1978

A Polynomial Distributed Lag Model with Stochastic Coefficients

Aman Ullah

Baldev Raj

RESEARCH REPORT 7825

A POLYNOMIAL DISTRIBUTED LAG MODEL WITH STOCHASTIC COEFFICIENTS

by

Aman Ullah
and
Baldev Raj

August, 1978

A POLYNOMIAL DISTRIBUTED LAG MODEL WITH STOCHASTIC COEFFICIENTS

by

Aman Ullah and Baldev Raj*

University of Western Ontario and Wilfrid Laurier University

Abstract

This paper considers the problem of estimating a polynomial distributed lag model under a more general stochastic specification on the coefficients than considered earlier by Shiller. An estimator, namely the Generalized Random Shiller (GRS) estimator, has been proposed. It has been shown that the Shiller, Almon and Bayesian Almon estimators, among others, can be considered as special cases of the GRS estimator. Finally, on the basis of the numerical estimates, it has been suggested that the GRS estimator is to be preferred as compared to the other estimators when the lag distribution is not expected to be smooth.

Contents

*The first authors wishes to thank the Econometric Institute, Erasmus University, for their research support during the summer of 1978. This work was also in part supported by a grant from the Canada Council. We are grateful to Brian Bentley and R.A.L. Carter for valuable suggestions on an earlier version of this paper.

A POLYNOMIAL DISTRIBUTED LAG MODEL WITH STOCHASTIC COEFFICIENTS

by

Aman Ullah and Baldev Raj

1. INTRODUCTION

A popular method of estimating the coefficients of a finite distributed lag model is the polynomial lag method proposed by Almon [1965]. This method assumes that the successive coefficients in the lag distribution lie on a low-degree polynomial. Further, it depends on fewer parameters and produces very plausible lag patterns that are often very appealing. However, it can be argued that the polynomial lag structure imposes strong constraints on the lag distribution even though no prior information exists to justify them. Shiller [1973] has argued that we specify this restriction not because we believe in it but because we believe that the lag distribution is smooth. He suggested a stochastic specification for the low-degree polynomial which implied a heteroscedastic covariance structure for the error term. He circumvented this problem by transforming his basic specification before introducing the additive error. However, it makes a difference whether we introduce the error term to the polynomial lag specification of Almon or to the transformed polynomial lag specification of Shiller (see e.g., Maddala [1977, pp. 384]). It has been shown that the Shiller method produces a special type of ridge estimator. Lindley and Smith [1972] have proposed a Bayesian estimator assuming a stochastic zero-degree polynomial for the lag distribution. Their estimator has been generalized for any-degree polynomial lag structure by Maddala [1977] and has been named as the Bayes Almon estimator. This estimator also reduces to a special type of ridge estimator. Like all ridge estimators the Shiller and Bayes Almon methods require estimation of the ridge coefficient for which Shiller has suggested a rule of thumb while Lindley and Smith have suggested an iterative method.

In this paper we have analyzed the polynomial distributed lag model when the coefficients of a lag distribution are assumed to be stochastic. A polynomial distributed lag model with fixed coefficients imposes strong constraints on the specification in the absence of any prior information to justify it. An alternative parametric specification in distributed lag models may be that parameters of distributed lags have both systematic and stochastic components. We also consider the case where the systematic component itself is stochastic. Thus our parametric specification is weaker than the 'weak' stochastic specification of Shiller. It is important to note that the stochastic coefficient approach has a similarity, even though superficial, with the analysis of a fixed coefficient model from the Bayesian viewpoint. In the former approach the distribution of coefficient is introduced as part of data-generating process while in the latter the distribution of coefficient is an expression of prior beliefs of the investigator.

In Section 2 of this paper the model and the parametric specifications are described. Section 3 presents some new estimators of the parameters in the model. Then a numerical example is given in Section 4. Finally, some conclusions and suggestions for further research are presented in Section 5.

2. SPECIFICATION OF A POLYNOMIAL DISTRIBUTED LAG MODEL WITH STOCHASTIC COEFFICIENTS

2.1 The Model and Its Assumptions

Let us write the finite distributed lag model with stochastic coefficients as

$$(2.1) \quad y_t = \sum_{i=0}^{p} \beta_{it} x_{t-i} + u_t, \qquad t = 1, \ldots, T$$

where y_t is the t^{th} observation on the dependent variable, x_{t-i} is the t^{th} observation of the i period lag value of x and u_t is the usual disturbance term in the equation corresponding to the t^{th} observation.[1] Further, β_{it} is a stochastically varying coefficient of the lag structure such that

$$(2.2) \quad \beta_{it} = \bar{\beta}_i + \epsilon_{it}$$

where ϵ_{it} is the unobservable disturbance term and considering η_i to be another random error we specify $\bar{\beta}_i$ as

$$(2.3) \quad \bar{\beta}_i = \delta_o + \delta_1 i + \ldots + \delta_r i^r + \eta_i \qquad i = o, \ldots, p.$$

If both ϵ_{it} and η_i are constant (could be zero) for all i and t then (2.2) and (2.3) reduce to an Almon specification. Further, if ϵ_{it} is constant but η_i stochastic, (2.2) and (2.3) reduce to the Shiller type stochastic specification. Thus (2.2) and (2.3) imply a class of weak specifications in the distributed lag models.

We make the following simplifying assumption:

Assumption 1:

(i) $Eu_t = E\,\epsilon_{it} = E\,\eta_i = 0$ for all t and i.

(ii) $Eu_t\,\epsilon_{it} = Eu_t\,\eta_i = E\,\eta_i\,\epsilon_{it} = 0$ for all t and i.

---

[1]The distribution term u_t in model (2.1) has been taken into account so that a comparison can be made with the conventional distributed lag model.

