Volume Title: Annals of Economic and Social Measurement, Volume 3, number 4

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm74-4

Publication Date: October 1974

Chapter Title: The Relative Efficiency of Instrumental Variables Estimators of Systems of Simultaneous Equations

Chapter Author: James M. Brundy, Dale W. Jorgenson

Chapter URL: http://www.nber.org/chapters/c10208

Chapter pages in book: (p. 679 - 700)

Annals of Econometric and Social Measurement, 3/4, 1974

THE RELATIVE EFFICIENCY OF INSTRUMENTAL VARIABLES ESTIMATORS OF SYSTEMS OF SIMULTANEOUS EQUATIONS

BY JAMES M. BRUNDY AND DALE W. JORGENSON

Consistent and efficient estimators of simultaneous equations by the method of instrumental variables require an initial consistent estimator of the structural form. Instrumental variables estimators that are consistent but not necessarily efficient can be employed for this purpose. The first objective of this paper is to measure the relative efficiency of alternative instrumental variables estimators proposed in the literature. The second objective is to assess the sensitivity of limited information efficient (LIVE) and full information instrumental variables efficient (FIVE) estimators to the choice of an initial consistent estimator.

1. INTRODUCTION

In previous papers (1971, 1973) we have provided a complete characterization of the class of consistent and efficient estimators of simultaneous equations by the method of instrumental variables. Our characterization of consistent and efficient instrumental variables estimators suggests two alternative approaches to the estimation of simultaneous equations:[1]

1. First estimate the *reduced* form by any consistent estimator. This approach underlies the methods of two- and three-stage least squares.
2. First estimate the *structural* form of the model by any consistent estimator; then derive a consistent estimator of the reduced form from the structural form estimator. This approach underlies the methods of limited information efficient (LIVE) and full information instrumental variables efficient (FIVE) proposed in our earlier paper.

The approach to simultaneous equations estimation based on consistent estimation of the structural form is easier to apply in practice. To obtain an initial consistent estimator of the structural form the method of instrumental variables provides a promising approach.[2] A number of alternative instrumental variables estimators have been proposed in the literature. Although these estimators differ in efficiency and in computational difficulty, all of them are consistent. The first objective of this paper is to compare these estimators with regard to computational difficulty and to evaluate their relative efficiency.

In the following sections we first outline the simultaneous equations model of econometrics. We then describe the statistical properties and computational requirements of alternative instrumental variables estimators. To evaluate the relative efficiency of the alternative estimators we compute estimates for Klein Model I.

The second objective of this paper is to assess the sensitivity of the LIVE and FIVE estimators to the choice of an initial consistent estimator of the structural

[1] A historical survey of these alternative approaches is given in our earlier paper (1973), pp. 215–219. The LIVE and FIVE estimators were proposed, independently, by Dhrymes (1971).

[2] The method of instrumental variables was originated by Reiersøl (1945) and Geary (1949). A definitive treatment of the method of instrumental variables for a single structural equation is given by Sargan (1958).

form. We also consider the effect of iteration of the LIVE and FIVE estimators. On the basis of our results we recommend the following approach to the estimation of simultaneous equations:

1. Estimate the structural form by ordinary least squares. This method is generally inconsistent.
2. Compute fitted values from the ordinary least squares estimators and use these as instruments for the corresponding jointly dependent variables. This method is consistent but generally inefficient.
3. Compute fitted values for the second round estimator and proceed to compute the LIVE or FIVE estimators proposed in our earlier paper.

The process described above can be truncated at the LIVE or FIVE estimators. Alternatively, the third step can be reiterated until the process converges. This iterative scheme coincides with Durbin's method for full information maximum likelihood estimation of simultaneous equations in the case of the FIVE estimator.[3] The scheme coincides with Lyttkens' iterative instrumental variables method in the case of the LIVE estimator.[4]

2. The Simultaneous Equations Model

We consider a simultaneous equations model with p equations; the structural form of the model is denoted:

$$Y\Gamma + X\mathrm{B} = \mathrm{E}, \tag{1}$$

with Y the matrix of observations on the p jointly dependent variables, X the matrix of n observations on the q predetermined variables, and E the matrix of random errors; the matrices $\{\Gamma, \mathrm{B}\}$ of structural coefficients are unknown parameters to be estimated. The reduced form of the model may be written:

$$Y = X\Pi + \Upsilon,$$

where the matrix $\Pi = -\mathrm{B}\Gamma^{-1}$ of reduced form coefficients is unknown and $\Upsilon = \mathrm{E}\Gamma^{-1}$ is a matrix of random errors.

Following the notation of Zellner and Theil (1962), we may denote the individual structural equations by:

$$y_j = Z_j\delta_j + \varepsilon_j, \quad (j = 1, 2, \ldots, p), \tag{2}$$

where

$$Z_j = [Y_j \;\; X_j], \qquad \delta_j = \begin{bmatrix} \gamma_j \\ \beta_j \end{bmatrix};$$

in this notation y_j is a vector of observations on the j-th column of Y; the structural coefficient of this variable is normalized at unity; Y_j is a matrix of observations on the other jointly dependent variables included in the equation, X_j is a matrix of

[3] See Durbin (1963) and Malinvaud (1970), pp. 686–687. The iterative process associated with Durbin's method is interpreted as an iterative application of the method of instrumental variables by Hausman (1974) and Lyttkens (1971).

[4] See Lyttkens (1970).

observations on the included predetermined variables, and ε_j is the j-th column of E. The vectors $\{\gamma_j, \beta_j\}$ are structural coefficients of the included jointly dependent variables (other than the variable with coefficient normalized at unity) and the included predetermined variables, respectively.

Combining the p equations into a system of simultaneous equations, we may denote the system by:

(3) $$y = Z\delta + \varepsilon,$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 & 0 & \dots & 0 \\ 0 & Z_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & Z_p \end{bmatrix}, \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}.$$

In this notation we write the reduced form as:

(4) $$y = [I \otimes X]\pi + \upsilon,$$

where $\otimes$ is the Kronecker product and

$$I \otimes X = \begin{bmatrix} X & 0 & \dots & 0 \\ 0 & X & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & X \end{bmatrix}, \quad \pi = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_p \end{bmatrix}, \quad \upsilon = \begin{bmatrix} \upsilon_1 \\ \upsilon_2 \\ \vdots \\ \upsilon_p \end{bmatrix}.$$

The vector π_j is the j-th column of Π and the vector υ_j is the j-th column of Υ.

The statistical specification of the simultaneous equations model, including the instrumental variables, is given by the following list of assumptions:

(i) X and W are random matrices.

(ii) $X'X$, $W'W$, and $W'X$ have ranks q, t and $\min(q, t)$, respectively, with probability one.

(iii) $\operatorname{plim}_{n\to\infty} n^{-1}X'X = \Sigma_{x'x}$, $\operatorname{plim}_{n\to\infty} n^{-1}W'W = \Sigma_{w'w}$, $\operatorname{plim}_{n\to\infty} n^{-1}W'X = \Sigma_{w'x}$, $\Sigma_{x'x}$ and $\Sigma_{w'w}$ positive definite and rank $\Sigma_{w'x} = \min(q, t)$.

(iv) $E(\varepsilon) = 0$.

