



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MONASH

WP 3/95

ISSN 1032-3813
ISBN 0 7326 0764 7

MONASH UNIVERSITY

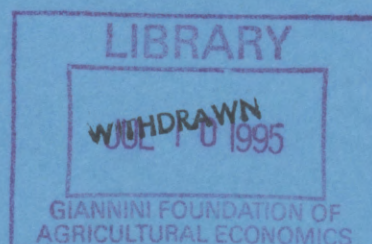


AUSTRALIA

**A PARSIMONIOUS AUTOCORRELATION CORRECTION
FOR SINGULAR DEMAND SYSTEMS**

Keith R. McLaren

**Working Paper 3/95
March 1995**



DEPARTMENT OF ECONOMETRICS

A PARSIMONIOUS AUTOCORRELATION CORRECTION FOR SINGULAR DEMAND SYSTEMS

Keith R. McLaren
Department of Econometrics
Monash University
Clayton, Vic. 3168
Australia

Email : Keith.McLaren@BusEco.Monash.Edu.Au

Abstract : The adding up condition of budget share equations is known to imply restrictions for the autoregressive structure of errors. The implications of these restrictions when estimation is in terms of additive normal errors or additive logistic normal errors is clarified, and a byproduct is a specification of the autocorrelation matrix with a structure consistent with the model, but with number of parameters equal to the number of goods. This is more appealing than the scalar diagonal matrix form, but more parsimonious than having number of parameters proportional to the square of the number of goods.

JEL Classification: C30

Keywords : Demand systems ; Singular systems ; Autoregressive errors.

Third Draft
March 7 1995

A PARSIMONIOUS AUTOCORRELATION CORRECTION FOR SINGULAR DEMAND SYSTEMS

Keith R. McLaren

1. Introduction

The deterministic specification of a system of budget share equations will satisfy a number of restrictions derived from economic theory. Similarly, the stochastic specification must also satisfy a number of restrictions. Two approaches appear to have emerged to handle these restrictions on the stochastic component in estimation. In what may be referred to as the "traditional" or additive errors approach, the stochastic term is added to the deterministic specification of the shares. Adding up then implies that these errors add identically to zero across equations, and this singularity is accommodated in estimation by deleting one of the equations. Barten (1969) and Powell (1969) have shown the conditions under which parameter estimates are invariant to the equation deleted. This adding up property was shown by Berndt and Savin (1975) to imply strong restrictions on any autocorrelation structure of a set of additive errors. In applications this has been taken to imply a choice between a diagonal system with a common scalar autoregressive parameter, or a "full" autoregressive system with number of parameters equal to the square of the number of independent equations. In the second approach, which may be referred to as the compositional data analysis (CODA) approach, (see, for example, Aitchison (1986)), or more usually in econometrics as the logit approach, the shares are transformed to logratios before the stochastic error is introduced, in order for the normality assumption to be consistent with the requirement that budget shares be constrained to the unit interval. In this approach singularity of the errors is implicit, because the system is usually specified by normalizing on a particular share. For this alternative stochastic specification of share equations, Chavas and Segerson (1986) have noted that such a specification has different implications for autocorrelation structure than the Berndt and Savin result for the case of additive errors.

This paper attempts to unify the results of these two approaches, and provide a parsimonious and easily estimated autocorrelation structure even for "standard" singular systems. In Section 2 the normalisation implicit in the CODA approach is applied in the additive errors framework, and it is shown how this normalization allows the specification of a non-scalar diagonal autoregressive structure for the transformed model. In Section 3 the errors of the CODA approach are specified in a way analogous to that used in the additive errors approach. The similarities are then clarified, and the Chavas-Segerson result is shown to be due to the particular interpretation of the set of unidentified stochastic variables. The issue of invariance is discussed in Section 4, and its meaning in these cases clarified. Finally, Section 5 proposes a general but parsimonious autocorrelation structure for such models.

