is that ε's are distributed independently of X. So, D-W statistic is strictly inapplicable for the cases where the regression equation contains lagged y_t's as regressors. In these cases, D-W statistic will be biased towards 2. For these cases, Durbin (1970) suggested a large sample test of serial independence of ε_t's.

Another assumption is that the regression should contain a constant term (i.e. $x_{1t} = 1 \; \forall \; t$).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Durbin-Watson test statistic is obviously closely related to Von-Neumann statistic defined in (1.3.1). Under H_o, $d \simeq 2$ and under H_1, $d < 2$.

Since, even under H_o, e_t's are autocorrelated among themselves, the determination of the sampling distribution of d under H_o has proved difficult. The sampling distribution of d under H_o depends on n,k and the data matrix

$$X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix}$$

Durbin and Watson were able to formulate only the bounds (d_ℓ, d_u) for each significance limit of d such that the limit lies in this interval whatever X may be.

If $d < d_\ell$, H_o is rejected in favour of H_1

$d > d_u$, H_o is accepted

$d_\ell < d < d_u$, the test is in conclusive.

A test against negative autocorrelation is obtained by replacing d by $(4 - d) \simeq 2(1 + r_1)$ which leads to significant values when r is negative and sufficiently large.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

For $n > 40$, Durbin and Watson (1951) described an approximation method for obtaining conclusive results when d falls in the inconclusive region. This method suggests using $\frac{1}{4}d$ as a Beta variable with parameter p and q determined by

$$p + q = \frac{E(d)\ (4 - E(d))}{var(d) - 1} \quad \text{and} \quad p = \frac{1}{4}(p + q)\ E(d) \tag{1.3.7}$$

Theil and Nagar (1961) also attempted a solution to remove the difficulty associated with the inconclusive region. The underlying assumption in his procedure is that the first and the second order differences of the regressors are small in absolute value compared to the range of the corresponding regressors themselves. On the basis of first four moments of d under H_o, given by (1.3.3), a Beta distribution was fitted to the sampling distribution of d in the range $(a, 4 - b)$ in the following manner.

Let $$w = \frac{d - a}{4 - (a + b)} \tag{1.3.8}$$

So, w is a Beta variable in the range $(0, 1)$ with density function

$$f(w) = \frac{1}{B(m, n)}\ w^p\ (1 - w)^q \tag{1.3.9}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Equating $g_1(d) = \frac{\mu_3(d)}{\mu_2^{3/2}(d)}$ to $g_1(w) = \frac{\mu_3(w)}{\mu_2^{3/2}(w)}$ (1.3.10)

and $g_2(d) = \frac{\mu_4(d)}{\mu_2^2(d)} - 3$ to $g_2(w) = \frac{\mu_4(d)}{\mu_2^2(w)} - 3$, (1.3.11)

(where $\mu_2(d)$, $\mu_3(d)$ and $\mu_4(d)$ are the second, third and fourth moments of d under H_o and $\mu_2(w)$, $\mu_3(w)$ and $\mu_4(w)$ are the corresponding moments of w) p and q are calculated. Next, equating

$$E(w) \text{ to } \frac{E(d) - a}{4 - (a+b)} \quad (1.3.12)$$

and

$$\mu_2(w) \text{ to } \frac{\mu_2(d)}{\{4 - (a+b)\}^2}, \quad (1.3.13)$$

a and b are calculated.

From (1.3.8),

$$d = w\{4 - (a+b)\} + a \quad (1.3.14)$$

Using this beta approximation, tables for 1 per cent and 5 per cent significance poinst of d for testing the null hypothesis H_o were constructed. The critical values were, on the average, found to be very close to d_u.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Incidentally, Hannan (1957) [see also Gregor (1960)] pointed out that to an order of accuracy higher than n^{-1}, true significance points of d are close to d_u in the case where the regression equation is a polynomial in time (t).

Malinvand (1966) pointed out that when each regressor is an approximate linear combination of a constant and a sinusoidal series with long periods, the critical values of d are close to d_u.

Theil and Nagar (1961) approach was pushed further by Henshaw (1966) who also fitted a Beta distribution to the sampling distribution of d under H_o on the basis of its first four moments. But this approach is more general because it is applicable also to those cases where the first and the second order differences of the regressors are large in absolute value compared to the range of the regressors themselves. This procedure is, however, computationally more laborious than Theil Nagar procedure.

Durbin (1970) suggested another procedure to be applied when d lies in the inconclusive region. He proposed a modified Durban-Watson statistic.

Let,

$$L = (\lambda_1, \lambda_2 \ldots \lambda_{k-1}). \qquad (1.3.15)$$

λ_i's are the characteristic vectors corresponding to(k-1)smallest

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

characteristic roots of

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix} \qquad (1.3.16)$$

zero characteristic roots have been ignored.

Let X denote the $n \times (k-1)$ matrix of $(k-1)$ regressors expressed as deviations from the sample means. Then one computes the least squares regression

$$Y = ai + Xb_1 + Lb_2 \qquad (1.3.17)$$

where $i = (1,1 \dots 1)'$ is an $n \times 1$ vector.

Let $P_1P_1' = (X_L' X_L)^{-1}$ and $P_2P_2' = (L_X' L_X)^{-1}$ (1.3.18)

where

$$X_L = X - L(L'L)^{-1} L'X$$

and

$$L_X = L - X(X'X)^{-1} X'L$$

Finally one computes

$$C = P_1P_2^{-1} b_2 \qquad (1.3.19)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Define

$$z = y - ai - Xb_1 - Lb_2 + X_L c \tag{1.3.20}$$

and the modified test statistic is

$$d' = \frac{\sum_{t=2}^{n} (z_t - z_{t-1})^2}{\sum_{t=1}^{n} z_t^2} \tag{1.3.21}$$

The critical value of d' are same as d_u.

Another statistic used for testing the randomness of disturbances (ε_t's) in the model

$$y_t = \beta_1 + \sum_{j=2}^{k} \beta_j x_{jt} + \varepsilon_t, \quad t = 1, 2, \ldots, n \tag{1.3.22}$$

(where x_{jt}'s are nonstochastic).

is the circular serial correlation coefficient.

$$r_1' = \frac{\sum_{t=1}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} \tag{1.3.23}$$

where e_t is the OLS residual. This is closely related to D-W statistic. Anderson (1942) obtained the sampling distribution of this statistic when only a mean correction has been made and tabulated the significance points.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Another important problem which often arises in practice is that the null hypothesis, is not the serial independence of ε_t, but rather

$$\varepsilon_t = \rho \, \varepsilon_{t-1} + u_t, \qquad (1.3.24)$$

where $|\rho| < 1$, u_t i.i.d. $N(0, \sigma^2)$.

The alternative hypothesis is

$$\varepsilon_t + \delta_1 \, \varepsilon_{t-1} + \delta_2 \, \varepsilon_{t-2} = u_t \qquad (1.3.25)$$

So, here H_0: $\delta_2 = 0$ (1.3.26)

This is tested approximately by the partial autocorrelation

$$r_{02.1} = \frac{r_2 - r_1^2}{1 - r_1^2} \qquad (1.3.27)$$

where $r_j = C_j / C_o$,

with $C_j = \dfrac{\sum_{t=1}^{n-j} e_t \, e_{t+j}}{n} = C_{-j}$

Daniels (1956), Jenkins (1954, 1956) and Watson (1956) made a detailed study of this type of statistic when only a mean correction has been made. Their exact derivations are however based on

$$C_j' = \frac{1}{n} \sum_{i=1}^{n} e_t \, e_{t+j} \quad \text{in place of} \quad C_j.$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

With the help of Fourier methods, Hannan and Terrel (1968) showed that the effect of the regression on the significance points of d depends substantially on the cross specta of the regressor vectors x_{jt} in (1.3.21). If the spectrum of x_{jt} is relatively very concentrated at the origin, the effect will be approximately allowed for by using the significance point for d_u as the true significance point. In the special cases such as a polynomial regression, this procedure will be accurate to order $\frac{1}{n}$. Hannan and Terrel also obtained $E(r_{02.1})$ and $E(r_{02.1}^2)$ upto order $\frac{1}{n}$. On the basis of these, they have suggested that when x_{jt} are series which have spectra concentrated at the origin $\{r_{02.1} + (\frac{k+1}{n})\}$ can be used as ordinary correlation from (n+2) pairs of observations.

One of the difficulties with least squares residuals is that even when elements of ε are spherical, the elements of e are not. Theil (1965) developed an estimate of ε which is best linear unbiased (BLU) and which has a scalar (S) covariance matrix. These estimators are called BLUS residuals. These obviously are more amenable to direct tests of autocorrelation and also of other hypotheses (e.g. tests of normality). The BLUS residuals are defined by

$$\tilde{e} = By \qquad (1.3.28)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $BX = 0$ (null matrix) (1.3.29)

and $BB' = I$ (1.3.30)

However (1.3.30) cannot be satisfied since by (1.3.29) the rank of BB' cannot exceed n-k. Thus, at best we can obtain a matrix on the right hand side of (1.3.29) with n-k unities on the principal diagonal and the remaining k elements zero. So, BLUS residual $\tilde{e}$ can estimate only n-k disturbances. Assuming that the first k disturbances are not estimated.

Let

$$B = \begin{pmatrix} 0 \\ \cdots\cdots \\ C \end{pmatrix} \qquad (1.3.31)$$

where 0 is of order $k \times n$ and C of order $(n-k) \times n$.

$$CX = 0 \text{ (null matrix)}$$

and $$CC' = I_{n-k} \qquad (1.3.32)$$

Let us partition C by the first k and the remaining n-k columns.

$$C = (C_1 \vdots C_2) \qquad (1.3.33)$$

and partition X by the first k and the remaining (n-k) rows.

$$X = \begin{pmatrix} X_1 \\ \cdots\cdots \\ X_2 \end{pmatrix} \qquad (1.3.34)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let $M_{11} = I - X_2(X'X)^{-1} X_2'$ (1.3.35)

Let $D^{1/2}$ be the diagonal matrix whose diagonal elements are the positive square roots of the latent roots of M_{11} and the columns of the matrix T are latent vectors of M_{11}. It can be shown that

$$C_2 = T D^{1/2} T$$

(1.3.36)

and $$C_1 = C_1 X_2(X_1)^{-1}$$

So, the BLUS estimator of the last $n-k$ disturbances is

$$\tilde{e}_2 = Cy \quad (1.3.37)$$

Let $$\tilde{e} = \begin{pmatrix} \tilde{e}_1 \\ \cdots \\ \tilde{e}_2 \end{pmatrix} .$$

Theil also proposed an alternative method of computing $\tilde{e}_1$ which requires computation of only k latent vectors. The basic result is

$$\tilde{e}_1 = e_1 - X_2(X_1)^{-1} \left(\sum_{i=1}^{k} \frac{d_i}{1 + d_i} f_i f_i' \right) e \quad (1.3.38)$$

where $e = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ (e being the OLS residuals),

d_i is the postive square root of the latent roots and f_i are the latent vectors of the matrix $X_1(X'X)^{-1} X_1'$. In general, all the k roots will be less than unity, but, if not,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the summation over i in (1.3.38) should be restricted to those which are less than unity.

Finally, the problem is, which of the k disturbances is to be omitted. Theil (1968) suggested that for testing against positive autocorrelation, the first m and the last $(k-m)$ observations (where m is a nonnegative integer $\leq k$) are to be delected. This gives a choice of $(k+1)$ different sets of k observations that might be deleted. According to Theil, the choice among these $(k+1)$ sets should be made as follows. For each set the matrix $X_1(X'X)^{-1}X_1'$ and the roots $d_1^2, d_2^2 \ldots d_k^2$ are to be computed. If one or more of the roots are zero, the set should be omitted because its X_1 matrix is singular. For remaining sets, the sum $d_1 + d_2 + \ldots + d_k$ is calculated and that set is selected for which the sum takes the largest value. This gives the set which minimises the expected sum of the estimation errors.

The usefulness of scalar covariance matrix of BLUS residuals is that Von-Neumann(1942(a),1942(b)) ratio Q can be directly applied to these residuals. Replacing u_t's in(1.3.1) by $\tilde{e}_t$ and n by $n-k$, the modified Von-Neumann ratio of $n-k$ successive BLUS residuals is

$$Q' = \frac{\sum_{t=2}^{n-k} (\tilde{e}_{t-1} - \tilde{e}_t)^2}{(n-k-1)s^2} \qquad (1.3.39)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $s^2 = \frac{\sum_{t=1}^{n} e_t^2}{n - k}$, and ϵ_t is the OLS residual defined in (1.3.6). Significance limits of Q' for $n-k = 2, \ldots, 60$ have been derived by Press and Brooks (1969) under the condition that $\tilde{e}_t$'s are i.i.d. $N(0, \sigma^2_{\tilde{e}})$. For $n - k > 60$, Q' may be approximated by a normal distribution with mean 2 and variance $\frac{4}{n-k}$.

Blattberg (1973) examined the power of Durbin-Watson (1950, 1951) test for situations where the disturbance follows processes other than Markov schemes. In large samples, power seems to be greater for a second order autoregressive process if ρ_2 (second order autocorrelation coefficient) > 0 than for a Markov scheme, but if $\rho_2 < 0$, the power seems to be less for the second order process. The power is about equally higher for a first order moving average process and for a Markov scheme when $0 < \rho_1 < .5$. However, in large samples, the power is greater for a Markov scheme than for a first order moving average process where $.5 < \rho_1 < 1.0$; ρ_1 being the first order autocorrelation coefficient.

Krishnaiah and Murthy (1966) set up simultaneous tests for trend and serial correlations for Gaussian Markov residuals. Let

$$x_t = \mu_t + \varepsilon_t, \quad t = 1, 2, \ldots, n \tag{1.3.40}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\mu_t = \alpha + \beta t, \quad t = 1, 2, \ldots, n$ (1.3.41)

$E(\varepsilon_t) = 0, \quad V(\varepsilon_t) = \sigma^2$ and $cov(\varepsilon_t, \varepsilon_{t+k}) = \sigma^2 \rho^k, \ \rho < 1$ (1.3.42)

ε_t's are normally distributed and form a stationary Gauss Markov process of order one.

Krishnaiah and Murthy developed a method of testing simultaneously the hypotheses

$H_{01} : \alpha = 0,$

$H_{02} : \beta = 0$

and H_{03} : Stationary Gaussian process $\{\varepsilon_t\}$ is a process of independent variates. A trivariate central F statistic with (1, n-3) degrees of freedom was used for the purpose of testing.

1.3.2 Estimation of regression coefficients.

So far we have discussed different tests of randomness. When the hypothesis of randomness of disturbances is rejected estimation of the regression coefficients poses a very serious problem. OLS estimates of regression coefficients are no longer BLUE. However, they are still unbiased. What is more serious, Ordinary least-squares formulae for sampling variances of the estimated regression coefficients give biased estimates. The

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

usual formulae for setting up interval estimates or for testing hypotheses regarding the regression coefficients are no longer valid. We shall also get inefficient predictions, that is, predictions with needlessly large sampling variances. Generalised least squares procedures are called for to overcome this hurdle.

Cochrane and Orcutt (1949) proposed estimating the regression coefficients assuming that the disturbances follow a Markov scheme and that the first order autocorrelation coefficient (ρ) of the disturbances is known. Starting from this idea, a two-step Cochrane-Orcutt method and a Cochrane-Orcutt iterative procedure for estimating the regression co-efficients (vide Johnson 1972) have been suggested when ρ is unknown.

The regression model is now

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + \varepsilon_t$$

where $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$, $|\rho| < 1$, and u_t is a random process with mean 0 and variance σ_u^2

(1.3.43)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In Cochrane-Orcutt iterative procedure, the estimates of $\beta_0, \beta_1, \ldots, \beta_k$ and ρ are obtained by minimising

$$\sum_{t=2}^{n} u_t^2 = \sum_{t=2}^{n} [(Y_t - \rho Y_{t-1}) - \beta_0(1-\rho) - \beta_1(X_{2t} - \rho X_{2,t-1})$$

$$- \ldots - \beta_k(X_{kt} - \rho X_{k,t-1})]^2 \qquad (1.3.44)$$

with respect to $\rho, \beta_0, \beta_1, \ldots, \beta_k$. First, for a fixed value of ρ, (1.3.44) is minimised with respect to $\beta_0, \beta_1, \ldots, \beta_k$. Then, using these estimates, (1.3.44) is again minimised with respect to ρ to get an estimate of ρ. This is continued until successive estimate differ only by a small amount.

Sargan (1964) showed that for this type of problem this iterative process will always converge to a stationary value of the sum of squares in (1.3.44). However, there is the possibility of the existence of several local minima, in which

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

case the process would converge to one of them depending on the starting point. In a large number of studies, conducted by Sargan (1964), no case of the occurence of multiple minima could be traced. 7/

In two-step Cochrane-Orcutt procedure (vide Johnston 1972) the first step is to estimate ρ as

---

7/ In order to control iteration, an alternative approach (vide, Johnston, 1972) is, first, to test the hypothesis of zero autocorrelation of ε by applying Durbin-Watson test to OLS residuals. If the hypothesis is rejected, the next step is to minimise

$$\sum_{t=2}^{n}\{(y_t-\beta_0-\beta_1 x_{1t}- \dots - \beta_k x_{kt}) - \rho(y_{t-1} - \beta_0 - \beta_{1,t-1} - \dots - \beta_k x_{k,t-1})\}^2$$

with respect to ρ for given values of $\beta_0, \beta_1, \dots, \beta_k$. Let r_1 be the estimate of ρ. Next, $\beta_0, \beta_1, \dots, \beta_k$ are estimated by OLS procedure from the equation (1.3.46 of page 91) using r_1 for ρ. With these estimates of ρ and $\beta_0, \beta_1, \dots, \beta_k$, D-W test is again applied to OLS residuals obtained from (1.3.46). Iteration is stopped when the hypothesis of zero autocorrelation is accepted.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t \, e_{t-1}}{\sum_{t=1}^{n} e_t^2} \qquad (1.3.45)$$

where e_t is the OLS residual defined in (1.3.6). Then $\beta_0, \beta_1, \ldots, \beta_k$ are estimated by OLS procedure from the equation

$$y_t - \hat{\rho} y_{t-1} = \beta_0(1 - \hat{\rho}) + \beta_1(x_{2t} - \hat{\rho} x_{2,t-1}) + \ldots + \beta_k(x_{kt} - \hat{\rho}\, x_{k,t-1})$$
$$+ \varepsilon_t - \hat{\rho}\varepsilon_{t-1} \qquad (1.3.46)$$

under fairly general conditions, it can be proved that $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ give consistent estimates of $\beta_0, \beta_1, \ldots, \beta_k$.

Durbin (1960) suggested a two step method which gives estimates with asymptotically the same mean vector and dispersion matrix as the least squares estimates obtained by direct minimisation $\sum_{t=2}^{n} u_t^2$ in (1.3.44). In this method the first step is to estimate ρ by regressing y_t on $y_{t-1}, x_{2t}, x_{2,t-1}, \ldots, x_{k,t}, x_{k,t-1}$. The OLS estimate of the coefficient of y_{t-1} is taken as an estimate of ρ. The second step is to use this estimate of ρ for $\hat{\rho}$ in equation (1.3.46) and estimate $\beta_0, \beta_1, \ldots, \beta_k$ by OLS method. Here also, it can be proved, under fairly general assumptions, Durbin's procedure yield consistent estimates.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

of $\beta_0, \beta_1, \ldots, \beta_k$ and ρ. Durbin's method extends quite simply to higher order autoregressive schemes.

In Prais and Winsten (vide Rao(1968)) method of estimation ρ is estimated by $\hat{\rho}$ in (1.3.45). Next, using this $\hat{\rho}$, $\beta_0, \beta_1, \ldots, \beta_k$ are estimated by generalised least squares method. Under fairly general conditions, this method also yields consistent estimators of $\beta_0, \beta_1, \ldots, \beta_k$ and ρ (vide Theil 1971, pp 398-408).

Another procedure is to apply OLS methods of estimations to the equation

$$y_t = \beta_0(1 - \rho) + \rho y_{t-1} + \beta_1 x_{2t} - \beta_1 \rho x_{2,t-1} + \ldots + \beta_k x_{kt} - \beta_k \rho x_{k,t-1} + \varepsilon_t - \rho \varepsilon_{t-1} \tag{1.3.47}$$

Using the restriction

$$\begin{aligned} \widehat{\beta_1 \rho} &= \hat{\beta}_1 \hat{\rho} \\ &\vdots \\ \widehat{\beta_k \rho} &= \hat{\beta}_k \hat{\rho} \end{aligned} \tag{1.3.48}$$

This method is iterative and computationally more expensive than the other four methods.

All these iterative or two-step estimates are computationally much more efficient than ordinary least squares. To examine the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

gain in efficiency and also the variation in the efficiency of various two step estimators in small samples, Griliches and Rao (1964) conducted a Monte-Carlo experiment with samples of size 20. They postulated the model,

$$y_t = \beta x_t + \varepsilon_t ,$$
$$x_t = \lambda x_{t-1} + v_t , \qquad (1.3.49)$$
$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t .$$

OLS was found to be less efficient than all the other estimators. This is specially true for $|\rho| > 0.3$. For low values of ρ, however there may be a little loss of efficiency in using the more complicated methods. Durbin's two-step method of estimating ρ appears to be better than the others and a GLS method using Durbin $\hat{\rho}$ seems to be the best over a wide range of parameters. Finally, it appears that the nonlinear method shows no improvement over the simpler two step procedures.

Wold (1950) gave an appropriate large samples formula for correcting the OLS expressions for standard errors of estimated regression coefficients when the disturbances are autocorrelated, assuming that all the explanatory variables are exogenous [vide also Wold and Jureen,1953,pp.209-215].Lyttkens(1964) generalised this formula to the case where correlation occurs between the regressors and lagged values of the residuals, for instance,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

when lagged value of y_{t-k} occurs in the regression equation of y_t.

1.3.4 Other references. For simultaneous equation models, considerable work has been done on the problem of autocorrelated disturbances, among others, by Sargan (1961), Amemiya (1966), Fair (1970, 1972), Hendry (1971), Dhrymes (1970), Guilkey and Schmidt (1972), Guilkey (1974) etc.

Finally, there are Bayesian methods (vide, Zellner 1971, Ch. 4) of estimation and testing of regression coefficients when the disturbances follow a first order autoregressive process.

1.4 Errors-in-variable models

Introduction

Lastly, we make a survey of the methods of estimation and hypotheses testing for models with errors in the variable (EVM).

Presence of errors in variables poses a serious problem to the econometricians. The classical linear regression model is widely used to draw inference and to make predictions with the help of data (time series and/or cross section). But such inferences and predictions are not valid when the regressors entering the regression equation are not free from errors of observation.

The problem of unobservable errors of measurement (or errors in variables) was first formulated by Frisch (1934) in

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the form particularly appropriate for econometric applications. Koopmans (1937) applied the weighted regression technique to handle the problem of errors-in variables. Durbin (1954), Madansky (1959) and Cochran (1968) wrote important review articles on the problem of errors in variables. More recent surveys are due to Griliches (1974) and Lankipalle (1975).

1.4.1 Effects of errors of observation on OLS estimation of regression coefficients.

Let us consider the simplest problem of errors-in-variables. Let y and x be true variables. The corresponding observed variables are y^* and x^*. So,

$$y_i^* = y_i + v_i$$

and $$x_i^* = x_i + u_i \qquad i = 1,2, \ldots, n \qquad (1.4.1)$$

Here v_i and u_i are the errors of measurement. Usually u_i and v_i are assumed to be random variables with zero mean and unknown constant variances σ_u^2 and σ_v^2. These errors are independent of the true variables and also mutually independent. The relationship between the true variables is

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1,2,\ldots, n \qquad (1.4.2)$$

u_i and v_i are assumed to be independent of ε_i for each i.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

ε_i's are mutually independent with mean 0 and variance $\sigma^2 \; \forall \; i$ and also independent of x_i's. The equation (1.4.2) can alternatively be written as

$$y_i^* = \alpha + \beta \, x_i^* + \xi_i \qquad (1.4.3)$$

where $\xi_i = \varepsilon_i + v_i - \beta \, u_i$.

Let us further assume that

$E(\varepsilon\varepsilon' / X^*) = \sigma^2 I$

where $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ and $X^* = \begin{pmatrix} 1 & x_1^* \\ 1 & x_2^* \\ \vdots & \vdots \\ 1 & x_n^* \end{pmatrix}$

$$\operatorname*{plim}_{n \to \infty} \frac{\varepsilon'\varepsilon}{n} = \sigma^2 \qquad (1.4.4)$$

$$\operatorname*{plim}_{n \to \infty} \frac{X^{*\prime} X^*}{n} = \Sigma_{X^*X^*} \quad \text{(a positive definite matrix)}$$

and $\operatorname*{plim}_{n \to \infty} \frac{X^{*\prime} \varepsilon}{n} = \underset{\sim}{0}$

It can be proved easily (vide Johnston (1972) Ch. 9, Theil (1971) Ch.-12, Sec.2) that the OLS regression of y^* on x^* provides inconsistent estimates of α and β in (1.4.3) because x_i and ξ_i are correlated for given i. This result can be easily generalised to the case of more than one explanatory variable.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(For the effect of correlated measurement errors on OLS estimat of regression coefficients see Chai (1971)).

Griliches and Ringstad (1970) examined the effect of using ordinary least squares method for estimating the regression coefficients of a nonlinear relationship when the regressors contain errors of observation.

Let the true model be

$$y_i = \alpha + \beta x_i + \gamma x_i^2 + \varepsilon_i, \quad i = 1,2,\ldots,n \qquad (1.4.5)$$

and

$$x_i^* = x_i + u_i \qquad i = 1,2,\ldots,n \qquad (1.4.6)$$

It was found that the OLS estimates of β and γ were biased downwards.

Richardson and Wu (1970) considered the functional relationship model

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1,2,\ldots,n \qquad (1.4.7)$$

$$x_i^* = x_i + u_i, \qquad i = 1,2,\ldots,n; \; n \geq 3 \qquad (1.4.8)$$

where u_i, ε_i and x_i are independent of each other. u_i and ε_i are assumed to be spherical with mean zero and variances σ^2 and σ_u^2 $\forall$ i. It is also assumed that ε_i and u_i are normally distributed. Under these assumptions, the authors calculated the exact bias and mean square error of the OLS estimate of β.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In the above set up, under the assumptions that for each i (ε_i, u_i) follow independent bivariate normal distribution with means 0 and the variance covariance matrix

$$\begin{pmatrix} \sigma^2 & \rho\sigma\sigma_u \\ \rho\sigma\sigma_u & \sigma_u^2 \end{pmatrix}$$

Halperin and Gurian (1971) obtained the exact bias and the mean square error of the OLS estimate of β.

Aigner (1973) examined the problem of getting consistent OLS estimates of regression coefficients when the regression equation contains a binary independent variable measured with error.

Maurice D. Levi (1973) studied the errors in the variables bias in the presence of correctly measured variables. He showed that if there is any one regressors measured with error and the others are measured correctly, under large sample assumptions, the the bias of the OLS estimate of the coefficient associated with the variables measured with error is downwards. The direction of bias of the estimated coefficients of the variables measured correctly can be determined if the variance-covariance matrix of observation is known.

Grether and Maddala (1973) considered the following model

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y_t = \beta\, x_t + \varepsilon_t$$
$$x^*_t = x_t + u_t \qquad t = 1,2,\ldots,n \qquad (1.4.9)$$

with the usual assumptions regarding the variables. The OLS residual is

$$e_t = \beta\, x^*_t + \varepsilon_t - \beta\, x_t \qquad (1.4.10)$$

where β is the OLS estimate of β. It has been observed that unless x and u are serially independent, e_t's will be autocorrelated even when ε_t's are serially independent. Similar results are reported for models with lagged endogenous variables and serially correlated errors. Considering the following errors-in-variables model.

$$\begin{aligned} y_t &= \alpha\, y_{t-1} + \beta\, x_t + \varepsilon_t \\ &= \alpha\, y_{t-1} + \beta\, x^*_t + (\varepsilon_t - \beta\, u_t) \end{aligned} \qquad (1.4.11)$$

Grether and Maddala proved that α and β are both inconsistent and the bias can be divided into two parts :— one due to measurement error and the other due to serial correlations of the series x_t and u_t.

Berkson (1950) (see also Lindley 1953) discussed an interesting situation in which it may be possible to set x^* at predetermined levels and corresponding to each observed x^* there

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

may be a number of x's that would lead to the same value of x^*. x^* is called a controlled variable. So, x is a random variable distributed about fixed x^* with an error u (i.e. $x = x^* + u$), which is independent of x.

In this case, if the model is

$$y_i = \alpha + \beta x_i, \quad i = 1,2, \ldots, n \tag{1.4.12}$$

One can write $y_i^* = \alpha + \beta x_i^* + (\beta u_i + v_i)$, $i = 1,2,\ldots,n$ (1.4.13)

where $y_i^* = y_i + v_i$

and $x_i^* = x_i + u_i$

Since both u and v are independent of x, OLS procedure will give consistent estimates of α and β.

Fedeorv (1974) generalised the Berkson case to k regressors. Under the assumption of existence of certain moments of u and v consistent estimators of the regression coefficients have been obtained.

For the extension of analysis of controlled variables to the non-linear case, see Geary (1953) and Scheffe (1950).

1.4.2 <u>Effect of errors on regression line</u>

Considering the relationship between the variables $y, x_1, x_2, \ldots, x_k$ to be linear, Lindley (1947) (See also

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Allen (1938), Fix (1949) and Laha (1958) established a condition under which the regression will continue to be linear when the variables are influenced with errors.

Let

$$E(y|x_1, x_2, \ldots, x_k) = \sum_{i=1}^{k} \beta_i x_i \qquad (1.4.14)$$

where the variables $y, x_1, x_2, \ldots, x_k$ are measured from their respective means.

Suppose,

$$x_i^* = x_i + u_i, \quad i = 1, 2, \ldots, k$$

$$\text{and} \quad y^* = y + v \qquad (1.4.15)$$

where $u_1, u_2, \ldots, u_k$ and v independently distributed with zero means ; furthermore, they are also independent of y and x_i's $(i = 1, 2, \ldots, k)$.

Lindley established the necessary and sufficient condition under which the result

$$E(y^*|x_1^*, x_2^*, \ldots, x_k^*) = \sum_{i=1}^{k} \beta_i^+ x_i^* \quad \text{8/} \qquad (1.4.16)$$

will hold good. The condition is

$$\sum_{i=1}^{k} (\beta_i - \beta_i^+) \frac{\chi(t_1, t_2, \ldots, t_k}{t_i} = \sum_{i=1}^{k} \beta_i^+ \frac{k_i(t_1, t_2, \ldots, t_k)}{t_i}$$

$$(1.4.17)$$

8/ β_i^+, $i=1,2,\ldots,k$ are the usual regression coefficients of of y^* on $x_1^*, x_2^*, \ldots, x_k^*$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $X(t_1, t_2, \ldots, t_k)$ is the cumulant generating function(cgf) of $X = (X_1, X_2, \ldots, X_k)$ and $k_i(t_1, t_2, \ldots, t_k)$ is the c.g.f. of u_i $(i = 1, 2, \ldots k)$.

The above condition is obviously satisfied of y and x_i $(i = 1, 2, \ldots, k)$ jointly follow multivariate normal distribution and further v and u_i $(i = 1, 2, \ldots k)$ have independent normal distributions so that y^* and x_i^* $(i = 1, 2, \ldots k)$ also jointly follow multivariate normal distribution. But the above condition may be satisfied for other types of distributions also.

Cochran (1970) dealt with the case where the standard linear regression model $y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1, 2, \ldots, n$ with ε and x independent and $E(\varepsilon) = 0$ is assumed to apply to a bivariate sample of pair (y, x). However, owing to the difficul ies is measuring th x values, a tually the bivariate sample used is of pair (y, x^*), where $x^* = x + u$, u being the error of measurement, independent of x, y and ε. Cochran examined the effect of departure from linearity of relationship between y and x^*. The examples he considered suggest that departure from linearity can be approximated by a quadratic in observed x^* if either u or x^* has a skewed distribution and by a cubic if either u or x has a symmetric distribution. The linear component dominates because the mean square deviation of y from linear component is only slightly larger than that from the exact regression of y on x^*.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

1.4.3 Classical methods of estimation

The classical solution to the problem of estimation of the regression coefficients in errors-in-variable models is based on normality assumption for u, v and ε and also sometimes for x's. Let us make such assumptions and examine the performance of maximum likelihood (ML) methods in estimating the regression coefficients in errors in variable models.

Let the true relationship be given by (1.4.2) and the errors-in-variable model be given by (1.4.3). Let $E(x_i^*)\ \mu\ \forall\ i$ and $V(x_i) = E(x_i - \mu)^2 = \sigma_x^2\ \forall\ i$. Then under the assumption of normality of x, u, v and ε, x^* and y^* variables have bivariate distribution with parameters

$$E(x_i^*) = \mu \quad \forall\ i$$

$$E(y_i^*) = \alpha + \beta\,\mu \quad \forall\ i$$

$$\sigma_{x^*}^2 = E(x_i^* - \mu)^2 = \sigma_x^2 + \sigma_u^2\ \forall\ i, \quad \text{where} \quad \sigma_u^2 = V(u_i)\ \forall\ i$$

$$\sigma_{y^*}^2 = E(y_i^* - \alpha - \beta\mu) = \beta^2\sigma_x^2 + \sigma^2 \quad \forall\ i \quad \text{where} \quad \sigma^2 = \sigma_v^2 + \sigma^2$$

$$\left.\begin{array}{l} \text{and } V(v_i) = \sigma_v^2, \\ V(\varepsilon_i) = \sigma^2 \end{array}\right\} \forall\ i \tag{1.4.18}$$

The corresponding sample statistics will be the maximum likelihood estimates of these parameters. Thus,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\begin{aligned}
\bar{x}^* &= \mu \\
\bar{y}^* &= \alpha + \beta\,\mu \\
m_{x^*x^*} &= \sigma_x^2 + \sigma_u^2 \\
m_{y^*y^*} &= \beta^2\,\sigma_x^2 + \sigma_\epsilon^2 \\
m_{x^*y^*} &= \beta\,\sigma_x^2
\end{aligned} \qquad (1.4.19)$$

where $m_{x^*x^*} = \frac{1}{n} \sum_{i=1}^{n} (x_i^* - \bar{x}^*)^2$

$m_{y^*y^*} = \frac{1}{n} \sum_{i=1}^{n} (y_i^* - \bar{y}^*)^2$

and $m_{x^*y^*} = \frac{1}{n} \sum_{i=1}^{n} (x_i^* - \bar{x}^*)y_i^*$.

On the left hand side of (1.4.19) we have five sample statistics and the right hand side contains six parameters. So, the model is not identified. But, if $\frac{\sigma_u^2}{\sigma_\epsilon^2}$ or σ_u^2 is known apriori, then the problem can be solved. In econometrics, it is possible to hazard an estimate of σ_u^2 . (For instance, for national income statistics, it is becoming a common practice to give some indication of the likely range of errors).

For the general case with k regressors, the true model is

$$Y = X\beta + \epsilon \qquad (1.4.20)$$

Y is an $n \times 1$ vector. X is an $n \times k$ matrix. β is a $k \times 1$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

vector and ε is an $n \times 1$ vector. Write

$$X^* = X + U$$

and (1.4.21)

$$Y^* = Y + U$$

where X^* is an $n \times k$ matrix of observations and U is an $n \times k$ matrix of unobservable errors, Y^* is an $n \times 1$ vector of observations and v is an $n \times 1$ vector of unobservable errors. The assumptions regarding X, U, v and ε are similar to those in (1.4.1), (1.4.2) and (1.4.3). Here also, for identification of the model, Σ_{uu}, the variance-covariance matrix of U should be known.

Next, one can allow for possible correlation between the errors u and v in (1.4.3). Thus, if in (1.4.3), u and v are correlated, writing $\varepsilon + v = \delta$ one can assume that u and δ jointly follow a bivariate normal distribution with means zero and variance covariance matrix

$$\begin{pmatrix} \sigma_\delta^2 & \rho_{uv}\sigma_u\sigma_\delta \\ \rho_{uv}\sigma_u\sigma_\delta & \sigma_u^2 \end{pmatrix} \qquad (1.4.22)$$

Moreover, if x also has a normal distribution which is independent of u and δ, it has been proved that unless any two of (σ_u^2, σ_δ^2 and ρ_{uv}) are known, the model remains unidentified, (vide Kendall and Stuart 1967, Ch. 29). See also Madansky (1951).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Instrumental variable method

While least squares estimators of α and β are inconsistent the ML estimators and methods of moments, require strong assumptions about the dispersion matrix of the observational errors. These difficulties do not arise with a class of estimators known as instrumental variable estimators.

Let the general errors in variable model be given by (1.4.20). The assumptions are also similar to those in (1.4.20). In the instrumental variable method, k variables $z_1, z_2, \ldots z_k$ are selected such that they are uncorrelated with ε, u and v and at the same time highly correlated with X. If any of the regressors in X be free from errors, it may be included among the z's. Let

$$Z = \begin{pmatrix} z_{11} & z_{21} & \cdots & z_{k1} \\ z_{12} & z_{22} & \cdots & z_{k2} \\ \vdots & \vdots & & \vdots \\ z_{1n} & z_{2n} & \cdots & z_{kn} \end{pmatrix}$$

So, the instrumental variable estimator of β is defined as

$$\hat{\beta}_{IV} = (z' X^*)^{-1} z'y \qquad (1.4.23)$$

Under the assumption that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} z' \xi\right) = 0$$

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} z' X^*\right) = \Sigma_{zX^*} \text{ exists and is nonsingular} \qquad (1.4.24)$$

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} z' z\right) = \Sigma_{zz} \text{ exists}$$

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}_{IV} = \beta \qquad (1.4.25)$$

$$\text{Asy var}(\hat{\beta}_{IV}) = \underset{n\to\infty}{\text{plim}} (\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)'$$

$$= \frac{\sigma_\xi^2}{n} \Sigma_{zX^*}^{-1} \Sigma_{zz} \Sigma_{X^*z}^{-1} \qquad (1.4.26)$$

where $\Sigma_{X^*z} = \Sigma'_{zX^*}$ and $\sigma_\xi^2 = E(\xi_i^2)\ \forall\ i$ and ξ_i is the i-th element in the vector $\varepsilon + v - U\beta$. Choice of instrumental variables is, however, a difficult job since there is no means of verifying that they are uncorrelated with the errors.

Now we shall discuss some instrumental variable estimators. Two well known grouping techniques now appear to be special cases of instrumental variable estimates.

The errors in variables model is given by (1.4.3).

(i) Wald's (1940) method of estimation.

Here,

$$Z = \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & +1 \\ \vdots & \vdots \\ 1 & -1 \end{pmatrix} \qquad (1.4.27)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where the elements in the second row are plus or minus according as the corresponding x^* is above or below the median. The underlying assumption is that the measurement errors u's are so small that the grouping of x^* values above and below the median will represent the same grouping for x values. (See also Neyman and Scott (1951)).

(ii) Bartlett's (1949) method of estimation.— Here the ranked x^* values are divided into three equal-sized groups, the first containing the smallest x^* values and the third containing the largest x^* values.

Here,

$$Z = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{1.4.28}$$

The elements in the second column of $\underset{\sim}{Z}$ are -1, 0 or 1 according as x^* belongs to the first group, or the central group or the third group respectively. Here also, the underlying assumption is that u's are so small that the grouping of x^* values represents the same grouping for x values.

Theil and Van Yzeren (1956) examined the optimal formation [9/] of the three groups in Bartlett approach for different types of

9/ By optimal grouping we mean that grouping for which the variances of α and β are minimum.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

distribution of x^*. See also Nair and Shrivastava (1942), Gibson and Jowett (1957a, 1957b), Hopper and Theil (1958).

Durbin (1954) suggested a more efficient method of estimation. Here

$$Z = \begin{pmatrix} 1 & r_1 \\ 1 & r_2 \\ \vdots & \vdots \\ 1 & r_n \end{pmatrix} \qquad (1.4.29)$$

where r_i is the rank of x_i.

The underlying assumption is that the ranking is unaffected by the errors of observation u's.

Stuart (1954) showed that when x follows a normal distribution, the efficiency of Durbin's estimator with respect to that of the OLS estimator is $E^2 = \frac{n-1}{n+1} \frac{3}{\pi}$; where n is the sample size. In large samples, $E^2 \simeq 0.96$. For samples of 20, the value drops to 0.86 and for samples of 5 to as low as 0.64.

If it is felt that the errors in x values are so large that rank ordering will be seriously affected by them, the x values may be arranged according to magnitude into k groups and the elements in the second row of Z may be i for all x's in the i-th group.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

When x is a random variable with an asymmetric probability distribution and u has a symmetric probability distribution, by taking the i-th element in the second row of Z as x_i^2, we get a consistent estimate of β.

1.4.5 Some recent developments in the method of estimation.

Sprent (1966) developed a method of estimation of coefficients of a linear E-V model by minimising the sum of squared residuals with weights inversely proportional to their variances. He considered the following exact functional relationship.

$$y_i = \beta x_i \qquad i = 1,2, \dots n \tag{1.4.30}$$

Let

$$\left.\begin{aligned} y_i^* &= y_i + v_i \\ x_i^* &= x_i + u_i \end{aligned}\right\} \qquad i = 1,2, \dots n \tag{1.4.31}$$

where u_i and v_i are independent of x_i and y_i for all i. u_i and v_i have zero means and the variance covariance matrix given by

$$\Sigma = \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix} \quad \forall \; i$$

Hence,

$$E(y_i^*) = y_i$$

and

$$E(x_i^*) = x_i \tag{1.4.32}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and the variance covariance matrix of (x^*, y^*) is

$$\Sigma = \begin{pmatrix} \sigma_{x^*x^*} = \sigma_u^2 & \sigma_{x^*y^*} = \sigma_{uv} \\ \sigma_{x^*y^*} = \sigma_{uv} & \sigma_{y^*y^*} = \sigma_v^2 \end{pmatrix} \qquad (1.4.33)$$

The corresponding sample variance covariance matrix of (x^*, y^*) is

$$S = \begin{pmatrix} s_{x^*x^*} & s_{x^*y^*} \\ s_{x^*y^*} & s_{y^*y^*} \end{pmatrix} \qquad (1.4.34)$$

Under the assumption that u_i and v_i are serially independent, Sprent showed that β can be estimated by minimising

$$Q = \frac{\beta^2 s_{x^*x^*} - 2\beta s_{x^*y^*} + s_{y^*y^*}}{\beta^2 \sigma_{x^*x^*} - 2\beta \sigma_{x^*y^*} + \sigma_{y^*y^*}} \qquad (1.4.35)$$

with respect to β. A solution to this problem is possible if Σ is known completely.

If $\sigma_{x^*y^*} = 0$,

$$\hat{\beta} = \frac{s_{y^*y^*} - \theta s_{x^*x^*} + \{(s_{y^*y^*} - \theta s_{x^*x^*}) + 4\theta s^2_{x^*y^*}\}^{\frac{1}{2}}}{2s_{x^*y^*}} \qquad (1.4.36)$$

where $\theta = \dfrac{\sigma_{y^*y^*}}{\sigma_{x^*x^*}}$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

when θ is unknown, they may be replaced by estimates obtained from replicated observations of x^* and y^* corresponding to given values of x_i and y_i of x and y.

The above method has been generalised to the cases where the errors of observations are serially dependent and when the number of regressors is more than one.

For a functional relationship model, Glesser and Watson (1973) attempted to provide maximum likelihood estimates of parameters of the model

$$y_i = B\,x_i + \varepsilon_i, \qquad i = 1,2, \ldots n \tag{1.4.37}$$

$$x_i^* = x_i + u_i, \qquad i = 1,2, \ldots n \tag{1.4.38}$$

and u' are independent of x's and ε's. B is a $p \times p$ matrix of unknown coefficients and each x_i^* is a $p \times 1$ vector having multivariate normal distribution with the common covariance matrix $\sigma^2\Sigma$. The underlying assumption is that Σ is completely known.

In a paper by Warren, White and Fuller (1974), for a structural relation model with errors-in-variables procedure is presented which provides estimators of regression coefficients and their sampling variances. The procedure is illustrated with an example of managerial role performance.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The model is

$$Y = X\beta + \varepsilon \qquad (1.4.39)$$

$$Y^* = Y + v$$
$$X^* = X + U \qquad (1.4.40)$$

where Y is an $n \times 1$ vector, X is an $n \times k$ matrix, β is a $k \times 1$ vector. v and ε are $n \times 1$ vectors and U is an $n \times k$ matrix. It is further assumed that U, v and ε are serially and mutually independent and U and v are independent of X and Y.

$$v \quad NID(0, \sigma_v^2)$$
and
$$U \quad N(0, \Sigma_U) \qquad (1.4.41)$$

Σ_U is known to be diagonal with the i-th diagonal element $\sigma_{u_i}^2$ $i = 1, 2, \ldots n$.

Estimates of $\sigma_{u_i}^2$ are assumed to be available which are mutually independent and independent of X, ε, v and u. The authors derived a consistent asymptotically normal estimator $\hat{\beta}$ of β and showed that this is a method of moment estimator.

For a two variable linear functional relationship model, when both variables are subject to errors, assuming the existence of first and the second order serial correlations ρ_1 and ρ_2 of the true regressors (X, y) of which $\rho_1 \neq 0$ and the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

measurement errors (u and v) to be serially uncorrelated, Karni and Weissman (1974) proposed a method of estimating the slope of the regression line. Earlier, Reiersol (1950) suggested that for such a situation one could take the lagged value of the observed regressor (x^*) as the instrument. Karni and Weissman proved that using $(x^*_{i-1} + x^*_{i+1})$ as an instrument of x^*_i, $i = 2, 3, \ldots n$, yields more efficient estimate than that given by Reiersol. They also proved that using an instrument

$$z_i = x^*_{i-h} + x^*_{i-h+1} + \ldots + x^*_{i-1} + x^*_{i+1} + \ldots + x^*_{i+h}$$

for certain h, one can get a still better estimate of β if $\rho_1 + \rho_2 + \ldots + \rho_h \neq 0$, ρ_k being the k-th lag correlation of the true x's.

For further details on the methods of estimation and their applications, see also Geary (1949), Sargan (1958), Liviatan (1961), Halperin (1961), Carlson, Sobel and Watson (1966), Clutton Brock (1967), Solari (1969), Mallios (1969), Ware (1971).

1.4.5 Other relevant problems.

On a criterion of minimum asymptotic coefficient bias (for OLS estimates), it was shown by Mc. Callum (1972) and Wicken (1972) that faced with a choice of using or discarding a proxy for a relevant unobservable independent variable which appears in a multiple regression model, one should always use a proxy.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Aigner (1974) expanded this analysis to consider the variance in addition to bias in the criterion function. He found that although inclusion of proxy is not always a better strategy, yet is is recommended for a broad range of situations. These results have a bearing on the problem of omission of relevant regressors.

Feldstein (1974) considered the following simple two variable linear relation model with independent errors in variables.

$$\begin{aligned} y_i &= \beta x_i + \varepsilon_i \\ x_i^* &= x_i + u_i \end{aligned} \qquad i = 1, 2, \ldots n \qquad (1.4.42)$$

The basic assumptions are similar to those in usual errors in variable models. The mean square error of the OLS estimator based on (y_i, x_i^*) has been compared with that of the instrumental variable estimator (IVE). An alternative estimator WAIVE (weighted average instrumental variable estimator) has been defined as

$$WAIVE = \lambda(OLSE) + (1 - \lambda)\, IVE \qquad (1.4.43)$$

where λ is chosen to minimise asymptotic MSE of WAIVE. It has been shown that the MSE of WAIVE is sometimes smaller than that of IVE though the two estimators are asymptotically

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

equivalent. On the basis of Monte-Carlo studies, the author concluded that WAIVE is preferable in general to OLSE or IVE.

Cochran (1970) considered the effect of errors in variables on measures of correlation. Assuming Lindley's conditions, linearity of the regression of y on the error affected variables may be exploited. The effect of these errors on the multiple correlation coefficient was examined by Cochran.

Lankipalle (1973) examined the effect of such errors on partial correlation coefficients.

Wu (1973) developed a set of useful tests for the independence of regressors and disturbances term with interesting applications to errors-in-variable models.

Sargan and Mikhail (1971) developed approximations of the Gram-Charlier type to the cumulative distribution function of the instrumental variable estimators.

Langasken and Ryckeghem (1974) presented a method for obtaining the variances of the measurement errors in the major components of national accounts.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Two independent estimates of some economic variable have been taken. Each estimate is composed of the true value of the variable and a measurement error.

$$x_1^* = x + u_1$$
$$x_2^* = x + u_2 \qquad (1.4.44)$$

where $E(u_1) = E(u_2) = 0$, $E(x_1^*) = E(x_2^*) = x$ and $E(x_1^* - x_2^*) = E(u_1 - u_2) = 0$.

For a small number of observations t-test can be used to verify if the sample differences $(x_1^* - x_2^*)$ differ on average significantly from zero. If the hypothesis is rejected, $\sigma_{u_1}^2$ and $\sigma_{u_2}^2$ the variances of the measurement errors u_1 and u_2 can be estimated as follows.

Assuming u_1 and u_2 to be independent and independent of the true value x,

$$\sigma^2_{(x_1^*-x_2^*)} = \sigma^2_{u_1} + \sigma^2_{u_1}$$
$$\sigma^2_{x_1^*} = \sigma^2_x + \sigma^2_{u_1} \qquad (1.4.45)$$
$$\sigma^2_{x_2^*} = \sigma^2_x + \sigma^2_{u_2}$$

Thus from (1.4.45) and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\sigma^2_{u_1} = \frac{\sigma^2_{(x_1^*-x_2^*)} + \sigma^2_{x_1^*} - \sigma^2_{x_2^*}}{2} = \frac{A}{2} \qquad (1.4.46)$$

$$\sigma^2_{u_2} = \frac{\sigma^2_{(x_1^*-x_2^*)} + \sigma^2_{x_2^*} - \sigma^2_{x_1^*}}{2} = \frac{B}{2} \qquad (1.4.47)$$

It is verified that $\sigma^2_{u_1} \leq \sigma^2_{u_2}$ implies $\sigma^2_{x_1^*} \leq \sigma^2_{x_2^*}$.

When x_2^* is a revised estimate of x_1^*, u_1 and u_2 may be expected to be correlated. Let ρ_{u_1,u_2} be the correlation coefficient of u_1 and u_2.

Then after certain algebraic manipulations on the equations

$$\sigma^2_{u_1} + \sigma^2_{u_2} = \sigma^2_{(x_1^*-x_2^*)} + 2\,\rho_{u_1u_2}\,\sigma_{u_1}\,\sigma_{u_2}$$

$$\sigma^2_{u_1} - \rho_{u_1u_2}\,\sigma_{u_1}\,\sigma_{u_2} = \frac{A}{2} \qquad (1.4.48)$$

and $\sigma^2_{u_2} - \rho_{u_1u_2}\,\sigma_{u_1}\,\sigma_{u_2} = \frac{B}{2}$

$$\sigma^2_{u_2} = \frac{B^2}{2B + (A-B)\,\rho^2_{u_1u_2} + \rho_{u_1u_2}\sqrt{4AB + (A-B)^2\rho_{u_1u_2}}} \qquad (1.4.49)$$

From this we get some information on the minimum value of $\rho_{u_1 u_2}$. This minimum $\rho_{u_1u_2}$ follows from

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\rho^2_{u_1u_2} = -\frac{4AB}{(A-B)^2} \tag{1.4.50}$$

A solution of $\sigma^2_{u_2}$ for this minimum $\rho_{u_1u_2}$ always exists. For alternative values of $\rho_{u_1u_2}$ ranging from the estimated minimum to the maximum value of one, we can then estimate $\sigma^2_{u_1}$ and $\sigma^2_{u_2}$.

Here, it is assumed that the formulae used to solve for the population variances $\sigma^2_{u_1}$ and $\sigma^2_{u_2}$ yield unbiased estimate for the sample values $s^2_{u_1}$ and $s^2_{u_2}$ if we replace $\sigma^2_{x^*_1}$, $\sigma^2_{x^*_2}$ and $\sigma^2_{(x^*_1-x^*_2)}$ by their respective unbiased estimates $s^2_{x^*_1}$, $s^2_{x^*_2}$ and $s^2_{(x^*_1-x^*_2)}$.

Lastly, the theoretical implication of the dependency of the measurement errors with the true variable has been examined.

In the field of simultaneous equation models with errors-in-variables, work has been done by Zellner (1970), Goldberger (1972), Robinson (1974) among others.

An extensive literature exists on Bayesian methods of estimation for errors-in-variable models (vide Zellner, 1971, Ch. 5).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 2

AUTOCORRELATED DISTURBANCES IN THE LIGHT OF SPECIFICATION ANALYSIS — PART I

2.1 Introduction

In the literature on single equation methods of econometrics [vide Cochrane and Orcutt, 1949; Johnston, 1972] it is generally recognised that autocorrelation among the disturbances is primarily caused by the omission of relevant regressors from the relationship between the variables. When the disturbances (ε's) are autocorrelated, they are generally assumed to follow the Markov scheme.

$$\varepsilon_t = \rho\, \varepsilon_{t-1} + u_t \quad \text{where } |\rho| < 1,\ E(u_t) = 0\ \forall\ t \text{ and } \operatorname{cov}(u_t, u_s) = \sigma_u^2\, \delta_{ts},\ \delta_{ts} \text{ being kronecker delta.}$$

In this case, the Ordinary Least Squares (OLS) formulae for estimating the sampling variances of the estimated regression coefficients tend to give serious underestimates in some important situations.

If, however, the effect of omission of regressors be examined following the approach of specification analysis due to Theil (1957), the usual formulae appear to overestimate the sampling variances of the estimated regression coefficients of the misspecified equation. So, we arrive at something like a contradiction.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Our first aim in this chapter is to examine this point carefully from the point of view of specification analysis.

We start with the model

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t \qquad t = 1,2,\dots,n \tag{2.1.1}$$

where $E(\varepsilon_t) = 0 \ \forall \ t$ and $\text{cov}(\varepsilon_t, \varepsilon_s) = \sigma^2 \delta_{ts}$, δ_{ts} being the kronecker delta.

The regressors in model (2.1.1) are assumed to be nonstochastic. One of the regressors may be 1 for all t, and the corresponding β will then represent the constant term. In matrix notation the above relationship may be written as

$$\begin{aligned} Y &= X\beta + \varepsilon \\ &= X^{+} \beta^{*} + X^{-} \beta^{**} + \varepsilon \end{aligned} \tag{2.1.2}$$

where $X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix}$

The rank of X is $k < n$. X^{+} is an $n \times m$ matrix containing the first m columns of X and X^{-} is an $n \times (k-m)$ matrix containing the remaining $(k-m)$ columns of X.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$$

$$\beta^* = (\beta_1, \beta_2, \ldots, \beta_m)'$$

and $$\beta^{**} = (\beta_{m+1}, \beta_{m+2}, \ldots, \beta_k)'$$

Now suppose that from (2.1.2), (k-m) regressors $X_{m+1}, x_{m+2}, \ldots, X_k$ have been omitted, that is to say, Y is regressed on X^+. The estimated regression coefficients of $X_1, \ldots, X_m$ will in general be biased estimates of $\beta_1, \ldots, \beta_m$ and may not be so meaningful. If, however, the interest lies mainly in predicting Y from the truncated set of regressors and not in estimating individual structural parameters, the equation may still be quite useful Estimating and testing the significance of individual regression coefficients of such misspecified equations seems to be of practical importance. In developing the methodology of doing these, it seems useful to redefine the regression coefficients associated with the truncated set of regressors, allowing them to capture as much of the partial influence of the omitted regressors on Y as possible, or in other words, to enable the regression function to approximate as closely as possible the systematic components of Y, that is, $\beta_1X_1 + \beta_2X_2 + \ldots + \beta_kX_k$. This is done in section 2 and this leads to a definition of the disturbance term of the misspecified equation. We also examine

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the nature of these disturbances and the effect of applying OLS procedures for point estimation of regression coefficients of the misspecified equation.