(v) $V(\varepsilon) = \Sigma \otimes I$, Σ positive definite.

(vi) The vectors of disturbances $(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{pn})$, each corresponding to a given observation $(i = 1, 2, \dots, n)$, are distributed independently and identically over observations, and are distributed independently of X and W.

(vii) The structural model is complete.

Under these assumptions:

$$\operatorname{plim} n^{-1}\mathrm{E}'\mathrm{E} = \Sigma,$$

$$\operatorname{plim} n^{-1}X'\mathrm{E} = 0,$$

$$\operatorname{plim} n^{-1}W'\mathrm{E} = 0.$$

Further, the vector $n^{-1/2}[I \otimes X]'\varepsilon$ is asymptotically normal with mean zero and variance–covariance matrix $\Sigma \otimes \Sigma_{x'x}$; the vector $n^{-1/2}[I \otimes W]'\varepsilon$ is asymptotically normal with mean zero and variance–covariance matrix $\Sigma \otimes \Sigma_{w'w}$.[5]

The most important properties of instrumental variables are that these variables are uncorrelated (asymptotically) with the errors E and that they are correlated (asymptotically) with the predetermined variables X. From the reduced form we can deduce that

$$\operatorname{plim} n^{-1}W'Y = \operatorname{plim} n^{-1}W'X\Pi + \operatorname{plim} n^{-1}W'\Upsilon = \Sigma_{w'x}\Pi,$$

so that under our assumptions the instrumental variables are correlated (asymptotically) with the jointly dependent variables.

An additional assumption:

(viii) ε is $N(0, \Sigma \otimes I)$,

is essential for consideration of the problem of efficient estimation. Under this assumption and the other assumptions we have made, the full information maximum likelihood estimator attains the Cramer–Rao lower bound for the asymptotic variance–covariance matrix of (essentially) any consistent estimator of the structural parameters. This bound, stated in terms of the asymptotic information matrix, depends on the likelihood function. Without an explicit likelihood function, such as that associated with a normal distribution of the errors, it is impossible to discuss efficient estimation. Of course, the asymptotic distribution theory we develop for instrumental variables estimators is valid whether or not the errors are normally distributed, provided that our other assumptions on the distribution of the errors are satisfied. While alternative estimators may be compared with regard to relative efficiency, no lower bound is available that would enable us to characterize any estimator as efficient in the class of consistent estimators.[6]

We consider estimation of the structural coefficients δ in the absence of restrictions on the variance–covariance matrix Σ of the errors; the full information maximum likelihood and three-stage least squares estimators are efficient. The asymptotic variance–covariance matrix of these estimators, also the Cramer–Rao bound, is $\{\Sigma'_{\bar{X}'z}(\Sigma \otimes \Sigma_{x'x})^{-1}\Sigma_{\bar{X}'z}\}^{-1}$, where $\bar{X} = I \otimes X$ and the matrix $\Sigma_{\bar{X}'z}$ has the form,

$$\Sigma_{\bar{X}'z} = \begin{bmatrix} \Sigma_{x'z_1} & 0 & \cdots & 0 \\ 0 & \Sigma_{x'z_1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \Sigma_{x'z_p} \end{bmatrix}$$

Further,

$$\Sigma_{x'z_j} = [\Sigma_{x'y_j} \quad \Sigma_{x'x_j}] = [\Sigma_{x'x}\Pi_j \quad \Sigma_{x'x_j}], \qquad (j = 1, \ldots, p).$$

[5] This specification of the simultaneous equations model is employed, for example, by Malinvaud (1970), pp. 250–253, 369–373.

[6] For further discussion of efficiency in simultaneous equations estimation, see Rothenberg (1974).

In this expression $X'X_j$ is a submatrix of $X'X$ and Π_j is a submatrix of Π corresponding to the reduced form equations,

$$Y_j = X\Pi_j + \Upsilon_j.$$

We also consider estimation of the structural coefficients of a single equation δ_j, subject only to the identifying restrictions for that equation; the limited information maximum likelihood and two-stage least squares estimators are efficient. The asymptotic variance–covariance matrix of these estimators, also the Cramer–Rao bound, is $\sigma_{jj}\{\Sigma'_{x'z_j}\Sigma^{-1}_{x'x}\Sigma_{x'x_j}\}^{-1}$, $(j = 1, 2, \ldots, p)$. This completes our discussion of the simultaneous equations model.

3. The Method of Instrumental Variables

The method of instrumental variables for estimation of a single equation in a system of simultaneous equations is the following: We suppose that r_j jointly dependent and s_j predetermined variables are included in the j-th equation and that a subset of $t_j = r_j + s_j - 1$ instrumental variables W_j is selected from the set of t instrumental variables W. The instrumental variables estimator d_j of δ_j is obtained by solving the equation:

$$W'_j y_j = W'_j Z_j d_j,$$

obtaining,

$$d_j = (W'_j Z_j)^{-1} W'_j y_j. \tag{5}$$

Examples of instrumental variables are:

1. The indirect least squares estimator,

$$d_j = (X'Z_j)^{-1} X'y_j,$$

where $t = p = r_j + s_j - 1$.

2 The two-stage least squares estimator:

$$d_j = \{Z'_j X(X'X)^{-1} X'Z_j\}^{-1} Z'_j X(X'X)^{-1} X'y_j,$$

where $W_j = X(X'X)^{-1}X'Z_j$, the fitted values from a regression of the right-hand-side variables in the equation Z_j on the matrix of predetermined variables X.

We first observe that any instrumental variables estimator d_j is consistent since:

$$\operatorname{plim} d_j = \delta_j + \operatorname{plim}(n^{-1}W'_j Z_j)^{-1} \operatorname{plim} n^{-1} W'_j \varepsilon_j = \delta_j,$$

where:

$$\operatorname{plim}(n^{-1}W'_j Z_j) = [\Sigma_{w'_j x}\Pi_j \quad \Sigma_{w'_j x_j}] = \Sigma_{w'_j z_j}$$

is a matrix of constants and:

$$\operatorname{plim}(n^{-1}W'_j \varepsilon_j) = 0.$$

Second, this estimator is asymptotically normal, since the vector $n^{-1/2}W'_j\varepsilon_j$ is asymptotically normal. Further, this distribution has expectation zero and variance–covariance matrix $\sigma_{jj}\{\Sigma'_{w_j z_j}\Sigma^{-1}_{w'_j w_j}\Sigma_{w'_j z_j}\}^{-1}$.

The principal results of estimation theory for instrumental variables methods can be embodied in two theorems presented here without proof. They are proved, and their implications discussed in detail, in our earlier paper [1971].

Theorem A. Under the assumptions given above, the estimator

$$d_j = (W_j'Z_j)^{-1}W_j'y_j \tag{6}$$

of the parameters δ_j of the equation (3) is asymptotically efficient if and only if the matrix of instrumental variables W_j can be transformed by means of a nonsingular matrix into a matrix that includes two subsets $W_j = [W_{j1}, W_{j2}]$ with the properties:

$$\text{plim}\, n^{-1}W_{j1}'X = \Pi_j'\Sigma_{x'x}$$
$$\text{plim}\, n^{-1}W_{j2}'X = I_j'\Sigma_{x'x}.$$

This theorem provides the necessary and sufficient conditions for consistent and efficient estimation of one equation from the model (1). The following theorem provides the same conditions for an estimation of all equations taken together.