2. Additive Errors Form

Consider the system of equations with the "traditional" stochastic specification

$$(2.1) \quad w_t = W(x_t, \beta) + u_t$$

where w_t is an $N \times 1$ vector of observed shares, W is an $N \times 1$ vector of share equations which are functions of the data vector x_t and parameter vector β , and u_t is an $N \times 1$ vector of multivariate normal errors, with mean zero and variance-covariance matrix Σ . Assuming that the w_t and W_t satisfy the adding up conditions $\mathbf{1}'w_t = 1$ and $\mathbf{1}'W_t = 1$, where $\mathbf{1}$ is the vector of ones, then it is well known that $\mathbf{1}'u_t = 0$ and Σ is singular of rank $n = N-1$. In addition, if the error term is autoregressive

$$(2.2) \quad u_t = Ru_{t-1} + e_t$$

then Berndt and Savin(1975) have shown that the adding up conditions imply that

$$\mathbf{1}'e_t = 0$$

and

$$\mathbf{1}'R = k\mathbf{1}'$$

where k is a scalar constant, implying that the columns of R sum to the same constant. In practice, this restriction has usually been taken to imply either of the following two cases :

(a) R is diagonal, with only one free parameter, the common diagonal element k , or

(b) R is "full", in which case R has $nN+1$ free parameters, of which only n^2 parameters are identified.

It would be convenient for empirical work to have available a parsimonious specification in which R had a number of free parameters of the order of the number of equations, n , and to have available a simple estimation method. Such an implied specification can be motivated and derived as follows.

In order to estimate (2.1) it is common to delete an equation, say the last. This corresponds to premultiplying (2.1) by the known $n \times N$ matrix

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

Following McLaren(1990) it is more transparent to consider the non-singular transformation induced by premultiplying by the (nonsingular) $N \times N$ matrix

$$B = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

which is A converted to a square matrix by the addition of a row of ones. This replaces the last (degenerate) stochastic equation by an identity, and makes transparent both the equivalence to (2.1) of the resulting specification, (being related by a nonsingular, known, transformation), and the legitimacy of deleting an arbitrary equation (for this particular transformation, the N^{th} , which has been mapped into an identity) for the purposes of estimation. All that is really happening with this transformation is that one particular singular N -dimensional representation of a nonsingular n -dimensional error process is being replaced by another.

Consider now the alternative transformation, motivated by the CODA approach, of normalizing on the N^{th} share, which can be generated by premultiplying by the $n \times N$ matrix (in Aitchison's notation)

$$F = \begin{bmatrix} 1 & 0 & 0 & \dots & -1 \\ 0 & 1 & 0 & \dots & -1 \\ & & \dots & & \\ 0 & 0 & & 1 & -1 \end{bmatrix}$$

i.e. each of the w_i is replaced by $w_i - w_N$, $i=1, \dots, n$, and each of the u_i is replaced by $u_i - u_N$, to give the model

$$(2.3) \quad w_i - w_N = W_i - W_N + u_i - u_N \quad (i = 1, \dots, n).$$

Again, this is merely an alternative representation of the error process, and is made transparent if the transformation is in terms of the nonsingular $N \times N$ matrix

$$C = \begin{bmatrix} 1 & 0 & \dots & 0 & -1 \\ 0 & 1 & \dots & 0 & -1 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

which preserves the last equation of (2.1) in the system. The system can then be represented in terms of the transformed variables $w_i - w_N$, the functions $W_i - W_N$, and the errors $u_i - u_N$ by the "structural" system :

$$(2.4) \quad \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & N \end{bmatrix} \begin{bmatrix} w_1 - w_N \\ w_2 - w_N \\ \dots \\ w_n - w_N \\ w_N \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & N \end{bmatrix} \begin{bmatrix} W_1 - W_N \\ W_2 - W_N \\ \dots \\ W_n - W_N \\ W_N \end{bmatrix} + \begin{bmatrix} u_1 - u_N \\ u_2 - u_N \\ \dots \\ u_n - u_N \\ 0 \end{bmatrix}$$

Defining D to be the matrix premultiplying the system above, the role of D has been to annihilate the last error from the system. Thus it is clear that the last equation can be deleted, and that there is a known nonsingular transformation back to (2.1), assuring invariance of parameter estimates based on maximum likelihood. Thus a nonsingular n dimensional system equivalent to deleting the last equation from (2.1) is obtained by successively premultiplying (2.1) by C , then D , and then deleting the last (clearly redundant) equation. When estimated in this form, the resulting system can be embedded in the above structural system, and mapped directly back to (2.1) in a one-to-one manner.