Performance of the Durbin-Watson (D-W) (1950, 1951) test of randomness of disturbances of the misspecified equation has been studied in section 3.

Section 4 deals with the performance of the Cochrane-Orcutt two step procedure (1949), the Durbin two step procedure (1960), and the Prais-Winsten method (vide Rao, 1968) of estimating the regression coefficients of the misspecified equation.

Section 5 concludes the chapter with some general observations.

2.2 Model after omission of regressors

The true model is given by (2.1.1). From this model, (k-m) regressors have been omitted. Here we seek to redefine the regression coefficients associated with the regressors $X_1, X_2, \ldots, X_m$ included in the misspecified equation. So, the observational equation may be written as

$$y_t = \beta_1^+ X_{1t} + \beta_2^+ X_{2t} + \ldots + \beta_m^+ X_{mt} + \varepsilon_t^+, \quad t = 1,2,\ldots,n \tag{2.2.1}$$

where β_i^+ is, in general, different from β_i, $i = 1,2,\ldots,m$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The definition of $\beta^+ = (\beta_1^+, \beta_2^+, \ldots, \beta_n^+)'$, and ε^+ will be given below.

In matrix notation (2.2.1) can be written as

$$y = X^+ \beta^+ + \varepsilon^+ \tag{2.2.2}$$

As stated earlier, β^+ should be so defined that X^+ may explain as much of the variation of y as possible. Now, it can be easily shown that

$$E\{(y - X^+\beta^+)' (y - X^+\beta^+)\}$$ will be minimised when

$$\beta^+ = P\beta, \quad \text{where } P = (X^{+\prime} X^+)^{-1} X^{+\prime} X \tag{2.2.3}$$

<u>Proof</u>: $E\{(y - X^+ \beta^+)' (y - X^+\beta^+)\}$

$$= E\{(X\beta - X^+ \beta^+)' + \varepsilon'\}\{(X\beta - X^+\beta^+) + \varepsilon\}$$

$$= (X\beta - X^+ \beta^+)' (X\beta - X^+\beta^+) + E(\varepsilon'\varepsilon)$$

$$= (X\beta - X^+ \beta^+)' (X\beta - X^+ \beta^+) + n\sigma^2.$$

Minimising this with respect to β^+ we get

$$X^{+\prime}(X\beta - X^+\beta^+) = \underset{\sim}{0}$$

or, $$(X^{+\prime}X)\beta - (X^{+\prime} X^+) \beta^+ = 0$$

or, $$\beta^+ = (X^{+\prime} X^+)^{-1} (X^{+\prime} X)\beta$$

$$= P\beta.$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Equation (2.2.3) defines β^+. For finite samples, one gets the best fit on average when β is chosen in this way. The disturbance ε^+ is accordingly defined as

$$\varepsilon^+ = y - X^+ \beta^+ = (X - X^+ P) \beta + \varepsilon \quad \text{1/} \qquad (2.2.4)$$

Now, this pseudo-disturbance ε^+ has some peculiar properties. Thus,

$$E(\varepsilon^+) = (X - X^+ P)\beta \neq \underset{\sim}{0} \quad \text{2/} \qquad (2.2.5)$$

although

$$X^{+\prime}(X - X^+ P) \beta = \underset{\sim}{0} \qquad (2.2.6)$$

Let us denote the t-th element of the vector $(X - X^+P)\beta$ by z_t. Then, if all the elements in the first column of X are equal to 1, (2.2.6) implies

---

1/ In some discussions on the effect of omitted regressors, $X^-\beta^{**} + \varepsilon$ has been taken as the disturbance of the misspecified equation [vide Ramsey (1969)]. The present approach seems to be more appropriate for problems kept in view in view in this chapter. Obviously, ε^+ would be nearer zero than than $X^- \beta^{**} + \varepsilon$ on the whole.

2/ Even when X is stochastic, $E(X - X^+P)\beta \neq \underset{\sim}{0}$. The reason is that $(X-X^+P)\beta = (X^- - X^{+\prime}X^+)^{-1} X^{+\prime}X^-)\beta^{**}$ and the elements of $(X^- - X^+(X^{+\prime}X^+)^{-1} X^{+\prime}X^-)$ are residuals from auxiliary regressions of X^- on X^+. Their expectations do not vanish if the true regressions are not strictly linear.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\sum_{t=1}^{n} z_t = 0 \tag{2.2.7}$$

$$E(\varepsilon^+ \varepsilon^{+\prime}) = (X - X^+P)\ \beta\beta'\ (X - X^+ P)' + \sigma^2 I_n$$

$$\neq \sigma^2 I_n \tag{2.2.8}$$

although the variance covariance matrix of ε^+ is

$$D(\varepsilon^+) = \sigma^2 I_n \tag{2.2.9}$$

From (2.2.5), (2.2.8) and (2.2.9), it is clear that ε_t^+ cannot follow a Markov scheme

$$\varepsilon_t^+ = \rho\ \varepsilon_{t-1}^+ + u_t\ , \tag{2.2.10}$$

where $|\rho| < 1$, $E(u_t) = 0\ \forall\ t$ and $cov(u_t, u_s) = \sigma_u^2\ \delta_{ts}$, δ_{ts} being the kronecker delta, because, (2.2.10) implies, among orther things,

$$E(\varepsilon^+) = \underset{\sim}{0} \tag{2.2.11}$$

and
$$D(\varepsilon^+) = \frac{\sigma_u^2}{(1-\rho^2)} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & & & & \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix} \tag{2.2.12}$$

In fact, if one wants to explain ε_t^+ by (2.2.10), one will get

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(e^{+\prime} e^{+}) = \beta'(X - X^{+}P)'(X - X^{+}P)\beta + (n - m)\sigma^2 \qquad (2.2.18)$$

Therefore, in general,

$$\frac{E(e^{+\prime} e^{+})}{n - m} > \sigma^2 \qquad (2.2.19)$$

which is a familiare result proved by Theil (1957). So, if we estimate $D(\hat{\beta}^{+})$ by $\frac{e^{+\prime} e^{+}}{n - m}(X^{+\prime} X^{+})^{-1}$, the sampling variance of $\hat{\beta}_i^{+}$ is overestimate for all i even though the omission of regressors leads to some apparent autocorrelation among the disturbances.

2.3 Performance of the D.W. test of randomness

Let us assume that in model (2.2.1), $x_{1t} = 1 \; \forall \; t$. One who is not aware of the true nature of disturbances ε^{+} of the model, would like to test their randomness in the usual way by using the Durbin-Watson (D-W) (1949, 1950) test statistic

$$d = \frac{\sum_{t=2}^{n} (e_t^{+} - e_{t-1}^{+})^2}{\sum_{t=1}^{n} e_t^{+2}} \qquad (2.3.1)$$

with a view to deciding upon the procedure of estimation.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(e^{+\prime}\, e^{+}) = \beta'(X - X^{+}P)'\,(X - X^{+}P)\beta + (n - m)\sigma^2 \qquad (2.2.18)$$

Therefore, in general,

$$\frac{E(e^{+\prime}\, e^{+})}{n - m} > \sigma^2 \qquad (2.2.19)$$

which is a familiare result proved by Theil (1957). So, if we estimate $D(\hat{\beta}^{+})$ by $\frac{e^{+\prime}\, e^{+}}{n - m}(X^{+\prime}\, X^{+})^{-1}$, the sampling variance of $\hat{\beta}_i^{+}$ is overestimate for all i even though the omission of regressors leads to some apparent autocorrelation among the disturbances.

2.3 Performance of the D.W. test of randomness

Let us assume that in model (2.2.1), $x_{1t} = 1 \;\forall\; t$. One who is not aware of the true nature of disturbances ε^{+} of the model, would like to test their randomness in the usual way by using the Durbin-Watson (D-W) (1949, 1950) test statistic

$$d = \frac{\sum_{t=2}^{n} (e_t^{+} - e_{t-1}^{+})^2}{\sum_{t=1}^{n} e_t^{+2}} \qquad (2.3.1)$$

with a view to deciding upon the procedure of estimation.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Now,

$$e^{+} = Y - X^{+} \hat{\beta}^{+}$$

$$= X\beta + \varepsilon - X^{+} \hat{\beta}^{+}$$

$$= (X\beta - X^{+} \hat{\beta}^{+}) + \varepsilon \qquad (2.3.2)$$

and $$\hat{\beta}^{+} = P\beta + (X^{+\prime} X^{+})^{-1} X^{+\prime} \varepsilon \qquad (2.3.3)$$

Under usual large sample assumption, that

$$\lim_{n \to \infty} (\frac{1}{n} X^{+\prime} X^{+}) \text{ exists and is nonsingular,} \qquad (2.3.4)$$

$$\operatorname{plim}_{n \to \infty} \hat{\beta}^{+} = P_{\infty} \beta \qquad (2.3.5)$$

where $P_{\infty} = \lim_{n \to \infty} P$.

So, as $n \to \infty$, e^{+} converges in distribution to $(X - X^{+}P_{\infty})\beta + \varepsilon$.

Let $z_{\infty t}$ be the t-th element of the vector $(X - X^{+} P_{\infty})\beta$. Then

$$d \xrightarrow{L} \frac{\sum_{t=2}^{n} [z_{\infty t} + \varepsilon_{t} - (z_{\infty, t-1} + \varepsilon_{t-1})]^2}{\sum_{t=1}^{n} (z_{\infty t} + \varepsilon_{t})^2}$$

By simple algebraic manipulations we get

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}}\ d = 2\left\{1 - \underset{n\to\infty}{\text{plim}} \frac{\frac{1}{n}\sum_{t=2}^{n}(z_{\infty t} + \varepsilon_t)(z_{\infty,t-1} + \varepsilon_{t-1})}{\frac{1}{n}\sum_{t=1}^{n}(z_{\infty t} + \varepsilon_t)^2}\right\}$$

$$= 2\left(1 - \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{t=2}^{n} z_{\infty t}\, z_{\infty,t-1}}{\frac{1}{n}\sum_{t=1}^{n} z_{\infty t}^2 + \sigma^2}\right) \qquad (2.3.6)$$

Since, for finite n,

$X^{+\prime}(X - X^{+}P)\beta = \underset{\sim}{0}$, it can be shown that

$$\lim_{n\to\infty} X^{+\prime}(X - X^{+} P_{\infty}) = \underset{\sim}{0} \qquad (2.3.7)$$

$$\therefore \lim_{n\to\infty} \sum_{t=1}^{n} z_{\infty t} = 0. \qquad (2.3.8)$$

$$\text{Let } \tilde{\rho} = \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{t=2}^{n} z_{\infty,t}\, z_{\infty,t-1}}{\frac{1}{n}\sum_{t=1}^{n} z_{0t}^2} \qquad (2.3.9)$$

$$\text{and } \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} z_{\infty,t}^2 = \sigma_0^2 \qquad (2.3.10)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\therefore \underset{n\to\infty}{\text{plim}} \; d = 2\left(1 - \frac{\tilde{\rho}}{1 + \frac{\sigma^2}{\sigma_o^2}}\right)$$

$$= 2(1 - \rho_o) \qquad (2.3.11)$$

where $\rho_o = \dfrac{\tilde{\rho}}{1 + \sigma^2/\sigma_o^2}$.

If $\tilde{\rho} = 0$, then $\rho_o = 0$. If, however, $\tilde{\rho}$ is positive and σ^2/σ_o^2 is such that ρ_o is appreciably greater than zero, the statistic d would often come out to be significantly smaller than 2 when the sample size is large.

2.4 Performance of some standard methods of estimation

Suppose that d has come out to be significantly less than 2. In this case the usual procedure is to re-estimate the regression coefficient β^+ making the usual assumption that the disturbances of the model (2.1.1) follow the Markov scheme

$$\varepsilon_t^+ = \rho_o \, \varepsilon_{t-1}^+ + u_t \qquad (2.4.1)$$

where $|\rho_o| < 1$, $E(u_t) = 0 \; \forall \; t$ and $\text{cov}(u_t, u_s) = \sigma_u^2 \, \delta_{ts}$, δ_{ts} being the kronecker delta.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Here the symbol ρ_o has been used for the autocorrelation coefficient in anticipation of subsequent results. The definition of ρ_o has been given in previous section. Now, our aim is to examine the performance of the following methods of estimation (which involve the estimation of ρ_o also).

(a) Cochrane-Orcutt two-step method (1949).

(b) Prais-Winsten method (*vide* Rao, 1968).

(c) Durbin two-step procedure (1960).

It will appear that in general, these estimators are inconsistent.

2.4.1 The two step Cochrane-Orcutt procedure

Here one first regresses y on X^+ and obtains the OLS residuals e^+. The next step is to estimate ρ_o by

$$\hat{\rho}_o = \frac{\sum_2^n e_t^+ e_{t-1}^+}{\sum_1^n e_t^{+2}} \tag{2.4.2}$$

The derivation of $\operatorname{plim}_{n\to\infty} d$ in section 3 shows that $\hat{\rho}_o$ gives a consistent estimate of ρ_o. Finally, one fits the equation

$$y_t - \hat{\rho}_o y_{t-1} = \beta_1^+(1 - \hat{\rho}_o) + \beta_2^+(X_{2t} - \hat{\rho}_o X_{2,t-1}) + \dots + \beta_m^+(X_{mt} - \hat{\rho}_o X_{m,t-1}) \tag{2.4.3}$$

by OLS method. This gives,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\hat{\beta}^{+}_{co} = \left(X^{+\prime}_{\hat{\rho}_o} X^{+}_{\hat{\rho}_o}\right)^{-1} X^{+\prime}_{\hat{\rho}_o} y_{\hat{\rho}_o}$$

$$= \left(X^{+\prime}_{\hat{\rho}_o} X^{+}_{\hat{\rho}_o}\right)^{-1} X^{+\prime}_{\hat{\rho}_o} (X^{+}_{\hat{\rho}_o} \beta^{+} + \varepsilon^{+}_{\hat{\rho}_o}) \tag{2.4.4}$$

where
$$X^{+}_{\hat{\rho}_o} = \begin{pmatrix} 1-\hat{\rho}_o & X_{22}-\hat{\rho}_o X_{21} & \cdots & X_{m2}-\hat{\rho}_o X_{m1} \\ 1-\hat{\rho}_o & X_{23}-\hat{\rho}_o X_{22} & \cdots & X_{m3}-\hat{\rho}_o X_{m2} \\ \vdots & & & \\ 1-\hat{\rho}_o & X_{2n}-\hat{\rho}_o X_{2,n-1} & \cdots & X_{mn}-\hat{\rho}_o X_{m,n-1} \end{pmatrix}$$

$$y_{\hat{\rho}_o} = [(y_2-\hat{\rho}_o y_1), (y_3-\hat{\rho}_o y_2) \cdots (y_n-\hat{\rho}_o y_{n-1})]'$$

and $$\varepsilon^{+}_{\hat{\rho}_o} = [(\varepsilon_2-\hat{\rho}_o \varepsilon_1), (\varepsilon_3-\hat{\rho}_o \varepsilon_2) \cdots (\varepsilon_n-\hat{\rho}_o \varepsilon_{n-1})]'$$

From (2.1.3) we get

$$\hat{\beta}^{+}_{co} = \beta^{+} + \left(X^{+\prime}_{\hat{\rho}_o} X^{+}_{\hat{\rho}_o}\right)^{-1} X^{+\prime}_{\hat{\rho}_o} \varepsilon^{+}_{\hat{\rho}_o} \tag{2.4.5}$$

Let us now define as $(n-1) \times n$ matrix $T_{\hat{\rho}_o}$ as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$T_{\hat{\rho}_o} = \begin{pmatrix} -\hat{\rho}_o & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\hat{\rho}_o & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\hat{\rho}_o & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\hat{\rho}_o & 1 \end{pmatrix} \tag{2.4.6}$$

and let

$$T'_{\hat{\rho}_o} T_{\hat{\rho}_o} = W^{-1}_{\hat{\rho}_o} \quad \text{(say)} \tag{2.4.7}$$

$$\therefore \hat{\beta}^{+}_{co} = \beta^{+} + (X^{+\prime} W^{-1}_{\hat{\rho}_o} X^{+})^{-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} \varepsilon^{+}$$

$$= \beta^{+} + (X^{+\prime} W^{-1}_{\hat{\rho}_o} X^{+})^{-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} [(X - X^{+}P)\beta + \varepsilon]$$

$$= (\tfrac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X^{+})^{-1} (\tfrac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X)\beta$$

$$+ (\tfrac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X^{+})^{-1} (\tfrac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} \varepsilon) \tag{2.4.8}$$

Since $\operatorname{plim}_{n\to\infty} \hat{\rho}_o = \rho_o$, $T_{\hat{\rho}_o}$ is a consistent estimator of

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$T_{\rho_o} = \begin{pmatrix} -\rho_o & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho_o & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho_o & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho_o & 1 \end{pmatrix}$$

So, $W^{-1}_{\hat{\rho}_o}$ is a consistent estimator of $W^{-1}_{\rho_o} = (T'_{\rho_o} T_{\rho_o})$.

Let us assume that [3/]

$$\lim_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\rho_o} X = C_{\rho_o} \text{ (exists)} \qquad (2.4.9)$$

$$\lim_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\rho_o} X^{+} = Q_{\rho_o} \text{ (is nonsingular)} \qquad (2.4.10)$$

$$\lim_{n\to\infty} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho_o} W^{-1}_{\rho_o} X^{+}) = \Sigma \text{ (exists)} \qquad (2.4.11)$$

Since $\operatorname{plim}_{n\to\infty} \hat{\rho}_o = \rho_o$, from (2.4.9), (2.4.10) we have,

$$\operatorname{plim}_{n\to\infty} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\hat{\rho}_o} X) = \operatorname{plim}_{n\to\infty} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho_o} X) \qquad (2.4.12)$$

Also, under the assumption that

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} x^2_{h,t-1}, \quad h = 1,2,\ldots,k \qquad (2.4.13)$$

3/ Compare Theil 1971, pp 370-371

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

exists,(which, in fact, exists by (2.3.4)), it can be proved (for proof see Appendix A, that

$$\operatorname*{plim}_{n\to\infty} \frac{1}{\sqrt{n-1}} X^{+\prime} (W^{-1}_{\hat{\rho}_o} - W^{-1}_{\rho_o}) \varepsilon = \underset{\sim}{0} \tag{2.4.14}$$

We also note that $E(\frac{1}{\sqrt{n-1}} X^{+\prime} W^{-1}_{\rho_o} \varepsilon) = \underset{\sim}{0}$ and variance-covariance matrix of $\frac{1}{\sqrt{n-1}} X^{+\prime} W^{-1}_{\rho_o} \varepsilon$ tends to $\sigma^2 \Sigma$ as $n \to \infty$. So, by Chebyshev's inequality (vide Theil 1971 pp 359-360)

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\rho_o} \varepsilon = \underset{\sim}{0}$$ and so, by (2.4.14),

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} \varepsilon = \underset{\sim}{0} \tag{2.4.15}$$

Thus, from (2.4.8), it can be shown that

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}^{+}_{co} = Q^{-1}_{\rho_o} C_{\rho_o} \beta = D_{\rho_o} \beta \text{ (say)} \neq P_{\infty} \beta \tag{2.4.16}$$

Hence, Cochrane-Orcutt two step method does not yield consistent estimator of β^+ in general,

2.4.2 Prais-Winsten method of estimation

Here also one makes the assumption (2.4.1). The variance-covariance matrix of ε^+ is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\frac{\sigma_u^2}{\left(1 - \rho_o^2\right)} V_{\rho_o} = \sigma^2 V_{\rho_o} \quad \text{(say)} \tag{2.4.17}$$

where $\sigma^2 = \dfrac{\sigma_u^2}{(1-\rho_o^2)}$

and
$$V_{\rho_o} = \begin{pmatrix} 1 & \rho_o & \rho_o^2 & \cdots & \rho_o^{n-1} \\ \rho_o & 1 & \rho_o & \cdots & \rho_o^{n-2} \\ \vdots & \vdots & \vdots & & \\ \rho_o^{n-1} & \rho_o^{n-2} & \rho_o^{n-3} & \cdots & 1 \end{pmatrix}$$

Here also (as in (2.4.2)) ρ_o is estimated by

$$\hat{\rho}_o = \frac{\sum_{t=2}^{n} e_t^+ e_{t-1}^+}{\sum_{t=1}^{n} e_t^{+2}}$$

and $\operatorname*{plim}_{n\to\infty} \hat{\rho}_o = \rho_o$.

So, a consistent estimator of V_o is given by

$$V_{\hat{\rho}_o} = \begin{pmatrix} 1 & \hat{\rho}_o & \hat{\rho}_o^{\,2} & \cdots & \hat{\rho}_o^{\,n-1} \\ \hat{\rho}_o & 1 & \hat{\rho}_o & \cdots & \hat{\rho}_o^{\,n-2} \\ \vdots & & & & \\ \hat{\rho}_o^{\,n-1} & \hat{\rho}_o^{\,n-2} & \hat{\rho}_o^{\,n-3} & \cdots & 1 \end{pmatrix} \tag{2.4.18}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In Prais-Winston method, β^{+} is estimated by

$$\hat{\beta}_{P}^{+} = (X^{+\prime} V^{-1}_{\hat{\rho}_o} X^{+})^{-1} (X^{+\prime} V^{-1}_{\hat{\rho}_o} y)$$

$$= (X^{+\prime} V^{-1}_{\hat{\rho}_o} X^{+})^{-1} X^{+\prime} V^{-1}_{\hat{\rho}_o} (X^{+} \beta^{+} + \varepsilon^{+})$$

$$= \beta^{+} + (X^{+\prime} V^{-1}_{\hat{\rho}_o} X^{+})^{-1} X^{+\prime} V^{-1}_{\hat{\rho}_o} [(X - X^{+}P) \beta + \varepsilon]$$

$$= (\tfrac{1}{n} X^{+\prime} V^{-1}_{\hat{\rho}_o} X^{+})^{-1} (\tfrac{1}{n} X^{+\prime} V^{-1}_{\hat{\rho}_o} X) \beta$$

$$+ (\tfrac{1}{n} X^{+\prime} V^{-1}_{\hat{\rho}_o} X^{+})^{-1} (\tfrac{1}{n} X^{+\prime} V^{-1}_{\hat{\rho}_o} \varepsilon) \qquad (2.4.19)$$

Let us assume that [4/]

$$\lim_{n \to \infty} \frac{1}{n} (X^{+\prime} V^{-1}_{\rho_o} X) = \tilde{C}_{\rho_o} \text{ (exists)} \qquad (2.4.20)$$

4/

$$V^{-1}_{\rho_o} = \frac{1}{(1-\rho_o^2)} \tilde{T}'_{\rho_o} \tilde{T}_{\rho_o} \text{ and } \tilde{T}_{\rho_o} = \begin{pmatrix} \sqrt{1-\rho_o^2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho_o & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho_o & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & -\rho_o & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\rho_o & 1 \end{pmatrix}$$

Comparing this with T_{ρ_o} defined in section (2.4.1), it can be proved easily that assumptions in (2.4.20), (2.4.21) and (2.4.22) are approximately equal to the assumptions in (2.4.9), (2.4.10) and (2.4.11) respectively.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\lim_{n\to\infty} \frac{1}{n} (X^{+\prime} V_{\rho_o}^{-1} X^{+}) = \tilde{Q}_{\rho_o} \quad \text{(is nonsingular)} \tag{2.4.21}$$

$$\lim_{n\to\infty} \frac{1}{n} (X^{+\prime} V_{\rho_o}^{-1} V_{\rho_o}^{-1} X^{+}) = \Omega \quad \text{(a positive definite)} \tag{2.4.22}$$

Since $\text{plim}_{n\to\infty} \hat{\rho}_o = \rho_o$, from (2.4.20) and (2.4.21) we have

$$\text{plim}_{n\to\infty} \frac{1}{n} (X^{+\prime} V_{\hat{\rho}_o}^{-1} X) = \lim_{n\to\infty} \frac{1}{n}(X^{+\prime} V_{\rho_o}^{-1} X) \tag{2.4.23}$$

Also, under the assumption (2.4.13), it can be proved as in (2.4.15) (see also the Appendix A) that

$$\text{plim}_{n\to\infty} \frac{1}{\sqrt{n}} X^{+\prime} (V_{\hat{\rho}_o}^{-1} - V_{\rho_o}^{-1})\varepsilon = \underset{\sim}{0} \tag{2.4.24}$$

As in section 2.4.1, here also, it can be proved easily by Chebychev's inequality that

$$\text{plim}_{n\to\infty} \frac{1}{n} X^{+\prime} V_{\rho_o}^{-1} \varepsilon = \underset{\sim}{0} \quad \text{and hence, from (2.4.24),}$$

$$\text{plim}_{n\to\infty} \frac{1}{n} X^{+\prime} V_{\hat{\rho}_o}^{-1} \varepsilon = \underset{\sim}{0} \tag{2.4.25}$$

Hence, from (2.4.19), it can be shown that

$$\text{plim}_{n\to\infty} \hat{\beta}_{PW}^{+} = \tilde{Q}_{\rho_o}^{-1} \tilde{C}_{\rho_o} \beta = \tilde{D}_{\rho_o} \beta \text{ (say)} \neq P_{\infty} \beta \text{ (in general)} \tag{2.4.26}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, Prais-Winsten procedure does not yield consistent estimator of β^+ in general.

2.4.3 Durbin's two-step procedure

Here we consider the equation

$$y_t = (1-\rho_o)\beta_1^+ + \rho_o\, y_{t-1} + \beta_2^+ X_{2t} - \rho_o\beta_2^+ X_{2,t-1} + \cdots$$
$$+ \beta_m^+ X_{mt} - \rho_o\beta_m^+ X_{m,t-1} + (\varepsilon_t^+ - \rho_o\, \varepsilon_{t-1}^+)$$
$$t = 2,\ldots,n \qquad (2.4.27)$$

In matrix notation, (2.4.27) can be written as

$$\bar{y} = \bar{X}^+ \bar{\beta}^+ + \varepsilon^+_{\rho_o} \qquad (2.4.28)$$

where $\bar{y} = (y_2, y_3,\ldots,y_n)'$

$$\text{and } \bar{X}^+ = \begin{pmatrix} 1 & y_1 & X_{22} & X_{21} & \cdots & X_{m2} & X_{m1} \\ 1 & y_2 & X_{23} & X_{22} & \cdots & X_{m3} & X_{m2} \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ 1 & y_{n-1} & X_{2n} & X_{2,n-1} & \cdots & X_{mn} & X_{m,n-1} \end{pmatrix}$$

$$\bar{\beta}^+ = [(1-\rho_o)\beta_1^+, \rho_o, \beta_2^+, -\rho_o\beta_2^+,\ldots,\beta_m^+, -\rho_o\,\beta_m^+]'$$

and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\varepsilon^+_{\rho_o} = [(\varepsilon^+_2 - \rho_o \varepsilon^+_1), (\varepsilon^+_3 - \rho_o \varepsilon^+_2), \ldots, (\varepsilon^+_n - \rho_o \varepsilon^+_{h-1})]'$$

using OLS procedure we get

$$\hat{\bar{\beta}}^+ = (\bar{X}^{+\prime} \bar{X}^+)^{-1} \bar{X}^{+\prime} \bar{y}$$

$$= \bar{\beta}^+ + (\bar{X}^{+\prime} \bar{X}^+)^{-1} \bar{X}^{+\prime} \varepsilon^+_{\rho_o} \qquad (2.4.29)$$

Now,

$$\varepsilon^+_{\rho_o} = \begin{pmatrix} z_2 - \rho_o z_1 \\ z_3 - \rho_o z_2 \\ \vdots \\ z_n - \rho_o z_{n-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_2 - \rho_o \varepsilon_1 \\ \varepsilon_3 - \rho_o \varepsilon_2 \\ \vdots \\ \varepsilon_n - \rho_o \varepsilon_{n-1} \end{pmatrix}$$

$$= z_{\rho_o} + \varepsilon_{\rho_o} \text{ (say)} \qquad (2.4.30)$$

Now, $z_{\rho_o} = (X_{\rho_o} - X^+_{\rho_o} P)\beta$

where

$$X_{\rho_o} = \begin{pmatrix} 1-\rho_o & X_{22} - \rho_o X_{21} & \cdots & X_{k2} - \rho_o X_{k1} \\ 1-\rho_o & X_{23} - \rho_o X_{22} & \cdots & X_{k3} - \rho_o X_{k2} \\ \vdots & \vdots & & \vdots \\ 1-\rho_o & X_{2n} - \rho_o X_{2,n-1} & \cdots & X_{kn} - \rho_o X_{k,n-1} \end{pmatrix}$$

and $X^+_{\rho_o}$ contains the first m columns of X_{ρ_o}.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

It can be shown that

$$(X_{\rho_o} - X^{+}_{\rho_o} P)\beta = \bar{X}\bar{\beta} - \bar{X}^{+}\bar{\beta}^{+} \qquad (2.4.31)$$

where $\bar{X} = \begin{pmatrix} 1 & y_1 & X_{22} & X_{21} & \cdots & X_{k2} & X_{k1} \\ 1 & y_2 & X_{23} & X_{22} & \cdots & X_{k3} & X_{k2} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & y_{n-1} & X_{2n} & X_{2,n-1} & \cdots & X_{kn} & X_{k,n-1} \end{pmatrix}$

$$= (\bar{X}^{+}, \bar{X}^{-})$$

and $\bar{\beta} = [(1-\rho_o), \rho_o, \beta_2, -\rho_o\beta_2, \ldots, \beta_k, -\rho_o\beta_k]'$

So

$$\hat{\bar{\beta}}^{+} = \bar{\beta}^{+} + (\bar{X}^{+\prime}\bar{X}^{+})^{-1}\bar{X}^{+\prime}(\bar{X}\bar{\beta} - \bar{X}^{+}\bar{\beta}^{+}) + (\bar{X}^{+\prime}\bar{X}^{+})^{-1}\bar{X}^{+\prime}\varepsilon_{\rho_o}$$

$$= (\bar{X}^{+\prime}\bar{X}^{+})^{-1}\bar{X}^{+\prime}\bar{X}\bar{\beta} + (\bar{X}^{+\prime}\bar{X}^{+})^{-1}\bar{X}^{+\prime}\varepsilon_{\rho_o}$$

$$= \bar{\beta}^{*} + (\bar{X}^{+\prime}\bar{X}^{+})^{-1}(\bar{X}^{+\prime}\bar{X}^{-})\bar{\beta}^{**} + (\bar{X}^{+\prime}\bar{X}^{+})^{-1}\bar{X}^{+\prime}\varepsilon_{\rho_o} \qquad (2.4.32)$$

where $\bar{\beta}^{*} = [(1-\rho_o)\beta_1, \rho_o, \beta_2, -\rho_o\beta_2, \ldots, \beta_m, -\rho_o\beta_m]'$

and $\bar{\beta}^{**} = (\beta_{m+1}, -\rho_o\beta_{m+1}, \ldots, \beta_k, -\rho_o\beta_k)'$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\therefore \underset{n\to\infty}{\text{plim}} \hat{\bar{\beta}}^{+} = \bar{\beta}^{*} + \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^{+}\right)^{-1} \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} X^{\prime +\prime} \bar{X}^{-}\right) \bar{\beta}^{**}$$

$$+ \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^{+}\right)^{-1} \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \varepsilon_{\rho_o}\right)$$

(2.4.33)

Now, since X is uncorrelated with ε under the assumption (2.3.4) (for proof see the Appendix 1B) only the second element of $\frac{1}{n-1}(\bar{X}^{+\prime} \varepsilon_{\rho_o})$ has now zero probability limit. The probability limit of this element is given by

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1}(y_1, y_2, \ldots, y_{n-1})' [(\varepsilon_2 - \rho_o \varepsilon_1), (\varepsilon_3 - \rho_o \varepsilon_2) \ldots (\varepsilon_n - \rho_o \varepsilon_{n-1})]'$$

$$= - \rho_o \sigma^2 \qquad (2.4.34)$$

provided the fourth order moment of ε_t's exists.

So,

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \varepsilon_{\rho_o}\right) = (0, -\rho_o \sigma^2, 0, 0, \ldots, 0)' = L \text{ (say)} \qquad (2.9.35)$$

Let us assume that the following stochastic limits exists :

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^{+}\right)^{-1} = N = (n_{ij})_{2m \times 2m}$$ is a positive definite matrix and

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^{-}\right) = C = (c_{ij})_{2m \times 2(k-m)}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\therefore \operatorname*{plim}_{n\to\infty} \hat{\bar{\beta}}^{+} = \bar{\beta}^{*} + N\,C\,\bar{\beta}^{**} + NL \qquad (2.4.36)$$

In Durbin's two step procedure, the second element of $\bar{\beta}^{+}$ is taken as ρ. Obviously from (2.4.36),

$$\operatorname*{plim}_{n\to\infty} \hat{\rho}_o \neq \rho_o \text{ (necessarily)} \qquad (2.4.37)$$

5/

Let, $$\operatorname*{plim}_{n\to\infty} \hat{\rho}_o = \rho_o + A \text{ (say)} \qquad (2.4.38)$$

where A is not necessarily zero.

Durbin's second step consists in fitting the equation

$$y_t - \hat{\rho}_o y_{t-1} = \beta_1^{+}(1 - \hat{\rho}_o) + \beta_2^{+}(X_{2t} - \hat{\rho}_o X_{2,t-1}) + \dots +$$

$$\beta_m^{+}(X_{mt} - \hat{\rho}_o X_{m,t-1}) + (\varepsilon_t - \hat{\rho}_o \varepsilon_{t-1}) \qquad (2.4.39)$$

by OLS method. This gives

$$\hat{\beta}_D^{+} = (X_{\hat{\rho}_o}^{+\prime} X_{\hat{\rho}_o}^{+})^{-1} X_{\hat{\rho}_o}^{+\prime} y_{\hat{\rho}_o}$$

$$= (X_{\hat{\rho}_o}^{+\prime} X_{\hat{\rho}_o}^{+})^{-1} X_{\hat{\rho}_o}^{+\prime}(X_{\hat{\rho}_o}^{+} \beta + \varepsilon_{\hat{\rho}_o}^{+}) \qquad (2.4.40)$$

5/ This $\hat{\rho}_o$ is, obviously different from $\hat{\rho}_o$ in (2.4.2).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $X^+_{\hat{\rho}_o}$, $\varepsilon^+_{\hat{\rho}_o}$ and $y_{\hat{\rho}_o}$ are defined in (2.4.4).

$W^{-1}_{\hat{\rho}_o} = T'_{\hat{\rho}_o} T_{\hat{\rho}_o}$ as defined before.

From (2.4.33), we get

$$\hat{\beta}_D^+ = \beta^+ + (X^{+\prime}_{\hat{\rho}_o} X^+_{\hat{\rho}_o})^{-1} X^{+\prime}_{\hat{\rho}_o} \varepsilon^+_{\hat{\rho}_o} \tag{2.4.41}$$

$$= \beta^+ + (X^{+\prime} W^{-1}_{\hat{\rho}_o} X^+)^{-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} \varepsilon^+$$

$$= \beta^+ + (X^{+\prime} W^{-1}_{\hat{\rho}_o} X^+)^{-1} X^{+\prime} W^{-1}_{\hat{\rho}_o}[(X - X^+P)\beta + \varepsilon]$$

$$=(\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X^+)^{-1}(\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X)\beta$$

$$+(\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} X^+)^{-1}(\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}_o} \varepsilon) \tag{2.4.42}$$

Since $\underset{n\to\infty}{\text{plim}}\, \hat{\rho}_o = \rho_o + A$(say), $T_{\hat{\rho}_o}$ is a consistant estimator of $T_{(\rho_o+A)}$ where $T_{(\rho_o+A)}$ is obtained by putting $(\rho_o + A)$ for ρ_o in T_{ρ_o}. So, $W^{-1}_{\hat{\rho}_o}$ is a consistent estimator of $W^{-1}_{(\rho_o+A)} = T'_{(\rho_o+A)} T_{(\rho_o+A)}$. Let us assume that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\lim_{n\to\infty} \left(\frac{1}{n-1} X^{+\prime} W^{-1}_{(\rho_0+A)} X\right) = C_{(\rho_0+A)} \text{ exists and}$$

$$\lim_{n\to\infty} \left(\frac{1}{n-1} X^{+\prime} W^{-1}_{(\rho_0+A)} X^{+}\right) = Q_{(\rho_0+A)} \text{ is nonsingular.}$$

Under the assumptions similar to those in section 2.4.1, it can be shown that

$$\begin{aligned} \operatorname*{plim}_{n\to\infty} \hat{\beta}_D^{+} &= Q^{-1}_{(\rho_0+A)} C_{(\rho_0+A)} \beta \\ &= D_{(\rho_0+A)} \beta \neq P_{\infty} \beta \text{ (in general)} \end{aligned} \qquad (2.4.43)$$

So, the Durbin two step procedure does not yield a consistent estimator of β^{+}.

2.4.4 An overview of three alternative methods of estimation.

From (2.4.16), we have,

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_{CO}^{+} = D_{\rho_0} \beta$$

and from (2.4.26),

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_{PW}^{+} = \tilde{D}_{\rho_0} \beta \text{ .}$$

From foot note 4 , it is also obvious that

$$D_{\rho_0} \simeq \tilde{D}_{\rho_0} \qquad (2.4.44)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So,

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{CO} \simeq \underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{PW} \qquad (2.4.45)$$

From (2.4.43),

$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{D} = D_{(\rho_o+A)}\ \beta$ and $D_{(\rho_o+A)}$ depends on

$W^{-1}_{(\rho_o+A)} = T'_{(\rho_o+A)}\ T_{(\rho_o+A)}$. $T_{(\rho_o+A)}$ is different from both

T_{ρ_o} and $\tilde{T}_{\rho_o}$. So,

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{D} \neq \underset{n\to\infty}{\text{plim}}\ \beta^{+}_{CO} \simeq \underset{n\to\infty}{\text{plim}}\ \beta^{+}_{PW} \qquad (2.4.46)$$

and, $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{CO}$, $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{PW}$ and $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{D}$ are all different

from $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+} = P_{\infty}\ \beta$ (in general).

5. Conclusion

The main results of this chapter may be briefly summarised as follows :

1. In the classical linear regression model with nonstochastic regressors, when some regressors have been omitted from an equation,

   (a) the OLS formulae for estimating sampling variances of the estimated regression coefficients will give over-estimates.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(b) The disturbances in the misspecified model will not follow a Markov scheme. They will be mutually uncorrelated with nonzero means. But the elements in the mean vector of these disturbances may be such as to give rise to some appearance of autocorrelation among the disturbances and as a result the D-W statistic may come out to be significantly less than 2 in large samples (provided the first order autocorrelation coefficient of the elements in the mean vector is positive).

(c) Cochrane-Orcutt two step, Prais-Winsten and Durbin two step estimators of the regression coefficients of the misspecified model are, in general, inconsistent. While the probability limits are approximately equal for the first two estimators, they are different from the Durbin two step estimator.

In the next chapter, we shall extend these results to the case where the true model has stochastic regressors and/or autocorrelated disturbances. (Such disturbances may occur due to autocorrelated errors in variables). On the basis of the above analysis it appears that some of the standard methods of testing and estimation mentioned in the literature on the problem of autocorrelated disturbances require modification if the autocorrelation is partly or wholly due to omission of non-stochastic regressors from the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

true model. Attempts are being made to extend the above results to the case of stochastic regressors and to the case where the algebraic form of the equation is misspecified.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 3

AUTOCORRELATED DISTURBANCES IN THE LIGHT OF SPECIFICATION ANALYSIS — PART II

3.1 Introduction

So far we have examined the effect of omission of relevant regressors from a regression equation where the regressors are nonstochastic and the disturbances are spherical. In this chapter we shall consider the case where the regressors are stochastic and the disturbances in the true model are themselves autocorrelated (possibly due to autocorrelated errors of observations) and examine in the same manner as in Chapter 2, the effects of omission of regressors from the regression equation. We shall also consider the following subcases :

1. The regressors are stochastic and the disturbances in the true model are spherical.

2. The regressors are nonstochastic and the disturbances in the true model are autocorrelated.

In section 2 we redefine the regression coefficients associated with the included regressors and also the disturbances of the misspecified equation; and we also study the nature of these disturbances. In section 3, we examine the performance of the OLS procedures for point estimation of the regression coefficients of the misspecified equation and also of the associated standard

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

errors. Performance of the Durbin-Watson (1950, 1951) test of randomness of disturbances of the misspecified equation has been studied in section 4. Section 5 deals with the performance of the Cochrane-Orcutt two step procedure (1949) the Durbin two-step procedure (Durbin, 1960) and the Prais-Winsten method (vide Rao, 1968) when applied to estimate the regression coefficient of the misspecified equation. Section 6 concludes the chapter with some general observations.

3.2 Model after omission of regressors

The true model is

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t, \quad t = 1,2,\dots,n \tag{3.2.1}$$

where $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$; $t = 2,3,\dots,n$;

$|\rho| < 1$ and u_t is the spherical disturbance term having mean zero and variance σ_u^2. ε's are assumed to be independent of X's. We further assume that

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} X'X\right) = \Sigma_{XX} \text{ (a positive definite matrix)} \tag{3.2.2}$$

The regressors in model (2.1) are assumed to be stochastic. One of the regressors may be 1 for all t ; and the corresponding β will then represent the constant term. In matrix notation (3.2.1) can be written as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y = X\beta + \varepsilon$$

$$= X^{+}\beta^{*} + X^{-}\beta^{**} + \varepsilon \qquad (3.2.3)$$

where X^{+} and X^{-}, β^{*} and β^{**} as same as in Chapter 1 and $X = (X^{+} \vdots X^{-})$.

Now, (2.1) implies that

$$E(\varepsilon) = \underset{\sim}{0} \qquad (3.2.4)$$

and $$E(\varepsilon\varepsilon') = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & & & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix} = \sigma^2 V \text{ (say)}$$

where $\sigma^2 = \dfrac{\sigma_u^2}{1-\rho^2}$ (3.2.5)

Suppose from the regression equation (3.2.1), the (k-m) regressors included in X^{-} have been omitted. In such a case, one tacitly takes $X^{-}\beta^{**} + \varepsilon$ as the disturbance term of the misspecified model [*vide* Ramsey (1969)]. But this is untenable. The regression coefficients associated with the included set of regressors should be redefined in such a way that they may capture as much of the partial influence of X^{-} on y as possible. The misspecified equation may be written as

$$y = \beta_1^{+} X_1 + \beta_2^{+} X_2 + \cdots + \beta_m^{+} X_m + \varepsilon^{+}$$

or, $$y = X^{+}\beta^{+} + \varepsilon^{+} \qquad (3.2.6)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where β_i^+ is, in general, different from β_i, $i = 1,2,\ldots,m$. The definition of $\beta^+ = (\beta_1^+, \beta_2^+,\ldots,\beta_m^+)'$ and ε^+ will be given below.

It can be easily shown that

$E\{(y - X^+ \beta^+)' (y - X^+\beta^+)\}$ will be minimised when

$$\beta^+ = \bar{P}\beta \tag{3.2.7}$$

where $\bar{P} = \{E(X^{+\prime} X^+)\}^{-1} E(X^{+\prime} X)$.

<u>Proof</u> : $E\{(y - X^+\beta^+)' (y - X^+\beta^+)\}$

$$= E\{(X\beta - X^+\beta^+)' (X\beta - X^+\beta^+)\} + n\sigma^2$$

Differentiating with respect to β^+ and equating the derivatives to zero, we get

$$-2E(X^{+\prime}X)\beta + 2E(X^{+\prime} X^+)\beta^+ = 0$$

$$\text{or,}\quad \beta^+ = \{E(X^{+\prime} X^+)\}^{-1} E(X^{+\prime}X)\beta .$$

$$= \bar{P}\beta.$$

So, the model after omission of regressors may be specified as

$$y = X^+\beta^+ + \varepsilon^+ \tag{3.2.8}$$

where β^+ is given by (3.2.7) and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\varepsilon^{+} = (y - X^{+}\beta^{+})$$
$$= (X - X^{+}\bar{P})\,\beta + \varepsilon \qquad (3.2.8)$$

$$\therefore\ E(\varepsilon^{+}) = E(X - X^{+}\bar{P})\beta \neq 0 \quad \text{in general} \qquad (3.2.9)$$

If however, the regression of X^{-} on X^{+} are strictly linear, then $E(X - X^{+}\bar{P})\beta$ vanishes and ε^{+} has zero mean. This can be proved in the following manner :

Suppose,

$$X^{-} = X^{+}\delta + \nu \qquad (3.2.10)$$

where δ is the matrix of regression coefficients associated with X^{+} and ν is the matrix of the disturbances with mean 0. If the regressions of X^{-} on X^{+} are strictly linear and is independent of X^{+},

$$E(\nu \mid X^{+}) = 0 \quad \text{(null matrix)} \qquad (3.2.11)$$

Then, it can be proved easily that the lines of best fit are given by

$$\delta = \bar{P}^{**} = \{E(X^{+\prime} X^{+})\}^{-1} E(X^{+\prime} X^{-}) \qquad (3.2.12)$$

$$\therefore\ \nu = (X^{-} - X^{+}\bar{P}^{**}) \qquad (3.2.13)$$

$$\therefore\ E(\nu \mid X^{+}) = 0 \quad \text{(null matrix)},$$

$$E\{(X^{-} - X^{+}\bar{P}^{**}) \mid X^{+}\} = 0 \qquad (3.2.14)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Now,

$$E(\varepsilon^+) = E(X - X^+\bar{P})\beta = E\{(X^+ \vdots X^-) - X^+(I \vdots \bar{P}^{**})\}(\beta^* \vdots \beta^{**})$$

$$= E(X^- - X^+ \bar{P}^{**})\beta^{**}$$

$$= E(\nu \beta^{**}) = \underset{\sim}{0} \quad \text{(by virtue of (3.2.14))} \qquad (3.2.15)$$

In general, however, ε^+ has nonzero mean given by (3.2. 9). In this case, the disturbances (ε^+'s) have some interesting properties : e.g. ;

$$E(\varepsilon^+ \varepsilon^{+\prime}) = E\{(X - X^+\bar{P})\beta\beta'(X - X^+\bar{P})'\} + \sigma^2 V \qquad (3.2.16)$$

and $$D(\varepsilon^+) = E[\{\varepsilon^+ - E(\varepsilon^+)\}\{\varepsilon^+ - E(\varepsilon^+)\}'] = \tilde{V} + \sigma^2 V \qquad (3.2.17)$$

where $\tilde{V} = E[\{\tilde{z} - E(\tilde{z})\}\{\tilde{z} - E(\tilde{z})\}']$

and $\tilde{z} = (\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_n)' = E(X - X^+ \bar{P})\beta$.

From previous discussions it follows easily that ε_t^+ cannot, in general, follow a Markov scheme of the form

$$\varepsilon_t^+ = \tilde{\rho}\, \varepsilon_{t-1}^+ + v_t \quad , \quad t = 2,3,\ldots,n \quad |\tilde{\rho}| < 1 \text{ and } v_t \text{ is spherical} \qquad (3.2.18)$$

(3.2.18) would imply

$$E(\varepsilon^+) = \underset{\sim}{0} \qquad (3.2.19)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{and} \quad D(\varepsilon^+) = \frac{\sigma_u^2}{(1-\tilde{\rho}^2)} \begin{pmatrix} 1 & \tilde{\rho} & \tilde{\rho}^2 & \cdots & \tilde{\rho}^{n-1} \\ \tilde{\rho} & 1 & \tilde{\rho}^2 & \cdots & \tilde{\rho}^{n-2} \\ \vdots & & & & \vdots \\ \tilde{\rho}^{n-1} & \tilde{\rho}^{n-2} & \tilde{\rho}^{n-3} & \cdots & 1 \end{pmatrix} \tag{3.2.20}$$

Further, if one tries to explain ε_t^+, by(3.2.17), one will obtain

$$\tilde{z}_t + \varepsilon_t = \tilde{\rho} \, (\tilde{z}_{t-1} + \varepsilon_{t-1}) + v_t$$

or,

$$\varepsilon_t - \tilde{\rho} \, \varepsilon_{t-1} = \tilde{\rho} \, \tilde{z}_{t-1} - \tilde{z}_t + v_t \tag{3.2.21}$$

This implies $E(\varepsilon_t - \tilde{\rho} \, \varepsilon_{t-1}) = E(\tilde{\rho} \, \tilde{z}_{t-1} - \tilde{z}_t) \neq 0$ (in general)

(3.2.22)

and

$$\text{cov}\{(\varepsilon_t^+ - \tilde{\rho} \, \varepsilon_{t-1}^+), (\varepsilon_{t-1}^+ - \tilde{\rho} \, \varepsilon_{t-2}^+)\}$$

$$= \text{cov}\{(\tilde{z}_t - \tilde{\rho} \, \tilde{z}_{t-1}), (\tilde{z}_{t-1} - \tilde{\rho} \, \tilde{z}_{t-2})\} + \sigma^2(1 - \rho \tilde{\rho}) (\rho - \tilde{\rho}) \neq 0$$

(in general) (3.2.23)

whereas (3.2.18) requires this covariance to be zero.

Although ε_t^+ cannot follow a Markov scheme, yet the elements of the vector ε^+ will be autocorrelated and the elements of the mean vector $E\{(X - X^+ \bar{P})\beta\}$ may also be autocorrelated.

Let us now consider the following cases.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Case 1. ε in model (3.2.1) is spherical : In this case, essentially all the results remain unaltered. Only V in (3.2.5) will be replaced by I_n.

Case 2. The regressors in model (3.2.1) are nonstochastic while ε's follow the Markov scheme : Here, since the regressors are nonstochastic,

$$E\{(y - X^+\beta^+)'(y - X^+\beta^+)\} \text{ will be minimised when}$$

$$\beta^+ = (X^{+\prime} X^+)^{-1} X^{+\prime} X\beta = P\beta \quad \text{(say)}$$

defined in (2.2.   ) of Chapter 2.

So, in the misspecified model,

$$\varepsilon^+ = (X - X^+P)\beta + \varepsilon$$
$$= z + \varepsilon \qquad (3.2.24)$$

where $z = (z_1, z_2, \ldots, z_n)' = (X - X^+P)\beta$.

In this case,

$$E(\varepsilon^+) = (X - X^+P)\beta = [X^- - X^+(X^{+\prime} X^+)^{-1} X^{+\prime} X^-]\beta^{**}$$
$$= (X^- - X^+ P^{**})\beta^{**} \neq \underset{\sim}{0} \quad \text{in general} \qquad (3.2.25)$$

$$E(\varepsilon^+ \varepsilon^{+\prime}) = (X-X^+P)\beta\beta'(X-X^+P)' + \sigma^2 V \qquad (3.2.26)$$

and $$D(\varepsilon^+) = \sigma^2 V \qquad (3.2.27)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Here also, it can be proved easily that ε_t^+ cannot follow a Markov scheme of the form (3.2.18). However, because of the autocorrelations of ε's and also because in practice the elements of the mean vector $(X - X^+P)\beta$ would appear to be autocorrelated if considered as a time series, the ε_t^+'s would, in general be autocorrelated.

3.3 Performance of OLS procedures

From now on, we shall assume that in equation (3.2.3), $x_{1t} = 1 \; \forall \; t$. So, the equation (3.2.6) can be written as

$$y = X_o^+ \beta_o^+ + \varepsilon \qquad (3.3.1)$$

where X_o^+ consists of n rows $(1, x_{2t}^o, \ldots, x_{mt}^o)$ and $x_{it}^o = x_{it} - \bar{x}_i$, $t = 1,2,\ldots,n$ and $i = 2,3,\ldots,m$, and $\bar{x}_i = \frac{\sum_{t=1}^{n} \bar{x}_{it}}{n}$

$$\beta_o^+ = [(\beta_1^+ + \sum_{i=2}^{m} \beta_i^+ \bar{x}_i), \beta_2^+, \ldots, \beta_m^+]'$$

$$= (\tilde{\beta}_1^+, \beta_2^+, \ldots, \beta_m^+)'.$$

The OLS estimate of β_o^+ is given by

$$\hat{\beta}_o^+ = (X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} y$$

$$= (X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} (X_o^+ \beta_o^+ + \varepsilon^+)$$

$$= \beta_o^+ + (X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} (\tilde{z} + \varepsilon) \quad (\because \varepsilon^+ = \tilde{z} + \varepsilon) \qquad (3.3.2)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\therefore \; E(\hat{\beta}_o^+) = \beta_o^+ + E\{(X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} \tilde{z}\}$$

$$\neq \beta_o^+ \quad \text{(in general)} \quad \text{since} \quad E(\tilde{z} \mid X^+) \neq \underset{\sim}{0} \quad \text{in general} \tag{3.3.3}$$

Again,

$$\underset{n\to\infty}{\text{plim}} \; \hat{\beta}_o^+ = \underset{n\to\infty}{\lim} \; \beta_o^+ + \underset{n\to\infty}{\text{plim}} \; (X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime}(\tilde{z} + \varepsilon)$$

$$= \underset{n\to\infty}{\lim} \; \beta_o^+ + \underset{n\to\infty}{\text{plim}} \; (X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} \tilde{z} \quad \text{(under fairly general conditions)}$$

$$= \underset{n\to\infty}{\lim} \; \beta_o^+ + \underset{n\to\infty}{\text{plim}} \; (\tfrac{1}{n} X_o^{+\prime} X_o^+)^{-1} \; \underset{n\to\infty}{\text{plim}} \; (\tfrac{1}{n} X_o^{+\prime} \tilde{z})$$

$$= \underset{n\to\infty}{\lim} \; \bar{P}\beta = \bar{P}_\infty \beta \text{ (say)} \tag{3.3.4}$$

where $\bar{P}_\infty = \underset{n\to\infty}{\lim} \; \bar{P}$.

