EFFICIENT ESTIMATION OF SIMULTANEOUS EQUATIONS
WITH AUTO-REGRESSIVE ERRORS
BY INSTRUMENTAL VARIABLES

Ray C. Fair

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 129
Revised April 1972

The research described in this paper
was supported by NSF Grant GS-2799.

ECONOMETRIC RESEARCH PROGRAM
Princeton University
207 Dickinson Hall
Princeton, New Jersey

EFFICIENT ESTIMATION OF SIMULTANEOUS EQUATIONS WITH AUTO-REGRESSIVE ERRORS BY INSTRUMENTAL VARIABLES

Ray C. Fair*

Department of Economics
Princeton University

1. Introduction

The purpose of this paper is to point out how the efficient instrumental-variables technique discussed by Brundy and Jorgenson [2] can be modified to take into account auto-regressive properties of the error terms. The limited-information and full-information estimators proposed in this paper are consistent and asymptotically efficient if the auto-regressive coefficients are known with certainty. For the case in which the auto-regressive coefficients are unknown and must be estimated, the estimators are consistent, but the asymptotic efficiency of the estimators has not been established. A result by Dhrymes [5] is available, however, that is encouraging as to the possible asymptotic efficiency of the estimators for the case in which the auto-regressive coefficients are unknown.

The full-information estimation of simultaneous equations models with auto-regressive errors has been discussed by Sargan [10], Hendry [8], Chow and Fair [4], and Dhrymes [5]. Sargan originally proposed the full-information maximum likelihood estimation of such models, and Hendry and Chow and Fair have recently developed computationally feasible methods for obtaining the maximum likelihood estimates. Hendry considered only the

---

* I would like to thank Gregory C. Chow and Phoebus J. Dhrymes for helpful discussions regarding certain parts of this paper.

case of completely unrestricted auto-regressive coefficient matrices (i.e., no zero elements), whereas Chow and Fair's method can handle the case of restricted auto-regressive coefficient matrices. Dhrymes has recently proposed the three-stage least squares estimator of simultaneous equations models with auto-regressive errors. Dhrymes also considered only the case of completely unrestricted auto-regressive coefficient matrices.

The limited-information estimation of simultaneous equations models with auto-regressive errors has been discussed by Sargan [10], Amemiya [1], and Fair [7], among others. Sargan proposed the limited-information maximum likelihood estimation of such models, and Amemiya and Fair considered various two-stage least squares estimators of such models. Most of the work on limited-information estimators has been concerned with the case of diagonal auto-regressive coefficient matrices.

Brundy and Jorgenson's criticism of the two- and three-stage least squares estimators, namely that the first stage involves estimating reduced form equations with a very large number of variables included in them, holds even more so for models with auto-regressive errors. For these models, the reduced form equations include not only all of the predetermined variables in the system but also all of the lagged endogenous and lagged predetermined variables. In fact, one of the main purposes of the work in [7] was to suggest ways in which the number of variables used in the first stage regression of two-stage least squares might be decreased with perhaps small loss of asymptotic efficiency. The advantage of the instrumental-variables techniques proposed in the Brundy-Jorgenson paper and in this paper is that the first stage regressions need not be run.

2. The Model

The model to be estimated is[1]

$$(1)\qquad Y\Gamma + XB = U \quad ,$$

where Y is a n x p matrix of endogenous variables, X is a n x q matrix of predetermined variables, U is a n x p matrix of error terms, and Γ and B are p x p and q x p coefficient matrices respectively. The X matrix may include lagged endogenous variables as well as exogenous variables. n is the number of observations. As distinct from the Brundy-Jorgenson paper, it is assumed here that the error terms in U follow a m^{th} order auto-regressive process:

$$(2)\qquad U = U_{-1}R^{(1)} + \ldots + U_{-m}R^{(m)} + E \quad ,$$

where the $R^{(k)}$ matrices are p x p coefficient matrices, E is a n x p matrix of error terms, and the subscripts denote lagged values of the terms of U. Combining (1) and (2) yields

$$(3)\qquad Y\Gamma + XB = Y_{-1}\Gamma R^{(1)} + X_{-1}BR^{(1)} + \ldots + Y_{-m}\Gamma R^{(m)} + X_{-m}BR^{(m)} + E \quad .$$

From (3) the reduced form for Y is

$$(4)\qquad Y = -XB\Gamma^{-1} + Y_{-1}\Gamma R^{(1)}\Gamma^{-1} + X_{-1}BR^{(1)}\Gamma^{-1} + \ldots + Y_{-m}\Gamma R^{(m)}\Gamma^{-1} + X_{-m}BR^{(m)}\Gamma^{-1} + E\Gamma^{-1} \quad ,$$

---

[1] The notation here follows as closely as possible the notation in Brundy and Jorgenson [2].

or

(5) $$Y = Q\Pi + V \quad ,$$

where $V = E\Gamma^{-1}$, $Q = [X\ Y_{-1}\ X_{-1}\ \ldots\ Y_{-m}\ X_{-m}]$, and Π is partitioned according to Q.