Theorem B. Under the assumptions given above, the estimator

$$d = (\overline{W}'Z)^{-1}\overline{W}'y \tag{7}$$

of the parameters δ in model (4) is asymptotically efficient if and only if the matrix of instrumental variables can be transformed by means of a nonsingular matrix into a matrix with typical submatrix $\overline{W}_{ij}$ $(W = \{W_{ij}\})$ that can be partitioned into two subsets $W_{ij} = [W_{ij1}, W_{ij2}]$ such that:

$$\text{plim}\, n^{-1}W_{ij1}'X = \sigma^{ij}\Pi_j'\Sigma_{x'x},$$
$$\text{plim}\, n^{-1}W_{ij2}'X = \sigma^{ij}I_j'\Sigma_{x'x},$$

where σ^{ij} is the (i, j) element of the matrix Σ^{-1}.

With the above results available, alternative techniques for choosing W_j and $\overline{W}$ can be considered.

Observe that for 2SLS and 3SLS, the instrumental variables for the included jointly dependent variables in a particular equation may be written, in the notation used above,

$$W_{j1} = X\hat{\Pi}_j \qquad \text{(LIVE)},$$

and

$$W_{ij1} = s_{ij}X\hat{\Pi}_j \qquad \text{(FIVE)}.$$

The estimator $\hat{\Pi}_j$ is a consistent estimator of the portion of the reduced form associated with the j-th structural equations. Any consistent estimator of these reduced form parameters can be used to generate estimators that are consistent and efficient. We have called estimators based upon a consistent estimator of the reduced form parameters Limited Information Instrumental Variables Efficient (LIVE) and Full Information Instrumental Variables Efficient (FIVE) estimators.

The instrumental variables used in LIVE and FIVE estimation are of the form $X\hat{\Pi}_j$, $\hat{\Pi}_j$ a consistent estimator of the reduced form parameters Π_j. Because model (1) is complete,

$$\Pi = -\mathrm{B}\Gamma^{-1}, \tag{8}$$

and a consistent estimator of Π is given by

$$\hat{\Pi} = -\hat{\mathrm{B}}\hat{\Gamma}^{-1}, \tag{9}$$

where $\hat{\mathrm{B}}$ and $\hat{\Gamma}$ are consistent estimators of the structural parameter matrices B and Γ of model (1).

Now the vectors $X\hat{\Pi}$ are "fitted values" from a consistent estimator of the reduced form; they need not be developed from least squares, as in 2SLS or 3SLS, but can be obtained from the derived reduced form estimates (9). The "fitted values" can be determined without solving explicitly for the derived reduced form by the following iterative algorithm:

Define the fitted values by

$$\hat{Y}_{r+1} = -\hat{Y}_r(\hat{\Gamma}^{-1} - I) - X\hat{\mathrm{B}}, \tag{10}$$

for any consistent estimators $\hat{\Gamma}$ and $\hat{\mathrm{B}}$, and iterate until $\hat{Y}_{r+1} = \hat{Y}_r + \mu$, μ arbitrarily small. At that point,

$$\hat{Y}_r = -X\hat{\mathrm{B}}\hat{\Gamma}^{-1} - \mu = X\hat{\Pi} - \mu. \tag{11}$$

The error in this iterative process (μ) must be zero if LIVE estimates obtained from (11) are to be asymptotically efficient.

A. "*True*" *Exogenous Instrumental Variables*

In estimating the Liu quarterly model (1963) used to illustrate LIVE and FIVE estimation in our earlier paper (1971), we used as instruments those predetermined variables which were not lagged values of jointly dependent variables. Instrumental variables do not need to be drawn from among the predetermined variables in the model, so long as the conditions stated above hold for the matrix of instrumental variables W.

In this method a set of instrumental variables W_j, which may differ from equation to equation, is chosen for consistently estimating the structural parameters according to

$$d_j = (W'_j Z_j)^{-1} W'_j y_j.$$

This preliminary consistent estimate provides the estimate of the reduced form from which instrumental variables for efficient estimation are obtained.

B. *Repeated Least Squares Estimators*

Theil [1958] and independently Basmann [1957] devised the method of two-stage least squares, in which the reduced form is estimated without constraint by ordinary least squares, and the resulting fitted values used as regressors in structural estimation. Zellner and Theil [1962] proposed the method of three-stage, least squares for estimating all equations simultaneously in a way equivalent to full information maximum likelihood. Klein [1955] and Madansky [1964] demonstrated that the 2SLS and 3SLS estimators were instrumental variables estimators.

In the typical econometric model, the number of predetermined variables is large in relation to the number of observations. Further, predetermined variables are often highly collinear. Computation of ordinary least squares estimator of the reduced form is difficult, and indeed, becomes impossible when the number of predetermined variables exceeds the number of observations. Computation of a derived reduced form estimator circumvents these difficulties.

By analogy with two- and three-stage least squares, a number of estimators have been proposed which employ a multistage procedure. In such estimators regressions are run upon a subset of the predetermined variables, and the fitted values from these regressions substituted for the included jointly dependent variables in second-stage regressions to estimate the structural parameters. The resulting estimator can be written

$$d_j = \begin{bmatrix} \hat{Y}_j'\hat{Y}_j & \hat{Y}_j'X_j \\ X_j'\hat{Y}_j & X_j'X_j \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_j' \\ X_j' \end{bmatrix} y_j, \tag{12}$$

where $\hat{Y}_j$ is the set of fitted values from the initial regressions:

$$\hat{Y}_j = V_j(V_j'V_j)^{-1}V_j'Y_j. \tag{13}$$

In this expression V_j is a set of predetermined variables chosen by any of the methods proposed above.

Cooper (1972) employed arbitrary subsets of the predetermined variables in a repeated least squares estimator. Fisher (1965) proposed a method that took *a priori* information in the model into account in selecting subsets of the predetermined variables through a stepwise regression procedure. A repeated least squares estimator based upon principal components of the predetermined variables was described by Kloek and Mennes (1960) and discussed by Amemiya (1966). Another version of the principal components estimator was applied by Evans and Klein (1968) to the Wharton model; this version has had widespread application.

Repeated least squares estimators must be designed with care if they are to be consistent or efficient. In our earlier paper, we showed that such estimators are consistent if they reduce to instrumental variables estimators, or if the initial regressions happen to estimate the relevant portions of the reduced form consistently. Repeated least squares estimators are efficient only if they provide consistent estimators of the reduced form parameters.

C. Structural Ordering Instrumental Variables (SOIV)

The SOIV method was proposed by Fisher (1965), was later corrected by Mitchell and Fisher (1970), and was applied to the Brookings model by Mitchell (1971). The method employs the following algorithm in choosing instruments for a jointly dependent variable:

1. To each predetermined variable, assign a vector with a number of elements equal to the maximum "order" of predetermined variables appearing in the model. The maximum "order" is determined in the course of further development of the algorithm.
2. Order one is assigned to each predetermined variable appearing in the equation defining the jointly dependent variable for which instruments are

being developed. Order two is assigned to predetermined variables appearing in equations defining the jointly dependent variables appearing in the first equation. Reappearance in the remaining equations of jointly dependent variables that have already been treated is ignored. Order three is assigned to predetermined variables appearing in the equations defining jointly dependent variables appearing in the second equations. The process continues until all of the equations to be estimated as a simultaneous block have been used in accumulating the ordering vector.

