THE ASSIMILATION OF "BEST-PRACTICE" ECONOMETRIC TECHNOLOGY

BY

JEHANBUX D. FERZANDI

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of Masters of Arts.

Department of Economics
McGill University
Montreal

ABSTRACT

The purpose of this study is threefold: Firstly, to provide a succinct overview of available "best-practice" technologies for estimation in linear regression. Secondly, we look at what economists do. Here we ask the important question of how quickly advances in statistical theory are assimilated into research activities by economists. Finally we consider certain areas in economics that would benefit from use of "best-practice" methods.

Our results show that there is an alarmingly large gap between advances in statistical theory and its assimilation into economic research.

RESUMÉ

Cette étude à trois buts principaux: d'abord pour faire l'inventaire des "meilleures techniques" (best-practice technologies) existantes pour l'estimation en régression lineáre. Ensuite, pour jetter un regard sur ce que font les économistes dans ce domaine. On se pose la question importante concernant la rapidité avec laquelle les économistes assimilent dans leurs activités de recherches les avances faites en théorie statistique. Finalement, nous considerons certain domaines dans la science économeque qui peuvent bénéficier le plus en utilisant ces meilleures techniques (best-practice methods).

Nos résultats démontrent qu'il éxiste un écart considerable entre les avances faites en théorie statistique et leur assimilation en recherche économique.

ACKNOWLEDGEMENTS

I would like to express my appreciation and thanks to Professor J.C.R. Rowley, for his thoughtful guidance and support throughout the preparation of this thesis. His continued encouragement and patient reading of the many drafts presented to him has made this work possible. To him I am sincerely grateful.

In addition, I extend my appreciation to my wife Sharlene, for her understanding and help in typing this thesis. I am also very grateful to R. Jain who in many ways aided in the final preparation. Finally I would like to thank Professor S. Gupta for reading my final draft and providing suggestions.

TABLE OF CONTENTS

The Assimilation of "Best-Practice" Econometric Technology

CHAPTER ONE

Introduction

In economic models involving heterogeneous capital, the spread of new techniques in production is often discussed. We cannot assume such techniques will be instantly adopted as soon as they appear. This would be contrary to available evidence on the spread of new techniques and it would neglect the real cost of adjustment as existing procedures are dislocated and amended. A superior approach is to view new techniques as being assimilated at rates that will vary through time according to their embodiment in new capital equipment and to the pace of the general movement in economic activity, especially as it impinges on capital accumulation. This process serves as a useful analogy for introducing the important issue of how quickly advances in statistical theory find expression in the research activities of economists. As with capital, the concern is not with small-scale activities in research laboratories but rather with the wider adoption of techniques outside the confines of initial developmental sources.

There are three significant elements in the present economic context that provide a tentative proposition concerning the speed of assimilation.The general momentum of empirical research in economics and the availability of software packages that are constantly being updated to include new statistical procedures would both seem to be indicative of a favourable environment for the assimilation of statistical advances. To these two characteristics, we should add the increasing stress being attached to statistical criteria for evaluating research studies in professional literature and in applications to funding agencies. Thus we might reasonably presume that rates of assimilation are rapid with small delays between the first appearance of new statistical techniques and their common use by economists in applied research. Unfortunately, evidence that we shall present below will reveal that this presumption is quite unwarranted. The true picture of econometric practice (through all of the economics profession rather than among the much smaller coterie of econometricians with narrower focus) is one of senility. Published studies reveal no sign of rapid advance in the assimilation of statistical techniques. Indeed they show considerable persistence with early methods, all of which stem from developments prior to the onset of the modern era in econometrics (about 1943 according to Klein) and from developments associated with the first attacks on autocorrelated errors in regression models (more than 30 years ago).

In 1983, the journal Technometrics contained an excellent review by Hocking of developments in the methodology of linear regression during the last quarter of a century. This survey reveals spectacular vitality in the expansion of the classical linear model. Although economists' textbooks often emphasize the simultaneous-equations model developed at the Cowles Commission from the pioneering efforts of Frisch, Haavelmo, and Mann and Wald in the decade of the 1940's, the predominant framework in much of empirical research by economists has remained the classical one of linear regression with some adjustments for autocorrelation, heteroscedasticity and structural instability. Hocking's review, therefore, provides a useful starting-point from which to begin our exploration of assimilation. The restriction to single-equation techniques will not seriously distort our findings. Indeed it could be argued that this restriction might make assimilation more rapid given relative levels of computational costs and mathematical complexity. (It should be recognized too that the two-stage least-squares estimator is the most discussed approach in the simultaneous equations model and yet this is already more than 25 years old.) Hocking's review was directed to an audience of applied statisticians but it can be read without making excessive demands of statistical or mathematical knowledge.

Our purposes are straightforward. We begin with a brief account of "best-practice" technologies for estimation in linear regression. These are already excellent descriptions provided by Belsley, Kuh and Welsch (1980), Leamer (1978), Bibby and Toutenburg (1977), and Cook and Weisberg (1982). These can be supplemented by parts of Judge et al. (1980), Judge et al. (1982), and Greenberg and Webster (1983) while the major listing of historical developments remains a series of papers by Harter (1974-1975) in the International Statistical Review. Given this collective backdrop, our account is slight. It contains the delineation of some well-defined categories, broad descriptions of particular techniques and citations for major references. The result is a skeletal outline only. Obviously we have been selective, giving more weight to advances that fit the economic context. Preference for techniques that have been shown computationally feasible and that have had their application thoroughly explored does not need much justification. We sought to make apparent certain developments and to give general guidance for understanding the maze of available literature. To provide these, we have augmented Hocking's review with additional material from sources closer to economists (including those cited above) and inadequately covered in his review.

Then we look at what economists do. An exploration on a larger scale might look at responses to a questionnaire sent to a significant proportion of the economics profession. Our alternative approach involves checking the contents of many leading journals for evidence of the use of the newer techniques cited in previous section. This approach is quicker, less costly and perhaps even more informative than that involving a questionnaire. In any case, the evidence on assimilation of best-practice methods is so clearcut that we have no grounds for believing other evidence might contradict our conclusions. Choice of journals for inclusion in our brief literature search was influenced by intangibles such as professional prestige as well as such mundane considerations as our estimates of the relative size of readership populations. From this search, we hope to explore the overall picture of assimilation and also the differences among the acceptance rates for the various categories. Some techniques may have been easier to implement or they might have enjoyed greater popularity on other grounds. The results of our search can be used to judge the quality of empirical research by economists. A slow rate of assimilation coupled with the characteristics cited earlier is indicative of an improper situation. Within this, unwarranted claims (or demands) that economics involves in a fundamental way the quantitative verification of its theoretically-based structures seems to be unmatched by statistical sophistication. Evaluations would thus be inadequate, misleading, or insincere.

In our final section, we turn from what is done to what might be done. Thus we consider certain areas where best-practice methods could be used, giving both justification and potential qualifications or constraints. The existence of techniques and a favourable attitude to their adoption in economics are insufficient conditions for actual use. The techniques must fit the economic context and their benefits, as compared to those linked with traditional approaches, must outweigh the usual increases in complexity and cost. Many adjustments to take account of statistical advances also involve implicit challenges to past habits of thought. For example, they often involve iteration, multiple uses of data, or different critical values so the Neyman-Pearson framework for statistical inference is not as simple as before. The comfortable aspects of dealing with stable structures of known form, with well- behaved errors, and reliable data disappear. They often cease to provide the given underpinnings of the newer techniques so that conclusions from research become more tentative. Robustness and sensitivity now appear as criteria for decisions at the expense of such static notions as unbiasedness or consistency.

Although we touch on these issues briefly, they are incidental to our primary objective. This remains throughout as the determination of the readiness of economists to assimilate statistical advances into their research practices. The readiness, in view of the availability of suitable software, is to be identified with actual use rather than with pious claims concerning the frequency of statistical analysis in economic research.A quality dimension (the "best" in best-practice) is essential.

CHAPTER TWO

The Statistical Background

The historical development of the least-squares principle for estimating parameters of linear equations in the familiar classical framework can be traced back beyond 1809. Those interested in assigning credit for its first introduction will be aware of one of the most famous priority disputes in statistics, with both Gauss and Legendre claiming the prestige for the inception of the least-squares principle. Papers by Plackett (1972) and Stigler (1981) have sought to resolve this issue of priority but some aspects remain unsettled. From our perspective of best-practice technology, this illustration is very similar to certain attempts in the economic literature on long waves of economic activity to locate dates of basic inventions (which might have provided the initial impetus for increased growth in critical sectors of national economies). Thus, in line with our focus on the assimilation of techniques, our concern is not with the priority issue for the least-squares principle but rather with what has happened to this basic approach since its inception. Leaving remote developments to the historical survey of Harter (1974-1975), we concentrate on the features of the last quarter century as changes in computational feasibility and research criteria made additional complexity less burdensome.

An emphasis on computational burden is easy to justify. The first one and half centuries after the inception of the least-squares method were a period of remarkable stability. Although there were important developments in the rapid evolution of theories of statistical inference, these were not matched by pronounced changes in real applications in economics. Practical difficulties affected the implementation of adjustments to least-squares regression methodology so innovations were severely restricted. From 1809 to 1959, the user of conventional regression methods had to be satisfied with analyses of simple models with few exogenous variables. The solution of normal equations with more than 3 or 4 variables remained difficult for most users. Obviously the computational hazards of this period meant that there was little incentive for users to forego simple linear equations of low dimension and to turn towards the asking of "interesting" questions. The quality of their efforts remained unchallenged from a technical viewpoint. Suitable tests of this quality were either undeveloped or not feasible. This situation was radically amended by the advent of high-speed computational facilities and by the increased accessability of appropriate software. Opportunities arose for extensions to larger and more complex models and for the search of diagnostic techniques to check the relevance of various assumptions within a framework of conditional specification. An inevitable concomitant was a reappraisal of purpose with the traditional picture of structural confirmation being augmented to include exploratory elements and sequential search.

Turning to the last quarter century, we would like to find a straightforward means of describing the changing stock of new techniques. Unfortunately the various developments in "best-practice" technologies have no clear pattern. Erratic shifts of interest in particular areas precludes, as yet, the presentation of an historical taxonomy. We are compelled to follow an alternative strategy which associates techniques with their potential incidence. To facilitate this, a particular paradigm is adopted here. This is represented by the contents of Table 1. Similar adoptions may be found in Box (1979-1980), Zellner (1975) and in Cook and Weisberg (1982). They are becoming increasingly evident in statistical textbooks as the advocacy for adoption of "best-practice" technologies becomes systematic. Our scheme was chosen to organize our brief account of them rather than for advocacy.

TABLE ONE

Schematic Outline For Model Fitting

2. ESTIMATION

(i) Least Squares

(ii) Biased Regression

(iii) Robust Estimation

1. FORMULATION

(i) Assumptions

(ii) Model Selection

(iii) Collinearity

3. INFERENCE

(i) Fitted Model Tests

(ii) Recursive Residuals

(iii) Jackknifing and Pseudo-Replication

4. CRITICISM

(i) Diagnostics

(ii) Transformation

The schematic outline of the table indicates four distinct stages that, for expositionary purposes, may be taken to be sequential in practice. The stages are termed formulation, estimation, inference and criticism. In fact, our study of best-practice does not exactly follow the flow direction indicated in the schematic outline of Table 1. Rather we only discuss available robust techniques at the end of this chapter. The reason for this minor adjustment is really quite simple. Since it is the criticism step which (through its assessment of the appropriateness of our assumptions) dictates to us the appropriateness of particular estimation procedures, it therefore makes sense to defer discussion on the availability of alternatives to least squares until the end of the chapter. In Table 2, we present the general organization of our account. The reader is advised to refer back to this table as he goes through our discussion on best-practice since it provides a succinct overview of the literature discussed and should prove useful in deciphering the material developed here. Table 2 also provides our reader with a select reading reference on each technique, thereby, providing him a quick reading list to the extensive literature.

We take as our starting point the classical linear framework for, if the classical assumptions are satisfied, least squares estimators are relatively efficient in a class containing many potential estimators. This "optimality" property of least squares within the classical framework is most often cited to justify its use in empirical work. We, therefore, take the classical framework as our "ideal" and direct our treatment of the availability of best-practice technologies towards the detection of any divergence from this ideal. This would seem to be consistent with the organization of many econometric textbooks which begin with the classical linear model and then present other ideal models, such as the Aitken framework or the simultaneous equations model, as amendments to this initial point of reference. The classical linear model, itself, is characterized by four or five assumptions, (A1)-(A5) inclusive, that are listed below. The fifth assumption of normally distributed errors is frequently not invoked and may be considered a valuable additional element when statistical tests are sought. Our list uses the familiar matrix notation.

ASSUMPTIONS OF CLASSICAL LINEAR MODEL

(A1) The observations on a measured variable y are a linear combination of the observations on a collection of carriers, with unknown fixed weights, and random errors;

$$y = XB + u$$

where y and u are vectors of length T (the number of observations), X is a matrix of size K by T, and B is a vector of length K.

(A2) The expected value of u is zero.

(A3) The errors have a constant finite variance σ^2 and are free from autocorrelation.

(A4) X is non-stochastic and of rank K.

(A5) The errors are normally distributed.

TABLE TWO

General Organization Of Best Practice

TOPIC		SUB TOPIC	SUGGESTED REFERENCE
A) EQUATION SELECTION TECHNIQUES			
		Stepwise Regression	Efroymson (1960)
		All Possible Regression	Garside (1965)
		Best Subset Regression	Hocking and Leslie (1967)
	Selection Criteria		
		R^2	Draper and Smith (1981)
		MSEP	Allen (1971a)
		RSS	Daniel and Wood (1980)
		PRESS	Allen (1971b)
		Cp	Mallows (1964)
B) MULTICOLLINEARITY			
	Detection	VIF	Marquardt (1970)
		Farrar/Glauber R.	Farrar and Glauber (1967)
		Eigen Values, Condition Index	Belsley et al. (1980)
		Decomposition Variance	Belsley et al. (1980)
	Estimation	Ridge Regression	Hoerl and Kennard (1970)
		Principal Component	Greenberg (1975)
		Bayesian	Theil (1963)
C) CRITICISM/INFERENCE			
		Diagnostics	Belsley et al. (1980)
		Transformations	Box and Cox (1964)
		Recursive Residuals	Brown et al. (1975)
		Jackknifing	Miller (1974)
D) ROBUST ESTIMATION			
		M	Huber (1964)
		R	Hodges and Lehmann (1963)
		L	Gastwirth (1966)
		Lp	Taylor (1974)
		Adaptive	Hogg (1974)

We turn in the following sections to the 4 basic topics; (A) equation selection techniques, (B) multicollinearity, (C) criticism and inference, and (D) robust estimation. These topics are split into 24 subcategories. Ultimately this survey will yield the basic list of best-practice statistical techniques that provided the framework for our review of the contents of 15 professional journals. This basic list is presented in our next chapter.

A) EQUATION SELECTION TECHNIQUES

i) Stepwise Regression

In many economic situation where there is an incomplete theoretical background from economics, the problem of selecting variables for a regression equation becomes an important one. Over the last 25 years, a large number of papers have been published on the subject of variable selection. The selection of a "best" equation based on a subset of the original set of predictor variables led to the development of "stepwise" regression. (See Draper and Smith (1981) for a definition of "best".)

Efroymson (1960) gave us a procedure for introducing the carrier variables one at a time. At each step, his procedure chooses the variable that gives us the greatest reduction in the residual sum of squares. This process is continued either until all the candidate variables are entered into the model or until it is stopped by an a priori selected tolerance level. There are two basic versions, forward selection and backward elimination.