To motivate the type of structure this allows on the autocorrelation term, consider a simple example with $N=3$, and how the transformation of variables can be accomplished. Then premultiplying (2.2) by F gives :

$$\begin{bmatrix} u_{1t} - u_{3t} \\ u_{2t} - u_{3t} \end{bmatrix} = \begin{bmatrix} R_{11} - R_{31} & R_{12} - R_{32} & R_{13} - R_{33} \\ R_{21} - R_{31} & R_{22} - R_{32} & R_{23} - R_{33} \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} - e_{3t} \\ e_{2t} - e_{3t} \end{bmatrix}.$$

This can be written as a difference equation in the transformed variables $u_i - u_N$ as

$$(2.5) \quad \begin{bmatrix} u_{1t} - u_{3t} \\ u_{2t} - u_{3t} \end{bmatrix} = \begin{bmatrix} R_{11} - R_{31} & R_{12} - R_{32} \\ R_{21} - R_{31} & R_{22} - R_{32} \end{bmatrix} \begin{bmatrix} u_{1t-1} - u_{3t-1} \\ u_{2t-1} - u_{3t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} - e_{3t} \\ e_{2t} - e_{3t} \end{bmatrix}$$

provided the following restrictions are satisfied:

$$\begin{aligned} R_{13} - R_{33} + R_{11} - R_{31} + R_{12} - R_{32} &= 0 \\ R_{23} - R_{33} + R_{21} - R_{31} + R_{22} - R_{32} &= 0. \end{aligned}$$

Now if the coefficient matrix in (2.5) is specified to be diagonal, the adding up restrictions do not require that these coefficients be equal, i.e. there are 2 ($=n$) free coefficients. If these two free coefficients are denoted a and b , then the elements of the original R matrix are further restricted as

$$\begin{aligned} R_{11} - R_{31} &= a \\ R_{12} - R_{32} &= 0 \\ R_{21} - R_{31} &= 0 \\ R_{22} - R_{32} &= b. \end{aligned}$$

By the results of Berndt and Savin, these restrictions are not sufficient to identify all of the elements of the R matrix. In particular, the scalar k is not identified, in common

with case (b) above. Setting k to zero for simplicity, the implied structure of the R matrix is

$$R = \begin{bmatrix} \frac{2a}{3} & \frac{-b}{3} & \frac{a+b}{3} & -a \\ -a & \frac{2b}{3} & \frac{a+b}{3} & -b \\ \frac{-a}{3} & \frac{-b}{3} & \frac{a+b}{3} & \\ \frac{-a}{3} & \frac{-b}{3} & \frac{a+b}{3} & \end{bmatrix}$$

(In general, $k/3$ can be added to each of the elements of R .)

For arbitrary N , this structure for R generalizes to

$$(2.6) \quad R = \left[\frac{1}{N} \begin{bmatrix} na_1 & -a_2 & -a_3 & \dots & \sum_{i=1}^n a_i - Na_1 \\ -a_1 & na_2 & -a_3 & \dots & \sum_{i=1}^n a_i - Na_2 \\ -a_1 & -a_2 & na_3 & \dots & \sum_{i=1}^n a_i - Na_3 \\ \dots & \dots & \dots & \dots & \dots \\ -a_1 & -a_2 & -a_3 & \dots & \sum_{i=1}^n a_i \end{bmatrix} \right]$$

Estimation is simple, being in terms of the transformed variables $w_i - w_N$ and the transformed deterministic components $W_i(z, \beta) - W_N(z, \beta)$, and lagged values of these for each i .

In general this structure can be represented by transforming (2.1) by premultiplying through by

$$E = DC = \begin{bmatrix} I & -1 \\ 1' & 1 \end{bmatrix}$$

to give

$$Ew_t = EW_t + Eu_t$$

where

$$\begin{aligned} Eu_t &= ERu_{t-1} + Ee_t \\ &= ERE^{-1}Eu_{t-1} + Ee_t \end{aligned}$$

or, selecting the first n rows (since D annihilates the N th error term),

$$v_t = Sv_{t-1} + f_t$$

where

$S =$ the $n \times n$ upper left block of ERE^{-1}

$$v_i = \begin{bmatrix} u_1 - u_N \\ u_2 - u_N \\ \dots \\ u_{N-1} - u_N \end{bmatrix} \quad \text{and} \quad f_i = \begin{bmatrix} e_1 - e_N \\ e_2 - e_N \\ \dots \\ e_{N-1} - e_N \end{bmatrix}$$

If R has the structure (2.6), then S will be diagonal with unequal diagonal terms.