It is convenient to write the structural equations in (1) in the form:

(6) $$y_j = Z_j\delta_j + u_j \quad , \quad j=1,2,\ldots,p \quad ,$$

where

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

As in Brundy and Jorgenson [2, p. 208], y_j is a vector of observations on the j^{th} column of Y whose structural coefficient has been normalized to one, Y_j is a matrix of observations on the other endogenous variables included in the equation, X_j is a matrix of observations on the predetermined variables included directly in the equation, u_j is the j^{th} column of U, and γ_j and β_j are structural coefficient vectors corresponding to Y_j and X_j respectively. The p equations in (6) can be combined to yield

(7) $$y = Z\,\delta + u \quad ,$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 & 0 & \dots & 0 \\ 0 & Z_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & Z_p \end{bmatrix}, \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{bmatrix}.$$

In order to implement the instrumental variables estimator in the auto-regressive case, it is necessary to transform (7) so that the error term on the right hand side is e rather than u, where

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix},$$

e_j being the j^{th} column of E. This transformation is:

(8) $$y - (R^{(1)'} \otimes I)y_{-1} - \dots -(R^{(m)'} \otimes I)y_{-m}$$

$$= [Z-(R^{(1)'} \otimes I)Z_{-1} - \dots -(R^{(m)'} \otimes I)Z_{-m}]\delta + e \quad ,$$

or

(9) $$y^* = Z^*\delta + e \quad ,$$

where I is an n x n identity matrix and the subscripts on y and Z denote lagged values.

3. The Full-Information Estimator

The basic idea of the Brundy-Jorgenson paper is that if a set of instrumental variables can be found that is based on a consistent estimate of Π, then using this set of instrumental variables will result in asymptotically efficient estimates (within the class of either limited-information or full-information methods). In the present case, Π in (5) is a function of the $R^{(k)}$ matrices as well as of Γ and B. Consequently, if consistent estimates of Γ, B, and the $R^{(k)}$ matrices are available, then a consistent estimate of Π in (5) is available. The equations in (5) can then be used to generate consistent predictions of the endogenous variables. Consistent estimates of Γ, B, and the $R^{(k)}$ matrices can also be used to obtain a consistent estimate of the variance-covariance matrix, Σ, of the error terms E in (3).

Assume, therefore, that initial consistent estimates of Γ, B, and the $R^{(k)}$ matrices are available[2] so that a consistent estimate of Π in (5) is available. The matrix Z consists of current endogenous variables as well as of predetermined variables. Since a consistent estimate of Π is assumed to be available, (5) can be used to generate consistent predictions of the endogenous variables in the model. Let $\hat{Z}$ denote the matrix Z except for the replacement of the current endogenous variables in Z by their predicted values from (5). Let $\bar{y}^*$ and $\bar{Z}^*$ denote the matrices y^* and Z^* respectively except for the use of consistent estimates of the $R^{(k)}$ matrices rather than the actual matrices to transform the variables. Also, let $\hat{\bar{Z}}^*$ denote the matrix $\bar{Z}^*$ except for the replacement of Z by $\hat{Z}$, and let $W = (\Sigma^{-1} \otimes I)\hat{\bar{Z}}^*$,

[2] It will be seen in Section 6 how initial consistent estimates of these matrices can be obtained.