These stepwise techniques may be useful variable-selection procedures. However, some problems are evident. First, they sometimes do not detect important variables. Second, the forward and backward procedures do not always give us the same conclusions. These and other problems have been discussed in papers by Pope and Webster (1972), Hocking (1976) and Thompson (1978a,b) and Hamaker (1962). Important papers discussing stepwise techniques are Hemmerle (1967), Jennrich (1977), Dixon (1964), Berk (1978), Beale (1970b), Allen (1971a), Pope (1969), Mantel (1970), Anscombe (1967), Beale et al. (1967), Bendel and Afifi (1977), Lund (1971) and Valiaho (1969).

ii) All Possible Regressions

With the availability of cheap high speed computational facilities, attention focused on the development of algorithms to examine and compute all possible equations from combinations of candidate variables. This method involves fitting all possible subset equations to a given body of data, thus with K candidate variable, we would fit 2^K equations. Included in this is an equation that has all the K variables and another that contains only the sample mean. Garside (1965) developed one of the early algorithms for this approach. Calculating all possible regressions gives the analyst maximum information but the number of equations to be analyzed is often large and not always feasible. For example, a set of 9 candidate variables implies looking at 512 equations. This drawback prompted the development of best subset regression discussed below. Available literature on all possible regression includes the book of Daniel and Wood (1980) and papers by Furnival (1971), Morgan and Tarter (1972) and Schatzoff et al. (1968).

iii) Best Subset Regression

Best subset regression methods try to reduce the computational burden of calculating all possible regressions. The branch- and-bound procedures developed by Hocking and Leslie (1967) and Lamote and Hocking (1970) are designed to find the "best" subset, without calculating all possible subsets. Furnival and Wilson (1974) introduced a similar method which is highly efficient and has recently gained a great deal of popularity. The papers by Beale (1970a), Lindley (1968), Anderson et al. (1970), Welsch and Peters (1978), Lamotte (1972), Beale et al. (1967), Allen (1971a, 1974), Mallows (1973), Aitkin (1974), Helms (1974), Rencher and Pun (1980), Hintze (1980), Young (1982), McKay (1979), Baskerville and Toogood (1982) and Hocking (1977) provide relevant accounts.

SELECTION CRITERIA

The three variable selection techniques introduced so far assess each regression equation according to some arbitrary criterion. In the case of "stepwise" regression, we generally use the value of R^2 from the least-square fits. In the last quarter century, several criteria have been considered, some of these have been discussed below.

i) R^2.

The use of R^2 as a selection criterion is straight forward and does not seem to need further comment. Draper and Smith (1981) provide an excellent description of its use. Lovell (1983) should be considered too.

ii) MSEP

Allen (1971a) introduces the mean square error of prediction MSEP criterion, for selecting an equation. This is given by $E(\hat{y}_p - p)$ when p carriers are involved. For a further discussion see Hocking (1972). Other closely related selection criteria have been suggested by Mallows (1966,1967), Rothman (1968), Lindley (1968) and Gorman and Toman (1966).

iii) RSS

Use of the Residual Mean Square RSS as a basis for selecting regression equations is described by Draper and Smith (1981). The relationship between RSS and the Cp (to be defined below) has been explored by Daniel and Wood (1980) and Chatterjee and Price (1977). For applications of RSS and similar criteria, see Beale et al. (1967), Hocking and Leslie (1967) and LaMotte and Hocking (1970).

iv) PRESS

Allen (1971b) suggested use of the predicted sum of squares PRESS criterion as an alternative to RSS. A complete assessment is provided by Thompson (1978a,b), Hocking (1972,1976), Younger (1979), Draper and Smith (1981), Anderson et al. (1972) Cook and Weisberg (1982).

v) Cp

The Cp statistic, introduced by Mallows (1964), has gained considerable acceptance as a good selection criterion. It is much easier to calculate than the PRESS statistic and may provide us with the opportunity for tests of significance. The statistic is defined as a simple sum of two components;

$$Cp = P + \frac{Sp - \sigma^2}{\sigma^2} (T-P),$$

where P is the number of parameters in the candidate model, Sp is the error mean square associated with the candidate model and σ^2 is the error mean square of the true best model. Now if our candidate model is the true model then $Sp - \sigma^2$ will be very close to zero and $Cp = P$. On the other hand, if our candidate model is not very good, then Sp will be far from σ^2 and Cp will not be near P. Clearly we are to choose that model that has a Cp closest to P. However, to use the Cp statistic, we must estimate σ^2 independently. The method most commonly used for this estimation is to take the Sp that is obtained from the model containing all of the candidate variables. For a more detailed explanation of the Cp statistic, see Daniel and Wood (1980). Other references to the Cp statistic include Draper and Smith (1981), Chatterjee and Price (1977), Younger (1979), Kennard (1971), Hocking (1972), Helms (1974), Searle (1971), Mallows (1973) and Thompson (1978a,b). The plot of Cp against P is discussed by Mallows (1973), Gorman and Toman (1966) and Draper and Smith (1981).

These 5 criteria assume that "collinear" relationships between the carrier variables are not severe. When ill conditioning of the data matrix is detected, the ridge regression technique of Hoerl and Kennard (1970a,b), provides us with a selection criterion whereby the ridge trace is used to eliminate variables from the equation. Rules of elimination are discussed by Chatterjee and Price (1977). Other important papers on selection criteria include Anderson et al. (1972), Darling and Tamura (1970), Forsythe et al. (1973), Gunst and Mason (1979), Haitovsky (1969 a), Kennedy and Bancroft (1971), Narula and Wellington (1979) and Nordberg (1982).

B) MULTICOLLINEARITY

DETECTION

Near-degeneracy of the signal matrix in a linear model is often described as the problem of multicollinearity. Mason et al.(1975), Wampler (1970), and Kumar (1975a) review some of the sources of multicollinearity, while Longley (1967) demonstrates some of its consequences. Well known consequences of collinear data include incorrect coefficient signs (see Mullet, 1976) and unstable parameter estimates. These and other consequences make the task of variable selection and inference very difficult . Below we list some of the new technologies that are available to detect degeneracy in the signal and some proposed remedies.

i) Variance Inflation Factor

The variance inflation factor VIFi was suggested by Marquardt (1970) and is defined as the reciprocal of (1-Ri), where Ri is the Multiple Correlation Coefficient of any given carrier Xi regressed on the remaining carriers. The problems with VIF summary statistic are two fold. First, it cannot distinguish between several coexisting near dependencies. Second, there is no set rule when considering which VIF is large. For a further discussion on the VIF measure see Snee (1983), Belsley et al. (1980) and Chatterjee and Price (1977). A closely related statistic is that developed by Farrar and Glauber, to which we now turn.

ii) Farrar and Glauber R Statistic.

Farrar and Glauber (1967) suggest that if the R^2 between one carrier and the other carrier exceeds the R^2 of the original equation then multicollinearity is a serious problem. Their diagnostic technique involves using the correlation matrix R or (X'X) to measure the degree of collinearity. They assume the data matrix X is a sample of size T from a K-variate normal distribution, with columns of X orthogonal. Then, in this context, transformation of the determinant of R is approximately distributed as chi-square giving us a test for collinearity. For a further discussion on this statistic and its drawbacks see Häitovsky (1969b), O'Hagan and McCabe (1975), Kumar (1975b) and Belsley et al. (1980).

iii) Eigenvalues, Eigenvectors and the Condition Index.

The use of eigenvalues and eigenvectors obtained from the correlation matrix of the carrier variables to detect collinearity is not new. A small eigenvalue λ_p is used to indicate a near perfect collinear relationship. See, for example, Kendall and Silvey (1969). The problem with this measure of collinearity is the fact that the econometrician is not informed as to what is "small". The tendency is, therefore, to take small as close to zero making this measure somewhat arbitrary. Chatterjee and Price (1977) suggested we compare the largest eigenvalue with the smallest but did not futher develop the idea. Belsley et al. (1980) take Chatterjee and Price's suggestion and provide a summary index which gives us a measure of the number of near dependencies there are among the columns of the data matrix X. This measure is termed the "condition index". This index is obtained by taking all the eigenvalues of the matrix X and then dividing the largest eigenvalue by all the others. Thus the Pth condition index of a data matrix X, or Kp, is the ratio of λ max to λp as p moves from 1 to K.

Belsley et al. suggest that the number of large values associated with this index indicate the number of near dependencies, with index values of say 5 or 10 being indicators of weak dependencies while values of 30 to 100 indicating moderate to strong dependencies. In this context, they can also define our overall measure statistic provided by the condition index called the "condition number" which is defined as the ratio of λ max to λ min and clearly either exceeds or is equal to unity. Their condition number is obviously the largest value of the condition index and provides a quick diagnosis on the conditioning of the data matrix. Other material on the subject is to be found in Golub and Styan (1973), Longley (1967, 1976), Householder (1964), Wilkinson (1965), Stewart (1973), Van der Sluis (1969), Kennedy and Gentle (1980) and Hocking (1983).

iv) Singular Value Decomposition and Variance Decompostion

The work of Golub and Kahan (1965), Businger and Golub (1969), Golub and Reinsch (1970), Golub (1969), Hanson and Lawson (1969) and Becker et al.(1974) gàve another approach. The data matrix X can be decomposed since there exist orthogonal matrices U and V and a diagonal matrix S having non- negative elements so that X is USV' and X'X is VS^2V' . The diagonal elements of S are called singular values or eigenvalues of X. The matrices U, V, and S contain only real numbers and the singular values of X are unique, with the number of nonzero singular values giving us the rank of X. The ratio of the largest to smallest singular value, which indicate the conditioning of X, is the condition number of X.

It can be shown that the estimated variance of each regression coefficient, can be decomposed into a sum of terms, each associated with a singular value. The "variance decomposition" enables us to determine the extent to which near-collinear relationship between the columns of X "degrade" each variance. For example, variance decomposition of the least squares estimate $\hat{B}$ is;

$$\mathrm{Var}\ (\hat{B}) = \sigma^2 (X'X)^{-1} = \sigma^2\ V\ S^{-2}\ V'$$

Belsley et al. (1980) have an excellent exposition on the use of this decomposition. Other references on the subject are Businger (1970), Golub (1968), Golub et al. (1976), Healy (1968), Longley (1976,1977), Van Loan (1976), Becker et al. (1974), Belsley (1976), Belsley and Klema (1974), Golub et al. (1980) and Kennedy and Gentle (1980). Having identified collinearity the next question is how should we proceed?

BIASED ESTIMATION

A number of alternatives to least squares have been recommended. These estimators are biased but may be preferable to least square estimators on various counts. Some corrective measures, of course, may be considered before one applies biased estimation techniques. Two of these are the elimination of variables and the introduction of new data. Both these measures have their drawbacks. Variable elimination often leads to poor estimates and new data are not always obtainable at reasonable cost.

i) Ridge Regression

The most popular of biased techniques is the ridge estimator introduced by Hoerl (1962) and Hoerl and Kennard (1970a, b). The ridge-regression estimator, with a single ridge parameter k is defined by the revised normal equations

$$(X'X + kI)\,\hat{B} = X'Y$$

or

$$\hat{B} = (X'X + kI)^{-1} X'Y$$

This estimator exhibits better mean squared error properties than the least squares estimator, with the "ridge-trace" providing graphical evidence of the effect of collinear data. The literature on ridge regression is very large. Bibliographies are provided by Hoerl and Kennard (1981,1982), Alldredge and Gilb (1976) and Draper and Smith (1981).

Papers dealing with the choice of k are Dempster et al. (1977) and Gibbons (1981), while papers critical of the ridge technique are Coniffe and Stone (1973), Smith and Campbell (1980), Smith (1980), Draper and Van Nostrand (1979). For the applications see Hoerl et al. (1975), Anderson and Scott (1974), Lawless and Wang (1976), Vinod (1976) and Mason and Brown (1975). Other illustrations include Holland (1973), Obenchain (1977,1978), Marquardt and Snee (1975), Vinod (1978), Stein (1956), Swamy et al. (1978), Newhouse and Oman (1971), and Dwivedi et al. (1980).

ii) Principal Component Estimation

This estimation technique utilizes linear combinations of the original data matrix, "principal components", to restate the linear regression model in terms of a set of orthogonal predictor variables. Details are available in the papers by Massy (1965), Jeffers (1967), Mitchell (1971), Press (1972), Marquardt (1970), Kendall (1957), Lott (1973), Hawkins (1973), Coniffe and Stone (1973), Hotelling (1933), Greenberg (1975), Hocking et al. (1976) and Marquardt and Snee (1975).

Closely related to this approach is a method proposed by Webster et al. (1974) called "Latent Root Regression". An application of this is given by Draper and Smith (1981) while technical aspects are discussed by White and Gunst (1979), Gunst et al.(1975), and Jackson and Hearne (1973).

iii) Bayesian Techniques

Any discussion of biased estimation techniques requires some mention of Bayesian methods. Notable among these are the use of prior information to specify the distribution of B's, the imposition of constraints on the B's based on prior information, and the introduction of dummy variables. These and other Bayesian techniques have been considered by Zellner (1971), Theil (1963), Lindley and Smith (1972), Box (1980), and Leamer (1973, 1978).

The Bayes-like method of "mixed estimation" developed by Theil and Golberger (1961) and Theil (1963), deserves comment since it has gained considerable popularity recently. It is easy to employ and does not require a a full specification of the prior distribution. Mixed estimation is useful when prior information is available but incomplete, Belsley et al. (1980) give an illustration using the consumption function.

C) CRITICISM AND INFERENCE

Our analysis has so far involved making some initial assumptions, selecting a model and checking for multicollinearity. Now we question the validity of our work so far, more specifically the appropriateness of our assumptions and quality of data. For this purpose, numerous multiple-regression diagnostic methods have been developed. Researchers are now able to examine the residuals that result from fitting a model. Pioneer papers in residual analysis are Anscombe (1961), Anscombe and Tukey (1963) and Tukey (1962). Anscombe (1961) presented techniques for the detection of outliers and violations of classical assumptions based upon examination of least-squares residuals. He developed statistical tests by considering the distribution of residuals and the relationship between the fitted value and the squares of residuals. The later paper by Anscombe and Tukey further develops some of these techniques. Tukey (1962) showed how to obtain information from cumulative residuals plotted on normal probability paper. These papers proved to be the catalyst for an explosion of material on the subject of residual analysis. In just 2 years after Anscombe's first paper, numerous other papers appeared. See, for example, Goldberger and Jochems (1961), Kabe (1963), Freund et al. (1961), Goldberger (1961) and Zyskind (1963). The role of residuals in revealing anomalous data later received a great deal of attention as can be seen in papers by Aigner (1974), Blomqvist (1972), Levi (L973), McCallum (1972), Wickens (1972) and Rao (1973).

i) Diagnostics

Within the linear model the vector of least-squares residuals is defined as

$$e = (e_i) = y - \hat{y}$$
$$= Py$$

where $P = (I-M)$ and $M=(Mij)$ or $X(X'X)^{-1}X'$. M is often called the "projection matrix" or "hat matrix" with well-known properties. The relationship between the residuals e and errors u is straightforward since e is the linear transformation Pu.

ia) The Projection Matrix

The projection matrix plays an important role in data analysis. The "leverage" of the ith data point is the ith diagonal element of the projection matrix M denoted by Mii. This gives us our starting point for revealing "multivariate outliers". Hoaglin and Welsch (1978) suggest we use a value of Mii in excess of 2K/T as an indication of high leverage. Belsley et al. (1980) suggest a similar diagnosis using an F test. Important properties of Mii, denoted as Hii by Belsley et al. (1980) and Vii by Cook and Weisberg (1982), have been discussed by Behuken and Draper (1972), Huber (1975), Davis and Hutton (1975), Box and Draper (1975), Velleman and Welsch (1981) and Hoaglin and Welsch (1978).

ib) Ordinary Residuals

The use of ordinary least-squares residuals to infer departures from classical assumptions is not new. They have long been used to detect autocorrelation, heteroscedasticity and non-normality of the errors. The correspondence between e and u is not invertable and often large outliers among the true errors can be reflected in residuals of modest size. This happens as the squared-error criterion weighs extreme values heavily.

ic) Studentized Residuals

The studentization of the least-squares residuals is a transformation to obtain a set of residuals with equal variances. The term "studentization" was first used by Margolin (1977). David (1981) makes a further distinction between "internal studentization" and "external studentization" which has been explored by Cook and Weisberg (1982). The studentized residual along with a scaled version called RSTUDENT introduced by Belsley et al. (1980) are defined below:

studentized residual or Ti

$$Ti = e_i / S\sqrt{1-Mii}$$

RSTUDENT or Ti*

$$Ti^* = e_i / S(i)\sqrt{1-Mii}$$

with S being the estimated standard error of the whole model, while S(i) is the estimated standard error with the ith row having been deleted. Other studies on the transformaton of residuals are Srikantan (1961), Anscombe and Tukey (1963), Ellenberg (1973, 1976) and Beckman and Trusell (1974). Welsch (1981) discusses the computional aspect of this and other diagnostics. Some non-graphical techniques using tables of critical values to test for outliers have been considered by Cook and Weisberg (1982), Stefansky (1972), Lund (1975), Prescott (1975), Miller (1966) and Weisberg (1980).

id) Diagnostics By Deletion

A large number of numerical diagnostics aimed at detecting outliers that have an unwarranted influence on the estimated coefficients have recently been developed. All of these use as their building-block Ti and Mii. The books by Cook and Weisberg (1982) and Belsley et al.(1980) have studied most of these techniques. It seems appropriate that we develop one such diagnostic technique as an illustration. A succinct summarization of these row deletion diagnostics is provided by Hocking (1983).

