Mean square error comparisons of the alternative estimators for the distributed lag models

Berrin Gültay[*] and Selahattin Kaçıranlar[†]

Abstract

The finite distributed lag models include highly correlated variables as well as lagged and unlagged values of the same variables. Some problems are faced for this model when applying the ordinary least squares (OLS) method or econometric models such as Almon and Koyck models. The primary aim of this study is to compare the performances of alternative estimators to the OLS estimator defined by combining the Almon estimator with some other estimators according to the mean square error (MSE) criterion. We use Almon [2] data to illustrate our theoretical results.

1. Introduction

Consider the finite distributed lag model,

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + u_t, \qquad t = p+1, \cdots, T$$

$$(1.1) \qquad = \sum_{i=0}^{p} \beta_i x_{t-i} + u_t$$

where u_t are $IN\left(0, \sigma_u^2\right)$. The coeffients β_i are called lag weights. The model in Eq.(1.1) can be written in the matrix notation as

$$(1.2) \qquad y = X\beta + u$$

---

[*]Faculty of Arts and Sciences, Çanakkale Onsekiz Mart University, Çanakkale, Turkey. Email: `berringultay@comu.edu.tr` Corresponding Author.

[†]Faculty of Arts and Sciences, Çukurova University, Adana, Turkey. Email:`skacir@cu.edu.tr`

where

$$y = \begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_T \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, X = \begin{bmatrix} x_{p+1} & x_p & \cdots & x_1 \\ x_{p+2} & x_{p+1} & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_T & x_{T-1} & \cdots & x_{T-p} \end{bmatrix}, u = \begin{bmatrix} u_{p+1} \\ u_{p+2} \\ \vdots \\ u_T \end{bmatrix}.$$

In case of estimating the model (1.1) by OLS, the following problems are encountered:

a) Multicollinearity problem among the explanatory variables may be occured. Because there are p lags of the same variables in the model.

b) The length of the lag, p, isn't known. Even if p is known, if this number is large and amount of the sample is small, it is unable to estimate the parameters.

To overcome these problems, some kind of distributed lag models have been suggested such as Koyck and Almon models (Yurdakul [21]). The most of these estimators require some prior information about the behavior of the β's in (1.1). In general, the two sources of prior information can be classified as nonstochastic and stochastic smoothness prior (Vinod and Ullah, [19]; Gujarati, [5]).

Irving Fisher [4] initially introduced nonstochastic smoothness prior information of the following type:

$$(1.3) \qquad \beta_i = (p + 1 - i)\alpha \qquad 0 \le i \le p$$
$$= 0 \qquad\qquad i > p$$

where α is any unknown parameter. Substituting (1.3) in (1.1) gives,

$$y_t = \left[\sum_{i=0}^{p} (p + 1 - i)x_{t-i} \right] \alpha + u_t$$

$$(1.4) \qquad = z_t \alpha + u_t$$

Thus the OLS estimate of α can be obtained from (1.4) and then using (1.3), the estimate of β_i can be obtained. A generalization of the linear nonstochastic prior on β_i can be written as

$$(1.5) \qquad \beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \ldots + \alpha_r i^r \qquad\qquad p \ge r \ge 0$$

which is a polynomial of the rth degree. This structure on lag weights β_i was proposed by Almon [2] and is known as the Almon polynomial lag. Again, substituting (1.5) in (1.1) we can get estimates of the α's and then using (1.5) we can obtain the estimates of β_i. Eq. (1.5) can be written in the matrix notation as

$$(1.6) \qquad \beta = A\alpha$$

where β is given before, and

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & p & p^2 & \vdots & p^r \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_r \end{bmatrix}$$

are $A : (p+1) \times (r+1)$ matrix and $\alpha : (r+1) \times 1$ vector. The ranks of matrices X and A are assume to be $(p+1) < (T-p)$ and $(r+1) < (p+1)$, respectively. If $r < p$, then the rank of A is $r+1$. We estimate β in (1.2), under the nonstochastic prior information on β is given by (1.6), using Almon estimation method. By substituting (1.6) in (1.2),

$$y = XA\alpha + u$$

$$(1.7) \qquad = Z\alpha + u, \qquad u \sim N\left(0, \sigma_u^2\right)$$

is obtained. This model can be called a linear Almon distributed lag model. Then, OLS estimator of α in model (1.7) is

$$(1.8) \qquad \widehat{\alpha}_A = \left(Z'Z\right)^{-1} Z'y = \left(A'X'XA\right)^{-1} A'X'y .$$

In this case,

$$(1.9) \qquad \widehat{\beta}_A = A\widehat{\alpha}_A$$

is the Almon estimator of β. $\widehat{\beta}_A$ is the best linear unbiased estimator (BLUE).

2. Alternative methods

In this section some alternative biased estimators to the Almon estimator are defined for the distributed lag model.

2.1. The Almon-modified ridge estimator. Hoerl and Kennard's ridge regression estimator has been discussed as an alternative approach to resolve problems encountered in due to some disadvantages of Almon estimator (Maddala [14], Vinod and Ullah [19], Chanda and Maddala [3]). Distributed lag estimation seems tractable only when prior information on the lag coefficients is incorporated. Ridge regression introduces yet another representation of such prior information and hence is a possible estimation procedure (Yeo and Trivedi [20]).