3 Logratio Form

It can be argued that a more appropriate statistical model for shares is the logratio form

$$(3.1) \quad y_i = \log(w_i / w_N) = \log(W_i / W_N) + v_i \quad (i = 1, \dots, n)$$

with v modelled as multivariate normal. For a discussion of this specification, see Fry, Fry and McLaren (1993), Aitchison (1986), Considine and Mount (1984)). In fact Chavas and Segerson (1986) argue that the errors in such a system can be more freely specified than the errors in (2.1), and relate these restrictions to what they call the "structural" model

$$(3.2) \quad w_i = \frac{E_i(z, \beta) \exp(u_i^*)}{\sum_{j=1}^N E_j(z, \beta) \exp(u_j^*)}$$

for which the errors u_i^* are unconstrained, and where the E_i indicate expenditures. The analogy with the results of Section 2 is, however, more transparent if (3.1) is compared with the *centered* logratio model

$$(3.3) \quad z_i = \log(w_i / \tilde{w}) = \log(W_i / \tilde{W}) + u_i, \quad (i = 1, \dots, N)$$

where \tilde{w} and \tilde{W} are the geometric means across elements i of the w and W . (This structure is provided in Aitchison (1986) pp. 78-79). In fact the u in this model are completely analogous in their properties to the errors in (2.1), the only difference being that the data and the deterministic terms add identically to zero rather than unity, and the v_i of (3.1) are constructed as $u_i - u_N$ from the errors of (3.3), so are analogous to the errors of (2.3). In particular, there is from the previous section a known one-to-one mapping between the system consisting of (3.1) and an appended identity and (3.3), in exactly the same way as the mapping between (2.1) and (2.4) can be carried out. (Aitchison (1986) notes (in a problem on p. 89) a similar, but not one-to-one, mapping between the logratios and the centered logratios : $y = Fz$ and

$$z = F'(FF')^{-1} y).$$

In terms of the errors of (3.2), note that if the numerator and denominator in (3.2) are multiplied by $\exp(\hat{u})$, where \hat{u} may be an independent error, (or indeed, \hat{u} could be any of the u_i) then the resulting model is observationally equivalent to (3.2). The errors of (3.2) are hence of full rank N , and while it may be of interest to note that the degeneracy comes from the ratio form of (3.2), the presence of this extra error term leads to an essential nonidentifiability. But this independent error disappears when differences are formed, so the v_i of (3.1) can just as well be identified either with $u_i - u_N$ or with $u_i^* - u_N^*$. Thus it can be argued that the apparent generality of the form (3.1) argued by Chavas and Segerson is in fact an artifact introduced by the relationship to the non-identified error term \hat{u} , and specification of autocorrelation structure in terms of this unidentified error. Comparison instead with the centered logratio form (3.3) clarifies the structure of the autocorrelation properties, and shows that they are perfectly analogous to the Berndt and Savin results.

4. Invariance

The expression "invariance to deleted equation" seems to have at least two interpretations in the literature, and the parameterization of the R matrix considered above is invariant in terms of one of these, but not the other. The first interpretation is most easily understood by its implicit use in Berndt and Savin, and essentially requires that it be possible to estimate a transformed model in such a way as to preserve parameter estimates and likelihood values. Invariance in this sense depends critically on the imposition of the condition that the R matrix have equal column sums. In this case, if S is parameterized as a diagonal matrix, then the model that results from normalizing on other than the N th equation would have to be estimated with the correspondingly transformed (nondiagonal) S matrix. In fact the R matrix structure considered above is quite analogous (in implications, not in structure) to the examples given by Berndt and Savin (compare their Models 2, 3, 4 and 5 on p.949 for which it is stated that parameter estimates are invariant to the equation deleted).

A second, and probably more appealing, interpretation of the concept of parameter invariance is most clearly explained in Aitchison (1986), p.95-96. Define a matrix P to be a permutation matrix if it is constructed by interchanging the columns of an identity matrix. Then applying P to a set of budget shares amounts to a reordering of these shares in the same way that the identity matrix was reordered to give P .

Define

$$w_p = Pw.$$

Statistical procedures are then said to be invariant if they are invariant to the application of such permutations, i.e. if the same parameter estimates and likelihood values result if the procedure is applied to w or w_p . Now if a particular structure such as the diagonality of a matrix is considered part of a statistical procedure, as distinct from a property of the particular element, then the question of such invariance is nontrivial. In general, if S is diagonal for w , it will not be diagonal for w_p . From (2.2)

$$Pu_i = PRP^{-1}Pu_{i-1} + Pe_i.$$

Then

$$EPu_t = EPRP^{-1}E^{-1}EPu_{t-1} + EPe_t$$

It can be shown by counter-example that ERE^{-1} having the top left $n \times n$ matrix diagonal does not ensure that $EPRP^{-1}E^{-1}$ has the same structure. Thus if the diagonal structure of the S matrix is considered part of the statistical model, there are N such models depending on the normalizing share. The problem with specification (2.6) is the asymmetric treatment of the N^{th} column of R .