The difference in the B coefficients caused by the deletion of the ith row is the measure called DFBETA by Belsley et al. (1980) and is defined as;

$$\text{DFBETA} = \hat{B} - \hat{B}(i) = (X'X)^{-1} x_i e_i / 1\text{-}Mii$$

where x_i is the ith row of the X matrix. $\hat{B}(i)$ denotes an estimate of

B when the ith row has been deleted. A scaled version of DFBETA called DFBETAS is also introduced with the suggestion that observations are influential if the absolute value of DFBETAS exceeds $2/\sqrt{T}$.

Other diagnostic summaries in this category are the COVRATIO, and DIFFITS considered by Belsley et al. (1980), the Andrews and Pregibon (1978) APi statistic, the Cook (1977) Di statistic and the Wilks (1963) Λ statistic. These and other multiple row diagnostics have been discussed in Belsley et al.(1980). Other interesting papers include Draper and John (1981) on the comparison between APi and Di, Bigam (1977) and Welsch and Peters (1978) on multiple row diagnostics, Velleman and Welsch (1981) and Velleman and Hoaglin (1980) on computational aspect of these diagnostics. Other recommended papers are Dempster and Gasko (1981), Brady and Hawkins (1982), Coleman (1977) and Welsch and Kuh (1977).

ie) Graphical Diagnostics

Graphical plots have long been used to identify violations of the classical assumptions. The standard plot of y_i against e_i generally diagnoses nonlinearity, autocorrelation and heteroscedasticity. However, in the multivariate case, these plots often fail to detect violations. Recently various other plots have been considered in providing better diagnosis. The use of studentized residual time and probability plots have been favoured by Andrews and Pregibon (1978), Behnken and Draper (1972) and Belsley et al. (1980) in detecting outlying data points and violations of our assumptions. The use of "partial-regression leverage plots" to help decide if the inclusion of a new variable enters linearly into a model and provide important information on the effects of outlying data points has gained considerable popularity with statisticians. Cook and Weisberg have called these plots "added variable plots". The use of these plots have been demonstrated by Belsley et al. (1980), Draper and Smith (1981), Anscombe (1967), Mosteller and Tukey (1977), Weisberg (1980) and Velleman and Welsch (1981). Two closely related plots serving the same purpose are the "partial residual plots" as indicated by Atkinson (1981,1982), and "residual plus component plots" as noted by Wood (1973) and Larson and McCleary (1972). The use of "probability plots" to check if the distribution of the error is normal is an important diagnostic tool. The shape of the probability plot will depend on the difference between the assumed distribution (in our case normal) and the sample distribution. If the sample distribution is short tailed and we assumed a normal the probability plot will tend to be S-shaped. A long tailed sample distribution on the other hand will give us an elongated S-shaped plot. Skewed sample distributions usually lead to a J-shaped normal probability plot. These plots also can be used to detect outliers in a particular sample. However, the proper use of probability plots does require some practice. A good starting point would be the training plots provided by Daniel and Wood (1980) and Daniel (1976). The use of probability plots have been discussed by

Draper and Smith (1980), Zahn (1975a,b), Andrews and Tukey (1973), Daniel (1959), Sparks (1970), Wilk and Gnanadesikan (1968), Belsley et al. (1980), Atkinson (1982) and Mallows (1982). A summary statistic for a probability plot has been provided by Shapiro and Francia (1972) and is further discussed by Weisberg and Bingham (1975). Also see Cook and Weisberg again.

Today numerous other plots are available. The techniques associated with exploratory data analysis (EDA) are described in Tukey (1977), Mosteller and Tukey (1977) and McNeil (1977). The Analysis Center at the Wharton School has prepared an extensive collection of programs designed for interactive analysis of data using EDA techniques. The package enables the analyst to obtain Box and Whisker plots, stem-and leaf plots, comparison box plots, and diagnostic plot for nonadditivity. (See Stein (n.d.) for more information.) Also see Anscombe (1967), Behnken and Draper (1972), Andrews(1972), Mallows (1982), Pasternack and Liuzzi (1965) for further discussions on plotting techniques.

ii) Transformations

The transformation of data is sometimes required so that our model has a constant error variance, approximately normal errors and a meaningful structure. The family of "power transformations" of the response variable was first studied by Box and Cox (1964). They worked with a parametric family of transformations from y to $y^{(\lambda)}$, where the parameter λ defined a particular transformation. They considered the following two examples:

$$y^{(\lambda)} = \begin{cases} (y - 1)/\lambda & (\lambda \neq 0) \\ & \qquad \text{for } y>0 \\ \ln\ y & (\lambda = 0) \end{cases}$$

and

$$y^{(\lambda_1,\lambda_2)} = \begin{cases} ((y + \lambda_2)^{\lambda_1} - 1)/\lambda_1 & (\lambda_1 \neq 0) \\ & \qquad \text{for } y + \lambda_2 > 0 \\ \ln\ (y + \lambda_2) & (\lambda_1 = 0) \end{cases}$$

Cook and Weisberg (1982) have a good generalization on this Box-Cox technique. More methods for assessing the need to transform the responses are provided by Atkinson (1973,1982) and Andrews (1971). They cite methods that are based on predictor variables developed from the original data set or "constructed variables". Their methods provide quick and efficient diagnostics. Papers by Draper and Hunter (1969) and Hill (1966) develop graphical techniques that help assess the need for a transformation. Alternatives to the Box-Cox family are modulus and folded power transformations suggested by John and Draper (1980) and Mosteller and Tukey (1977). These prove superior under certain conditions which have been outlined by Cook and Weisberg (1982). Generalized versions of the Box-Cox transformations have been considered by Bickel and Doksum (1981), Carroll (1980), Carroll and Ruppert (1981) and Hinkley (1975).

The transformation of variables is not just limited to that of transforming the responses. Box and Tidwell (1962) proposed a procedure to select transformations of the carriers. They assume that the response y can be written in the form;

$$y_i = \sum_{j=1}^{K} B_i \; x_{ij}^{(\lambda_j)} + u_i$$

where $x_{ij}^{(\lambda_j)}$ is the transformation of the jth carrier. Their method has been further explored in papers by Dolby (1963) and Box and Draper (1982).

iii) Recursive Residuals

In economic analyses, the question about the stability of regression relationship over time is generally an important one. Two important approaches that test this stability are that of Chow (1960) and the "Cusum" and "Cusum squares" tests proposed by Brown, Durbin and Evans (1975), a preliminary account of which is given by Brown and Durbin (1969). This latter test is useful when departures of the B's from constancy are characterized by many jumps with unknown step-points. The computation of this test can be carried out using the TIMVAR statistical package, a guide to which is provided by Evans (1973). The rth recursive residual is defined as

$$Rr = \frac{y_r - x_r \hat{B}_{r-1}}{(1 + x_r'(X_{r-1}' X_{r-1}) x_r)} \quad \text{For } r=k+1,\ldots,T$$

where Xr-1 is the data matrix containing all observations until the (r-1)th time period. The Cusum test examines the plot of

$$Wr = \sum_{j=k+1}^{r} Rj / \hat{\sigma}$$

against r for r = k+1,.......,T to check for stability. Brown et al. propose the use of a pair of lines lying symmetrically above and below the line Wr = 0 since large departures of Wr from E (Wr) = 0 indicates instability.

The use of recursive residuals is not limited to stability testing. Hadayat and Robson (1970) and Harvey and Phillips (1974) use these residuals to test for heteroscadisticity and autocorrelation. Other papers on the subject are Farebrother (1976) and Schweder (1976).

iv) Jackknifing and Pseudo-Replication

The technique referred to as the jackknife can be traced back to Quenouille (1949,1956). This technique assumes that bias and sample size are reciprocally related, with the bias being represented by a power series B + (1/T) + ($1/T^2$)....., . Here T is the sample size.

The jackknifing technique gives us an estimate for the variance of B as noted by Tukey (1958), Hinkley (1977a) and Efron and Stein (1981). It also provides us with bias reduction techniques as stressed by Schucany et al. (1971) and Hinkley (1978). Miller (1974) is a good survey of jackknifing. Other recent papers on the basic theory include Finifter (1972), Efron (1979a,b), Fox et al. (1979), Duncan (1978) and Mosteller and Tukey (1968). To calculate the "exact jackknife", we begin by splitting our sample $y_1 \ldots\ldots y_T$ into g groups each of size h. Let $\hat{B}$ be the least-squares estimator based on the complete sample. Let $\hat{B}(i)$ be the corresponding estimator based on the reduced sample of size (g-1)h where the ith group of observations has been left out. The consequence of generating $\hat{B}(i)$ with this method is that the g estimates $\hat{B}(i)$ are usually correlated. These estimates can be made approximately independent by a linear transformation to obtain "Pseudo-values" by the formula

$$\tilde{B}(i) = g\,\hat{B} - ((g-1)\,\hat{B}(i)) \qquad \text{For } i = 1\ldots\ldots,g.$$

The mean of these pseudo-values is the approximately unbiased jackknife estimate with reduction in bias by (1/T) and calculated as

$$\bar{\tilde{B}} = \left(\sum_{i=1}^{g} \tilde{B}(i) \right)/g$$

The jackknife estimate of variance of $\hat{B}$ can now easily be computed as

$$S^2 = \frac{\sum\{\tilde{B}(i)^2\} - (1/g)(\sum\tilde{B}(i))^2}{g-1}$$

D) ROBUST ESTIMATION

Robust estimators were developed so as to provide desired protection against departures from the Gaussian framework by imposing relatively less weight on outlying data points which have an unwarranted effects on least-square estimates. Some of these consequences have been discussed in Finney (1974), Fox (1972), Wilks (1963) and Barnett and Lewis (1978). Developments in robust estimation can be categorized into 5 distinct groups, and labelled M, R, L, Lp and adaptive. Here we discuss their basic differences and provide the necessary references.

i) M-Estimation

M estimators are a shorthand for maximum likelihood type estimators. Given the usual linear model a robust estimate of B is obtained by minimizing

$$\sum_{i=1}^{T} \rho((y - x B)/\hat{\sigma})$$

where ρ is a suitably selected loss function and $\hat{\sigma}$ is a robust scale estimate that may be determined previously. A large number of alternative loss functions have been suggested for use by different authors. Huber (1964) suggested that we use for our robust estimator the maximum likelihood estimator of the location parameter associated with a density that is like a normal in the middle, but a double exponential in the tails. His ρ function is given by,

$$\rho(e) = e^2/2, \qquad |e| \leq k$$

$$= k\,|e| - k/2, \quad |e| > k$$

The Princeton study undertaken by Andrews et al. (1972) analysed the Huber ρ function considering various alternative k values. Further references are Andrews (1974), Huber (1972,1973), Anscombe (1967) and Bickel (1975). Clearly the choice of ρ is an important one and this choice depends on the type of distribution one assumes likely for the errors. A choice of $\rho(e) = e^2/2 + c$ (the normal loss function) will lead to least- squares estimates. Other loss functions are suggested by Hampel (1971), Andrews (1974), Hinich and Talwar (1975). Applications of some of these are given by Denby and Larsen (1977) and Mallows (1979). Hogg (1977,1979) are appropriate review papers.

ii) R-Estimation

R estimators are also a modification of the least-square procedure. In least-squares we minimize

$$\sum_{i=1}^{T} (y_i - x_i B)^2$$

Now we denote the rank of $(y_i - x_i B)$ by Ri, which is a function of B, and then minimize

$$\sum_{i=1}^{T} (y_i - x_i B) Ri$$

Different generalizations of the choice of Ri lead to different R estimates. Common choices are the Wilcoxon and Median scores. Other modifications of R estimates have been discussed by Hodges and Lehmann (1963), Wagman and Carroll (1977), Hettmansperger and McKean (1977), Jaeckel (1972), Jureckova (1971) and Policello (1976).

iii) L-Estimation

Estimates derived by taking linear combinations of order statistics are called L estimators. Examples of L estimators are sample median, trimmed mean and weighted averages. Some common estimators in this category are the "Gastwirth estimator", by Gastwirth (1966), which is a weighted average of the 33rd, 50th and 66th percentiles with respective weights of .3,.4, and .3. Other well known L estimates are provided by Tukey (1962) and Jaeckel (1971a). These and other L estimates have been discussed by Chernoff et al. (1967) and Andrews et al. (1972).

iv) Lp-Estimation

Lp estimation requires solving the problem of minimizing

$$\sum_{i=1}^{T} |y - x B|^{p} \qquad \text{for various fixed } P$$

With P equal to 1 we have the L1 estimator which has been referred to (in the literature) by variety of names. Common among these are; minimum or least sum of absolute errors (MSAE,LSAE), minimum or least absolute deviations (MAD,LAD), minimum absolute errors (MAE), least absolute value (LAE) and least absolute residual (LAR). Use of the double exponential function within the M estimation framework with $\rho(e)$ given by $|e| + c$ yields none other than our L1 estimator.

When P equals 2 we have least-squares. Comparison of estimates obtained by estimating P above and below 2 will provide knowledge of unwarranted leverage. When we increase the value of P above 2 we put greater weight on larger residuals while as we fix values of P below 2 we reduce the weight on large residuals. Thus large changes in estimated coefficients for a range of values of P indicate outliers. L1 estimates can be solved iteratively using generalized least-squares algorithms. The idea being straight-forward. Obtain an initial estimate of $\hat{B}$. One choice could be least-squares. Then consider the weight chosen as $wi = (y_i - x_i B)$ and then minimize by choice of b

$$\sum_{i=1}^{T} wi^{-1} |y_i - x_i b|^2$$

If continued iteratively, this procedure gives us the L1 estimate. Iteration is usually stopped when changes in the estimated coefficents are small. Alternatively we could choose other weights. Papers on Lp estimation are by Taylor (1974), Bassett (1978), Hill (1977),Harle and Spath (1974), Schlossmacher (1973) and Holland and Welsch (1977) who further describe this technique. The bibliographies by Gentle (1977) and Kennedy and Gentle (1980) should also be consulted.

v) Adaptive Robust Estimators

Hogg (1974) suggests we select our robust estimation procedure based on observing the sample data. In our discussion so far on robust estimation, selection of a ρ function like Huber's choice and the level of trimming were fixed by prior considerations before the sample was observed. Adaptive procedures provide several selection criterion that help us better select the appropriate robust procedure after a sample has been assembled. Jaeckel (1971b) suggests that the level of trimming (γ) be selected so as to minimize the standard error of $m(\gamma)$, the γ-trimmed mean. Hogg (1974) further suggests some test statistics that help determine the choice of trimming based on the sample kurtosis and two statistics called Q and Q1, which use order statistics of the sample to identify tail length. These adaptive robust estimators may give better results than non-adaptive approaches as indicated by the Monte-Carlo studies undertaken by Wagman and Carroll (1977). The Princeton study reviewed some of these adaptive versions but failed to consider the possibility of short tailed distributions. Hogg (1974) sheds some light on the use of adaptive versions in this context too. Papers by Takeuchi (1969,1971), Von Eeden (1970) and Shorack (1971) provide further research on adaptive robust estimators.

Concluding Remarks

Finally as suggested by developments in this chapter, econometric analysis is an iterative process in search of a useful model with the number of iterations depending on our satisfaction with the assumed model. As Snee (1983,p.232) interestingly puts it.