(iii) $E\,\eta_i\,\eta_j = \sigma_\eta^2$ if $i = j$

$= 0$ otherwise

(iv) $Eu_t\,u_{t'} = \sigma_u^2$ if $t = t'$

$= 0$ otherwise

(v) $E\,\epsilon_{it}\,\epsilon_{i't'} = \sigma_{ii}$ if $t = t'$ and $i = i'$

$= 0$ otherwise[2]

(vi) The x_{t-i} is an exogenous variable distributed independently of ϵ_{it}, η_i and u_t for all t and i.

Substituting (2.2) in (2.1) we get

$$(2.4)\quad y_t = \sum_{i=o}^{p} \bar{\beta}_i\, x_{t-i} + w_t$$

where

$$(2.5)\quad w_t = u_t + \sum_{i=o}^{p} \epsilon_{it}\, x_{t-i}$$

such that by Assumption 1, $E\,w_t = 0$ for all t and

(2.5a) $E\,w_t w_{t'} = \omega_{tt}$ if $t = t'$

$= 0$ otherwise;

the ω_{tt} is given by

$$(2.5b)\quad \omega_{tt} = \sigma_u^2 + \sum_{i=o}^{p} x_{t-i}^2\, \sigma_{ii} > 0.$$

The model (2.4) and restriction (2.3) can be written, respectively, in a compact form as

---

[2]The analysis of this paper can be extended to the case where $E\,\epsilon_{it}\,\epsilon_{i't}$ is non-zero for all i and i'. We do not consider in this paper for the sake of simplicity in exposition.

(2.6) $y = X\bar{\beta} + w$

and

(2.7) $\bar{\beta} = A\delta + \eta$

where

$$(2.8)\quad y = \begin{bmatrix} y_{p+1} \\ \vdots \\ y_T \end{bmatrix}, \quad \bar{\beta} = \begin{bmatrix} \bar{\beta}_o \\ \vdots \\ \bar{\beta}_p \end{bmatrix}, \quad w = \begin{bmatrix} w_{p+1} \\ \vdots \\ w_T \end{bmatrix}, \quad \delta = \begin{bmatrix} \delta_o \\ \vdots \\ \delta_r \end{bmatrix}, \quad \eta = \begin{bmatrix} \eta_o \\ \vdots \\ \eta_p \end{bmatrix}$$

and

$$(2.9)\quad X = \begin{bmatrix} x_{p+1} & \dots x_1 \\ x_{p+2} & \dots x_2 \\ \vdots & \vdots \\ x_T & \dots x_{T-p} \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 \\ \vdots & & & \\ 1 & p & \dots & p^r \end{bmatrix}.$$

We observe that y and w are vectors of order (T-p)x1; η is a (p+1)x1 stochastic vector, and $\bar{\beta}$ and δ are unknown vectors of order (p+1)x1 and (r+1)x1, respectively. Further, X and A are (T-p) x (p+1) and (p+1) x (r+1) matrices of known values, respectively.

The rank of the matrix X is assumed to be (p+1). Further it follows from (2.5a) and (2.5b) that Ew = 0 and

$$(2.10)\quad E\,ww' = \begin{bmatrix} \omega_{p+1,\,p+1} & & \dots 0 \\ \vdots & \ddots & \\ 0 & & \omega_{TT} \end{bmatrix} = \Omega$$

where Ω is a positive definite matrix of order (T-p) x (T-p). It is noted that if $\sigma_{oo} = \dots = \sigma_{pp} = 0$ and $\sigma_\eta^2 = 0$, then model (2.6), along with (2.7), reduces to the Almon [1965] polynomial distributed lag model with fixed coefficients. In this case the covariance matrix (2.10) reduces to a scalar times an identity matrix, i.e., $\sigma_u^2\, I$ and (2.7) becomes

(2.11) $\bar{\beta} = A\delta.$

In the following section, using (2.10), we discuss various methods of estimating the model (2.6) both under the stochastic specification of $\bar{\beta}$ in (2.7) and nonstochastic specification in (2.11).

3. METHODS OF ESTIMATION

3.1 Estimation of $\bar{\beta}$ when Ω is Known

The model (2.6), assuming Ω is known, may be estimated by applying Aitken's generalized least squares (GLS). However, in order to avoid the problems of multicollinearity and degrees of freedom, some other methods of estimation discussed below may be used. We note that while the estimators in I do not incorporate (2.7) or (2.11), the estimators in II are under the restriction (2.11) and those in III are under the restriction (2.7).

I. Random Ridge Estimator

A simple solution to the dual problems of multicollinearity and heteroscedasticity in model (2.6) requires that we minimize $(y - X\bar{\beta})' \Omega^{-1} (y - X\bar{\beta})$ subject to the condition $\bar{\beta}'\bar{\beta} = r^2 \leq \infty$. The value of $\bar{\beta}$ so obtained is given by

(3.1) $$\bar{b}_{RR} = (X' \Omega^{-1} + \mu I)^{-1} X' \Omega^{-1} y$$

and its variance-covariance matrix is given as

(3.2) $$V(\bar{b}_{RR}) = (X' \Omega^{-1} X + \mu I)^{-1} X' \Omega^{-1} X (X' \Omega^{-1} X + \mu I)^{-1}.$$

The estimator $\bar{b}_{RR}$, which may be termed as the Random Ridge (RR) estimator, is a generalization of the ordinary ridge estimator developed by Hoerl and Kennard [1970]. We note that this estimator exists even when the rank of X is less than p + 1. We may now write (3.1) as

(3.3) $$\bar{b}_{RR} = a^* \bar{b}_{GLS}$$

where

$$a^* = [I + \mu(X' \, \Omega^{-1} \, X)^{-1}]^{-1}$$

and $\bar{b}_{GLS}$ is the generalized least squares (GLS) estimator of $\bar{\beta}$ in (2.6), i.e.,

(3.4) $\bar{b}_{GLS} = (X' \, \Omega^{-1} \, X)^{-1} \, X' \, \Omega^{-1} \, y.$

It is clear from (3.3) that $\bar{b}_{RR}$ is a "matrix" weighted average of the null and $\bar{b}_{GLS}$ vectors. The correction factor a*, which is different for different elements, shrinks the estimated value of $\bar{\beta}$ a fixed percentage away from $\bar{b}_{GLS}$ towards the null vector. It can be seen that the Random Ridge estimator is biased. Further, there exists some $\mu > 0$ such that the average value of the square distance of $\bar{b}_{RR}$ from $\bar{\beta}$ is smaller than the corresponding distance of $\bar{b}_{GLS}$ from $\bar{\beta}$. Thus, the point estimate of $\bar{\beta}$, obtained by the generalized ridge procedure is on an average closer to $\bar{\beta}$, than the corresponding point estimate of $\bar{\beta}$ obtained by the generalized least squares procedure.