<u>Proof</u> : Let us first consider

$$\underset{n\to\infty}{\text{plim}} \; \frac{1}{n} (X^{+\prime} \tilde{z})$$

$$= \underset{n\to\infty}{\text{plim}} \; \frac{1}{n} X^{+\prime} (X - X^+ \bar{P}) \beta$$

$$= \Big[\underset{n\to\infty}{\text{plim}} \; (\tfrac{1}{n} X^{+\prime} X) - \underset{n\to\infty}{\text{plim}} \; (\tfrac{1}{n} X^{+\prime} X^+) \cdot \underset{n\to\infty}{\lim} \Big\{E\Big(\frac{X^{+\prime} X^+}{n}\Big)\Big\}^{-1} \times \underset{n\to\infty}{\lim} \; E\Big(\frac{X^{+\prime} X}{n}\Big) \Big] \tag{3.3.5}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Since by (3.2.2), $\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} X'X\right) = \Sigma_{XX}$,

So,

$$\lim_{n\to\infty} E\left(\frac{1}{n} X'X\right) = \Sigma_{XX} \quad \text{(\underline{vide} Goldberger (1968, pp. 118-119)} \tag{3.3.6}$$

Now,

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} X^{+\prime}X\right) = \underset{n\to\infty}{\text{plim}} \left\{\frac{1}{n}(X^{+\prime}X^{+} \vdots X^{+\prime}X^{-})\right\} = (\Sigma_{X^+X^+} \vdots \Sigma_{X^+X^-}) \tag{3.3.7}$$

where $\Sigma_{X^+X^+}$ and $\Sigma_{X^+X^-}$ are the submatrices of the matrix

$$\Sigma_{XX} = \begin{pmatrix} \Sigma_{X^+X^+} & \Sigma_{X^+X^-} \\ \Sigma'_{X^+X^-} & \Sigma_{X^-X^-} \end{pmatrix}$$

Using these, we get from (3.3.5)

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} X^{+\prime} \tilde{z}\right) = \underset{\sim}{0} \tag{3.3.8}$$

Now, note that the first column of X^+ is $(1,1,\ldots,1)'$. So,

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n}(1,1,\ldots,1)' \tilde{z} = \underset{\sim}{0} \tag{3.3.9}$$

Next, let us consider

$$\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} X_o^{+\prime} \tilde{z}\right)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$X_o^+ = \begin{pmatrix} 1 & x_{21} - \bar{x}_2 & \cdots & x_{m1} - \bar{x}_m \\ 1 & x_{22} - \bar{x}_2 & \cdots & x_{m2} - \bar{x}_m \\ \vdots & \vdots & & \vdots \\ 1 & x_{2n} - \bar{x}_2 & \cdots & x_{mn} - \bar{x}_m \end{pmatrix}$$

$$= (\, 1 \quad x_2 - \bar{x}_2 \quad \cdots \quad x_m - \bar{x}_m) \qquad \text{(say)}$$

Since the first column of X_o^+ is $(1,1,\ldots,1)'$, the first element in the vector $\frac{1}{n}(X_o^{+\prime} \tilde{z})$ has zero probability limit by (3.3.8). It suffices to show that the second element in the vector $\frac{1}{n}(X_o^{+\prime} \tilde{z})$ converges to 0. This element is given by

$$\frac{1}{n}(x_2 - \bar{x}_2)' \tilde{z} = \frac{1}{n} x'_2 \tilde{z} - \frac{1}{n} \bar{x}_2 (1,1,\ldots,1)' \tilde{z} \qquad (3.3.10)$$

Now, $\underset{n\to\infty}{\text{plim}} \frac{1}{n} x'_2 \tilde{z} = \underset{\sim}{0}$ (by (3.3.8)

and since, $\underset{n\to\infty}{\text{plim}} \frac{1}{n}(1,1,1,\ldots,1)' \tilde{z} = \underset{\sim}{0}$,

we have

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} (x_2 - \bar{x}_2)' \tilde{z} = \underset{\sim}{0} \qquad (3.3.11)$$

So, $$\underset{n\to\infty}{\text{plim}} \hat{\beta}_o^+ = \lim_{n\to\infty} \beta_o^+ = \tilde{P}_\infty \beta \qquad (3.3.12)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Hence $\hat{\beta}_o^+$ is consistent but (generally) biased estimate of β_o^+.

$$D(\hat{\beta}_o^+) = E[\{\beta_o^+ - E(\beta_o^+)\}\{\beta_o^+ - E(\beta_o^+)\}']$$

$$= E[\{(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} z - E(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} z\}\{(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} z - E(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} z\}']$$

$$+ E\{(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime} \varepsilon\varepsilon' X_o^+ (X_o^{+\prime} X_o^+)^{-1}\}$$

$$= V^+ + E[E\{(X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} \varepsilon\varepsilon' X_o^+(X_o^{+\prime}X_o^+)^{-1} | X_o^+\}]$$

$$= V^+ + \sigma^2 E(X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} V X_o^+(X_o^{+\prime} X_o^+)^{-1} \qquad (3.3.13)$$

where V^+ is the dispersion matrix of

$$(X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} \tilde{z} \qquad (3.3.14)$$

Let us consider the term

$$\sigma^2(X_o^{+\prime} X_o^+)^{-1} X_o^{+\prime} V X_o^+(X_o^{+\prime} X_o^+)^{-1} \qquad (3.3.15)$$

Let $(X_o^{+\prime} X_o^+)^{-1} = \frac{1}{\Delta} (a_{ij})$ where $a_{1j} = a_{j1} = 0$

for $j = 2,3,\ldots,m$ and $\Delta = \mathrm{Det}(X_o^{+\prime} X_o^+)$.

Let $C_i = x_i - \bar{x}_i$ $\quad i = 2,3,\ldots,m$

and $C_1 = (1,1,1,\ldots,1)'$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The i-th diagonal term in the matrix (3.3.15) is

$$\frac{\sigma^2}{\Delta^2} \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ji} \; a_{ki} \; C_j' V \, C_k \qquad i = 1,2,\ldots,m \qquad (3.3.16)$$

So, from (3.3.15) the first diagonal element of (3.3.15) given by

$$E\left\{ \frac{\sigma^2}{\Delta^2} \; a_{11}^2 \; C_1' V \; C_1 \right\}$$

$$= E\left(\frac{a_{11} \; \sigma^2}{\Delta}\right) + 2E\left\{\frac{a_{11}^2}{\Delta^2} \sum_{j=1}^{n-1} \rho^j (n-j)\right\} \sigma^2 \qquad (3.3.17)$$

For $i = 2$, the expression in (3.3.16) becomes

$$\frac{\sigma^2}{\Delta^2} \Big(a_{22}^2 \, C_2' V \, C_2 + 2a_{22}a_{32}C_2'VC_3 + 2a_{22}a_{42}C_2'V \, C_4 + \ldots + 2a_{22}a_{m2}C_2'VC_m$$

$$+ a_{32}^2 C_3'VC_3 + 2a_{32}a_{42}C_3'VC_4 + \ldots\ldots + 2a_{32}a_{m2}C_3'V \, C_m$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ a_{m2}^2 \, C_m'V \, C_m \Big) \qquad (3.3.18)$$

in (3.3.17) the term independent of ρ is given by

$$\frac{\sigma^2}{\Delta^2} \left(a_{22}^2 C_2'VC_2 + a_{22}a_{32} \, C_2' \, V \, C_3 + \ldots + a_{22} \, a_{m2}C_2'VC_m\right)$$

$$= \frac{\sigma^2}{\Delta^2} \, a_{22}\Big[a_{22}\Big\{ \sum_{i=1}^{n} (x_{2i}-\bar{x}_2)^2 + a_{32} \sum_{i=1}^{n} (x_{2i}-\bar{x}_2)(x_{3i}-\bar{x}_3)$$

$$+ a_{42} \sum_{i=1}^{n} (x_{2i}-\bar{x}_2)(x_{4i}-\bar{x}_4) + \ldots + a_{m2} \sum_{i=1}^{n} (x_{2i}-\bar{x}_2)(x_{mi}-\bar{x}_m) \Big\}\Big]$$

$$= \frac{\sigma^2 a_{22} \Delta}{\Delta^2} = \frac{\sigma^2 a_{22}}{\Delta} \qquad (3.3.19)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Again, in (3.3.16), if we consider expressions of the type

$$(a_{32}\, a_{22}\, C_3' V C_2 + a_{32}^2\, C_3' V C_3 + a_{32} a_{42} C_3' V C_4 + \cdots + a_{32}\, a_{m2}\, C_3' V C_m),$$

it can be shown that here the term independent of ρ is zero. Proceeding in this manner, it can be shown that the second diagonal element in (3.3.15) is

$$\frac{a_{22}\sigma^2}{\Delta} + \frac{2\sigma^2}{\Delta^2}\left(\rho \sum_{i=1}^{n-1} a_i^{(2)} a_{i+1}^{(2)} + \rho^2 \sum_{i=1}^{n-2} a_i^{(2)} a_{i+2}^{(2)} + \cdots + \rho^{n-1} a_i^{(2)} a_n^{(2)}\right)$$

where (3.3.20)

$$a_i^{(2)} = a_{22}(x_{2i}-\bar{x}_2) + a_{23}(x_{3i}-\bar{x}_3) + \cdots + a_{2m}(x_{mi}-\bar{x}_m).$$

In this way we can obtain all the diagonal elements in the matrix (3.3.14). So, form (3.3.13),

$$V(\hat{\tilde{\beta}}_1^{+}) = (\text{the } (1,1) \text{ element in the matrix } V^{+}) + \mu_1\sigma^2$$

$$+ 2\sigma^2 E\left\{\frac{a_{11}^2}{\Delta^2} \sum_{j=1}^{n-j} \rho^j (n-j)\right\} \qquad (3.3.21)$$

where $\mu_1 = E(\frac{a_{11}}{\Delta})$

$$V(\hat{\beta}_2^{+}) = (\text{the } (2,2) \text{ element in the matrix } V^{+}) + \mu_2\sigma^2$$

$$+ 2\sigma^2(\rho q_1 + \rho^2 q_2 + \cdots + \rho^{n-1} q_{n-1}) \qquad (3.3.22)$$

where $\mu_2 = E(\frac{a_{22}}{\Delta})$ and $q_i = E\left(\frac{\sum_{j=1}^{n-i} a_j^{(2)} a_{j+i}^{(2)}}{\Delta^2}\right)$.

Now, the standard OLS formulae for $V(\hat{\tilde{\beta}}_1^{+})$ and $V(\hat{\beta}_2^{+})$ are

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$V_{OLS}(\hat{\tilde{\beta}}_1^{+}) = \mu_1 \sigma^2 \tag{3.3.23}$$

$$V_{OLS}(\hat{\beta}_2^{+}) = \mu_2 \sigma^2 \tag{3.3.24}$$

If in (3.3.21), $\rho > 0$, then $V(\hat{\tilde{\beta}}_1^{+})$ is greater that $V_{OLS}(\hat{\tilde{\beta}}_1^{+})$ given by (3.3.23). Similarly, if in (3.3.22), $q_1, q_2, \ldots, q_{n-1}$ are positive, then for $\rho > 0$ $V(\hat{\beta}_2^{+})$ given by (3.3.22) will be greater than $V_{OLS}(\hat{\beta}_2^{+})$ given by (3.3.24). This is true also $\hat{\beta}_3^{+}$, $\hat{\beta}_4^{+}$ etc.

Now, σ^2 is estimated by $\frac{e^{+\prime}e^{+}}{n-m}$ in OLS procedure, where

$$e^{+} = y - X_o^{+} \hat{\beta}_o^{+}$$

$$= [I - X_o^{+}(X_o^{+\prime}X_o^{+})^{-1} X_o^{+\prime}] (\tilde{z} + \varepsilon) \tag{3.3.25}$$

Under the assumption the z 's are homoscedastic, with variance $\sigma^2_{\tilde{z}}$,

$$E(e^{+\prime}e^{+}) = E(\tilde{z}^{\prime}\tilde{z} + \varepsilon^{\prime}\varepsilon) - E[(\tilde{z}+\varepsilon)^{\prime}X_o^{+}(X_o^{+\prime}X_o^{+})^{-1}X_o^{+\prime}(\tilde{z}+\varepsilon)]$$

$$= n(\sigma^2_{\tilde{z}} + \sigma^2) - E\ \text{trace}[(\tilde{z}+\varepsilon)^{\prime}X_o^{+}(X_o^{+\prime}X_o^{+})^{-1}X_o^{+\prime}(\tilde{z}+\varepsilon)]$$

$$= n(\sigma^2_{\tilde{z}} + \sigma^2) - \text{trace}\ E[(X_o^{+\prime}X_o^{+})^{-1}X_o^{+\prime}(\tilde{z}+\varepsilon)(\tilde{z}+\varepsilon)^{\prime}X_o^{+}] \tag{3.3.26}$$

$$\therefore E(e^{+\prime}e^{+}) = n(\sigma^2_{\tilde{z}} + \sigma^2) - \text{trace}\ W - \sigma^2 E\ \text{trace}[(X_o^{+\prime}X_o^{+})^{-1}X_o^{+\prime}V X_o^{+}] \tag{3.3.27}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $W = E[(X_o^{+\prime} X_o^{+})^{-1} X_o^{+\prime} \tilde{z} \tilde{z}' X_o^{+}]$ (3.3.28)

Let us consider the last term in (3.3.26). It can be shown that

$$\sigma^2 \text{ trace}[(X_o^{+\prime} X_o^{+})^{-1} X_o^{+\prime} V X_o^{+}]$$

$$= \frac{\sigma^2}{\Delta} \left(\sum_{i=1}^{m} a_{1i} C_i' V C_1 + \sum_{i=1}^{m} a_{2i} C_i' V C_2 + \cdots + \sum_{i=1}^{m} a_{mi} C_i' V C_m \right)$$

$$= m\sigma^2 + \frac{2\sigma^2}{\Delta} [a_{11} \sum_{j=1}^{n-1} \rho^j (n-j) + a_{22} \sum_{j=1}^{n-1} \rho^j \sum_{i=1}^{n-j} (x_{2i} - \bar{x}_2)(x_{2,i+j} - \bar{x}_2)$$

$$+ a_{23} \sum_{j=1}^{n-1} \rho^j \left\{ \sum_{i=1}^{n-j} (x_{2i} - \bar{x}_2)(x_{3,i+j} - \bar{x}_3) \right.$$

$$\left. + \sum_{i=1}^{n-j} (x_{3i} - \bar{x}_3)(x_{2,i+j} - \bar{x}) \right\} + \cdots$$

$$+ a_{mm} \sum_{j=1}^{n-1} \rho^j \sum_{i=1}^{n-j} \rho^j \sum_{i=1}^{1-j} (x_{mi} - \bar{x}_m)(x_{m,i+j} - \bar{x}_m)]$$

$$= m\sigma^2 + S_o \quad (3.3.29)$$

So, from (3.3.28),

$$E(e^{+\prime} e^{+}) = (n-m)\sigma^2 + n\sigma_z^2 - E(S_o) \text{ trace } W \quad (3.3.30)$$

If $\rho = 0$, $E(S_o)$ vanishes. The term $n\sigma_z^2$ - trace W may be taken as the effect of omission on the residual sum of squares. Whether $\frac{e^{+\prime}e^{+}}{n-m}$ is an overestimate, or underestimate

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

or unbiased estimate of σ^2 will depend on the relative magnitudes of $n\sigma_2^2$ - trace W and $E(\delta_o)$. So, the standard OLS formulae for estimating $V(\hat{\tilde{\beta}}_1^+)$, $V(\hat{\beta}_i^+)$, $i = 2,3,\ldots,m$ can be in error in two respects. The estimate of σ^2 may not be unbiased and at the sametime, many terms in the expression for $V(\hat{\tilde{\beta}}_1^+)$ or $V(\hat{\beta}_i^+)$, $i = 2,3,\ldots,m$ may be neglected. This result is, however, not surprising. A similar result holds for the OLS estimation when the disturbances in a correctly specified equation are autocorrelated, and there is no omission of regressors.

Let us now consider the special case when $E(\tilde{z}|X^+) = \underset{\sim}{0}$

$$\therefore \quad E(\tilde{z}|X^+) = \underset{\sim}{0}, \quad E(\tilde{z}) = \underset{\sim}{0} \qquad (3.3.31)$$

We have mentioned in section 2 that this is possible when the regressions of X^- on X^+ are perfectly linear. Consider the following set of regression equations

$$X^- = X^+\delta + \vartheta \quad \text{given by (3.2.10)}$$

We have further assumed that ϑ is independent of X^+. So,

$$E(\vartheta|X^+) = E(\vartheta|X_o^+) = \underset{\sim}{0} \Rightarrow E(\vartheta) = \underset{\sim}{0} \qquad (3.3.32)$$

From (3.2.14) and (3.2.15), since $\delta = \bar{P}^{**}$ and $\tilde{z} = \vartheta\beta^{**}$,

$$E(\tilde{z}|X_o^+) = E(\vartheta|X_o^+)\beta^{**} = \underset{\sim}{0} \qquad (3.3.33)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $E(\tilde{z}\tilde{z}'|X_o^+) = E(\nu\beta^{**}\beta^{**'}\nu'|X_o^+) = E(\nu\beta^{**}\beta^{**'}\nu')$ (3.3.34)

Now, we may write

$$\nu = \begin{pmatrix} \nu_{m+1,1} & \nu_{m+2,1} & \cdots & \nu_{k,1} \\ \nu_{m+1,2} & \nu_{m+2,2} & \cdots & \nu_{k,2} \\ \vdots & & & \vdots \\ \nu_{m+1,n} & \nu_{m+2,n} & \cdots & \nu_{kn} \end{pmatrix} \quad (3.3.35)$$

Here we have (k-m) column vectors of disturbances. We get some interesting results if we assume these vectors to be mutually independent. This means any two regressors in X^- have zero partial correlation if influence of X^+ has been eliminated. We further assume that each vector of disturbances is homoscedastic and the i-th vector follows the Markov scheme given by

$$\nu_{m+i,t} = \rho_i \nu_{m+i,t-1} + \eta_{m+i,t}; \quad |\rho_i| < 1, \quad i = 1,2,\ldots,(k-m)$$
$$t = 2,3,\ldots,n \quad (3.3.36)$$

$\eta_{m+i,t}$ is the spherical disturbance term with

$E(\eta_{m+i,t}) = 0 \ \forall\ t$ and $E(\eta_{m+i,t}, \eta_{m+i,s}) = \sigma^2_{m+i,\eta}\,\delta_{ts}$,

δ_{ts} being the kronecker delta.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(\nu\beta^{**}\beta^{**\prime}\nu') = \begin{pmatrix} \sigma^{*2} & \sum_{i=1}^{k-m}\rho_i\sigma_i^{*2} & \sum_{i=1}^{k-m}\rho_i^2\sigma_i^{*2} & \cdots & \sum_{i=1}^{k-m}\rho_i^{n-1}\sigma_i^{*2} \\ \sum_{i=1}^{k-m}\rho_i\sigma_i^{*2} & \sigma^{*2} & \sum_{i=1}^{k-m}\rho_i\sigma_i^{*2} & \cdots & \sum_{i=1}^{k-m}\rho_i^{n-2}\sigma_i^{*2} \\ \vdots & & & & \\ \sum_{i=1}^{k-m}\rho_i^{n-1}\sigma_i^{*2} & \sum_{i=1}^{k-m}\rho_i^{n-2}\sigma_i^{*2} & \sum_{i=1}^{k-m}\rho_i^{n-3}\sigma_i^{*2} & \cdots & \sigma^{*2} \end{pmatrix}$$

$$= V_o \qquad (3.3.37)$$

where $\sigma_i^{*2} = \sigma^2_{m+i,\nu} \quad \beta^{**2}_{m+i}$

$$\sigma^{*2} = \sum_{i=1}^{k-m} \sigma_i^{*2} \qquad i = 1,2,\ldots (k-m) \qquad (3.3.38)$$

and $$\sigma^2_{m+i,\nu} = \frac{\sigma^2_{m+i,\eta}}{1-\rho_i^2}$$

Since here, $E(\tilde{z}|X^+) = \underset{\sim}{0}$, it can be shown easily that $\hat{\beta}^+$ is an unbiased estimator of $\beta^+ = \bar{\bar{I}}\,\beta$. Again, it can be proved easily that

$$D(\hat{\beta}_o^+) = E\{(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime}V_oX_o^+(X_o^{+\prime}X_o^+)^{-1}\}+\sigma^2E\{(X_o^{+\prime}X_o^+)^{-1} X_o^{+\prime}VX_o^+(X_o^{+\prime}X_o^+)^{-1}\} \qquad (3.3.39)$$

So, $$V(\hat{\tilde{\beta}}_1^+) = E(\frac{a_{11}^2}{\Delta^2} C_1'V_oC_1) + E(\frac{a_{11}^2\, C_1'V\,C_1}{\Delta^2})\,\sigma^2$$

$$= \mu_1(\sigma^{*2}+\sigma^2) + 2E[\frac{a_{11}^2}{\Delta^2} \sum_{j=1}^{n-j} \{\sum_{i=1}^{k-m}(\rho_i)^j\ \sigma_i^{*2}\}(n-j)]$$

$$+2E\,[\frac{a_{11}^2}{\Delta^2} \sum_{j=1}^{n-j} \rho^j(n-j)]\ \sigma^2 \qquad (3.3.40)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$V(\hat{\beta}_2^+) = \mu_2(\sigma^2+\sigma^{*2}) + 2\,[(\sum_{i=1}^{k-m} \rho_i \sigma_i^{*2} + \rho\,\sigma^2)\,q_1$$

$$+(\sum_{i=1}^{k-m} \rho_i^2\,\sigma_i^{*2} + \rho^2\sigma^2)q_2 + \dots + (\sum_{i=1}^{k-m} \rho_i^{n-1}\sigma_i^{*2} + \rho^{n-1}\sigma^2)q_{n-1}]$$

(3.3.41)

The usual OLS variances of $\hat{\beta}_1^+$ and $\hat{\beta}_i^+$ $i = 2,\dots,m$ are given by (3.3.23) and (3.3.24). As in the general case, here also σ^2 will be estimated by $\frac{e^{+\prime}e^+}{n-m}$. The expression for $E(e^{+\prime}e^+)$ will remain the same as in (3.3.27). But here trace W can be calculated and σ_z^2 will be replaced by σ^{*2}. It can be easily shown that

$$\text{trace } W = m\sigma^{*2} + \frac{2}{\Delta} E[a_{11}\{\sum_{j=1}^{n-1} (\sum_{i=1}^{k-m} \rho_i^j\,\sigma_i^{*2})\,(n-j)\}$$

$$+ a_{22} \sum_{j=1}^{n-1} (\sum_{i=1}^{k-m} \rho_i^j\,\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{2t}-\bar{x}_2)(x_{2,t+j}-\bar{x}_2)\}$$

$$+ a_{23} \sum_{j=1}^{n-1} (\sum_{i=1}^{k-m} \rho_i^j\,\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{2t}-\bar{x}_2)(x_{3,t+j}-\bar{x}_3)$$

$$+ \sum_{t=1}^{n-j}(x_{3t}-\bar{x}_3)\,(x_{2,t+j}-\bar{x}_2)\}$$

$$+ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$+ a_{mm} \sum_{j=1}^{n-1} (\sum_{i=1}^{k-m} \rho_i^j\,\sigma_i^{*2})\{\sum_{t=1}^{n-j}(x_{mt}-\bar{x}_m)(x_{m,t+j}-\bar{x}_m)\}]$$

$$= m\sigma^{*2} + E(S_o^*) \quad \text{(say)}$$

(3.3.42)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, from (3.3.25), we have

$$E(e^{+\prime}e^{+}) = (n-m)\ (\sigma^2 + \sigma^{*2}) - E(S_o+S_0^*) \qquad (3.3.43)$$

Here the term $(n-m)\sigma^{*2} - E(S_o^*)$ may be taken as the effects of omission of regressors and the term $E(S_o)$ is the effect of autocorrelation on the residual sum of squares. From the above results we find that conclusions regarding bias in estimating the sampling variances of $\hat{\tilde{\beta}}_1^+$ and $\hat{\beta}_i^+$'s $i = 2,\ldots,m$ by the OLS method are similar to those in the general case.

Let us now consider the following subcases.

Case 1. The regressors are stochastic and ε's spherical. Here all the results can be obtained putting $\rho = 0$ in the results of this section. So, in this case,

(i) The OLS procedure will give biased estimator of β^+; but if $E(\tilde{z}\,|X^+) = \underset{\sim}{0}$, the OLS estimator is unbiased.

(ii) Putting $\rho = 0$ in (3.3.20) and (3.3.21) we get

$$V(\hat{\tilde{\beta}}_1^+) \approx [\text{the } (1.1) \text{ element of } V^+ + \mu_1\sigma^2] \geq \mu_1\sigma^2 = V_{OLS}(\hat{\beta}_1^+) \qquad (3.3.44)$$

and

$$V(\hat{\beta}_i^+) = [\text{the } (i,i) \text{ element of } V^+ + \mu_i\sigma^2] \geq \mu_i\sigma^2 = V_{OLS}(\hat{\beta}_i^+), \quad i = 2,3,\ldots,m \qquad (3.3.45)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Also, from (3.3.29), putting $\rho = 0$, we have

$$E(e^{+\prime}e^{+}) = (n-m)\sigma^2 + E[\{I-X_o^+(X_o^{+\prime}X_o^+)^{-1}X_o^{+\prime}\} z]'[\{I-X_o^+(X_o^{+\prime}X_o^+)^{-1}X_o^{+\prime}\} z]$$

$$\geq (m-m)\sigma^2 \qquad (3.3.46)$$

So, if in $\mu_1\sigma^2$ we use $\frac{e^{+\prime}e^{+}}{n-m}$ as an estimate of σ^2, $V(\hat{\beta}_1^+)$, $V(\hat{\beta}_i^+)$, $i = 2,3,\ldots,m$ may be overestimated, underestimated or estimated unbiasedly depending upon the relative effects of the first term in (3.3.44) or (3.3.45) and of the second term in (3.3.46).

Next, let us derive the asymptotic variance of the OLS estimator of β^+ when the regressions of X^- on X^+ are strictly linear i.e.

$$X^- = X^+\delta + \nu$$

where
$$\delta = \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1,k-m} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2,k-m} \\ \vdots & \vdots & & \vdots \\ \delta_{m1} & \delta_{m2} & \cdots & \delta_{m,k-m} \end{pmatrix}$$
and ν has been defined in (3.2.11).

So,

$$x_{m+i,t} = x_{1t}\delta_{1i} + x_{2t}\delta_{2i} + \cdots + x_{mt}\delta_{mi} + \nu_{i,t}, \quad i = 1,2,\ldots(k-m)$$

$$\text{and} \quad t = 1,2,\ldots n \qquad (3.3.47)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The OLS estimator of β^+ in the equation (3.2.6) is

$$\hat{\beta}^+ = (X^{+\prime}X^+)^{-1} X^{+\prime}y$$

$$= \beta^* + (X^{+\prime}X^+)^{-1}(X^{+\prime}X^-)\beta^{**} + (X^{+\prime}X^+)^{-1} X^{+\prime}\varepsilon \qquad (3.3.48)$$

$$\therefore V(\hat{\beta}^+) = V\{(X^{+\prime}X^+)^{-1}(X^{+\prime}X^-)\}\beta^{**} + \sigma^2 E\{(X^{+\prime}X^+)^{-1}\} \qquad (3.3.49)$$

So, the asymptotic variance 1/ of $\hat{\beta}^+$ is given by

$$\therefore \lim_{n\to\infty} V(\hat{\beta}^+) = \lim_{n\to\infty} V\{(X^{+\prime}X^+)^{-1}(X^{+\prime}X^-)\}\beta^{**} + \sigma^2 \lim_{n\to\infty} E\{(X^{+\prime}X^+)^{-1}\}$$

or,

$$\bar{V}(\hat{\beta}^+) = \bar{V}\{(X^{+\prime}X^+)^{-1}(X^{+\prime}X^-)\}\beta^{**} + \sigma^2 \bar{E}\{(X^{+\prime}X^+)^{-1}\} \qquad (3.3.50)$$

Let $\hat{\delta}_{hi}$ be the OLS estimate of δ_{hi}, $h = 1,2,\ldots,m$, in equation (3.3.47). So, from (3.3.47) and (3.3.48), the j-th element in the vector

$$(X^{+\prime}X^+)^{-1}(X^{+\prime}X^-)\beta^{**} \quad \text{is} \quad \beta^{**\prime}\hat{\delta}_j$$

where $\hat{\delta}_j = (\delta_{j1}, \delta_{j2}, \ldots, \delta_{j,k-m})'$.

Now, $\delta_{j\lambda}$ is the j-th element in the vector

$$(X^{+\prime}X^+)^{-1} X^{+\prime} x_{m+\lambda}, \quad \lambda = 1,2,\ldots,k-m$$

$$= \delta^{(\lambda)} + (X^{+\prime}X^+)^{-1} X^{+\prime} \nu_\lambda \qquad (3.3.51)$$

1/ The existence of such asymptotic mean and variance of β^+ has been proved in Appendix 2C.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\delta^{(\ell)} = (\delta_{1\ell}, \delta_{2\ell}, \ldots, \delta_{m\ell})'$ and $\nu_\ell = (\nu_{\ell 1}, \ldots, \nu_{\ell m})'$.

Under the assumption that $\nu_{\ell i}$'s are homoscedastic and serially uncorrelated, $V(\hat{\delta}_{j\ell}) = V(\nu_\ell)$, the j-th element of $E(X^{+\prime}X^{+})^{-1}$ (since ν's are indepent of X^{+}'s)

$$\therefore \bar{V}(\hat{\delta}_{j\ell}) = \bar{V}\,\frac{(\nu_\ell)}{n}\,.\ \text{the j-th element of } \bar{E}\left(\frac{X^{+\prime}X^{+}}{n}\right)^{-1} \quad (3.3.52)$$

Now, it can be shown that

$$\bar{V}(\beta^{**\prime}\,\hat{\delta}_j) = \beta^{**\prime}\,E\left(\frac{\nu'\nu}{n}\right)\beta^{**}\left[\text{the j-th element of } \bar{E}\left(\frac{X^{+\prime}X^{+}}{n}\right)^{-1}\right] \quad (3.3.53)$$

From (3.3.6), we have

$$E\left(\frac{X'X}{n}\right) = \bar{E}\begin{pmatrix} \frac{X^{+\prime}X^{+}}{n} & \frac{X^{+\prime}X^{-}}{n} \\ \frac{X^{-\prime}X^{+}}{n} & \frac{X^{-\prime}X^{-}}{n} \end{pmatrix} = \Sigma_{XX} = \begin{pmatrix} \Sigma_{X^{+}X^{+}} & \Sigma_{X^{+}X^{-}} \\ \Sigma'_{X^{+}X^{-}} & \Sigma_{X^{+}X^{+}} \end{pmatrix}$$

$$\text{So, } \bar{E}\left(\frac{\nu'\nu}{n}\right) = (\Sigma_{X^{-}X^{-}} - \Sigma'_{X^{+}X^{-}}\,\Sigma^{-1}_{X^{+}X^{+}}\,\Sigma_{X^{+}X^{-}}) \quad (3.3.54)$$

So, from (3.3.52) and (3.3.53),

$$\bar{V}(\beta^{**\prime}\,\hat{\delta}_j) = \frac{1}{n}\,\beta^{**\prime}(\Sigma_{X^{-}X^{-}} - \Sigma'_{X^{+}X^{-}}\,\Sigma^{-1}_{X^{+}X^{+}}\,\Sigma_{X^{+}X^{-}})\,\beta^{**}.(\text{the j-th element in } \Sigma^{-1}_{X^{+}X^{+}}) \quad (3.3.55)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, from (3.3.   ),

$$\bar{V}(\hat{\beta}_j^{+}) = \frac{1}{n}\{\beta^{**\prime}(\Sigma_{X^-X^-} - \Sigma'_{X^+X^-}\Sigma^{-1}_{X^+X^+}\Sigma_{X^+X^-})\beta^{**} + \sigma^2\}(\text{the j-th element in } \Sigma^{-1}_{X^+X^+}) \quad (3.3.56)$$

The usual OLS formula for $\bar{V}(\hat{\beta}_j^{+})$ is, however,

$$\bar{V}_{OLS}(\hat{\beta}_j^{+}) = \frac{\sigma^2}{n}(\text{the j-th element in } \Sigma^{-1}_{X^+X^+}) \quad (3.3.57)$$

and σ^2 is estimated by $\frac{e^{+\prime}e^{+}}{n - m}$. From (3.3.46),

$$e^{+} = y - X^{+}\hat{\beta}^{+}$$

$$= X\beta + \varepsilon - X^{+}\beta^{*} - X^{+}(X^{+\prime}X^{+})^{-1}(X^{+\prime}X^{-})\beta^{**} + \{I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}\varepsilon$$

$$= \{I - X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}X^{-}\beta^{**} + \{I - X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}\varepsilon \quad (3.3.58)$$

$$\therefore E(e^{+\prime}e^{+}) = E[\beta^{**\prime}X^{-\prime}\{I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}'\{I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}X^{-}\beta^{**}]$$

$$+ E[\varepsilon'\{I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}'\{I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime}\}\varepsilon]$$

$$= E[\beta^{**\prime}X^{-\prime}(I-X^{+}(X^{+\prime}X^{+})^{-1}X^{+\prime})X^{-}\beta^{**}] + (n-m)\sigma^2 \quad (3.3.59)$$

$$\therefore \lim_{n\to\infty} E\frac{(e^{+\prime}e^{+})}{n} = \beta^{**\prime}(\Sigma_{X^-X^-} - \Sigma_{X^-X^+}\Sigma^{-1}_{X^+X^+}\Sigma_{X^+X^-})\beta^{**} + \sigma^2 \quad (3.3.60)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $V(\hat{\beta}_j^+) = \frac{e^{+\prime}e^+}{n} \{\text{the j-th element of } (X^{+\prime}X^+)^{-1}\}$ (3.3.61)

will give assymptotically an unbiased estimate of $\bar{V}(\hat{\beta}_j^+)$ derived in (3.3.56).

<u>Case 2.</u> Regressors are nonstochastic and the disturbances autocorrelated (ε's follow the Markov scheme given by (3.2.1)). Here,

(i) the OLS estimator $\hat{\beta}^+$ gives an unbiased estimate of $\beta^+ = P\beta$ (defined in Section 2).

(ii) In the expression for $D(\hat{\beta}_o^+)$ (<u>vide</u> equation (3.3.10)), the first term will be absent and here we need not take expectation on X. So,

$$V(\hat{\tilde{\beta}}_1^+) = (\text{the 1,1) element of } V^+) + \frac{a_{11}}{\Delta}\sigma^2 \quad (3.3.62)$$

and $V(\hat{\beta}_i^+) = (\text{the (i,i) element of } V^+) + \frac{a_{ii}}{\Delta}\sigma^2$, $i=2,3,\ldots m$ (3.3.63)

Also,

$$V_{OLS}(\hat{\tilde{\beta}}_1^+) = \frac{a_{11}}{\Delta}\sigma^2 \quad (3.3.64)$$

and $V_{OLS}(\hat{\beta}_i^+) = \frac{a_{ii}}{\Delta}\sigma^2$, $i = 2,3,\ldots m$ (3.3.65)

Thus (3.3.64) and (3.3.65) neglect some of the terms in (3.3.62) and (3.3.63).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(iii) σ^2 will be estimated by $\frac{e^{+\prime}e^{+}}{n - m}$, and since $X^{+\prime}z = 0$, from (3.26) and (3.27), we have

$$E(e^{+\prime}e^{+}) = (n - m)\sigma^2 + z'z - S_o \tag{3.3.66}$$

Here $z'z$ is the effect of omission of regressors and S_ρ is the effect of autocorrelation among the disturbances. Hence the conclusions regarding the biasedness of OLS estimates of the sampling variances are similar to those in the general case of stochastic regressor with autocorrelated disturbances.

3.4 Performance of the D-W test of randomness

Let us consider the model (3.3.1). The D-W statistic (1950, 1951) is

$$d = \frac{\sum_{t=2}^{n} (e_t^+ - e_{t-1}^+)^2}{\sum_{t=1}^{n} e_t^{+2}} \tag{3.4.1}$$

where $e^+ = y - X_o^+ \hat{\beta}_o^+$

$= y - X^+ \hat{\beta}^+$

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}^+ = \lim_{n\to\infty} \bar{P} \beta = \bar{P}_\infty \beta .$$

As $n \to \infty$, e^+ converges in distribution to

$$(X - X^+ \bar{P}_\infty) \beta + \varepsilon = z_\infty + \varepsilon \tag{3.4.2}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\tilde{z}_{\infty} = (\tilde{z}_{\infty,1}, \tilde{z}_{\infty,2} \dots \tilde{z}_{\infty,n})'$.

Hence,

$$d \xrightarrow{L} 2 \left\{ 1 - \frac{\sum_{t=2}^{n} (\tilde{z}_{\infty,t} + \varepsilon_t)(\tilde{z}_{\infty,t-1} + \varepsilon_{t-1})}{\sum_{t=1}^{n} (\tilde{z}_{\infty,t} + \varepsilon_t)^2} \right\}$$

$$\therefore \underset{n\to\infty}{\text{plim}}\, d = 2 \left\{ 1 - \frac{\underset{n\to\infty}{\text{plim}} \frac{1}{n} \left(\sum_{t=2}^{n} \tilde{z}_{\infty,t} \tilde{z}_{\infty,t-1} + \sum_{t=2}^{n} \varepsilon_t \varepsilon_{t-1} \right)}{\underset{n\to\infty}{\text{plim}} \frac{1}{n} \left(\sum_{t=1}^{n} \tilde{z}^2_{\infty,t} + \sum_{t=1}^{n} \varepsilon_t^2 \right)} \right\} \quad (3.4.3)$$

Now, let us assume that $\tilde{z}$'s follow a covariance stationary process (*vide* Dhrymes, 1970, pp 385-386).

So,

$$E(\tilde{z}_{\infty,t}) = \mu_t, \qquad t = 1,2,\dots,n \quad (3.4.4)$$

[since, $\lim_{n\to\infty} E\{X'(X - X^{+} \bar{P}_{\infty})\beta\} = \underset{\sim}{0}$,

$$\lim_{n\to\infty} E\left(\sum_{t=1}^{\infty} \tilde{z}_{\infty,t}\right) = 0 .$$

So, $E(\tilde{z}_{\infty,t})$ cannot be same for all t.]

$$E(\tilde{z}_{\infty,t} - \mu_t)^2 = \sigma^2_{\tilde{z}} \quad \forall\, t \quad (3.4.5)$$

and $$E(\tilde{z}_{\infty,t} - \mu_t)(\tilde{z}_{\infty,t-\tau} - \mu_{t-\tau}) = c_{\tau} \quad (3.4.6)$$

Let $c_1 = \rho_{\tilde{z}}\, \sigma^2_{\tilde{z}}$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\rho_{\tilde{z}}$ is the first order serial correlation coefficient of $\tilde{z}$.

So, it can be shown easily that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} \sum_{t=2}^{n} \tilde{z}_{\infty,t}\, \tilde{z}_{\infty,t-1} = \rho_{\tilde{z}}\, \sigma_{\tilde{z}}^{2} + \lim_{n\to\infty} \frac{1}{n} \sum_{t=2}^{n} \mu_t \mu_{t-1} \text{ (provided the limit exists)} \quad (3.4.7)$$

and

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} \sum_{t=1}^{n} \tilde{z}_{\infty,t}^{2} = \sigma_{\tilde{z}}^{2} + \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mu_t^{2} \text{ (provided the limit exists)} \quad (3.4.8)$$

provided

(i) the dependence between the distant values of $\tilde{z}_{\infty}$'s wears off rapidly as the distance increases,

(ii) $\tilde{z}_{\infty}$'s have a finite fourth order moment,

and (iii) $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mu_t^{4}$, exist.

So, under the assumption that u_t's in (3.2.1) have a finite fourth order moment [vide Goldberger, 1964, pp 143-155], it can be shown that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}}\ d = 2\left\{1 - \frac{\rho_{\tilde{z}}\sigma^2_{\tilde{z}} + \rho\,\sigma^2 + \lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{t=2}^{n}\mu_t\,\mu_{t-1}}{\sigma^2_{\tilde{z}} + \sigma^2 + \lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{t=1}^{n}\mu_t^2}\right\}$$

$$= 2(1-\rho^*) \quad \text{(say)} \qquad (3.4.9)$$

In the special case, where $E(\tilde{z}\mid X^+) = \underset{\sim}{0}$, $\mu_t = 0, \forall\, t$ (3.4.10)

$$\therefore \quad \underset{n\to\infty}{\text{plim}}\ d = 2\left\{1 - \frac{\rho_{\tilde{z}}\sigma^2_{\tilde{z}} + \rho\,\sigma^2}{\sigma^2_{\tilde{z}} + \sigma^2}\right\} \qquad (3.4.11)$$

From the above it is clear that for $\rho > 0$, $\underset{n\to\infty}{\text{plim}}\ d$ is likely to be less than 2 unless in large samples,

$$\frac{\rho_{\tilde{z}}\,\sigma^2_{\tilde{z}}\ \frac{1}{n}\sum\limits_{t=1}^{n}\mu_t^2}{\sigma^2_{\tilde{z}} + \frac{1}{n}\sum\limits_{t=1}^{n}\mu_t^2} \text{ in (3.4.10) or } \rho_{\tilde{z}} \text{ in (3.4.11)}$$

is negative and sufficiently large in magnitude.

<u>Case 2.</u> Regressors are stochastic and ε's spherical. Here, all the results can be obtained by putting $\rho \equiv 0$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Case 3. Regressors are nonstochastic while ε's follow the Markov scheme given by (3.2.1). Here e^+ converges in distribution to

$$(X - X^+ P_\infty)\beta + \varepsilon ,$$

$$= z_\infty + \varepsilon \qquad (3.4.12)$$

where $z_\infty = (z_{\infty,1},\ z_{\infty,2} \dots z_{\infty,n})$ and P_∞ has been defined in (2.3.5) of Chapter 2.

Under the assumption that u_t's in (3.2.1) have a finite fourth order moment, (vide Goldberger 1964, pp 143-155) it can be shown that

$$\operatorname*{plim}_{n\to\infty} d = 2\left(1 - \frac{\rho_z\sigma_o^2 + \rho\sigma^2}{\sigma_o^2 + \sigma^2}\right) = 2(1 - \rho^*)$$

where $\rho_z = \lim\limits_{n\to\infty} \dfrac{\sum\limits_{t=2}^{n} z_{\infty,t}\, z_{\infty,t-1}}{\sum\limits_{t=1}^{n} z_{\infty,t}^2}$

and $\lim\limits_{n\to\infty} \dfrac{1}{n} \sum\limits_{t=1}^{n} z_{\infty,t}^2 = \sigma_o^2 \qquad (3.4.13)$

From (3.4.13) it is clear that for $\rho > 0$, $\operatorname*{plim}_{n\to\infty} d$ is likely to be less than 2, unless ρ_z is negative and sufficiently large numerically.

In all the above cases, d is likely to be significantly below 2 in large samples more often than the stated risk of first kind of error (.05, or .01).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

3.5. Performance of alternative methods of estimation

Suppose that d has come out significantly less than 2. In this case the usual procedure is to re-estimate the regression coefficient β^+ under the assumption that the disturbance in the model (3.2.6) follow the Markov scheme

$$\varepsilon_t^+ = \rho^* \varepsilon_{t-1}^+ + u_t \qquad (3.5.1)$$

where $|\rho^*| < 1$, $E(u_t) = 0 \ \forall \ t$

and $\text{cov}(u_t, u_s) = \sigma_u^2 \delta_{ts}$, δ_{ts} being Kronecker delta.

Here the symbol ρ^* has been used in anticipation of subsequent result. The expression for ρ^* has been given in (3.4.9). Our aim is to examine the consistency of the estimate of β^+ given by the following methods (which require the estimation of ρ^*) :

(i) Cochrane-Orcutt two-step procedure (1949).

(ii) Prais-Winsten method (vide Rao, 1968)

(iii) Durbin two-step procedure (Durbin 1960)

5.1 Cochrane-Orcutt two step method. Here the first step is to estimate β^+ by OLS method from the regression equation (3.2.6). Then ρ^* is consistently estimated by [vide (3.4.9)]

$$\hat{\rho}^* = \frac{\sum_{t=2}^{n} e_t^+ \; e_{t-1}^+}{\sum_{t=1}^{n} e_t^{+2}} \qquad (3.5.2)$$

From section 4 it is clear that

$$\underset{n \to \infty}{\text{plim}} \ \hat{\rho}^* = \rho^* \qquad (3.5.3)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The next step is to fit the equation

$$y_t - \hat{\rho}^* y_{t-1} = \beta_1^+(1-\hat{\rho}^*)+\beta_2^+(x_{2t}-\hat{\rho}^* x_{2,t-1})+\dots+\beta_m^+(X_{mt}-\hat{\rho}^* x_{m,t-1})$$
$$+ (\varepsilon_t - \hat{\rho}^* \varepsilon_{t-1}) \qquad (3.5.4)$$

by OLS method. This gives

$$\hat{\beta}_{co}^+ = (X^{+\prime}_{\hat{\rho}^*} X^+_{\hat{\rho}^*})^{-1} X^{+\prime}_{\hat{\rho}^*} y_{\hat{\rho}^*}$$

$$= (X^{+\prime}_{\hat{\rho}^*} X^+_{\hat{\rho}^*})^{-1} X^{+\prime}_{\hat{\rho}^*} (X^+_{\hat{\rho}^*}\beta^+ + \varepsilon^+_{\hat{\rho}^*})$$

$$= \beta^+ + (X^{+\prime}_{\hat{\rho}^*} X^+_{\hat{\rho}^*})^{-1}(X^{+\prime}_{\hat{\rho}^*} \varepsilon^+_{\hat{\rho}^*}) \qquad (3.5.5)$$

where
$$X^+_{\hat{\rho}^*} = \begin{pmatrix} 1-\hat{\rho}^* & X_{22}-\hat{\rho}^* X_{21} & \cdots & X_{m2}-\hat{\rho}^* X_{m1} \\ 1-\hat{\rho}^* & X_{23}-\hat{\rho}^* X_{22} & \cdots & X_{m3}-\hat{\rho}^* X_{m2} \\ \vdots & \vdots & & \vdots \\ 1-\hat{\rho}^* & X_{2n}-\hat{\rho}^* X_{2,m-1} & \cdots & X_{mn}-\hat{\rho}^* X_{m,n-1} \end{pmatrix}$$

$$y_{\hat{\rho}^*} = [(y_2-\hat{\rho}^* y_1), (y_3-\hat{\rho}^* y_2) \; \dots \; (y_n-\hat{\rho}^* y_{n-1})]'$$

and $\varepsilon^+_{\hat{\rho}^*} = [(\varepsilon_2^+ - \hat{\rho}^* \varepsilon_1^+), (\varepsilon_3^+ - \hat{\rho}^*\varepsilon_2^+), \dots, (\varepsilon_n^+ - \hat{\rho}^* \varepsilon_{n-1}^+)]'$

Let us now define as $(n-1) \times n$ matrix $T_{\hat{\rho}^*}$ as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$T_{\hat{\rho}^*} = \begin{pmatrix} -\hat{\rho}^* & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\hat{\rho}^* & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\hat{\rho}^* & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\hat{\rho}^* & 1 \end{pmatrix} \qquad (3.5.6)$$

and let $(T'_{\hat{\rho}^*} \; T_{\hat{\rho}^*}) = W^{-1}_{\hat{\rho}^*}$ (say) (3.5.7)

$$\therefore \hat{\beta}^+_{co} = \beta^+ + (X^{+\prime} W^{-1}_{\hat{\rho}^*} X^+)^{-1} (X^{+\prime} W^{-1}_{\hat{\rho}^*} \varepsilon^+)$$

$$= \beta^+ + (X^{+\prime} W^{-1}_{\hat{\rho}^*} X^+)^{-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} [(X - X^+ \bar{P}) \beta + \varepsilon]$$

$$= (\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} X^{+\prime})^{-1} (\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} X) \beta + (\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} X^+)^{-1}$$

$$\times (\frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} \varepsilon) \qquad (3.5.8)$$

Since, $\underset{n\to\infty}{\text{plim}} \hat{\rho}^* = \rho^*$, $T_{\hat{\rho}^*}$ is a consistent estimator of T_{ρ^*} obtained by putting ρ^* for $\hat{\rho}^*$ in $T_{\hat{\rho}^*}$. So, $W^{-1}_{\hat{\rho}^*}$ is a consistent estimator of $W^{-1}_{\rho^*}$.

Let us assume that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} X) = C_{\rho^*} \quad \text{(exists)} \qquad (3.5.9)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} X^{+}) = Q_{\rho^*} \quad \text{(is nonsingular)} \tag{3.5.10}$$

$$\underset{n\to\infty}{\lim} E\left\{\frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} V W^{-1}_{\rho^*} X^{+})\right\} = \Omega_1 \quad \text{(is positive definite)} \tag{3.5.11}$$

Since $\hat{\rho}^*$ is a consistent estimator of ρ^*, from (3.5. 8) and (3.5. 9),

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} X) = \underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} X) \tag{3.5.12}$$

Also, under the assumption that

$$\underset{n\to\infty}{\lim} E\left(\frac{1}{n-1} \sum_{t=1}^{n-1} x^2_{ht}\right) \tag{3.5.13}$$

exists, (which, in fact, exists by virtue of (3.2.2)), it can be proved that (for proof, see Appendix 2 A),

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n-1}} X^{+\prime} (W^{-1}_{\hat{\rho}^*} - W^{-1}_{\rho^*}) \varepsilon = \underset{\sim}{0} \tag{3.5.14}$$

We further note that,

$$E \frac{1}{\sqrt{n-1}} (X^{+\prime} W^{-1}_{\rho^*} \varepsilon) = \underset{\sim}{0}$$

and

$$\underset{n\to\infty}{\lim} V \left\{\frac{1}{\sqrt{n-1}} (X^{+\prime} W^{-1}_{\rho^*} \varepsilon)\right\} = \sigma^2 \Omega_1$$

Thus, by Chebyshev's inequality

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} \varepsilon) = \underset{\sim}{0} \tag{3.5.15}$$

and hence, from (3.5.14),

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho^*} \varepsilon) = \underset{\sim}{0} \tag{3.5.16}$$

So, from (3.5.6),

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}^{+}_{co} = Q^{-1}_{\rho^*} C_{\rho^*} \beta = D_{\rho^*} \beta \neq \bar{P}_{\rho^*} \beta \text{ (in general)} \tag{3.5.17}$$

So, $\hat{\beta}^{+}_{co}$ is not, in general, a consistent estimator of β^{+}.

<u>Special case</u> : Let us consider the case where the regressions of X^{-} on X^{+} are strictly linear. So, $E(\tilde{z}|X^{+}) = E(\tilde{z}) = \underset{\sim}{0}$. From (3.5.8),

$$\hat{\beta}^{+}_{co} = \beta^{+} + (X^{+\prime} W^{-1}_{\hat{\rho}^*} X^{+})^{-1} X^{+\prime} W^{-1}_{\hat{\rho}^*} (\tilde{z} + \varepsilon) \tag{3.5.18}$$

In addition to the assumptions (3.5.9), (3.5.10) and (3.5.11), here we further assume that

$$\lim_{n\to\infty} E(\frac{1}{n-1} X^{+\prime} W^{-1}_{\rho^*} V_{\tilde{z}} W^{-1}_{\rho^*} X^{+}) = \Omega_1(\tilde{z}) \text{ (is positive definite)} \tag{3.5.19}$$

where $V_{\tilde{z}} = E(\tilde{z}\,\tilde{z}')$.

We note that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$X^{+\prime} W^{-1}_{\hat{\rho}*} \tilde{z}$$

$$= X^{+\prime} W^{-1}_{\hat{\rho}*} (X - X^{+} \bar{P})\beta$$

$$= (X^{+\prime} W^{-1}_{\hat{\rho}*} X - X^{+\prime} W^{-1}_{\hat{\rho}*} X^{+} \bar{P})\beta \qquad (3.5.20)$$

using (3.5.9), (3.5.10) and (3.5.12), it can be proved that

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\hat{\rho}*} \tilde{z}) = \operatorname*{plim}_{n\to\infty} \frac{1}{n-1} (X^{+\prime} W^{-1}_{\rho*} z) \qquad (3.5.21)$$

Again, $E(\frac{1}{n-1} X^{+\prime} W^{-1}_{\rho*} \tilde{z}) = \underset{\sim}{0}$

and $\lim_{n\to\infty} V(\frac{1}{n-1} X^{+\prime} W_{\rho*} \tilde{z})$

$$= \lim_{n\to\infty} \frac{1}{n-1} \Omega_{1}(\tilde{z}) = 0 \text{ (null matrix)}$$

So, by Chebyshev's inequality,

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\rho*} \tilde{z} = 0 \qquad (3.5.22)$$

and hence, from (3.5.19),

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}*} \tilde{z} = \underset{\sim}{0} \qquad (3.5.23)$$

Thus, from (3.5.18), it follows that

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}^{+}_{co} = \bar{P}_{\infty} \beta \qquad (3.5.24)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, in this case, $\hat{\beta}^{+}_{co}$ is a consistent estimator of β^{+}.

3.5.2. Prais- Winsten method

Under the assumption (3.5.1), the variance-covariance matrix of ε^{+} is given by

$$D(\varepsilon^{+}) = \frac{\sigma_u^2}{(1-\rho^{*2})} \begin{pmatrix} 1 & \rho^{*} & \rho^{*2} & \cdots & \rho^{*n-1} \\ \rho^{*} & 1 & \rho^{*} & \cdots & \rho^{*n-2} \\ \vdots & & & & \\ \rho^{*n-1} & \rho^{*n-2} & \rho^{*n-3} & \cdots & 1 \end{pmatrix}$$

$$= \frac{\sigma_u^2}{(1-\rho^{*})^2} V_{\rho^{*}} \quad \text{(say)} \qquad (3.5.25)$$

ρ^{*} is estimated by $\hat{\rho}^{*}$ given in (3.5.2) and $\hat{\rho}^{*}$ is a consistent estimator of ρ^{*}. So, $V_{\rho^{*}}$ is consistently estimated by $V_{\hat{\rho}^{*}}$ which can be obtained by putting $\hat{\rho}^{*}$ for ρ^{*} in $V_{\rho^{*}}$.

In Prais-Winstent method of estimation, β^{+} is estimated by

$$\hat{\beta}^{+}_{PW} = (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} X^{+})^{-1} (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} y)$$

$$= (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} X^{+})^{-1} (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} X)\beta + (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} X^{+})^{-1} (X^{+\prime} V^{-1}_{\hat{\rho}^{*}} \varepsilon) \qquad (3.5.26)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let us assume that 2/

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} (X^{+\prime} V^{-1}_{\rho^*} X) = C_{\rho^*} \quad \text{(exists)} \tag{3.5.27}$$

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} (X^{+\prime} V^{-1}_{\rho^*} X^{+}) = \tilde{Q}_{\rho^*} \quad \text{(nonsingular)} \tag{3.5.28}$$

$$\lim_{n\to\infty} E\left\{ \frac{1}{n}(X^{+\prime} V^{-1}_{\rho^*} V V^{-1}_{\rho^*} X^{+}) \right\} = \Omega_2 \quad \text{(exists and is positive definite)} \tag{3.5.29}$$

As in (3.5.12), here also, it can be proved from (3.5.27) and (3.5.28) that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} (X^{+\prime} V^{-1}_{\rho^*} X) = \underset{n\to\infty}{\text{plim}} \frac{1}{n} (X^{+\prime} V^{-1}_{\rho^*} X) \tag{3.5.30}$$

Also, under the assumption (3.5.12), it can be proved (See Appendix 2 A) that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} X^{+\prime} (V^{-1}_{\hat{\rho}^*} - V^{-1}_{\rho^*}) \varepsilon = \underset{\sim}{0} \tag{3.5.31}$$

and thus, as in (3.5.16), here also, it can be proved similarly that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} (X^{+\prime} V^{-1}_{\hat{\rho}^*} \varepsilon) = \underset{\sim}{0} \tag{3.5.32}$$

2/ As in Chapter 2, here also, the assumptions in (3.5.27), (3.5.28) and (3.5.29) are asymptotically same as the assumptions in (3.5.9), (3.5.10) and (3.5.11).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, from (3.5.26), it can be proved that

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^{+}_{PW} = \tilde{Q}^{-1}_{\rho^*}\ \tilde{C}_{\rho^*}\ \beta = \tilde{D}_{\rho^*}\ \beta \neq \bar{P}_{\infty}\ \beta \text{ (in general)} \tag{3.5.33}$$

So, $\hat{\beta}^{+}_{PW}$ is not, in general, a consistent estimator of β^{+}.

<u>Special case.</u> In the special case when the regressions of X^{-} on X^{+} are strictly linear, $E(\tilde{z}|X^{+}) = E(\tilde{z}) = \underset{\sim}{0}$. From (3.5.26),

$$\hat{\beta}^{+}_{PW} = \beta^{+} + (X^{+\prime}\ V^{-1}_{\hat{\rho}^*}\ X^{+})^{-1}\ X^{+\prime}\ V^{-1}_{\hat{\rho}^*}\ (\tilde{z} + \varepsilon) \tag{3.5.34}$$

In addition to the assumptions in (3.5.27), (3.5.28) and (3.5.29), we further assume that

$$\lim_{n\to\infty} E\left\{\frac{1}{n}(X^{+\prime} V^{-1}_{\rho^*}\ V_{\tilde{z}}\ V^{-1}_{\rho^*}\ X^{+})\right\} = \Omega_2(\tilde{z}) \text{ (exists and is positive definite)} \tag{3.5.35}$$

As in (3.5.21), here also, it can be proved from (3.5.30) that

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{n}(X^{+\prime} V^{-1}_{\hat{\rho}^*}\ \tilde{z}) = \underset{n\to\infty}{\text{plim}}\ \frac{1}{n}(X^{+\prime}\ V^{-1}_{\rho^*}\ \tilde{z}) \text{ (which exists by (3.5.27) and (3.5.28))} \tag{3.5.36}$$

We further note that

$$E\left(\frac{1}{n}\ X^{+\prime}\ V^{-1}_{\rho^*}\ \tilde{z}\right) = \underset{\sim}{0}$$

and $\lim_{n\to\infty} V\left\{\frac{1}{n}(X^{+\prime}\ V^{-1}_{\rho^*}\ \tilde{z})\right\} = \lim_{n\to\infty} \frac{1}{n}\ \Omega_2(\tilde{z}) = 0$ (null matrix).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, by Chebyshev's inequality,

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} X^{+\prime} V_{\rho^*}^{-1} \tilde{z} = \underset{\sim}{0} \qquad (3.5.37)$$

and hence, from (3.5.29),

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} X^{+\prime} V_{\rho^*}^{-1} \tilde{z} = \underset{\sim}{0} \qquad (3.5.38)$$

Thus, from (3.5.38),

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}^{+}_{PW} = \bar{P}_{\infty} \beta \qquad (3.5.39)$$

So, in this case, $\hat{\beta}^{+}_{PW}$ is a consistent estimator of β^{+}.