3. At each occurrence of a predetermined variable in the previous procedure, an entry is made in the next unused location in the ordering vector for that variable. Entries from the equation defining the jointly dependent variable for which instruments are sought are defined to have order one, and the first unused entry is assumed to have indefinite value.
4. The ordering vectors determine a lexicographic ordering on the predetermined variables. A variable with the ordering vector $(1, 3, \infty)$ precedes one with the vector $(1, 4, 5, \infty)$, and follows one with the ordering vector $(1, 2, 3, \infty)$. Frequently, two or more variables will have the same ordering vector; these variables are treated together as if they were a single variable for purposes of the following analysis.
5. The significance of a variable or set of variables in explaining the variance of the jointly dependent variable for which instruments are sought is tested by a stepwise procedure. A regression is performed of the jointly dependent variable upon the largest possible number of variables or sets in the ordering for which moment matrices are nonsingular, omitting only the variables lowest in the lexicographic ordering, if any must be omitted. Then, the variable or set of variables appearing lowest in the ordering is dropped from the list of regressors, and the resulting regression calculated. If the sum of squared residuals from the latter regression is SSE_0, while that from the former set is SSE, the test statistic

$$F = SSE_0/SSE \quad [(n-q)/(n-q_0)],$$

where q is the number of regressors in the former regression, and q_0 is the number in the latter, is distributed as F. A level of significance is chosen, and the hypothesis that the variable(s) omitted from the second regression is significant is tested. If the variable is significant, it forms part of the instrumental variables for the jointly dependent variable under study. The process is repeated until all variables or sets in the "structural order" have been tested. At each stage, significant variables (s) are retained in the set of regressors used for testing further sets.

If the resulting subset of the predetermined variables is denoted X_0, the Mitchell–Fisher (1970) estimator is

$$d_j = (\hat{Z}_j'\hat{Z}_j)^{-1}\hat{Z}_j'y_j, \tag{14}$$

where

$$\hat{Z}_j = X_0(X_0'X_0)^{-1}X_0'Z_j.$$

This estimator is an instrumental variables estimator, since it can be written

$$d_j = (\hat{Z}'_j Z_j)^{-1} \hat{Z}'_j y_j.$$

The estimator then is clearly consistent. It is efficient only if $(X'_0 X_0)^{-1} X'_0 Z_j$ is a consistent estimator of the portion of the reduced form associated with the included jointly dependent variables.

D. Principal Components Instrumental Variables (PCIV)

The justification for the use of principal components of the predetermined variables as instruments is that the first $j < q$ principal components capture, in a sense, much of the variation existing in the full set of q predetermined variables. Hence, principal components appear to provide an ingenious way of eliminating very little information from X, while significantly reducing the size of the estimation problem.

If A is the $q \times q$ matrix of characteristic vectors of the moment matrix $X'X$ of predetermined variables, then the principal components are the set of orthogonal linear forms F satisfying

$$F = XA$$

and

$$F'F = \Lambda,$$

where

$$A'A = I,$$

and Λ is the diagonal matrix whose nonzero elements are the characteristic roots of $X'X$.

The Kloek–Mennes (1960) procedure seeks a subset of the principal components which has a maximum correlation with the jointly dependent variables for which the principal components are to be instruments, and yet has as low a correlation as possible with the included predetermined variables in the equation to be estimated. To select a subset with the latter property, Kloek and Mennes noted that for any principal component f_k

$$f'_k X_j (X'_j X_j)^{-1} X'_j f_k = V_k R_k^2 \qquad (k = 1, 2, \ldots, q).$$

V_k is the characteristic root associated with the principal component f_k, and R_k^2 is the multiple determination coefficient for the regression of f_k on the included predetermined variables X_j. Kloek and Mennes employ standardized and normalized variables in all of their computations. To select principal components, define θ_k by

$$\theta_k = V_k(1 - R_k^2),$$

and rank the principal components in decreasing order of the value of θ_k. To obtain m instrumental variables, choose the first m principal components in that ordering. No *a priori* rule exists for selecting the number m of principal components. Kloek and Mennes suggest beginning with a set that exactly identifies the equation to be estimated. If the resulting estimates have undesirable standard errors, the number of components is increased by using the next component in the ordering, until all components or all degrees of freedom are used without the standard errors increasing.

The Kloek and Mennes method results in a choice of V_j for equation (13) that has the form

$$V_j = [X_j F_j],$$

so that the substitution estimator based upon the method reduces to an instrumental variables method. As a result, the estimator always is consistent, but it will be efficient only in the most unlikely of circumstances.

The principal components estimator developed by Evans and Klein (1968) is a substitution estimator in which

$$V_j = F_j.$$

This estimator will be consistent and efficient only if all principal components are used in F_j, in which case 2SLS should be applied directly, or if the resulting estimate of the reduced form is consistent, a very unlikely event.

Taylor (1962) used principal components directly as instrumental variables, not in a repeated least squares estimator, in estimating the structural equations. While this method always is consistent, it cannot be efficient unless the conditions of Theorem A, above, are fulfilled.

E. Iterated Instrumental Variables (IIV)

The method of iterated instrumental variables was originated by Lyttkens (1970) and has been employed by Dutta and Lyttkens (1974). This method is initiated by estimating the structural equations by ordinary least squares and deriving (inconsistent) estimates of the reduced form parameters for the purpose of computing reduced form fitted values. These fitted values then are used as instruments for consistent estimation of the structural parameters, and the process is iterated until the parameter estimates cease to vary upon iteration. Clearly, only one additional iteration is required to produce LIVE estimates.

Durbin (1963) described a method for estimating the linear structural model by full-information maximum likelihood that required that the following set of normal equations be solved by iteration:

$$[I \otimes \overline{W}'Z][S^{-1} \otimes I]\delta = [I \otimes \overline{W}]'(S^{-1} \otimes I)y.$$

In this expression, $\overline{W}$ is the set of instrumental variables obtained by fitting the derived reduced form estimates produced in the previous iteration and combining these with the predetermined variables appearing in the equations. Beginning from some consistent set of parameters, $\overline{W}$ and δ are calculated and the equation is solved for δ. Iterating the process produces a new $\overline{W}$ and a new δ at each iteration, by using the implied estimator of the derived reduced form and the moment matrix of the residuals. At convergence, the estimator of δ is the maximum-likelihood estimator.

4. Empirical Comparison of Instrumental Variables Estimators

In view of the fact that only large-sample results are available about the statistical properties of estimators for the linear simultaneous equations model, it is of some interest to consider comparisons among consistent estimators of the covariance matrices of the instrumental variables estimators discussed earlier. It has been shown above that a number of consistent estimators of the parameters exist which are not efficient. Comparisons of consistent estimators of the asymptotic covariance matrices of these parameter estimators might provide some information helpful in selecting a method for choosing instrumental variables for consistent estimation. Intuitively, at least, it is appealing to argue that a method for choosing instrumental variables which is estimated to have smaller asymptotic covariance is superior to one having greater. Of greater importance are empirical evaluations of the effect that choices of instrumental variables for the initial consistent estimation have upon the resulting LIVE and FIVE parameter estimates. If LIVE and FIVE are relatively insensitive to the initial choice of instrumental variables, that choice becomes of much less consequence.