where $\hat{\Sigma}$ is a consistent estimate of Σ. Then the "full-information instrumental variables efficient"[3] estimator in the auto-regressive case (say, FIVER) can be defined to be:

$$(10) \qquad d = (\bar{W}'\bar{Z}^*)^{-1}\bar{W}'\bar{y}^* \quad .$$

It is easy to show that the FIVER estimator is consistent. From (8) or (9) and the definition of $\bar{y}^*$ and $\bar{Z}^*$, it follows that

$$(11) \qquad \bar{y}^* = \bar{Z}^*\delta + [(R^{(1)'} - \hat{R}^{(1)'}) \otimes I + \ldots + (R^{(m)'} - \hat{R}^{(m)'}) \otimes I]u_{-1} + e \quad ,$$

where the $\hat{R}^{(k)}$ matrices are consistent estimates of the $R^{(k)}$ matrices. Substituting (11) into (10) yields:

$$(12) \qquad d = \delta + (\bar{W}'\bar{Z}^*)^{-1}\bar{W}' \; [(R^{(1)'} - \hat{R}^{(1)'}) \otimes I + \ldots + (R^{(m)'} - \hat{R}^{(m)'}) \otimes I]u_{-1}$$
$$+ (\bar{W}'\bar{Z}^*)^{-1}\bar{W}'e \quad .$$

Assuming that plim $n\,(\bar{W}'\bar{Z}^*)^{-1}$ exists and is finite, it follows that plim $d = \delta$, since plim $n^{-1}\bar{W}'e = 0$ because of the inclusion in $\bar{W}$ of only predetermined variables or linear functions of predetermined variables and since the $R^{(k)}$ matrices are consistently estimated.[4]

---

[3] Brundy and Jorgenson [2], p. 214.

[4] See Dhrymes [5], Lemma 8, for a detailed proof of the consistency of the three-stage least squares estimator of the above model. It can be shown that the formula for the three-stage least squares estimator that Dhrymes presents, formula (30), is the same as the formula in (10), except that for the three-stage least squares formula $\bar{W}$ includes predicted values from the first stage regressions rather than values generated from (5) using a consistent estimate of Π.

It is also easy to see that the FIVER estimator is asymptotically efficient if the $R^{(k)}$ matrices are known with certainty. In this case the model in (9) is merely a standard simultaneous equations model in y* and Z*. Brundy and Jorgenson have shown that asymptotic efficiency is preserved if one changes the three-stage least squares estimator by using generated predictions of the endogenous variables rather than predictions from the first stage regressions. The formula in (10) is the same as the three-stage least squares formula (30) in Dhrymes [5] (both assuming known $R^{(k)}$ matrices)[5] except for the use of generated predictions rather than predictions from the first stage regressions. Therefore, since the three-stage least squares estimator of the standard simultaneous equations model is asymptotically efficient, the FIVER estimator is asymptotically efficient in the case of known $R^{(k)}$ matrices.

The asymptotic efficiency of the FIVER estimator has not been established for the case in which the $R^{(k)}$ matrices are unknown and must be estimated along with the structural coefficients. With respect to the three-stage least squares estimator of the above model, however, Dhrymes [5] has shown that if one iterates back and forth between estimates of δ and estimates of the $R^{(k)}$ matrices and if convergence is reached, then asymptotically the set of equations that is solved by this procedure is the same set of equations that the full-information maximum likelihood estimator satisfies. This result cannot be extended directly to the FIVER estimator because, unlike three-stage least squares, the FIVER estimator is

---

[5] Dhrymes actually considered only the case of a first-order auto-regressive process, but it is easy to generalize his arguments and formulas to higher-order processes. Dhrymes also considered only the case of completely unrestricted auto-regressive coefficient matrices, but his formula (30) is also valid for the case of restricted matrices.

not based on the minimization of anything. Dhrymes' analysis is also based on the assumption of unrestricted auto-regressive coefficient matrices, an assumption that is not likely to be made in practice. Dhrymes' result is, however, at least encouraging as to the possibility that the FIVER estimator, based on iterating back and forth between estimates of δ and estimates of the $R^{(k)}$ matrices until convergence is reached, is asymptotically efficient even when the $R^{(k)}$ matrices are unknown.

The asymptotic variance-covariance matrix of d is:

$$\text{(13)} \qquad \text{asy.var-cov } d = n^{-1} \text{plim } n\ (\bar{W}'\bar{Z}^*)^{-1}\bar{W}'ee'\bar{W}(\bar{Z}^{*\prime}\bar{W})^{-1} \quad .$$

From the fact that $\hat{\Sigma}$ is a consistent estimate of Σ, that $\hat{\bar{Z}}^*$ differs from $\bar{Z}^*$ merely by the replacement of the endogenous variables in $\bar{Z}^*$ by predictions of the endogenous variables based on a consistent estimate of Π in (5), that the variance-covariance matrix of e is $\Sigma \otimes I$, and that $\bar{W} = (\hat{\Sigma}^{-1} \otimes I)\hat{\bar{Z}}^*$, it can be shown from (13) that the asymptotic variance-covariance matrix of d is $n^{-1} \text{plim } n\ (\bar{W}'\bar{Z}^*)^{-1}$. The asymptotic variance-covariance matrix of d can thus be estimated as $(\bar{W}'\bar{Z}^*)^{-1}$.