3.5.2 Durbin's two-step procedure

Here we consider the equation

$$y_t = (1-\rho^*)\beta_1^{+} + \rho^* y_{t-1} + \beta_2^{+} x_{2t} - \rho^*\beta_2^{+} x_{2,t-1} + \cdots$$
$$+ \beta_m^{+} x_{mt} - \rho^*\beta_m^{+} x_{m,t-1} + (\varepsilon_t^{+} - \rho^*\varepsilon_{t-1}^{+}) \quad t=2,3,\ldots,n$$

(3.5.40)

In matrix notation (3.5.40) can be written as

$$\bar{y} = \bar{X}^{+} \bar{\beta}^{+} + \varepsilon^{+}_{\rho^*} \qquad (3.5.41)$$

where $\bar{y} = (y_2, y_3, \ldots, y_n)'$.

$\bar{X}^{+}$ is a $(n-1) \times 2m$ matrix in which the t-th row is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$(1, y_t, x_{2,t+1}, x_{2t}, \ldots, x_{m,t+1}, x_{mt})$ $t = 1,2,\ldots,n-1.$

$\bar{\beta}^{+} = [(1 - \rho^*)\ \beta_1^{+}, \rho^*, \beta_2^{+}, - \rho^*\ \beta_2^{+}, \ldots, \beta_m^{+}, - \rho^* \beta_m^{+}]'$

and $\varepsilon_{\rho^*}^{+}$ is a column vector in which the t-th element is

$(\varepsilon_{t+1}^{+} - \rho^*\ \varepsilon_t^{+})$, $t = 1,2,\ldots,n-1.$

Using OLS procedure, we get

$$\bar{\beta}^{+} = (\bar{X}^{+\prime}\ \bar{X}^{+})^{-1}\ \bar{X}^{+\prime}\ \bar{y}$$

$$= \bar{\beta}^{+} + (\bar{X}^{+\prime}\ \bar{X}^{+})^{-1}\ \bar{X}^{+\prime}\ \varepsilon_{\rho^*}^{+} \qquad (3.5.42)$$

Now it can be easily proved that

$$\varepsilon_{\rho^*}^{+} = \bar{X}\ \bar{\beta} - X^{+}\ \bar{\beta}^{+} + \varepsilon_{\rho^*} \qquad (3.5.43)$$

where $\bar{X}$ is an $(n-1) \times 2k$ matrix in which the t-th row is

$(1, y_t, x_{2,t+1}, x_{2t}, \ldots, x_{m,t+1}, x_{m,t}, \ldots, x_{k,t+1}, x_{k,t})$

$t = 1,2, \ldots, n-1.$

$\bar{X} = (\bar{X}^{+}, \bar{X}^{-})$ where $\bar{X}^{-}$ is an $(n-1) \times 2(k-m)$ matrix containing the last $2(k-m)$ columns of $\bar{X}$.

$\bar{\beta} = [(1-\rho^*)\beta_1, \rho^*, \beta_2, - \rho^*\beta_2, \beta_3, - \rho^*\beta_3, \ldots, \beta_m, - \rho^*\beta_m, \beta_{m+1}$

$- \rho^*\ \beta_{m+1}, \ldots, \beta_k, - \rho^*\beta_k]'$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\varepsilon_{\rho^*} = [(\varepsilon_2 - \rho^*\varepsilon_1), (\varepsilon_3 - \rho^*\varepsilon_2), \ldots (\varepsilon_n - \rho^*\varepsilon_{n-1})]'$$

From (3.5.42) it can be shown that

$$\underset{n\to\infty}{\text{plim}}\ \hat{\bar{\beta}}^+ = \bar{\beta}^* + \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^+\right)^{-1} \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^-\right) \bar{\beta}^{**}$$

$$+ \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^+\right)^{-1} \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \varepsilon_{\rho^*}\right) \qquad (3.5.44)$$

where $\bar{\beta}^*$ contains the first 2m element and $\bar{\beta}^{**}$ the last 2(k-m) elements of $\bar{\beta}$.

Under the assumption (3.2.2) and also under the assumption that u_t's in (3.5.1) have finite fourth order moment, it can be proved (see Appendix 2B), that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (\bar{X}^{+\prime} \varepsilon_{\rho^*}) = [0, (\rho - \rho^*)\sigma^2, 0, 0, \ldots 0]' = L_1 \text{ (say)} \qquad (3.5.45)$$

Thus, it can be shown from (3.5.44) that

$$\underset{n\to\infty}{\text{plim}}\ \hat{\bar{\beta}}^+ = \bar{\beta}^* + N^{-1} C \bar{\beta}^{**} + N^{-1} L_1 \qquad (3.5.46)$$

where $\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^+\right) = N$ (say) assumed to be nonsingular

and $\underset{n\to\infty}{\text{plim}} \left(\frac{1}{n-1} \bar{X}^{+\prime} \bar{X}^-\right) = C$ (say) assumed to exist.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In Durbin's two-step procedure, the second element of $\hat{\beta}^+$ is taken as an estimate of ρ^*. Obviously, from (3.5.46),

$$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}^* \neq \rho^* \text{ necessarily}^{3/} \tag{3.5.47}$$

Let, $$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}^* = \rho^* + A_1 \text{(say)}, \quad A_1 \neq 0 \tag{3.5.48}$$

Durbin's second step consists in fitting the equation

$$y_t - \hat{\rho}^* y_{t-1} = \beta_1^+ (1 - \hat{\rho}^*) + \beta_2^+ (x_{2t} - \hat{\rho}^* x_{2,t-1}) + \dots$$
$$+ \beta_m^+ (x_{mt} - \hat{\rho}^* x_{m,t-1}) + (\varepsilon_t^+ - \hat{\rho}^* \varepsilon_{t-1}^+) \tag{3.5.49}$$

by OLS method. This gives

$$\hat{\beta}_D^+ = (X_{\hat{\rho}^*}^{+\prime} X_{\hat{\rho}^*}^+)^{-1} X_{\hat{\rho}^*}^{+\prime} y_{\hat{\rho}^*} \tag{3.5.50}$$

where the t-th row of $X_{\hat{\rho}^*}^+$ is given by

$$[(1 - \hat{\rho}^*), (x_{2,t+1} - \hat{\rho}^* x_{wt}), \dots, (x_{m,t+1} - \hat{\rho}^* x_{mt})]$$

$$t = 1, 2, \dots (n-1)$$

$\varepsilon_{\hat{\rho}^*}^+$ can be obtained from $\varepsilon_{\rho^*}^+$ by writing $\hat{\rho}^*$ for ρ^* and $y_{\hat{\rho}^*} = [(y_2 - \hat{\rho}^* y_1), (y_3 - \hat{\rho}^* y_2), \dots, (y_n - \hat{\rho}^* y_{n-1})]'$

3/ Obviously this $\hat{\rho}^*$ is different from $\hat{\rho}^*$ in (3.5.2).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

From (3.5.50), it can be shown that

$$\hat{\beta}^+_D = \beta^+ + (X^{+\prime} W^{-1}_{\hat{\rho}*} X^+)^{-1} (X^{+\prime} W^{-1}_{\hat{\rho}*} \varepsilon^+) \qquad (3.5.51)$$

where $W^{-1}_{\hat{\rho}*} = (T'_{\hat{\rho}*} T_{\hat{\rho}*})$.

Now, $W^{-1}_{\hat{\rho}*}$ is a consistent estimator of $W^{-1}_{\rho*+A_1}$ which can be obtained by putting $(\rho* + A_1)$ for $\rho*$ in $W^{-1}_{\rho*}$. Then, following the arguments as in Section (3.5.1), it can be shown from (3.5.51), that under certain assumptions similar to those stated in Section (3.5.1),

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_D = Q^{-1}_{(\rho*+A_1)} C_{(\rho*+A_1)} \beta$$

$$= D_{(\rho*+A_1)} \beta \neq \bar{P}_\infty \beta \quad \text{(necessarily)} \qquad (3.5.52)$$

where $C_{(\rho*+A_1)} = \underset{n\to\infty}{\text{plim}} (\frac{1}{n-1} X^{+\prime} W^{-1}_{(\rho*+A_1)} X)$ assumed to exist

and $Q_{(\rho*+A_1)} = \underset{n\to\infty}{\text{plim}} (\frac{1}{n-1} X^{+\prime} W^{-1}_{(\rho*+A_1)} X^+)$ assumed to be non-singular. So, β^+_D gives inconsistent estimate of β^+ in general.

Special case : Let us consider the case where the regressions of X^- on X^+ are strictly linear, i.e., $E(\tilde{Z} | X^+) = 0$.

From (3.5.42), it can be shown that OLS estimator of $\bar{\beta}$ is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\hat{\bar{\beta}}^{+} = \bar{\beta}^{+} + (\bar{X}^{+\prime} \bar{X}^{+})^{-1} \bar{X}^{+\prime} (\tilde{z}_{\rho *} + \varepsilon_{\rho *}) \qquad (3.5.53)$$

where, the t-th element of the vector $z_{\rho *}$ is

$$(\tilde{z}_{t+1} - \rho^{*} \tilde{z}_{t}), \quad t = 1,2,\ldots (n-1).$$

Since, in this case,

$$E(\tilde{z}/X^{+}) = E(\tilde{z}) = \underset{\sim}{0} ,$$

so,
$$E(\tilde{z}_{\rho *}/X^{+}) = E(\tilde{z}_{\rho *}) = \underset{\sim}{0} \qquad (3.5.54)$$

Since, in $\bar{X}^{+}$, all the columns excepting the second one contain elements of the matrix X^{+}, under the assumption (3.2.2), it can be proved that (see Appendix 2 C)

$$\underset{n\to\infty}{\text{plim}} \frac{X^{+\prime} \tilde{z}_{\rho *}}{n - 1} = [0, (\rho_{\tilde{z}} - \rho^{*})\sigma^{2}_{\tilde{z}} , 0, 0, \ldots 0]' = L_{2}(\text{say}) \qquad (3.5.55)$$

provided

(a) $\tilde{z}$'s are assumed to be stationary with finite fourth order moment,

(b) $\forall_{\tau}$, $\tau = 1,2,\ldots,t-1$, the covariance between $\tilde{z}_{t}$ and $\tilde{z}_{t+\tau}$ tends to 0 as $n \to \infty$. (<u>vide</u> Goldberger 1964).

Also, from (3.5.45),

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \frac{X^{+\prime}\ \varepsilon_{\rho}{}'}{n-1} = L_1 .$$

So, $$\underset{n\to\infty}{\text{plim}}\ \hat{\bar{\beta}}^{+} = \bar{\beta}^{+} + N^{-1}(L_1 + L_2) \tag{3.5.56}$$

Since the second element in the vectors $\bar{\beta}^{+}$ is ρ^{*},

$$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}^{*} \neq \rho^{*} \quad \text{(in general)} \tag{3.5.57}$$

Let $$\underset{n\to\infty}{\text{plim}}\ \hat{\rho}^{*} = \rho^{*} + A_2 \text{ (say)} \tag{3.5.58}$$

The next step is to fit the equation

$$y_t - \hat{\rho}^{*}y_{t-1} = \beta_1^{+}(1-\hat{\rho}^{*}) + \beta_2^{+}(x_{2t} - \hat{\rho}^{*}x_{2,t-1}) + \dots + \beta_m^{+}(x_{mt} - \hat{\rho}^{*}x_{m,t-1}) + (\varepsilon_t^{+} - \hat{\rho}^{*}\varepsilon_{t-1}) \tag{3.5.59}$$

by OLS method.

Now, as in the general case, here also, it can be shown easily that

$$\hat{\beta}_D^{+} = \beta^{+} + (X^{+\prime} W_{\hat{\rho}^*}^{-1} X^{+})^{-1} (X^{+\prime} W_{\hat{\rho}^*}^{-1} \varepsilon^{+}) \dots$$

$$= \beta^{+} + (X^{+\prime} W_{\hat{\rho}^*}^{-1} X^{+})^{-1} \{X^{+\prime} W_{\hat{\rho}^*}^{-1} (\tilde{z} + \varepsilon)\} \tag{3.5.60}$$

Following the same arguments as in Section (3.5.1) and making similar assumptions, it can be shown that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} X^{+\prime} W^{-1}_{\hat{\rho}*} \tilde{z} = \underset{n\to\infty}{\text{plim}} \frac{1}{n-1} X^{+\prime} W^{-1}_{\rho*} \tilde{z} = \underset{\sim}{0} \qquad (3.5.61)$$

∴ from (3.5.60)

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}^{+}_{D} = \bar{P}_{\infty} \beta \qquad (3.5.62)$$

So, here $\hat{\beta}^{+}_{D}$ gives consistent estimate of β^{+}.

In the case where X's are stochastic and ε's spherical, it can be shown easily from the results derived already that $\hat{\beta}^{+}_{co}$, $\hat{\beta}^{+}_{PW}$ and $\hat{\beta}^{+}_{D}$ give, in general, in consistent estimate of β^{+} (= $\bar{P}$ β), when $E(\tilde{z} \mid X^{+}) \neq \underset{\sim}{0}$; and give consistent estimates of $\beta^{+} = (\bar{\bar{P}} \beta)$ when the regressions of X^{-} on X^{+} are strictly linear (i.e., $E(\tilde{z} \mid X^{+}) = \underset{\sim}{0}$).

When X's are nonstochastic and ε's follow a Markov scheme, it can be proved easily that $\hat{\beta}^{+}_{co}$, $\hat{\beta}^{+}_{PW}$ and $\hat{\beta}^{+}_{D}$, give inconsistent estimates of β^{+}(= Pβ).

3.5.4 An overview of three alternative methods of estimation.

From (3.5.17) we find that in the general case,

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}^{+}_{co} = D_{\rho*} \beta \text{ .}$$

Again, from (3.5.33)

$$\underset{n\to\infty}{\text{plim}} \hat{\beta}^{+}_{PW} = D_{\rho*} \beta \text{ .}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

From footnote 2, it is obvious that

$$D_{\rho^*} = \tilde{D}_{\rho^*}$$

and thus, $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_{co} = \underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_{PW}$.

Also, from (3.5.52),

$$\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_D = D_{(\rho^*+A)}\ \beta$$

and $D_{(\rho^*+A)}\ \beta$ depends on $W^{-1}_{(\rho^*+A)} = T'_{(\rho^*+A)}\ T_{(\rho^*+A)}$.

So, $\underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_D \neq \underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_{co} = \underset{n\to\infty}{\text{plim}}\ \hat{\beta}^+_{PW}$.

Moreover, all these three methods of estimation give inconsistent estimates of $\bar{P}_\infty\ \beta$.

If, however, $E(\tilde{z} \mid X^*) = 0$, all the three methods give consistent estimate of β^+.

These conclusions remain unchanged in other two subcases mentioned in Section 1.

6. Conclusions

The main results of this chapter, unfortunately most of which are negative, may be briefly summarised as follows :

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(I)(a) When some regressors have been omitted from a regression equation, in which the regressors are stochastic and the disturbances follow a Markov scheme, the OLS procedure gives biased but consistent estimates of the regression coefficients (appropriately defined) of the regressors included in the misspecified equation. But the usual OLS formulae for estimating the sampling variances of these estimates may give underestimates, overestimates or unbiased estimates depending on the relative effects of (i) the omission of regressors, (ii) the autocorrelation among the disturbances of the true model and (iii) the autocorrelations of the included regressors.

(b) In the case of a model with nonstochastic regressors and the disturbances following a Markov scheme, the OLS procedure gives unbiased estimates of the regression coefficients (appropriately defined) in the misspecified model. But the performances of the OLS formulae for estimating the sampling variances of these estimated regression coefficients are the same as described above.

(c) If, on the other hand, the disturbances in the true model with stochastic regressors are spherical, the OLS procedure, again, gives biased but consistent estimates of the regression coefficients (appropriately defined) in the misspecified equation. But the usual OLS formulae for estimating the sampling variances of the estimated regression coefficients will give underestimates,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

overestimates or unbiased estimates depending on the effect of omission of regressors.

(d) In the special case where the regression equations of the omitted regressors on the included regressors are strictly linear, the OLS procedure gives unbiased estimates of the appropriately defined regression coefficients (in the misspecified equation), in all the cases discussed in this chapter.

(II) For the cases I(a), I(b) and I(c), the disturbances in the misspecified model will have, in general, nonzero means. However, in case (d) the disturbances will have zero means. But, the disturbances in all the cases [(a) to (d)] will not follow a Markov scheme. But these disturbances will be autocorrelated and if the first order autocorrelation coefficients of the disturbances (ε's) in the tru model and/or the autocorrelation coefficient of the vector ε^{+} (considered as a time series) in the misspecified model is positive, then in large samples the D-W statistic may be significantly less than 2 with higher probability than the chosen level of significance.

(III) For all the three cases [I(a) to I(c)], the usual methods of re-estimation (e.g., Cochrane-Orcutt two step procedure, Prais Winsten method and Durbin two-step procedure) fail to give consistent estimates of the re-defined regression coefficients of

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the included regressors in the misspecified model. Moreover, although the probability limits of $\hat{\beta}^{+}_{CO}$ and $\hat{\beta}^{+}_{PW}$ are approximately equal, these are different from the probability limit of $\hat{\beta}^{+}_{D}$. However, in the special case, mentioned in I(d) above, all the methods of re-estimation give consistent estimates of the re-defined regression coefficients in the misspecified equation.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 4

MSE CRITERION IN THE CONTEXT OF SPECIFICATION ERROR ANALYSIS WITH STOCHASTIC REGRESSORS *

4.1 Introduction

It is well known from Theil's (1967) specification analysis that when some regressors are omitted from a regression equation, the OLS procedure gives biased estimates of the regression coefficients in the misspecified equation. But these biased estimators have been used to generate optimal estimators in the case of different types of econometric problems, e.g., the problem of grouped observations dealt by Griliches (1970) and Haitovsky (1973).

Also, it is well known that imposition of exact linear restrictions, even if incorrect, may reduce the sampling variances of the restricted estimators. Feldstein (1973, 1974), using the concept of a trade-off between bias and sampling variance on mean square error (MSE) criterion, has obtained a class of pretest estimators to aid the problems of multicollinearity and errors-in-the variables in linear models. A familiar result in this field of work has been given by Toro-Vizcarrondo and Wallace (1968,1969). They have proved that in a two-regressor model, the MSE of the estimate of the parameter of interest (β_1) can be reduced by omitting another variable x_2 if and only if the absolute value of the "true t_{β_2} statistic" (i.e., the ratio of the regression coefficient associated with x_2 to the true standard error of its

---

* Some of the results of this chapter have been derived jointly with Lahiri (1976).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

estimate) is less than one.

Toro-Vizcarrondo and Wallace (1968, 1969), in particular, also suggested a way of investigating the problem of multicollinearity through the MSE criterion. Their test based on MSE allows one to examine the hypothesis about the effect of a particular restriction upon the comparative properties of the restricted and the OLS estimators. Very often, the restriction that is used to remove the problem of multicollinearity is the zero restriction, i.e., dropping a regressor or a set of regressors from the relationship. The zero restrictions is, however, the worst possible choice. Because, if one knew that a particular variable had zero coefficient, one would not have included that variable in the relationship. However, by imposing restrictions one may under certain conditions, obtain better estimators with smaller MSE. Toro-Vizcarrondo and Wallace in particular, have given a test procedure for accepting or rejecting these zero restrictions.

The aim of the present chapter is to compare the asymptotic variance and MSE of the omitted variable (OV)[1/] estimator with those

---

1/ When some of the regressors have been omitted from a regression equation, the OLS estimator of the regression coefficients in the misspecified model will be denoted by <u>OV estimator</u> in this chapter.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

of the OLS[2/] (f) estimator of regression coefficients for linear regression models with stochastic regressors. This we have done in section 2 for a two regressor model. There we have also obtained a condition under which the OV estimator has asymptotically smaller MSE than the OLS(f) estimator. This result has also been extended to the case of $k(\geq 2)$ regressors in section 3. In Section 4, the question of MSE dominance of the OV estimator over proxy variable estimator has been examined. This latter issue was originally discussed by Aigner (1974). Section 5 concludes the chapter with some general observations.

2. Comparison of asymptotic variances and MSE's of the OV estimator with those of the OLS estimator in the fully specified model.

2.1 Comparison of asymptotic variances

Let us consider the following two regressor linear model with stochastic regressors :

$$y = \beta x + \beta_2 x_2 + \varepsilon \qquad (4.2.1),$$

where ε is the spherical disturbance term with mean 0 and variance σ_{nn}. ε is distributed independently of x_1 and x_2. x_1 and x_2 are $(n \times 1)$ column vectors of stochastic regressors. Also suppose that each regressor has been measured from its respective mean. We assume that

2/ Here, the word 'OLS(f) estimator' has been used to denote the OLS estimator of regression coefficients in the fully specified model.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\bar{E}^{3/}\left(\frac{1}{n}\sum_{i=1}^{n} x_{1i}\varepsilon_i\right) = \bar{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_{2i}\ \varepsilon_i\right) = 0$$

$$\bar{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_{2i}^2\right) = \sigma_{22},\quad \bar{E}\left(\frac{1}{n}\sum_{i=1}^{n}(x_{2i}x_{1i})\right) = \sigma_{12},\quad \bar{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_{1i}^2\right) = \sigma_{11} \qquad (4.2.2)$$

and $\bar{E}\left(\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2\right) = \sigma^2$

Suppose, from the above model, the regressor z has been omitted. So, the misspecified model is

$$y = \beta_1^+ x_1 + \varepsilon^+ \qquad (4.2.3)$$

where β_1^+ is the redefined regression coefficient associated with the regressor x_1 in the misspecified model and ε^+ is the corresponding disturbance term. Let $\hat{\beta}_1^+$ be the OLS estimator of β_1^+ (i.e., $\hat{\beta}_1^+$ is the OV estimator). Then

$$\hat{\beta}_1^+ = (x_1' x_1)^{-1} x_1' y$$

$$= \beta + (x_1' x_1)^{-1} x_1'(x_2\ \beta_2 + \varepsilon) \qquad (4.2.4)$$

where $x = (x_{11}, x_{12}, \ldots, x_{1n})'$ and $z = (x_{21}, x_{22}, \ldots, x_{2n})'$.

3/ $\bar{E}(x^{(n)}) = \lim_{n\to\infty} E(x^{(n)})$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Thus

$$\bar{E}(\hat{\beta}_1^{+}) = \beta_1 + \bar{E}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} x_{1i}\, x_{2i}}{\frac{1}{n}\sum_{i=1}^{n} x_{1i}^2}\ \beta_2 \right)$$

$$= \beta_1 + \frac{\sigma_{12}}{\sigma_{11}}\beta_2$$

$$= \beta_1 + \delta_{21}\beta_2$$

$$= \beta_1^{+} \quad \text{(say)}\ ^{4/} \tag{4.2.5}$$

where δ_{21} is the population regression coefficient in the auxiliary regression of x_2 on x_1.

Now, asymptotic variance of β_1^{+} is given by

$$\bar{V}(\hat{\beta}_1^{+}) = n^{-1} \lim_{n\to\infty} E\{n(\beta_1^{+} - \beta_1^{+})(\beta_1^{+} - \beta_1^{+})'\}$$

$$= n^{-1}\bar{E}\{n(\beta_1^{+} - \beta_1^{+})(\beta_1^{+} - \beta_1^{+})'\}$$

$$= n^{-1}\bar{E}[\{n(x_1'x_1)^{-1}x_1'x_2 - \frac{\sigma_{12}}{\sigma_{11}}\beta_2\}\{(x_1'x_1)^{-1}x_1'x_2 - \frac{\sigma_{12}}{\sigma_{11}}\beta_2\}']$$

$$+ n^{-1}\bar{E}\{n(x_1'x_1)^{-1}x_1'\varepsilon\}\{(x_1'x_1)^{-1}x_1'\varepsilon\}'$$

$$= \beta_2^2\,\bar{V}\frac{\sum_{i=1}^{n} x_{1i}\,x_{2i}}{\sum_{i=1}^{n} x_{1i}^2} + n^{-1}\bar{E}[n(x_1'x_1)^{-1}x_1'\varepsilon\varepsilon'x_1(x_1'x_1)^{-1}$$

$$= \beta_2^2\,\bar{V}(\hat{\delta}_{21}) + n^{-1}\frac{\sigma^2}{\bar{E}(\frac{1}{n}\sum_{i=1}^{n} x_{1i}^2)}$$

$$= \beta_2^2\,\bar{V}(\delta_{21}) + n^{-1}\frac{\sigma^2}{\sigma_{11}} \tag{4.2.6}$$

4/ Existence of such asymptotic mean and variance of $\hat{\beta}_1^{+}$ has been proved in Appendix 2d.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Again, if the model is fully specified, the OLS(f) of β_1 is $\hat{\beta}_1$ = the first element in $(X'X)^{-1} X'y$; Here $\bar{E}\,\beta_1 = \beta_1$ and

$$\bar{V}(\beta_1) = \text{the (1,1) element of } n^{-1}\,\bar{E}[n(X'X)^{-1} X'\varepsilon\varepsilon'X(X'X)^{-1}] \quad (4.2.7)$$

where $X = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{pmatrix}$

From (4.2.5), we have

$$\bar{V}(\hat{\beta}_1) = \text{the (1,1) element of } \frac{\sigma^2}{n}\,\bar{E}\left(\frac{X'X}{n}\right)^{-1}$$

$$= \frac{\sigma^2}{n} \cdot \frac{\sigma_{22}}{\sigma_{11}\,\sigma_{22} - (\sigma_{12})^2}$$

$$\frac{\sigma^2}{n} \cdot \frac{1}{\sigma_{11}(1-\rho_{12}^2)} \quad (4.2.8)$$

where $\rho_{12}^2 = \dfrac{\sigma_{12}^2}{\sigma_{11}\,\sigma_{22}}$.

Comparing (4.2.6) with (4.2.8), we find that it is not always possible,to say that

$\bar{V}(\hat{\beta}_1^+)$ is greater than equal to or less than $\bar{V}(\hat{\beta}_1^-)$ in all the cases. So, we try to develop a condition under which $\bar{V}(\hat{\beta}_1^+)$ will be less than $\bar{V}(\hat{\beta}_1^-)$,

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\bar{V}(\hat{\beta}_1^+) < \bar{V}(\hat{\beta}_1)$ if and only if

$$\beta_2^2 \bar{V}(\hat{\delta}_{21}) + \frac{\sigma^2}{n\,\sigma_{11}} < \frac{\sigma^2}{n} \frac{\sigma_{12}}{\sigma_{12}\,\sigma_{22} - (\sigma_{12})^2}$$

or, $$\beta_2^2 \bar{V}(\hat{\delta}_{21}) < \frac{\sigma^2}{n} \frac{\sigma_{12}^2}{(\sigma_{11}\,\sigma_{22} - \sigma_{12}^2)\,\sigma_{11}}$$

$$= \frac{\sigma^2}{n} \cdot \frac{1}{\sigma_{22}(1 - \rho_{12}^2)} \cdot \delta_{12}^2$$

$$= \bar{V}(\hat{\beta}_2)\,\delta_{21}^2$$ (where β_2 is the OLS(f) of β_2 in (4.2.1))

or, $$\frac{\beta_2^2}{\bar{V}(\beta_2)} < \frac{\delta_{21}^2}{\bar{V}(\hat{\delta}_{21})}$$

or, $$\theta_{\beta_2}^2 < \theta_{\delta_{21}}^2 \text{ (say)} \qquad (4.2.9)$$

Remark 1. If y_1, x_1, x_2 jointly follow a trivariate normal distribution, then x_1 and x_2 will jointly follow a bivariate normal distribution (vide Anderson 1957). Then it can be shown that

$$\bar{V}(\hat{\delta}_{21}) = \frac{\sigma_{22}(1 - \rho_{12}^2)}{n\,\sigma_{11}} \qquad (4.2.10)$$

So, from (4.2.9), we get

$$\theta_{\beta_2}^2 < \frac{n\,\rho_{12}^2}{1 - \rho_{12}^2} \qquad (4.2.11)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

In the special case where x_1 and x_2 are uncorrelated in the limit,

$$\rho_{12} = 0 \text{ . So, } \bar{V}(\hat{\delta}_{21}) = \frac{\sigma_{22}}{n\,\sigma_{11}} \tag{4.2.12}$$

$$\text{Here} \quad \bar{V}(\hat{\beta}_1^+) = \frac{\beta_2^2\,\sigma_{22} + \sigma^2}{n\,\sigma_{11}} \tag{4.2.13}$$

$$\text{again, } \bar{V}(\hat{\beta}_1) = \frac{\sigma^2}{n\,\sigma_{11}} < \bar{V}(\hat{\beta}_1^+) \tag{4.2.14}$$

Remark 2. With nonstochastic regressors, however,

$$\bar{V}(\hat{\beta}_1^+) = \frac{\sigma^2}{n\,\sigma_{11}} \tag{4.2.15}$$

which is always less than the sampling variance of β^+ in the case of stochastic regressors given in (4.2.12).

4.2.2 Comparison of asymptotic MSE's

Our next attempt is to find out a condition under which asymptotically the OV estimator has a smaller MSE than the OLS estimator in the fully specified model.

$$\text{Now,} \quad MSE(\hat{\beta}_1^+) = E(\hat{\beta}_1^+ - \beta_1)^2$$

$$= E(\hat{\beta}_1^+ - \beta_1^+ - \beta_1^+ - \beta_1)^2 \tag{4.2.16}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, for large n,

$$\lim_{n\to\infty} MSE(\hat{\beta}_1^{+}) = \bar{E}(\hat{\beta}_1^{+} - \beta_1^{+})^2 + (\beta_1^{+} - \beta_1)^2 \qquad (4.2.17)$$

$$= \bar{V}(\hat{\beta}_1^{+}) + (\beta_1^{+} - \beta_1)^2$$

$$= \beta_2^2\,\bar{V}(\delta_{12}) + \frac{\sigma^2}{n\,\sigma_{11}} + \frac{\sigma_{12}^2}{\sigma_{11}^2}\beta_2^2 \quad \text{(from (4.2.5) and (4.2.6))} \qquad (4.2.18)$$

when the model is fully specified.

Thus the $\lim_{n\to\infty} MSE(\hat{\beta}_1) = \bar{V}(\hat{\beta}_1)$

$$= \frac{\sigma^2}{n\,\sigma_{11}(1-\rho_{12}^2)} \qquad (4.2.19)$$

Thus the OV estimator has asymptotically less MSE than the OLS estimator in the fully specified model if and only if

$$\frac{\sigma^2}{n\,\sigma_{11}(1-\rho_{12}^2)} > \beta_2^2\,\bar{V}(\hat{\delta}_{21}) + \frac{\sigma^2}{n\,\sigma_{11}} + \frac{\sigma_{12}^2}{\sigma_{11}^2}\beta_2^2$$

or, $$\frac{\sigma^2}{\beta_2^2\, n\,\sigma_{22}(1-\rho_{12}^2)} > \frac{\bar{V}(\hat{\delta}_{21})}{\delta_{21}^2} + 1$$

or, $$\frac{\bar{V}(\hat{\beta}_2)}{\beta_2^2} > \frac{\bar{V}(\hat{\delta}_{21})}{\delta_{21}^2} + 1$$

or, $$\frac{\beta^2}{\bar{V}(\hat{\beta}_2)} < \frac{1}{1+\dfrac{\bar{V}(\hat{\delta}_{21})}{\delta_{21}^2}}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

or, $$\theta^2_{\beta_2} < \frac{\theta^2_{\delta_{21}}}{\theta^2_{\delta_{21}}+1} \qquad (4.2.20)$$

Remark 1. With nonstochastic regressors, the condition becomes stringent. In that case, the condition becomes $\theta^2_{\beta_2} < 1$.

Remark 2. With $\sigma_{12} = 0$, the OV estimator will always have a larger MSE than the OLS estimator in the fully specified model.

Remark 3. If x, y and z follow a trivariate normal distribution,

$$\theta^2_{\delta_{21}} = \frac{n\,\rho^2_{12}}{1-\rho^2_{12}} \qquad (4.2.21)$$

So, $$\frac{\theta^2_{\delta_{21}}}{\theta^2_{\delta_{21}}+1} = \frac{n\,\rho^2_{12}}{(n-1)\,\rho^2_{12}+1} \qquad (4.2.22)$$

So, the condition (4.2.20) reduced to

$$\theta^2_{\beta_2} < \frac{n\,\rho^2_{12}}{(n-1)\,\rho^2_{12}+1} \qquad (4.2.23)$$

So, the pretest criterion is

$$t^2_{\beta_2} < \frac{n\,r^2_{12}}{(n-1)\,r^2_{12}+1} \qquad (4.2.24)$$ 5/

5/ Even when the joint distribution of y, x_1 and x_2 is not normal, one can get the conditions (4.2.11) and (4.2.23) if the regression of x_2 on x_1 is strictly linear.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $t^2_{\beta_2}$ is the square of the t statistic for testing $H_o : \beta_2=0$ associated with the regression equation of y on x and x_2 ; and r_{12} is the sample correlation between x_1 and x_2.

Thus, if (4.2.24) holds, the estimator to be used is $\hat{\beta}_1$+(OV estimator) ; otherwise, the estimator is $\hat{\beta}$ (OLS(f)).

4.3 Comparison of the asymptotic MSE of the OV estimator with that of OLS estimator in the fully specified model for the k(k > 2) regressor case

In this section we shall generalise the result in (4.2.20) to the k(k > 2) regressor case.

Let the regression equation be

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \qquad (4.3.1)$$

where all the variables are measured from respective means. Assumptions regarding ε are same as in Section 2. y_1 $x_1, x_2, \dots, x_k$ are $n \times 1$ column vectors of stochastic regressors. Suppose from the equation (4.3.1), x_k has been omitted. So, the misspecified model is

$$y = \beta^+_1 x_1 + \beta^+_2 x_2 + \dots + \beta^+_{k-1} x_{k-1} + \varepsilon^+$$

$$= \beta^+ X^+ + \varepsilon^+ \qquad (4.3.2)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $X^+ = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k-1,1} \\ x_{12} & x_{22} & \cdots & x_{k-1,2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{k-1,n} \end{pmatrix} = (x_1, x_2, \ldots, x_{k-1})$

and $\beta^+ = (\beta_1^+, \beta_2^+ \cdots \beta_{k-1}^+)'$ and $\varepsilon^+ = (\varepsilon_1^+, \varepsilon_2^+ \cdots \varepsilon_n^+)'$.

The OLS estimate of β^+ is

$$\hat{\beta}^+ = (X^{+\prime} X^+)^{-1} X^{+\prime} y$$

$$= \beta^* + (X^{+\prime} X^+)^{-1} X^{+\prime} x_k \beta_k + (X^{+\prime} X^+)^{-1} X^{+\prime} \varepsilon \qquad (4.3.3)$$

where $\beta^* = (\beta_1, \beta_2, \ldots, \beta_{k-1})'$.

$$\therefore \quad E(\hat{\beta}^+) = \beta^* + \beta_k E\{(X^{+\prime} X^+)^{-1} X^{+\prime} x_k\}$$

$$= \beta^* + \beta_k \delta \qquad (4.3.4)$$

where δ is the vector of expectations of sample regression coefficients of x_k on X^+.

Next, let us assume that

$$\lim_{n\to\infty} E\left(\frac{X^{+\prime} X^+}{n}\right) = \Sigma^+ \text{ (a positive definite matrix) exists} \qquad (4.3.5)$$

So, $$\lim_{n\to\infty} E(\hat{\beta}^+) = \beta^* + \beta_k \lim_{n\to\infty} E\{(X^{+\prime} X^+)^{-1} X^{+\prime} x_k\}$$

$$= \beta^* + \beta_k \delta_o \text{ (say)}. \qquad (4.3.6)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Also, it can be shown easily that

$$\bar{V}(\hat{\beta}^+) = \beta_k^2 \, \bar{V} \{(X^{+\prime} X^+)^{-1} X^{+\prime} x_k\}$$

$$+ n^{-1} \bar{E} \{n(X^{+\prime} X^+)^{-1} X^{+\prime} \varepsilon \varepsilon' X^+ (X^{+\prime} X^+)^{-1}\}$$

$$= \beta_k^2 \, \bar{V} \{(X^{+\prime} X^+)^{-1} X^{+\prime} x_k\} + \frac{\sigma^2}{n} \bar{E} \left(\frac{X^{+\prime} X^+}{n}\right)^{-1}$$

$$= \beta_k^2 \, \bar{V} \{(X^+ X'^+)^{-1} X^{+\prime} x_k\} + \frac{\sigma^2}{n} (\Sigma^+)^{-1} \qquad (4.3.7)$$

So, $\bar{V}(\hat{\beta}_1^+) = \beta_k^2$ [(1,1) element in $\bar{V} \{(X^{+\prime} X^+)^{-1} X^{+\prime} X_k\}$]

$+ \frac{\sigma^2}{n} \{$(1,1) element in $(\Sigma^+)^{-1}\}$ (4.3.8)

So, $\lim_{n\to\infty} MSE(\hat{\beta}_1^+) = \beta_k^2$ [(1,1) element in $\bar{V}\{(X^{+\prime} X^+)^{-1} X^{+\prime} x_k\}$]

$+ \frac{\sigma^2}{n} \{$(1,1) element in $(\Sigma^+)^{-1}\}$

$+ \beta_k^2$ (the 1st element in $\delta_o)^2$ (4.3.9)

Let $\hat{\beta}_1$ be the OLS(f) estimator of β_1 in (4.3.1).

So, $\hat{\beta}_1$ = (the first element of the vector $(X'X)^{-1} X'y$ (4.3.10)

It is also known that

$$\bar{E}(\hat{\beta}_1) = \beta_1$$

and $\bar{V}(\hat{\beta}_1) = \frac{\sigma^2}{n} \{$the(1,1) element in $\Sigma^{-1}\} = \lim_{n\to\infty} MSE(\hat{\beta}_1)$ (4.3.11)

where $\Sigma = \bar{E}(\frac{X'X}{n})$, and $X = (X^+, x_k)$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{Now,}\quad (X'X)^{-1} = \begin{pmatrix} x_1' x & x_1' X_2 \\ X_2' x & X_2' X_2 \end{pmatrix}$$

$$\text{where } X_2 = \begin{pmatrix} x_{21} & \cdots & x_{k1} \\ x_{22} & \cdots & x_{k2} \\ \vdots & & \vdots \\ x_{2n} & \cdots & x_{kn} \end{pmatrix} \text{ and } x_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}.$$

$$\text{So,}\quad \lim_{n\to\infty} MSE(\hat{\beta}_1) = \frac{\sigma^2}{n}\bar{E}\,\frac{x_1'x_1 - x_1'X_2(X_2'X_2)^{-1}X_2'x_1}{n} \qquad (4.3.12)$$

Let us consider the following regression equation

$$x_1 = \delta_{12.34\ldots k}\, x_2 + \delta_{13.24\ldots k}\, x_3 + \cdots + \delta_{1k\cdot 23\ldots k-1} x_k + \varepsilon_1 \qquad (4.3.13)$$

ε_1 is the spherical disturbance term having mean 0 and is independent of $x_2, x_3, \ldots, x_k$.

$$\text{So,}\quad S^2_{1.23\ldots k} = \sum_{i=1}^{n} e^2_{1i} = x_1'x_1 - x_1' X_2(X_2' X_2)^{-1} X_2' x_1 \qquad (4.3.14)$$

Let

$$\sigma^2_{1.23\ldots k} = \bar{E}\left(\frac{1}{n}\sum_{i=1}^{n} e^2_{1i}\right) = \bar{E}\left\{\frac{x_1'x_1 - x_1'X_2(X_2' X_2)^{-1} X_2'x_1}{n}\right\} \qquad (4.3.15)$$

So, from (4.3.12),

$$\lim_{n\to\infty} MSE(\hat{\beta}_1) = \frac{\sigma^2}{n\cdot\ \sigma^2_{1.23\ldots k}} \qquad (4.3.16)$$

Again, the (1,1) element of $(\Sigma^+)^{-1}$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\bar{E}\left\{\frac{x_1'x_1 - x_1'X_2^+(X_2^{+\prime}X_2^+)^{-1}X_2^{+\prime}x_1}{n}\right\}^{-1} \tag{4.3.17}$$

where X_2^+ contains only the first (k-2) columns of X_2.

Considering the regression equation

$$x_1 = \delta_{12.34\ldots k-1}x_2 + \delta_{13.24\ldots k-1}x_3 + \ldots + \delta_{1,k-1,23\ldots k-2}x_{k-1} + \varepsilon_2 \tag{4.3.18}$$

we get

$$S^2_{1.23\ldots k-1} = \sum_{i=1}^{n} e^2_{2i} = x_1'x_1 - x_1'X_2^+(X_2^{+\prime}X_2^+)^{-1}X_2^{+\prime}x_1 \tag{4.3.19}$$

Let $\sigma^2_{1.23\ldots k-1} = \bar{E}(\frac{1}{n}\sum_{i=1}^{n} e^2_{2i}) = \bar{E}\left\{\frac{x_1'x_1 - x_1'X_2^+(X_2^{+\prime}X_2^+)^{-1}X_2^{+\prime}x_1}{n}\right\}$ (4.3.20)

So, from (4.3.9),

$$\lim_{n\to\infty} MSE(\hat{\beta}_1^+) = \beta_k^2[(1,1) \text{ element in } \bar{V}\{(X^{+\prime}X^+)^{-1}X^{+\prime}x_k\}] + \frac{\sigma^2}{n\,\sigma^2_{1.23\ldots k-1}} + \beta_k^2\,\delta^2_{k1.23\ldots k-1} \tag{4.3.21}$$

$$= \beta_k^2[(1,1) \text{ element in } \bar{V}\{(X^{+\prime}X^+)^{-1}X^{+\prime}x_k\}] + \frac{\sigma^2}{n.\sigma^2_{1.23\ldots k-1}} + \beta_k^2\,\rho^2_{1k.23\ldots k-1}\frac{\sigma^2_{k.23\ldots k-1}}{\sigma^2_{1.23\ldots k-1}} \tag{4.3.22}$$

$\rho_{1k.23\ldots k-1}$ is the partial correlation coefficient between x_1 and x_k.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $\lim_{n\to\infty} MSE(\hat{\beta}_1^+) \leq \lim_{n\to\infty} MSE(\hat{\beta}_1)$ if

$$\beta_k^2[(1.1) \text{ element in } \bar{V}\{(X^{+\prime}X^+)^{-1}X^{+\prime}X_k\}] + \frac{\sigma^2}{n\,\sigma^2_{1.23\ldots k-1}} + \beta_k^2\,\rho^2_{1k.23\ldots k-1}\frac{\sigma^2_{k.23\ldots k-1}}{\sigma^2_{1.23\ldots k-1}} \leq \frac{\sigma^2}{n\,\sigma^2_{1.23\ldots k}}$$

or,
$$\frac{\frac{\sigma^2}{n}\left(\frac{1}{\sigma^2_{1.23\ldots k}} - \frac{1}{\sigma^2_{1.23\ldots k-1}}\right)}{\beta_k^2\,\rho^2_{1k.23\ldots k-1}\frac{\sigma^2_{k.23\ldots k-1}}{\sigma^2_{1.23\ldots k-1}}} \geq 1 + \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}} \qquad (4.3.23)$$

where
$$\theta^2_{\delta_{k.123\ldots k-1}} = \frac{\delta^2_{k1.23\ldots k-1}}{\bar{V}(\hat{\delta}_{k1.23\ldots k-1})} = \frac{\rho^2_{1k.23\ldots k-1}\frac{\sigma^2_{k.23\ldots k-1}}{\sigma^2_{1.23\ldots k-1}}}{\bar{V}(\hat{\delta}_{k1.23\ldots k-1})}$$

where $\hat{\delta}_{k1.23\ldots k-1}$ is the OLS estimate of $\delta_{k1.23\ldots k-1}$.

Now, from (4.3.23), we have,

$$\frac{\frac{\sigma^2}{n\,\beta_k^2}\,\frac{1}{\sigma^2_{1.23\ldots k-1}}\left(\frac{1}{1-\rho^2_{1k.23\ldots k-1}} - 1\right)}{\rho^2_{1k.23\ldots k-1}\frac{\sigma^2_{k.23\ldots k-1}}{\sigma^2_{1.23\ldots k-1}}} > 1 + \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{or,}\quad \frac{\sigma^2}{n\,\beta_k^2\,\sigma^2_{k.23\ldots k-1}(1-\rho^2_{1k.23\ldots k-1})} \geq 1 + \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}$$

$$\text{or,}\quad \frac{\sigma^2}{n\,\beta_k^2\,\sigma^2_{k1.23\ldots k-1}} \geq 1 + \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}$$

$$\text{or,}\quad \theta^2_{\beta_k} \leq \frac{1}{1+\dfrac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}}$$

$$= \frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{1+\theta^2_{\delta_{k1.23\ldots k-1}}} \quad \text{6/} \qquad (4.3.24)$$

<u>Remark 1</u>. Let $y, x_1, x_2, \ldots, x_k$ jointly follow a(k+1) variate normal distribution. So, $x_1, x_2, \ldots, x_k$ also jointly follow a k variate normal distribution. Then.

$$\bar{V}(\hat{\delta}_{k1.23\ldots k-1}) = \frac{\sigma^2_{k.23\ldots k-1}(1-\rho^2_{1k.23\ldots k-1})}{n\,\sigma^2_{1.23\ldots k-1}} \qquad (4.3.25)$$

So,

$$\theta^2_{\delta_{k1.23\ldots k-1}} = \frac{n\,\rho^2_{1k.23\ldots k-1}}{1-\rho^2_{1k.23\ldots k-1}} \qquad (4.3.26)$$

6/ For nonstochastic regressors, this condition reduces to $\theta^2_{\beta_k} \leq 1$ (vide Rao, 1971).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, the condition (4.3.24) reduces to

$$\theta^2_{\beta_k} \leq \frac{n\,\rho^2_{1k.23\ldots k-1}}{(n-1)\,\rho^2_{1k.23\ldots k-1}+1} \tag{4.3.27}$$

So, the pretest criterion is

$$t^2_{\beta_k} \leq \frac{n\,r^2_{1k.23\ \ldots\ k-1}}{(n-1)\,r^2_{1k.23\ \ldots k-1}+1}\ ^{7/} \tag{4.3.28}$$

where $r_{1k.23\ \ldots\ k-1}$ is the sample partial correlation coefficient between x_1 and x_k. If (4.3.28) holds, the estimator to be used is $\hat{\beta}_1^{+}$; otherwise, the estimator to be used is $\hat{\beta}_1$.

4.4 MSE dominace of OLS with errors of observations

Let us consider the two variable regression equation given by (4.2.1). Suppose, the variable x_2 is measured with error

i.e. $$x^*_{2i} = x_{2i} + u_i, \quad i = 1, 2, \ldots, n \tag{4.4.1}$$

So, x^*_2 is an observable proxy for x_2. u_i's are i.i.d. $N(0, \sigma^2_u)$ and indepenent of all other variables in the model.

7/ Condition (4.3.28) may hold even when the joint distribution of $y, x_1, x_2, \ldots, x_k$ is not normal. The only requirement for this condition to hold is that the regression of x_k on $x_1, x_2, \ldots, x_{k-1}$ is strictly linear, which is obvious when the joint distribution of $y, x_1, x_2, \ldots, x_k$ is (k+1) variate normal.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Mc.Callum (1972) and Wickens (1972) proved, on the basis of the asymptotic bias of the regression coefficients on the left-in variables, that it is always better to use the proxy variable x_2^* instead of omitting x_2 altogether. Aigner (1974), however, considered the problem from the standpoint of the MSE criterion. He derived a condition under which the MSE of the estimated coefficient of x_1 obtained by omitting x_2 completely is less than that obtained by using the proxy variable x_2^*. The important results in his paper supports the use of a proxy variable in most empirical situations.

Mc Callum (1972) derived the expression for the asymptotic bias of $\hat{\beta}_{1P}$ (the OLS estimate of β when x_2^* has been used in place of x_2 in the regression equation) as

$$\underset{n\to\infty}{\text{plim}}\ (\hat{\beta}_{1P} - \beta_1) = \frac{\sigma_{12}}{\sigma_{11}}\left(\frac{\sigma_u^2}{\sigma_u^2 + M_{21}^2}\right) \tag{4.4.2}$$

where $M_{21}^2 = \sigma_{22}\left(1 - \frac{\sigma_{21}^2}{\sigma_{11} \cdot \sigma_{22}}\right)$ = the residual variance in the regression of x_2 on x_1. Since $\frac{\sigma_u^2}{\sigma_u^2 + M_{21}^2} < 1$, the asymptotic bias of $\hat{\beta}_{1P}$ is less than that of the OV estimator $\hat{\beta}_1^+$.

Aigner (1974) derived the expression for the asymptotic sampling variance of $\hat{\beta}_{1P}$. He assumed that y, x and x_2 jointly follow a trivariate normal distribution with the variance-

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

covariance matrix

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \qquad (4.4.3)$$

The sampling variance [8/] of β_{1P} derived by Aigner is

$$V(\hat{\beta}_{1P}) = \frac{\sigma^2}{n}\left(\frac{\sigma_{22}+\sigma_u^2}{\phi}\right) + \beta_2^2 \left\{\frac{\sigma_u^2(\sigma_{22}\,\sigma_{11} - \sigma_{12}^2)}{\phi}\right\}\left(\frac{\sigma_{22}+\sigma_u^2}{\phi}\right) \qquad (4.4.4)$$

where $\phi = (\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2$. Thus,

$$\text{MSE}(\hat{\beta}_{1P}) = \beta_2^2\left(\frac{\sigma_{21}\;\sigma_u^2}{\phi}\right)^2 + \frac{\sigma^2}{n}\left(\frac{\sigma_{22}+\sigma_u^2}{\phi}\right) + \frac{\beta_2^2}{n}\left\{\frac{\sigma_u^2(\sigma_{22}\;\sigma_{11}-\sigma_{12}^2)}{\phi}\right\}\left(\frac{\sigma_{22}+\sigma_u^2}{\phi}\right) \qquad (4.4.5)$$

The MSE($\hat{\beta}_1^+$) is given by

$$\text{MSE}(\hat{\beta}_1^+) = \beta_2^2\,\frac{\sigma_{12}^2}{\sigma_{11}^2} + \frac{\sigma^2}{n\,\sigma_{11}} \qquad (4.4.6)$$

Aigner compared (4.4.5) with (4.4.6) and obtained a sufficient condition for MSE($\hat{\beta}_{1P}$) $\geq$ MSE($\hat{\beta}_1^+$) as

$$\frac{\{1 - (1 - \lambda n)\,\rho_{12}^2\}}{\{1 - (1 - \lambda)\,\rho_{12}^2\}^2} \cdot \frac{\lambda}{n} \geq \rho_{12}^2 \qquad (4.4.7)$$

where $\lambda = \dfrac{\sigma_u^2}{\sigma_{22}+\sigma_u^2}$.

8/ These results (MSE $\hat{\beta}_{1P}$ and MSE $\hat{\beta}_1^+$) are all asymptotic. But in the paper by Aigner (1974) has claimed them to be exact finite sample results.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Now, MSE($\hat{\beta}_1^+$) as given by Aigner in (4.4.6) is true only for nonstochastic regressors. Since Aigner deals with stochastic regressors, MSE($\hat{\beta}_{1P}$) should actually be compared with MSE($\hat{\beta}_1^+$) for stochastic regressors. In this section we do this for a k-variable regression equation in which one variable has been measured with error.

Let us consider the following regression equation

$$y = x_1\beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \varepsilon \quad (4.4.8)$$

Suppose that x_k has been measured with error. So, instead of having observations on x_k, we have observations on

$$x_k^* = x_k + u \quad (4.4.9)$$

where u is assumed to follow a normal distribution with mean 0 and variance σ_u^2. u is also assumed to be independent of $x_1, x_2, \dots, x_k$. Moreover, y, x_1, $x_2, \dots, x_k$ are assumed to follow jointly a (k+1) variate normal distribution. ε is $N(0, \sigma^2)$ and distributed independently of $x_1, x_2, \dots, x_k$.

Thus the E-V model is given by

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}^* + (\varepsilon_i - \beta_k u_i), \quad i=1,2,\dots,n \quad (4.4.10)$$

In matrix notation this can be written as

$$y = X^*\beta + \varepsilon - \beta_k u \quad (4.4.11)$$

where $y = (y_1, y_2, \dots, y_n)'$ and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{and} \quad X^* = \begin{pmatrix} x_{11} & x_{21} & \cdots & x^*_{k1} \\ x_{12} & x_{22} & \cdots & x^*_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & & x^*_{kn} \end{pmatrix}$$

$\beta = (\beta_1, \beta_2, \dots \beta_k)'$, $\varepsilon - \beta_k u = (\varepsilon_1 - \beta_k u_1, \varepsilon_2 - \beta_k u_2, \dots, \varepsilon_n - \beta_k u_n)'$.

The OLS estimate of β is

$$\hat{\beta}_{1P} = (X^{*\prime} X^*)^{-1} X^{*\prime} y$$

$$= \beta + (X^{*\prime} X^*)^{-1} X^{*\prime} \varepsilon - \beta_k (X^{*\prime} X^*)^{-1} X^{*\prime} u \qquad (4.4.12)$$

$$\therefore \bar{E}(\hat{\beta}_{1P}) = \beta + \bar{E}\{(X^{*\prime} X^*)^{-1} X^{*\prime} \varepsilon\} - \beta_k \bar{E}\{(X^{*\prime} X^*)^{-1} X^{*\prime} u\}$$

$$= \beta - \beta_k \bar{E}\{(X^{*\prime} X^*)^{-1} X^{*\prime} u\} \qquad (4.4.13)$$

$$\bar{V}(\hat{\beta}_{1P}) = \bar{V}\{(X^{*\prime} X^*)^{-1} X^{*\prime} \varepsilon\} + \beta_k^2 \bar{V}\{(X^{*\prime} X^*)^{-1} X^{*\prime} u\}$$

$$= \frac{\sigma^2}{n} \bar{E}\left(\frac{X^{*\prime} X^*}{n}\right)^{-1} + \beta_k^2 \bar{V}\{(X^{*\prime} X^*)^{-1} X^{*\prime} u\} \qquad (4.4.14)$$

Next, considering the term $(X^{*\prime} X^*)^{-1} X^{*\prime} u$, we shall obtain the probability limit and the asymptotic variance of $\hat{\beta}_{1P}$ (the first element of $\hat{\beta}_{1P}$).