> "Models are always wrong because we will never know the true state of nature. The relevant question is, is the model useful? Practioners and methodology developers alike should keep this view in mind as they evaluate the results of their analysis and statistical research".

Here we have outlined some of the available best-practice technologies that could help us evaluate the correctness and consistency of our model and data. Before ending, we would like to draw attention to the ever increasing support from statisticians in favour of robust estimation technologies. Criticism employing diagnostic checks may be insufficient because some discrepancies are not easily detectable. It is on these grounds so many argue, suggesting that when developing models, one should robustify them against such contingencies. This suggestion does not in any way diminish the important role played by criticism. We must remember that criticism and robustness are not substitutes but complementary in nature. Both tools should be considered in practice.

CHAPTER THREE

Survey of Journals for use of Best-Practice (1974-1982)

The primary goal of this study is to explore the assimilation of best-practice statistical techniques in the empirical research of economists. Our approach is a simple one. We consider the contents of 15 leading economic and statistical journals in the light of the brief outline of statistical techniques that was provided in Chapter 2. The choice of journals to be appraised in this exploration was influenced by our own ranking of them by prestige. It was also affected by the rankings given by Hawkins et al. (1973), Oster (1980) and others. The 15 journals are listed in Table 1, which indicates the mnemonic coding that we have followed in the two later tables presenting the results of our survey. In the first column of Table 1, our basic ranking is given. We differentiate between economic (E) and statistical (S) journals and suggest two levels of quality. The superior economic journals are the Quarterly Journal of Economics, the Journal of Political Economy, the American Economic Review, the Review of Economic Studies and Econometrica. The second group of economic journals includes the Journal of Business, the Journal of Monetary Economics, the European Economic Review, the Journal of Money, Credit and Banking, the Journal of Econometrics, and the Review of Economics and Statistics. These 11 economic journals seem to cover most of the areas of economics. We have not tried to rank journals within the two groups but such rankings are given by

Hawkins et al. and Oster. Among 8 journals identified by Oster, the highest prestige and familiarity ratings were to be found in the orders that are recorded in Table 2. Oster only considered American publications. Her lists give strong support for our choice of superior journals. Similar support comes from the rankings of Hawkins et al. that are presented in Table 3. Our list does not include the Economic Journal, Economica and the Journal of Economic Literature which are recorded as in the top 10 according to their prestige by Hawkins et al. The list also does not include the Harvard Buisiness Review, the Oxford Economic Papers, the Journal of Finance or the Southern Economic Journal which scored highly in terms of familiarity. These omissions, taken as a whole, are likely to lead to an exaggeration of best-practice use especially when we recognize the alternative inclusion of newer journals such as the Journal of Monetary Economics, the European Economic Review and the Journal of Econometrics with which Hawkins et al. were unfamiliar a decade ago. The choice of the Journal of the American Statistical Association and the three journals of the Royal Statistical Society does not seem to call for much explanation.

In the final column of Table 1, we provide an amended grouping that might be appropriate if the incidence of best-practice statistical techniques in papers were to be given high weight in determining the quality of journals. A justification of potential changes in ranking will be given later. Some of these changes may be surprising. Among them was the need, as we saw it, to introduce a further category of E3.

Our inquiry surveyed the contents of the 15 journals throughout the period extending from 1974 to 1982, searching for evidence of usage of best-practice statistical techniques by economists. Results of this survey are described in Tables 4 and 5. From the background that was described in Chapter 2, we took a basic collection of 36 statistical techniques. These were assigned to 8 broad categories with headings of (1) robust estimators, (2) diagnostic plots, (3) diagnostic summary statistics, (4) equation selection, (5) recursive residuals, (6) transformations, (7) jackknifing, and (8) multicollinearity and imprecise estimates. The techniques are listed on the left hand side of Table 4. Within the main body of this table, we use an "X" to indicate whether a specific technique was used or suggested for use in an article. The code for the journal, its year, and the initial page of each paper is indicated at the top of the table for reference. Two illustrations can establish how the contents of the table should be read. The first journal in our list, the Quarterly Journal of Economics, contains only 4 papers that use best-practice methods during 1974-1982. One paper was published in 1974 and this involved the generalized Box-Cox technique. The same method was also used in a paper published in the journal during 1977. In two other papers during the reference period, two other best-practice methods (Box-Cox transformation, recursive residuals) were found. Turning to our eleventh journal, Econometrica, we find isolated papers in 1974, 1975, 1976, 1977 (two),1978 (two), 1979, 1980, 1981 and 1982 (two) that use such methods. One paper in the 1978 volume (beginning at page 33) uses robust estimators including M-, L- and Lp- methods.

Although the paucity of entries in Table 4 speaks for itself, we felt that a list of findings might be useful. We turn to these now.

Findings:

1) From 1974 to 1982, the general trend in assimilation of best-practice statistical techniques is upward but the relative change in the extent of use between 1974 and 1982 is not very large. This view is supported by the tendencies revealed in the entries of Table 5, which record the yearly incidence of usage for the 36 techniques. The bottom row of Table 5 indicates an increase from 20 cases of use in 1974 to 33 cases in 1982. Taking 3-year averages for 1974-1976 and 1980-1982, we detect about a 30 percent increase.

2) The entries in the bottom row of Table 5 should also be linked to the number of journals. It is clear that, even in the "best" year, the incidence of papers involving any of the 36 statistical techniques is about two papers per journal per year.

3) Table 5 also provides a summary of the use of individual techniques in its final column. It is clear that there is a considerable range of individual entries here. Recursive residuals and Box-Cox transformations are most in evidence whereas many other

techniques are almost never cited. The individual entries, however, should also be accumulated to form block totals. These are 45, 19, 10, 33, 28, 37, 5 and 43 for the 8 groups.

4) Robust estimation techniques seem to have gained some ground through the period. Our survey found 45 instances of their use with M-estimates more popular than the other ones. There is little discernible trend toward increased usage, however, if we ignore the first two years 1974-1975. Two-thirds of the cited instances are found in the four statistical journals of our list. This leaves the economic journals to share an average of just a little more than one paper each in 9 years.

5) Precursors of the Chow test for structural stability can be traced back about 35 years. Recursive residuals provide a more recent alternative. They are simple to explore and represent a straightforward extension of the Chow test. We found 28 instances of the use of recursive residuals, about 3 a year on average. This is surprising in view of the general availability of the TIMVAR software package. Indeed, although one of the more frequently used techniques in our collection, its assimilation into economics is still inadequate considering the numerous publications that persist with the Chow test.

6) The use of the jackknife and pseudo-replication seems to have eluded the economic community totally. We found only 5 instances of the use of the jackknife, of which only one was in an economic journal. This neglect fits with the predominance of significance testing and the general introduction of normality by economists even

when the range of economic variables is obviously constrained (so as to prelude normality). It also fits with the neglect of plots of residuals, partial regression leverage elements and partial residuals. These members of our second group of statistical techniques are seldom found in economic journals.

7) Multicollinearity is often cited both in econometric textbooks and in actual applied research as a major problem affecting estimation by introducing bias and reducing efficiency. Yet the use of the available best-practice techniques to detect multicollinearity and to make appropriate adjustments is rare among economists. 46 instances of use were found. It may be that there is growing acceptability of biased estimates among statisticians and econometricians but there is not basis for this view in our data for incidence. We can detect increased use of ridge regression but a total of 10 instances in the last 5 years of our sample is clearly insufficient to justify optimism with the rate of assimilation.

8) Among the techniques that we considered, the transformations seem to have been assimilated most. Early budget studies dealing with consumption expenditures discussed the choice of mathematical forms for the variables that they involved in fitted equations. The linear, logarithmic and semi-logarithmic forms are familiar to readers of this early literature. Since Box and Cox suggested their family of power-series transformations over two decades ago, we should expect them to be a common feature of economic models. Our results find evidence that use of the Box-Cox transformations is present in the

journals with an average of about 4 papers per year involving such transformations. However, the poor levels of incidence in 1979 and 1981 suggest that assimilation by economists has yet to advance still further.

9) Two of the remaining groups of techniques are diagnostic summary statistics and equation selection methods. The first of these may be too "new" with more rapid assimilation to be expected only as the efforts of Cook and Weisberg, Andrews and others become better recognized. Our survey ends in the year of publication of the textbook by Cook and Weisberg, which could serve as a major instrument in the spread of these diagnostic statistics and the acceptance of "criticism" as an integral part of the research process. The second group is, we suspect, used but not reported. It is very unlikely that many economists have not used stepwise regression as imbedded in software packages. Stepwise and stagewise methods may be hidden from our sight because the users of software are unaware of the implications of the choices that are presented to them in software manuals.

10) Our survey of available graphical diagnostics provided us with further evidence on the limited assimilation of best-practice into econometric analysis. Of the 5 techniques surveyed, normal probability plots recorded the largest incidence of use(10), while the other 4 have low rates of usage. Considering the important information provided by these plots one would expect a far greater rate of assimilation than that indicated by our survey. Again it is

likely that economists, in recent years, have used some plotting techniques but have not reported this practice in their papers.

11) Table 5 summarizes the incidence of best-practice techniques in the 11 economic and the 4 statistical journals. The evidence provided by this table points conclusively to the fact that the quality of empirical research undertaken by economists is one of senility. The assimilation of best-practice into economics is negligable and the level of statistical sophistication embarrassingly low. Although some techniques like recursive residuals and the Box-Cox power transformations have gained relatively higher rates of assimilation than others, the overall situation is one that is disturbingly slow. Finally our results clearly point out that the gap between theory and practice is alarmingly large and economists must do something about this discrepancy if they want to sustain credibility.

12) In the final column of Table 1, we provide an amended ranking based on incidence of use of best-practice statistical techniques in the 15 journals surveyed. These rankings have been calculated giving high weight to the number of times best-practice statistical techniques were used in each journal. The scheme used in this ranking is as follows; a rank of 3 was introduced for journals having less than 6 incidence of use, a rank of 2 was given to journals having 6 to 10 incidence of use, while journals having more than 10 incidences have been given a top ranking of 1. For example, the <u>Quarterly Journal of Economics</u> which is the first journal surveyed in table 4 has in 9 years published only 4 articles using those

best-practice statistical techniques considered by our survey. Based on this we have given this economics journal a rank of E3 in the final column of table 1. We strongly recommend that this survey ranking be considered when assessing the quality of the various publications surveyed.

TABLE ONE

Mnemonic Code For Journals

OUR RANKING	JOURNAL	CODE	SURVEY RANKING
E1	Quarterly Journal of Economics	J1	E3
E1	Journal of Political Economy	J2	E1
E1	American Economic Review	J3	E2
S2	Applied Statistics	J4	S1
E2	Journal of Business	J5	E3
E2	Journal of Monetary Economics	J6	E3
E2	European Economic Review	J7	E3
E2	Journal of Money, Credit and Banking	J8	E3
E1	Review of Economic Studies	J9	E3
E2	Journal of Econometrics	J10	E1
E1	Econometrica	J11	E1
E2	Review of Economics and Statistics	J12	E1
S2	Journal of the Royal Statistical Society, Ser.A	J13	S2
S1	Journal of the Royal Statistical Society, Ser.B	J14	S1
S1	Journal of the American Statistical Association.	J15	S1

TABLE TWO

Rankings of Journals by Prestige and Familiarity, Oster (1980)

Prestige:

American Economic Review (E1, J3)

Econometrica (E1, J11)

Journal of Political Economy (E1, J2)

Quarterly Journal of Economics (E1, J1)

Review of Economics and Statistics (E2, J12)

International Economic Review

Southern Economic Journal

Economic Inquiry

Familiarity:

American Economic Review (E1, J3)

Journal of Political Economy (E1, J2)

Quarterly Journal of Economy (E1, J1) /Econometrica (E1,J11)

Southern Economic Journal

Review of Economics and Statistics (E2, J12)

Economic Inquiry

International Economic Review

Note: For mnemonics,ser Table 1.

Source: Oster, S. (1980), "Optimal Order", American Economic Review, Vol. 70, No. 3, June, pp. 444-448.

TABLE THREE

Rankings of Journals by Prestige and Familiarity; Hawkins et al. (1973)

Prestige:

1. American Economic Review (E1, J3)
2. Journal of Political Economy (E1, J2)
3. Econometrica (E1, J11)
4. Quarterly Journal of Economics (E1, J1)
6. Review of Economics and Statistics (E2, J12)
9. Review of Economic Studies (E1, J9)
   Journal of the American Statistical Association (S1, J15)
21. Journal of Business (E2, J5)
22. Journal of Money, Credit and Banking (E2, J8)

Familiarity:

1. American Economic Review (E1, J3)
2. Journal of Political Economy (E1, J2)
3. Quarterly Journal of Economics (E1, J1)/Econometrica (E1,J11)
10. Review of Economics and Statistics (E2, J12)
11. Journal of the American Statistical Association(S1,J15)

Note: For mnemonics, see Table 1.

Source: Hawkins, R.G., L.S. Ritter and I. Walter (1973), "What Economists Think of Their Journals", Journal of Political Economy, Vol. 70,No. 3, pp. 1017-1032.

TABLE FOUR

Incidence of Usage, Selected Journals

1974-1982

	Journal Code	J1				J2							
	Year	74	77	77	81	74	75	75	76	76	77	79	79
	Page-Number	502	59	359	477	1205	407	571	207	299	741	109	179
1	M-Estimators												
	L-Estimators												
	R-Estimators												
	L^P-Estimators				X		X			X			
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i—Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression												
	Mallows-Cp												
5	Recursive Residuals					X						X	X
6	Box-Cox Transformations		X								X		
	Box-Tidwell Transformations										X		
	Generalized-Box-Cox	X		X									
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												
	Condition-Index							X					
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation								X				
	Ridge-Regression											X	
	Principal-Component												

	Journal Code	J2			J3						J4		
	Year	79	80	82	77	77	77	81	82	82	74	74	74
	Page-Number	522	771	1295	445	452	653	825	498	829	11	51	278
1	M-Estimators												
	L-Estimators												
	R-Estimators												
	L^P-Estimators	X							X				
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	$M_{i,i}$ Criterion												
	D_i—Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i—Andrews-Pregibon												
4	Best-Subset-Regression							X			X		
	All-Possible Subset-Regression							X					
	Stepwise-Regression											X	
	Mallows-Cp											X	
5	Recursive Residuals	X	X	X									
6	Box-Cox Transformations				X	X	X			X			X
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques						X						
	Mixed-Estimation												
	Ridge-Regression												
	Principal-Component												

	Journal Code	J4											
	Year	74	74	76	76	76	76	77	77	77	77	78	78
	Page-Number	323	405	47	154	210	266	34	173	227	246	10	52
1	M-Estimators						X						
	L-Estimators						X						
	R-Estimators						X						
	L^P-Estimators	X			X		X						
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots			X									
	Half-Normal-Plots		X										
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	H_{ii} Criterion												
	D_i—Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic											X	
	Studentized Residuals											X	
	AP_i—Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression										X		
	Stepwise-Regression		X					X		X			
	Mallows-Cp		X										
5	Recursive Residuals												
6	Box-Cox Transformations								X				X
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value					X							
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression												
	Principal-Component												

J4

	Journal Code												
	Year	78	78	78	78	80	80	80	80	80	82	82	82
	Page-Number	67	242	251	291	5	142	190	231	246	1	14	149
1	M-Estimators		X						X	X	X		X
	L-Estimators		X										
	R-Estimators		X										
	L^p-Estimators									X			
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots			X				X	X				
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	H_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
	Best-Subset-Regression				X								
	All-Possible Subset-Regression				X								
	Stepwise-Regression				X								
	Mallows-Cp												
5	Recursive Residuals	X					X						
6	Box-Cox Transformations					X		X				X	
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates										X		
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value					X							
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression					X							
	Principal-Component												

	Journal Code	J5					J6			J7			
	Year	74	75	78	78	82	78	79	81	74	77	80	81
	Page-Number	186	411	653	687	401	103	301	395	33	19	159	247
1	M-Estimators												
	L-Estimators												
	R-Estimators												
	L^P-Estimators												
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots		X	X									
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression	X	X		X	X							
	Mallows-Cp					X							
5	Recursive Residuals						X	X					
6	Box-Cox Transformations		X								X	X	
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor								X				
	R-Farrar and Glauber												
	Eigen Value								X				
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation									X			X
	Ridge-Regression								X				
	Principal-Component												