Bayesian Interpretation of the Random-Ridge Estimator

Assumption 2: The (p + 1) x 1 vector $\bar{\beta}$ is a random variable whose prior distribution is normal with mean vector 0 and variance-covariance matrix $\sigma^2_{\bar{\beta}} \, I$.

Suppose that we add Assumption 2 to the model (2.6), and consider w in (2.6) to be normally distributed, then we can provide a Bayesian interpretation to the Random Ridge estimator. It can be shown that the RR estimator in (3.1) is the mean of the posterior distribution of $\bar{\beta}$ with $\mu = \frac{1}{\sigma^2_{\bar{\beta}}}$ (see Zellner [1971, p. 76]). Further, following Lindley and Smith [1972], we can obtain an iterative estimate of $\sigma^2_{\bar{\beta}}$ starting from the generalized least square estimates of $\bar{\beta}$'s in (3.4). That is

$$\hat{\sigma}^2_{\bar{\beta}} = \frac{1}{p} \sum_{i=o}^{p} (\bar{b}_{i,\ GLS} - \bar{\bar{b}})^2$$

where $\bar{\bar{b}}$ is the mean of the $\bar{b}_{i,\ GLS}$. The RR estimator for $\mu = \frac{1}{\hat{\sigma}^2_{\bar{\beta}}}$ then provides new estimates for $\bar{\beta}$'s. A revised estimate of $\sigma^2_{\bar{\beta}}$ may be obtained based on the new estimates of $\bar{\beta}$'s and the procedure repeated until convergence in μ is obtained.

Modified Random Ridge Estimator

Since the RR estimator implies a prior distribution for $\bar{\beta}$ with zero mean, which may not be a plausible assumption to make in many situations, we may obtain a modified Random Ridge (MRR) estimator of $\bar{\beta}$, given as

$$\bar{b}_{MRR} = (X' \ \Omega^{-1} \ X + \mu \ I)^{-1} \ (X' \ \Omega^{-1} \ y + \mu\alpha)$$

The MRR estimator is, in fact, the mean of the posterior distribution of $\bar{\beta}$ with $\mu = \frac{1}{\sigma^2_{\bar{\beta}}}$ when the prior distribution of $\bar{\beta}$ is assumed to be $N(\alpha, \ \sigma^2_{\bar{\beta}} \ I)$.

II. Random Almon Estimator

The generalized ridge procedure restricts $\bar{\beta}$ to be in the hypersphere of radius r^2, which is an unduly restrictive condition. Another difficulty with the RR method is that it ignores the degrees of freedom problem associated with the estimation process for the distributed lag model. It is well known that when long lags are specified there may be very few degrees of freedom left for the estimation process. This difficulty may be resolved if we use some a priori structure on the $\bar{\beta}$ such as (2.11).

Writing the restriction on $\bar{\beta}$ in (2.11) as

(3.5) $S\bar{\beta} = 0$

where

(3.6) $S = I - A(A' \ A)^{-1} \ A'$

is an idempotent matrix of order $(p + 1) \times (p + 1)$ and of rank $(p - r)$; we may obtain the restricted generalized least square estimator of $\bar{\beta}$ in (2.6) which minimizes $(y - X\bar{\beta})' \, \Omega^{-1}(y - X\bar{\beta})$ subject to the condition (3.5). This is

$$(3.7) \quad \bar{b}_{RA} = c^* \bar{b}_{GLS}$$

where $\bar{b}_{GLS}$ is the unconstrained generalized least squares estimator of $\bar{\beta}$ defined in (3.4) and c^* is the correction factor defined as (+ represents generalized inverse)

$$c^* = I - (X' \Omega^{-1} X)^{-1} S' [S(X' \Omega^{-1} X)^{-1} S']^{+} S.$$

Its variance-covariance matrix is given as

$$(3.8) \quad V(\bar{b}_{RA}) = c^* (X' \Omega^{-1} X)^{-1} c^{*\prime}.$$

The estimator $\bar{b}_{RA}$, which may be termed as the Random Almon estimator, is a generalization of the Almon [1965] estimator. An alternative form of $\bar{b}_{RA}$, useful for numerical evaluations, is given at the end of this section in the Remark 4. The Almon estimator of $\bar{\beta}$ in (2.6) may be obtained by substituting $\Omega = \sigma_u^2 I$ in (3.7), and is given as

$$(3.9) \quad \bar{b}_A = d^* \bar{b}_{LS}$$

where

$$(3.10) \quad d^* = I - (X' X)^{-1} S' [S (X' X)^{-1} S']^{+} S$$

and $\bar{b}_{LS}$ is the unconstrained least squares estimator of $\bar{\beta}$ in (2.9) and is given as

$$(3.11) \quad \bar{b}_{LS} = (X' X)^{-1} X' y.$$

The variance-covariance matrix of $\bar{b}_A$ is given as

$$(3.12) \quad V(\bar{b}_A) = d^* (X' X)^{-1} X' \Omega X (X' X)^{-1} d^{*\prime}.$$