Since $x_1, x_2, \dots, x_k^*, u$ jointly follow a (k+1) variate normal distribution, the distribution of u (conditional on $x_1, x_2, \dots, x_k^*$) is a univariate normal distribution with conditional mean

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(u/x_1, x_2, \ldots, x_k^*) = a_1x_1 + a_2x_2 + \ldots + a_kx_k^* \tag{4.4.15}$$

where $$A = (a_1, a_2, \ldots, a_k)' = \Sigma_{22}^{-1}\,\Sigma_{12} \tag{4.4.16}$$

$$\Sigma_{22} = E\left\{ \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ki}^* \end{pmatrix} (x_{1i}, x_{2i}, \ldots, x_{ki}^*) \right\}$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & & & \\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk} + \sigma_u^2 \end{pmatrix} \quad \forall\, i \tag{4.4.17}$$

and $$\Sigma_{12} = E(u_i) \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ki}^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \sigma_u^2 \end{pmatrix} \quad \forall\, i \tag{4.4.18}$$

Thus, one can consider the following regression equation

$$u_i = a_1x_{1i} + a_2x_{2i} + \ldots + a_kx_{ki}^* + \eta_i, \quad i = 1, 2, \ldots, n \tag{4.4.19}$$

where η_i's are spherical disturbances with zero means and common variance σ_η^2. η's are distributed independently of $x_1, x_2, \ldots, x_k^*$. The OLS estimate of A is given by

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$A = (X^{*\prime} X^*)^{-1} X^{*\prime} u$$

$$= A + (X^{*\prime} X^*)^{-1} X^{*\prime} \eta \qquad (4.4.20)$$

$$\therefore \quad E(\hat{A}) = A \qquad (4.4.21)$$

$$V(\hat{A}) = E(\hat{A} - A)(\hat{A} - A)' = E(X^{*\prime}X^*)^{-1}X^{*\prime}\eta\eta' X(X^{*\prime}X^*)^{-1}$$

$$= E\{(X^{*\prime}X^*)^{-1}X^{*\prime}E(\eta\eta' | X^*)X^*(X^{*\prime}X^*)^{-1}\}$$

$$= \frac{1}{n} \sigma_\eta^2 \, E\left(\frac{X^{*\prime}X^*}{n}\right)^{-1} \qquad (4.4.22)$$

$$\therefore \quad \overline{V}(\hat{A}) = \frac{1}{n} \sigma_\eta^2 \, \overline{E}(X^{*\prime}X^*)^{-1} \qquad (4.4.23)$$

Since, under quite general conditions, sample moments give consistent estimates of population moments,

$$\underset{n \to \infty}{\text{plim}} \left(\frac{X^{*\prime}X^*}{n}\right)^{-1} = \Sigma_{22}^{-1} = \overline{E}(X^{*\prime}X^*)^{-1} \qquad (4.4.24)$$

$\therefore$ from (4.4.24)

$$\overline{V}(A) = \frac{1}{n} \sigma_\eta^2 \, \Sigma_{22}^{-1} \qquad (4.4.25)$$

where $\sigma_\eta^2 = (\sigma_u^2 - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$ (4.4.26)

and $\Sigma_{21} = \Sigma_{12}'$ (4.4.27)

Next, for obtaining $\overline{E}(\hat{\beta}_{1P})$ and $\overline{V}(\hat{\beta}_{1P})$, we consider the first element in the vector $\Sigma_{22}^{-1} \Sigma_{12}$ and the (1,1) element in

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

the matrix $\frac{1}{n}(\sigma_u^2 - \Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21})\,\Sigma_{22}^{-1}$. The first element in the vector $\Sigma_{22}^{-1}\,\Sigma_{12}$ is

$$\frac{\sigma_u^2 \text{ cofactor of } \sigma_{1k} \text{ in } \Sigma_{22}}{\text{Det } \Sigma_{22}} \tag{4.4.28}$$

Now, it can be shown that cofactor of σ_{1k} in Σ_{22} is

$$\sigma_{1k} \text{ cofactor of } \sigma_{11} \text{ in } \Delta_1 + \sigma_{2k} \text{ cofactor of } \sigma_{12} \text{ in } \Delta_1 + \dots$$
$$+ \sigma_{k-1,k} \text{ cofactor of } \sigma_{1,k-1} \text{ in } \Delta_1 \tag{4.4.29}$$

where $$\Delta_1 = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1,k-1} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \cdots & \sigma_{2,k-1} \\ \vdots & & & & \\ \sigma_{1,k-1} & \sigma_{2,k-1} & \sigma_{3,k-1} & \cdots & \sigma_{k-1,k-1} \end{pmatrix} \tag{4.4.30}$$

$$\text{Det } \Sigma_{22} = |\Sigma_{22}| = \text{Det } \bar{\Delta} + \sigma_u^2 \text{ cofactor of } \sigma_{kk} \text{ in } \bar{\Delta}$$
$$= |\bar{\Delta}| + \sigma_u^2\,|\Delta_1| \tag{4.4.31}$$

where $$\bar{\Delta} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & & & \\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk} \end{pmatrix} \tag{4.4.32}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The (1,1) element in the matrix $\frac{1}{n}(\sigma_u^2 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\Sigma_{22}^{-1}$ is given by

$$\frac{1}{n}(\sigma_u^2 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) \frac{\text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22}}{|\Sigma_{22}|} \qquad (4.4.33)$$

$$\text{Now, } \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \frac{\sigma_u^2 \text{ cofactor of } (\sigma_{kk}+\sigma_u^2) \text{ in } \Sigma_{22}}{|\Sigma_{22}|} \qquad (4.4.34)$$

$$\text{So, } \bar{V}(\hat{a}_1) = \frac{\sigma_u^2}{(|\Sigma_{22}|)^2} (|\Sigma_{22}| - \sigma_u^2 \text{ cofactor of } (\sigma_{kk}+\sigma_u^2) \text{ in } \Sigma_{22}) \times \text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22}$$

$$= \frac{\sigma_u^2}{(|\Sigma_{22}|)^2} |\bar{\Delta}| \times \text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22} \qquad (4.4.35)$$

So, from (4.4.28) and (4.4.35)

$$\lim_{n\to\infty} \text{MSE}(\hat{\beta}_{1P}) = \frac{\sigma^2}{n} \frac{\text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22}}{|\bar{\Delta}| + \sigma_u^2|\Delta_1|} + \frac{\beta_k^2}{n} \frac{\sigma_u^2}{|\bar{\Delta}|+\sigma_u^2|\Delta_1|} |\bar{\Delta}|$$

$$\times \text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22} + \beta_k^2 \frac{\sigma_u^2}{n} \frac{(\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22})^2}{(|\bar{\Delta}|+\sigma_{ee}\,|\Delta_1|)^2}$$

$$(4.4.36)$$

Now, consider the case where x_k has been totally omitted from the model (4.4.8). Following the derivations given in

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Section 3,

$$\lim_{n\to\infty} MSE(\beta_1^+) = \frac{\sigma^2}{n} \frac{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|} + \beta_k^2 \{(1,1)\text{element in}$$

$$\overline{V}(X^{+\prime}X^+)^{-1} X^{+\prime}X_k\} + \beta_k^2 (\text{the 1st element of } \Delta_1^{-1}\overline{\Sigma}_{12}) \quad (4.4.37)$$

where $X^+ = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k-1,1} \\ x_{12} & x_{22} & \cdots & x_{k-1,2} \\ \vdots & & & \\ x_{1n} & x_{2n} & \cdots & x_{k-1,n} \end{pmatrix}$,

$$X_k = (x_{k1}, x_{k2}, \ldots, x_{kn})'$$

$$\overline{\Sigma}_{12} = (\sigma_{1k}, \sigma_{2k}, \ldots, \sigma_{k-1,k})'.$$

$$\text{The 1st element of } \Delta_1^{-1}\overline{\Sigma}_{12} = \frac{\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22}}{|\Delta_1|} \quad (4.4.38)$$

$$(1,1)\text{element in } \overline{V}\ (X^{+\prime}X^+)^{-1}X^{+\prime}X_k = \left(\frac{\sigma_{kk} - \overline{\Sigma}_{12}\Delta_1^{-1}\overline{\Sigma}_{12}'}{n}\right) \times \frac{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|}$$

(4.4.39)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $\lim_{n\to\infty} MSE(\hat{\beta}_{1P}) \geq \lim_{n\to\infty} MSE(\hat{\beta}_1^{+})$ if and only if

$$\frac{\sigma^2}{n} \frac{\text{cofactor of } \sigma_{11} \text{ in } \Sigma_{22}}{|\bar{\Delta}| + \sigma_u^2 |\Delta_1|} + \frac{\beta_k^2}{n} \frac{\sigma_u^2}{(|\bar{\Delta}| + \sigma_u^2 \; |\Delta_1|)^2} |\bar{\Delta}| \text{ cofactor of } \sigma_{11} \text{ in } \Sigma_{22}$$

$$+ \beta_k^2 \sigma_u^2 \frac{(\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22})^2}{(|\bar{\Delta}| + \sigma_u^2 \; |\Delta_1|)^2} \geq \frac{\sigma^2}{n} \frac{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|}$$

$$+ \frac{\beta_k^2}{n} \frac{(\sigma_{kk} - \bar{\Sigma}_{12} \Delta_1^{-1} \bar{\Sigma}_{12}') \text{ cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|}$$

$$+ \beta_k^2 \frac{(\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22})^2}{(|\Delta_1|)^2} \qquad (4.4.40)$$

Now, let us partition the matrix $\bar{\Delta}$ in the following way.

$$\bar{\Delta} = \begin{pmatrix} \Delta_1 & \vdots & \bar{\Sigma}_{12} \\ \cdots & \cdots & \cdots \\ \bar{\Sigma}_{12}' & \vdots & \sigma_{kk} \end{pmatrix} \qquad (4.4.41)$$

$$\therefore |\bar{\Delta}| = |\Delta_1| (\sigma_{kk} - \bar{\Sigma}_{12} \Delta_1^{-1} \bar{\Sigma}_{12}') \qquad (4.4.42)$$

also, cofactor of σ_{11} in Σ_{22} = cofactor of σ_{11} in $\bar{\Delta}$

$$+ \sigma_u^2 \text{ cofactor of } \sigma_{11} \text{ in } \Delta_1 \qquad (4.4.43)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\therefore$ from (4.4.40), using (4.4.42) and (4.4.43) we get the condition

$$\frac{\sigma^2}{n\,\beta_k^2}\;\frac{(|\Delta_1|\text{ cofactor of }\sigma_{11}\text{ in }\bar{\Delta} - |\bar{\Delta}|\text{ cofactor of }\sigma_{11}\text{ in }\Delta_1)}{(|\Delta_1|)^2\left(\frac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2\right)\;\frac{(\text{cofactor of }\sigma_{1k}\text{ in }\Sigma_{22})^2}{(|\Delta_1|)^2}}$$

$$+\frac{\sigma_u^2}{n}\;\frac{|\Delta_1|\,(\sigma_{kk}-\bar{\Sigma}_{12}\,\Delta_1^{-1}\bar{\Sigma}_{12}')\text{ cofactor of }\sigma_{11}\text{ in }\Sigma_{22}}{(|\Delta_1|)^2\left(\frac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2\right)\left(\frac{\text{cofactor of }\sigma_{1k}\text{ in }\Sigma_{22})^2}{(|\Delta_1|)^2}\right.}$$

$$+\frac{\sigma_u^2}{\left(\frac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2\right)^2} \geq 1+\frac{1}{n}\;\frac{(\sigma_{kk}-\bar{\Sigma}_{12}\,\Delta_1^{-1}\,\Sigma_{12}')\text{ cofactor of }\sigma_{11}\text{ in }\Delta_1}{|\Delta_1|\;\frac{(\text{cofactor of }\sigma_{1k}\text{ in }\Sigma_{22})^2}{(|\Delta_1|)^2}},$$

(4.4.44)

or, $$\frac{\sigma^2}{n\,\beta_k^2}\;\frac{|\bar{\Delta}|}{|\Delta_1|}\;\frac{\left(\frac{\text{cofactor of }\sigma_{11}\text{ in }\bar{\Delta}}{|\bar{\Delta}|}-\frac{\text{cofactor of }\sigma_{11}\text{ in }\Delta_1}{|\Delta_1|}\right)}{\left(\frac{\text{cofactor of }\sigma_{1k}\text{ in }\Sigma_{22}}{|\Delta_1|}\right)^2}$$

$$\geq \frac{\left(\frac{|\bar{\Delta}|}{|\Delta_1|}\right)^2+2\frac{|\bar{\Delta}|}{|\Delta_1|}\,\sigma_u^2}{\frac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2}+\frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}\;\frac{\sigma_u^2\text{ cofactor of }\sigma_{11}\text{ in }\Sigma_{22}}{\left(\frac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2\right)^2.\text{cofactor of }\sigma_{11}\text{ in }\Delta_1}$$

(4.4.45)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\theta^2_{\delta_{k1.23\ldots k-1}} = \frac{\delta^2_{k1.23\ldots k-1}}{V(\hat{\delta}_{k1.23\ldots k-1})}$

and $\hat{\delta}_{k1.23\ldots k-1}$ and $\delta_{k1.23\ldots k-1}$ have been defined in Section 3.

$\therefore$ from (5.4.45), we get

$$\frac{\sigma^2}{n\,\beta_k^2}\,\frac{|\bar{\Delta}|}{|\Delta_1|}\;\frac{\left(\frac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{|\bar{\Delta}|} - \frac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}_1}{|\Delta_1|}\right)}{\left(\frac{\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22}}{|\Delta_1|}\right)^2}\,.$$

$$\geq \frac{\left(\frac{|\bar{\Delta}|}{|\Delta_1|}\right)^2 + 2\,\frac{|\bar{\Delta}|}{|\Delta_1|}\,\sigma_u^2}{\frac{|\bar{\Delta}|}{|\Delta_1|} + \sigma_u^2}$$

$$+ \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}}\;\frac{\left(\left(\frac{|\bar{\Delta}|}{|\Delta_1|}\right)^2 + 2\sigma_u^2\,\frac{|\bar{\Delta}|}{|\Delta_1|} - \sigma_u^2\,\frac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}\right)}{\frac{|\bar{\Delta}|}{|\Delta_1|} + \sigma_u^2}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{or, } \frac{\sigma^2}{n\,\beta_k^2} \; \frac{\left(\dfrac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{|\bar{\Delta}|} - \dfrac{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|}\right)}{\left(\dfrac{\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22}}{|\Delta_1|}\right)}$$

$$\geq \frac{\dfrac{|\bar{\Delta}|}{|\Delta_1|} + 2\sigma_u^2}{\dfrac{|\bar{\Delta}|}{|\Delta_1|} + \sigma_u^2}$$

$$+ \frac{1}{\theta^2_{\delta_{k1.23\ldots k-1}}} \; \frac{\left(\left(\dfrac{|\bar{\Delta}|}{|\Delta_1|}\right)^2 + 2\sigma_u^2 \dfrac{|\bar{\Delta}|}{|\Delta_1|} - \sigma_u^2 \dfrac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}\right.}{\left(\dfrac{|\bar{\Delta}|}{|\Delta_1|}\right)^2 + \sigma_u^2 \dfrac{|\bar{\Delta}|}{|\Delta_1|}} \qquad (4.4.46)$$

Now, from Section (4.3),

$$\frac{\sigma^2}{n}\left(\frac{\left(\dfrac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{|\bar{\Delta}|} - \dfrac{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}{|\Delta_1|}\right)}{\left(\dfrac{\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22}}{|\Delta_1|}\right)^2}\right. = \bar{V}(\hat{\beta}_k) \qquad (4.4.47)$$

So, from (4.4.46),

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\frac{\beta_k^2}{\bar{V}(\hat{\beta}_k)} \leq \frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{\theta^2_{\delta_{k1.23\ldots k-1}}\left(1+\dfrac{\sigma_u^2}{\dfrac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2}\right)+1+\dfrac{\sigma_u^2}{\dfrac{|\bar{\Delta}|}{|\Delta_1|}+\sigma_u^2}\left(1-\dfrac{\dfrac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}}{\dfrac{|\bar{\Delta}|}{|\Delta_1|}}\right)} \qquad (4.4.48)$$

Again, $$1-\frac{\dfrac{\text{cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}}{\dfrac{|\bar{\Delta}|}{|\Delta_1|}}$$

$$= \frac{|\bar{\Delta}| \text{ cofactor of } \sigma_{11} \text{ in } \Delta_1 - |\Delta_1| \text{ cofactor of } \sigma_{11} \text{ in } \bar{\Delta}}{|\bar{\Delta}| \text{ cofactor of } \sigma_{11} \text{ in } \Delta_1}$$

$$= -\frac{|\Delta_1|}{|\bar{\Delta}|}\,\frac{|\Delta_1|}{\text{cofactor of } \sigma_{11} \text{ in } \Delta_1}\cdot\frac{(\text{cofactor of } \sigma_{1k} \text{ in } \Sigma_{22})^2}{(|\Delta_1|)^2}$$

$$= -\frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{n} \qquad \text{(using (4.4.42))} \qquad (4.4.49)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\therefore$ for large n,

$$\theta^2_{\beta_k} < \frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{-\,\theta^2_{\delta_{k1.23\ldots k-1}}\left(1 + \frac{\sigma_u^2}{\frac{|\bar{\Delta}|}{|\Delta_1|} + \sigma_u^2}\right) + 1}\,,$$

$$= \frac{\theta^2_{\delta_{k1.23\ldots k-1}}}{\theta^2_{\delta_{k1.23\ldots k-1}}\left(1 + \sigma_u^2/(\sigma_{kk} - \bar{\Sigma}_{12}\,\Delta_1^{-1}\,\bar{\Sigma}'_{12}) + \sigma_u^2\right) + 1} \qquad (4.4.50)$$

Now, $\theta^2_{\beta_k} = \beta_k^2 \Big/ \left(\frac{\sigma^2}{n\sigma^2_{k.123\ldots k-1}}\right)$ [vide (4.3.24)]

$$= \beta_k^2 \Big/ \left\{\frac{\sigma^2}{n\sigma^2_{k.23\ldots k-1}(1 - \rho^2_{1,k.23\ldots k-1})}\right\} \qquad (4.4.51)$$

[vide (4.4.42)]

and $\frac{|\bar{\Delta}|}{|\Delta_1|} = \sigma^2_{k.123\ldots k-1}$

$$= \sigma^2_{k.123\ldots k-1}\,(1 - \rho^2_{1k.23\ldots k-1}) \qquad (4.4.52)$$

So, the condition in (4.4.50) can also be written as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\theta^2_{\beta_k} = \beta_k^2 \Big/ \left\{ \frac{\sigma^2}{n\sigma^2_{k.23...k-1}(1-\rho^2_{1k.23...k-1})} \right\}$$

$$\leq \frac{\theta^2_{\delta_{1k.23...k-1}}}{\theta^2_{\sigma_{1k.23...k-1}} \left\{ 1 + \frac{\sigma_u^2}{\sigma^2_{k.23...k-1}(1-\rho^2_{1k.23...k-1}) + \sigma_u^2} \right\}} \qquad (4.4.53)$$

where, for large n, $\frac{\theta^2_{\delta_{1k.23...k-1}}}{n} = \frac{\rho^2_{1,k.23...k-1}}{1 - \rho^2_{1k.23...k-1}}$

Next, Comparing (4.4.53) with (4.3.29), we find that the condition to omit x_k^* is stronger than the condition to omit x_k. This, however, does not admit any easy intuitive explanation.

For $k = 2$, the condition for the asymptotic MSE of β_1 to be greater than that of β_1^+ has been derived.

Here, the true model is

$$y = x_1\beta_1 + x_2\beta_2 + \varepsilon \qquad (4.4.54)$$

and

$$x_2^* = x_2 + u \qquad (4.4.55)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The assumptions regarding ε and u are same as in the general case given by (4.4.8) and (4.4.9), y, x_1 and x_2 jointly followed a normal distribution. x_1 and x_2 have zero means and the variance-covariance matrix given by

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \tag{4.4.56}$$

Here, $$\lim_{n\to\infty} \text{MSE}(\hat{\beta}_{1P}) = \frac{\sigma^2}{n} \left\{ \frac{\sigma_{22} + \sigma_u^2}{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2} \right\}$$

$$+ \frac{\beta_2^2}{n} \left\{ \frac{\sigma_u^2(\sigma_{22}\sigma_{11} - \sigma_{12}^2)}{\sigma_{22}^2 + \sigma_u^2)\,\sigma_{11}^2 - \sigma_{12}^2} \right\} \left\{ \frac{\sigma_{22} + \sigma_u^2}{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2} \right\}$$

$$+ \beta_2^2 \left\{ \frac{\sigma_{12}\,\sigma_u^2}{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2)} \right\}^2 \tag{4.4.57}$$

and $$\lim_{n\to\infty} \text{MSE}(\hat{\beta}_1^+) = \frac{\sigma^2}{n\,\sigma_{11}} + \frac{\beta_2^2}{n}\,\frac{\sigma_{22}(1 - \rho_{12}^2)}{\sigma_{11}} + \beta_2^2\,\frac{\sigma_{12}^2}{\sigma_{11}^2} \tag{4.4.58}$$

So, $\lim_{n\to\infty} \text{MSE}(\hat{\beta}_{1P}) > \lim_{n\to\infty} \text{MSE}(\hat{\beta}_1^+)$ if and only if

$$\theta_{\beta_2}^2 = \frac{\theta_{\delta_{21}}^2}{\theta_{\delta_{21}}^2 \left(1 + \frac{\sigma_u^2}{M_{21}^2 + \sigma_u^2}\right)} \tag{4.4.59}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where for large n, $\sqrt{\theta^2_{\delta_{21}}} = \frac{(1 - \rho^2_{12})}{n \rho^2_{12}}$, ρ^2_{12} being the squared population correlation coefficient between x_2 and x_1.

However, for this simple case, the exact results for $MSE(\hat{\beta}_{1P})$ and $MSE(\hat{\beta}^+_1)$ are available.

$$V(\hat{\beta}^+_1) = \sigma^2 E\left(\frac{1}{\sum_{i=1}^{n} x^2_{1i}}\right) + \beta^2_2 \; V(\delta_{21}) \quad (4.4.60)$$

and $$V(\hat{\delta}_{21}) = \sigma_{22}(1 - \rho^2_{12}) \; E\left(\frac{1}{\sum_{i=1}^{n} x^2_{1i}}\right) \quad (4.4.61)$$

Now, $$\frac{\sum_{i=1}^{n} x^2_{1i}}{\sigma_{11}} \sim \chi^2_n \text{ (chisquare with } n \text{ degrees of freedom)} \quad (4.4.62)$$

So, $$E\left(\frac{1}{\sum_{i=1}^{n} x^2_{1i}}\right) = \frac{1}{(n-2)\sigma_{11}} \quad (4.4.63)$$

So, $$MSE(\hat{\beta}^+_1) = \frac{\sigma^2}{(n-2)\sigma_{11}} + \beta^2_2 \; \frac{\sigma_{22}(1-\rho^2_{12})}{(n-2)\sigma_{11}} + \beta^2_2 \; \frac{\sigma^2_{12}}{\sigma^2_{11}} \quad (4.4.64)$$

Also, $$V(\hat{\beta}_{1P}) = \sigma^2 \; E\{(1,1) \text{ element in } (X^{*\prime} X^*)^{-1}\}$$

$$+ \; \beta^2_2 \; \frac{\sigma^2_u \; (\sigma_{11} \; \sigma_{22}) - \sigma^2_{12}}{(\sigma^2_{22} + \sigma^2_u) \; \sigma^2_{11} - \sigma^2_{12}} \; E\{(1,1) \text{element in } (X^{*\prime}X^*)^{-1}\} \quad (4.4.65)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $X^* = \begin{pmatrix} x_{11} & x^*_{21} \\ x_{12} & x^*_{21} \\ \vdots & \vdots \\ x_{1n} & x^*_{2n} \end{pmatrix}$

Now, $E\{(1,1) \text{ element in } (X^{*\prime}X^*)^{-1}\}$ = the (1,1) element in

$$\Sigma^{*-1} \cdot E \frac{1}{\chi^2_{n-1}} \quad (\underline{\text{vide}} \text{ Anderson(1957), pp. 170-171})$$

$$= \text{the}(1,1) \text{ element in } \Sigma^{*-1} \cdot \frac{1}{n-3} \qquad (4.4.66)$$

where $\Sigma^* = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} + \sigma_u^2 \end{pmatrix}$

So, $$\text{MSE}(\hat{\beta}_{1P}) = \frac{\sigma^2}{n-3} \frac{\sigma_{22} + \sigma_u}{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2} + \frac{\beta_2^2}{n-3} \frac{\sigma_u^2(\sigma_{22}\sigma_{11} - \sigma_{12}^2)(\sigma_{22} + \sigma_u^2)}{\{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2\}^2}$$

$$+ \beta_2^2 \left\{ \frac{\sigma_{12}\,\sigma_u^2}{(\sigma_{22} + \sigma_u^2)\,\sigma_{11} - \sigma_{12}^2} \right\}^2 \qquad (4.4.67)$$

A sufficient condition for $\text{MSE}(\hat{\beta}_{1P}) \geq \text{MSE}(\hat{\beta}_1^+)$ is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\frac{1}{n-3}\ \frac{\sigma_u^2(\sigma_{22}\sigma_{12}-\sigma_{12}^2)(\sigma_{22}+\sigma_u^2)}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}+\frac{\sigma_{12}^2\,\sigma_u^4}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}$$

$$\geq \frac{\sigma_{22}(1-\rho_{12}^2)}{(n-2)\,\sigma_{11}}+\frac{\sigma_{12}^2}{\sigma_{11}} \qquad (4.4.68)$$

For the above inequality to hold, it is sufficient that

$$\frac{1}{n-2}\ \frac{\sigma_u^2(\sigma_{22}\sigma_{11}-\sigma_{12}^2)(\sigma_{22}+\sigma_u^2)}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}+\frac{\sigma_{12}^2\,\sigma_u^4}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}$$

$$\geq \frac{\sigma_{22}(1-\rho_{12}^2)}{(n-2)\,\sigma_{11}}+\frac{\sigma_{12}^2}{\sigma_{11}} \qquad (4.4.69)$$

or, $$\frac{1}{n-2}\ \frac{\lambda(1-\rho_{12}^2)\cdot\sigma_{22}\,\sigma_{11}}{\sigma_{11}^2\left\{1-\dfrac{\sigma_{12}^2}{\sigma_{11}(\sigma_{22}+\sigma_u^2)}\right\}^2}+\frac{\sigma_{12}^2\,\lambda^2}{\sigma_{11}^2\left\{1-\dfrac{\sigma_{12}^2}{\sigma_{11}(\sigma_{22}+\sigma_u^2)}\right\}^2}$$

$$-\frac{\sigma_{22}\,(1-\rho_{12}^2)}{(n-2)\,\sigma_{11}} \geq \frac{\sigma_{12}^2}{\sigma_{11}}$$

or, $$\frac{1}{n-2}\ \frac{\lambda(1-\rho_{12}^2)}{\{1-(1-\lambda)\,\rho_{12}^2\}^2}+\frac{\lambda^2\,\rho_{12}^2}{\{1-(1-\lambda)\,\rho_{12}^2\}^2}-\frac{(1-\rho_{12}^2)}{n-2} \geq \rho_{12}^2$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\text{or,} \quad \frac{1-\{1-(n-2)\lambda\}\rho_{12}^2}{\{1-(1-\lambda)\rho_{12}^2\}^2} \cdot \frac{\lambda}{n-2} - \frac{(1-\rho_{12}^2)}{n-2} \geq \rho_{12}^2 \qquad (4.4.70)$$

where $\lambda = \dfrac{\sigma_u^2}{\sigma_{12} + \sigma_u^2}$.

It can be shown easily that for $n'(=n-2) = 5, 10, 15, \ldots 50$, $\rho_{12}^2 = .1, .2, \ldots, .7$, and $\lambda = .1, .2, \ldots, .9$, the inequality in (4.4.70) is not satisfied. This, however, does not mean necessarily that $MSE(\hat{\beta}_{1P}) < MSE(\hat{\beta}_1^+)$.

The sufficient condition for $MSE(\hat{\beta}_{1P}) < MSE(\hat{\beta}_1^+)$ is

$$\frac{\dfrac{1}{n-3}\,\dfrac{\sigma_u^2(\sigma_{22}\sigma_{11}-\sigma_{12}^2)(\sigma_{22}+\sigma_u^2)}{\{(\sigma_{22}+\sigma_u^2)\sigma_{11}-\sigma_{12}^2\}^2} + \dfrac{\sigma_{12}^2\,\sigma_u^4}{\{(\sigma_{22}+\sigma_u^2)\sigma_{11}-\sigma_{12}^2\}^2}}{\dfrac{1}{(n-2)\sigma_{11}}}$$

$$< \frac{\dfrac{\sigma_{22}(1-\rho_{12}^2)}{(n-2)\,\sigma_{11}} + \dfrac{\sigma_{12}^2}{\sigma_{11}}}{\dfrac{1}{(n-3)}\,\dfrac{\sigma_{22}^2+\sigma_u^2}{\{(\sigma_{22}+\sigma_u^2)\sigma_{11}-\sigma_{12}^2\}}} \qquad (4.4.71)$$

For the above inequality to hold, it is sufficient that

$$\frac{\sigma_{11}}{n-3}\;\frac{\sigma_u^2(\sigma_{22}\sigma_{12}-\sigma_{12}^2)(\sigma_{22}+\sigma_u^2)}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}+\frac{\sigma_{12}^2\;\sigma_u^4}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}^2}$$

$$<\frac{\left\{\dfrac{\sigma_{22}(1-\rho_{12}^2)}{(n-2)\;\sigma_{11}}+\dfrac{\sigma_{12}^2}{\sigma_{11}^2}\right\}n_0}{\dfrac{\sigma_{22}+\sigma_u^2}{\{(\sigma_{22}+\sigma_u^2)\,\sigma_{11}-\sigma_{12}^2\}}}\qquad \text{where } n_0=\frac{n-3}{n-2}$$

or,

$$\frac{\lambda\,\sigma_u^2\,\sigma_{12}^2\left\{\lambda+\dfrac{(1-\rho_{12}^2)}{(n-3)\;\rho_{12}^2}\right\}}{\sigma_{11}\left(\sigma_u^2-\lambda\dfrac{\sigma_{12}^2}{\sigma_{11}}\right)^2}<\frac{\left\{\dfrac{\sigma_{12}^2}{\sigma_{11}}+\dfrac{\sigma_{22}}{(n-2)}(1-\rho_{12}^2)\right\}\left(\sigma_u^2-\lambda\dfrac{\sigma_{12}^2}{\sigma_{11}}\right)\cdot n_0}{\sigma_u^2}$$

or,

$$\frac{\dfrac{\lambda^3}{(1-\lambda)^2}\,\rho_{12}^2\left\{\lambda+\dfrac{(1-\rho_{12}^2)}{(n-3)\;\rho_{12}^2}\right\}}{\left(\dfrac{\lambda}{1-\lambda}-\lambda\,\rho_{12}^2\right)^2}<\frac{\left\{\rho_{12}^2+\dfrac{(1-\rho_{12}^2)}{n-2}\right\}\left(\dfrac{\lambda}{1-\lambda}-\lambda\,\rho_{12}^2\right)\cdot n_0}{\dfrac{\lambda}{1-\lambda}}$$

or,

$$\frac{\lambda}{n-3}\;\frac{[1-\{1-(n-3)\,\lambda\,\rho_{12}^2\}]}{\{1-(1-\lambda)\,\rho_{12}^2\}^3}-\frac{n_0(1-\rho_{12}^2)}{n-2}<\rho_{12}^2\,n_0 \qquad (4.4.72)$$

For n'' $(=n'-1)$ = 5, 20, 25, ..., 50, ... ρ_{12}^2 = .1, ..., .4, $n''\geq 20$, $\rho_{12}^2=.5$, $n''\geq 40$, $\rho_{12}^2=.6$ and λ = .1, .2, ..., .9 and for n'' = 20, 25, 30, ..., 50, ..., ρ_{12}^2 = .6 and

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\lambda = .1, .2, \ldots, .8$ the above inequality is always satisfied. Only for $\rho_{12}^2 = .7$, $\lambda = .7$ this inequality is ~~only~~ satisfied for ~~any~~ large value of n''. ~~= 10, 15, ..., 50 (see Appendix 3).~~ In fact, for smaller values of λ, lower values of n'' will satisfy the inequality (Appendix 3).

Based on these results we conclude that there is evidence to broadly support the use of the proxy in a multiple regression on x_1 and x_2^* 9/.

Moreover, if one is able and willing to specify β_2 in addition, the equation (4.4.57) and (4.4.58) can be compared directly.10/

4.5. Conclusions

The main results of this chapter may be summarised as follows.

We have considered the problem of omitting regressors for the regression models with stochastic explanatory variables. We have

---

9/ This conclusion also easily apply to the case of k regressors discussed at the beginning of this section.

10/ In the book by Griliches and Ringsterd (1971), sufficient data available for an examination of the effect of capital measurement error on the estimates of the returns to scale with a Cobb - Douglas production function. With $\log(\frac{V}{L}) = \beta_1 \log L + \beta_2 \log (\frac{K}{L}) + u$, the upper bound of λ for log K has been taken to be 0.13 (p. 99). Then using the factor share estimates of β_2, Aigner (1974), in his paper, compared $MSE(\hat{\beta}_{1P})$ and $MSE(\hat{\beta}_1^+)$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

derived the condition under which the omitted variables (OV) estimator of a regression coefficient may be better than the OLS (f) estimator by mean square error criterion. Comparing this with the case of nonstochastic regressors, we find that the condition here becomes stringent.

Next, considering a regression model with errors-in-the variables (E - V) we have derived the condition under which the proxy variable estimator is better than the OV estimator. We obtain results analogous to but more general than those derived by Aigner (1974).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 5

HANDLING OF ERRORS IN VARIABLES MODELS WITH TRENDING OR AUTOCORRELATED ERRORS

5.1 Introduction

In the literature on econometric methods, it is shown that when the regressors in a single equation regression model are measured with error, the estimation of regression coefficients poses a serious problem. In the standard model with errors in variables, the Ordinary Least Squares (OLS) method of estimation gives inconsistent estimates of the regression coefficients. Here we can get consistent estimators either by maximum likelihood methods or by instrumental variable methods. However, for the estimators based on maximum likelihood methods, one has to make strong assumptions about the distribution of errors and also about the dispersion matrix of the measurement errors. The technique of instrumental variables can yield consistent estimates provided the instruments are suitably chosen. Here the underlying assumptions are that the instruments are highly correlated with the true values of the regressors but uncorrelated with the measurement errors in all the regressors. Wald (1940), Bartlett (1949) and Durbin (1954) have proposed different types of instruments to get consistent estimates of the regression coefficients. However, underlying assumption in all these methods is that the measurement errors are too small to affect the grouping or ranking of the regressor-values ; or, in other words, the grouping or ranking of the true values is the same as that of the observed values.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

There is extensive literature on the grouping methods like those due to Wald and Bartlett. It has been found that the efficiency of Bartletts estimator is higher than that of Wald's estimator. Theil and Van Yzeren (1954) examined how the efficiency of the Bartlett type estimator depends on the grouping of the observations. Of these three estimators mentioned above, Durbin's estimator is likely to have the highest efficiency [vide Johnston 1972,p.285].

In general, in the errors-in-variable models, errors in the regressors are assumed to be spherical with zero means. But this assumption may be violated frequently in econometric practice. The regressors may be measured in such a way that the errors entering them may contain some systematic element ; e.g., the mean of the errors may be nonzero, or the mean error may have a time trend; also the errors may form autocorrelated time series. This is illustrated by the data presented in the following table reproduced from Mukherjee and Chatterjee (1974).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table : Two series of official estimates of private consumption expenditure in India at current market prices : Rs.10^9.

Year	Revised Series	Conventional Series	Differences : col.(3)-col.(2)
(1)	(2)	(3)	(4)
1954-55	81.1	88.2	7.1
55-56	82.1	88.3	6.2
56-57	95.2	101.7	6.5
57-58	98.4	105.4	7.0
58-59	109.5	116.3	6.8
59-60	110.2	117.2	7.0
60-61	118.8	127.4	8.6
61-62	125.4	132.7	7.3
62-63	131.1	136.1	5.0
63-64	147.8	148.9	1.1
64-65	176.1	179.5	3.4
65-66	178.4	179.1	0.7
66-67	206.3	206.3	0.0
67-68	248.7	241.4	-7.3
68-69	242.4	235.9	-6.5

If we examine the differences between the revised series and the conventional series of estimates for India, both emanating from the Central Statistical Organisation, India and both based on the same kind of methodology and material, broadly speaking ; we find that these differences do not strictly follow a random series with zero mean. In fact, they reveal a clear trend over time. In the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

beginning, the revised series is lower than the conventional series, but the sign of the difference changes during the period. If the revised series be taken as true, for the sake of argument, the errors in the conventional series appear to have a clear time trend. Of course, even the revised series is not perfect, and the illustration is only to be taken as suggestive.

The aim of the present chapter is to suggest some methods of estimation for regression models with eoors-in-variables showing systematic movements in errors over time. In section 2 we have considered some models for the errors-in- variables problems where the errors contain linear or exponential trends over time and examined the efficiency of some instrumental variable (IV) estimators suggested in this paper. In section 3 we have examined the small sample bias of both the IV estimator and the OLS estimator. We have studied these in the simple case of two variable linear regression relationship. In section 4 we have briefly extended some of the results of section 2 to relationships between more than two variables. Section 5 considers some methods of estimation for situations where the errors in the variables are autocorrelated. Section 6 concludes the chapter with some general observations.

5.2 Model with trending errors

The standard two-variable errors-in-variables model (E-V-M)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

is

$$y_t^* = Y_t + v_t \tag{5.2.1}$$

$$x_t^* = x_t + u_t \tag{5.2.2}$$

where Y_t and x_t (considered stochastic) are true values, y_t^* and x_t^* the corresponding observations at time t, and v_t and u_t are random disturbances having zero means and variances σ_v^2 and σ_u^2 respectively. These disturbances are serially and mutually independent and also independent of the true values of the variables. Also,

$$Y_t = \alpha + \beta\, x_t + \varepsilon_t \tag{5.2.3}$$

where ε_t is the spherical disturbance term with variance σ_t^2 These ε_t's are independent of u, v and x. The model to be discussed here relaxes some of these assumptions and lays down that

$$u_t = c_1 + c_2 t + \tilde{u}_t \tag{5.2.4}$$

and

$$v_t = d_1 + d_2 t + \tilde{v}_t \tag{5.2.5}$$

where $\tilde{u}_t$ and $\tilde{v}_t$ are spherical errors with zero mean and variances $\sigma_{\tilde{u}}^2$ and $\sigma_{\tilde{v}}^2$ respectively. These $\tilde{u}_t$' and $\tilde{v}_t$'s are serially and mutually independent and independent of x and y. These are also independent of ε_t's. So, the relationship among the observed variables can be formed as follows :

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$
\begin{aligned}
y_t^* &= \alpha + \beta\, x_t + \varepsilon_t + v_t \\
&= \alpha + \beta(x_t^* - u_t) + \varepsilon_t + d_1 + d_2 t + \tilde{v}_t \\
&= \alpha + \beta\, x_t^* - \beta(c_1 + c_2 t) + d_1 + d_2 t + (\varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t) \\
&= (\alpha + d_1 - \beta\, c_1) + \beta\, x_t^* + (d_2 - \beta\, c_2)t + (\varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t) \\
&= \tilde{\alpha} + \beta x_t^* + \gamma t + \xi_t \text{ (say)} \qquad (5.2.6)
\end{aligned}
$$

where $\tilde{\alpha} = (\alpha + d_1 - \beta c_1)$

$\gamma = (d_2 - \beta c_2)$

and $\xi_t = (\varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t)$

5.2.1 <u>The I-V estimator</u>

From (5.2.6), we find

$$
\begin{aligned}
\operatorname{cov}(x_t^*, \xi_t) &= \operatorname{cov}(x_t + u_t, \varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t) \\
&= \operatorname{cov}(x_t + c_1 + c_2 t + \tilde{u}_t, \varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t) = -\beta\sigma^2_{\tilde{u}} \neq 0 \qquad (5.2.7)
\end{aligned}
$$

Hence, here even if one regresses y on x and t, the standard OLS method of estimation will fail to give consistent estimates of $\tilde{\alpha}$, β and γ. Maximum likelihood methods need assumptions about the distribution of the error variables. They also need strong assumptions about the covariance matrix of the measurement errors (<u>vide</u> Kendall and Stuart, Volume 2, pp.375-418).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, we may try to find some suitable instruments to get consistent estimates of the regression coefficients in model (5.2.6) by the E-V technique.[1/]

Following Durbin (1959) one obvious choice of instruments is the $n \times 3$ matrix where the j-th row is given by $(1, r(x_j^*), j)$ where $r(x_j^*)$ is the rank of x_j^* in ascending order (say). But, for this, we have to assume that the measurement errors in x^* are so small that the rank ordering is not affected by these errors.

Another choice of instruments is represented by the $n \times 3$ matrix whose j-th row is given by $(1, j, j^2)$, $j = 1, 2, \ldots, n$. Here we need not assume anything about the magnitude of measurement errors. But to achieve higher efficiency, we may exploit the serial correlation of the $\{x_t\}$ series. This was done by Reiersol (1941), for a two-variable case when the relationship is

$$y_t = \beta x_t + \varepsilon_t \qquad (5.2.8)$$

so that the relationship among the observed variables is

$$y_t^* = \beta x_t^* + (\varepsilon_t - \beta u_t) \quad (v_t \text{ merged with } \varepsilon_t) \qquad (5.2.9)$$

If x_t's are serially correlated then the lagged value of the

---

1/ If $c_2 = d_2 = 0$, the term involving t is absent from model (5.2.6). Here we get almost the standard E-V-M as given in the text books (vide Johnston 1972).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

observed x_t^*'s can be taken as instrument [2/]. So for our problem we can use the following matrix of instrumental variables

$$Z_{(1)} = \begin{pmatrix} 1 & x_1^* & 2 \\ 1 & x_2^* & 3 \\ 1 & x_3^* & 4 \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}^* & n \end{pmatrix} \qquad (5.2.10)$$

The equation (5.2.6) can be written as

$$y_t^* = (\alpha + \beta\,\bar{x}^* + \gamma\,\bar{t}) + \beta(x_t^* - \bar{x}^*) + \gamma\,(t - \bar{t}) + \xi_t$$

$$= \alpha' + \beta(x_t^* - \bar{x}^*) + \gamma\,(t - \bar{t}) + \xi_t, \quad t = 1,2,\ldots,n \qquad (5.2.11)$$

where $\bar{x} = \dfrac{\sum_{t=2}^{n} x_t}{n-1}$ and $\bar{t} = \dfrac{\sum_{t=2}^{n} t}{n-1}$.

So, the matrix of observations can be written as

$$X_{0(1)}^* = \begin{pmatrix} 1 & x_2^* - \bar{x}^* & 2 - \bar{t} \\ 1 & x_3^* - \bar{x}^* & 3 - \bar{t} \\ \vdots & \vdots & \vdots \\ 1 & x_n^* - \bar{x}^* & n - \bar{t} \end{pmatrix}$$

2/ This instrument cannot, however, be used if $\tilde{u}_t$'s are serially correlated. In that case the instrument will not be independent of the observational errors $\tilde{u}_t$'s. This case of serially correlated observational errors has been considered in section 5.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Let $\delta = (\alpha', \beta, \gamma)'$.

Let

$$y^*_{(1)} = (y^*_2, y^*_3, \ldots, y^*_n)' \quad \text{and} \quad \xi_{(1)} = (\xi_2, \xi_3, \ldots, \xi_n)'. \qquad (5.2.12)$$

The instrumental varable (I-V) estimator of δ is

$$\begin{aligned} \hat{\delta}_{I(1)} &= (Z'_{(1)} X^*_{0(1)})^{-1} (Z'_{(1)} y^*_{(1)}) \\ &= \delta + (Z'_{(1)} X^*_{0(1)})^{-1} (Z'_{(1)} \xi_{(1)}) \text{ 3/} \end{aligned} \qquad (5.2.13)$$

$$\text{Now, } (Z'_{(1)} X^*_{0(1)}) = \begin{pmatrix} n-1 & 0 & 0 \\ \sum_{t=1}^{n-1} x^*_t & c & \sum_{t=1}^{n-1} x^*_t(t+1-\bar{t}) \\ \sum_{t=1}^{n-1} (t+1) & \sum_{t=1}^{n-1} (x^*_{t+1}-\bar{x}^*)(t+1) & \sum_{t=1}^{n-1} (t+1-\bar{t})^2 \end{pmatrix} \qquad (5.2.14)$$

3/ Since t is included as a regressor in model (5.2.6) and since x^*_t also contains a trend component, it is not possible to write

$$\underset{n\to\infty}{\text{plim}}\ \hat{\delta}_{I(1)} = \delta + \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} Z'_{(1)} X^*_{0(1)}\right)^{-1} \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} Z'_{(1)} \xi_{(1)}\right)$$

because the probability limits of the terms $(\frac{1}{n} Z'_{(1)} X^*_{0(1)})^{-1}$ and $(\frac{1}{n} Z'_{(1)} \xi_{(1)})$ do not exist separately. So, to obtain the probability limit of $\hat{\delta}_{I(1)}$ we consider the term $(Z'_{(1)} X^*_{0(1)})^{-1} (Z'_{(1)} \xi_{(1)})$ as a whole.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $c = \sum_{t=1}^{n-1} x_t^*(x_{t+1}^* - \bar{x}^*)$

So,

$$\mathrm{Det}(Z'_{(1)}\ X^*_{0(1)}) = (n-1)\Big[c \sum_{t=1}^{n-1}(t+1-\bar{t})^2 - \Big\{\sum_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1)\Big\}\Big\{\sum_{t=1}^{n-1} x_t^*(t+1-\bar{t})\Big\}\Big] \tag{5.2.15}$$

So, $(\hat{\beta}_{I(1)} - \beta)$ = the 2nd element in vector

$$(Z'_{(1)}\ X^*_{0(1)})^{-1}(Z'_{(1)}\ \xi_{(1)}) \tag{5.2.16}$$

Now, $(Z'_{(1)}\ X^*_{0(1)})^{-1}$

$$= \frac{\begin{pmatrix} \Big[c \sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2 - \Big\{\sum\limits_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1)\Big\}\Big\{\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t})\Big\}\Big] & 0 & 0 \\ -\Big[\Big(\sum\limits_{t=1}^{n-1} x_t^*\Big)\sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2 - \sum\limits_{t=1}^{n-1}(t+1)\Big\{\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t})\Big\}\Big] & (n-1)\sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2 & -(n-1)\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t}) \\ \Big[\sum\limits_{t=1}^{n-1} x_t^*\Big\{\sum\limits_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1) - \sum\limits_{t=1}^{n-1}(t+1)\,c\Big]\Big. & -(n-1)\sum\limits_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1) & (n-1)c \end{pmatrix}}{\mathrm{Det}(Z'_{(1)}\ X^*_{0(1)})} \tag{5.2.17}$$

So, $(\hat{\beta}_{I(1)} - \beta)$

$$= \frac{\Big[-\Big\{\Big(\sum\limits_{t=1}^{n-1} x_t^*\Big)\sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2 - \sum\limits_{t=1}^{n-1}(t+1)\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t})\Big\}\sum\limits_{t=2}^{n}\xi_t + (n-1)\sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2\Big(\sum\limits_{t=2}^{n} x_{t-1}^*\xi_t\Big) - (n-1)\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t})\sum\limits_{t=2}^{n}(t+1)\xi_t\Big]}{(n-1)\Big[c\sum\limits_{t=1}^{n-1}(t+1-\bar{t})^2 - \Big\{\sum\limits_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1)\Big\}\Big\{\sum\limits_{t=1}^{n-1} x_t^*(t+1-\bar{t})\Big\}} \tag{5.2.18}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Now, $x_t^* = x_t + c_1 + c_2 t + \tilde{u}_t$.

So, the expression in the numerator of (5.2.18) becomes

$$-\left\{\sum_{t=1}^{n-1}(x_t+c_1+c_2t+\tilde{u}_t)\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1}(t+1)\sum_{t=1}^{n-1}(x_t+c_1+c_2t+\tilde{u}_t)(t+1-\bar{t})\right\}$$

$$\times\sum_{t=2}^{n}\xi_t + (n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2\left\{\sum_{t=2}^{n-1}(x_{t-1}+c_1+c_2t+\tilde{u}_t)\xi_{t+1}\right\}$$

$$+(n-1)\left\{\sum_{t=1}^{n-1}(x_t+c_1+c_2t+\tilde{u}_t)(t+1-\bar{t})\right\}\sum_{t=2}^{n}t\,\xi_t \qquad (5.2.19)$$

As $n \to \infty$, the expression becomes

$$-\left\{\sum_{t=1}^{n-1}x_t\sum_{t=1}^{n-1}(t+1-\bar{t})^2+\sum_{t=1}^{n-1}\tilde{u}_t\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1}x_t(t+1)\sum_{t=1}^{n-1}x_t(t+1-\bar{t})\right.$$

$$\left.-\sum_{t=1}^{n-1}\tilde{u}_t(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\right\}\sum_{t=2}^{n}\xi_t+\left\{(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2\sum_{t=2}^{n}(x_t+\tilde{u}_t)\xi_t\right\}$$

$$-(n-1)\left\{\sum_{t=1}^{n-1}x_t(t+1-\bar{t})\sum_{t=2}^{n}(t+1)\xi_t+(n-1)\sum_{t=1}^{n-1}\tilde{u}_t(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\xi_t\right\}$$

$$(5.2.20)$$

The expression in the denominator of (5.2.18) for large n

$$(n-1)\left\{c\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1}(x_{t+1}^*-\bar{x}^*)(t+1)\sum_{t=1}^{n-1}x_t^*(t+1-\bar{t})\right\}$$

$$=(n-1)\left[\sum_{t=1}^{n-1}(x_t+\tilde{u}_t)\left\{(x_{t+1}-\bar{x})+(\tilde{u}_{t+1}-\bar{\tilde{u}})\right\}\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right.$$

$$\left.-\sum_{t=1}^{n-1}\left\{(x_t-\bar{x})+(\tilde{u}_t-\bar{\tilde{u}})\right\}(t+1)\left\{\sum_{t=1}^{n-1}(x_t+u_t)(t+1-\bar{t})\right\}\right]$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Under the assumptions that

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} \begin{pmatrix} 1 & x_1+\tilde{u}_1 & \frac{2}{n} \\ 1 & x_2+\tilde{u}_2 & \frac{3}{n} \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}+\tilde{u}_{n-1} & \frac{n}{n} \end{pmatrix}' \begin{pmatrix} \xi_2 \\ \xi_3 \\ \vdots \\ \xi_n \end{pmatrix} = \underset{\sim}{0} \tag{5.2.21}$$

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} \begin{pmatrix} 1 & x_1+\tilde{u}_1 & \frac{2}{n} \\ 1 & x_2+\tilde{u}_2 & \frac{3}{n} \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}+\tilde{u}_{n-1} & \frac{n}{n} \end{pmatrix}' \begin{pmatrix} 1 & (x_1-\bar{x})+(\tilde{u}_1-\bar{\tilde{u}}) & \frac{2}{n}-\frac{\bar{t}}{n} \\ 1 & (x_2-\bar{x})+(\tilde{u}_2-\bar{\tilde{u}}) & \frac{3}{n}-\frac{\bar{t}}{n} \\ \vdots & \vdots & \vdots \\ 1 & (x_{n-1}-\bar{x})+(\tilde{u}_{n-1}-\bar{\tilde{u}}) & \frac{n}{n}-\frac{\bar{t}}{n} \end{pmatrix} \tag{5.2.22}$$

exists and is nonsingular,

and

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n} \tilde{u}_t \frac{t}{n} = 0, \tag{5.2.23}$$

dividing both the numerator and the denominator of (5.2.17) by n^5, it can be proved that

$$\underset{n\to\infty}{\text{plim}} (\hat{\beta}_{I(1)} - \beta) = 0 \tag{5.2.24}$$

Similarly, it can be proved that

$$\underset{n\to\infty}{\text{plim}} (\alpha_{I(1)} - \alpha) = 0 \tag{5.2.25}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\underset{n\to\infty}{\text{plim}} \; (\hat{\gamma}_{I(1)} - \gamma) = 0$ (5.2.26)

So, $\hat{\delta}_{I(1)}$ is a consistent estimator of δ .

5.2.1 Comparative efficiency of IV and OLS estimators

Next we consider the case where OLS gives consistent estimators, i.e., where $\tilde{u}_t = 0 \; \forall \; t$. In this case we compare the efficiency of the instrumental variable estimator of β with that of the corresponding OLS estimator. The asymptotic covariance matrix of $\hat{\delta}_{I(1)}$ is given by

$$\bar{V}_{\hat{\delta}_{I(1)}} = \lim_{n\to\infty} E\left\{(Z'_{(1)} X^*_{0(1)})^{-1} Z'_{(1)} \tilde{\varepsilon} \tilde{\varepsilon}' Z_{(1)} (X^{*\prime}_{0(1)} Z_{(1)})^{-1}\right\} \tag{5.2.27}$$

where $\tilde{\varepsilon} = \varepsilon + \tilde{v}$ and

$$(Z'_{(1)} \tilde{\varepsilon})(\tilde{\varepsilon}' Z_{(1)}) =$$

$$\begin{pmatrix} (\sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1})^2 & \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} x_t \tilde{\varepsilon}_{t+1} & \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} (t+1) \tilde{\varepsilon}_{t+1} \\ \sum_{t=1}^{n-1} x_t \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1} & \sum_{t=1}^{n-1} (x_t \tilde{\varepsilon}_{t+1})^2 & \sum_{t=1}^{n-1} x_t \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} (t+1) \tilde{\varepsilon}_{t+1} \\ \sum_{t=1}^{n-1} \frac{t+1}{n} \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1} & \sum_{t=1}^{n-1} \frac{t+1}{n} \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} x_t \tilde{\varepsilon}_{t+1} & (\sum_{t=1}^{n-1} (t+1) \tilde{\varepsilon}_{t+1})^2 \end{pmatrix} \tag{5.2.28}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$(Z'_{(1)}X^*_{0(1)}) = \begin{pmatrix} n-1 & 0 & 0 \\ \sum_{t=1}^{n-1} x^*_t & c & \sum_{t=1}^{n-1} x^*_t(t+1-\bar{t}) \\ \sum_{t=1}^{n-1} (t+1) & \sum_{t=1}^{n-1} (x^*_{t+1}-\bar{x}^*)(t+1) & \sum_{t=1}^{n-1} (t+1-\bar{t})^2 \end{pmatrix} \tag{5.2.29}$$

where $c = \sum_{t=1}^{n-1} x^*_t (x^*_{t+1} - \bar{x}^*)$.