	Journal Code	J8		J9	J10								
	Year	81	82	81	74	75	76	76	77	77	78	79	80
	Page-Number	273	252	489	307	297	147	285	103	117	1	315	41
1	M-Estimators												
	L-Estimators	X						X					
	R-Estimators												
	L^p-Estimators												
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots	X											
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression												
	Mallows-Cp												
5	Recursive Residuals				X	X			X				
6	Box-Cox Transformations		X	X						X	X		
	Box-Tidwell Transformations										X		
	Generalized-Box-Cox										X		
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												X
	Condition-Index											X	
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression						X						
	Principal-Component												

	Journal Code	J10					J11						
	Year	80	80	82	82	82	74	75	76	77	77	78	78
	Page-Number	143	319	31	89	125	87	719	1259	1263	1741	33	1273
1	M-Estimators											X	
	L-Estimators											X	
	R-Estimators												
	L^P-Estimators											X	
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots								X				
	Half-Normal-Plots												
	Partical-Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic									X			
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression												
	Mallows-Cp												X
5	Recursive Residuals		X	X	X	X							
6	Box-Cox Transformations						X	X			X		
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression	X											
	Principal-Component												

	Journal Code	J11					J12						
	Year	79	80	81	82	82	74	74	75	75	76	76	76
	Page-Number	1	1333	193	1597	43	58	294	214	508	129	356	463
1	M-Estimators		X	X									
	L-Estimators			X									
	R-Estimators			X									
	L^P-Estimators			X	X	X							
	Adaptive-Estimators			X									
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression												
	Mallows-Cp												
5	Recursive Residuals	X											
6	Box-Cox Transformations												X
	Box-Tidwell Transformations									X	X	X	
	Generalized-Box-Cox												
7	Jackknife Estimates								X				
8	Variance-Inflation-Factor												
	R-Farrar and Glauber						X	X					
	Eigen Value												
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression												
	Principal-Component												

	Journal Code	J12									J13		
	Year	77	77	78	82	82	82	82	82	81	74	74	75
	Page-Number	389	465	121	104	117	307	458	584	233	131	200	385
1	M-Estimators												
	L-Estimators												
	R-Estimators												
	L^P-Estimators												
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression										X		X
	Mallows-Cp												
5	Recursive Residuals		X			X			X	X		X	
6	Box-Cox Transformations				X		X						
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												
·	Condition-Index	X											
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression			X				X					
	Principal-Component			X									

	Journal Code	J13					J14						
	Year	75	75	76	80	82	74	75	76	76	77	78	78
	Page-Number	463	504	318	383	1	284	149	132	205	371	85	186
1	M-Estimators				X								
	L-Estimators				X								
	R-Estimators				X								
	L^p-Estimators												
	Adaptive-Estimators				X								
2	Studentized-Residual-Plots			X									
	Normal-Probability-Plots											X	
	Half-Normal-Plots				X								
	Partical Regression-Leverage-Plots				X								
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic			X									
	Studentized Residuals											X	
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression					X							
	All-Possible Subset-Regression												
	Stepwise-Regression					X			X		X		
	Mallows-Cp												
5	Recursive Residuals							X		X			
6	Box-Cox Transformations				X								
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												X
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques	X					X			X			
	Mixed-Estimation												
	Ridge-Regression	X			X		X						
	Principal-Component		X										X

	Journal Code	J14						J15					
	Year	78	79	80	80	82	82	74	74	75	75	75	76
	Page-Number	313	313	71	347	1	370	154	350	113	194	769	491
1	M-Estimators		X							X			
	L-Estimators												
	R-Estimators												
	L^p-Estimators												
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots			X									
	Half-Normal-Plots					X							
	Partical Regression-Leverage-Plots					X							
	Partial-Residual-Plots					X							
3	M_{ii} Criterion												
	D_i-Cooks Distance					X							
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals						X						
	AP_i-Andrews-Pregibon					X							
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression												
	Mallows-Cp												
5	Recursive Residuals	X											X
6	Box-Cox Transformations			X		X			X				
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates				X	X							
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value												
	Condition-Index											X	
	Variance-Decompositon												
	Bayesian-Techniques							X					
	Mixed-Estimation												
	Ridge-Regression											X	
	Principal-Component										X		

J15

	Journal Code												
	Year	77	77	77	77	78	78	79	79	79	80	80	80
	Page-Number	46	54	97	608	122	488	140	169	703	16	74	525
1	M-Estimators				X			X			X		
	L-Estimators												
	R-Estimators												
	L^P-Estimators				X								
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots												
	Half-Normal-Plots												
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion								X				
	D_i-Cooks Distance								X				
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals								X				
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression					X							
	All-Possible Subset-Regression												
	Stepwise-Regression	X		X									
	Mallows-Cp	X		X									
5	Recursive Residuals		X										
6	Box-Cox Transformations						X					X	
	Box-Tidwell Transformations												
	Generalised-Box-Cox												
7	Jackknife Estimates												
8	Variance-Inflation-Factor										X		
	R-Farrar and Glauber												
	Eigen Value								X				
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression			X								X	
	Principal-Component									X		X	

	Journal Code				J15								
	Year	80	80	81	81	82	82	82	82				
	Page-Number	801	839	312	766	52	103	262	351				
1	M-Estimators		X					X					
	L-Estimators												
	R-Estimators												
	L^p-Estimators												
	Adaptive-Estimators												
2	Studentized-Residual-Plots												
	Normal-Probability-Plots					X							
	Half-Normal-Plots	X											
	Partical Regression-Leverage-Plots												
	Partial-Residual-Plots												
3	M_{ii} Criterion												
	D_i-Cooks Distance												
	DFBETA												
	DFFITS												
	Covratio												
	Wilks Λ Statistic												
	Studentized Residuals												
	AP_i-Andrews-Pregibon												
4	Best-Subset-Regression												
	All-Possible Subset-Regression												
	Stepwise-Regression								X				
	Mallows-Cp				X								
5	Recursive Residuals												
6	Box-Cox Transformations						X						
	Box-Tidwell Transformations												
	Generalized-Box-Cox												
7	Jackknife Estimates			X									
8	Variance-Inflation-Factor												
	R-Farrar and Glauber												
	Eigen Value		X		X								
	Condition-Index												
	Variance-Decompositon												
	Bayesian-Techniques												
	Mixed-Estimation												
	Ridge-Regression				X								
	Principal-Component												

TABLE FIVE

Annual Patterns of Incidence

1974-1982

YEAR	74	75	76	77	78	79	80	81	82	TOTAL
1) M-Estimators	0	1	2	1	2	2	5	2	3	18
L-Estimators	0	0	2	0	2	0	1	2	0	7
R-Estimators	0	0	1	0	1	0	1	1	0	4
Lp-Estimators	1	1	3	1	1	1	1	2	3	14
Adaptive-Estimators	0	0	0	0	0	0	1	1	0	2
2) Studentized Residual Plots	0	0	1	0	0	0	0	0	0	1
Normal Probability Plots	0	1	2	0	3	0	3	0	1	10
Half Normal Plots	1	0	0	0	0	0	2	0	1	4
Partial Regression Leverage Plots	0	0	0	0	0	0	1	0	1	2
Partial Residual Plots	0	0	0	0	0	0	0	1	1	2
3) Mii-Criterion	0	0	0	0	0	1	0	0	0	1
Di-Cooks Distance	0	0	0	0	0	1	0	0	1	2
DFBETA	0	0	0	0	0	0	0	0	0	0
DFFITS	0	0	0	0	0	0	0	0	0	0
COVRATIO	0	0	0	0	0	0	0	0	0	0
Wilks Statistic	0	0	1	1	1	0	0	0	0	3
Studentized Residuals	0	0	0	0	2	0	0	0	1	3
APi-Andrews Pregibon Statistic	0	0	0	0	0	0	0	0	1	1
4) Best Subset Regression	1	0	0	0	2	0	0	1	1	5
All Possible Subset Regression	0	0	0	1	1	0	0	1	0	3
Stepwise Regression	4	2	1	5	2	0	0	1	3	18
Mallows-Cp	2	0	0	2	1	0	0	1	1	7
5) Recursive Residuals	3	2	2	3	3	5	3	1	6	28
6) Box-Cox Transformation	3	2	1	9	3	0	6	1	4	29
Box-Tidwell Transformation	0	1	2	1	1	0	0	0	0	5
Generalized Box-Cox	1	0	0	1	1	0	0	0	0	3
7) Jackknife Estimates	0	1	0	0	0	0	1	1	2	5
8) Variance Inflation Factor	0	0	0	0	0	0	1	1	0	2
R-Farrar and Glauber	2	0	0	0	0	0	0	0	0	2
Eigen Value	0	0	1	0	1	1	3	2	0	8
Condition Index	0	2	0	1	0	1	0	0	0	4
Variance Decomposition	1	0	0	0	0	0	0	0	0	1
Bayesian Techniques	1	1	1	1	0	0	0	0	1	5
Mixed Estimation	0	0	1	0	0	0	0	1	0	2
Ridge Regression	0	2	1	1	1	2	3	2	2	14
Principal Component	0	2	0	0	1	1	1	0	0	5
TOTAL	20	18	22	28	29	15	33	22	33	220

CHAPTER FOUR

Suggestions for Assimilation of Best-Practice Methods in Economics

Certain areas in economics would obviously benefit from the use of the best-practice techniques outlined in chapter 2. We cannot point out all circumstances in economics where this use is beneficial. Hence, discussion should be limited to specific situations in practice which seem to demand using "best-practice" methods because their benefits (as compared to traditional approaches) outweigh the usual increases in complexity and cost. Our discussion encompasses four basic areas i) spread and precision, ii) robustness, iii) multicollinearity and iv) criticism. The use of equation selection and recursive residuals techniques in economics does not require further justification for assimilation. Since their purpose is straight forward, we can refrain from further discussion.

i) Spread and Precision

Most econometric work involves using finite samples to estimate economic hypothesis and infer their consequences. What economists very often fail to realize is that finite sample estimates have infinite variances and therefore inference based on estimates of spread generally provide misleading inferential conclusions. This realization also carries forward to systems. It is a well known fact that the commonly used two-stage least squares estimates have infinite variances when finite samples are involved. We therefore recommend that economists use robust estimates of spread, obtained by using the jackknife and pseudoreplication techniques outlined in chapter 2 to generate confidence intervals using the sample at hand. The use of jackknife and pseudoreplication techniques is not just limited to obtaining estimates of spread. Often non-normality and non-symmetry of errors can be checked by a plot of a distribution generated by pseudoreplication techniques using the available sample. This method also provides the economist with prior knowledge of what kind of distribution to hypothesize.

ii) Robustness

Economic data are often characterized by non-normality. In such situations least-square estimates often lose their attractiveness. We therefore recommend that robust estimators be considered in all econometric practice with both least-squares and robust estimates being computed. A comparison of the two will provide important information on the distribution of the errors. If the two estimates obtained differ to a large extent, the practioner is at once warned of large leverage. Robustification does not guarantee optimality. Instead these estimators try to ensure that they will be fairly good over a wide range of possible distributions likely to be encountered in practice. Economists cannot under any pretext rule out the existence of short and long tailed characterization of their errors.

This is clearly so when economic variables like consumption, investment and participation in the labour force are of a constrained range making it impossible for the errors to be adequately characterized by normality. This fact clearly sheds some light on the quality of estimates for consumption and investment functions that are obtained by economists using the least-squares method. Finally, if in doubt as to what distribution for the error is to be specified the prudent course would be to use robust techniques at the cost of a small premium.

iii) Multicollinearity

The problem of collinear carriers has long plagued econometric estimation. This becomes even more acute when the practioner is not aware of the problem or the degree of damage caused by its presence. We therefore, strongly recommend that economists put the collinearity diagnostics provided by Belsley et al. into common practice. The variance decomposition of the estimated parameters into a sum of terms each associated with a singular value is an excellent tool. It provides the economist with an accurate assessment as to the degree of collinearity and the potential damage to his estimates. The obvious solution for estimating parameters with collinear carriers is to use the biased estimation techniques outlined in chapter 2. Economists might also consider extending these biased techniques to systems which could lead to further theoretical developments.

iv) Criticism

Econometric research undertaken by economists is a dynamic process involving a number of iterations, a fact clearly brought out by our discussion in the early part of chapter 2. Static analysis has no part to play, since a great deal of important information is always obtained through this dynamic process. Part of this dynamic process involves the important role played by criticism which provides an effective mechanism by which to check if the initially hyphothesised assumptions are in fact true. The role of diagnostics using available best-practice plots and residual analysis to detect outliers should be routine practice in all econometric work. Economists should also be aware of the pre-test bias affecting interpretation of significance test routinely undertaken. Finally economists must incorporate both criticism and robustification in all empirical work since criticism performed may not always provide an effective check for model performance. We, therefore, recommend robustification to support analysis with potential failures of criticism.

Ultimately, economists must remember that use of available best-practice is in no way limited to our findings which consider some general areas of possible use. The potential for "best-practice" methods in economic research is vast and a rapid assimilation of these technologies will help foster more credible economic research.

BIBLIOGRAPHY

AIGNER, D.J. (1974), "MSE Dominance of Least Squares with Errors-of-Observation", Journal of Econometrics, 2, pp. 364-372.

AITKIN, M.A. (1974), "Simultaneous Inference and the Choice of Variable Subsets in Multiple Regression", Technometrics, 16, pp. 221-227.

ALLDREDGE, J.R., and GILB, N.A. (1976), "Ridge Regression: An Annotated Bibliography", International Statistical Review, 44, pp. 355-360.

ALLEN, D.M. (1971a), "Mean Square Error of Prediction as a Criterion for Selecting Variables", Technometrics, 13, pp. 469-475.

ALLEN, D.M. (1971b), "The Prediction Sum of Squares as a Criterion for Selecting Prediction Variables", Technical Report 23, Department for Statistics, University of Kentucky.

ALLEN, D.M. (1974), "The Relationship Between Variable Selection and Data Angmentation and a Method for Prediction", Technometrics, 16, pp. 125-127.

ANDERSON, D.A., and SCOTT, R.G. (1974), "The Application of Ridge Regression Analysis to a Hydrologic Target-Control Model", Water Resources Bulletin, 10, pp. 680-690.

ANDERSON, R. L., ALLEN, D.M. and CADY, F.B. (1970), "Selection of Predictor Variables in Linear Multiple Regression", Statistical Papers in Honor of George Snedecor, Ed. T.A. Bancroft (Iowa State University Press).

ANDREWS, D.F. (1971), "A Note on the Selection of Data Transformations", Biometrika, 58, pp. 249-254.

ANDREWS, D.F. (1972), "Plots of High Dimensional Data", Biometrics, 28, pp. 125-136.

ANDREWS, D.F. (1974), "A Robust Method for Multiple Linear Regression", Technometrics, 16, pp. 523-531.

ANDREWS, D.F., BICKEL, P.J., HAMPEL, F.R., HUBER, P.J., ROGERS, W.H., and TUKEY, J.W. (1972), Robust Estimates of Location (New Jersey: Princeton University Press).

ANDREWS, D.F., and PREGIBON, D. (1978), "Finding Outliers That Matter", Journal of the Royal Statistical Society, Ser. B., 40, pp. 85-93.

ANDREWS, D.F., and TUKEY, J.W. (1973), "Teletypewriter Plots for Data Analysis can be Fast: 6-Line Plots Including Probability Plots", Applied Statistics, 22, pp. 192-202.

ANSCOMBE, F.J. (1961), "Examination of Residuals", Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1, pp. 1-36.

ANSCOMBE, F.J. (1967), "Topics in the Investigation of Linear Relations Fitted by the Method of Least Squares", (with discussions) Journal of the Royal Statistical Society, Ser. B, 29, pp. 1-52.

ANSCOMBE, F.J., and TUKEY, J.W. (1963), "The Examination and Analysis of Residuals," Technometrics, 5, pp. 141-160.