Rothenberg (1947) presented measures of the gains in efficiency resulting from taking into account more *a priori* information. We employ similar measures in comparing the efficiency of IV, LIVE, and FIVE estimators. The results on minimum variance bounds given above imply that efficiency must increase as the estimation method moves from IV to LIVE to FIVE. Empirical estimation of the asymptotic covariance matrices of the estimators provides a means of evaluating the efficiency gain. While it is quite inexpensive to solve for reduced form fitted values when proceeding from IV to LIVE, LIVE estimates are at least twice as expensive to produce as IV estimates. The increase in cost arises from the need to obtain reduced form fitted values and then to re-estimate the model. FIVE estimates are very substantially more expensive, since they involve not only computation of the estimate of the structural covariance matrix Σ, but also estimates of the coefficients. Preparation of these estimates requires inversion of a matrix whose order is the number of structural parameters. Even when this is possible, it is extremely expensive, since the computational effort involved in most algorithms for inversion increases with the cube of the order of the matrix to be inverted. The value of incurring such a computational burden can be weighed against the estimated increase in efficiency from proceeding from LIVE to FIVE.

Forecasting with an econometric model makes use of the reduced form of the model. Therefore it is of interest to make the same inquiries about the estimates of the reduced form that have just been described for the structural form:

1. What is the relative efficiency of alternative methods of developing reduced form estimates, using different choices of instrumental variables?
2. How sensitive are the reduced form parameter estimates derived from IV, LIVE, and FIVE estimates of the structural parameters to alternative choices of the instrumental variables for initial IV estimation?
3. How significant is the gain in efficiency from proceeding from IV to LIVE to FIVE in terms of the efficiency of the reduced form estimators derived from the structural estimator?

A. Klein Model I and Covariance Comparisons

We employ an annual model of the United States economy 1921–1941 developed by Klein (1950), commonly referred to as Klein Model I, in measuring the relative efficiency of alternative estimators of the structural and of the reduced form parameters. Klein Model I has three behavioral equations and three identities, and is linear. It determines six jointly dependent variables from eight predetermined variables, including a dummy variable of constant unit value corresponding to the intercept in each behavioral equation. The model, as estimated by 2SLS, is given in Table 1. The parameter estimates given there are in accord with previous

TABLE 1
KLEIN MODEL I 2SLS ESTIMATES
All Predetermined Variables as Instruments

1. Consumption

$$C = \underset{(0.118)}{0.017P} + \underset{(0.107)}{0.216P_{-1}} + \underset{(0.040)}{0.810W} + \underset{(1.32)}{16.6}$$

$R^2 = 0.977 \qquad SE = 0.223$

2. Investment

$$I = \underset{(0.173)}{0.150P} + \underset{(0.163)}{0.616P_{-1}} - \underset{(0.036)}{0.158K_{-1}} + \underset{(7.54)}{20.3}$$

$R^2 = 0.885 \qquad SE = 0.257$

3. Private Wages

$$W^* = \underset{(0.036)}{0.439E} + \underset{(0.039)}{0.147E_{-1}} + \underset{(0.029)}{0.130T} + \underset{(1.15)}{1.50}$$

$R^2 = 0.987 \qquad SE = 0.151$

4. Corporate Profits

$$P = C + I + G - X - W$$

5. Wages

$$W = W^* + W^{**}$$

6. Private Product

$$E = X + P + W^*$$

Exogenous Variables:

a. P_{-1} = Profits, lag 1	e. X = Indirect taxes
b. K_{-1} = Capital stock, lag 1	f. W^{**} = Government wage bill
c. E_{-1} = Private product, lag 1	g. G = Government expenditures
d. T = Time trend, 1921, = −10	

estimates, for example, those reported by Rothenberg and Leenders (1964), although the computer program used for estimates treated 2SLS as a type of instrumental variables estimator.

For the purpose of quantitative comparisons of the relative efficiency of alternative estimators, we employ consistent estimates of the covariance matrices of the alternative estimators, and compare these estimates against consistent estimates of the minimum variance bounds. The assumptions stated above imply that

$$V_{2S} = [Z_i'X(X'X)^{-1}X'Z_i]^{-1}[Z_i'X(X'X)^{-1}X'Z_j] \cdot [Z_j'X(X'X)^{-1}X'Z_j]^{-1} \cdot s_{ij} \tag{15}$$

is a consistent estimator of the minimum variance bound (9) for the covariance matrices of the estimators of the distinct equations which are efficient in the class of limited information estimators. The same assumptions, together with the assumptions about the properties of the instrumental variables, ensure that

$$(16) \qquad V_{IV} = [Z_i'W_i(W_i'W_i)^{-1}W_i'Z_i]^{-1} \cdot [Z_i'W_i(W_i'W_i)^{-1}W_i'W_j(W_j'W_j)^{-1}W_j'Z_j] \cdot [Z_j'W_j(W_j'W_j)^{-1}W_j'Z_j]^{-1} \cdot s_{ij}$$

is a consistent estimator of the covariance matrix of the estimators of two distinct equations, when those estimators are consistent but not efficient. Because (15) is a consistent estimator of the minimum variance bound (MVB) for any consistent estimators of two distinct equations, (16) must differ from (15) by a positive semidefinite matrix, when the estimators for which (16) is the estimate of the covariance matrix are not efficient.

The MVB for estimators from the class of full information estimators (10) can be estimated consistently, under the assumptions given above, by

$$(17) \qquad V_{3S} = [S^{-1} \otimes Z'X(X'X)^{-1}X'Z]^{-1}.$$

The MVB for limited information estimators of all parameters, considered together, differs from the MVB for full information estimators by a positive semidefinite matrix, and the same property holds for comparisons between a consistent estimator of the covariance matrix of the limited information estimators and the consistent estimator of the covariance matrix for (efficient) full information estimators (17).

Comparisons of the relative efficiency of alternative (derived) estimates of the reduced form depend upon the estimates of the covariance matrices for the structural estimators just stated. The comparisons make use of the expression

$$(18) \qquad \left[\Gamma^{-1} \otimes \begin{pmatrix} \Pi' \\ I \end{pmatrix}\right]' \Delta \left[\Gamma^{-1} \otimes \begin{pmatrix} \Pi' \\ I \end{pmatrix}\right]$$

for the covariance matrix of the derived reduced form, where Δ is the covariance matrix of the structural estimator from which the reduced form estimates are derived. Expression (18) was obtained from Goldberger, Nagar, and Odeh (1961); a more direct derivation is available in Dhrymes (1970). The covariance matrix (18) is estimated consistently by replacing Γ, Π, and Δ with consistent estimators.

The covariance matrix of the unconstrained reduced form (8) is estimated consistently by

$$(19) \qquad V_{URF} = \hat{\Omega} \otimes (X'X)^{-1},$$

where $\hat{\Omega}$ is a consistent estimator of the reduced form covariance matrix Ω. The estimator of the covariance matrix of the unconstrained reduced form (19) differs from any consistent estimator of the derived reduced form (18) by a positive semidefinite matrix.