III. Random Shiller Estimator

We observe that while in I the Random Ridge and generalized least squares estimators of $\bar{\beta}$ in (2.6) are obtained without using (2.7) or (2.11), the Random Almon and Almon estimators in II are obtained by considering the restriction $\bar{\beta} = A\delta$. It has been, however, argued by Shiller [1973] that this restriction is often specified not because we believe in it but because we believe the lag distribution to be smooth. Since the latter may not be true we now develop estimators by considering a stochastic specification of the form given in (2.7) viz,

$$\bar{\beta} = A\delta + \eta.$$

The above specification can alternatively be written as

(3.13) $R\bar{\beta} = R\eta$

where R is a p - (r - 1) x (p + 1) matrix defined below

(3.14) $$R = \begin{bmatrix} (-1)^o\binom{r+1}{o} & (-1)^1\binom{r+1}{1} & \cdots & (-1)^{r+1}\binom{r+1}{r+1} & 0 & 0 \cdots & 0 \\ 0 & (-1)^o\binom{r+1}{o} & \cdots & (-1)^r\binom{r+1}{r} & (-1)^{r+1}\binom{r+1}{r+1} & & \\ \vdots & & & & & & \\ 0 & 0 \cdots & (-1)^o\binom{r+1}{o} & & \cdots & & (-1)^{r+1}\binom{r+1}{r+1} \end{bmatrix}$$

such that

(3.15) $RA\delta = 0$

and $R\eta$ is a random vector with (using Assumption 1) $ER\eta = 0$ and $ER\eta\eta' R' = \sigma^2_\eta RR'$; $\sigma^2_\eta = \sigma^2_{\bar{\beta}}$ according to (2.7).

We note here that Shiller's [1973] specification is less general than that given in (3.13) in the sense that $R\eta$ in his specification is considered as a random vector with mean zero and covariance matrix σ^2_η I. In fact he begins with

$\bar{\beta} = A\delta$ which implies $R\bar{\beta} = 0$ and then puts a stochastic vector, say η, to specify $R\bar{\beta} = \eta$. In addition, in his case $\Omega = \sigma_u^2 I$.

Now to obtain the mixed estimator of $\bar{\beta}$ we combine (2.6) and (3.13), and apply the Aitken theorem. This provides

(3.16) $\bar{b}_{GRS} = (X' \Omega^{-1} X + \mu R' (RR')^{-1} R)^{-1} X' \Omega^{-1} y$

where $\mu = \frac{1}{\sigma_{\bar{\beta}}^2}$, $\sigma_{\bar{\beta}}^2 = \sigma_{\eta}^2$. The variance-covariance matrix of $\bar{b}_{GRS}$ is

(3.17) $V(\bar{b}_{GRS}) = [X' \Omega^{-1} X + \mu R' (RR')^{-1} R]^{-1}$.

The estimator (3.16) may be termed as the Generalized Random Shiller (GRS) estimator. As would be expected, the GRS estimator becomes RA estimator given in (3.7) when σ_{η}^2 tends to zero. This follows by noting that

(3.18) $R' (RR')^{-1} R = S = I - A(A'A)^{-1} A'$,

where S is an idempotent matrix given in (3.6), and using the result

(3.19) $(X' \Omega^{-1} X + \mu S)^{-1} = (X' \Omega^{-1} X)^{-1} [I - S' \{S(X' \Omega^{-1} X)^{-1} S' + \frac{1}{\mu} I\}^{-1} S]$.

It is also interesting to note that if $\sigma_{\eta}^2 \to 0$ and $\sigma_{ii} = 0$, i.e., $\Omega = \sigma_u^2 I$ then the GRS estimator becomes the Almon estimator given in (3.9).

Now, if we assume that $R\eta$ is distributed with zero mean and the variance-covariance matrix $\sigma_{\eta}^2 I$ then the mixed-estimator for $\bar{\beta}$ in (2.6) under the restriction (3.13) is

(3.20) $\bar{b}_{RS} = (X' \Omega^{-1} X + \mu R' R)^{-1} X' \Omega^{-1} y$

and its variance-covariance matrix is

(3.21) $V(\bar{b}_{RS}) = (X' \Omega^{-1} X + \mu R' R)^{-1} (X' \Omega^{-1} X + \mu (R' R)^2)(X' \Omega^{-1} X + \mu R' R)^{-1}$.

The estimator $\bar{b}_{RS}$ may be termed as the Random Shiller (RS) estimator. The RS estimator is a generalization of the Shiller (S) estimator which considers $\sigma_{ii} = 0$, i.e., $\Omega = \sigma_u^2 I$ (see Maddala [1977]) and it is given as

(3.22) $\bar{b}_S = (X' X + \lambda R' R)^{-1} X' y$, $\lambda = \sigma_u^2 \mu$.

The variance-covariance matrix of $\bar{b}_S$ is

(3.23) $V(\bar{b}_S) = (X' X + \lambda R'R)^{-1} (X' \Omega X + \lambda \sigma_u^2 (R' R)^2)(X' X + \lambda R'R)^{-1}$

<u>Bayesian Interpretation of the GRS Estimator</u>

<u>Assumption 3</u>: The (p + 1) vector of $\bar{\beta}$ is a random variable whose prior distribution is normal with the mean $A\delta$ and variance covariance $\sigma_{\bar{\beta}}^2 I$. Also, δ follows a diffused prior.

Suppose we add Assumption 3 to Assumption 1 and consider w to be normally distributed, then we can provide a Bayesian interpretation to the GRS estimator. In fact, following Maddala [1977, p. 385] it can be shown that the mean of the posterior distribution of $\bar{\beta}$ is

(3.24) $(X' \Omega^{-1} X + \mu S)^{-1} X' \Omega^{-1} y$

which is GRS estimator, as given in (3.17), because of (3.18). We may also call GRS estimator as Bayes Random Almon (BRA) estimator. This is because (3.24), using the Assumption 3, is the mean of the posterior distribution of $\bar{\beta}$ for the model (2.6) under the restriction (3.5).