ATKINSON, A.C. (1973), "Testing Transformations to Normality", Journal of the Royal Statistical Society, Ser. B, pp. 473-479.

ATKINSON, A.C. (1981), "Robustness, Transformations and Two Graphical Displays for Outlying and Influential Observations in Regressions," Biometrika, 68, pp. 13-20.

ATKINSON, A.C. (1982), "Regression Diagnostics, Transformations and Constructed Variables", Journal of the Royal Statistical Society, Ser. B, 44, pp. 1-36.

BAILEY, B. (1977), "Tables on the Bonferroni t-Statistic", Journal of the American Statistical Association, 72, pp. 469-478.

BARNETT, V., and LEWIS, T. (1978), Outliers in Statistical Data, (New York: John Wiley).

BARTLETT, M.S. (1974), "The Use of Transformations", Biometrics, 3, pp. 39-52.

BASKERVILLE, J.D., and TOOGOOD, J.H. (1982), "Guided Regression Modelling for Predicting and Exploration of Structure with Many Explanatory Variables", Technometrics, 24, pp. 9-17.

BASSETT, G.W., and KOENKER, R.W. (1978), "The Asymptotic Distribution of the Least Absolute Error Estimator," Journal of the American Statistical Association, 73, pp. 618-622.

BEALE, E.M.L. (1970a), Selecting an Optimal Subset, Integer and Nonlinear Programming (Amsterdam: North Holland Publishing Co.).

BEALE, E.M.L. (1970b), "Note on the Procedures for Variable Selection in Multiple Regression", Technometrics, 12, pp. 909-914.

BEALE, E.M.L., KENDALL, M.G. and MANN, D.W. (1967), "The Discarding of Variables in Multivariate Analysis," Biometrica, 54, pp. 357-366.

BEATON, A.E., KENDALL, M.G., and BARONE, J.L. (1976), "The Acceptability of Regression Solutions: Another Look at Computational Accuracy", Journal of the American Statistical Association, 71, pp. 158-168.

BECKER, R., AKDEN, N., and KLEMA, V. (1974), "The Singular Value Analysis in Matrix Computation," Working Paper 46, Computer Research Center, National Bureau of Economic Research, Cambridge, Mass.

BEKMAN, R., and TRUSSELL, H. (1974), "The Distribution of an Arbitrary Studentized Residual and the Effect of Updating in Multiple Regression", Journal of the American Statistical Association, 69, pp. 199-201.

BEHNKEN, D.W. and DRAPER, N.R. (1972), "Residuals and Their Variance Patterns", Technometrics, 14, pp. 101-111.

BELSLEY, D.A. (1976), "Multicollinearity: Diagnosing Its Presence and Assessing the Potential Damage it Causes Least-Squares Estimation", Working Paper No. 154, Computer Research Center National Bureau of Economic Research, Cambridge, Mass.

BELSLEY, D.A. and KLEMA, V.C. (1974), "Detecting and Assessing the Problems Caused by Multicollinearity: A use of the Singular Value Decomposition", Working Paper No. 66, Computer Research Center, National Bureau of Economic Research, Cambridge, Mass.

BELSLEY, D.A., KUH, E., and WELSCH, R.E. (1980), Regression Diagnostics: Identifying Influential Data and Sources of Collinearity (New York:John Wiley).

BENDEL, R.B., and AFIFI, A.A. (1977), "Comparison of Stopping Rules in Forward "Stepwise" Regression", Journal of the American Statistical Society, 72, pp. 46-53.

BERK, K.N. (1978), "Comparing Subset Regression Procedures", Technometrics, 20, pp. 1-6.

BIBBY, J., and TOUTENBURG, H. (1977), Prediction and Improved Estimation in Linear Models, (New York: John Wiley).

BICKEL, P.J. (1975), "One-Step Huber Estimates in the Linear Model", Journal of the American Statistical Association, 70, pp. 428-434.

BICKEL, P.J. and DOKSUM, K. (1981), "An Analysis of Transformations Revisited", Journal of the American Statistical Association, 76, pp. 296-311.

BINGHAM, C. (1977), "Some Identities Useful in the Analysis of Residuals from Linear Regressions", Technical Report No. 300, School of Statistics, University of Minnesota.

BLOMQUIST, A.G. (1972), "Approximating the Least Square Bias in Multiple Regression with Errors in Variables", Review of Economics and Statistics, 54, pp. 202-204.

BOX, G.E.P. (1979), "Strategy of Scientific Model Building" in Launer, R.L. and Wilkinson, G.N., (eds.), Robustness in Statistics (New York: Academic Press).

BOX, G.E.P. (1980), "Sampling and Bayes Inference in Scientific Modelling and Robustness", Journal of Royal Statistical Society, Ser. A, 143, pp. 383-404.

BOX, G.E.P. and COX, R.D. (1964), "An Analysis of Transformations", Journal of the Royal Statistical Society, Ser. B, 26, pp. 211-243.

BOX, G.E.P. and DRAPER, N.R. (1975), "Robust Designs", Biometrika, 62, pp. 347-352.

BOX, G.E.P. and TIDWELL, P.W. (1962), "Transformations of the Independent Variables", Technometrics, 16, pp. 385-389.

BRADLEY, R.A. and SRIVASTAVA, S.S. (1979), "Correlation in Polynomial Regression", The American Statistician, 33, pp. 11-14.

BRADU, D. and HAWKINS, D.M. (1982), "Location of Multiple Outliers in Two-Way Tables, Using Tetrads", Technometrics, 24, pp. 103-107.

BROWN, R.L. and DURBIN, J. (1968), "Methods for Investigating Whether a Regression Relationship is Constant Over Time", Selected Statistical Papers, European Meeting, Mathematical Center Tracts No. 26, Amsterdam.

BROWN, R.L., DURBIN, J. and EVANS, J.M. (1975), "Techniques for Testing the Constancy of Regression Relationships", Journal of the Royal Statistical Society, Ser. B, 37, pp. 149-163.

BUSHNELL, R.C. and HUETTNER, D.A. (1973), "Multicollinearity, Orthogonality and Ridge Regression Analysis", unpublished mimeo, presented to the Decmeber 1973 meetings, Econometric Society, New York.

BUSINGER, P.A. (1970), "Updating a Singular Value Decomposition", Nordisk Tidskrift for Information-behandling, 10, pp. 376-385.

BUSINGER, P.A. and GOLUB, G.H. (1969), "Algorithm 358: Singular Value Decomposition of a Complex Matrix", Communications of the Association for Computing Machinery, 12, pp. 564-565.

CARROLL, R. (1980), "A Robust Method for Testing Transformations to Achieve Approximate Normality", Journal of the Royal Statistical Society, Ser. B, 42, pp. 71-78.

CARROLL, R. and ROPPERT, D. (1981), "On Prediction and the Power Transformation Family", Biometrika, 68, pp. 609-616.

CHATTERJEE, S. and PRICE, B. (1977), Regression Analysis by Example (New York: John Wiley).

CHERNOFF, H., GASTWIRTH, J. and JOHNS, M.V. (1967), "Asymptotic Distributions of Linear Combinations of Functions of Order Statistics with Applications to Estimation", Annals of Mathematical Statistics, 38, pp. 52-72.

CHOW, G.C. (1957), Demand for Automobiles in the United States (Amsterdam: North Holland).

CHOW, G.C. (1960), "Test for Equality Between Sets of Coefficients in Two Linear Regressions", Econometrica, 28, pp. 591-605.

COLEMAN, D.E. (1977), "Finding Leverage Groups", Working Paper 195, Computer Research Center, National Bureau of Economic Research, Cambridge, Mass.

CONIFFE, D. and STONE, J. (1973), "A Critical View of Ridge Regression", The Statistician, 22, pp. 181-187.

COOK, R.D. (1977), "Detection of Influential Observations in Linear Regression", Technometrics, 41, pp. 311-341.

COOK, R.D. and WEISBERG, S. (1982), Residuals and Influence in Regression (New York: Chapman and Hall).

DALING, R.J. and TAMURA, H. (1970), "Use of Orthogonal Factors for Selection of Variables in a Regression Equation", Applied Statistics, 19, pp. 260-268.

DANIEL, C. (1959), "Use of Half Normal Plots in Interpreting Factorial Two Level Experiments", Technometrics, 1, pp. 311-341.

DANIEL, C. and WOOD, F.S. (1980), Fitting Equations to Data, 2nd edition (New York: John Wiley).

DAVID, H.A. (1981), Order Statistics (New York: John Wiley).

DAVIES, R.B. and HUTTON, B. (1975), "The Effects of Errors in the Independent Variables in Linear Regression", Biometrika, 62, pp. 383-391.

DEMPSTER, A.P. and GASKO-GREEN, M. (1981), "New Tools for Residual Analysis," Annals of Statistics, 9, pp. 945-959.

DEMPSTER, A.P., SCHATZOFF, M. and WERMUTH, N. (1977), "A Simulation Study of Alternatives to Least Squares", Journal of the American Statistical Association, 72, pp. 77-106.

DENBY, L. and LARSEN, W.A. (1977), "Robust Regression Estimators Compared Via Monte Carlo", Commuinications in Statistics, A6, pp. 335-362.

DENBY, L. and MALLOWS, C.L. (1977), "Two Diagnostic Displays for Robust Regression Analysis", Technometrics, 19, pp. 1-13.

DENT, W.T. and CAVANDER, D.C. (1977), "More on Computational Accuracy in Regression", Journal of the American Statistical Association, 72, pp. 598-600.

DIXON, W.J. (1964), BMD: Biomedical Computer Programs, Los Angeles: UCLA Student Store. See BMDOZR, pp. 233-257.

DOLBY, J.L. (1963), "A Quick Method for Choosing a Transformation", Technometrics, 5, pp. 317-325.

DRAPER, N.R. and COX, D.R. (1969), "On Distributions and their Transformations to Normality", Journal of the Royal Statistical Society, Ser. B, 31, pp. 472-476.

DRAPER, N.R. and HUNTER, W.G. (1969), "Transformations: Some Examples Revisted", Technometrics, 11, pp. 23-40.

DRAPER, N.R. and JOHN, J.A. (1981), "Influential Observations and Outliers in Regression", Technometrics, 23, pp. 21-26.

DRAPER, N.R. and SMITH, H. (1981), Applied Regression Analysis (New York: John Wiley).

DRAPER, N.R. and VAN NOSTAND, R.C. (1979), "Ridge Regression and James-Stein Estimation: Review and Comments", Technometrics, 21, pp. 451-466.

DUNCAN, G.T. (1978), "An Empirical Study of Jackknife-Constructed Confidence Regions in Nonlinear Regression", Technometrics, 20, pp. 123-129.

DURBIN, J. and WATSON, G.S. (1950), "Testing for Serial Correlation in Least Squares Regression", Biometrika, 37, pp. 409-428.

DWIVEDI, T.D., SRIVASTAVA, V.K. and HALL, R.L. (1980), "Finite Sample Properties of Ridge Estimators", Technbometrics, 22, pp. 205-212.

EFRON, B. (1979a), "Bootstrap Methods: Another Look at the Jackknife", Annals of Statistics, 6, pp. 1-26.

EFRON, B. (1979b), "Computers and the Theory of Statistics: Thinking the Unthinkable", Siam Rev, 21, pp. 460-480.

EFRON, B. and STEIN, C. (1981), "The Jackknife Estimate of Variance", Annals of Statistics, 9, pp. 586-596.

EFROYMSON, M.A. (1960), "Multiple Regression Analysis", In Mathematical Methods for Digital Computers, Ralston, A. and Wilf H. (Eds.), pp. 191-203 (New York: John Wiley).

ELLENBERG, J.H. (1973), "The Joint Distribution of the Standardized Least Squares Residuals from a General Linear Model", Journal of the American Statistical Association, 68, pp. 941-943.

ELLENBERG, J.H. (1976), "Testing for a Single Outlier from a General Linear Regression Model", Biometrics, 32, pp. 637-645.

EVANS, J.M. (1973), "Users' Guide to TIMVAR", Research Exercise Note 10/73 of the Central Statistical Office.

FAREBROTHER, R.W. (1976), "Recursive Residuals -- a Remark on Algorithm A75: Basic Procedures for Large, Sparse or Weighted Least Squares Problems", Applied Statistics, 25, pp. 323-24.

FARRAR, D.E. and GLAUBER, R.R. (1967), "Multicollinearity in Regression Analysis: The Problem Revisited", Review of Economics and Statistics, 49, pp. 92-107.

FINIFTER, B. M. (1972), "The Generation of Confidence: Evaluating Research Findings by Random Subsample Replication". Sociological Methodology 1972 , pp. 112-175.

FINNEY, D.J. (1974) "Problems, Data and Inference", Journal of the Royal Statistical Society, Ser. B, 36, pp. 1-24.

FORSYTHE, A.B., ENGELMAN, L., JENNRICH, R. and MAY, P.R.A. (1973), "A Stopping Rule for Variable Selection in Multiple Regression", Journal of the American Statistical Association, 68, pp. 75-77.

FORSYTHE, G.E. and STRAUS, E.G. (1955), "On Best Conditioned Matrices", Proceedings of the American Mathematical Association, 83, pp. 303-345.

FOX, A.J. (1972) "Outliers in Time Series", Journal of the Royal Statistical Society, Ser. B, 34 pp. 350-363.

FOX, T., HINKLEY, D. and LARNTZ, K. (1980), "Jackknifing in Nonlinear Regression", Technometrics, 22, pp. 29-33.

FREUND, R.J., VAIL, R.W. and CLUNIS-ROSS, C.W. (1961), "Residual Analysis", Journal of the American Statistical Association, 56, pp. 98-104.

FURNIVAL, G.M. (1971), "All Possible Regression with Less Computation", Technometrics, 13, pp. 403-408.

FURNIVAL, G.M. and WILSON, R.W. (1974), "Regression by Leaps and Bounds", Technometrics, 16, pp. 499-511.

GARSIDE, J.J. (1965), "The Best Subset in Multiple Regression Analysis", Applied Statistician, 14, pp. 196-200.

GASTWIRTH, J.L. (1965), "Percentile Modifications of Two Sample Rank Test", Journal of the American Statistical Association, 60, pp. 1127-1141.

GASTWIRTH, J.L. (1966), "On Robust Procedures", Journal of the American Statistical Association, 61, pp. 929-948.

GAUSS, C.F. (1809), "Theoria Motus Corporum Coelestium", Perthes, Hamberg. Translation reprinted as: Theory of the Motions of the Heavenly Bodies Moving About the Sun in Conic Sections (New York: Dover), 1963.

GENTLE, J.E. (1977), "Least Absolute Values Estimation: An Introduction", Communications in Statistics, (B) 6, pp. 313-328.

GIBBONS, D.G. (1981), "A Simulation Study of Some Ridge Estimators", Journal of the American Statistical Association, 76, pp. 131-139.

GOLDBERGER, A.S. (1961), "Stepwise Least Squares: Residual Analysis and Specification of Error", Journal of the American Statistical Association, 56, pp. 998-1000.

GOLDBERGER, A.S. and JOCHEMS, D.B. (1961), "Note on Stepwise Least Squares", Journal of the American Statistical Association, 56, pp. 105-110.

GOLUB, G.H. (1968), "Least Squares, Singular Values, and Matrix Approximations", Aplikace Matematiky, 13, pp. 44-51.

GOLUB, G.H. (1969),"Matrix Decompositions and Statistical Calculations",In Statistical Computation Milton,R. and Nelder J.(Eds.),(New York: Academic Press), pp. 365-397.

GOLUB, G., KLEMA, V. and PETERS, S. (1980), "Rules and Software for Detecting Rank Degeneracy", Journal of Econometrics, 12, pp. 41-48.

GOLUB, G., KLEMA, V. and STEWART, G.W. (1976), "Rank Degeneracy and Least-Squares Problems", Computer Science Technical Report Series #TR-456, University of Maryland.

GOLUB, G.H. and REINSCH, C. (1970), "Singular Value Decompositions and Least Squares Solutions", Numerische Mathematik, 15, pp. 403-420.