5. A Fully Invariant Specification

Diagonality of the R matrix is analogous to diagonality of the variance-covariance matrix. Because of adding up, the variance-covariance matrix cannot be diagonal, and hence diagonality cannot be equated with independence. Similarly, adding up ensures certain structure on R , and diagonality of R is not the neutral case. Consider now the following specification of R , which unlike that of Section 2, treats the N^{th} column in a symmetric way. For N parameters a_i , let $a = \sum_{i=1}^N a_i$.

Define

$$(5.1) \quad R = \begin{bmatrix} na_1 & a - a_2 & \dots & a - a_N \\ a - a_1 & na_2 & \dots & a - a_N \\ & & \dots & \\ a - a_1 & a - a_2 & \dots & na_N \end{bmatrix}$$

Such a structure for R is clearly invariant to permutations, and contains N free parameters, of the same order as the number of equations. By the reasoning above this structure could be applied either to the additive error structure, (2.1), or to the centered logratio form, (3.3). While there is no obvious transformation to make estimation simple, as with the structure in Section 2, estimation by constrained maximum likelihood would be straightforward.

The structure for R in (5.1) is clearly inconsistent with the result from Chavas and Segerson (1986) that the rows of R should sum to the same constant. Two points can be made to clarify this relationship. First, the structure in (5.1) applies to the stochastic process u_t which has dimension N but rank n , not to u_t^* which has dimension *and* rank N . It is this degeneracy that has been used in Section 2 in setting up the transformation from u_t to the process v_t that has the same rank. Secondly, the Chavas and Segerson result is based on the relationship in their equation (20a), which in the notation of this paper is

$$(5.2) \quad SF = FR.$$

This relation implies nN linear equations among the n^2 elements of S , and the N^2 elements of R . Clearly for any given S there are many possible solutions for R , and for arbitrary R there is not necessarily a solution for S (this relates to the discussion on identification in Berndt and Savin (1975)), but if an unrestricted (rank N) process

$$u_t = Ru_{t-1} + e_t$$

is to be transformed to the rank n process

$$v_t = Sv_{t-1} + f_t$$

then rank reducing restrictions will have to apply in some form. In general, apart from the fact that the particular structure of F implies that any solution R of (5.2) is determined only up to an additive constant, any solution to (5.2) can be written in the form

$$R = F^\oplus SF$$

where F^\oplus is a generalized inverse of F (one possibility is $F'(FF')^{-1}$). But this implies that $R1 = 0$, since $F1 = 0$, which is effectively a rank reducing transformation, and another interpretation of Lemma 1 of Chavas and Segerson. With this reduced rank structure for R it is then clear that Ru_{t-1} can be rewritten as R_1v_{t-1} where R_1 is the first n columns of R . Thus the implied structure for R is really just another aspect of the non-identifiability of the u^* , the rank N error process.

6. Conclusion

It has been argued that the apparent difference in implications of an autocorrelation structure applied to the additive errors model or to the logratio model is more a matter of interpretation than substance. As a byproduct, two alternative parsimonious specifications of the autocorrelation structure in such models have been suggested, which allow the number of autocorrelation parameters to be proportional to the number of equations. One of these specifications satisfies only a limited definition of invariance, but the other is fully invariant.

References

- Aitchison, J. (1986), *The Statistical Analysis of Compositional Data*, Chapman and Hall, New York.
- Barten, A. (1969), "Maximum Likelihood Estimation of a Complete System of Demand Equations", *European Economic Review*, 1, 7-73.
- Berndt, E. and Savin, N. (1975), "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances", *Econometrica*, 43/5-6, 937-957.

- Chavas, J-P. and Segerson, K. (1986), "Singularity and Autoregressive Disturbances in Linear Logit Models", *Journal of Business and Economic Statistics*, 4/2, 161-169.
- Considine, T. J. and Mount, T. D. (1984), "The Use of Linear Logit Models for Dynamic Input Demand Systems", *Review of Economics and Statistics*, 66, 434-443.
- Fry, J. M., Fry, T. R. L. and McLaren, K. R. (1993), "The Stochastic Specification of Demand Share Equations: Restricting Budget Shares to the Unit Simplex", Working Paper 3/93, Department of Econometrics, Monash University. Forthcoming in *Journal of Econometrics*.
- McLaren, K. R. (1990), "A Variant on the Arguments for the Invariance of Estimators in a Singular System of Equations", *Econometric Reviews*, 9/1,91-102.
- Powell, A. (1969), "Aitken Estimators as a Tool in Allocating Predetermined Aggregates", *Journal of the American Statistical Association*, 64, 913-922.