4. The Limited-Information Estimator

In this section the limited-information case will be analyzed under the assumption that the $R^{(k)}$ matrices are diagonal. A brief description of how one can estimate models with non-diagonal $R^{(k)}$ matrices by limited-information techniques is presented in Section 7.

If the $R^{(k)}$ matrices are diagonal, then equation (6) can be transformed as:

(14) $$y_j - r_{jj}^{(1)} y_{j-1} - \ldots - r_{jj}^{(m)} y_{j-m} = [Z_j - r_{jj}^{(1)} Z_{j-1} - \ldots - r_{jj}^{(m)} Z_{j-m}]\delta_j + e_j \quad , \quad j=1,2,\ldots,p \quad ,$$

or

(15) $$y_j^* = Z_j^* \delta_j + e_j \quad ,$$

where the subscripts on y_j and Z_j denote lagged values and where $r_{jj}^{(k)}$ is the j^{th} diagonal element of $R^{(k)}$ $(k=1,\ldots,m)$. Let $\hat{Z}_j$ denote the matrix Z_j except for the replacement of the current endogenous variables in Z_j by their predicted values from (5). Let $\bar{y}_j^*$ and $\bar{Z}_j^*$ denote the matrices y_j^* and Z_j^* respectively except for the use of consistent estimates of the $r_{jj}^{(k)}$ coefficients rather than the actual coefficients to transform the variables. Also, let $\bar{W}_j$ denote the matrix $\bar{Z}_j^*$ except for the replacement of Z_j by $\hat{Z}_j$. Then the "limited-information instrumental variables efficient"[6] estimator in the auto-regressive case (say, LIVER) is:

(16) $$d_j = (\bar{W}_j' \bar{Z}_j^*)^{-1} \bar{W}_j' \bar{y}_j^* \quad .$$

The discussion of the asymptotic properties of the LIVER estimator is similar to the discussion of the asymptotic properties of the FIVER estimator and need not be repeated. The LIVER estimator is consistent, and within the class of limited-information estimators, the estimator is asymptotically efficient if the $r_{jj}^{(k)}$ coefficients are known with certainty.

---

[6] Brundy and Jorgenson [2], p. 211.

The asymptotic efficiency of the estimator has not been established for the case in which the $r_{jj}^{(k)}$ coefficients are unknown. The asymptotic variance-covariance matrix of the estimator is $n^{-1}\sigma_{jj}$ plim $n(\bar{W}_j'\bar{Z}_j^*)^{-1}$, which can be estimated as $\hat{\sigma}_{jj}(\bar{W}_j'\bar{Z}_j^*)^{-1}$, where $\hat{\sigma}_{jj}$ is the j^{th} diagonal element of $\hat{\Sigma}$.

5. Estimates of the $R^{(k)}$ Matrices

Given consistent estimates of the Γ and B matrices, consistent estimates of the error matrices $U, U_{-1},\ldots,U_{-m}$ can be obtained from the current and lagged versions of (1). Let $\hat{U}$ denote any consistent estimate of U, and let $\hat{\bar{U}}$ denote any consistent estimate of $\bar{U}$, where $\bar{U} = (U_{-1}\ldots U_{-m})$. Also, let $\bar{R}' = (R^{(1)'} \ldots R^{(m)'})$ and write (2) as

$$(17) \qquad U = \bar{U}\,\bar{R} + E \quad .$$

Now, for known values of U and $\bar{U}$, (17) can be interpreted as a Zellner "seemingly unrelated regression" model. If the $\bar{R}$ matrix is completely unrestricted, then (17) is, of course, merely the standard multivariate linear regression model. Since consistent estimates of U and $\bar{U}$ are available, for the full-information case $\bar{R}$ can be estimated as

$$(18) \qquad \hat{\bar{R}} = (\hat{\bar{U}}'\hat{\Sigma}^{-1}\hat{\bar{U}})^{-1}\hat{\bar{U}}'\hat{\Sigma}^{-1}\hat{U} \qquad ,$$

where $\hat{\Sigma}$ is a consistent estimate of Σ. For the case in which $\bar{R}$ is completely unrestricted, the full-information estimator is merely $(\hat{\bar{U}}'\hat{\bar{U}})^{-1}\hat{\bar{U}}'\hat{U}$. This is the case analyzed by Hendry [8] and Dhrymes [5]. For the limited-information case, the (diagonal) elements of the $R^{(k)}$ matrices can be estimated by merely regressing each column of $\hat{U}$ on the corresponding columns of $\hat{U}_{-1},\ldots,\hat{U}_{-m}$. For the limited-information case, information about $\hat{\Sigma}$ is ignored.