Let $\Delta = \text{Det}(Z'_{(1)}X^*_{0(1)}) = (n-1)\{c \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \sum_{t=1}^{n-1} (x^*_{t+1}-\bar{x}^*)(t+1) \sum_{t=1}^{n-1} x^*_t(t+1-\bar{t})\}$

Let $x^T_t = x_t + c_1 + c_2 t$. (5.2.30)

Substituting $\tilde{u}_t = 0$ in (5.2.27), (5.2.28) and (5.2.29), we have

asy var$(\hat{\beta}_{I(1)})$ = the (2,2) element in $\bar{V}_{\hat{\delta}_{I(1)}}$

4/

$$= \frac{\bar{E}}{\bar{E}\Delta^2} [\{-\sum_{t=1}^{n-1} x^T_t \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x^T_t(t+1-\bar{t}) \sum_{t=1}^{n-1} (t+1)\} \times (\sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1})^2$$

$$+(n-1) \sum_{t=1}^{n-1} (t+1-\bar{t})^2 \sum_{t=1}^{n-1} x^T_t \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1}$$

$$-(n-1) \sum_{t=1}^{n-1} x^T_t(t+1-\bar{t}) \sum_{t=1}^{n-1} (t+1) \tilde{\varepsilon}_{t+1} \sum_{t=1}^{n-1} \tilde{\varepsilon}_{t+1}]$$

$$\times -\{\sum_{t=1}^{n-1} x^T_t \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x^T_t(t+1-\bar{t}) \sum_{t=1}^{n-1} (t+1)\}$$

4/ The definition of $\bar{E}$ is same as in Chapters 3 and 4.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$+\frac{\bar{E}}{\bar{E}\Delta^2}[\{-\sum_{t=1}^{n-1} x_t^T \sum_{t=1}^{n-1}(t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t}) \sum_{t=1}^{n-1}(t+1)\}(\sum_{t=1}^{n-1}\tilde{\varepsilon}_{t+1}\sum_{t=1}^{n-1} x_t^T \tilde{\varepsilon}_{t+1})$$

$$+(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2(\sum_{t=1}^{n-1} x_t^T \tilde{\varepsilon}_{t+1})^2$$

$$-(n-1)\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\tilde{\varepsilon}_{t+1}\sum_{t=1}^{n-1} x_t^T \tilde{\varepsilon}_{t+1}](n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2$$

$$+\frac{\bar{E}}{\bar{E}\Delta^2}[\{-\sum_{t=1}^{n-1} x_t^T \sum_{t=1}^{n-1}(t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\}(\sum_{t=1}^{n-1}\tilde{\varepsilon}_{t+1}\sum_{t=1}^{n-1}(t+1)\tilde{\varepsilon}_{t+1})$$

$$+(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2 \sum_{t=1}^{n-1} x_t^T \tilde{\varepsilon}_{t+1}\sum_{t=1}^{n-1}(t+1)\tilde{\varepsilon}_{t+1}$$

$$-(n-1)\sum_{t=1}^{n-1} x_t(t+1-\bar{t})(\sum_{t=1}^{n-1} t\,\tilde{\varepsilon}_{t+1})^2] \times -(n-1)\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})$$

$$=(n-1)\tilde{\sigma}^2\frac{\bar{E}}{\bar{E}\Delta^2}[\{-\sum_{t=1}^{n-1} x_t^T \sum_{t=1}^{n-1}(t+1-\bar{t})^2 + \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\}$$

$$+\{\sum_{t=1}^{n-1}(t+1-\bar{t})^2\sum_{t=1}^{n-1} x_t^T - \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\}]$$

$$\times -\{\sum_{t=1}^{n-1} x_t^T \sum_{t=1}^{n-1}(t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)\}]$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$+\frac{\tilde{\sigma}^2}{\bar{E}\,\Delta^2}\bar{E}\Big\{-\Big(\sum_{t=1}^{n-1} x_t^T\Big)^2 \sum_{t=1}^{n-1}(t+1-\bar{t})^2 + \sum_{t=1}^{n-1} x_t^T \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t}) \sum_{t=1}^{n-1}(t+1)$$

$$+(n-1)\sum_{t=1}^{n-1} x_t^{T2} \sum_{t=1}^{n-1}(t+1-\bar{t})^2 - (n-1)\sum_{t=1}^{n-1} x_t^T(t+1) \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\Big\}(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2$$

$$+\frac{\tilde{\sigma}^2}{\bar{E}\,\Delta^2}\bar{E}\Big\{-\Big(\sum_{t=1}^{n-1} x_t^T\Big)\sum_{t=1}^{n-1}(t+1-\bar{t})^2 \sum_{t=1}^{n-1}(t+1) + \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\Big(\sum_{t=1}^{n-1}(t+1)\Big)^2$$

$$+(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2 \sum_{t=1}^{n-1} x_t^T(t+1) - (n-1)\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t}) \sum_{t=1}^{n-1}(t+1)^2\Big\}$$

$$\times -(n-1)\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})$$

[where $\tilde{\sigma}^2 = \mathrm{var}(\varepsilon_t)\ \forall\ t$]

$$= \frac{\tilde{\sigma}^2}{\bar{E}\,\Delta^2}\bar{E}\Big[\Big\{-\Big(\sum_{t=1}^{n-1} x_t^T\Big)^2 \sum_{t=1}^{n-1}(t+1-\bar{t})^2 + n\sum_{t=1}^{n-1} x_t^{T2} \sum_{t=1}^{n-1}(t+1-\bar{t})^2$$

$$-n\sum^{n-1} x_t^T(t+1-\bar{t}) \sum_{t=1}^{n-1}(x_t^T - \bar{x}^{T'})(t+1-\bar{t})\Big\}(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2$$

$$+\Big\{(n-1)\sum_{t=1}^{n-1}(t+1-\bar{t})^2 \sum_{t=1}^{n-1}(x_t^T - \bar{x}^{T'})(t+1-\bar{t})$$

$$-(n-1)\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t}) \sum_{t=1}^{n-1}(t+1-\bar{t})^2\Big\} \times -(n-1)\sum_{t=1}^{n-1}(x_t^T(t+1-\bar{t})\Big]$$

where $\bar{x}^{T'} = \dfrac{\sum_{t=1}^{n} x_t^T}{n-1}$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$= \frac{\tilde{\sigma}^2 (n-1)^2}{\bar{E}\ \Delta^2} \bar{E}[\{\sum_{t=1}^{n-1} (x_t^T - \bar{x}^{T'})^2 \sum_{t=1}^{n-1} (t+1-\bar{t})^2\} - \{\sum_{t=1}^{n-1} (x_t^T - \bar{x}^{T'})(t+1-\bar{t})\}^2] \times \sum_{t=1}^{n-1} (t+1-\bar{t})^2$$

$$= \frac{\sigma^2 (n-1)^2 \bar{E}[\sum_{t=1}^{n-1} \{(x_t - \bar{x}') + c_2(t-\bar{t}')\}^2 \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \{(\sum_{t=1}^{n-1} (x_t - \bar{x}') + c_2(t-\bar{t}'))(t+1-\bar{t})\}^2] \sum_{t=1}^{n-1} (t+1-\bar{t})^2}{(n-1)^2 \bar{E}[\sum_{t=1}^{n-1} \{x_t(x_{t+1} - \bar{x}) + c_2^2 t(t+1-\bar{t})\} \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \{(\sum_{t=1}^{n-1} (x_t - \bar{x}') + c_2(t+1-\bar{t})\}(t+1) \sum_{t=1}^{n-1} (x_t + c_2 t)(t+1-\bar{t})\}]^2}$$

where $\bar{t}' = (\sum_{t=1}^{n-1} t)/_{n-1}$ (5.2.31)

Under the assumption that $\{x_t\}$ is stationary,

$$\text{asy var}(\hat{\beta}_{I(1)}) = \frac{\tilde{\sigma}^2\ \bar{E} \sum_{t=1}^{n-1} (x_t - \bar{x})^2 \{\sum_{t=1}^{n-1} (t+1-\bar{t})^2\}^2}{\bar{E}\{\sum_{t=1}^{n-1} x_t (x_{t+1} - \bar{x}) \sum_{t=1}^{n-1} (t+1-\bar{t})^2\}^2}$$

$$= \frac{\tilde{\sigma}^2}{n\ \rho_1^2 \theta^2} \qquad (5.2.32)$$

where $\rho_1 = \bar{E} \dfrac{\sum_{t=1}^{n-1} (x_t - \bar{x}')(x_{t+1} - \bar{x}')}{\sum_{t=1}^{n-1} (x_t - \bar{x}')^2}$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\theta^2 = \bar{E}\ \frac{1}{n} \sum_{t=1}^{n-1} (x_t - \bar{x}')^2.$

Again, we know that the asymptotic variance of the OLS estimator of δ is given by

$$\bar{V}(\hat{\delta}_{OLS}) = \frac{\sigma^2}{n} \left\{ \bar{E} \left(\frac{1}{n} X_o^{T'} X_o^T\right) \right\}^{-1} \qquad (5.2.33)$$

where $X_0^T = \begin{pmatrix} 1 & x_1^T - \bar{x}''^T & 1-\bar{t}'' \\ 1 & x_2^T - \bar{x}''^T & 2-\bar{t}'' \\ \vdots & \vdots & \vdots \\ 1 & x_n^T - \bar{x}''^T & n-\bar{t}'' \end{pmatrix}$, where $\bar{x}''^T = \frac{\sum_{t=1}^{n} x_t}{n}$, $\bar{t}''^T = \frac{\sum_{t=1}^{n} t}{n}$

$\bar{E}(\frac{1}{n} X_o^{T'} X_o^T)$ is assumed to exist and is nonsingular.

$$\text{So, } \bar{V}(\hat{\beta}_{OLS}) = \tilde{\sigma}^2 \bar{E} \frac{n \sum_{t=1}^{n-1} (t - \bar{t}'')^2}{n \sum_{t=1}^{n} (x_t^T - \bar{x}''^T)^2 \sum_{t=1}^{n} (t - \bar{t}'')^2 - \sum_{t=1}^{n} (x_t^T - \bar{x}''^T)(t - \bar{t}'')\}^2} \qquad (5.2.34)$$

Under the assumption that $\{x_t\}$ is stationary, it can be shown that the expression in (5.2.34) becomes

$$\frac{\tilde{\sigma}^2}{n} \frac{1}{\bar{E}\ \frac{1}{n} \sum_{t=1}^{n} (x_t - \bar{x}'')^2} = \frac{\sigma^2}{n\ \theta^2} \qquad (5.2.35)$$

So, assymptotic efficiency E_1 of the instrumental variable estimator of β with respect to that of OLS estimator is ρ_1^2

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.2.2 Exponential trend in the error component

In the model considered so far, the error components (u_t and v_t) in x_t and y_t are assumed to contain linear trends. But, when sample size (n) is large, this assumption may not be correct. Because, this implies that as n increases, the systematic parts of v_t and u_t increase or decrease without limit and as a result the numerical values of errors go on increasing indefinitely.

To do away with this difficulty, a more realistic assumption is that the systematic parts of u_t and v_t decrease to zero with time. For example, suppose

$$u_t = A_u \exp[-d_u t] + \tilde{u}_t \tag{5.2.36}$$

$$v_t = A_v \exp[-d_v t] + \tilde{v}_t \tag{5.2.37}$$

The true model is given by

$$Y_t = \alpha + \beta x_t + \varepsilon_t \tag{5.2.38}$$

The relation between x_t^* and y_t^* is

$$y_t^* = \alpha + \beta x_t^* - \gamma e^{-d_u t} + \delta e^{-d_v t} + \xi_t \tag{5.2.39}$$

where $\gamma = \beta A_u$

$\delta = A_v$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $\xi_t = \varepsilon_t + \tilde{v}_t - \beta\, \tilde{u}_t$

Although this model may be more realistic than the previous one for practical purposes, it is not easily estimable unless d_u and d_v are known. So, let us assume that d_u and d_v are known[5]. Moreover, for convenience, let us also assume that $d_u = d_v = d$ (The general case can be treated without much additional difficulty). The model in (5.2.39) now becomes

$$y = \alpha + \beta x_t^* + \tilde{\gamma}\, e^{-dt} + \xi_t \qquad (5.2.40)$$

where $\tilde{\gamma} = (\gamma - \delta)$

The matrix of instrumental variables can be taken as

$$\tilde{Z}_{(1)} = \begin{pmatrix} 1 & x_1^* & 2 \\ 1 & x_2^* & 3 \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}^* & n \end{pmatrix} \qquad (5.2.41)$$

Again, the model (5.2.39) can be written as

$$y_t = (\alpha + \beta\, \bar{x}^* + \tilde{\gamma}\, m) + \beta(x_t^* - \bar{x}^*) + \tilde{\gamma}\,(e^{-dt} - m) + \xi_t$$

$$= \alpha' + \beta(x_t^* - \bar{x}^*) + \tilde{\gamma}\,(e^{-dt} - m) + \xi_t \;;\; t = 1,2,3,\ldots,n$$

(5.4.42)

[5] One can suggest search procedures for estimating d_u and d_v; but it would be difficult to prove that they are reasonable.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\bar{x}^* = \frac{\sum_{t=2}^{n} x_t^*}{n-1} = m = \frac{\sum_{t=2}^{n} e^{-dt}}{n-1}$.

So, the matrix of observations can be written as

$$\tilde{X}^*_{0(1)} = \begin{pmatrix} 1 & x_2^* - \bar{x}^* & (e^{-2d} - m) \\ 1 & x_3^* - \bar{x}^* & (e^{-3d} - m) \\ \vdots & \vdots & \vdots \\ 1 & x_n^* - \bar{x}^* & (e^{-nd} - m) \end{pmatrix}$$

Then, the IV estimator is

$$\hat{\delta}_{I(1')} = (Z'_{(1)} \tilde{X}^*_{0(1)})^{-1} (Z'_{(1)} y_{(1)}) \qquad (5.2.43)$$

As in the case of $\hat{\delta}_{I(1')}$, here also, it can be proved that

$$\underset{n\to\infty}{\text{plim}} \ \hat{\delta}_{I(1')} = \delta = (\alpha', \beta, \tilde{\gamma}) \qquad (5.2.44)$$

As in the previous case, if $\tilde{u}_t = 0$ (i.e., OLS is consistent), it can be shown that the asymptotic efficiency of $\beta_{I(1')}$ relative to that of $\hat{\beta}_{OLS}$ for the model (5.2.39) is given by

$$E_1 = \rho_1^2 \text{ (provided } \{x_t\} \text{ is stationary)} \qquad (5.2.45)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.2.3 An alternative IV estimator

Recently, Karni and Weissman (1974) have shown that in the standard E-V-M, when $\{x_t\}$ is serially correlated, then using $(x_{t+1} + x_{t-1})$ instead of x_{t+1} or x_{t-1} as instrument, we get more efficient estimates of the regression coefficients. We exploit this idea to suggest an alternative choice of the matrix of instrumental variables.

Let us consider the model (5.2.12). The matrix of instrumental variable may be

$$Z_{(2)} = \begin{pmatrix} 1 & x_3^* + x_1^* & 2 \\ 1 & x_4^* + x_2^* & 3 \\ \vdots & \vdots & \vdots \\ 1 & x_n^* + x_{n-2}^* & (n-1) \end{pmatrix} \qquad (5.2.46)$$

The matrix of observations can be written as

$$X_{0(2)}^* = \begin{pmatrix} 1 & x_2^* - \bar{x}^* & 2 - \bar{t} \\ 1 & x_3^* - \bar{x}^* & 3 - \bar{t} \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}^* - \bar{x}^* & n-1-\bar{t} \end{pmatrix} \qquad (5.2.47)$$

where, now $\bar{x}^* = \dfrac{\sum_{t=2}^{n-1} x_t^*}{n-2}$ and $\bar{t} = \dfrac{\sum_{t=2}^{n-1} t}{n-2}$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $\hat{\delta}_{I(2)} = (Z'_{(2)} X^*_{0(2)})^{-1} (Z'_{(2)} y^*_{(2)})$ (5.2.48)

is the IV estimator of δ ; where $y^*_{(2)} = (y^*_2, y^*_3 \ldots y^*_{n-1})'$.
As in section (5.2.1), here also, it can be proved similarly that $\hat{\delta}_{I(2)}$ is a consistent estimator of δ.

Under the assumption that $\{x_t\}$ is stationary, it can be shown that

$$\text{asy } V(\hat{\beta}_{I(2)}) = \frac{\sigma^2}{n} \frac{\bar{E} \sum_{t=2}^{n-1} \{(x_{t+1}-\bar{x}_+) + (x_{t-1} - \bar{x}_-)\}^2}{\bar{E} \sum_{t=2}^{n-1} [\{(x_{t+1}-\bar{x}_+) + (x_{t-1} - \bar{x}_-\} x_t]^2} \tag{5.2.49}$$

where $\bar{x}_+ = \dfrac{\sum_{t=2}^{n-1} x_{t+1}}{n-2}$ and $\bar{x}_- = \dfrac{\sum_{t=2}^{n-1} x_{t-1}}{n-2}$.

The expression on the right hand side of (5.2.49) is

$$\frac{\sigma^2}{n} \frac{2(1+\rho_2)}{\rho_1^2 \theta^2} \tag{5.2.50}$$

where $\rho_2 = \dfrac{\bar{E} \sum_{t=2}^{n-2} [(x_{t+1} - \bar{x}_+)(x_{t-1} - \bar{x}_-)}{\theta^2}$

θ^2 and ρ_1 have been defined in (5.2.32).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, the asymptotic efficiency of $\hat{\beta}_{I(2)}$ with respect to that of the OLS estimator of β is

$$E_2 = 2\rho_1^2/(1+\rho_2) > E_1 = \rho_1^2 \qquad (5.2.51)$$

We can obtain similar results for the model (5.2.40) also.

5.3 Small sample bias of I-V and OLS estimators

In the previous section we have examined the efficiency of some I-V estimators of β in models (5.2.6) and (5.2.40) relative to corresponding OLS estimators in the special case where the OLS estimator will be consistent.

It is well known that asymptotic results may not be a perfectly reliable indicator of small sample performance (e.g., small sample properties of TSLS, LIML and OLS studied through Monte-Carlo experiments in simultaneous equation systems). It is therefore of some interest to examine the exact (small sample) bias of the I-V and the OLS estimators of β in models (5.2.6) and (5.2.40). We do this in the present section. Here we have also obtained the bias of the corresponding OLS estimator when the time variable has been omitted from the models (5.2.6) and (5.2.40)(i.e.,y^* is regressed on x^* alone). Although we have been able to obtain the exact bias of the OLS estimator (both including and excluding t), for the I-V estimator, we have been able to evaluate the bias term only upto order $\frac{1}{n}$(n being the sample size).

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.3.1 Small sample bias of IV estimator

First let us consider the model (5.2.6). From (5.2.13),

$$\hat{\delta}_{I(1)} = (Z'_{(1)} X^*_{0(1)})^{-1} Z'_{(1)} y_{(1)}$$

$$= \delta + (Z'_{(1)} X^*_{0(1)})^{-1} Z'_{(1)} \xi_{(1)} .$$

Now, $(Z'_{(1)} X^*_{0(1)})^{-1} Z'_{(1)} \xi$

$$= [(Z'^{T}_{(1)} + \bar{\eta}') (X^T_{0(1)} + \eta)]^{-1} (Z^T_{(1)} + \bar{\eta}') \xi_{(1)} \tag{5.3.1}$$

where $X^T_{0(1)} = \begin{pmatrix} 1 & x^T_2 - \bar{x}^T & 2 - \bar{t} \\ 1 & x^T_3 - \bar{x}^T & 3 - \bar{t} \\ \vdots & \vdots & \vdots \\ 1 & x^T_n - \bar{x}^T & n - \bar{t} \end{pmatrix}$

and $x^T_t = x_t + c_1 + c_2 t$.

$$\eta = \begin{pmatrix} 0 & \tilde{u}_2 - \bar{\tilde{u}} & 0 \\ 0 & \tilde{u}_3 - \bar{\tilde{u}} & 0 \\ \vdots & \vdots & \vdots \\ 0 & \tilde{u}_n - \bar{\tilde{u}} & 0 \end{pmatrix}, \quad \bar{\tilde{u}} = \frac{\sum_{i=2}^{n} \tilde{u}_i}{n - 1}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$Z^T_{(1)} = \begin{pmatrix} 1 & x_1^T & 2 \\ 1 & x_2^T & 3 \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1}^T & n \end{pmatrix} \quad \text{and} \quad \bar{\eta} = \begin{pmatrix} 0 & u_1 & 0 \\ 0 & u_2 & 0 \\ \vdots & \vdots & \vdots \\ 0 & u_{n-1} & 0 \end{pmatrix}$$

Now, the r.h.s. of (5.3.1) can be written as

$$(Z^{T'}_{(1)} X^T_{0(1)})^{-1} [I + (Z^{T'}_{(1)} \eta)(Z^{T'}_{(1)} X^T_{0(1)})^{-1} + \bar{\eta}' X^T_{0(1)} (Z^T_{(1)} X^T_{0(1)})^{-1}$$

$$+ \bar{\eta}'\eta (Z^{T'}_{(1)} X^T_{0(1)})^{-1}]^{-1} (Z^{T'}_{(1)} + \bar{\eta}') \xi_{(1)} \tag{5.3.2}$$

Expanding the expression in the third bracket of (5.3.2) as $(I + A)^{-1}$, we have the expression in (5.3.2) as

$$(Z^{T'}_{(1)} X^T_{0(1)})^{-1} [I - (Z^{T'}_{(1)} \eta)(Z^{T'}_{(1)} X^T_{0(1)})^{-1} - \bar{\eta}' X^T_{0(1)} (Z^{T'}_{(1)} X^T_{0(1)})^{-1}$$

$$- \bar{\eta}' \eta (Z^{T'}_{(1)} X^T_{0(1)})^{-1} + \text{other terms}] (Z^{T'}_{(1)} + \bar{\eta}') \xi_{(1)} \tag{5.3.3.}$$

So,

$$E(Z'_{(1)} X^*_{0(1)})^{-1} Z'_{(1)} \xi_{(1)} = -E[(Z^{T'}_{(1)} X^T_{0(1)})^{-1} (Z^{T'}_{(1)} \eta)(Z^{T'}_{(1)} X^T_{0(1)})^{-1} \times (Z^{T'}_{(1)} \xi_{(1)})]$$

$$- E[(Z^{T'}_{(1)} X^T_{0(1)})^{-1} \bar{\eta}' \eta (Z^{T'}_{(1)} X^T_{0(1)})^{-1} \bar{\eta}' \xi_{(1)}]$$

$$+ \text{other terms} \tag{5.3.4}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Here the underlying assumption is that $\tilde{u}_i$'s are mutually independent and have symmetric distribution.

$$\text{Now,}\quad \bar{\eta}'\eta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sum_{t=1}^{n-1} \tilde{u}_t(\tilde{u}_{t+1}-\bar{\tilde{u}}) & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{and}\quad \bar{\eta}'\xi = \begin{pmatrix} 0 \\ \sum_{t=1}^{n-1} u_t \xi_{t+1} \\ 0 \end{pmatrix}$$

$$(Z^{T'}_{(1)} X^T_{0(1)}) = \begin{pmatrix} n-1 & 0 & 0 \\ \sum_{t=1}^{n-1} x_t^T & \sum_{t=1}^{n-1} x_t^T(x^T_{t+1}-\bar{x}^T) & \sum_{t=1}^{n-1} x_t^T(t+1-\bar{t}) \\ \sum_{t=1}^{n-1}(t+1) & \sum_{t=1}^{n-1}(t+1)(x^T_{t+1}-\bar{x}^T) & \sum_{t=1}^{n-1}(t+1)(t+1-\bar{t}) \end{pmatrix}$$

$$= \begin{pmatrix} a & 0 & 0 \\ b & c & d \\ e & f & g \end{pmatrix} \text{ say}$$

$$\text{So,}\quad (Z'^T_{(1)} X^T_{0(1)})^{-1} = \begin{pmatrix} cg-fd & 0 & 0 \\ -(bg-ed) & ag & -ad \\ bf-ec & -af & ac \end{pmatrix} \Bigg/ a(cg-fd)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$(Z^{T'}_{(1)}\xi_{(1)}) = \begin{pmatrix} \sum_{t=1}^{n-1} \xi_{t+1} \\ \sum_{t=1}^{n-1} x_t^T \xi_{t+1} \\ \sum_{t=1}^{n-1} (t+1)\, \xi_{t+1} \end{pmatrix}$$

$$Z'^{T}_{(1)} \eta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sum_{t=1}^{n-1} x_t^T(\tilde{u}_{t+1} - \bar{\tilde{u}}) & 0 \\ 0 & \sum_{t=1}^{n-1} (t+1)(\tilde{u}_{t+1} - \bar{\tilde{u}}) & 0 \end{pmatrix}$$

So, $(Z^{T'}_{(1)}\, X^T_{0(1)})^{-1}\, (Z^{T'}_{(1)}\, \eta)\, (Z^{T'}_{(1)}\, X^T_{0(1)})^{-1}\, (Z'^{T}_{(1)}\, \xi_{(1)})$

$$= \frac{\begin{pmatrix} 0 \\ \nu(-bg+ed) \sum_{t=1}^{n-1} \xi_{t+1} + \nu ag \sum_{t=1}^{n-1} x_t^T \xi_{t+1} - \nu ad \sum_{t=1}^{n-1} x_t^T \xi_{t+1} \\ \tilde{\nu}(-bg+ed) \sum_{t=1}^{n-1} \xi_{t+1} + \tilde{\nu}\, ag \sum_{t=1}^{n-1} x_t^T \xi_{t+1} - \tilde{\nu}\, ad \sum_{t=1}^{n-1} (t+1)\xi_{t+1} \end{pmatrix}}{a^2(cg - fd)^2} \qquad (5.3.5)$$

where $\nu = ag \sum_{t=1}^{n-1} x_t^T(\tilde{u}_{t+1} - \bar{\tilde{u}}) - ad \sum_{t=1}^{n-1} (t+1)(\tilde{u}_{t+1} - \bar{\tilde{u}})$

$$\tilde{\nu} = -af \sum_{t=1}^{n-1} x_t^T(\tilde{u}_{t+1} - \bar{\tilde{u}}) + ac \sum_{t=1}^{n-1} (t+1)(\tilde{u}_{t+1} - \bar{\tilde{u}})$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(\sqrt{\ } ag \sum_{t=1}^{n-1} x_t^T \xi_{t+1}) / a^2(cg - fd)^2$$

$$= -\beta\sigma_{\hat{u}}^2 E \left[\frac{\sum_{t=1}^{n-1} (x_t^T - \bar{x}'^T)^2 \left\{ \sum_{t=1}^{n-1} (t+1-\bar{t})^2 \right\}^2 - \sum_{t=1}^{n-1} (t+1)(x_t^T - \bar{x}'^T) \times \sum_{t=1}^{n-1} (t+1)(x_t^T - \bar{x}'^T) \sum_{t=1}^{n-1} (t+1-\bar{t})^2}{\left\{ \sum_{t=1}^{n-1} x_t^T (x_{t+1}^T - \bar{x}^T) \sum_{t=1}^{n-1} (t+1-\bar{t})^2 - \sum_{t=1}^{n-1} x_t^T (t+1-\bar{t}) \sum_{t=1}^{n-1} (x_{t+1}^T - \bar{x}^T)(t+1) \right\}^2} \right] \quad (5.3.8)$$

Now,

$$\sum_{t=1}^{n-1} (x_t^T - \bar{x}'^T)^2 = \sum_{t=1}^{n-1} (x_t - \bar{x}')^2 + c_2^2 \sum_{t=1}^{n-1} (t - \bar{t}')^2 + 2c_2 \sum_{t=1}^{n-1} (x_t - \bar{x}')(t - \bar{t}') \quad (5.3.9)$$

where $\bar{t}' = \dfrac{\sum_{t=1}^{n-1} t}{n-1}$

$$\left\{ \sum_{t=1}^{n-1} (t+1)(x_t^T - \bar{x}^T) \right\}^2 = \left\{ \sum_{t=1}^{n-1} (t+1)(x_t - \bar{x}') + c_2 \sum_{t=1}^{n-1} (t+1)(t - \bar{t}') \right\}^2 \quad (5.3.10)$$

$$\sum_{t=1}^{n-1} x_t^T (x_{t+1}^T - \bar{x}^T) = \sum_{t=1}^{n-1} x_t (x_{t+1} - \bar{x}) + c_2^2 \sum_{t=1}^{n-1} t(t+1-\bar{t}) \quad (5.3.11)$$

$$\sum_{t=1}^{n-1} x_t^T (t+1-\bar{t}) = \sum_{t=1}^{n-1} (x_t^T - \bar{x}^T) t \quad (5.3.12)$$

and

$$\sum_{t=1}^{n} (x_{t+1}^T - \bar{x}^T)(t+1) = \sum_{t=1}^{n-1} (x_{t+1} - \bar{x})(t+1) + c_2^2 \sum_{t=1}^{n-1} (t+1-\bar{t})(t+1) \quad (5.3.13)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, the expression on the r.h.s. of (5.3.7) becomes

$$-\beta\sigma_u^2 E\left[\frac{\sum_{t=1}^{n-1}(x_t-\bar{x}')^2\left\{\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}^2-\left\{\sum_{t=1}^{n-1}(t+1)(x_t-\bar{x}')\right\}^2\sum_{t=1}^{n-1}(t+t-\bar{t})^2}{\left\{\sum_{t=1}^{n-1}x_t(x_{t+1}-\bar{x})\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1}(t+1)(x_t-\bar{x}')\sum_{t=1}^{n-1}(x_{t+1}-\bar{x})\right\}^2}\right]$$

$$=\frac{-\beta\sigma_u^2\left[\frac{1}{n}\sum_{t=1}^{n-1}(x_t-\bar{x}')^2\left\{\frac{1}{n^3}\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}^2-\left\{\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1)(x_t-\bar{x}')\right\}^2\left\{\frac{1}{n^3}\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}\right]}{\left\{\frac{1}{n}\sum_{t=1}^{n-1}x_t(x_{t+1}-\bar{x})\frac{1}{n^3}\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1)(x_t-\bar{x}')\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1)(x_{t+1}-\bar{x})\right\}^2}$$

$$= B_1 \quad \left(\text{of order } \frac{1}{n}\right) \tag{5.3.14}$$

$$E(-ad)\sum_{t=1}^{n-1} t\,\varepsilon_{t+1}$$

$$= -\beta\,\sigma_u^2 E\left[\frac{\left\{\sum_{t=1}^{n-1}(x_t^T-\bar{x}'^T)(t+1)ag-\sum_{t=1}^{n-1}(t+1-\bar{t})^2 ad\right\}\times -ad}{a^2(cg-fd)^2}\right]$$

$$-\beta\sigma_u^2 E\left[\frac{\left\{\sum_{t=1}^{n-1}x_t^T(t+1-\bar{t})\right\}^2\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\left\{\sum_{t=1}^{n-1}x_t^T(t+1-\bar{t})\right\}\sum_{t=1}^{n-1}(t+1-\bar{t})^2}{a^2(cg-fd)^2}\right]$$

$$= 0 \tag{5.3.15}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Again, it can be shown easily that the second element in the vector

$$E[(Z^{T'}_{(1)}\ X^{T}_{0(1)})^{-1}\ \bar{\eta}'\eta\ (X^{T'}_{0}\ Z^{T}_{(1)})^{-1}\ \bar{\eta}'\xi] \quad \text{is}$$

$$E\left[\frac{(n-1)^2\left\{\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}^2 \times E\left\{\sum_{t=1}^{n-1} u_t(u_{t+1}-\bar{u})\sum_{t=1}^{n-1} u_t\,\xi_{t+1}\right.}{(n-1)^2\left\{\sum_{t=1}^{n-1} x_t^T(x_{t+1}^T-\bar{x}^T)\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1} x_t^T(t+1-\bar{t})\sum_{t=1}^{n-1}(t+1)(x_{t+1}^T-\bar{x}^T\right\}^2}\right]$$

$$= -(n-2)\beta\sigma_{\tilde{u}}^4 E\left[\frac{\left\{\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}^2}{\left\{\sum_{t=1}^{n-1} x_t(x_{t+1}-\bar{x})\sum_{t=1}^{n-1}(t+1-\bar{t})^2-\sum_{t=1}^{n-1} x_t(t+1-\bar{t})\sum_{t=1}^{n-1}(x_{t+1}-\bar{x})(t+1)\right\}^2}\right]$$

$$= -(n-2)\beta\sigma_{\tilde{u}}^4 E\left[\frac{\left\{\frac{1}{n^3}\sum_{t=1}^{n-1}(t+1-\bar{t})^2\right\}^2}{\left\{\frac{1}{n}\sum_{t=1}^{n-1} x_t(x_{t+1}-\bar{x})\cdot\frac{1}{n^3}\sum_{t=1}^{n-1}(t+1-\bar{t})^2 - \frac{1}{n^2}\sum_{t=1}^{n-1} x_t(t+1-\bar{t})\cdot\frac{1}{n^2}\sum_{t=1}^{n-1}(x_{t+1}-\bar{x})(t+1)\right\}^2}\right]$$

$$= -B_2(\text{of order } \tfrac{1}{n}). \tag{5.3.16}$$

So, $$E(\hat{\beta}_{I(1)}-\beta) = -B_1+B_2+\text{other terms.} \tag{5.3.17}$$

It can be proved easily that other terms will be of order lower than $\frac{1}{n}$.

Let us now consider the model (5.2.40). Here, proceeding in the same way as in the previous case, it can be shown easily

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

that the expression in (5.3.7) is again 0. The expression in (5.3.14) now becomes

$$
-\frac{\beta\sigma_u^2}{n}E\left[\frac{\frac{1}{n}\sum_{t=1}^{n-1}(x_t-\bar{x}')^2\left\{\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1-\bar{t})(e^{-d(t+1)}-m)\right\}^2-\left\{\frac{1}{n}\sum_{t=1}^{n-1}(e^{-d(t+1)}-m)(x_t-\bar{x}')\right\}^2\left\{\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1-\bar{t})(e^{-d(t+1)}-m)\right\}}{\left\{\frac{1}{n}\sum_{t=1}^{n-1}(x_{t+1}-\bar{x})\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1-\bar{t})(e^{-d(t+1)}-m)-\frac{1}{n}\sum_{t=1}^{n-1}(e^{-d(t+1)}-m)(x_t-\bar{x}')\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1)(x_{t+1}-\bar{x}')\right\}^2}\right]
$$

$= B_{1'}$ (is of order $\frac{1}{n}$) (5.3.18)

The expression in (5.3.16) becomes

$$
-\frac{(n-2)}{n^2}\beta\sigma_u^2 E\left[\frac{\left\{\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1-\bar{t})(e^{-d(t+1)}-m)\right\}^2}{-\left\{\frac{1}{n}\sum_{t=1}^{n-1}x_t(x_{t+1}-\bar{x})\frac{1}{n^2}\sum_{t=1}^{n-1}(t+1-\bar{t})(e^{-d(t+1)}-m)-\frac{1}{n}\sum_{t=1}^{n-1}x_t(e^{-d(t+1)}-m)\frac{1}{n}\sum_{t=1}^{n-1}(x_{t+1}-\bar{x})(t+1)\right\}^2}\right]
$$

$= -B_{2'}$ (is of order $\frac{1}{n}$) (5.3.19)

So, $E(\hat{\beta}_{I(1')}-\beta) = -B_{1'} + B_{2'}$ + terms of order lower than $\frac{1}{n}$

(5.3.20)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

5.3.2 Small sample bias of OLS estimator

(a) Let us now compute the small sample bias of the OLS estimator when t is included as regressor. So, the model considered is (5.2.6).

$$\hat{\delta}_{OLS} = (X_o^{*\prime} X_o^*)^{-1} X_o^{*\prime} y \tag{5.3.21}$$

$$\text{Now, } (X_o^{*\prime} X_o^*)^{-1} = \begin{pmatrix} BD-C^2 & 0 & 0 \\ 0 & AD & -AC \\ 0 & -AC & AB \end{pmatrix} \Big/ A(BD-C^2) \tag{5.3.22}$$

where $A = n$, $B = \sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime})^2$, $C = \sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime})(t-\bar{t}')$,

$D = \sum_{t=1}^{n} (t-\bar{t}'')^2$, $\bar{x}'' = \frac{\sum_{t=1}^{n} x_t^*}{n}$ and $\bar{t}'' = \frac{\sum_{t=1}^{n} t}{n}$

$$\text{So, } \hat{\delta}_{OLS} = \frac{\sum_{t=1}^{n} (t-\bar{t}'')^2 \sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime}) y_t^* - \sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime})(t-\bar{t}') \sum_{t=1}^{n} (t-\bar{t}'') y_t^*}{\sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime})^2 \sum_{t=1}^{n} (t-\bar{t}'')^2 - \left\{ \sum_{t=1}^{n} (x_t^* - \bar{x}^{*\prime\prime})(t-\bar{t}'') \right\}^2} \tag{5.3.23}$$

$\therefore E(\hat{\beta}_{OLS} | x_1^*, x_2^*, \ldots, x_n^* ; x_1^T, x_2^T, \ldots, x_n^T)$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$= \frac{\left\{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^*-\bar{x}^{*''})(\beta x_t^T+\gamma t) - \sum_{t=1}^{n}(x_t^*-\bar{x}''^*)(t-\bar{t}'')\left\{\sum_{t=1}^{n}(t-\bar{t}'')(\beta x_t^T+\gamma t)\right\}\right.}{\sum_{t=1}^{n}(x_t^*-\bar{x}''^*)^2 \sum_{t=1}^{n}(t-\bar{t}'')^2 - \left\{\sum_{t=1}^{n}(x_t^*-\bar{x}^{*''})(t-\bar{t}'')\right\}^2}$$

$$= \frac{\beta\left\{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^*-\bar{x}''^*)x_t^T - \sum_{t=1}^{n}(x_t^*-\bar{x}''^*)(t-\bar{t}'')\sum_{t=1}^{n}(t-\bar{t}'')x_t^T\right\}}{\sum_{t=1}^{n}(x_t^*-\bar{x}''^*)^2 \sum_{t=1}^{n}(t-\bar{t}'')^2 - \left\{\sum_{t=1}^{n}(x_t^*-\bar{x}^{*''})(t-\bar{t}'')\right\}^2} \quad (5.3.24)$$

Let $\lambda_t = \dfrac{x_t^* - x_t^T}{\sigma_u}$, $t = 1, 2, \ldots, n$.

Now, let us consider the following orthogonal transformation of $u_1, u_2, \ldots, u_n$ to $w_1, w_2, \ldots, w_n$, where

$$w_1 = \frac{1}{\sqrt{n}}(\lambda_1 + \lambda_2 + \ldots + \lambda_n) \quad (5.3.25)$$

$$w_2 = \frac{\sum_{t=1}^{n}(t-\bar{t}'')\lambda_t}{\sqrt{\sum_{t=1}^{n}(t-\bar{t}'')^2}} \quad (5.3.26)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$w_3 = \frac{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)\lambda_t - \sum_{t=1}^{n}(t-\bar{t}'')x_t^T \sum_{t=1}^{n}(t-\bar{t}'')\lambda_t}{\sqrt{\sum_{t=1}^{n}(t-\bar{t}'')^2}\sqrt{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2 - \left\{\sum_{t=1}^{n}(t-\bar{t}'')x_t^T\right\}^2}} \tag{5.3.27}$$

w_4

$\vdots$

w_n

Under the assumption that u_i's are i.i.d. $N(0, \sigma_u^2)$, each of w_i's $i = 1,2,\ldots,n$ is $N(0,1)$ and w_i's are mutually independent.

It can be easily shown that

$$E(\hat{\beta}_{OLS}|x_1^T, x_2^T, \ldots, x_n^T, x_1^*, x_2^*, \ldots, x_n^*)$$

$$= \beta \frac{(w_3 + \sqrt{\tau})\sqrt{\tau}}{(w_3 + \sqrt{\tau})^2 + \sum_{i=4}^{n} w_i^2} \tag{5.3.28}$$

$$\text{where } \tau = \frac{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^T - \bar{x}''^T)^2 - \left\{\sum_{t=1}^{n}(t-\bar{t}'')x_t^T\right\}^2}{\sum_{t=1}^{n}(t-\bar{t}'')^2 \ \sigma_u^2} \tag{5.3.29}$$

and $x_t^T = x_t + c_1 + c_2 t$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Gurian and Halperin (1971)[6] obtained

$$E\left\{ \frac{(w_2' + \sqrt{\tau_o})\sqrt{\tau_o}}{(w_2' + \sqrt{\tau_o})^2 + \sum_{i=4}^{n} w_i'^2} \right\} = \frac{\tau_o}{n-1} e^{-\tau_o/2} {}_1F_1\left(\frac{n-1}{2}, \frac{n-1}{2} + 1, \tau_o/2\right)$$

where w_i''s are $N(0,1)$ and w_i's are mutually independent and τ_o is a fixed constant.

So,

$$E(\hat{\beta}_{OLS} | x_1^T, x_2^T, \ldots, x_u^T) = \beta \frac{\tau}{n-2} e^{-\tau/2} {}_1F_1\left(\frac{n-2}{2}, \frac{n-2}{2} + 1, \tau/2\right) \quad (5.3.30)$$

So,

$$E(\beta_{OLS}) = \beta\, E\left\{ \frac{\tau}{n-2} e^{-\tau/2} {}_1F_1\left(\frac{n-2}{2}, \frac{n-2}{2} + 1, \tau/2\right)\right\} \quad (5.3.31)$$

For large positive value of τ, the asymptotic expansion of the confluent hypergeometric function ${}_1F_1\left(\frac{n-2}{2}, \frac{n-2}{2} + 1, \tau/2\right)$ is given by (vide Slater (1960))

6/ Richardson and Wu (1970) considered the model $y_i = \alpha + \beta x_i^* + \varepsilon_i$, $x_i^* = x_i + u_i$. ε_i and u_i are serially and mutually independent. For each i, ε_i and u_i are distributed $N(0, \sigma^2)$ and $N(0, \sigma_u^2)$ respectively. For such a model, Richardson and Wu (1970) obtained the exact bias of the OLS estimator of β. Gurian and Halperin (1971) extended this result to the case where ε and u jointly follow a bivariate normal distribution $N_2(\underset{\sim}{0}, \Sigma)$, where $\underset{\sim}{0} = (0,0)'$ and

$$\text{and } \Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\sigma_u \\ \rho\sigma\sigma_u & \sigma_u^2 \end{pmatrix}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$ {}_1F_1\left(\frac{n-2}{2}, \frac{n-2}{2}+1, \tau/2\right) \simeq \frac{\Gamma\left(\frac{n-2}{2}+1\right)}{\Gamma\left(\frac{n-2}{2}\right)} \frac{e^{\tau/2}}{(\tau/2)} \left[1+ \frac{\left(1-\frac{n-2}{2}\right)}{1!} \frac{2}{\tau} + \frac{\left(1-\frac{n-2}{2}\right)\left(2-\frac{n-2}{2}\right)}{2!}\left(\frac{2}{\tau}\right)^2 + \ldots\right] \tag{5.3.32} $$

$$ \therefore E(\hat{\beta}_{OLS}-\beta) = \beta\left[1-\frac{\left(1-\frac{n-2}{2}\right)}{1!} E\left(\frac{2}{\tau}\right) + \left(1-\frac{n-2}{2}\right)\left(2-\frac{n-2}{2}\right)E\left(\frac{2}{\tau}\right)^2 + \ldots\right] \tag{5.3.33} $$

For the model (5.2.40), it can be shown that

$$ E(\hat{\beta}_{OLS}) = \beta\, E\left\{\frac{\tilde{\tau}}{n-2}\, e^{-\tilde{\tau}/2}\, {}_1F_1\left(\frac{n-2}{2}, \frac{n-2}{2}+1, \tilde{\tau}/2\right)\right\} \tag{5.3.34} $$

where

$$ \tilde{\tau} = \frac{\sum_{t=1}^{n}(e^{-dt}-m)^2 \sum_{t=1}^{n}(x_t^T - \bar{x}''^T)^2 - \left\{\sum_{t=1}^{n}(e^{-dt}-m)x_t^T\right\}^2}{\sum_{t=1}^{n}(e^{-dt}-m)^2\, \sigma_{\tilde{u}}^2} \tag{5.3.35} $$

and here $x_t^T = x_t + A_u e^{-dt}$

(b) Let us now consider the case where t has been omitted from the model (5.2.6). So the misspecified E-V model is

$$ y^* = \alpha^+ + \beta^+ x^* + \xi^+ \tag{5.3.36} $$

where ξ^+ is the disturbance term in the misspecified model. y^* and x^* are same as in (5.2.1).

Equation (5.3.36) can be written as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y^* = (\alpha - \beta^+ \bar{x}^{*''}) + \beta^+ (x^* - \bar{x}^{*''}) + \xi^+ \qquad (5.3.37)$$

So, the OLS estimator of β^+ is

$$\hat{\beta}^+ = \frac{\sum_{t=1}^{n} (x_t^* - \bar{x}^{*''})\, y_t^*}{\sum_{t=1}^{n} (x_t^* - \bar{x}^{*''})^2} \qquad (5.3.38)$$

$$\therefore \quad E(\hat{\beta}^+ \mid x_1^T, x_2^T, \ldots, x_u^T ; \; x_1^*, x_2^*, \ldots, x_n^*)$$

$$= \gamma \frac{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})^t}{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})^2} + \beta \frac{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})\; x_t^T}{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})^2} \qquad (5.3.39)$$

Let us consider the term

$$\gamma \frac{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})^t}{\sum_{t=1}^{n} (x_t^* - \bar{x}^{''*})^2} \qquad (5.3.40)$$

This term can be written as

$$\gamma \frac{\sum_{t=1}^{n} (x_t^T - \bar{x}^{''T})t + \sum_{t=1}^{n} (\tilde{u}_t - \bar{\tilde{u}}^{''})t}{\sum_{t=1}^{n} (x_t^T - \bar{x}^{T''})^2 - 2\sum_{t=1}^{n} (x_t^T - \bar{x}^{''T})\, \tilde{u}_t + \sum_{t=1}^{n} (\tilde{u}_t - \bar{\tilde{u}}^{''})^2}$$

(5.3.41)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Since $\sum_{t=1}^{n}(\tilde{u}_t - \bar{\tilde{u}}'')t$ and $\sum_{t=1}^{n}(x_t - \bar{x}'')\tilde{u}_t$ jointly follow a bivariate normal distribution,

$$E\left(\sum_{t=1}^{n}(\tilde{u}_t - \bar{\tilde{u}}'')t \,\Big|\, \sum_{t=1}^{n}(x_t^T - \bar{x}''^T)\tilde{u}_t\right)$$

$$= \left\{\frac{\sum_{t=1}^{n} t(x_t^T - \bar{x}''^T)}{\sqrt{\sum_{t=1}^{n}(t-\bar{t}'')^2 \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2}} \times \frac{\sqrt{\sum_{t=1}^{n}(t-\bar{t})^2}}{\sqrt{\sum_{t=1}^{n}(x_t^T - \bar{x}''^T)^2}}\right\} \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)\tilde{u}_t^2$$

$$= \frac{\sum_{t=1}^{n} t(x_t^T - \bar{x}''^T)}{\sum_{t=1}^{n}(x_t^T - \bar{x}''^T)^2}\left\{\sum_{t=1}^{n}(x_t^T - \bar{x}''^T)\tilde{u}_t\right\} \qquad (5.3.42)$$

$$\therefore E\left[\gamma\frac{\sum_{t=1}^{n}(x_t^*-\bar{x}''^*)t}{\sum_{t=1}^{n}(x_t^*-\bar{x}''^*)^2} \,\Big|\, x_1, x_2, \ldots x_n, \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)\tilde{u}_t, \sum_{t=1}^{n}(\tilde{u}_t-\bar{\tilde{u}}'')^2\right]$$

$$= \gamma\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)t}{\sum_{t=1}^{n}(x_t^T - \bar{x}''^T)^2}\left[\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2 + \sum_{t=1}^{n}(x_t^T-\bar{x}''^T)\tilde{u}_t}{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2 - 2\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)\tilde{u}_t + \sum_{t=1}^{n}(\tilde{u}_t-\bar{\tilde{u}}'')^2}\right]$$

$$= \gamma\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)t}{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2} \cdot \frac{\sum_{t=1}^{n}(x_t^T - \bar{x}''^T)x_t^*}{\sum_{t=1}^{n}(x_t^* - \bar{x}''^*)^2} \qquad (5.3.43)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Gurian and Halperin (1971) obtained

$$E\left[\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)x_t^*}{\sum_{t=1}^{n}(x_t^*-\bar{x}''^*)^2}\,\middle|\,x_1,x_2,\ldots,x_n\right]=\frac{\tau'}{n-1}\,e^{-\tau'/2}\;{}_1F_1\left(\frac{n-1}{2},\frac{n-1}{2}+1,\frac{\tau'}{2}\right) \tag{5.3.44}$$

where

$$\tau'=\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2}{\sigma_u^2}=\frac{\sum_{t=1}^{n}(x_t-\bar{x}'')^2-2c_2\sum_{t=1}^{n}(x_t-\bar{x}'')t+c_2^2\sum_{t=1}^{n}(t-\bar{t}')^2}{\sigma_u^2} \tag{5.3.45}$$

So, from (5.3.39)

$$E(\hat{\beta}^+)=E\left[\left\{\gamma\frac{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)t}{\sum_{t=1}^{n}(x_t^T-\bar{x}''^T)^2}+\beta\right\}\frac{\tau'}{n-1}\,e^{-\tau'/2}\;{}_1F_1\left(\frac{n-1}{2},\frac{n-1}{2}+1,\frac{\tau'}{2}\right)\right] \tag{5.3.46}$$

As in (5.3.31), here also for large positive value of τ', asymptotic expansion of (5.3.46) can be obtained.

When the regressor e^{-dt} is omitted from the model (5.2.40),

$$E(\beta^+)=E\left[\left\{\gamma\frac{\sum_{t=1}^{n}(x_t-\bar{x}'')e^{-dt}}{\sum_{t=1}^{n}(x_t-\bar{x}'')^2}+\beta\right\}\frac{\tilde{\tau}'}{n-1}\,e^{-\tilde{\tau}'/2}\;{}_1F_1\left(\frac{n-1}{2},\frac{n-1}{2}+1,\frac{\tilde{\tau}'}{2}\right)\right] \tag{5.2.47}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\tilde{\tau}' = \frac{\sum_{t=1}^{n}(x_t-\bar{x}'')^2 - 2c_2 \sum_{t=1}^{n}(x_t-\bar{x}'')e^{-dt} + c_2^2 \sum_{t=1}^{n}(t-\bar{t}')^2}{\sigma_u^2}$ (5.2.48)

5.4 Extension to multiple regression

It is not at all difficult to extend the results of Section 5.2 to the case of multiple regression equations. Let the true relationship be

$$Y = \beta_o + x_1\beta_1 + x_2\beta_2 + \dots + x_k \beta_k + \varepsilon \quad (5.4.1)$$

where the disturbance term ε has mean zero and variance σ^2. Moreover, ε's are independent of x's which are stochastic. Now, suppose that y_t is measured with an error v_t, i.e.,

$$y_t^* = y_t + v_t \quad (5.4.2)$$

Let m of the k regressors be measured with errors $u_{i,t}$, i = 1,2,...,m. So,

$$x_{it}^* = x_{it} + u_{it} \quad (5.4.3)$$

Let $v_t = d_1 + d_2 t + \vec{v}_t$ (5.4.4)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and $u_{it} = c_{i1} + c_{i2}t + \tilde{u}_{it}, \quad i = 1,2,\ldots,m$ (5.4.5)

where $\tilde{u}_{it}$ and $\tilde{v}_t$, $t = 1,2,\ldots,m$ have zero means and variances $\sigma^2_{\tilde{v}}$ and $\sigma^2_{\tilde{u}_i}$, $i = 1,2,\ldots,m$ respectively. These errors are independent of true values and also independent of themselves.

So, the relationship among the observed variables is

$$y^*_t = \tilde{\beta}_o + \beta_1 x^*_{1t} + \beta_2 x^*_{2t} + \ldots + \beta_m x^*_{mt} + \beta_{m+1} x_{m+1,t} + \ldots + \beta_k x_{kt} + \tilde{\gamma} t + \xi_t \quad (5.4.6)$$

where $\tilde{\beta}_o = \beta_o - \sum_{i=1}^{m} \beta_i c_{i1} + d_1$

$$\tilde{\gamma} = d_2 - \sum_{i=1}^{m} \beta_i c_{i2}$$

and $\xi_t = \epsilon_t - \tilde{v}_t - \sum_{i=1}^{m} \beta_i \tilde{u}_{it}$

Here also the OLS regression of y^* on $x^*_1,\ldots,x^*_m, x_{m+1},\ldots,x_k$ gives inconsistent estimates because x^*_{it}, $i = 1,2,\ldots,m$, are correlated with $\tilde{\xi}_t$ for $t = 1,2,\ldots,n$.

Under the assumption that x_{it}'s are serially correlated and $\tilde{u}_{it}$'s and $\tilde{v}_t$ are serially independent, the matrix of instrumental variables can be taken as

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$Z=\begin{pmatrix} 1 & x^*_{11} & \cdots & x^*_{m1} & x_{m+1,1} & \cdots & x_{k1} & 2 \\ 1 & x^*_{12} & \cdots & x^*_{m2} & x_{m+2,2} & \cdots & x_{k2} & 3 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 1 & x^*_{1,n-1} & \cdots & x^*_{m,n-1} & x_{m+1,n-1} & \cdots & x_{k,n-1} & n \end{pmatrix} \quad (5.4.7)$$

The regression equation may be written as

$$y^*_t = (\tilde{\beta}_o-\beta_1\bar{x}^*_1-\beta_2\bar{x}^*_2-\cdots-\beta_m\bar{x}^*_m-\beta_{m+1}\bar{x}_{m+1}-\cdots-\beta_k\bar{x}_k-\tilde{\gamma}\,\bar{t})$$

$$+\beta_1(\bar{x}^*_{1t}-\bar{x}^*_1)+\beta_2(x^*_{2t}-\bar{x}^*_2)+\cdots+\beta_m(x^*_{mt}-\bar{x}^*_m)+\beta_{m+1}(x_{m+1}-\bar{x}_{m+1})$$

$$+\cdots+\beta_k(x_{kt}-\bar{x}_k)+n\tilde{\gamma}(t-\bar{t})+\zeta_t, \quad t=2,3,\ldots,n \quad (5.4.8)$$

where $\bar{x}^*_i = \dfrac{\sum_{t=2}^{n} x^*_{it}}{n-1}$, $i=1,2,\ldots,m$; $\bar{x}_j = \dfrac{\sum_{t=2}^{n} x_{jt}}{n-1}$, $j=m+1,\ldots,k$

and $\bar{t} = \sum_{t=2}^{n} t$.

So, the matrix of observations will be

$$X^*_o=\begin{pmatrix} 1 & x^*_{12}-\bar{x}^*_1 & x^*_{22}-\bar{x}^*_2 & \cdots & x^*_{m2}-\bar{x}^*_m & x_{m+1,2}-\bar{x}_{m+1} & \cdots & x_{k2}-\bar{x}_k & (2-\bar{t}) \\ 1 & x^*_{13}-\bar{x}^*_1 & x^*_{23}-\bar{x}^*_2 & \cdots & x^*_{m3}-\bar{x}^*_m & x_{m+1,3}-\bar{x}_{m+1} & \cdots & x_{k3}-\bar{x}_k & (3-\bar{t}) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 1 & x^*_{1n}-\bar{x}^*_1 & x^*_{2n}-\bar{x}^*_2 & \cdots & x^*_{mn}-\bar{x}^*_m & x_{m+1,n}-\bar{x}_{m+1} & \cdots & x_{kn}-\bar{x}_k & (n-\bar{t}) \end{pmatrix} \quad (5.4.9)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

As in (5.2.40), here also, we can consider exponential trends for $\tilde{v}_t$ and $\tilde{u}_{it}$'s.

Like (5.2.46), here also we can consider $x^*_{i,t-1}+x^*_{i,t+1}$, $i = 1,2,\ldots,m$ and $(x_{j,t-1} + x_{j,t+1})$, $j = m+1,\ldots,k$ as instruments for x^*_{it}'s and x_{jt}'s.

5.5 <u>The case of autocorrelated errors in variables.</u> Grether and Maddala (1973) considered the case of autocorrelated errors in variables. They examined the consequences of OLS methods of estimation in the presence of autocorrelated measurement errors. Here also the method leads to inconsistent estimates of the regression coefficients. They observed that for models with no lags or finite lags, measurement errors in exogeneous variables may lead to appearance of spurious long lags in adjustment. Autocorrelated measurement errors may produce autocorrelation among the true disturbances even when they are not originally autocorrelated. Generalised least squares method of estimation which is used to increase efficiency when the true disturbances are autocorrelated are likely to result in increased inconsistency.

This section intends to give some consistent methods of estimation when the measurement errors are serially correlated.

Let the true relationship be given by

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y_t = \alpha + \beta x_t + \varepsilon_t \tag{5.5.1}$$

where ε_t's, $t = 1,2,\ldots,n$ have mean zero and variance σ^2.

Let

$$y_t^* = y_t + v_t$$

and

$$x_t^* = x_t + u_t \tag{5.5.2}$$

where $E(u_t) = E(v_t) = 0 \; \forall \; t$. u_t and v_t are mutually independent and independent of true values. But both the series $\{u_t\}$ and $\{v_t\}$ are serdally correlated.