GOLUB, G.H. and STYAN, G.P.H. (1973), "Some Aspects of Numerical Cmputations for Linear Models", Proceedings of Computer Science and Statistics: Seventh Annual Symposium on the Interface, edited by W.J. Kennedy, pp. 189-192.

GORMAN, J.W. and TOMAN, R.J. (1966), "Selection of Variables for Fitting Equations to Data", Technometrics, 8, pp. 27-51.

GREENBERG, E., and WEBSTER, C.E. (1983), Advanced Econometrics: A Bridge to the Literature, (New York: John Wiley).

GREENBERG, E. (1975), "Minimum Variance Properties of Principal Component Regression", Journal of the American Statistical Association, 70, pp. 194-197.

GUNST, R.F., and Mason R.L. (1977), "Biased Estimation in Regression: An Evaluation using Mean Square Error", Journal of the American Statistical Association, 72, pp. 616-627.

GUNST, R.F. and MASON, R.L. (1979), "Some Considerations in the Evaluation of Alternative Prediction Equations", Technometrics, 21, pp. 55-63.

GUNST, R.F., WEBSTER, J.T. and MASON, R.L. (1975), "A Comparison of Regression Estimators-Least Squares Versus Latent Root", Presented at Eastern Regional Statistics Meetings, St. Paul, Minn.

HAITOVSKY, Y. (1969a), "A Note on the Maximization of R^2", American Statistician, 23, pp. 20-21.

HAITOVSKY, Y. (1969b), "Multicollinearity in Regression Analysis: Comment", _Review of Economics and Statistics_, 69, pp. 486-489.

HAMAKER, H.C. (1962), "On Multiple Regression Analysis", _Statistic Neerlandica_, 16, pp. 31-56 (in English).

HAMPEL, F. (1968), "Contributions to the Theory of Robust Estimation", _Ph.D. Thesis_, Berkeley.

HAMPEL, F. (1971), "A General Qualitative Definition of Robustness", _Annals of Mathematical Statistics_, 42, pp. 1887-1896.

HANSON, R.J. and LAWSON, C.L. (1969), "Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems", _Mathematics of Computation_, 23, pp. 787-812.

HARTER, H.L (1974a), "The Method of Least Squares and Some Alternatives-Part I", _International Statistical Review_, Vol. 42, No. 2, pp. 147-174.

HARTER, H.L. (1974b), "The Method of Least Squares and Some Alternatives-Part II", _International Statistical Review_, Vol.42, No. 3, pp.235-264.

HARTER, H.L. (1975a), "The Method of Least Squares and Some Alternatives-Part III", International Statistical Review, Vol.43, No.1, pp. 1-44.

HARTER, H.L. (1975b), "The Method of Least Squares and Some Alternatives-Part IV", International Statistical Review, Vol.43, No. 2, pp. 125-190.

HARTER, H.L. (1975c), "The Method of Least Squares and Some Alternatives-Part V", International Statistical Review, Vol. 43, No.2, pp. 269-278.

HARVEY, A.C. and PHILLIPS, G.D.A. (1974), "A Comparison of the Power of Some Tests for Heteroscedasticity in the General Linear Model", Journal of Econometrics, 2, pp. 307-316.

HAUSMAN, J.A. (1978), "Specification Tests in Econometrics", Econometrica, 46, pp. 1251-1271.

HAWKINS,R., RITTER,L., and WALTER,I. (1973), "What Economists Think of Their Journals", Journal of Political Economics, 81, pp. 1017-1032.

HAWKINS, D.M. (1973), "On the Investigation of Alternative Regressions by Principal Component Analysis", Applied Statistics, 22, pp. 275-286.

HEALEY, M.J.R. (1968), "Triangular Decomposition of a Symmetric Matrix", Applied Statistics, 17, pp. 195-197.

HELMS, R.W. (1974), "The Average Estimated Variance Criterion for the Selection of Variables Problem in General Linear Models", Technometrics, 16, pp. 261-273.

HEMMERLE, W.J. (1967), Statistical Computations on a Digital Computer (Waltham, Mass.: Blaisdell).

HETTMANSPERGER, T.P. and McKEAN, J.W. (1977), "A Robust Alternative Based on Ranks to Least Squares in Analyzing Linear Models", Technometrics, 19, pp. 275-284.

HEYADAT, A. and ROBSON, D.S. (1970), "Independent Stepwise Residuals for Testing Homoscedasticity", Journal of the American Statistical Association, 65, pp. 1573-1581.

HILL, W.J. (1966), "Statistical Techniques for Model Building", Ph.D. Thesis, University of Wisconsin.

HILL, R.W. and HOLLAND, P.W. (1977), "Two Robust Alternatives to Least-Squares Regression", Journal of the American Statistical Association, 72, pp. 828-833.

HILL, W.J. (1966), "Statistical Techniques for Model Building", Ph.D. Thesis, University of Wisconsin.

HILL, R.W. and HOLLAND, P.W. (1977), "Two Robust Alternatives to Least-Squares Regression", Journal of the American Statistical Association, 72, pp. 828-833.

HINICH, M.J. and TALWAR, P.P. (1975), "A Simple Method for Robust Regression", Journal of the American Statistical Association, 70, pp. 113-119.

HINKLEY, D.V. (1975), "On Power Transformations to Symmetry", Biometrika, 62, pp. 101-112.

HINKLEY, D.V. (1977a), "Jackknife Confidence Intervals Using Student-t Approximations", Biometrika, 64, pp. 21-28.

HINKLEY, D.V. (1977b), "Jackknifing in Unbalanced Situations", Technometrics, 19, pp. 285-292.

HINKLEY, D.V. (1978), "Improving the Jackknife with Special Reference to Correlation Estimation", Biometrika, 65, pp. 13-21.

HINTZE, H.L. (1980), "On the Use of Elemental Analysis in Multivariate Variable Selection", Technometrics, 22, pp. 609-612.

HOAGLIN, D.C. and WELSCH, R.E. (1978), "The Hat Matrix in Regression and ANOVA", The American Statistician, 32, pp. 17-22.

HOCKING, R.R. (1972), "Criteria for the Selection of a Subset Regression Which one Should we use", Technometrics, 14, pp. 967-970.

HOCKING, R.R. (1976), "The Analysis of Selection of Variables in Linear Regression", Technometrics, 18, pp. 425-438.

HOCKING, R.R. (1977), "Selection of the Best Subset of Regression Variables", In Statistical Methods for Digital Computers, edited by K. Enslein, et al (New York: Wiley), pp. 39-57.

HOCKING, R.R. (1983), "Developments in Linear Regression Methodology: 1959-1982", Technometrics, 25, pp. 219-230.

HOCKING, R.R. and LESLIE, R.N. (1967), "Selection of Best Subset in Regression Analysis", Technometrics, 9, pp. 531-540.

HOCKING, R.R., SPEED, F.M. and LYNN, M.J. (1976), "A Class of Biased Estimators in Linear Regression", Technometrics, 18, pp. 425-438.

HODGES, J.L., Jr. and LEHMANN, E.L. (1963), "Estimates of Location Based upon Rank Tests", Annals of Mathematical Statistics, 34, pp. 598-611.

HOERL, A.E. (1962), "Application of Ridge Analysis to Regression Problems", Chemical Engineering Progress, 58, pp. 54-59.

HOERL, A.E. and KENNARD, R.W. (1970a), "Ridge Regression: Biased Estimation for Non Orthogonal Problems", Technometrics, 12, pp. 55-68.

HOERL, A.E. and KENNARD, R.W. (1970b), "Ridge Regression: Application for Non Orthogonal Problems", Technometrics, 12, pp. 69-92.

HOERL, A.E., KENNARD, R.W. and BALDWIN, K.F. (1975), "Ridge Regression: Some Simultations", Communications in Statistics, 4, pp. 105-123.

HOERL, A.R. and KENNARD, R.W. (1976), "Ridge Regression: Iterative Estimation of the Biasing Parameter", Communications in Statistics, (A), 5, pp. 77-88.

HOERL, A.E. and KENNARD, R.W. (1981), "Ridge Regression 1980: Advanced Algorithms and Applications", American Journal of Mathematical and Management Science, 1, pp. 5-83.

HOERL, A.E. and KENNARD, R.W. (1982), "Ridge Regression Bibliography Update", Accession Report 16487, Dupont Co., Wilmington, Del.

HOGG, R.V. (1974), "Adaptive Robust Procedures: A Partial Review and Some Suggestions for Future Applications and Theory", Journal of the American Statistical Association, 69, pp. 909-923.

HOGG, R.V. (1977), "An Introduction to Robust Procedures", Communications in Statistics, A6, pp. 789-794.

HOGG, R.V. (1979), "Statistical Robustness: One View of its Use in Applications Today", The American Statistician, 33, pp. 108-115.

HOLLAND, P.W. (1973), 'Weighted Ridge Regression: Combining Ridge and Robust Regression Methods", Working Paper 11, Computer Research Center, National Bureau of Economic Research, Cambridge.

HOTELLING, H. (1933), "Analysis of a Complex of Statistical Variables into Principal Components", Journal of Educational Psychology, 24, pp. 417-441, 489-520.

HOUSEHOLDER, A.S. (1964), The Theory of Matrices in Numerical Analysis (New York: Blaisdell).

HUBER, P.J. (1964), "Robust Estimation of a location parameter", Annals of Mathematical Statistics, 35, pp. 73-101.

HUBER, P.J. (1972), "Robust Statistics: A Review", Annals of Mathematical Statistics, 43, pp. 1041-1067.

HUBER, P.J. (1973), "Robust Regression: Asymptotics, Conjectures, and Monte Carlo", Annals of Statistics, 1, pp. 799-821.

HUBER, P. (1975), "Robustness and Designs", In Srivastava, J.N. (ed.) A Survey of Statistical Design and Linear Models (Amsterdam: North Holland).

HUANG, C.L., MOON, L.C. and CHANG, H.S. (1978), "A Computer Program Using the Box-Cox Transformation Technique for the Specification of Functional Form", American Statistician, 32, p. 144.

JACKSON, J.E. and HEARNE, F.T. (1973), "Relationships Among Coefficients of Vectors Used in Principal Components", Technometrics, 15, pp. 601-610.

JAECKEL, L.A. (1971a), "Robust Estimates of Location: Symmetry and Asymmetric Contamination", Annals of Mathematical Statistics, 42, pp. 1020-1034.

JAECKEL, L.A. (1971b), "Some Flexible Estimates of Location", Annals of Mathematical Statistics, 42, pp. 1540-1552.

JAECKEL, L.A. (1972), "Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals", Annals of Mathematical Statistics, 43, pp. 1449-1458.

JEFFERS, J.N.R. (1967), "Two Case Studies in the Application of Principal Component Analysis", Applied Statistics, 3, pp. 225-236.

JENNRICH, R.I. (1977), Stepwise Regressions, in Statistical Methods for Digital Computers, edited by K. Enslein, A. Ralston and H.S. Wilf (New York: John Wiley), pp. 58-75.

JOHN, J.A. and DRAPER, N.R. (1980), "An Alternative Family of Transformations", Applied Statistics, 29, pp. 190-197.

JOLLIFFE, I.T. (1972-73), "Discarding Variables in a Principal Component Analysis, 1: Artificial Data", Applied Statistics, 21, pp. 160-173.

JUDGE, G.G., GRIFFITHS, W.E., LEE, T.C. (1980), The Theory and Practice of Econometrics, (New York: John Wiley).

JUDGE, G.G., HILL, R.C., GIFFITHS, W.E., LUTKEPOHL, H. and LEE, T.C. (1982), Introduction to the Theory and Practice of Econometrics, (New York: John Wiley).

JURECKOVA, J. (1971), "Nonparametric Estimates of the Regression Coefficients", Annals of Mathematical Statistics, 42, pp. 1328-1338.

KABE, D.G. (1963), "Stepwise Multivariate Linear Regression", Journal of American Statistical Association, 58, pp. 770-773.

KENDALL, M.G. (1957), A Course in Multivariate Analysis (London: Charles Griffin).

KENNARD, R.W. (1971), "A Note on the Cp Statistics", Technometrics, 13, pp. 899-900.

KENNEDY, W.J. and BANCROFT, T.A. (1971), "Model Building for Prediction in Regression Based upon Repeated Significance Tests", Annals of Mathematical Statistics, 42, pp. 1273-1284.

KENNEDY, W.J., Jr. and GENTLE, J.E. (1980), Statistical Computing (New York and Basel: Marcel Dekker, Inc.).

KLOEK, T. and MENNES, L.B.M. (1960), "Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables", Econometrica, 28, pp. 45-61.

KRASKER, W.S. (1980), "Estimation in Linear Regression Models with Disparate Points", Econometrica, 48, pp. 1333-1346.

KRUSKAL, J.B. (1978), "Statistical Analysis, Special Problems of...", In International Encyclopedia of Statistics, Vol. 2 (New York: The Free Press), pp. 1044-1056.

KUMAR, T.K. (1975a), "The Problem of Multicollinearity: A Survey:, Unpublished Mimeo (Cambridge, Mass.: Abt Associates, Inc.).

KUMAR, T.K. (1975b), "Multicollinearity in Regression Analysis", The Review of Economics and Statistics, 57, pp. 365-366.

LAMOTTE, L.R. (1972), "The SELECT Routines: a Program for Identifying Best Subset Regression", Applied Statistics, 21, pp. 92-93.

LAMOTTE, L.R. and HOCKING, R.R. (1970), "Computational Efficiency in the Selection of Regression Variables", Technometrics, 12, pp. 83-93.

LARSEN, W.A. and McCLEARY, S.A. (1972), "The use of Partial Residual Plots in Regression Analysis", Technometrics, 14, pp. 781-790.

LAWLESS, J.F. and WANG, P. (1976), "A Simulation Study of Ridge and other Regression Estimators", Communications in Statistical Theory Methods, A5, pp. 307-323.

LEAMER, E.E. (1973), "Multicollinearity: A Bayesian Interpretation", Review of Economics and Statistics, 55, pp. 371-380.

LEAMER, E.E. (1978), Specification Searches: Ad Hoc Inference With Nonexperimental Data, (New York: John Wiley).

LEVI, M.D. (1973), "Errors in Variables Bias in the Presence of Correctly Measured Variables", Econometrica, 41, pp. 985-986.

LINDLEY, D.V. (1968), "The Choice of Variables in Multiple Regression", Journal of the Royal Statistical Society, Soc. (B) 30, pp. 31-53.

LINDLEY, D.V. and SMITH, A.F.M. (1972), "Bayes Estimates for the Linear Model", Journal of the Royal Statistical Society, Ser. B, 34, pp. 1-34.

LONGLEY, J.W. (1967), "An Appraisal of Least Squares Problems for the Electronic Computer from the Point of View of the User", Journal of the American Statistical Association, 62, pp. 819-841.

LONGLEY, J.W. (1976), "A Subroutine for Singular Value Decomposition by normalized Modified Gram-Schmidt Algorithm for the Statistical Computing Section, American Statistical Association, pp. 207-212.

LONGLEY, J.W. (1977), "Least Squares Computations and the Condition of the Matrix", Paper Presented at the International symposium on Extremal Methods and Systems Analysis, Austin, Univ. of Texas.

LOTT, W.F. (1973), "The Optimal Set of Principal Component Restrictions on a Least Squares Regression", Communications in Statistics, 2, pp. 449-464.

LOVELL, M.C (1983), "Data Mining", The Review of Economics and Statistics, 65, pp. 1-12.

LUND, I.A. (1971), "An Application of Stagewise and Stepwise Regression Procedures to a Problem of Estimating Precipitation in California", Journal of Applied Meteorol, 10, pp. 892-902.

LUND, R.E. (1975), "Tables for an Approximate Test for Outliers in Linear Models", Technometrics, 17, pp. 473-476.

MALLOWS, C.L. (1964), "Choosing Variables in Linear Regression: A Graphical Aid", Presented at the Central Regional Meeting of the Institute of Mathematical Statistics, Manhattan, KS.

MALLOWS, C.L. (1966), "Choosing a Subset Regression", Presented at the Joint Statistical Meetings, Los Angeles, California.

MALLOWS, C.L. (1967), "Choosing a Subset Regression", Unpublished Report. Bell Telephone Laboratories.

MALLOWS, C.L. (1973), "Some Comments on Cp", Technometrics, 15, pp. 661-675.