Similar statements can be made about the asymptotic properties of the estimators of the $R^{(k)}$ matrices as were made about the estimators of the δ vector. In the full-information case, for example, $\hat{\bar{R}}$ is consistent and asymptotically efficient if the error matrices are known with certainty. The asymptotic efficiency of the estimator has not been established for the case in which the error matrices are unknown and must be estimated. The asymptotic variance-covariance matrix of $\hat{\bar{R}}$ can be estimated as $(\hat{\bar{U}}'\hat{\Sigma}^{-1}\hat{\bar{U}})^{-1}$, or as $(\hat{\bar{U}}'\hat{\bar{U}})^{-1}$ if $\bar{R}$ is completely unrestricted. In the limited information case, an estimate of the variance-covariance matrix of the $r_{jj}^{(k)}$ coefficient estimates is merely the estimate of the variance-covariance matrix computed from each of the least squares regressions.

The result by Dhrymes mentioned above indicates that asymptotic efficiency is likely to be gained by iterating back and forth between estimates of the $R^{(k)}$ matrices and estimates of δ, although there is no guarantee of convergence from following this procedure. For the full-information maximum likelihood case, iterating back and forth between estimates of the $R^{(k)}$ matrices and estimates of B and Γ (or δ) will result in the maximum likelihood estimates of B, Γ, and the $R^{(k)}$ matrices. To the extent that the separate maximization problems can be solved, convergence to the maximum likelihood estimates is guaranteed from iterating -- see Chow and Fair [4].

6. Obtaining Initial Consistent Estimates

There are many ways in which initial consistent estimates of Γ, B, and the $R^{(k)}$ matrices can be obtained. One general technique is as follows: Treat all lagged endogenous variables (as well as endogenous variables) as endogenous, and estimate each equation of (1) by instrumental variables ignoring the auto-regressive properties of the error terms. This will result in consistent estimates of Γ and B as long as only exogenous and lagged exogenous variables are used as instruments. Use these consistent estimates to compute consistent estimates of the residuals $U, U_{-1},\ldots,U_{-m}$. Then for each equation, regress the unlagged estimated residuals on the appropriate lagged estimated residuals. The set of lagged estimated residuals will in general include both lagged estimated residuals of the particular equation being estimated as well as lagged estimated residuals of other equations of the model. This procedure will yield consistent estimates of the $R^{(k)}$ matrices since the residuals are consistently estimated. In special cases (such as diagonal $R^{(k)}$ matrices) there are, of course, other techniques that can be used to obtain initial consistent estimates. For example, in the first-order case with a diagonal $R^{(1)}$ matrix, the technique described in [7] can be used.[7]

---

[7] Dhrymes, Berner, and Cummins [6] have also considered the estimation of the first-order auto-regressive model with a diagonal $R^{(1)}$ matrix. The estimator that they propose is similar to, but is not, a LIVER estimator. Dhrymes, Berner, and Cummins first obtain consistent estimates of Γ and B in (1) by an instrumental-variables technique treating lagged endogenous variables as endogenous. They then use these estimates in the reduced from of (1) -- ignoring the auto-regressive process of U -- and obtain a set of instrumental variables by dynamic simulation (i.e., using generated values of the lagged endogenous variables as opposed to the actual values). They also use the estimates of Γ and B to estimate U and U_{-1} and from the estimates of U and U_{-1} to obtain estimates of the diagonal elements of $R^{(1)}$ by ordinary least squares. They then use the set of instrumental variables and the estimates of the elements of $R^{(1)}$ to obtain new estimates of Γ and B. Johnston [9] has shown that the estimator is not asymptotically efficient within the class of limited-information estimators.