Let

$$u_t = \rho_1 u_{t-1} + w_{1t}$$

and

$$v_t = \rho_2 v_{t-1} + w_{2t} \tag{5.5.3}$$

where $|\rho_1| < 1$ and $|\rho_2| < 1$. w_{1t} and w_{2t} are serially and mutually independent disturbance terms with zero means and variances $\sigma^2_{w_1}$ and $\sigma^2_{w_2}$. The equation (5.5.1) can be written as

$$y_t^* = \alpha + \beta x_t^* + (v_t - \beta u_t + \varepsilon_t)$$

$$= \alpha + \beta x_t^* + \xi_t \tag{5.5.4}$$

Let ε_t also follow a Markov scheme given by

$$\varepsilon_t = \rho_o \varepsilon_{t-1} + w_{ot}, \quad |\rho_o| < 1 \tag{5.5.5}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

and w_{ot}'s are again serially independent with a common mean 0 and variance $\sigma^2_{w_o}$. OLS will obviously yield inconsistent estimate of β. IV techniques, however, holds greater promise.

Suppose, as in Durbin's (1954) method, u's are so small that rank ordering of the observed x^*s gives the rank ordering of the true x's. Under the above assumptions, let us use the following matrix Z as the matrix of instrumental variables.[7/]

$$Z = \begin{pmatrix} 1 & r_1 \\ 1 & r_2 \\ \vdots & \vdots \\ 1 & r_n \end{pmatrix} \qquad (5.5.6)$$

where r_i is the rank of x_i^*.

Here the covariance matrix of ξ_t is

$$V=E(\xi\xi') = \sigma^2_\xi \begin{pmatrix} 1 & \frac{\rho_o\sigma^2+\rho_1\beta^2\sigma_u^2+\rho_2\sigma_v^2}{\sigma^2_\xi} & \frac{\rho_o^2\sigma^2+\rho_1^2\beta^2\sigma_u^2+\rho_2^2\sigma_v^2}{\sigma^2_\xi} & \cdots & \frac{\rho_o^{n-1}\sigma^2+\rho_1^{n-1}\beta^2\sigma_u^2+\rho_2^{n-1}\sigma_v^2}{\sigma^2_\xi} \\ \frac{\rho_o\sigma^2+\rho_1\beta^2\sigma_u^2+\rho_2\sigma_v^2}{\sigma^2_\xi} & 1 & \frac{\rho_o\sigma^2+\rho_1\beta^2\sigma_u^2+\rho_2\sigma_v^2}{\sigma^2_\xi} & \cdots & \frac{\rho_o^{n-2}\sigma^2+\rho_1^{n-2}\beta^2\sigma_u^2+\rho_2^{n-2}\sigma_v^2}{\sigma^2_\xi} \\ \frac{\rho_o^{n-1}\sigma^2+\rho_1^{n-1}\beta^2\sigma_u^2+\rho_2^{n-1}\sigma_v^2}{\sigma^2_\xi} & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

$$= \sigma^2_\xi V_o \text{ (say)} \qquad (5.5.7)$$

7/ Here, even if $\{x_t\}$ series is serially correlated, we cannot use lagged values of x_t^* as instruments as suggested by Reiersol (1941). The reason is that, here u_t's are serially correlated and so, the lagged values of x_t^* will not be independent of u_t.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where $\sigma_\xi^2 = \sigma^2+\beta^2\sigma_u^2+\sigma_v^2$ and $\sigma^2 = \dfrac{\sigma_{w_o}^2}{1-\rho_o^2}$,

$$\sigma_u^2 = \frac{\sigma_{w_1}^2}{1-\rho_1^2} \quad \text{and} \quad \sigma_v^2 = \frac{\sigma_{w_1}^2}{1-\rho_2^2}.$$

Now, the usual IV estimator of δ is given by

$$\delta_{I(1)} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (Z' X^*)^{-1} Z' y^* \tag{5.5.8}$$

where $X^* = \begin{pmatrix} 1 & x_1^* \\ 1 & x_2^* \\ \vdots & \vdots \\ 1 & x_n^* \end{pmatrix}$

Now, since Z is independent of ξ, it can be shown easily that

$$\underset{n\to\infty}{\text{plim}} \ \hat{\delta}_{I(1)} = \left(\delta\right) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{5.5.9}$$

and asymptotic covariance matrix of $\hat{\delta}_{I(1)}$ is given by

$$\bar{V}\ \hat{\delta}_{I(1)} \quad \underset{n\to\infty}{\text{plim}} \left[(Z' X^*)^{-1} Z' V Z (X^{*\prime} Z)^{-1} \right]$$

$$= \frac{1}{n} \ \Sigma_{ZX^*}^{-1} \ \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} Z'VZ\right) \ \Sigma_{X^*Z}^{-1} \tag{5.5.10}$$

provided

$$\Sigma_{ZX^*} = \underset{n\to\infty}{\text{plim}} \left(\frac{1}{n} Z'X^*\right) \text{ exists and is nonsingular.} \tag{5.5.11}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\Sigma^{-1}_{X^*Z} = (\Sigma'_{ZX^*})^{-1}$$

and $\underset{n\to\infty}{\text{plim}} \ (Z' \ V \ Z)$ exists (5.5.12)

So, $\bar{V}(\delta^*_{I(1)})$ is estimated by

$$\hat{\bar{V}}(\hat{\delta}_{I(1)}) = \frac{1}{n}(\frac{1}{n} Z'X^*)^{-1}(\frac{1}{n} Z' \hat{V} Z)(\frac{1}{n} X^{*\prime}Z)^{-1} \quad (5.5.13)$$

where $\hat{V}$ is a consistent estimator of V.

Let us now discuss the method of obtaining $\hat{V}$. Let

$$\hat{\xi}_t = y^*_t - \hat{\alpha} - \hat{\beta} x^*_t \quad (5.5.14)$$

$\because$ $\underset{n\to\infty}{\text{plim}} \ \hat{\alpha} = \alpha$ and $\underset{n\to\infty}{\text{plim}} \ \hat{\beta} = \beta$,

$$\hat{\xi}_t \xrightarrow{L} \xi_t \quad (5.5.15)$$

So, $$\frac{1}{n} \sum_{t=i+1}^{n} \hat{\xi}_t \hat{\xi}_{t-i} \xrightarrow{L} \frac{1}{n} \sum_{t=i+1}^{n} \xi_t \xi_{t-i}$$

(5.5.16)

and $$\frac{1}{n} \sum_{t=1}^{n} \hat{\xi}^2_t \xrightarrow{L} \frac{1}{n} \sum_{t=1}^{n} \xi^2_t$$

Now, $$\underset{n\to\infty}{\text{plim}} \ \frac{1}{n} \sum_{t=i+1}^{n} \xi_t \xi_{t-i} = E(\frac{1}{n} \sum_{t=i+1}^{n} \xi_t \xi_{t-i}) \quad (5.5.17)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

provided the covariance between $(\xi_t \, \xi_{t-i})$ and $(\xi_{t+s} \quad \xi_{t-i+s})$, $s = 2,3,\ldots,n$ has the limiting value 0(*vide* Goldberger 1964, p 148)

and, $$\operatorname*{plim}_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \xi_t^2 = E(\frac{1}{n} \sum_{t=1}^{n} \xi_t^2) = \sigma_\xi^2 \qquad (5.5.18)$$

provided the covariance between ξ_t^2 and ξ_{t+s}^2, $s = 2,3,\ldots,n$ has the limiting value zero (*vide* Goldberger 1964, p 146).

Thus,

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n} \left(\sum_{t=1}^{n} \xi_t{}^2 \right) = \sigma_\xi^2$$

and

$$\operatorname*{plim}_{n\to\infty} \frac{\frac{1}{n} \sum_{t=i+1}^{n} \xi_t \, \xi_{t-i}}{\frac{1}{n} \sum_{t=1}^{n} \xi_t^2} = \frac{E(\xi_t \, \xi_{t-i})}{\sigma_\xi^2}, \quad t = 2,3,\ldots,n; \; i = 1,2,\ldots,n \qquad (5.5.19)$$

Obtaining $\hat{V}$ in this way, we can also suggest a new type of IV estimator analogous to Generalised Least Squares methods of estimation. The estimator is given by

$$\begin{aligned} \hat{\delta}_{I(2)} &= (Z' \hat{V}^{-1} X^*)^{-1} (Z' \hat{V}^{-1} y^*) \\ &= \delta + (Z' \hat{V}^{-1} X^*)^{-1} (Z' \hat{V}^{-1} \xi) \end{aligned} \qquad (5.5.20)$$

Since $\operatorname*{plim}_{n\to\infty} \hat{V} = V$, under the assumption that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

(a) $\underset{n\to\infty}{\text{plim}}\left(\frac{1}{n}\, Z'\, V^{-1}\, X^*\right)$ exists, it can be shown that

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\,(Z'\,\hat{V}^{-1}\,X^*) = \underset{n\to\infty}{\text{plim}}\ \frac{1}{n}\,(Z'\,V^{-1}\,X^*) \qquad (5.5.21)$$

Also, under the assumption that the covariance matrix of $\frac{1}{\sqrt{n}}(Z'\,V^{-1}\xi)$ exists, i.e.

(b) $\frac{1}{n}\,(Z'\,V\,Z)$ exists,

it can be shown that

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{\sqrt{n}}\,(Z'\,\hat{V}^{-1}\xi) = \underset{\sim}{0} \qquad (5.5.22)$$

Since, under the assumption that $\bar{E}(\frac{1}{n}\,Z'Z)$ exists,

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{\sqrt{n}}(Z'\,\hat{V}^{-1}\xi) = \underset{n\to\infty}{\text{plim}}\ \frac{1}{\sqrt{n}}\,(Z'V^{-1}\xi) \qquad (\underline{\text{vide}}\ \text{Appendix 2(b)})$$

(5.5.23)

$\therefore$ from (5.5.20),

$$\underset{n\to\infty}{\text{plim}}\ \frac{1}{\sqrt{n}}\,(Z'\,\hat{V}^{-1}\xi) = \underset{\sim}{0} \qquad (5.5.24)$$

$$\therefore\ \underset{n\to\infty}{\text{plim}}\ \hat{\delta}_{I(2)} = \delta \qquad (5.5.25)$$

The asymptotic variance of $\hat{\delta}_{I(2)}$ is given by

$$\bar{V}(\hat{\delta}_{I(2)}) = \frac{1}{n}\,\bar{E}\left\{\left(\frac{1}{n}\,Z'\,V^{-1}X^*\right)^{-1}\left(\frac{1}{n}\,Z'\,V^{-1}Z\right)\left(\frac{1}{n}\,X^{*\prime}\,V^{-1}Z\right)^{-1}\right\}$$

$$= \frac{1}{n}\,\bar{E}\,\left(\frac{1}{n}Z'\,V^{-1}X^*\right)^{-1}\,\bar{E}\left(\frac{1}{n}Z'\,V^{-1}Z\right)\,\bar{E}\left(\frac{1}{n}\,X^{*\prime}\,V^{-1}Z\right)^{-1} \qquad (5.5.26)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Though we have not been able to prove whether the variance of $\hat{\delta}_{I(1)}$ is smaller that that of $\hat{\delta}_{I(2)}$ or not, yet we can, however, give a very simple example where the variance of $\hat{\delta}_{I(2)}$ may be smaller than that of $\hat{\delta}_{I(1)}$.

Let us consider the simple model

$$y_t = \beta\, x_t + \varepsilon_t \tag{5.5.27}$$

where the assumptions about ε_t are same as in (5.5.1)

$$y^*_t = y_t + v_t$$

and

$$x^*_t = x_t + u_t \tag{5.5.28}$$

where the assumptions about v_t and u_t are same as in (5.5.2). Let us further assume that $\rho_o = \rho_1 = \rho_2 = \rho$. So, here

$$\bar{V}(\hat{\delta}_{I(1)}) = \frac{\sigma^2_\varepsilon}{n} \frac{\bar{E}\{\frac{1}{n}\sum_{t=1}^{n} z_t^2(1+2\rho\tilde{\rho}_1+2\rho^2\tilde{\rho}_2 + \dots + 2\rho^{n-1}\tilde{\rho}_n)\}}{\{\bar{E}\,\frac{1}{n}(\sum_{t=1}^{n} z_t\, x^*_t)\}^2} \tag{5.5.29}$$

where $\tilde{\rho}_i = \bar{E}\, \dfrac{\sum_{t=1}^{n-i} z_t\, z_{t+i}}{\sum_{t=1}^{n} z_t^2}$, $i = 1,2,\dots,n$

and $\bar{E}(\sum_{t=1}^{n} z_t\, x^*_t) = \bar{E}(\sum_{t=1}^{n} z_t\, x_t)$ since z_t is independent of $u_t\ \forall\ t$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\bar{V}(\hat{\delta}_{I(2)}) = \frac{\sigma^2}{n} \frac{\bar{E}\{\frac{1}{n} z_1^2(1-\rho^2)+\frac{1}{n}\sum_{t=2}^{n}(z_t-\rho z_{t-1})^2\}}{\bar{E}\{\frac{1}{n} z_1 x_1(1-\rho^2)+\frac{1}{n}\sum_{t=2}^{n}(z_t-\rho z_{t-1})(x_t-\rho x_{t-1})\}^2}$$

$$\simeq \frac{\bar{E}\{\sum_{t=2}^{n}(z_t-\rho z_{t-1})^2\}}{\{\bar{E}\sum_{t=2}^{n}(z_t-\rho z_{t-1})(x_t-\rho x_{t-1})\}^2}$$

$$= \frac{\bar{E}(\sum_{t=2}^{n} z_t^2)(1+\rho^2-2\rho\rho_1)}{\{\bar{E}(\sum_{t=2}^{n} z_t x_t)\}^2\{1+\rho^2-\rho(\rho_{zx\_}+\rho_{xz\_})\}^2} \qquad (5.5.30)$$

where

$$\rho_{zx\_} = \bar{E}\frac{\sum_{t=2}^{n} z_t x_{t-1}}{\sum_{t=2}^{n} z_t x_t} \qquad (5.5.31)$$

and

$$\rho_{xz\_} = \bar{E}\frac{\sum_{t=2}^{n} x_t z_{t-1}}{\sum_{t=2}^{n} z_t x_t}$$

both $\rho_{zx\_}$ and $\rho_{xz\_}$ are assumed to be less than 1.

Under the assumption that $\rho_{zx\_} + \rho_{xz\_} < 2\tilde{\rho}_1$, then from (5.5.30),

$$\bar{V}(\hat{\delta}_{I(1)}) \leq \frac{\sigma^2}{n} \frac{\bar{E}(\frac{1}{n}\sum_{t=2}^{n} z_t)^2}{\ldots} \qquad (5.5.42)$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$= \bar{E}\left(\sum_{t=2}^{n} z_t^2\right)(1+\rho^2 - 2\rho\tilde{\rho}_1)^{-1} \cdot \left\{\bar{E}\left(\sum_{t=2}^{n} z_t x_t\right)\right\}^{-2}$$

$$\leq \frac{1}{n}\bar{E}\left(\frac{1}{n}\sum_{t=2}^{n} z_t^2\right)(1-\rho^2)^{-1}\left\{\bar{E}\left(\frac{1}{n}\sum_{t=2}^{n} z_t x_t\right)\right\}^2 \quad \text{if} \quad \rho \geq \tilde{\rho}_1 \tag{5.5.33}$$

If $\rho_i \geq \tilde{\rho}_1 \ \forall i$,

$$\bar{V}(\hat{\delta}_{I(1)}) \geq \frac{1}{n}\frac{\bar{E}\left(\frac{1}{n}\sum_{t=1}^{n} z_t^2\right)(1+2\rho^2+2\rho^4+\ldots)}{\left\{\bar{E}\left(\frac{1}{n}\sum_{t=1}^{n} z_t x_t\right)\right\}^2} \geq \bar{V}(\delta_{I(2)}) \tag{5.5.34}$$

Next, let us consider the methods of obtaining V in some special cases :-

Case (a) $\rho_0 = 0.$

Here we consider the following four eqtions in four unknowns $\beta^2\sigma_u^2$, σ_v^2, ρ_1 and ρ_2.

$$\beta^2\rho_1\sigma_u^2 + \rho_2\sigma_v^2 = \frac{\sum_{t=2}^{n}\hat{\xi}_t\hat{\xi}_{t-1}}{n} = c_1 \quad \text{(say)} \tag{5.5.35}$$

$$\beta^2\rho_1^2\sigma_u^2 + \rho_2^2\sigma_v^2 \quad \frac{\sum_{t=3}^{n}\hat{\xi}_t\hat{\xi}_{t-2}}{n} = c_2 \quad \text{(say)} \tag{5.5.36}$$

$$\beta^3\rho_1^3\sigma_u^2 + \rho_2^3\sigma_v^2 \quad \frac{\sum_{t=4}^{n}\hat{\xi}_t\hat{\xi}_{t+3}}{n} = c_3 \quad \text{(say)} \tag{5.5.37}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\beta^2 \rho_1^4 \sigma_u^2 + \rho_2^4 \sigma_v^2 = \frac{\sum_{t=5}^{m} \hat{\xi}_t \, \hat{\xi}_{t-4}}{n} = c_4 \text{ (say)} \qquad (5.5.38)$$

Multiplying (5.5.35) by ρ_1 and subtracting from (5.5.36), we get

$$\rho_1 = \frac{\rho_2^2 \, \sigma_v^2 - c_2}{\rho_2 \, \sigma_v^2 - c_1} \qquad (5.5.39)$$

Similarly, multiplying (5.5.36) by ρ_1 and subtracting from (5.5.37) we get

$$\rho_1 = \frac{\rho_2^3 \, \sigma_v^2 - c_3}{\rho_2^2 \, \sigma_v^2 - c_2} \qquad (5.5.40)$$

Again, multiplying (5.5.37) by ρ_1 and subtracting from (5.5.38) we get

$$\rho_1 = \frac{\rho_2^4 \, \sigma_v^2 - c_4}{\rho_2^3 \, \sigma_v^2 - c_3} \qquad (5.5.41)$$

$\therefore$ From (5.5.39) and (5.5.40), we have

$$\sigma_v^2 = \frac{c_1 \, c_3 - c_2^2}{c_3 \rho_2 + c_1 \rho_2^3 - 2c_2 \rho_2^2} \qquad (5.5.42)$$

From (5.5.37) and (5.5.41) we get

$$\frac{\rho_2^3 \, \sigma_v^2 - c_3}{\rho_2^2 \, \sigma_v^2 - c_2} = \frac{\rho_2^4 \, \sigma_v^2 - c_4}{\rho_2^3 \, \sigma_v^2 - c_3}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

or, $(\rho_2^3\sigma_v^2 - c_3)^2 = (\rho_2^4\sigma_v^2 - c_4)\,(\rho_2^2\sigma_v^2 - c_2)$ (5.5.43)

Substituting the value of σ_v^2 from (5.5.42) in (5.5.43), we get

$$c_2(c_2^2-c_1c_3)\rho_2^4 + (c_1c_2c_4+c_1c_3^2-2c_3c_2^2)\rho_2^3 - (c_4c_2^2-2c_3^2c_2+c_1c_3c_4)\rho_2^2 + c_3(c_2c_4-c_3^2)\rho_2 = 0$$

Since $\rho_2 \neq 0$,

$$c_2(c_2^2-c_1c_3)\rho_2^3 + (c_1c_2c_4+c_1c_3^2-2c_3c_2^2)\rho_2^2 - (c_4c_2^2-2c_3^2c_2+c_1c_3c_4)\rho_2 + c_3(c_2c_4-c_3^2) = 0 \qquad (5.5.44)$$

Solving this cubic equation by Cardan's method we get three roots of ρ_2. We shall consider these roots only which lie in the region (-1,1). Substituting values of ρ_2 in (5.5.39), we get values of σ_v^2. We shall consider that value only for which $\sigma_v^2 > 0$. From (5.5.37), we get estimate of ρ_1. Here also we shall consider those values of ρ_1 which lie in the region (-1,1). From (5.5.38), we get an estimate $\beta^2\sigma_u^2$. In the expression for V in (5.5.7), putting $\rho_o = 0$, taking $\sigma_\xi^2 = \frac{1}{n}\sum_{t=1}^{n} \xi_t^2$, and using the estimates of $\rho_1, \rho_2, \beta^2\sigma_u^2$ and σ_v^2, $\hat{V}$ can be obtained. This $\hat{V}$ will give a consistent estimate of V.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

This method is, however, not fully worked out. Empirical trial is necessary to justify the workability of this method.

These results can be extended to the case where $\rho_0 \neq 0$. But that will be algebraically more laborious.

Case (b). $\rho_0 = 0, \rho_2 = 0$

$$\text{Here } V = \begin{pmatrix} \sigma_\xi^2 & \beta^2\rho_1\sigma_u^2 & \beta^2\rho_1^2\sigma_u^2 & \cdots & \beta^2\rho_1^{n-1}\sigma_u^2 \\ \beta^2\rho_1\sigma_u^2 & \sigma_\xi^2 & \beta^2\rho_1\sigma_u^2 & & \beta^2\rho_1^{n-2}\sigma_u^2 \\ \vdots & \vdots & \vdots & & \vdots \\ \beta^2\rho_1^{n-1}\sigma_u^2 & \beta^2\rho_1^{n-2}\sigma_u^2 & \beta^2\rho_1^{n-3}\sigma_u^2 & & \sigma_\xi^2 \end{pmatrix} \quad (5.5.45)$$

$$\text{So, } \quad \hat{\rho}_1 = \frac{\sum_{t=3}^{n} \hat{\xi}_t \hat{\xi}_{t-2}}{\sum_{t=2}^{n} \hat{\xi}_t \hat{\xi}_{t-1}} \quad (5.5.46)$$

$$\text{and } \quad \widehat{\beta^2 \sigma_u^2} = \frac{\sum_{t=2}^{n} \hat{\xi}_t \hat{\xi}_{t-1}}{\hat{\rho}_1} \quad (5.5.47)$$

$$\text{and } \quad \widehat{\sigma_\xi^2} = \sum_{t=1}^{n} \hat{\xi}_t^2 / n \quad (5.5.48)$$

These are, clearly consistent estimates of the population parameters.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Case (c). $\dfrac{\sigma_u^2}{\sigma_v^2} = \lambda$ (known), $\rho_o = 0$ and $\rho_1 = \rho_2 = \rho^*$.

Here,

$$V = \begin{pmatrix} \sigma_\xi^2 & \rho^*\sigma_v^2(1+\beta^2\lambda) & \rho^{*2}\sigma_v^2(1+\beta^2\lambda)\cdots & \rho^{*n-1}\sigma_v^2(1+\beta^2\lambda) \\ \rho^*\sigma_v^2(1+\beta^2\lambda) & \sigma_\xi^2 & \rho^*\sigma_v^2(1+\beta^2\lambda)\cdots & \rho^{*n-2}\sigma_v^2(1+\beta^2\lambda) \\ \vdots & \vdots & \vdots & \vdots \\ \rho^{*n-1}\sigma_v^2(1+\beta^2\lambda) & \rho^{*n-2}\sigma_v^2(1+\beta^2\lambda) & \rho^{*n-3}\sigma_v^2(1+\beta^2\lambda)\cdots & \sigma_\xi^2 \end{pmatrix}$$

where $\sigma_\xi^2 = \sigma_\varepsilon^2 + \sigma_v^2\ (1 + \beta^2\lambda)$ (5.5.49)

$$\hat{\rho}^* = \frac{\sum_{t=3}^{n} \hat{\xi}_t\ \hat{\xi}_{t-2}}{\sum_{t=2}^{n} \hat{\xi}_t\ \hat{\xi}_{t-1}} \qquad (5.5.50)$$

$$\widehat{\sigma_v^2\ (1+\beta^2\lambda)} = \frac{\sum_{t=2}^{n} \hat{\xi}_t\ \hat{\xi}_{t-1}}{\hat{\rho}^*} \qquad (5.5.51)$$

$$\widehat{\sigma_\xi^2} = \frac{\sum_{t=1}^{n} \hat{\xi}_t^2}{n} - \widehat{\sigma_v^2\ (1 + \beta^2\lambda)} \qquad (5.5.52)$$

These are all consistent estimates of population parameters.

Let us again consider case (b) which is also somewhat tractable. Here we can obtain instrumental variable estimator of δ .

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The matrix of instrumental variables can be taken as

$$Z = \begin{pmatrix} 1 & y_2^* \\ 1 & y_3^* \\ \vdots & \vdots \\ 1 & y_n^* \end{pmatrix} \qquad (5.5.53)$$

Note that $y_t^* = \alpha + \beta x_t + \varepsilon_t + v_t$,

and $\xi_{t-1} = (\varepsilon_{t-1} + v_{t-1} - \beta u_{t-1})$ are uncorrelated.

The instrumental variable estimator of δ is

$$\hat{\delta}_I = (Z' X^*)^{-1} Z' y^* \qquad (5.5.54)$$

where $$X^* = \begin{pmatrix} 1 & x_1^* \\ 1 & x_2^* \\ \vdots & \vdots \\ 1 & x_{n-1}^* \end{pmatrix}$$

Under usual assumptions similar to those in (5.5.11) and (5.5.12) $\hat{\delta}_I$ is a consistent estimator of δ .

5.6 Concluding remarks

Although the literature on errors-in-variable models is quite extensive, yet some simple problems still remain to be investigated. In this chapter, we have mainly considered the cases where the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

errors of observations contain a trend part. Two simple cases of linear and exponential trends have been taken into account. Other types of trends may also exist. For such errors-in-variable models, instrumental variable estimators have been proposed. These estimators can be used easily even when the term involving the time variable is insignificant.

In this chapter, we have also considered the case where the errors of observations are autocorrelated. This situation is much more complicated than the previous one with trending errors and our study in this respect is very much insufficient. Here also a few instrumental variable estimators have been suggested, but their workability remains to be experimentally verified.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

CHAPTER 6

CONCLUDING OBSERVATIONS

6.1 Introduction

In this chapter we take an overview of the entire investigation and stress the results reached and their significance. We also point out to the directions in which further researches are needed. It appears that many of the results on the problems of testing and estimation relating to single equation econometric models from which some of the regressors have been omitted (discussed in Chapters 2 and 3) are negative in the sense that they point to weaknesses of standard techniques which cannot be easily removed. Results in Chapters 4 and 5 are, however, to some extent, positive. In Chapter 4 we have derived conditions under which the OLS estimator of a regression coefficient in a misspecified model (misspecification due to omission of regressors) with stochastic regressors may have a smaller MSE than that in a fully specified model. In Chapter 5 problems of estimation relating to errors in variable models with trending or autocorrelated errors of observations have been discussed and Instrumental variable technique have been adopted for handling such problems.

6.2 Problems of omission of regressors from a single equation econometric model.

The study on the problems arising out of the omission of relevant regressors and autocorrelation of disturbances in

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Chapters 2 and 3 reveals that these problems require more careful investigation than is usually made by researchers. Techniques available for handling autocorrelated disturbances do not seem to be safe when regressors have been omitted. It is very widely known in econometrics that when some relevant regressors are omitted from an econometric model, the disturbances in the misspecified model become autocorrelated. For testing the hypothesis of randomness of disturbances, one performs the Durbin-Watson (1950, 1951) test with the OLS residuals. If the test statistic comes out to be significant one generally proceeds by assuming a Markov scheme for the disturbances in the misspecified model. This method, which is available in all the text books in econometrics and which is also frequently used in practice, seems to be wrong if we proceed from the point of view of Theil's (1957) specification analysis.

It has been observed that when some relevant regressors are omitted from a regression equation (with stochastic or nonstochastic regressors), the disturbances (ε^+) of the misspecified equation may have nonzero means and the disturbances do not follow a first order Markov scheme of the form

$$\varepsilon_t^+ = \rho\ \varepsilon_{t-1}^+ + u_t \qquad (6.2.1)$$

where $|\rho| < 1$ and u_t is the spherical disturbance term with mean zero and variance $\sigma_u^2\ \forall\ t$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

However, the disturbance vector may appear to be autocorrelated even when the true disturbances are spherical.

If the regressors are nonstochastic, the OLS method of estimation gives unbiased estimates of the suitably defined regression coefficients in the misspecified equation. For the case of stochastic regressors, the OLS method gives consistent estimates of the suitably defined regression coefficients in the misspecified equation. But in all the cases, whether the regressors are stochastic or not and whether the disturbances in the true model are spherical or follow a Markov scheme, the standard OLS estimates of the standard errors of the estimated regression coefficients in the misspecified model are biased.

Since the disturbances in the misspecified equation have nonzero means, the D-W (1950, 1951) statistic for testing the randomness of disturbances in the true model are not at all autocorrelated and the apparent autocorrelation is due to omission of regressors. The D-W test would lead researchers to estimate the regression coefficient by some technique based on generalised least squares estimation. The standard methods of such estimation in the presence of autocorrelated disturbances (e.g. Cochrane-Orcutt (1949) two step procedure, Prais-Winsten (vide Rao, 1968) method and Durbin (1960) two-step procedure) are found to give inconsistent estimates of suitably defined regression coefficient in the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

misspecified model. Only in the special case where the regressions of the omitted regressors on the included ones are strictly linear, the above two-step procedure give consistent estimates of the suitably defined regression coefficients in the misspecified model. But in practice one may not be able to test the hypothesis of linearity of such regressions owing to lack of observations, inadequacy of sample size etc. If, however, no regressors have been omitted, and the disturbances follow some AR process, convenient methods are available for estimation and testing in the context of single equation econometric models.

So, need is felt for some proper method of testing whether the autocorrelation among the disturbances is due to omission of relevant regressors and/or due to autocorrelation among the true disturbances. It appears that discrimination between the two causes of autocorrelation could be quite difficult at least for moderate sized samples. Proper methods of estimating the regression coefficients (suitably defined) in the misspecified model need also be developed for the case where there has been some omission of regressors; it is also necessary to find unbiased and consistent estimates of sampling variances of such estimates.

6.2.1 Tests of misspecification due to omission of regressors.

Ramsey (1969) proposed different tests (RESET, RASET and COMSET) of omission of relevant regressors from single equation

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

regression models. These tests are based on BLUS residuals of the misspecified equation. Later, Ramsey and Schmidt (1976) simplified the test RESET using the OLS residuals of the misspecified equation. These tests have been discussed in Chapter 1. Ramsey (1969) and Ramsey and Schmidt (1976), however, considered the case where the disturbances in the true model are spherical.

If, however, it is known beforehand that the disturbances in the true model follow a Markov scheme with the first order autoregression coefficient ρ (assumed to be known), then, in the following way, the above test can be easily modified to yield more general tests of misspecification. Let the true model be

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t \tag{6.2.2}$$

and

$$\varepsilon_t = \rho \, \varepsilon_{t-1} + u_t , \tag{6.2.3}$$

$|\rho| < 1$ and u_t is the spherical disturbance term with mean 0 and variance $\sigma_u^2 \; \forall \; t$.

From (6.1.2) and (6.1.3) we have

$$y_t - \rho y_{t-1} = \beta_1 (x_{1t} - \rho x_{1,t-1}) + \beta_2 (x_{2t} - \rho x_{2,t-1}) + \dots + \beta_k (x_{k,t} - \rho x_{k,t-1}) + u_t \tag{6.2.4}$$

Considering $(x_{1t} - \rho \, x_{1,t-1})$, $(x_{2t} - \rho \, x_{2,t-1})$ etc., to be the

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

regressors and $y_t - \rho\, y_{t-1}$ to be the regressand, the tests can be used as tests of omission of regressors.

But in most of the cases, nothing is actually known about the nature of ε_t. So, what is necessary is the simultaneous test of autocorrelation of disturbances in the true model and omission of regressors from the true model. But, it seems very difficult to develop such tests.

Some work by Sarkar (1978) currently underway at I. S. I., Calcutta, seem to have thrown up useful results in this area. Sarkar tried to discriminate between autocorrelation among OLS residuals arising due to omission of regressors and pure autocorrelation among the disturbances of the true model. The true model is given by (6.1.2) and ε_t's follow the Markov scheme given by (6.1.3). The misspecified model is given by.

$$y_t = \beta_1^+ x_{1t} + \beta_2^+ x_{2t} + \ldots + \beta_m^+ x_{mt} + \varepsilon_t^+ \quad (m < k) \tag{6.2.5}$$

where β_i^+, $i = 1,2,\ldots,m$ and ε^+ have been defined in Chapter 3. Sarkar assumed that

$$\tilde{z}_t = \rho_{\tilde{z}}\, \tilde{z}_{t-1} + v_t \tag{6.2.6}$$

where $\tilde{z} = E(\varepsilon^+)$, $|\rho_{\tilde{z}}| < 1$ and v_t is the spherical disturbance term with mean zero and variance σ_v^2 $\forall$ t.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$E(\tilde{z}_t) = 0 \; \forall \; t$$

and $$\underset{n \to \infty}{\text{plim}} \left(\frac{1}{n} \tilde{z}' \tilde{z}\right) = \sigma^2_{\tilde{z}} \tag{6.2.7}$$

Next, Sarkar defined

$$\theta_t = e^+_t - \hat{\rho}^* e^+_{t-1} \quad (t = 2,3,\ldots,n) \tag{6.2.8}$$

where $$\hat{\rho}^* = \frac{\sum_{t=3}^{n} e^+_t e^+_{t-2}}{\sum_{t=2}^{n} e^+_t e^+_{t-1}}$$

e^+_t being the OLS residual in the misspecified equation. A large sample test based on the autocovariances of θ_t of different orders has been developed in an attempt to discriminate among the following cases :

1. $\rho = 0, \quad \rho_{\tilde{z}} \neq 0, \quad \sigma^2_{\tilde{z}} > 0$.
2. $\rho \neq 0, \quad \rho_{\tilde{z}} = 0, \quad \sigma^2_{\tilde{z}} > 0$.
3. $\rho = 0, \quad \sigma^2_{\tilde{z}} = 0$.
4. $\rho = 0, \quad \rho_{\tilde{z}} = 0, \quad \sigma^2_{\tilde{z}} > 0$.
5. $\rho \neq 0, \quad \sigma^2_{\tilde{z}} = 0$.
6. $0 \neq \rho \neq \rho_{\tilde{z}} \neq 0, \quad \sigma^2_{\tilde{z}} > 0$.
7. $\rho = \rho_{\tilde{z}} = 0, \quad \sigma^2_{\tilde{z}} > 0$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Sarkar, however, has been able to discriminate among the following four groups only.

Group 1	Cases 1 and 2
Group 2	Cases 3 and 4
Group 3	Cases 5 and 7
Group 4	Cases 6 .

Following generalised least squares method of estimation, Sarkar (1978) also obtained consistent estimate of β^+. But in doing this, he implicitly assumed that $\tilde{z}$ is independent of X^+ or $E(\tilde{z}|X^+) = 0$.

However, in practice, $\tilde{z}_t$ need not follow a Markov scheme. In fact, $E(\tilde{z}_t) = 0 \; \forall \; t$ and $\tilde{z}$ is independent of X^+ are oversimplified assumptions. If these could be assumed, then, all the standard methods of estimation in the presence of autocorrelated disturbances will give consistent estimates of $\hat{\beta}^+$ under fairly general conditions.

6.2.2 Estimation when some regressors have been omitted.

When the D-W (1950,1951) statistic comes out to be significantly less than 2, one generally fits a Markov scheme of the form (6.2.1) to ε_t^+ and tries to re-estimate the regression coefficients (β^+) of the misspecified model by various two-step methods of estimation.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

It has been observed in chapters 2 and 3 that these methods of estimation (e.g. Cochrane-Orcutt two-step procedure, Prais-Winsten method and Durbin two-step procedure) give inconsistent estimates of the regression coefficients β^+. These results are, however, large sample results. In actual calculation with numerical data, the large sample bias may be negligibly small and the usual two-step methods of estimation can then be used for estimation without much risk. Monte-Carlo experiments should be conducted to study the small sample bias of these methods of estimation in case of misspecification as a result of omission of regressors. Theoretical and experimental work is also needed on the suitability of estimates of sampling variances associated with these methods. If these investigations point to serious deficiencies of available methods, new methods may have to be developed for unbiased estimates of regression coefficients of the misspecified model and the sampling variances of these estimated regression coefficients.

For developing his tests Ramsey (1969) suggested a method for approximating the mean of the disturbances in the misspecified model. The approximation is

$$E(\varepsilon_t^+) \simeq \alpha_0 + \alpha_1\mu_{01t} + \alpha_2\mu_{02t} + \alpha_3\mu_{03t} + \cdots \qquad (t = 1,2,\ldots,n) \tag{6.2.9}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

where μ_{0jt}, $j = 1,2,\ldots$ are the j-th moments about the origin of $\hat{y}_t$ (the OLS estimate of y_t is the misspecified model) given X^+, where

$$X^+ = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{pmatrix}$$

$\mu_{0j} = (\mu_{0j1}, \mu_{0j2}, \ldots \mu_{0jn})'$ is estimated unbiasedly by $\hat{y}^j = (\hat{y}_1^j, \hat{y}_2^j, \ldots \hat{y}_n^j)'$. This work of Ramsey suggests that by fitting the model

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_m x_{mt} + \alpha_2 \hat{y}_t^2 + \alpha_3 \hat{y}_t^3 + \cdots + \delta_t \quad (6.2.10)$$

(where δ_t is the spherical error term with mean zero and variance $\sigma_\delta^2 \;\forall\; t$) we may get better results. However, the validity of the approximation (6.2.8) suggested by Ramsey is not very clear. Ramsey considered the true model to be

$$y = x_1\beta_1 + x_2\beta_2 + \varepsilon \quad (6.2.11)$$

where y, x_1, x_2 and ε are $(n \times 1)$ vectors and $E(\varepsilon) = 0$, $E(\varepsilon\varepsilon') = \sigma^2 I_n$.

The misspecified model is

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$y = x_1\beta_1 + \nu \qquad (6.2.12)$$

when $x_2 = x_1^2$, Ramsey obtained

$$E(\hat{\nu}|x_1) = \alpha_0 + \alpha_1 E(y^2|x_1) \qquad (6.2.13)$$

Following this, we tried to examine the case where the true model is

$$y = x_1\beta + x_2\beta_2 + x_3\beta_3 + \varepsilon$$

y, x_1, x_2, x_3 and ε and $(n \times 1)$ vectors and the properties of ε are same as in (6.2.11).

The misspecified model is

$$y = x_1\gamma_1 + x_2\beta_2 + \nu \qquad (6.2.14)$$

x_3 is not orthogonal to x_1 and x_2.

In this case, it appears that the approximation given by Ramsey for estimating $E(\hat{\nu}|x_1, x_2)$, is not as good as it is in the case (6.2.13). Also the properties of the OLS estimates based on models like (6.2.10) are not very clear and the applicability of such models should be carefully studied for estimation purposes.

6.2.3 Further problems connected with omission of regressors.

In Chapter 4 we have derived conditions under which the OLS estimator of a regression coefficient in a mispecified model (only omission of regressors has been considered) with stochastic

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

regressors may have a smaller MSE than that in a fully specified model. Note that in the former case the OLS estimate is biased for the true regression coefficient and this bias is considered in computing the MSE. These conditions can be easily verified under the assumptions that the regressors in the true model have a multivariate normal distribution.

In this chapter we have also considered by means of the MSE criterion the problem of using or discarding a (perhaps poor)proxy for a relevant regressor which appears in a multiple regression model. This problem was originally considered by Aigner (1974) for a two-variable case. The model considered by Aigner was

$$y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad (6.2.15)$$

where y, x_1 and x_2 jointly follow a trivariate normal distribution, ε $N(0, \sigma^2)$ and the variance-covariance matrix of x_1 and x_2 is given by

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

x_2 has been measured with error ; the observation called proxy variable is x_2^* where

$$x_2^* = x_2 + u \qquad (6.2.16)$$

$u \sim N(0, \sigma_u^2)$ and u is independent of all other basic variables.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

When x_2 has been omitted from (6.2.14), the model can be written as

$$y = \beta_1^+ x_1 + \varepsilon^+ \qquad (6.2.17)$$

The OLS estimator of β_1^+ is given by $\hat{\beta}_1^+$. Let $\hat{\beta}_{1p}$ be the OLS estimator of β_1 when the proxy variable x_2^* has been included in the model (6.2.15) instead of x_2. Aigner compared MSE($\hat{\beta}_1^+$) with that of MSE($\hat{\beta}_{1p}$). But his result for MSE($\hat{\beta}_1^+$) was not fully correct. It ignored some of the terms in the true expression for MSE($\hat{\beta}_1^+$). We have compared the correct expression for MSE($\hat{\beta}_1^+$) with that of MSE($\hat{\beta}_{1p}$). The conclusion is to include the proxy in most of the cases. Our examination covers a wider range of values of ρ_{12} (the correlation coefficient between x_1 and x_2), n (the sample size) and $\lambda = \dfrac{\sigma_u^2}{\sigma_{22} + \sigma^2}$ than that considered by Aigner. Some attempts have also been made to examine this question in the case of k regressors (k > 2).

6.3 Handling of trending or autocorrelated errors in variables.

The fifth chapter discusses some problems of estimation relating to errors in variables. Though there is extensive literature on the problem of errors in variables, many obvious practical problems have not been considered at all. For example, the errors in the variables in a regression model may have some

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

trend component or the errors may themselves be autocorrelated. We have considered a few problems of estimation in such cases.

In searching for methods of estimation for these types of errors in variable models, we found it necessary to adopt instrumental variable techniques as OLS is fundamentally inapplicable. The true model is

$$y_t = \alpha + \beta x_t + \epsilon_t \qquad (6.3.1)$$

ϵ_t is the spherical disturbance term with mean zero and variance σ^2 for all t. ϵ_t's are also independent of x_t's. In the case of errors with linear trend, the E-V model is

$$y_t^* = \alpha + \beta x_t^* + \gamma t + \xi_t \qquad (6.3.2)$$

where

$$y_t^* = y_t + v_t \qquad (6.3.3)$$

$$x_t^* = x_t + u_t \qquad (6.3.4)$$

$$u_t = c_1 + c_2 t + \tilde{u}_t \qquad (6.3.5)$$

$$v_t = d_1 + d_2 t + \tilde{v}_t \qquad (6.3.6)$$

$\tilde{u}_t$ and $\tilde{v}_t$ have zero means and variances $\sigma^2_{\tilde{u}}$ and $\sigma^2_{\tilde{v}}$. $\tilde{u}_t$ and $\tilde{v}_t$ are serially and mutually independent and independent of x_t, y_t and ϵ_t.

$$\xi_t = \epsilon_t - \beta \tilde{u}_t + \tilde{v}_t .$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

For exponential trend,

$$u_t = A_u \exp(-d_u t) + \tilde{u}_t \qquad (6.3.7)$$

$$v_t = A_v \exp(-d_v t) + \tilde{v}_t \qquad (6.3.8)$$

and the E-V model is

$$y_t^* = \alpha + \beta x_t^* - \gamma e^{-d_u t} + \delta e^{-d_v t} + \xi_t \qquad (6.3.9)$$

where $\gamma = \beta A_u$ and $\delta = A_v$.

Under the assumption that

$$d_u = d_v = d,$$

$$y_t^* = \alpha + \beta x_t^* + \tilde{\gamma} e^{-dt} + \xi_t \qquad (6.3.10)$$

where $\tilde{\gamma} = (\delta - \gamma)$.

Under the assumption that the true regressors are themselves serially correlated, the methods of using the lagged regressors or a linear combination of lagged regressors (vide Reierso (1941) and Karni and Weissman (1974)) as instruments have been found to be suitable. These methods are, however, not applicable for serially correlated $\tilde{u}_t$'s. We suggest that these methods may be used rather routinely even though the errors may not have any trend component.

On the basis of instrumental variable (IV) methods of estimation of β and γ in (6.3.2), estimated standard error is obtained. Using this, the significance of γ is tested. If γ is non-significant, we may drop the trend component. It is afer even then to include t as a regressor. For the case of exponential

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

trends, also, we may get similar results. Next problem is to choose between the linear and exponential trend. The suggestion is to fit the E-V model by I-V technique once by assuming a linear trend in errors and then by using an exponential trend in errors and choose that one which has smaller residual sum of squares.

In the case where the errors are autocorrelated, we have proposed Durbin's (1954) wellknown I-V technique which uses the ranks of the observed regressors under the assumption that the errors of observations are so small that ranking on the basis of the observed variables do not affect the rankings of the true variables..But in the case where the errors of observations are large, we do not get any solution and the above I-V technique is also not applicable.

In order to carry out the estimation process in this case, the first task is to decide whether the errors are autocorrelated or not. This is, however, not so easy. The autocorrelation in the errors of observations can only by suspected on the basis of intimate knowledge of constructing the data series.

The discussions in Chapter 5 can be regarded as just a beginning of the investigation with some of the practical problems relating to errors in variable models. In fact we may have more general E-V models. The errors may have a trend component and an autocorrelated error term at the same time. The problems of estimation become much more complicated in such cases. These problems require detailed investigation and further researches should be carried out on these topics.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

APPENDIX 1

A. Here we shall prove the results (2.4.14) and (2.4.24) of pages 136 and 139 respectively. Result (2.4.14) is

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n-1}} X^{+\prime} (W^{-1}_{\hat{\rho}_o} - W^{-1}_{\rho_o}) \varepsilon = \underset{\sim}{0} \quad \text{under the assumption}$$

$$\text{that} \quad \lim_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} x^2_{h,t-1}, \quad h = 1,2,\ldots,n \quad \text{exists.}$$

<u>Proof</u> : Let us consider the h-th element of the vector

$\frac{1}{\sqrt{n-1}} X^{+\prime}(W^{-1}_{\hat{\rho}_o} - W^{-1}_{\rho_o}) \varepsilon$. This is given by

$$\frac{1}{\sqrt{n-1}} \sum_{t=2}^{n} [(x_{ht} - \hat{\rho}_o x_{h,t-1})(\varepsilon_t - \hat{\rho}_o \varepsilon_{t-1}) - (x_{ht} - \rho_o x_{h,t-1})(\varepsilon_t - \rho_o \varepsilon_{t+1})$$

$$= (\hat{\rho}_o - \rho_o)[\frac{1}{\sqrt{n-1}} \sum_{t=2}^{n} x_{h,t-1} \varepsilon_t - \frac{\rho_o}{\sqrt{n-1}} \sum_{t=2}^{n} x_{h,t-1} \varepsilon_{t-1}$$

$$+ \frac{1}{\sqrt{n-1}} \sum_{t=2}^{n} (x_{ht} - \rho_o x_{h,t-1}) \varepsilon_{t-1} - \frac{\hat{\rho}_o - \rho_o}{\sqrt{n-1}} \sum_{t=2}^{n} x_{h,t-1} \varepsilon_{t-1}] \tag{1}$$

The first term in the bracket is a random variable with mean 0 and variance $\frac{\sigma^2}{n-1} \sum_{t=2}^{n} x^2_{h,t-1}$. Since we have assumed that $\lim_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} x^2_{h,t-1}$ exists, the term has a bounded variance and hence the multiplication by $\hat{\rho}_o - \rho_o$ ensures that it converges

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

in probability to zero. Similar arguments apply to the second term and the fourth term. The third term in the bracket has zero mean and variance $\frac{\sigma^2}{n-1} \sum_{t=2}^{n} (x_{ht} - \rho_o x_{h,t-1})^2$. Since by assumption (a), $\lim_{n \to \infty} \frac{1}{n-1} \sum_{t=2}^{n} (x_{ht} - \rho_o x_{h,t-1})^2$ exists, this term also has bounded variance. Hence, the multiplication by $(\hat{\rho}_o - \rho_o)$ ensures that it converges in probability to zero. Similarly, it can be shown that other elements in the vector $\frac{1}{\sqrt{n-1}} X^{+\prime}(W^{-1}_{\hat{\rho}_o} - W^{-1}_{\rho_o})\varepsilon$ converge in probability to zero. So,

$$\operatorname*{plim}_{n \to \infty} \frac{1}{\sqrt{n-1}} X^{+\prime} (W^{-1}_{\hat{\rho}_o} - W^{-1}_{\rho_o}) \varepsilon = \underset{\sim}{0} .$$

Next, we shall prove the result (2.4.21). The result is, under the assumption (2.4.13),

$$\operatorname*{plim}_{n \to \infty} \frac{1}{\sqrt{n}} X^{+\prime}(V^{-1}_{\hat{\rho}_o} - V^{-1}_{\rho_o}) \varepsilon = \underset{\sim}{0} .$$

<u>Proof</u>: Since for large n, $\frac{n}{n-1} \simeq 1$, and $\operatorname*{plim}_{n \to \infty} \hat{\rho}_o = \rho_o$, the probability limit of the h-th element in the vector

$\frac{1}{\sqrt{n}} X^{+\prime}(V^{-1}_{\hat{\rho}_o} - V^{-1}_{\rho_o}) \varepsilon$ is

$$\frac{1}{(1-\rho_o^2)} \operatorname*{plim}_{n \to \infty} \left[\frac{1}{\sqrt{n}}\{(1-\hat{\rho}_o^2) - (1-\rho_o^2)\} x_{h1}\varepsilon_1 + \text{the expression in (1)}\right] \qquad (2)$$

$$= \frac{1}{(1-\rho_o^2)} \underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} [\{(1-\hat{\rho}_o^2) - (1-\rho_o^2)\} x_{h1}\, \epsilon_1] \qquad (3)$$

Since, $\frac{1}{\sqrt{n}} x_{h,1}\, \epsilon_1$ has mean 0 and variance $\frac{\sigma^2}{n} x_{h1}^2$ and since the term $\frac{1}{\sqrt{n}}\{(1-\hat{\rho}_o^2) - (1-\rho_o^2)\} x_{h1}\, \epsilon_1$ converges in probability to zero. Similar arguments apply to other elements of the vector $\frac{1}{\sqrt{n}} X^{+\prime} (V_{\hat{\rho}_o}^{-1} - V_{\rho_o}^{-1})\, \epsilon$. Hence,

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} X^{+\prime}(V_{\hat{\rho}_o}^{-1} - V_{\rho_o}^{-1})\, \epsilon = \underset{\sim}{0}\,.$$

B. Here we shall prove the result (2.4.35) of page 143. The result is :

Under the assumption (2.3.4), $\underset{n\to\infty}{\text{plim}} \frac{1}{n-1}(\bar{X}^{+\prime} \epsilon_{\rho_o}) = (0, -\rho_o\, \sigma^2, 0, 0 \ldots 0)'$

<u>Proof</u> : All the elements excepting the second one in the vector $\frac{1}{n-1} \bar{X}^{+\prime}\, \epsilon_{\rho_o}$ has mean zero. The first element of this vector is

$$\frac{1}{n-1} \sum_{t=2}^{n} (\epsilon_t - \rho_o\, \epsilon_{t-1}) \qquad (4)$$

The variance of this element is

$$\frac{1}{(n-1)^2} \{(n-1)(1+\rho_o^2)\, \sigma^2 - 2(n-2)\, \rho_o\, \sigma^2\} \to 0 \quad \text{as} \quad n \to \infty$$

(5)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Next, let us consider the (2h-1) th ($h \neq 1$) element of the vector $\frac{1}{n-1}(\bar{X}' \; \varepsilon_{\rho_o})$. This element is

$$\frac{1}{n-1} \sum_{t=2}^{n} x_{ht} \, (\varepsilon_t - \rho_o \, \varepsilon_{t-1}) \tag{6}$$

The variance of this element is

$$\frac{\sigma^2}{(n-1)^2} \left\{ (1 + \rho_o^2) \sum_{t=2}^{n} x_{ht}^2 - 2 \, \rho_o \sum_{t=2}^{n} x_{ht} \, x_{h,t-1} \right\} \tag{7}$$

Since, by (2.3.3), $\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} x_{ht}^2$ exists, and since $\frac{1}{n-1} \sum_{t=2}^{n} x_{ht} \, x_{h,t-1}$ cannot be greater than $\frac{1}{n-1} \sum_{t=2}^{n} x_{ht}^2$ in absolute value, the expression in (7) tends to 0 as $n \to \infty$. Similar arguments apply to the variance of 2h-th ($h \neq 1$) element of of the vector $\frac{1}{n-1} (\bar{X}^{+\prime} \cdot \varepsilon_{\rho_o})$.

Now, let us consider the element of the vector $\frac{1}{n-1}(\bar{X}^{+\prime} \; \varepsilon_{\rho_o})$. This element is

$$\frac{1}{n-1} \sum_{t=2}^{n} y_{t-1} (\varepsilon_t - \rho_o \, \varepsilon_{t-1})$$

$$= \frac{1}{n-1} \sum_{t=2}^{n} \left\{ (\beta_1 + \beta_2 \, x_{2,t-1} + \beta_3 x_{3,t-1} + \dots + \beta_k x_{k,t-1}) + \varepsilon_{t-1} \right\} (\varepsilon_t - \rho_o \varepsilon_{t-1})$$

$$= \frac{1}{n-1} \left\{ \sum_{t=2}^{n} M_t (\varepsilon_t - \rho_o \varepsilon_{t-1}) + \sum_{t=2}^{n} \varepsilon_{t-1} \, \varepsilon_t - \rho_o \sum_{t=2}^{n} \varepsilon_{t-1}^2 \right\} \tag{8}$$

where $M_{t-1} = \beta_1 + \beta_2 \, x_{2,t-1} + \beta_3 \, x_{3,t-1} + \dots + \beta_k \, x_{k,t-1}$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

The term $\frac{1}{n-1} \sum_{t=2}^{n} M_t \ (\varepsilon_t - \rho_o \varepsilon_{t-1})$ has mean 0 and variance

$$\frac{\sigma^2}{(n-1)^2} \{(1+\rho_o^2) \sum_{t=2}^{n} M_{t-1}^2 - 2\rho_o \sum_{t=2}^{n} M_{t-1}, M_t\} \tag{9}$$

Now, $\frac{1}{n-1} \sum_{t=2}^{n} M_{t-1}^2 = \frac{1}{n-1} \sum_{t=2}^{n} (\beta_1^2 + \beta_2^2 x_{2,t-1}^2 + \cdots + \beta_k^2 x_{k,t-1} + 2\beta_1\beta_2 x_{2,t-1}$

$$+ \cdots + 2\beta_1\beta_k x_{k,t-1} + 2\beta_2\beta_3 x_{2,t-1} x_{3,t-1}$$

$$+ \cdots + \beta_{k-1}\beta_k x_{k-1,t-1} x_{k,t-1}) \tag{10}$$

By assumption (2.3.2), the expression on the right hand side of (10) is bounded. Also, $\frac{1}{n-1} \sum_{t=2}^{n} M_{t-1} M_t$ cannot exceed $\frac{1}{n-1} \sum_{t=2}^{n} M_t^2$ in absolute value. So, as $n \to \infty$, the expression in (9) tends to 0.