Whenever the covariance matrices of the structural and reduced forms Σ and Ω, or any of their elements, appear in this analysis, they must be replaced by consistent estimators. These matrices are estimated consistently by the moment matrices of the residuals from consistent estimates of the structural and reduced forms, respectively. Throughout the subsequent analysis, Γ, Π, and Σ are replaced by the estimates developed from two-stage least squares, while Ω is estimated from the moment matrix of the residuals from ordinary least squares estimation of the reduced form.

To reduce the covariance matrices which are being compared to scalar measures of efficiency, three functions of the covariance matrices have been computed: 1) the sum of the elements of a covariance matrix; 2) the trace of a covariance matrix; and 3) the determinant of the covariance matrix. Each of these measures preserves the ranking of the matrices in terms of relative efficiency, that is, if B_1 is the covariance matrix of an estimator that is more efficient than one with covariance B_0, then any one of the measures has the property that $M(B_0) > M(B_1)$, where M denotes the fact that the measures are functions of the elements of the covariance matrices.

Table 2 contains the results of estimating Klein Model I consistently by the four inefficient instrumental variables methods we have presented. The coefficient estimates are very similar to those prepared by two-stage least squares in Table 1.

TABLE 2
KLEIN MODEL I. CONSISTENT STRUCTURAL ESTIMATES BY INSTRUMENTAL VARIABLE

Equation Variables	Coefficients (Standard Errors)			
	SOIV	PCIV	TEIV	IIV
1. Consumption				
Constant	16.43	16.62	16.05	16.91
	(1.35)	(1.42)	(1.34)	(1.33)
*Profits	0.0667	−0.0003	0.0611	−0.1499
	(0.142)	(0.165)	(0.140)	(0.134)
Profits, Lag 1	0.1790	0.2305	0.1650	0.3388
	(0.124)	(0.145)	(0.122)	(0.117)
*Wages	0.8078	0.8100	0.8248	0.8214
	(0.040)	(0.040)	(0.041)	(0.041)
2. Investment				
Constant	21.22	25.77	22.84	21.41
	(7.73)	(11.3)	(9.49)	(8.04)
*Profits	0.1197	−0.0281	0.0671	0.1136
	(0.182)	(0.325)	(0.255)	(0.195)
Profits, Lag 1	0.6421	0.7690	0.6874	0.6474
	(0.169)	(0.287)	(0.229)	(0.180)
Capital Stock, Lag 1	−0.1620	−0.1827	−0.1694	−0.1629
	(0.037)	(0.053)	(0.044)	(0.038)
3. Private Wages				
Constant	1.571	1.846	1.686	1.351
	(1.15)	(1.15)	(1.15)	(1.15)
*Private Product	0.4253	0.3732	0.4035	0.4673
	(0.036)	(0.041)	(0.039)	(0.040)
Private Product, Lag 1	0.1594	0.2087	0.1801	0.1198
	(0.039)	(0.043)	(0.041)	(0.042)
Time Trend	0.1337	0.1464	0.1390	0.1235
	(0.029)	(0.030)	(0.029)	(0.029)

* Indicates an endogenous variable.

The coefficients are relatively stable across alternative choices of instrumental variables, with the PCIV estimates varying slightly more from the 2SLS estimates than the others.

Parameter estimates derived from the use of each of the methods of consistent structural estimation to develop fitted values for use as instrumental variables in preparing LIVE estimates are given in Table 3. As might be expected, the parameter

TABLE 3

KLEIN MODEL I. STRUCTURAL PARAMETERS ESTIMATED BY LIMITED INFORMATION EFFICIENT INSTRUMENTAL VARIABLES (LIVE) METHODS

Equation Variable	Coefficients (Standard Error)				
	2SLS	SOIV	PCIV	TEIV	IIV
1. Consumption					
Constant	16.55 (1.32)	16.79	16.74	16.77	16.80
*Profits	0.0173 (0.118)	−0.1182	−0.1097	−0.1153	−0.1156
Profits, Lag 1	0.2162 (0.107)	0.3133	0.3056	0.3105	0.3121
*Wages	0.8102 (0.040)	0.8214	0.8222	0.8217	0.8205
2. Investment					
Constant	20.28 (7.54)	21.50	21.52	21.63	21.60
*Profits	0.1502 (0.173)	0.1105	0.1098	0.1062	0.1076
Profits, Lag 1	0.6159 (0.163)	0.6500	0.6506	0.6537	0.6527
Capital Stock, Lag 1	−0.1578 (0.036)	−0.1633	−0.1634	−0.1639	−0.1638
3. Private Wages					
Constant	1.500 (1.15)	1.552	1.607	1.571	1.601
*Private Product	0.4389 (0.036)	0.4290	0.4186	0.4254	0.4197
Private Product, Lag 1	0.1467 (0.039)	0.1561	0.1658	0.1593	0.1648
Time trend	0.1304 (0.029)	0.1328	0.1353	0.1337	0.1351

* Indicates an endogenous variable

Note: All limited information instrumental variables efficient estimators have asymptotic variances and covariances equal to those of two-stage least squares.

estimates for each of the LIVE estimates except 2SLS are very similar, and they do not vary significantly from the 2SLS estimates. Thus, the choice of instrumental variables for LIVE does not make appreciable difference to the resulting coefficient estimates.

The outcome of structural estimation by full-information methods is shown in Table 4, for each of the methods of choosing instrumental variables, and for Durbin's method of obtaining full information maximum likelihood estimates. Once again, the parameter estimates show little effect of the choice of instrumental variables. The full information maximum likelihood estimates appear to differ slightly from the others, but none of the variation in coefficients is significant.

TABLE 4

KLEIN MODEL I. STRUCTURAL PARAMETERS ESTIMATED BY FULL INFORMATION EFFICIENT INSTRUMENTAL VARIABLES (FIVE) METHODS

Equation	Coefficients					
Variable	(Standard Errors)					
	3SLS	SOIV	PCIV	TEIV	IIV	FIML
1. Consumption						
Constant	16.61 (1.30)	16.60	16.44	16.54	16.55	16.52
*Profits	0.0557 (0.108)	0.0514	0.0583	0.0532	0.0744	0.0569
Profits, Lag 1	0.2240 (0.100)	0.2260	0.2186	0.2234	0.2134	0.2336
*Wages	0.7902 (0.038)	0.7913	0.7955	0.7932	0.7883	0.7881
2. Investment						
Constant	25.78 (6.79)	25.73	25.12	25.44	25.62	27.99
*Profits	−0.0169 (0.162)	−0.0158	−0.0136	−0.0131	−0.0207	−0.1733
Profits, Lag 1	0.7514 (0.153)	0.7506	0.7462	0.7473	0.7529	0.8536
Capital Stock Lag 1	−0.1822 (0.032)	−0.1820	−0.1788	−0.1805	−0.1812	−0.1884
3. Private Wages						
Constant	1.972 (1.12)	1.965	2.052	1.996	2.047	2.507
*Private Product	0.3886 (0.032)	0.3900	0.3802	0.3866	0.3782	0.3555
Private Product, Lag 1	0.1905 (0.034)	0.1893	0.1980	0.1922	0.2001	0.2157
Time Trend	0.1579 (0.028)	0.1572	0.1588	0.1577	0.1614	0.1694

* Indicates an endogenous variable.