It is interesting to note that when $\sigma_{ii} = 0$, i.e., $\Omega = \sigma_u^2 I$ the GRS or BRA estimator becomes

(3.25) $\bar{b}_{BA} = (X' X + \lambda S)^{-1} X' y$

which is Bayes Almon (BA) estimator discussed in Maddala [1977] (also see Lindley and Smith [1972]). Its variance covariance matrix is given by

(3.26) $V(\bar{b}_{BA}) = (X' X + \lambda S)^{-1} (X' \Omega X + \lambda \sigma_u^2 S)(X' X + \lambda S)^{-1}$.

It is clear from (3.16) and (3.25) that both the BRA (GRS), and BA estimators are also ridge type estimators. Further, the BRA and BA estimators treat $\bar{\beta}$ as a random variable while the RA and A estimators regard $\bar{\beta}$ as a fixed parameter.

The ridge coefficients in (3.16) and (3.25) may be obtained by an iterative method similar to the procedure for μ in the RR method, explained in I.

It is clear from the GRS and RS estimators that it makes a difference whether we postulate that the variance-covariance matrix of $R\eta$ is $\sigma_\eta^2 RR'$ or $\sigma_\eta^2 I$. In fact, both the RS estimator which is based on the misspecification of $R\eta$ and the Shiller estimator will be less efficient compared to GRS.

At this point a few remarks on the above discussion may be useful.

Remark 1: The above exposition has been in terms of a single explanatory variable x with a total lag length of p, but the analysis can easily be extended to allow for more explanatory variables with different total lag lengths (Tinsley [1967] and Almon [1968]).

Remark 2: In some situations it may be more desirable to assume that

$$\beta_{it} = \bar{\beta}_i + \bar{\gamma}_i z_{t-i} + \epsilon_{it}, \qquad i = 0, 1, \ldots, p.$$

Thus the size of the effect of x_{t-i} on y_i depends on not only the lapse of time and the specific x value but also on the value of the variable z. Under this assumption the model becomes varying parameters distributed lag model with two explanatory variables (see Almon [1968]).

Remark 3: In other situations it may be appropriate to postulate that

$$\beta_{it} = \bar{\beta}_i + \bar{\gamma}_{1i} D_{1t-i} + \bar{\gamma}_{2i} D_{2t-1} + \bar{\gamma}_{3i} D_{3t-i}$$

where D_1, D_2 and D_3 are seasonal dummies. Assuming $\bar{\beta}_1$, $\bar{\gamma}_{1i}$, $\bar{\gamma}_{2i}$ and $\bar{\gamma}_{3i}$ lie on low degree polynomials, the model has four Almon type explanatory variables (see Pesando [1972]).

Remark 4: It is noted that Ω is a diagonal matrix whose inverse is easily found. Therefore, the computation in the BRA (GRS) and other methods may be quite easily and inexpensively handled. Further, the computation for the RA and the Almon (A) estimators may be handled with even greater ease if we obtain these estimators of $\bar{\beta}$ using an alternative approach. Substituting the condition

(3.5) in the model (2.6) we may obtain $y = Z\delta + u$ where $Z = XA$. Then the generalized least squares estimator of δ is $d = (Z' \Omega^{-1} Z)^{-1} Z' \Omega^{-1} y$ and its variance-covariance matrix is $V(d) = (Z' \Omega^{-1} Z)^{-1}$. We may write the RA estimator of $\bar{\beta}$ by substituting d for δ in the condition (3.5) as:

$$\bar{b}_{RA} = A(Z' \Omega^{-1} Z)^{-1} Z' \Omega^{-1} y$$

and its variance-covariance matrix is

$$V(\bar{b}_{RA}) = A(Z' \Omega^{-1} Z)^{-1} A'.$$

The Almon estimator and its variance-covariance may be obtained by substituting $\Omega = \sigma_u^2 I$ in $\bar{b}_{RA}$ and $V(\bar{b}_{RA})$, respectively.

Remark 5: The polynomial lag structure (2.3) excludes the possibility of any long tail lag distribution. However, this problem can easily be solved by using a piecemeal polynomial.

Remark 6: The Random Almon method is applicable only to a finite lag structure whose length must be specified a priori. Some researchers have suggested trying different values of p and choosing the optimal value on the basis of either the minimum standard error or $\bar{R}^2$ criteria.

Remark 7: The end point restrictions not dealt with here are generally imposed a priori, but one can test these restrictions because they are linear hypotheses which can be verified through standard tests. Schmidt and Waud [1973] have argued against imposition of these end point constraints. In addition, the question of tests for restrictions in the set of equations (2.11) or (3.13) could easily be handled on the lines of similar tests for the A and S estimators (see Maddala [1977] test B-10, p. 458).

3.2 Estimation of $\bar{\beta}$ when Ω is not Known

In situations when Ω is not known, we may replace it by its consistent estimator and obtain approximate estimators corresponding the alternative estimators of $\bar{\beta}$ discussed above. From the large-sample point of view, all that is required is any consistent estimator of Ω because the asymptotic distributions under some general regularity conditions of the alternative estimators are not affected (see Fuller and Battese [1973] and Anderson [1971]).

Now, we may discuss a few consistent methods of estimating the unknown parameters in Ω.

If $\sigma = (\sigma_u^2, \sigma_{oo}, \ldots, \sigma_{rr})'$ and $M = I - Z(Z'Z)^{-1}Z'$, then $Z = XA$, following Hildreth and Houck [1968], the LS estimator of σ is

$$(3.27) \quad \hat{\sigma} = [(\dot{M}\dot{X}^*)'(\dot{M}\dot{X}^*)]^{-1}(\dot{M}\dot{X}^*)'\dot{\hat{w}}$$

where $\hat{w} = My$ and $X^* = (\iota, X)$; ι being the vector of unit elements. Further $\dot{M} = M{*}M$, $\dot{X}^* = X^* {*} X^*$, $\dot{\hat{w}} = \hat{w} {*} \hat{w}$ are the Hadamard products (see Rao [1970, p. 30]). A necessary and sufficient condition for σ to be identified is that the rank of $\dot{M}\dot{X}^*$ is $(p + 2)$. Once the estimate of σ is obtained from (3.27) the estimate of Ω immediately follows.