So, $\operatorname{plim}_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} M_{t-1}(\varepsilon_t - \rho_o \varepsilon_{t-1}) = 0$ (by Chebyshev's inequality)

(11)

The term $\frac{1}{(n-1)} \sum_{t=2}^{n} \varepsilon_t \varepsilon_{t-1}$ has mean 0 and variance given by

$$\frac{1}{(n-1)^2} V(\sum_{t=2}^{n} \varepsilon_t \varepsilon_{t-1}) = \frac{1}{(n-1)^2} \{\sum_{t=2}^{n} V(\varepsilon_t^2 \varepsilon_{t-1}^2) +$$

$$2 \sum_{t \neq t'} \operatorname{cov}(\varepsilon_t \varepsilon_{t-1}), (\varepsilon_{t'} \varepsilon_{t'-1})\}$$

$$= \frac{\sigma^4}{(n-1)} \to 0 \text{ as } n \to \infty \tag{12}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} \sum_{t=2}^{n} \varepsilon_t \varepsilon_{t-1} = 0$ (by Chebyshev's inequality) (13)

The term $-\frac{\rho_o}{n-1} \sum_{t=2}^{n} \varepsilon_{t-1}^2$ has mean $-\rho_o \sigma^2$ and variance given by

$$= \frac{\rho_o^2}{(n-1)^2} \left\{ \sum_{t=2}^{n} E(\varepsilon_{t-1}^4) + \sum_{t \neq t'} E(\varepsilon_{t-1}^2 \varepsilon_{t'-1}^2) \right\} - \rho_o^2 \sigma^4$$

$$= \frac{\rho_o^2}{(n-1)^2} \sum_{t=2}^{n} E(\varepsilon_{t-1}^4) + \frac{\rho_o^2 \sigma^4 (n-1)(n-2)}{(n-1)^2} - \rho_o^2 \sigma^4$$

$$\longrightarrow \lim_{n\to\infty} \frac{\rho_o^2}{(n-1)^2} \sum_{t=2}^{n} E(\varepsilon_{t-1}^4) \quad \text{as} \quad n \to \infty \qquad (14)$$

Under the assumption that $E(\varepsilon_t^4) = \mu_4 \ \forall\ t$ exists, the expression in (12) is equal to zero. So,

$$-\rho_o \underset{n\to\infty}{\text{plim}} \frac{1}{n-1} \sum_{t=2}^{n} \varepsilon_{t-1}^2 = -\rho_o \sigma^2 \qquad (15)$$

Hence, $\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} X^{*\prime} \varepsilon_{\rho_o} = (0, -\rho_o \sigma^2, 0 \ldots 0)'$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

APPENDIX 2

A. Here we shall prove the result (3.5.14) of page 185 and the result (3.5.31) of page 199.

The result (3.5.14) is that under the assumption

$$\lim_{n\to\infty} E\left(\frac{1}{n-1}\sum_{t=1}^{n-1} x_{ht}^2\right) \text{ exists,}$$

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n-1}} X^{+\prime}(W_{\hat{\rho}^*}^{-1} - W_{\rho^*}^{-1})\, \varepsilon = \underset{\sim}{0}$$

<u>Proof</u>: As in Appendix 1A, here also, the h-th (h =1,2,...,m) element of $\frac{1}{\sqrt{n-1}} X^{+\prime}(W_{\hat{\rho}^*}^{-1} - W_{\rho^*}^{-1})\, \varepsilon$ is given by

$$-(\hat{\rho}^* - \rho^*)\left[\frac{1}{\sqrt{n-1}}\sum_{t=2}^{n} x_{h,t-1}\,\varepsilon_t - \frac{\rho^*}{\sqrt{n-1}}\sum_{t=2}^{n} x_{h,t-1}\,\varepsilon_{t-1}\right.$$

$$\left. + \frac{1}{\sqrt{n-1}}\sum_{t=2}^{n}(x_{h,t} - \hat{\rho}^* x_{h,t-1})\,\varepsilon_{t-1} - \frac{\hat{\rho}^* - \rho^*}{\sqrt{n-1}}\sum_{t=2}^{n} x_{h,t-1}\,\varepsilon_{t-1}\right] \quad (1)$$

The first term in the bracket has mean 0 and the following variance:

$$\frac{1}{n-1} E\left\{\sum_{t=2}^{n}\sum_{t'=2}^{n} x_{h,t-1}\, x_{h,t'-1}\, E(\varepsilon_t\, \varepsilon_{t'} \mid X)\right\} \quad (2)$$

$$= \frac{\sigma^2}{n-1} E\left\{\sum_{t=2}^{n}\sum_{t'=2}^{n} \rho^{|t-t'|}\, x_{h,t-1}\, x_{h,t'-1}\right\}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$= \frac{\sigma^2}{n-1} E\{\sum_{t=2}^{n}[x_{h,t-1}^2 + \rho(x_{h,t-1}\, x_{h,t} + x_{h,t-1}\, x_{h,t-2})$$

$$+ \rho^2(x_{h,t-1}\, x_{h,t+1} + x_{h,t-1}\, x_{h,t-3}) + \cdots]\} \qquad (3)$$

$$\leq \frac{\sigma^2}{n-1} E \sum_{t=2}^{n} x_{h,t-1}^2 (1 + 2\rho + 2\rho^2 + \cdots + 2\rho^{n-1}) \qquad (4)$$

So, as $n \to \infty$, the variance in (2) cannot exceed

$$\sigma^2 \cdot \lim_{n\to\infty} E(\frac{1}{n-1} \sum_{t=2}^{n} x_{h,t-1}^2)(1 + 2\rho + 2\rho^2 + \cdots)$$

$$= \sigma^2(1 + \frac{2\rho}{1-\rho}) \cdot \lim_{n\to\infty} E(\frac{1}{n-1} \sum_{t=2}^{n} x_{h,t-1}^2) \qquad (5)$$

So, under the assumption that $\lim_{n\to\infty} E(\frac{1}{n-1} \sum_{t=2}^{n} x_{h,t-1}^2)$ exists, the first term in the bracket has bounded variance. Since $\operatorname{plim}_{n\to\infty} \hat{\rho}^* = \rho^*$, the multiplication by $\hat{\rho}^* - \rho^*$ ensures that it converges in probability to zero. A similar argument applies for the second term and the fourth term. With the help of similar algebraic manipulations as in (2), (3), (4) and (5), it can be proved that for large n, the second term in the bracket will have bounded variance if, $\lim_{n\to\infty} E \frac{1}{n-1} \sum_{t=2}^{n} (x_{h,t} - \hat{\rho}^* x_{h,t-1})^2$ exists. This, in fact, exists since by assumption (3.5.0)

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} \sum_{t=2}^{n} (x_{h,t} - \hat{\rho}^* x_{h,t-1})^2$ exists. So, multiplication by $\hat{\rho}^* - \rho^*$ ensures that this term also converges in probability to zero. Similar arguments apply for other elements of the vector $\frac{1}{\sqrt{n-1}} X^{+\prime}(W^{-1}_{\hat{\rho}^*} - W^{-1}_{\rho^*}) \varepsilon$. Hence,

$$\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n-1}} X^{+\prime}(W^{-1}_{\hat{\rho}^*} - W^{-1}_{\rho^*}) \varepsilon = \underset{\sim}{0} .$$

Next, we shall prove the result (3.5.31) which states that $\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} X^{+\prime}(V^{-1}_{\hat{\rho}^*} - V^{-1}_{\rho^*}) \varepsilon = \underset{\sim}{0}$ under the assumption that $\lim_{n\to\infty} E\left\{\frac{1}{n-1} \sum_{t=2}^{n} x^2_{h,t-1}\right\}$ exists.

<u>Proof</u> : Since, for large n, $\frac{n}{n-1} \simeq 1$, the h-th $(h = 1,2,\ldots,m)$ element of the vector $\frac{1}{\sqrt{n}} X^{+\prime} (V^{-1}_{\hat{\rho}^*} - V^{-1}_{\rho^*}) \varepsilon$ is given by

$$= \frac{1}{\sqrt{n}} \{(1-\hat{\rho}^{*2})-(1-\rho^{*2})\} x_{h1} \varepsilon_1 + \text{the expression in (1)} \qquad (6)$$

$$E\left(\frac{x_{h1} \varepsilon_1}{\sqrt{n}}\right) = 0 \quad \text{and} \quad V\left(\frac{x_{h1} \varepsilon_1}{\sqrt{n}}\right) = \frac{\sigma^2}{n} E(x^2_{h1}).$$

So, $\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} \{(1-\hat{\rho}^{*2}) - (1-\rho^{*2})\} x_{h1} \varepsilon_1 = 0$. Thus, using (5),

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

it can be shown that the expression in (6) has probability limit zero. This can be proved also for other elements of the vector.

Hence, $\underset{n\to\infty}{\text{plim}} \frac{1}{\sqrt{n}} X^{+\prime} (V^{-1}_{\hat{\rho}^*} - V^{-1}_{\rho^*}) \varepsilon = \underset{\sim}{0}$.

B. Here we shall prove the result (3.5.40) of page 193. The result is

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (\bar{X}^{+\prime} \varepsilon_{\rho^*}) = [0, (\rho - \rho^*)\sigma^2, 0, 0, \ldots 0]' = L_1$$

<u>Proof</u>: All the elements in the vector $\frac{1}{n-1} (\bar{X}^{+\prime} \varepsilon_{\rho^*})$ excepting the second one has mean zero. The first element of this vector is

$$\frac{1}{n-1} \sum_{t=2}^{n} (\varepsilon_t - \rho^* \varepsilon_{t-1}) \tag{7}$$

This term has mean 0 and variance

$$\frac{\sigma^2}{(n-1)^2} \left\{(n-1)(1+\rho^{*2} - 2\rho\rho^*) + 2 \sum_{t=2}^{n} \sum_{\substack{t'=2 \\ t \neq t'}}^{n} \left(\rho^{|t-t'|} - 2\rho^*\rho^{|t-t'+1|} + \rho^{*2}\rho^{|t-t'|}\right)\right\} \tag{8}$$

as $n \to \infty$, the first term in the bracket tends to 0. The second term in the bracket is

$$(n-1)\{2\rho + 2\rho^2 + \ldots\} = \frac{2\rho(n-1)}{1-\rho} \tag{9}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, as $n \to \infty$, $\frac{\sigma^2}{(n-1)} \frac{2\rho}{1-\rho} \to 0$. Similarly, it can be shown that other terms in (8) also tend to zero as $n \to \infty$.

So, the expression in (7) converges in probability to zero.

The (2h-1) th, h = 2,3,...m element of the vector $\frac{1}{n-1} \bar{X}^{+\prime} \varepsilon_{\rho^*}$ is given by

$$\frac{1}{(n-1)} \sum_{t=2}^{n} x_{ht}(\varepsilon_t - \rho^* \varepsilon_{t-1}) \tag{10}$$

$$= \frac{1}{n-1} \left\{ \sum_{t=2}^{n} x_{ht} \varepsilon_t - \rho^* \sum_{t=2}^{n} x_{ht} \varepsilon_{t-1} \right\} \tag{11}$$

The terms in the bracket have 0 means. The variance of the first term in the bracket is

$$\frac{\sigma^2}{(n-1)^2} E \left\{ \sum_{t=2}^{n} x_{ht}^2 + \sum_{\substack{t=2 \\ t \neq t'}}^{n} \sum_{t'=2}^{n} x_{ht} x_{ht'} \rho^{|t-t'|} \right. \tag{12}$$

$$\leq \frac{\sigma^2}{(n-1)^2} E \sum_{t=2}^{n} x_{ht}^2 \left\{ 1 + 2\rho + 2\rho^2 + \ldots + 2\rho^{n-1} \right\} \tag{13}$$

So, as $n \to \infty$, the expression in (13) becomes

$$\lim_{n\to\infty} \frac{\sigma^2}{n-1} \cdot \lim_{n\to\infty} \left\{ \frac{1}{n-1} E \sum_{t=2}^{n} x_{h,t}^2 (1 + \frac{2\rho}{1-\rho}) \right\} \tag{14}$$

Since (3.2.2) implies that $\lim_{n\to\infty} E(\frac{1}{n-1} \sum_{t=2}^{n} x_{ht}^2)$ exists, the expression in (14) tends to zero.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Similarly, it can be shown that

$$\rho^* \operatorname*{plim}_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} x_{ht} \, \varepsilon_{t-1} = 0 \tag{15}$$

Similarly, the 2h-th (h = 2,3,...,m) element of the vector $\frac{1}{n-1} \overline{X}^{+\prime} \varepsilon_{\rho^*}$ also converges in probability to zero.

Lastly, let us consider the second element in the vector $\frac{1}{n-1} (\overline{X}^{+\prime} \varepsilon_{\rho^*})$. This element is given by

$$\frac{1}{n-1} \sum_{t=2}^{n} y_{t-1} (\varepsilon_t - \rho^* \varepsilon_{t-1})$$

$$= \frac{1}{n-1} \sum_{t=2}^{n} \{(\beta_1+\beta_2 x_{2,t-1}+\beta_3 x_{3,t-1}+\dots+\beta_k x_{k,t-1})+\varepsilon_{t-1}\}(\varepsilon_t - \rho^* \varepsilon_{t-1})$$

$$= \frac{1}{n-1} \sum_{t=2}^{n} (M_{t-1} + \varepsilon_{t-1}) (\varepsilon_t - \rho^* \varepsilon_{t-1}) \tag{16}$$

(definition of M_t has been given in Appendix 1B).

Now, $\frac{1}{n-1} E(\sum_{t=2}^{n} M_{t-1} \varepsilon_t) = 0$;

$$V\left(\frac{\sum_{t=2}^{n-1} M_{t-1} \varepsilon_t}{n-1}\right) = \frac{\sigma^2}{(n-1)^2} E(\sum_{t=2}^{n} M_{t-1}^2)$$

$$+ \frac{\sigma^2}{(n-1)^2} E \sum_{t=2}^{n} \sum_{t'=2}^{n} M_{t-1} M_{t'-1} \rho^{|t-t'|}$$

$$\leq \frac{\sigma^2}{(n-1)} \left\{\frac{1}{(n-1)} E(\sum_{t=2}^{n-1} M_{t-1}^2)\right\}(1+2\rho + 2\rho^2 + \dots + 2\rho^{n-1})$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, $$\lim_{n\to\infty} V\left(\frac{1}{n-1}\sum_{t=2}^{n-1} M_{t-1}\varepsilon_t\right) = \lim_{n\to\infty}\frac{\sigma^2}{(n-1)^2}\lim_{n\to\infty} E\sum_{t=2}^{n-1} M_{t-1}^2\left(1+\frac{2\rho}{1-\rho}\right) \quad (17)$$

Since assumption (3.2.2) implies that $\lim_{n\to\infty} E\left(\frac{1}{n-1}X'X\right)$ exists, it can be shown easily that $\lim_{n\to\infty} E\frac{1}{n-1}\sum_{t=2}^{n-1} M_{t-1}^2$ exists. Hence, from (17), we see that $V\left(\frac{1}{n-1}\sum_{t=2}^{n-1} M_{t-1}\varepsilon_t\right) \to 0$ as $n \to \infty$.

So, $$\operatorname{plim}_{n\to\infty}\frac{1}{n-1}\sum_{t=2}^{n} M_{t-1}\varepsilon_t = 0 \quad (18)$$

Under the assumption that the fourth order movement of u_t's in (3.5.1) exists, it can be shown that the probability limit of the other term in (16) is,

$$\operatorname{plim}_{n\to\infty}\frac{1}{n-1}\sum_{t=2}^{n}\varepsilon_{t-1}(\varepsilon_t - \rho^*\varepsilon_{t-1}) = (\rho - \rho^*)\sigma^2 \quad (19)$$

(for proof, vide Goldberger 1968, pp 142-155)

Hence, $$\operatorname{plim}_{n\to\infty}\frac{1}{n-1}\sum_{t=2}^{n}\overline{X}^{+}\varepsilon_{\rho^*} = (0,\ (\rho-\rho^*)\sigma^2,\ 0, 0 \ldots 0)'.$$

C. Now, we shall prove the result (3.5.53) that

$$\operatorname{plim}_{n\to\infty}\frac{1}{n-1}\overline{X}^{+\prime}\tilde{z}_{\rho^*} = [0,\ (\rho_{\tilde{z}} - \rho^*)\,\sigma^2_{\tilde{z}},\ 0, 0, \ldots 0]'$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

when the regressions of X^- on X^+ are strictly linear. In such a case, $\tilde{z}$ is independent of X^+. So, $\tilde{z}_{\rho*}$ is also independent of $\bar{X}^+$; and $E(\tilde{z}|X^+) = E(\tilde{z}) = 0$. The first element in the vector $\frac{1}{n}\bar{X}^{+\prime}\tilde{z}_{\rho*}$ is given by

$$\frac{1}{n-1}\sum_{t=2}^{n}(\tilde{z}_t - \rho^*\tilde{z}_{t-1}) \tag{20}$$

$$E\left\{\frac{1}{n-1}\sum_{t=2}^{n}(\tilde{z}_t - \rho^*\tilde{z}_{t-1})\right\} = 0$$

and $V\left\{\frac{1}{n-1}\sum_{t=2}^{n}(\tilde{z}_t - \rho^*\tilde{z}_{t-1})\right\} \to 0$ as $n \to \infty$,

under the assumptions (a) and (b) of page 196, (for proof, See Goldberger 1964, pp 142-149).

So, $$\underset{n\to\infty}{\text{plim}}\ \frac{1}{n-1}\sum_{t=2}^{n}(\tilde{z}_t - \rho^*\tilde{z}_{t-1}) = 0 \tag{21}$$

The (2h-1) th, $h = 2,3,\ldots,m$, element of the vector $\frac{1}{n-1}\bar{X}^{+\prime}\tilde{z}_{\rho*}$ is

$$\frac{1}{n-1}\sum_{t=2}^{n}x_{ht}(\tilde{z}_t - \rho^*\tilde{z}_{t-1}) \tag{22}$$

This term has mean 0 and variance

$$\frac{1}{(n-1)^2}\left\{\left(E\sum_{t=2}^{n}x_{ht}^2\right).(\tilde{z}_t - \rho^*\tilde{z}_{t-1})^2 + 2E\sum_{\substack{t=2\\t\neq t'}}^{n}\sum_{t'=2}^{n}x_{ht}x_{ht'}\,E(\tilde{z}_t - \rho^*\tilde{z}_{t-1})(\tilde{z}_{t'} - \rho^*\tilde{z}_{t'-1})\right\} \tag{23}$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Since, by (3.2.4), $\lim_{n\to\infty} E(\frac{1}{n-1} \sum_{t=2}^{n} x_{ht}^2)$ exists, and since, $\frac{1}{n} \sum_{t=2}^{n} \sum_{\substack{t'=2 \\ t\neq t'}}^{n} x_{ht}\, x_{ht'}$ cannot exceed $\frac{1}{n} \sum_{t=2}^{n} x_{ht}^2$, in absolute value, by (a) and (b) of page 196, the expression in (23) tends to zero as $n \to \infty$. Thus,

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n-1} \sum_{t=2}^{n} x_{ht}\,(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) = 0, \quad h = 2,3,\ldots m \tag{24}$$

Similarly, the 2h-th ($h = 2,3,\ldots m$) element of the vector $\frac{1}{n-1} \bar{X}^{+\prime}\, \tilde{z}_{\rho^*}$ also converges in probability to zero.

Lastly, let us consider the second element in the vector $\frac{1}{n-1} (\bar{X}^{+\prime}\, \tilde{z}_{\rho^*})$. This element is given by

$$\frac{1}{n-1} \sum_{t=2}^{n} y_{t-1}(\tilde{z}_t - \rho^* \tilde{z}_{t-1})$$

$$= \frac{1}{n-1} \sum_{t=2}^{n} (\beta_1^+ + \beta_2^+ x_{2t-1} + \ldots + \beta_m^+ x_{m,t-1} + \tilde{z}_{t-1} + \varepsilon_{t-1})(\tilde{z}_t - \rho^* \tilde{z}_{t-1})$$

$$= \frac{1}{n-1} \sum_{t=2}^{n} (M_{t-1}^+ + \tilde{z}_{t-1} + \varepsilon_{t-1})(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) \tag{25}$$

where $M_t^+ = \beta_1^+ + \beta_2^+ x_{2t} + \ldots + \beta_m^+ x_{mt}$

Under the assumptions (3.2.2) and (a) and (b) of page 196, it can be proved easily that

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} \sum_{t=2}^{n} M_{t-1}^{+} (\tilde{z}_t - \rho^* \tilde{z}_{t-1}) = 0 \tag{26}$$

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} \sum_{t=2}^{n} \tilde{z}_{t-1}(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) = (\rho_{\tilde{z}} - \rho^*)\, \sigma_{\tilde{z}}^2 \tag{27}$$

Lastly, the term

$$\frac{1}{n-1} \sum_{t=2}^{n} \varepsilon_{t-1} (\tilde{z}_t - \rho^* \tilde{z}_{t-1}) \tag{28}$$

has mean 0 and the variance

$$\frac{\sigma^2}{(n-1)^2} \Big\{ \sum_{t=2}^{n} V(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) + \sum_{t=2}^{n} \underset{t \neq t'}{\sum_{t'=2}^{n}} \rho^{|t-t'|} \times$$

$$\times \operatorname{cov}(\tilde{z}_t - \rho^* \tilde{z}_{t-1})(\tilde{z}_{t'} - \rho^* \tilde{z}_{t'-1}) \Big\} \tag{29}$$

For large n, the expression in (29) cannot exceed

$$\lim_{n\to\infty} \Big\{ \frac{\sigma^2}{(n-1)^2} \sum_{t=2}^{n} V(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) \Big\} (1 + 2\rho + 2\rho^2 + \ldots)$$

$$= \sigma^2 \left(1 + \frac{2\rho}{1-\rho}\right) \lim_{n\to\infty} \frac{1}{(n-1)^2} \sum_{t=2}^{n} V(\tilde{z}_t - \rho^* \tilde{z}_{t-1}) \tag{30}$$

= 0 under the assumptions (a) and (b) of page 196.

Thus,

$$\underset{n\to\infty}{\text{plim}} \frac{1}{n-1} (\bar{X}^{+\prime} \tilde{z}_{\rho^*}) = [0, (\rho_{\tilde{z}} - \rho^*)\sigma_{\tilde{z}}^2, 0, 0, \ldots, 0]'$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

9. Now we shall prove the existence of asymptotic mean vector and the variance covariance matrix of $\hat{\beta}^+$ given by

$$\hat{\beta}^+ = \beta^+ + (X^{+\prime} X^+)^{-1} X^{+\prime}(\tilde{z} + \varepsilon). \tag{31}$$

For this, let us first desire the asymptotic distribution of

$$\sqrt{n}\,\{\hat{\beta}^+ - \beta^+ - (X^{+\prime} X^+)^{-1} X^{+\prime}\tilde{z}\} = \{\tfrac{1}{n}(X^{+\prime} X^+)\}^{-1} \frac{1}{\sqrt{n}} (X^{+\prime} \varepsilon).$$

Let $\frac{1}{\sqrt{n}} X^{+\prime} = D = (d_1, d_2, \ldots d_n)$

Now, since $\bar{E}\{\frac{1}{n}(X^{+\prime} X^+)\} = \Sigma$, the elements of $E(DD')$ are bounded and so, the m elements of $E(d_n)$ are of order $\frac{1}{\sqrt{n}}$ [for proof, vide Theil 1971, pp.380-381].

Under the assumption that ε's are independently and identically distributed with zero mean and variance σ^2, the vector $d_n \varepsilon_n$ has zero mean and covariance matrix $\sigma^2 E(d_n d_n')$ which is of order $\frac{1}{n}$. So, the characteristic function of $d_n \varepsilon_n$ is

$$1 - \tfrac{1}{2}\sigma^2 t' E(d_n d_n')t + o(\tfrac{1}{n})^{1/}. \tag{32}$$

Thus, the vector $\frac{1}{\sqrt{n}} X^{+\prime}\varepsilon = d_1 \varepsilon_1 + \ldots + d_n \varepsilon_n$ has the following characteristic function :

1/ $o(\frac{1}{n})$ means of order lower than $\frac{1}{n}$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$\phi_n(t) = \prod_{i=1}^{n} [1 - \frac{1}{2} \sigma^2 t' E(d_i d_i')t + o(\frac{1}{n})]. \quad (33)$$

Taking logarithms and expanding for large n, we have

$$\log \phi_n(t) = \sum_{i=1}^{n} \log [1 - \frac{1}{2} \sigma^2 t' E(d_i d_i')t + o(\frac{1}{n})]$$

$$= -\frac{1}{2} \sigma^2 t' E(\sum_{i=1}^{n} d_i d_i')t + o(1)$$

$$= -\frac{1}{2} \sigma^2 t' E(\frac{1}{n} X^{+\prime} X^{+})t + o(1).$$

So, $$\lim_{n\to\infty} \log \phi_n(t) = -\frac{1}{2} \sigma^2 t' \bar{E} (\frac{1}{n} X^{+\prime} X^{+})t$$

$$= -\frac{1}{2} \sigma^2 t' \Sigma t \quad (34)$$ 2/

This corresponds asymptotically to the multinormal distribution with 0 mean vector and variance-covariance matrix $-\frac{1}{2} \sigma^2 t' \Sigma t$.

Since $\operatorname{plim}_{n\to\infty} (\frac{1}{n} X^{+\prime} X^{+}) = \Sigma$ (by (3.2.2)) (a positive definite matrix)

$\therefore$ $(\frac{1}{n} X^{+\prime} X^{+})^{-1} (\frac{1}{\sqrt{n}} X^{+\prime} \varepsilon)$ converges in distribution to

Σ^{-1}(asymptotic distribution of $\frac{1}{\sqrt{n}} X^{+\prime} \varepsilon$) (*vide* Theil 1971, pp 370-371).

2/ $e^{-\frac{1}{2} \sigma^2 t'\Sigma t}$ is continuous at $t = 0$ as is required.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

So, asymptotic $E\left\{\frac{1}{n}(X^{+\prime}X^{+})^{-1}\left(\frac{1}{\sqrt{n}}X^{+\prime}\varepsilon\right)\right\} = \Sigma^{-1}\cdot\underset{\sim}{0} = \underset{\sim}{0} = \overline{E}\left\{\left(\frac{1}{n}X^{+\prime}X^{+}\right)^{-1} \times\right.$

$\left.\times\left(\frac{1}{\sqrt{n}}X^{+\prime}\varepsilon\right)\right\}$ (35)

and asymptotic $V\left\{\frac{1}{n}(X^{+\prime}X^{+})^{-1}\left(\frac{1}{\sqrt{n}}X^{+\prime}\varepsilon\right)\right\} = \sigma^2\Sigma^{-1} = \sigma^2\,\overline{E}\left\{\left(\frac{1}{n}X^{+\prime}X^{+}\right)^{-1}\right\}$

$= \overline{V}\left\{\frac{1}{n}(X^{+\prime}X^{+})^{-1} \times\right.$

$\left.\times\left(\frac{1}{\sqrt{n}}X^{+\prime}\varepsilon\right)\right\}$ (36)

So, from (35),

asy $E\left\{\hat{\beta}^{+}-\beta^{+}-(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\}$

$= \overline{E}\left\{\hat{\beta}^{+}-\beta^{+}-(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\} = \underset{\sim}{0}$

$\therefore\ \overline{E}(\hat{\beta}^{+}-\beta^{+}) = \overline{E}\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\} = \overline{E}\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}(X-X^{+}\overline{P})\right\}\beta$

$= \overline{E}\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}X\right\} - \overline{E}(X^{+\prime}X^{+})^{-1}\overline{E}(X^{+\prime}X)\beta$

$= \overline{E}(X^{+\prime}X^{+})^{-1}\overline{E}(X^{+\prime}X) - \overline{E}(X^{+\prime}X^{+})^{-1}\overline{E}(X^{+\prime}X)\beta = \underset{\sim}{0}$

(37)

asy $V\left\{\hat{\beta}^{+}-\beta^{+}-(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\}$

$= \overline{V}\left\{\hat{\beta}^{+}-\beta^{+}-(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\}$

$= \frac{\sigma^2}{n}\Sigma^{-1}$ (38)

$\therefore\ \overline{V}(\hat{\beta}^{+}-\beta^{+}) = \frac{\sigma^2}{n}\Sigma^{-1} - \overline{V}\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\} + 2\,\overline{\text{cov}}\left[(\beta^{+}-\beta^{+})\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\}\right]$ [3]

$= \frac{\sigma^2}{n}\Sigma^{-1} + \overline{V}\left\{(X^{+\prime}X^{+})^{-1}X^{+\prime}\tilde{z}\right\}$ (using the expression for $(\hat{\beta}^{+}-\beta^{+})$ in (31)). (39)

3/ $\overline{\text{cov}}(a^{(n)}) = \lim_{n\to\infty}$ covariance $(a^{(n)})$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

APPENDIX 3

Let us consider the term

$$\left[\frac{\lambda\{1-(1-n\lambda)\rho_{12}^2\}}{n\{1-(1-\lambda)\rho_{12}^2\}^3}-\frac{(1-\rho_{12}^2)n}{(n+1)^2}\right] \qquad (1)$$

$$\frac{\partial}{\partial\lambda}\left[\frac{\lambda\{1-(1-n\lambda)\rho_{12}^2\}}{n\{1-(1-\lambda)\rho_{12}^2\}^3}-\frac{(1-\rho_{12}^2)n}{(n+1)^2}\right]$$

$$=\frac{1-(1-n\lambda)\rho_{12}^2}{n\{1-(1-\lambda)\rho_{12}^2\}^3}+\frac{n\lambda\rho_{12}^2}{n\{1-(1-\lambda)\rho_{12}^2\}^3}-\frac{3\lambda\{1-(1-n\lambda)\rho_{12}^2\}\rho_{12}^2}{n\{1-(1-\lambda)\rho_{12}^2\}^4}$$

$$=\frac{(1-\rho_{12}^2+2n\lambda\rho_{12}^2)\{1-(1-\lambda)\rho_{12}^2\}-3\lambda\rho_{12}^2(1-\rho_{12}^2+n\lambda\rho_{12}^2)}{n\{1-(1-\lambda)\rho_{12}^2\}^4}$$

$$=\frac{(1-\rho_{12}^2)[\{1-(1-\lambda)\rho_{12}^2\}+\rho_{12}^2(1-\rho_{12}^2)(2n\lambda-3\lambda)-n\lambda^2\rho_{12}^4]}{n\{1-(1-\lambda)\rho_{12}^2\}^4}$$

$$>0 \quad \text{for} \quad n>2 \qquad (2)$$

So, for $n > 2$, the expression in (1) increases as λ increases

$$\frac{\partial}{\partial n}\left[\frac{\lambda\{1-(1-n\lambda)\rho_{12}^2}{n\{1-(1-\lambda)\rho_{12}^2\}^3}-\frac{(1-\rho_{12}^2)n}{(n+1)^2}\right]$$

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

$$= \frac{\lambda^2 \rho_{12}^2}{n\{1-(1-\lambda)\rho_{12}^2\}^3} - \frac{\lambda\{1-(1-n\lambda)\rho_{12}^2\}}{n^2\{1-(1-\lambda)\rho_{12}^2\}^3} + \frac{(1-\rho_{12}^2)(n-1)}{(n+1)^2}$$

$$= \frac{-(1-\rho_{12}^2)[\lambda - \tilde{n}\{1-(1-\lambda)\rho_{12}^2\}^3]}{n^2\{1-(1-\lambda)\rho_{12}^2\}^3}, \quad \tilde{n} = \frac{n^2(n-1)}{(n+1)^2} \qquad (3)$$

When, $\lambda = .9$, this is negative for $\rho = .1, .2, .3$ ~~and~~ upto certain value of n and ~~to~~ after this value it is position. The expression in (1) is negative for $\rho = .4, .5, \ldots, .7$.

So, for $\lambda = .9$, the expression in (1) decreases as n increases upto certain value of n and then it begins to increase as n increases for $\rho = .1, .2, .3$ and decreases as n increases for $\rho = .4, .5, \ldots, .7$.

The sufficient condition for $MSE(\hat{\beta}_{1p}) < MSE(\hat{\beta}_1^+)$ is

$$\frac{\lambda\{1-(1-n\lambda)\rho_{12}^2\}}{n\{1-(1-\lambda)\rho_{12}^2\}^3} - \frac{n(1-\rho_{12}^2)}{(n+1)^2} < \rho_{12}^2 \frac{n}{n+1} \qquad (4)$$

To verify this, we have to calculate the expression in (1) only for the following values of n, ρ_{12}^2 and λ.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Table showing the value of the expression in (1) for different values of ρ_{12}, λ and n.

$\lambda = .9$

n	ρ_{12}^2	Expression in (1)	$\frac{n}{n+1}\rho_{12}^2$
(50, 15, 10)	.1	(.083, .088, .093)	(.098, .093, .091)
(50, 15, 10)	.2	(.172, .176, .183)	(.196, .187, .182)
(50, 15, 10)	.3	(.267, .272, .277)	(.294, .281, .273)
(15, 10)	.4	(.372, .378)	(.375, .364)
(15, 20)	.5	(.479, .476)	(.468, .476)
(30, 40)	.6	(.584, .586)	(.599, .594)
100	.7	.704	.693
~~80~~	~~.8~~	~~.808~~	

For $\lambda = .7$, $n = 100$, $\rho_{12}^2 = .7$, the expression in (1) is .686

For $\lambda = .8$, $n = 20$, and $\rho_{12}^2 = .6$, the value of the expression in (1) becomes $.566 \leq \frac{n}{n+1}\rho_{12}^2 = .571$

Since for $n > 2$, the expression in (1) increases as λ increases and since for $\rho_{12}^2 = .1, .2, .3$ and $\lambda = .9$, the expression in (1) decreases as n increases upto certain value of n and then it ~~in (1)~~ increases as n increases and for $\rho_{12}^2 = .4, .5, \ldots .7$ and for $\lambda = .9$, the expression in (1) decreases as n increases, from the above table, we find that the condition (4) holds for values of $n = 15, 20, 25, \ldots, 50$, $\rho_{12}^2 = .1, .2, .4$, $n \geq 20$ $\rho_{12}^2 = .5$ and $n \geq 40$ $\rho_{12}^2 = .6$ and $\lambda = .1, \ldots .9$ and $n = 25, 29, \ldots, 50$, $\lambda = .1, .2, \ldots, .8$, $\rho_{12}^2 = .6$.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

REFERENCES

Aigner, D.J. (1973). Regression with a binary independent variable subject to errors of observations. Journal of Econometrics, 1, pp 49-60.

_______ (1974). MSE dominance of least squares with errors of observations. Journal of Econometrics, 2, pp. 365-372.

Aitken, A.C. (1934/35). On least squares and linear combination of observations. Proceeding of Royal Statistical Society, Edinburgh, 55, pp 42-48.

Allen, H.V. (1938). A theorem concerning the linearity of regression. Statistical Research Memoire 2, pp 60-68.

Almon, S. (1965). The distributed lag between capital appropriation and expenditures, 30, pp 178-196.

_______ (1968). Lags between investment decisions and their causes. Review of Economics and Statistics, 50, pp 193-206.

Amemiya, T. (1966). Specification analysis in the estimation of parameters of a simultaneous equation model with autoregressive residuals, Econometrical 34, pp. 283-306.

Anderson, R.L. (1941). Serial correlation in the analysis of time series. Unpublished thesis, Iowa State College, Ames, Iowa.

_______ (1942). Distribution of the serial correlation coefficient. Annals of Mathematical Statistics, 13, pp 1-13.

Anderson, T.W. (1957). An introduction to multivariate Statistical Analysis. John Wiley and Sons, Inc.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Anderson, W.H.L. (1964). Corporate finance and fixed investment, an econometric study. Boston : Division of Research, Graduat School of Business Administration, Harvard University, 1964.

________ (1967). Business fixed investment : a marriage of fact and fancy, in determinants of investment behaviour, ed. by R. Ferber. Universities-National Bureau Conference Series, No.18, New York : Columbia University Press for the National Bureau of Economic Research.

Bartlett, M.S. (1935). Some aspects of the time correlation problem in regard to tests of significance, Journal of Royal Statistical Society, 98, pp 536-543.

________ (1949). Fitting a straight line when both variables are subject to error, Biometrics 5, pp 207-212.

Berkson, J. (1950). Are there two regressions ? Journal of the American Statistical Association, 45, pp 164-180.

Blattberg, R.C. (1973). Evaluation of the power of the Durbin-Watson statistics for non-first order serial correlation alternatives. Review of Economics and Statistics, 56, pp 508-516.

Byron, R.P. (1972). Testing for misspecification in econometric systems using full information. International Economic Review, 13, pp 745-756.

Carlson, F.D., Sobel, E. and Watson, G.S. (1966). Linear functional relationships between variables affected by errors, Biometrics, 22, pp 252-267.

Chai, J. (1971). Correlated measurement errors and least squares estimators of the regression coefficient. Journal of the American Statistical Association, 66, pp 478-483.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Champernowne, D.G. (1948). Sampling theory applied to auto-regressive sequences. Journal of the Royal Statistical Society, Series B, 10, 204-242.

Chaudhuri (Mukherjee), M. (1977). Autocorrelated disturbances in the light of specification analysis : Journal of Econometrics 5, 301-313.

Clutton, Brock, M. (1967). Maximum likelihood distribution for estimation of functions when both variables are subject to error. Technometrics, 9, 261-269.

Cochran, W.G. (1968). Errors of measurement in statistics, Technometrics, 10, 637-666.

__________ (1970). Some effects of errors of measurement on multiple correlation. Journal of the American Statistical Association, 65, 22-34.

__________ (1971). Some effects of errors of measurement on linear regression. Proceedings of the Sixth Berkley Symposium on Mathematical Statistics and Probability, 527-539.

Cochrane, D. and Orcutt, G.H. (1949). Applications of least squares regressions to relationships containing auto-correlated errors terms. Journal of the American Statistical Association, 44, 32-61.

Copas, J.B. (1966). Monte Carlo results for estimation in a suitable Markov series. Journal of Royal Statistical Society, Series A, 129, 110-116.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Cragg, J.G. (1968). Some effects of incorrect specification on the small sample properties of several simultaneous equation estimators. International Economic Review, 9, 63-86.

Daniels, H.E. (1956). The approximate distribution of serial correlation coefficients. Biometrika, 43, 169-185.

Dixon, W.J. (1944). Further contributions to the problem of serial correlation. Annals of Mathematical Statistics, 15, 119-144.

Dhrymes, P.J. (1970). Full information estimation of dynamic simultaneous equation models with autoregressive errors, Discussion paper no. 203. The Whartor School of Finance and Commerce, University of Pennsylvania.

______ (1970). Econometrices. Statistical foundations & Applications. Harper International Edition, 385-386.

Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression I. Biometrika 37, 409-428.

______ (1951). Testing for serial correlation in least squares regression II. Biometrika 38, 159-178.

Durbin, J. (1954). Testing for serial correlation in systems of simultaneous regression equations. Biometrika 44, 370-377.

______ (1960). Estimation of parameters in time series regression models. Journal of the Royal Statistical Society, Series B, 22, 139-153.

______ (1970). An alternative to the bounds test for testing serial correlation in least squares regression. Econometrica, 38, 422-429.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Eisner, R. (1962). Investment plans and realizations. American Economic Review, 52, 190-203.

Fair, Ray, C. (1970). The estimation of simultaneous equation models with lagged endogenous variables and first order serially correlated errors, Econometrica 38, 507-516.

_________ (1972). Efficient estimation of simultaneous equation with autoregressive errors by instrumental variables. Review of Economics and Statistics, 54, 444-449.

Federov, V.V. (1974). Regression problems with controllable variables subject to error. Biometrika 61, 49-56.

Feldstein, M.S. (1973). Multicollinearity and the mean square error of alternative estimators, Econometrica 41, 337-348.

_________ (1974). Errors in variables : A consistent estimator with smaller MSE in finite samples. Journal of the American Statistical Association 69, 990-996.

Fisher, F.M. (1966). The relative sensitivity to specification error of different k class estimators. Journal of the American Statistical Association, 61, 345-356.

Fix, E. (1949). Distributions which lead to linear regressions. Proceedings (first) of Berkley Symposium on Mathematical Statistics and Probability, University of California Press, 79.

Frissch, R. (1934). Statistical confluence analysis by means of complete regression systems, OSLO University Economics Institute Publication no. 5.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Geary, R.C. (1949). Determination of linear relations between systematic parts of variables with errors of observations the variances of which are unknown. Econometrica 17,30-58.

______ (1953). Nonlinear functional relationship between two variables when one veriable is controlled. Journal of the American Statistical Association, 48, 94-103.

Geisser, S. (1956). The modified mean square successive difference and related statistics. Annals of Mathematical Statistics, 27, 819-924.

Gibson W. M and Jowett G.H (1957(a)) Three group regression analysis part I: Applied statistics 6, 114
(1957(b)) —— part II —— 6, 18

Gilbert, R.F. (1969). The demand for money : an analysis of specification error, Unpublished Ph.D. Thesis, Michigan State University, East Lansing, Michigan.

Glahe, and Hunt, (1970). Small sample properties of simultaneous equation least absolute estimators vis-a-vis least squares estimators, Econometrica 38, 742-753.

Glesser, L.J. and Watson, G.S. (1973). Estimation of a linear transformation, Biometrika 60, 525-534.

Goldberger, A.S. (1964). Econometric theory. John Wiley and Sons. Inc.

______ (1972). Maximum likelihood estimation of regression models containing unobserable variables :International Economic Review 13, 1-15.

Gregor, M.C.(Jr.) (1960). An appropriate test for serial correlation inpolynomial regression. Biometrika 47, 111-120.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Grether, D.W. and Maddala, G.S. (1973). Errors in variables and serially correlated disturbances in distributed lags. Econometrica, 41, 255-262.

Griliches, Z. (1957). Specification bias in estimates of production functions . Journal of Farm Economics, 39, 8-20.

_______ (1961). A note on serial correlation bias in estimates of distributed lags. Econometrica, 29, 65-73.

_______ (1967). Distributed lags : A survey. Econometrica, 35, 16-49.

_______ (1970). Notes on the role of education in production functions and growth accounting in. W. Lee Hansen, ed. Education, Income and Human Capital (National Bureau of Economic Research, New York).

_______ (1974). Errors in variables and other unobserables, Econometrica, 42, 971-998.

Griliches, Z. and Rao, P. (1964). Small sample properties of several two stage regression methods in the context of autocorrelated errors. Journal of the American Statistical Association, 64, 253-272.

Griliches, Z. and Ringstad, V. (1970). Errors in the variables bias in nonlinear contexts. Econometica 38, 368-370.

__________ and _______ (1971). Economics of scale and the form of the production function North-Holland, Amsterdam .

Guilkey, D.K. (1974). Alternative tests for a first order vector autoregressive error specification. Journal of Econometrics, 2, 95-104.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Guilkey, D.K. and Schmidt, P. (1973). Estimation of seemingly unrelated regressions with vector autoregressive errors. Journal of the American Statistical Association, 68, 642-647.

Haitovsky, Y. (1973). Regression estimation from grouped oberservations. Hafner, New York .

Halperin, M. (1961). Fitting of straight lines and prediction when both variables are subject to error. Journal of the American Statistical Association, 56, 657-689.

Halperin, M. and Gurian, J. (1971). Estimation in straight line regression when both variables are subject to error. Journal of the American Statistical Association,66, 587-589.

Hannan, E.J. (1957). Testing for serial correlation in least squares regression, Biometrika, 44, 57-66.

Hannan, E.J. and Terrel, R.D. (1968). Testing for serial correlation after least squares regression, Econometrica, 36, 133-150.

Hart, B. I. (1942(a)). Tabulation of the probability of the ratio of the mean square difference to the variance. Annals of Mathematical Statistics, 13, 207-214.

________ (1942(b)). Significance levels for the ratio of mean square successive difference to the variance. Annals of Mathematical Statistics, 13, 445-447.

Hendry, D.F. (1971). Maximum likelihood estimation of systems of regression equations with errors generated by a vector autoregressive process. International Economics Review, 12, 257-272.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Henshaw, R.C., Jr. (1966). Testing single equation least squares regression models for autocorrelated disturbances, Econometica, 34, 646-660.

Hoch, I. (1955). Estimation of production function parameters and testing for efficiency. Econometrica, 28, 326.

Hooper J.W and Theil H (1958). The extension of Wald's method of fitting straight lines to multiple regression. Review of International Statistical Institute, 26, 37-47

Jenkins, G.M. (1954). Tests of hypotheses in the linear autoregressive model I: null hypotheses distributions in Yule scheme. Biometrika, 41, 405-419.

________ (1956). Tests of hypothesis in the linear autoregressive model II: null distribution for higher order schemes: non-null distributions. Biometrika 42, 186-199.

Johnston, J. (1972). Econometric methods. 2nd edition, McGraw Hill Book Company.

Jorgenson, D.W., Hunter, J. and Nadiri, M.I. (1970). A comparison of alternative econometric models of quarterly investment behaviour. Econometrica, 38, 187-212.

Karni, Edi and Weissman, Istiay (1974). A consistent estimator of a slope in a gression in errors in variables, Journal of American Statistical Association, 69, 211-213.

Kendall, M.G. and Stuart, A. (1967). The advanced theory of statistics, Vol.2, Charles Griffin and Co., London.

Kloek, T. (1961). Note on convenient matrix notations in multivariate statistical analysis and in the theory of linear aggregation. International Economic Review 2, 351-360.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Klbek, T. (1975). Note on a large sample result in specification analysis. Econometrica, 43, 933-937.

Koerts, J. and Abrahamse, A.P. (1970). The correlation coefficient in the general linear model. European Economic Review 1, 401-427.

Koopmans, T. (1937). Linear regression analysis of economic time series Haarlem. De Erven F. Bohn N.V.

Koyck, L.M. (1954). Distributed lags and investment analysis, Amsterdam : North-Holland Publishing Company.

Krishhaiah, P.R. and Murthy, V.K. (1966). Simultaneous tests for trend and serial correlation for Gaussian Markov residuals, Econometrica 34, 472-480.

Laha, R.G. (1956). On a characterization of the stable law with finite expectation. Annals of Mathematical Statistics, 27, 187-195.

Langaskens, Y. and Van Ryckeghem, W. (1974). A new method to estimate measurement errors in national account statistics: The Belgian case. International Statistical Review 42, 291-298.

Lankiralle, K.N. (1974). Some effects of errors in variables on correlations. Presented at the Thirteenth Conference of Indian Econometric Society.

______ (1975). Errors in variables stochastic linear regression model. Presented at the Fourteenth Conference of Indian Econometric Society.

Lee, K.W. (1972). An international study of manufacturing production functions. Unpublished Doctoral Thesis, Michigan State University, East Lansing, Michigan.

Leeuw, F. de (1962). The demand for capital goods by manufacturers, A study of quarterly time series. Econometrica 30, 178-196.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Levi, M.D. (1973). Errors in the variables bias in the presence of correctly measured variables. Econometrica 41, 985-986.

Lindley, D.V. (1947). Regression lines and the linear functional relationship. Journal of the Royal Statistical Society, Supplementary 9, 219-244.

________ (1953). Estimation of a functional relationship. Biometrika, 40, 47-49

Liviatan, N. (1961). Errors in variables and Engel curve analysis. Econometrica 29, 336-362.

________ (1963). Consistent estimation of distributed lags. International Economic Review 4, 44-52.

Loeb Peter, D. (1976). Specification error tests and investment functions. Econometrica 44, 185-194.

Lund, P.J. and Holden, K. (1968). An econometric study of private sector gross fixed capital formation in United Kingdom, 1923-1938, Oxford Economic Papers 20, 56-73.

Lyttkens, E. (1964). Standard errors of regression coefficients by autocorrelated residuals : Chapter 4, Econometric model building : Essays on the causal chain approach. Amsterdam : North Holland Publishing Company.

Madansky, A. (1959). Fitting of straight lines when both variables are subject to error. Journal of the American Statistical Association, 54, 133-205.

Malinvaud, E. (1966). Statistical methods of econometrics. North Holland, Amsterdam, 425.

Mallios, U.S. (1969). A general application of instrumental variables estimation of straight line relations when both variables are subject to error. Technometrics 11, 255-263.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Marriot, F.H.C. and Pope, J.A. (1954). Bias in the estimation of autocorrelations. Biometrica 41, 390-402.

Mc.Callum, B.T. (1972). Relative asymptotic bias from errors of omission and measurement, Econometrica, 40, 757-758.

Meyer, J.R. and Glauber, R.R. (1974). Investment decisions, Economic forecasting and public policy. Boston. Division of Research, Graduate School of Business Administration, Harvard University.

Morton, B.R. (1964). Numerical approximation. Routledge and Kegan Paul, London, 4.

Mukherjee, M. (1974). A reconsideration of the problem of autocorrelated disturbances from the point of view of specification analysis : Presented at the 14th Indian Econometric Conference. Technical Report No. Econ/18/74, Indian Statistical Institute, Calcutta.

Mukherjee, M. and Lahiri, K. (1975). Specification analysis and autocorrelated variables : Presented at the 15th Indian Econometric Conference. Technical Report No. Econ/12/75, Indian Statistical Institute, Calcutta.

Mukherjee, M. and Chatterjee, G.S. (1974). On the validity of NSS estimates of consumption expenditure. Sankhyā, 36, Series C, 139-147.

Mundlak, Y. (1961). Empirical production function free of management bias. Journal of Farm Economics, 43, 44-56.

Nair K.R and Shrivastava M.P (1942). On a simple method of curve fitting Sankhya 6, 121

Neumann, J. Von (1941). Distribution of the ratio of the mean square successive difference to the variance. Annals of Mathematical Statistics, 12, 367-395.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Neumann, J. Von (1942). A further remark concerning the distribution of the ratio of mean square successive difference to the variance. Annals of Mathematical Statistics, 13, 86-88.

Neyman, J. and Scott, E.L. (1951). On certain methods of estimating the linear structural relations. Annals of Mathematical Statistics, 22, 352-361.

Orcutt, G. H. and Winokur, H.S. Jr. (1969). First order autoregression : Inference, estimation and prediction. Econometrica 37, 1-14.

Press, S.J. and Brooks, R.B. (1969). Testing for serial correlation in regression. Report no. 6911, Centre for Mathematical Studies in Business and Economics. The University of Chicago.

Ramsey, J.B. (1968). Tests for specification errors in classical linear least squares regression analysis. Ph.D. Thesis, University of Wisconsin, Madison, Wisconsin.

________ (1969). Tests for specification errors in classical linear least squares regression analysis. Journal of Royal Statistical Society, Series B, Part 2, 350-371.

________ (1971). Limiting forms for demand functions : Preliminary tests of some specific hypotheses. Paper presented at the meeting of the Western Economics Association, Vancouver, B.C.

________ (1974). Classical model selection through specification error tests : in Frontiers of Econometrics edited by Paul Zarembka, Academic Press, 13-47.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Ramsey, J.B. and Gilbert, R. (1969). A monte carlo study of some small properties of tests for specification error : Econometrics Working Paper no.6813, Michigan State University, East Lansing, Michigan.

________ and ________ (1972). Some small sample properties of tests for specification error. Journal of American Statistical Association 67, 180-186.

Ramsey, J.B. and Zarembka, P. (1971). Specification error tests and alternative functional form of aggregate production function. Journal of the American Statistical Association 66, 471-477.

Ramsey, J.B. and Schmidt, P. (1976). Some further results on the use of OLS and BLUS residuals in specification error tests. Journal of the American Statistical Association, 71, 389-390.

Rao, K.K. (1968). A transformation to circumvent the problems of autocorrelation. Econometrica, 36, 93-96.

Rao, P. (1971). Some notes on misspecification in multiple regressions. American Statistician 25, no.5, 37-39.

Reiersol, O. (1941). Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 9, 1-24.

Richardson, D.H. and Wu De-Min (1970). Least square and grouping method estimators in the errors in variable model. Journal of the American Statistical Association 65, 724-748.

Robinson, P.M. (1974). Identification, estimation and large sample theory for regressions containing unobservable variables. International Economic Review 15, 680-692.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Rubin, H. (1945). On the distribution of the serial correlation coefficient. Annals of Mathematical Statistics 16, 211-215.

Sargan, J.D. (1958). The estimation of economic relationships using instrumental variables. Econometrica 26, 393-415.

________ (1961). The maximum likelihood estimation of economic relationships with autoregressive residuals. Econometrica, 29, 414-426.

Sargan, J.D. and Mikhail, W.M. (1971). A general approximation to the distribution of instrumental variables estimates. Econometrica 39, 131-169.

Sarkar, N. (1978). A large sample test for identification of the causes of autocorrelation and a general method of estimation. Presented at the 16th Indian Econometric Conference. Technical Report No. ERU/7/79, Indian Statistical Institute, Calcutta.

Scheffe, H. (1958). Fitting straight lines when one variable is controlled. Journal of the American Statistical Association 53, 106-117.

Schmidt, P. (1974). A note on Theil's minimum standard error criterion when the disturbances are autocorrelated. Review of Economics and Statistics 56, 122-123.

Shenton, L.R. and Johnson, W.L. (1965). Moments of a serial correlation coefficient. Journal of Royal Statistical Socieity, 27, Series B, 308-320.

Slater, L.J. (1960). Confluent hypergeometric functions, Cambridge University Press, Cambridge.

Solari, M.E. (1969). The maximum likelihood solution of the problem of estimating a linear functional relationship. Journal of the Royal Statistical Socieity, B, 31, 372-375.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Sprent, P. (1966). A generalised least squares approach to linear functional relationships. Journal of the Royal Statistical Society, B, 28, 278-297.

Stuart, A. (1954). The correlation between variate values and ranks in samples from à continuous distribution. British Journal of Statistical Psychology 7, 37-44.

Summers, R. (1965). A capital intensive approach to the small sample properties of various simultaneous equation estimators. Econometrica 33, 1-41.

Taylor, L.D. and Wilson, T.A. (1964) : Three-pass least squares : A method for estimating models with a lagged dependent variables : Review of Economics and Statistics, 46;329-346.

Theil, H. (1954). Linear aggregation of economic relations : Amsterdam : North Holland Publishing Company.

________ (1965). The analysis of disturbances in regression analysis. Journal of the American Statistical Association, 60, 1067-1079.

________ (1968). A simplification of the BLUS procedure for analysing regression disturbances. Journal of the American Statistical Association, 63, 242-251.

________ (1971). Principles of Econometrics. John Wiley and Sons. Inc.

Theil, H. and Nagar, A.L. (1961). Testing the independence of regression disturbances. Journal of the American Statistical Association 56, 793-806.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Theil, H. and Van Yzeren, J. (1956). On the efficiency of Wald's method of fitting straight lines : Review of the International Statistical Institute 24, 17-26.

Tinsely, P. A. (1967). An application of variable weight distributed lags. Journal of American Statistical Association 62, 1277-1289.

Toro-Vizerrondo, C. and Wallace, T.C. (1968). A test of the mean square error criterion for restrictions in linear regressions. Journal of the American Statistical Association 63, 558-572.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. Annals of Mathematical Statistics 11, 284-300.

Wallace T.C. and Toro-Vizcarrondo, C. (1969). Tables for the mean square error test for exact linear restrictions in regression. Journal of the American Statistical Association 64, 1649-1663.

Wallis, K.P. (1967). Lagged dependent variables and serially correlated errors : A reappraisal of three-pass least squares. Review of Economics and Statistics 49, 555-567.

Ware, S.H. (1972). Fitting straight lines when both variables are subject to error and ranks of the mean are known. Journal of American Statistical Association 67, 891-897.

Warren, R.D., White, J.K. and Fuller, W.A. (1974). An errors in variables analysis of managerial role performance. Journal of the American Statistical Association 64, 886-893.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Watson, G.S. (1956). On the j int distribution of the circular serial correlation coefficients. Biometrika, 43, 101-168.

Waud, R.N. (1966). Small sample bias due to misspecification in the partial adjustment and adaptive expectations models. Journal of the American Statistical Association 61, 1130-1152.

White, J.S. (1957). Approximate moments for the serial correlation coefficient. Annals of Mathematical Statistics 28, 798-802.

_______ (1961). Asymptotic expansions for the mean and variance of the serial correlation coefficient. Biometrika, 48, 85-94.

White, H. (1976). Using least squares to approximate unknown regression functions. Discussion Paper, 76-17 Department of Economics, University of Rochester.

Wickens, M.R. (1972). A note on the use of proxy variables, Econometrica, 40, 759-761.

Wold, H. (1950). On least square regression with autocorrelated variables and residuals. Bulletine de l' Institut International de Statistique 32, No.2.

Wold, H. and Jureen, L. (1953). Demand analysis, New York : John Wiley and Sons, 209-215.

Wu, De-Min (1973). Alternative tests of independence between stochastic regressors and disturbances, Econometrica, 41, 733-750.

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor

Yule, G.U. (1926). Why do we sometimes get nonsense correlations between time series etc. ? Journal of the Royal Statistical Socieity, 89, 1-64.

Zellner, A. (1970). Estimation of regression relationships containing unobservable independent variables. International Economic Review, 11, 441-454.

__________ (1971). An introduction to Bayesian inference in econometrics. John Wiley and Sons. Inc.

Zellner, A. and Geisel, M.S. (1968). Analysis of distributed lag models with applications to consumption function estimation. Paper presented to European meeting of the Econometric Society.

INDIAN STATISTICAL INSTITUTE LIBRARY
No. T 032
Date 2 FEB 1984
CALCUTTA

PDF compression, OCR, web optimization using a watermarked evaluation copy of CVISION PDFCompressor