Note: All full information instrumental variables efficient estimates have asymptotic variances and covariance equal to those of three-stage least squares.

The measures of relative asymptotic efficiency of structural estimates are given in Table 5. Performance of the estimators varies from equation to equation. Except for the consumption function, however, the method of structural ordering is the most efficient.

TABLE 5

KLEIN MODEL I. VARIANCE MEASURES FOR ALTERNATIVE STRUCTURAL ESTIMATES

Equation						
Measure	SOIV	PCIV	TEIV	IIV	2SLS	3SLS
1. Consumption						
Sum	1.700	1.892	1.695	1.665	1.643	1.599
Trace	1.846	2.068	1.839	1.802	1.772	1.725
Generalized	0.479	0.695	0.474	0.425	0.324	0.239
Variance	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$
2. Investiment						
Sum	58.83	126.5	88.61	63.71	56.07	45.47
Trace	59.77	128.9	90.16	64.75	56.95	46.26
Generalized	0.667	2.13	1.31	0.772	0.607	0.476
Variance	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$	$\times 10^{-8}$
3. Private Wages						
Sum	1.306	1.318	1.312	1.314	1.305	1.232
Trace	1.322	1.333	1.327	1.330	1.321	1.248
Generalized	0.680	0.886	0.781	0.831	0.668	0.460
Variance	$\times 10^{-11}$	$\times 10^{-11}$	$\times 10^{-11}$	$\times 10^{-11}$	$\times 10^{-11}$	$\times 10^{-11}$
Entire Model						
Sum	64.65	137.5	97.45	70.74	62.66	51.42
Trace	62.94	132.2	93.32	67.88	60.04	49.18
Generalized	1.26	4.82	1.72	1.58	0.746	0.220
Variance	$\times 10^{-29}$	$\times 10^{-29}$	$\times 10^{-29}$	$\times 10^{-29}$	$\times 10^{-29}$	$\times 10^{-29}$

The principal components method performed uniformly least well. Differences among the methods other than principal components (PCIV) are relatively small, in most cases. In view of the very considerable expense of developing SOIV estimates, IIV would appear to be the method of choice for estimating Klein Model I by instrumental variables.

A gain in efficiency is achieved by employing limited information efficient estimators in place of any consistent but inefficient ones. Table 6 shows the gain in efficiency, based on the trace measure, from using 2SLS (or any other LIVE estimator) in place of the four consistent but inefficient methods. The efficiency gain is defined as the percentage increase from the trace of the covariance matrix of the LIVE estimator to the trace of the covariance matrices of the consistent but

TABLE 6

EFFICIENCY GAIN FROM USE OF MORE INFORMATION STRUCTURAL ESTIMATES

(In Percent, Based on Trace)

Estimation Method	Equation			
IV to LIVE	CNS	INV	W*	MODEL
SOIV	4.2	4.9	0.0	4.8
PCIV	16.7	126.3	0.9	120.2
TEIV	3.8	58.3	0.5	55.4
IIV	1.7	13.7	0.7	12.4
LIVE to FIVE	2.7	29.8	5.9	22.1

inefficient estimators. Since SOIV estimates are relatively the most efficient of the consistent but inefficient methods, the gain is least pronounced for this method. A further gain in efficiency is obtained by employing a full information estimator in place of a limited information one. The trace measure used to construct Table 6 makes it difficult to tell whether a greater gain is achieved by employing a full information method instead of a limited information method or by using a limited information method in preference to a consistent but inefficient one. This result is not in accord with results we reported earlier (1971) for the Liu model using generalized variance as the measure but agrees with the result reported for Klein Model I by Rothenberg (1974).

B. Comparison of Structural and Reduced Form Estimators

In this section, it is our objective to compare alternative estimators of the structural and reduced form which employ different algorithms for selecting instrumental variables, and which use different amounts of information. While we have not reproduced the tables of reduced form parameter estimates which are analogous to Tables 3, 4, and 5 above, the reduced form parameter estimates are no less stable than the structural estimates, with respect to choice of estimation method and choice of instrumental variables.

TABLE 7

KLEIN MODEL I. EFFICIENCY MEASURES FOR REDUCED FORM ESTIMATES

Equation	Estimation Method						
Measure	SOIV	PCIV	TEIV	IIV	2SLS	3SLS	OLSQ
1. Consumption							
Trace	85.70	108.59	85.12	75.88	67.98	63.49	829.85
2. Investment							
Trace	43.02	47.96	44.89	42.45	42.01	39.74	423.72
3. Private Wages							
Trace	52.00	62.00	54.55	50.78	48.26	45.63	576.47
4. Model*							
Trace	535.69	650.14	550.76	496.46	467.81	441.91	5444.30

* In addition to measures for the three structural equations given above, the measure for the entire model includes the results for the three identities.

Table 7 gives the values of the efficiency measures for estimates of the reduced form parameters derived from the consistent but inefficient estimation methods we have discussed, and for three efficient methods using differing amounts of information. For the derived reduced form too, principal components is uniformly the least efficient method. In this case, iterated instrumental variables provides unequivocally more efficient estimates than the other consistent but inefficient methods.

Table 8 depicts the gain in efficiency of estimate of the derived reduced form that is obtained by employing an efficient structural estimator for each equation in place of a consistent but inefficient method, for each of the four instrumental

TABLE 8

EFFICIENCY GAIN FROM STRUCTURAL ESTIMATION DERIVED REDUCED FORM
(In Percent, Based upon Trace)

Estimation Method	Equation			
	CNS	INV	W*	MODEL
IV to LIVE				
SOIV	26.1	2.4	7.7	14.5
PCIV	59.7	14.2	28.5	39.0
TEIV	25.2	6.8	13.0	17.7
IIV	11.6	1.1	5.2	6.1
LIVE to FIVE	7.1	5.7	5.8	5.9

variables methods we have described. The table also reflects the results of employing FIVE instead of a LIVE estimator. The percentage gain from the use of LIVE methods is substantial except for iterated instrumental variables, indicating that for Klein Model I the IIV method is nearly as efficient as LIVE. The gain from using structural information about the covariance matrix in obtaining derived reduced form estimates is modest.

5. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH

The results reported above lead to the following conclusions, on the basis of alternative estimation of Klein Model I:

A. The most appealing method for consistent estimation of the linear structural model is to begin with ordinary least squares applied to the structural form. The fitted values of jointly dependent variables can be used as instrumental variables to obtain a consistent estimator.

B. For efficient estimation the fitted values of the jointly dependent variables from an initial consistent estimator should be used as instruments in obtaining a LIVE estimator. The expense of computing FIVE estimates may well outweigh the benefits.

C. Choice of instrumental variables for obtaining a consistent estimator appears to have little impact on the resulting LIVE or FIVE structural or reduced form parameter estimates. The implication is that the initial instrumental variables should be chosen so as to minimize computational difficulty.

D. There appears to be no advantage and great computational difficulty associated with iteration of the method of instrumental variables beyond LIVE or FIVE estimators.