It is well known that the LS estimator $\hat{\sigma}$ is an unbiased, consistent but inefficient estimator of σ. Further, it can be shown that $E\eta\eta' = 2\dot{\Sigma} = 2M\Omega M{*}M\Omega M$ when $\eta = \dot{\hat{w}} - \dot{M}\dot{X}^*\sigma$ is normally distributed. Thus one may obtain the GLS estimator of σ as

$$(3.28) \quad \tilde{\sigma} = [(\dot{M}\dot{X}^*)\dot{\Sigma}^{-1}(\dot{M}\dot{X}^*)]^{-1}(\dot{M}\dot{X}^*)'\dot{\Sigma}^{-1}\dot{\hat{w}}.$$

It is not clear whether we should use $\hat{\sigma}$ or $\tilde{\sigma}$ in place of the true values of σ in the alternative estimators of $\bar{\beta}$ depending upon Ω. However, since $\hat{\sigma}$ is computationally easier and less expensive we use it in this paper. Further, any element of $\hat{\sigma}$ or $\tilde{\sigma}$ can be negative with positive probability. Since we cannot use a negative value for a positive parameter, it has been suggested that one may replace negative elements of $\hat{\sigma}$ or $\tilde{\sigma}$ by zeros. Following this suggestion, one may obtain truncated estimators of $\bar{\beta}$. In simpler situations the truncated estimators are found to be preferable in terms of mean square error to untruncated estimators. In the numerical example below we have obtained the truncated estimators.

4. AN EXAMPLE: ALMON MODEL

To illustrate the GLS, RA, RS and GRS estimators developed in the previous section, we utilize the Almon data (see Almon [1968]). A stochastic coefficients version of the model with eight period lag is briefly described below:

$$(4.1)\quad y_t = \sum_{i=0}^{8} \beta_{it}\, x_{t-i} + u_t$$

where

y_t = capital expenditure in period t

x_{t-i} = capital appropriation in period t-i.

We further assume that

$$(4.2)\quad \beta_{it} = \bar{\beta}_i + \epsilon_{it}$$

and $\bar{\beta}_i$ lies on a second-degree polynomial, i.e.,

$$(4.3)\quad \bar{\beta}_i = \delta_0 + \delta_1 i + \delta_2 i^2 + \eta_i \quad i=0, 1, \ldots, 8.$$

Our objective is to estimate the mean of β's on the basis of quarterly data for the period 1953-1967, using the LS estimates of σ's. The fixed coefficient version of the Almon model has been estimated by Almon [1968] and others. We present the GLS, RA, RS and BRA (GRS) estimates for the stochastic version of Almon model in Table 1 along with the usual LS, A, S, and BA estimates.

A plot of the alternative estimates in columns 2 to 5 in Figure 1 reveals that the LS estimator produces an erratic lag pattern while the A, S and BA estimators produce a quadratic smooth lag pattern. Further, it appears that the choice among the A, S and BA methods, purely on the basis of the shape of lag distributions, is difficult.

In order to provide a comparison among alternative methods of estimation in the distributed lag model with stochastic coefficients, purely on the basis of shape of lag distributions, we have plotted the alternative estimates in columns 2, 3 and 6 to 9 in Figure 2. An interesting point to note in Figure 2 is that the A method seems to produce a more smooth lag pattern compared to the RA method. Further, the GRS and RS methods are to be preferred over RA estimator if a less smooth lag pattern is expected.

The number of iterations required for convergence for the ridge type estimators and the values of the ridge coefficients are also given in rows 12 and 13 respectively of Table 1. In all cases the convergence was obtained in nine or less iterations.

The estimates of the model (4.1) were also obtained by the RR method. The convergence for this method was obtained in fifteen iterations for the ridge coefficient value of 2055. The RR estimates (.12141, .12780, .13085, .13687, .11795, .09750, .09769, .08332, .06877), appear to produce an over smoothed lag pattern although they are not equal, see Maddala [1977, p. 386].

5. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

In this paper we studied the problem of estimating a polynomial distributed lag model under more general stochastic specification on the coefficients than considered earlier by Shiller [1973]. We first developed an efficient estimator, namely the Random Almon estimator, for the stochastic parameters distributed lag model when the mean of the stochastic coefficient is regarded fixed. The Almon estimator is a special case of this estimator. Further, when the mean of the stochastic coefficient is random we proposed the estimator, namely the Generalized Random Shiller (GRS) estimator. It has been indicated that the Shiller estimator, which is based on a misspecified stochastic disturbance term, will be less efficient compared to the GRS. Also we have shown that the Shiller, Almon, Bayesian Almon and Random Almon estimators, among others, can be considered as special cases of the GRS estimator. Finally, on the basis of the numerical estimates, it has been suggested that the GRS is to be preferred as compared to the other estimators when the lag distribution is not expected to be smooth.

Several suggestions for further research can be made. First, a modification of the distributed lag model with stochastic coefficients when an Almon lag is combined with the Koyck lag, may be done on the lines of Schmidt [1974] work. Second, there is a need to extend the model when errors are autocorrelated, and to compare its performance in the extended form with the Hannan Inefficient estimator discussed in Hannan [1967].