7. Limited-Information Estimation of Models with Non-Diagonal $R^{(k)}$ Matrices

In this section it will be shown how limited information techniques can be used to estimate models with non-diagonal $R^{(k)}$ matrices. Assume without the loss of generality that the following equation is to be estimated:

$$(19) \qquad y_j = Z_j\delta_j + u_j \quad ,$$

where

$$Z_j = [Y_j \; X_j] \; , \; \delta_j = \begin{bmatrix} \gamma_j \\ \beta_j \end{bmatrix}, \quad u_j = r_{jj}^{(1)}u_{j-1} + r_{ij}^{(1)}u_{i-1} + e_j \quad ,$$

$r_{jj}^{(1)}$ and $r_{ij}^{(1)}$ being elements of $R^{(1)}$. Equation (19) can be rewritten as

$$(20) \qquad y_j - r_{ij}^{(1)}u_{i-1} = Z_j\delta_j + r_{jj}^{(1)}u_{j-1} + e_j$$

$$= Z_j\delta_j + r_{jj}^{(1)}y_{j-1} - r_{jj}^{(1)}Z_{j-1}\delta_j + e_j$$

or

$$(21) \qquad y_j^* = Z_j^*\delta_j + e_j \quad ,$$

where

$$y_j^* = y_j - r_{ij}^{(1)}u_{i-1} - r_{jj}^{(1)}y_{j-1} \; , \quad Z^* = Z_j - r_{jj}^{(1)}Z_{j-1} \quad .$$

Equation (21) is in a form like (15) except for the inclusion of the $-r_{ij}^{(1)}u_{i-1}$ term in y_j^*. If consistent estimates of r_{ij} and the residual vector u_{i-1} are available, however, then the estimation of (21) by the LIVER technique can procede like the estimation of (15). All that has been done is the subtraction of a consistent estimate of $r_{ij}^{(1)}u_{i-1}$ from y_j^* in (15).

8. Conclusion

One of the main advantages of the estimators proposed in this paper is that first-stage, reduced-form regressions do not have to be run. For the single-equation case, a disadvantage is that an entire model must be specified and consistently estimated in order to obtain efficient estimates of any single equation. In at least some practical applications this may be a serious disadvantage, and for these cases one might wish to resort to a less efficient estimator like the two-stage least squares estimator proposed in [7], which does not require the specification and estimation of the entire model.

Since in the full-information case it is now feasible to estimate models with auto-regressive errors by the maximum likelihood method, it might be desirable to attempt to estimate a model by full-information maximum likelihood (FIML) before resorting to the FIVER estimator. As discussed in Chow [3], there are some methodological reasons for preferring the FIML estimator over other asymptotically efficient estimators. It is, however, likely that the FIVER estimator will be able to handle larger models than the FIML estimator can.

REFERENCES

[1] Amemiya, Takeshi: "Specification Analysis in the Estimation of Parameters of a Simultaneous Equation Model with Autoregressive Residuals", Econometrica, XXXIV (April, 1966), 283-306.

[2] Brundy, James M. and Dale W. Jorgenson: "Efficient Estimation of Simultaneous Equations by Instrumental Variables", The Review of Economics and Statistics, LIII (August, 1971), 207-224.

[3] Chow, Gregory C.: "A Comparison of Alternative Estimators for Simultaneous Equations", Econometrica, XXXII (October, 1964), 532-553.

[4] Chow, Gregory C. and Ray C. Fair: "Maximum Likelihood Estimation of Linear Equation Systems with Auto-Regressive Residuals", Research Memorandum No. 118, Econometric Research Program, Princeton University, revised September, 1971.

[5] Dhrymes, Phoebus J.: "Full Information Estimation of Dynamic Simultaneous Equations Models with Autoregressive Errors", mimeographed.

[6] Dhrymes, P. J., R. Berner, and R. Cummins: "Limited Information Estimation of Simultaneous Equations Models with Lagged Endogenous Variables and Autoregressive Errors", Discussion Paper No. 183, The Wharton School of Finance and Commerce, University of Pennsylvania, October 1970.

[7] Fair, Ray C.: "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors", Econometrica, XXXVIII (May, 1970), 507-516.

[8] Hendry, D. F.: "Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with Errors Generated by a Vector Autoregressive Process", International Economic Review, XII (June, 1971).

[9] Johnston, H. N.: Ph.D. Thesis, Department of Economics, University of Pennsylvania.

[10] Sargan, J. D.: "The Maximum Likelihood Estimation of Economic Relationships with Autoregressive Residuals", Econometrica, XXIX (July, 1961), 414-426.