E. Where structural estimation is possible, it is to be preferred, as a means of deriving an estimate of the reduced form, to unconstrained estimation of the reduced form.

The principal shortcoming of the research reported here is that the results apply with assurance only to one simple model. In the absence of general results on the small sample properties of estimators for the linear simultaneous equations model, there is no way to avoid this difficulty. It would be of considerable value to have available further results based on the application of our methods to larger models.

The results could also be strengthened by using Monte Carlo techniques to develop repeated samples for estimating the model. By repeating the analysis many times, based upon repeated samples from the same population, the degree of dependence of coefficient stability and relative efficiency upon the data used to estimate the model could be investigated. Direct comparisons of the methods could be made with respect to the error of estimate, and some additional hypotheses could be tested. Such Monte Carlo experiments would be expensive and as open to criticism for dependence on a particular model as the results reported here.

The estimation techniques used here are not computationally dependent upon the linearity of the model in the variables. Where nonlinearities can be expressed in terms of identities defining variables appearing in stochastic equations, the instrumental variables methods can be applied. However, the statistical properties of such estimates are unknown; the results given above certainly do not apply. Since models used for practical purposes of economic analysis and forecasting make extensive use of nonlinearities, the most pressing theoretical problem in simultaneous equations econometrics is to find estimators with desirable statistical properties that are suitable for nonlinear models.

The model we have used assumes that no autocorrelation is exhibited by any of the errors on any equation. A useful generalization of the LIVE and FIVE methods would be to modify them for the case of autocorrelated residuals.[7]

Federal Reserve Bank of San Francisco
Harvard University

References

Amemiya, T. (1966), "On the Use of Principal Components of Independent Variables in Two-Stage Least-Squares Estimation," *International Economic Review*, Vol. 7, No. 3, September, pp. 283–303.

Basmann, R. (1957), "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," *Econometrica*, Vol. 25, No. 1, January, pp. 77–83.

Brundy, J. and D. Jorgenson (1971), "Efficient Estimation of Simultaneous Equations by Instrumental Variables," *Review of Economics and Statistics*, Vol. 53, No. 3, August, pp. 207–224.

——— and ——— (1973), "Consistent and Efficient Estimation of Systems of Simultaneous Equations," in P. Zarembka (ed.), *Frontiers in Econometrics*, New York, Academic Press, 1973, pp. 215–244.

Cooper, R., (1972), "The Predictive Performance of Quarterly Econometric Models of the United States," in B. Hickman (ed.), *Econometric Models of Cyclical Behavior*, Studies in Income and Wealth, No. 36, New York, Columbia University Press, pp. 813–947.

Dhrymes, P. (1970), *Econometrics: Statistical Foundations and Applications*, New York, Harper & Row.

Durbin, J. (1963), "Maximum-Likelihood Estimation of the Parameters of a System of Simultaneous Regression Equations," unpublished manuscript.

Dutta, M. and E. Lyttkens (1974), "Iterative Instrumental Variables Method and Estimation of a Large Simultaneous System," unpublished manuscript.

Evans, M. and L. Klein (1968). *The Wharton Econometric Forecasting Model*, Philadelphia, Economic Forecasting Unit, Department of Economics, Wharton School of Finance and Commerce.

Fisher, F. (1965a). "The Choice of Instrumental Variables in the Estimation of Economy-Wide Econometric Models," *International Economic Review*, Vol. 6, No. 3, September, pp. 245–274.

Fisher, F. (1965b), "Dynamic Structure and Estimation in Economy-Wide Econometric Models," in J. Duesenberry, G. Fromm, L. Klein, and E. Kuh (eds.), *The Brookings Quarterly Econometric Model of the United States*, Chicago, Rand McNally, pp. 589–636.

[7] Fair (1973) has proposed a modification of the LIVE and FIVE estimators that takes autocorrelation into account. No analysis of the efficiency of the resulting estimators is currently available.

Geary, R. (1949), "Determination of Linear Relations Between Systematic Parts of Variables with Errors of Observation the Variances of Which are Unknown," *Econometrica*, Vol. 17, No. 1, January, pp. 30–58.

Goldberger, A., A. Nager and A. Odeh, (1961), "The Covariance Matrixes of Reduced-Form Coefficients and of Forecasts for a Structural Econometric Model," *Econometrica*, Vol. 29, No. 4, October, pp. 556–573.

Hausman, J. (1974), "An Instrumental Variable Approach to Full-Information Estimators for Linear and Non-Linear Econometric Models," *Econometrica*, forthcoming.

Klein, L. (1950), *Economic Fluctuations in the United States, 1921–1941*, New York, John Wiley and Sons, Inc.

——— (1955), "On the Interpretation of Theil's Method of Estimating Economic Relationships," *Macroeconomica*, Vol. 7, No. 3, December, pp. 147–153.

Kloek, T. and L. Mennes (1960), "Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables," *Econometrica*, Vol. 28, No. 1, January, pp. 45–61.

Liu, T. C. (1963), "An Exploratory Quarterly Econometric Model of Effective Demand in the Postwar U.S. Economy," *Econometrica*, Vol. 31, No. 3, July, pp. 301–348.

Lyttkens, E. (1970), "Symmetric and Asymmetric Estimation Methods," in E. Mosbaek and H. Wold, *Interdependent Systems*, Amsterdam, North-Holland pp. 434–459.

——— (1971), "The Iterative Instrumental Variables Method and the Full Information Maximum Likelihood Method for Estimating Interdependent Systems," unpublished manuscript.

Madansky, A. (1964), "On the Efficiency of Three-Stage Least-Squares Estimation," *Econometrica*, Vol. 32, No. 1–2, January–April, pp. 51–56.

Mitchell, B. (1971), "Estimation of Large Econometric Models by Principal Component and Instrumental Variable Methods," *Review of Economics and Statistics*, Vol. 53, No. 2, May, pp. 140–146.

Mitchell, B. and F. Fisher (1970), "The Choice of Instrumental Variables in the Estimation of Economy-Wide Econometric Models: Some Further Thoughts," *International Economic Review*, Vol. 11, No. 2, June, pp. 226–234.

Reiersøl, O. (1945), "Confidence Analysis by Means of Instrumental Sets of Variables," *Arkiv for Mathematik, Astronomi och Fysik*, Vol. 32.

Rothenberg, T. (1974), *Efficient Estimation with A Priori Information*, New Haven, Yale University Press.

Rothenberg, T. and C. Leenders (1964), "Efficient Estimation of Simultaneous Equation Systems," *Econometrica*, Vol. 32, No. 1–2, January–April, pp. 57–76.

Sargan, J. (1968), "The Estimation of Economic Relationships Using Instrumental Variables," *Econometrica*, Vol. 26, No. 3, July, pp. 393–415.

Taylor, L. D. (1962), *The Principal-Component-Instrumental-Variable Approach to the Estimation of Systems of Simultaneous Equations*, unpublished, Ph.D. dissertation, Harvard University.

Theil, H. (1971), "A Simple Modification of the Two-Stage Least-Squares Procedure for Undersized Samples," Report 7107, Center for Mathematical Studies in Business and Economics, University of Chicago, February, 22 pp.

Zellner, A. and H. Theil (1962), "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica*, Vol. 30, No. 1, January, pp. 54–78.