MALLOWS, C.L. (1979), "Robust Methods - Some Examples of Their Use", The American Statistician, 33, pp. 179-184.

MALLOWS, C.L. (1982), "Discussion on Atkinsons", Journal of the Royal Statistical Society, Ser. B, 44, p. 29.

MANTEL, N. (1970), "Why Stepdown Procedures in Variable Selection", Technometrics, 12, pp. 463-481.

MARGOLIN, B.H. (1977), "The Distribution of Internally Studentised Statistics via Laplace Transform Inversion", Biometrika, 64, pp. 573-582.

MARQUARDT, D.W. (1970), "Generalized Inverse, Ridge Regression, Biased Linear Estimation", Technometrics, 12, pp. 591-612.

MARQUARDT, D.W. and SNEE, R.D. (1975), "Ridge Regression in Practice", The American Statistician, 29, pp. 3-19.

MASON, R. and BROWN, W.G. (1975), "Multicollinearity Problems and Ridge Regression in Sociological Models", Social Science Research, 4, pp. 135-149.

MASON, R.L., WEBSTER, J.T. and GUNST, R.F. (1975), "Source of Multicollinearity in Regression Analysis", Communication in Statistics, 4, pp. 277-292.

MASSY, W.F. (1965), "Principal Components Regression in Exploratory Statistical Reearch", Journal of the American Statistical Association, 60, pp. 234-246.

MAYOR, L.S. and WILLKE, T.A. (1973), "On Biased Estimation in Linear Models", Technometrics, 15, pp. 497-508.

McCALLUM, B.T. (1972), "Relative Asymptotic Bias from Errors of Omission and Measurement", Econometrica, 40, pp. 757-758.

McDONALD, G.. and GALARNEAU, D.I. (1975), "A Monte Carlo Evaluation of Some Ridge-Type Estimators", Journal of the American Statistical Association, 70, pp. 407-416.

McKAY, R.J. (1979), "The Adequacy of Variable Subsets in Multivariate Regression", Technometrics, 21, pp. 475-480.

McNEIL, D.R. (1977), Interactive Data Analysis (New York: Wiley).

MERLE, G. and SPATH, H. (1974), "Computational Experiences with Discrete Lp Approximation", _Computing_, 12, pp. 315-321.

MILLER, R. (1966) _Simultaneous Inference_ (New York: McGraw Hill).

MILLER, R.G. (1974), "The Jackknife: a Review", _Biometrika_, 61, pp. 1-15.

MITCHELL, B. (1971), "Estimation of Large Econometric Models by Principal Components and Instrumental Variable Methods", _Review of Economics and Statistics_, 53, pp. 140-146.

MORGAN, J.A. and TATAR, J.F. (1972), "Calculating Residual Sum of Squares for all Possible Regressions", _Technometrics_, 14, pp. 317-325.

MOSES, L. (1979), "Charts for Finding Upper Percentage Points of Students' t in the Range 0.01 to 0.00001", _Technometrics_, 14, pp. 317-325.

MOSTELLER , F., and TUKEY, J.W. (1968), "Data Analysis, Including Statistics". In G. Lindzey and E. Aronson (Eds.), _Handbook of Social Psychology_. Vol. 2. (Reading Mass: Addison-Wesley).

MOSTELLER, F. and TUKEY, J.W. (1977), Data Analysis and Regression (Reading, Mass.: Addison-Wesley).

MULLET, G.M. (1976), "What Regression Coefficients have the Wrong Sign", Journal of Quality Technology, 8, pp. 121-126.

NARULA, S.C. and WELLINGTON, J.F. (1977), "Selection of Variables in Linear Regression Using the Minimum Sum of Weighted Absolute Errors Criterion", Technometrics, 21, 299-306; Correction, 22, pp. 452.

NEWHOUSE, J.P. and OMAN, S.D. (1971), "An Evaluation of Ridge Estimators", Rand Report, No. R-716-RP, pp. 1-28.

NORDBERG, L. (1982), "On Variable Selection in Generalized Linear and Related Regression Models", Communications in Statistics, 11, pp. 2427-2450.

OBENCHAIN, R.L. (1977), "Classical F Test and Confidence Regions for Ridge Regression", Technometrics, 19, pp. 429-439.

OBENCHAIN, R.L. (1978), "Good and Optimal Ridge Estimators", Annals of Statistics, 6, pp. 1111-1121.

O'HAGAN, J. and McCABE, B. (1975), Tests for the Severity of Multicollinearity in Regression Analysis: A Comment", The Review of Economics and Statistics, 57, pp. 369-370.

OSTER, S. (1980), "The Optimal Order for Submitting Manuscripts", The American Economic Review, 70, pp. 444-448.

PASTERNACK, B. and LIUZZI, A. (1965), "Patterns in Residuals: A Test for Regression Model Adequacy in Radionuclide Assay", Technometrics, 7, pp. 603-621.

PLACKETT, R.L. (1972), "The Discovery of the Method of Least Squares". Biometrika, 59, pp. 239-251.

POLICELLO, G.E. 11., and HETTMANSPERGER, T.P. (1976), "Adaptive Robust Procedures for the one Sample Location Problem", Journal of the American Statistical Association, 71, pp. 624-633.

POPE, P.T. and WEBSTER, J.T. (1972), "The Use of an F Statistic in Stepwise Regression Procedures", Technometrics, 14, pp. 327-349.

PRESCOTT, P. (1975), "An Approximate Test for Outliers in Linear Models", Technometrics, 17, pp. 129-132.

PRESS, S.J. (1972), Applied Multivariate Analysis (New York: Holt, Rinehart and Winston).

QUENOUILLE, M.H. (1949), "Approximate Tests of Correlation in Time Series", Journal of the Royal Statistical Society, Ser. B, 11, p. 68-84.

QUENOUILLE, M.H. (1956), "Notes on Bias in Estimation", Biometrika, 43, pp. 353-360.

RAO, C.R. (1973) Linear Statistical Inference and its Applications 2nd ed., (New York: John Wiley).

RAO, P. (1973), "Some Notes on the Error-in-Variables Model", The American Statistician, 27, pp, 217-218.

RENCHER, A.C. and PUN, F.C. (1980), "Inflation of R-Squared in Best Subset Regression", Technometrics, 22, pp. 49-54.

ROTHMAN, D. (1968), "Letter to the Editor", Technometrics, 10, p. 432.

SCHATZOFF, M., FIENBERG, S. and TSAO, R. (1968), "Efficient Calculation of all Possible Regressions", Technometrics, 10, pp. 769-779.

SCHEFFE, H. (1959), The Analysis of Variance (New York: John Wiley).

SCHLOSSMACHER, E.J. (1973), "An Iterative Technique for Absolute Deviations Curve Fitting", Journal of the American Statistical Association, 68, pp. 857-859.

SCHUCANY, W.R., GRAY, H.L. and OWEN, D.B. (1971), "On Bias Reduction in Estimation", Journal of the American Statistical Association, 66, pp. 524-533.

SCHWEDER, T. (1976), "Optimal Methods to Detect Structural Shift or Outliers in Regression", Journal of the American Statistical Association, 71, pp. 491-501.

SCLOVE, S.L. (1968), "Improved Estimators for Coefficients in Linear Regression", Journal of the American Statistical Association, 63, pp. 597-606.

SEARLE, S.R. (1971), Linear Models (New York: John Wiley).

SHAPIRO, S.S. and FRANCIA, R.S. (1972), "An Approximate Analysis of Variance Test for Normality", Journal of the American Statistical Association, 67, pp. 215-216.

SHORACK, G.R. (1971), "Random Means", Technical Report No. 26, Dept. of Mathematics, Univ. of Washington.

SMITH, G. (1980), "An Example of Ridge Regression Difficulties", Canadian Journal of Statistics, 8, pp. 217-225.

SMITH, G. and CAMPBELL, F. (1980), "A Critique of some Ridge Regression Methods", Journal of the American Statistical Association, 75, pp. 74-81.

SMITH, V.K. (1976), "The Estimated Power of Several Tests for Autocorrelation with Non-first-order Alternatives", Journal of the American Statistical Association, 71, pp. 879-883.

SNEE, R.D. (1983), "Discussion on Hocking's Paper," Technometrics, 25, pp. 230-237.

SPARKS, D.N. (1970), "Half Normal Plotting", Applied Statistics, 19, pp. 192-199.

SRIKANTAN, K.S. (1961), "Testing for a Single Outlier in a Regression Model", Sankhya, A, 23, pp. 251-260.

STEWART, G.W. (1973), Introduction to Matrix Computations (New York: Academic Press).

STIGLER, S.M. (1981), "Gauss and the Invention of Least Squares", The Annals of Statistics, 3, pp. 465-474.

STEIN, C.M. (1960), "Multiple Regression", Contributions to Probability and Statistics, Essays in Honor of Harold Hotelling ed., I. Olkin, Stanford University Press, pp. 424-443.

STINE, R.A., Director of Computing, The Analysis Center, 3609 Locust Walk C9, University of Pennsylvania, Philadelphia, PA, 11104.

SWAMY, P.A.V.B., MEHTA, J.A. and RAPPORT, P.N. (1978), "Two Methods of Evaluating Hoerl and Kennard's Ridge Regression", Communications in Statistics-Theory and Methods, A7(12), pp. 1133-1155.

TAKEUCHI, K. (1969), "A Uniformly Asymptotically Efficient Robust Estimator of Location", NYU, Courant Institute Technical Report, IMM 375.

TAKEUCHI, K. (1971), "A Uniformly Asymptotically Efficient Estimator of a Location Parameter", Journal of the American Statistical Association, 66, pp. 292-301.

TAYLOR, L.D. (1974), "Estimating by Minimizing the Sum of Absolute Errors", in _Frontiers in Econometrics_, edited by P.Zarembka (New York: Academic Press), pp. 169-190.

THOMPSON, L. (1978a), "Selection of Variables in Multiple Regression", Part 1, _International Statistical Review_, 46, pp. 1-20.

THOMPSON, M.L. (1978b), "Selection of Variables in Multiple Regression", Part ||, _International Statistical Review_, 46, pp. 129-146.

THEIL, H. (1963), "On the Use of Incomplete Prior Information in Regression Analysis", _Journal of the American Statistical Association_, 58, pp. 401-414.

THEIL, H. (1971), _Principles of Econometrics_ (New York: John Wiley).

THEIL, H. and GOLDBERGER, A.S. (1961), "On Pure and Mixed Statistical Estimation in Economics", _International Economic Review_, 2, pp. 65-78.

TIETJEN, G.L., MOORE, R.H. and BECKMAN, R.J. (1973), "Testing for a Single Outlier in Simple Linear Regression", _Technometrics_, 15, pp. 717-721.

TUKEY, J.W. (1949), "One Degree of Freedom for Nonadditivity", Biometrics, 5, pp. 232-242.

TUKEY, J.W. (1957), "On the Comparative Anatomy of Transformations", Annals of Mathematical Statistics, 28, pp. 602-632.

TUKEY, J.W. (1958), "Bias and Confidence in Not-Quite Large Samples (Abstract)", Annals of Mathematical Statistics, 29, p. 614.

TUKEY, J.W. (1962), "The Future of Data Analysis", Annals of Mathematical Statistics, 33, pp. 1-67.

TUKEY, J.W. (1977), Exploratory Data Analysis (Reading, Mass.: Addison-Wesley).

VALIAHO, H. (1969), "A Synthetic Approach to Stepwise Regression Analysis", Communications in Physics and Mathematics, 34, pp. 31-131.

VAN DER SLUIS, A. (1969), "Condition Numbers and Equilibration of Matrices", Numerische Mathematik, 14, pp. 14-23.

VAN EEDEN, C. (1970), "Efficiency-Robust Estimation of Location", Annals of Mathematical Statistics, 41, pp. 172-181.

VAN LOAN, C.F. (1976), "Generalizing the Singular Value Decomposition", SIAM Journal of Numerical Analysis, 13, pp. 76-83.

VELLEMAN, P.F. and HOAGLIN, D.C. (1980), Applications, Basic and Computing for Exploratory Data Analysis (Mass.: Duxbury Press).

VELLEMAN, P. and WELSCH, R. (1981), "Efficient Computing of Regression Diagnostics", American Statistician, 35, pp. 234-242.

VINOD, H.D. (1976), "Application of New Ridge Regression Methods to a Study of Bell System Scale Economics", Journal of the American Statistical Association, 71, pp. 835-841.

VINOD, H.D. (1978a), "Survey of Ridge Regression and Related Techniques for Improvements Over Ordinary Least Squares", Review of Economics and Statistics, 60, pp. 121-131.

VINOD, H.D. (1978b), "Equivariance of Ridge Estimators Through Standardization, a Note", Communications in Statistics: Theory and Methods, A7 (12) pp. 1157-1161.

VON HOHENBALKEN, B. and RIDDELL, W.C. (1978), "Wedge Regression: A Ridge Equivalent Procedure That Forecasts and Generalized Differently", Unpublished Mimeo, Department of Economics, Univ. of Alberta, Canada.

WAMPLER, R. (1970), "A Report on the Accuracy of Some Widely Used Least Squares Computer Programs", Journal of the American Statistical Association, 65, pp. 549-565.

WEBSTER, J.T., GUNST, R.F. and MASON, R.L. (1974), "Latent Root Regression Analysis", Technometrics, 16, pp. 513-523.

WESMAN, E.J. and CARROLL, R.J. (1977), "A Monte Carlo Study of Robust Estimators of Location", Communications in Statistics, A6, pp. 795-812.

WEISBERG, S. (1980), Applied Linear Regression (New York: John Wiley).

WEISBERG, S. and BINGHAM, C. (1975), "An Analysis of Variance Test for Normality Suitable for Machine Calculation", Technometrics, p. 133.

WELSCH, R.E. and KUH, E. (1977), "Linear Regression Diagnostics", Working Paper No. 173, National Bureau of Economic Research.

WELSCH, R.E. and PETERS, S.C. (1978), "Finding Influential Subsets of Data in Regression Models", Proceedings of the Computer Science and Statistics: Eleventh Annual Symposium on the Interface, edited by A.R. Gallant, et al.,Institute of Statistics, North Carolina State University, Raleigh, pp. 240-244.

WHITE, J.W. and GUNST, R.F. (1979), "Latent Root Regression: Large Sample Analysis", Technometrics, 21, pp. 481-488.

WILKS, S.S. (1963), "Multivariate Statistical Outliers", Sankhya, A, 25, pp. 407-426.

WILK, M.D. and GNANADESIKAN, R. (1968), "Probability Plotting Methods for the Analysis of Data", Biometrika, 55, pp. 1-17.

WILKINSON, J.H. (1965), The Algebraic Eigenvalue Problem (Oxford: Clarendon Press).

WILLAN, A.R. and WATTS, D.G. (1978), "Meaningful Multicollinearity Measures", Technometrics, 20, pp. 407-412.

WICKENS, M.R. (1972), "A Note on the Use of Proxy Variables", Econometrica, 40, pp. 759-761.

WOOD, F.S. (1973), "The Use of Individual Effects and Residuals in Fitting Equations to Data", Technometrics, 15, pp. 677-695.

YOUNG, A.S. (1982), "The Bivar Criterion for Selecting Regressions", Technometrics, 24, 181-189.

YOUNGER, M.S. (1979), A Handbook for Linear Regression (Mass.: Duxbury Press).

ZAHN, D. (1975a), "Modifications of and Revised Critical Values for the Half Normal Plot", Technometrics, 17, pp. 189-200.

ZAHN, D. (1975b), "An Empirical Study of the Half-Normal Plot", Technometrics, 17, pp. 201-211.

ZAREMBKA, P. (1974), "Transformation of Variables in Econometrics", in Frontiers in Econometrics, Academic Press, pp. 81-103.

ZELLNER, A. (1971), An Introduction to Bayesian Inference in Econometrics (New York: John Wiley).

ZELLNER, A (1975), "Time Series Analysis and EconometricModel Construction", Applied Statistics (Ed.) R.P. Gupta, pp. 374-398.

ZYSKIND, G. (1963), "A Note of Residual Analysis", Journal of the American Statistical Association, 58, pp. 1125-1132.