TABLE

Alternative Estimates of the Means of Stochastic Coefficients in Almon's Model: 1953Q1-67Q4

Lag	Coefficient Estimates							
	LS	A	S	BA	GLS	RA	RS	BRA (GRS)
0	.07233	.10484	.10225	.10451	.06141	.12020	.08960	.10458
1	.08219	.12646	.12721	.12645	.14017	.13177	.14301	.13326
2	.23065	.13931	.14132	.13970	.16718	.13716	.16290	.14749
3	.18609	.14337	.14484	.14366	.22503	.13638	.15305	.15924
4	.13388	.13864	.13854	.13857	.09550	.12941	.12418	.11877
5	.01338	.12514	.12355	.12472	.02605	.11628	.09745	.09538
6	.13535	.10284	.10091	.10273	.11689	.09696	.08244	.10238
7	.06429	.07177	.07120	.07179	.13309	.07147	.07146	.07992
8	.06924	.03191	.03463	.03218	.02030	.03981	.05799	.04014
sum	.98740	.98428	.98445	.98431	.98562	.97944	.98208	.98116
No. of Iterations	-	-	4	4	-	-	6	9
Ridge Coefficient	-	-	832.156	824.756	-	-	785.411	855.060

ALTERNATIVE LAG DISTRIBUTIONS

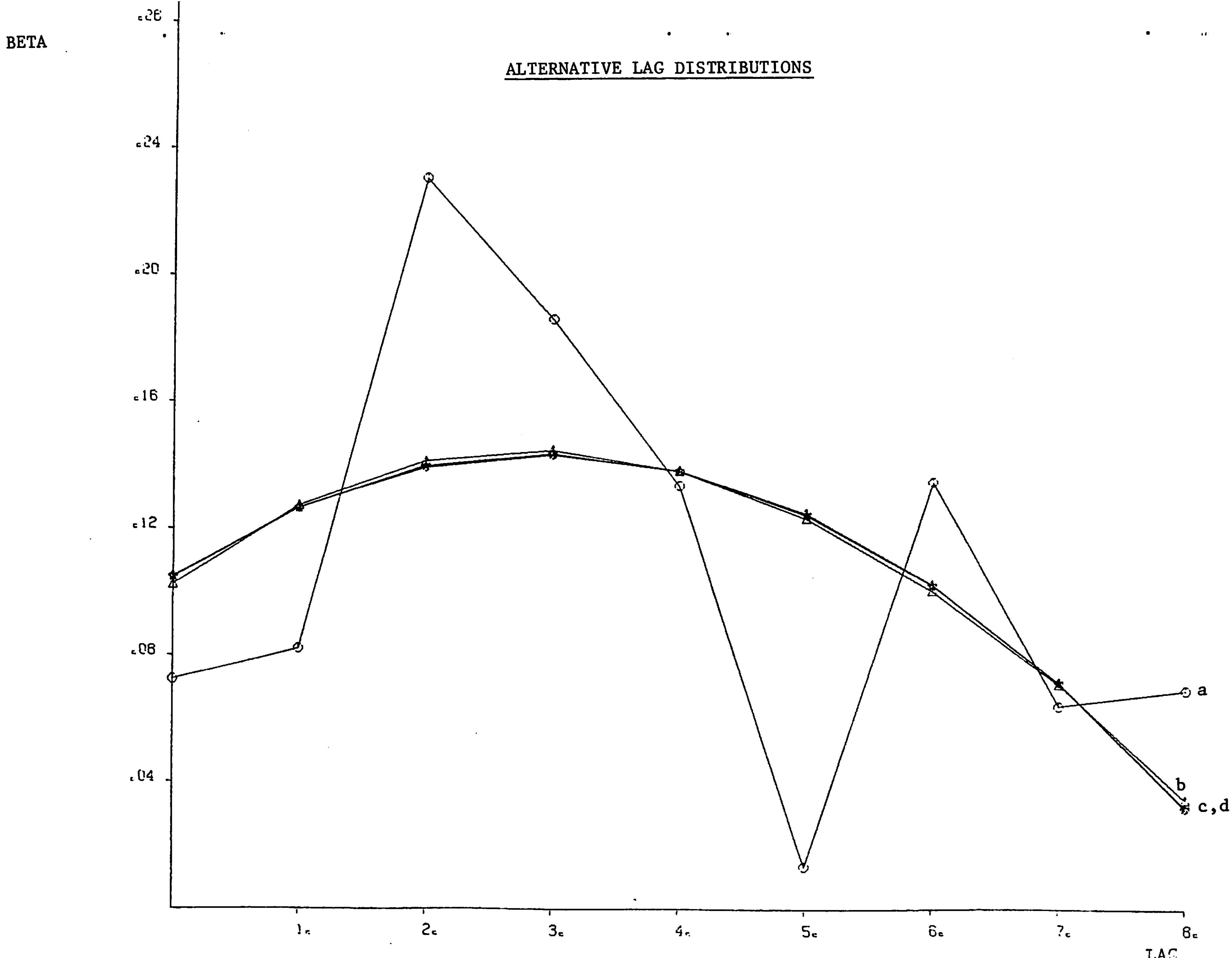

Figure 1

a - Least Squares Estimates
b - Shiller Estimates
c - Almon Estimates
d - Bayes-Almon Estimates

ALTERNATIVE LAG DISTRIBUTIONS

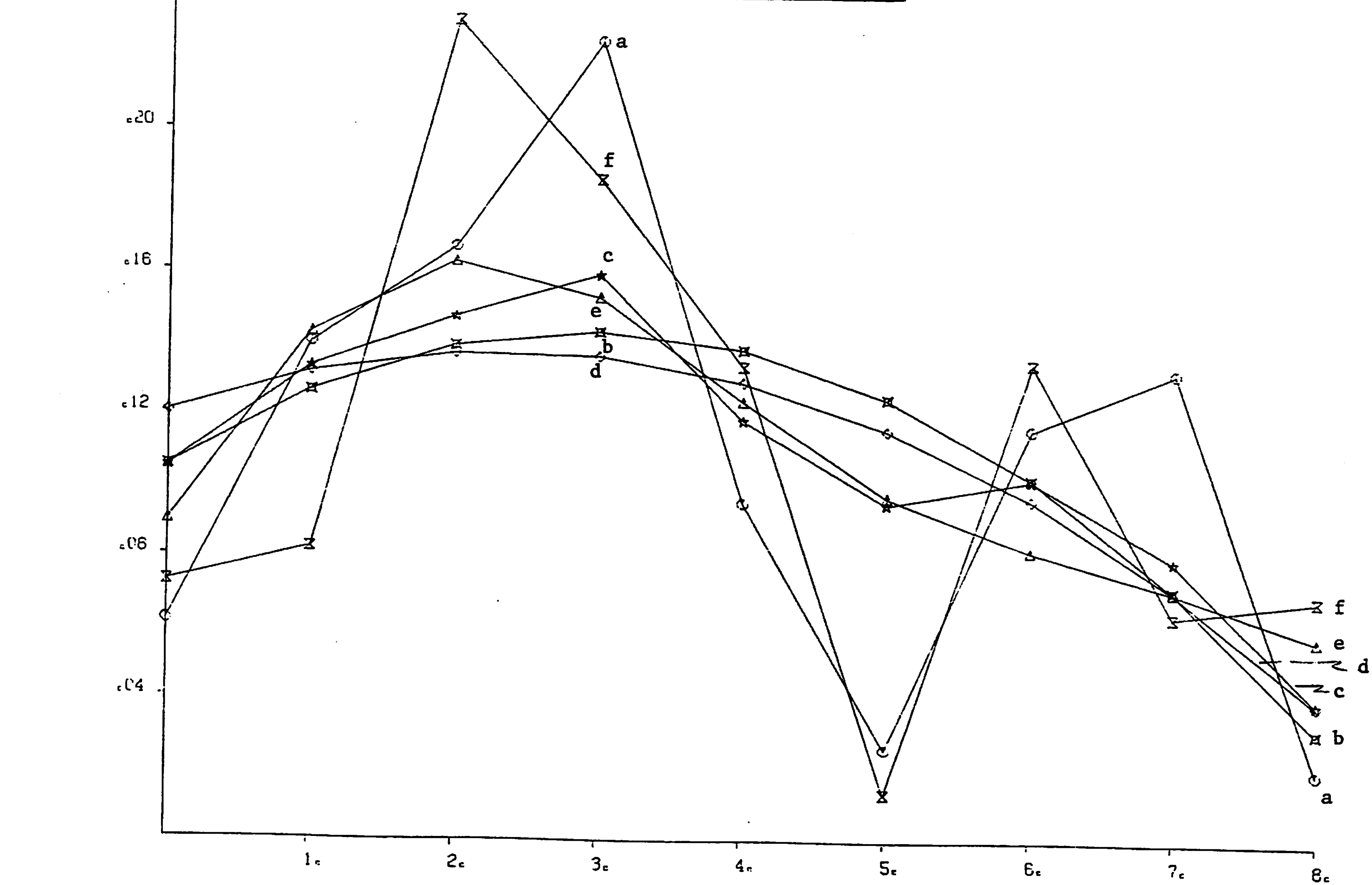

Figure 2

a - 'GLS' Estimates
b - 'A' Estimates
c - 'GRS' Estimates
d - 'RA' Estimates
e - RS Estimates
f - 'LS' Estimates

REFERENCES

Almon, S., "The Distributed Lag Between Capital Appropriations and Net Expenditures," Econometrica (January 1965), pp. 178-196.

Almon, S., "Lags Between Investment Decisions and Their Causes," Review of Economics and Statistics, 1968, pp. 193-206.

Anderson, T. W., The Statistical Inference of Time Series (New York: Wiley, 1971).

Fuller, W. A., and G. E. Battese, "Transformations for Estimation of Linear Models with Nested-Error Structure," Journal of the American Statistical Association (September 1973), pp. 626-632.

Hannan, E. J., "The Estimation of a Lagged Regression Relation," Biometrica, (1967), pp. 409-418.

Hildreth, C., and J. Houck, "Some Estimates for a Linear Model with Random Coefficients," Journal of the American Statistical Association (June 1968), pp. 584-595.

Hoerl, A. E., and R. W. Kennard, "Ridge Regression: Biased Estimation for Non-Orthogonal Problems, Technometrics, 1970, pp. 55-67.

Lindley, D. V., and A. F. M. Smith, "Bayes Estimates for the Linear Model (with Discussion)," Journal of the Royal Statistical Society, B Series, 1972, pp. 1-41.

Maddala, G. S., Econometrics (New York: McGraw-Hill Book Company, 1977).

Pesando, J. E., "Seasonal Variability in Distributed Lag Models," Journal of the American Statistical Association (June 1972), pp. 311-312.

Rao, C. R., Linear Statistical Inference and Its Applications, 2nd Edition (New York: Wiley, 1973).

Schmidt, P., and R. N. Wand, "The Almon Lag Technique and the Monetary Versus Fiscal Policy Debate," Journal of the American Statistical Association (March 1973), pp. 11-19.

Schmidt, P., "A Modification of the Almon Distributed Lag," Journal of the American Statistical Association (September 1974), pp. 679-681.

Shiller, R. J., "A Distributed Lag Estimator Derived from Smoothness Priors," Econometrica (July 1973), pp. 775-788.

Swamy, P. A. V. B., "Criteria, Constraints and Multicollinearity in Random Coefficient Regression Models," Annals of Economic and Social Measurement (1973), pp. 429-449.

Swamy, P. A. V. B., and J. S. Mehta, "Bayesian and Non-Bayesian Analysis of Switching Regressions and of Random Coefficient Regression Models," Journal of the American Statistical Association (September 1975), pp. 593-602.

Tinsley, P. A., "An Application of Variable Weight Distributed Lags," Journal of the American Statistical Association (1967), pp. 1277-1289.

Zellner, A., An Introduction to Bayesian Influence in Econometrics (New York: John Wiley and Sons, 1971).