The Almon-ridge estimator of α in model (1.7) is

$$(2.1) \qquad \widehat{\alpha}_k = \left(Z'Z + kI\right)^{-1} Z'y$$

$$= \left(A'SA + kI\right)^{-1} A'X'y \qquad k > 0$$

where $S = X'X$. Thus

$$(2.2) \qquad \widehat{\beta}_k = A\widehat{\alpha}_k$$

is the Almon-ridge estimator for the model (1.2). However, the ridge estimator and the extension given by Lindley and Smith [12] are not as promising for the distributed lag models (Maddala, [14]). They tried various values of the k. But they are not satisfied the

results of some empirical examples with this method. Because the selection of k reveals several problems. Therefore, alternative estimation methods must be considered.

Swindel [16] introduced a modified ridge estimator based on prior information b_0. Almon-modified ridge estimator of α in model (1.7) is defined,

$$(2.3) \qquad \widehat{\alpha}(k, b_0) = \left(Z'Z + kI\right)^{-1} \left(Z'y + kb_0\right).$$

As pointed out by Swindel [16], it seems more useful and reasonable in the applications to consider the prior information. To overcome multicollinearity problem, if we take $b_0 = \widehat{\alpha}_k$, (2.3) is reduced to

$$\widehat{\alpha}_m(k) = \left(Z'Z + kI\right)^{-1} \left(Z'y + k\widehat{\alpha}_k\right)$$

$$= T_k\widehat{\alpha}_A + k\left(Z'Z + kI\right)^{-1} \widehat{\alpha}_k$$

$$(2.4) \qquad\qquad = T_k\widehat{\alpha}_A + \left(I - T_k\right)\widehat{\alpha}_k$$

where $T_k = \left(Z'Z + kI\right)^{-1} Z'Z$. Substituting $\widehat{\alpha}_k$ for b_0, it is expected that $\widehat{\alpha}_m(k)$ has advantage according to the Almon-ridge and Almon estimators. Thus, Almon-modified ridge estimator of β in model (1.2) is $\widehat{\beta}_m(k) = A\widehat{\alpha}_m(k)$. In application b_0 might well be chosen to reflect as well as possible the prior information or restricted on β.

2.2. The Almon-modified Liu estimator. In order to overcome the multicollinearity problem, ridge estimator that we have discussed before is widely used in practice, but selection of k poses some problems. To overcome this problem an estimator is defined by combining Ridge and Stein type estimators in Liu [13]. This estimator was called Liu estimator in Akdeniz and Kaçıranlar [1]. The advantage of Liu estimator over ridge estimator is a linear function of d and therefore selection of d is easier. Liu estimator of β in (1.2) is

$$\widehat{\beta}_d = \left(X'X + I\right)^{-1} \left(X'y + db\right)$$

$$(2.5) \qquad = \left(X'X + I\right)^{-1} \left(X'X + dI\right) b, \qquad 0 < d < 1$$

where b is the OLS estimator for model (1.2). To overcome multicollinearity problem, if we take $\widehat{\alpha}_A$ instead of b, Almon-Liu estimator of α in model (1.7) is

$$\widehat{\alpha}_d = \left(Z'Z + I\right)^{-1} \left(Z'y + d\widehat{\alpha}_A\right)$$

$$(2.6) \qquad = \left(A'SA + I\right)^{-1} \left(A'X'y + d\widehat{\alpha}_A\right)$$

obtained. This estimator can be given,

$$\widehat{\alpha}_d = \left(Z'Z + I\right)^{-1} \left(Z'Z + dI\right) \widehat{\alpha}_A$$

$$= \left(A'SA + I\right)^{-1} \left(A'SA + dI\right) \widehat{\alpha}_A$$

(2.7) $= F_d \widehat{\alpha}_A$

where $F_d = (Z'Z + I)^{-1}(Z'Z + dI)$. Thus, the Almon-Liu estimator of β is

$\widehat{\beta}_d = A\widehat{\alpha}_d$. Comparison of $\widehat{\alpha}_A$ with $\widehat{\alpha}_d$ and selection of d are given in Kaçıranlar [9].

Li and Yang [11] introduced a modified Liu estimator based on prior information similar to (2.3). Almon-modified Liu estimator of α in model (1.7) is defined,

(2.8) $\widehat{\alpha}(d, b_0) = (Z'Z + I)^{-1}(Z'Z + dI)\widehat{\alpha}_A + (1 - d)(Z'Z + I)^{-1}b_0.$

To overcome multicollinearity problem, if we take $b_0 = \widehat{\alpha}_d$, (2.8) is reduced to

(2.9) $\widehat{\alpha}_m(d) = F_d\widehat{\alpha}_A + (I - F_d)\widehat{\alpha}_d.$

Substituting $\widehat{\alpha}_d$ for b_0, it is expected that Almon-modified Liu estimator has advantage according to the Almon-Liu and the Almon estimators.

3. Matrix mean square error comparisons

Bias and variance of an estimator $\tilde{\beta}$ are measured simultaneously by the MSE matrix,

$$MSE\left(\tilde{\beta}\right) = E\left[\left(\tilde{\beta} - \beta\right)\left(\tilde{\beta} - \beta\right)'\right] = V\left(\tilde{\beta}\right) + Bias(\tilde{\beta})Bias(\tilde{\beta})'$$

where

$$V\left(\tilde{\beta}\right) = E\left[\left((\tilde{\beta} - E(\tilde{\beta})\right)\left((\tilde{\beta} - E(\tilde{\beta})\right)'\right]$$

and

$$Bias\left(\tilde{\beta}\right) = E\left(\tilde{\beta}\right) - \beta.$$

For a given value of β, $\tilde{\beta}_2$ is preferred to an alternative estimator, $\tilde{\beta}_1$, when $MSE\left(\tilde{\beta}_1\right) - MSE\left(\tilde{\beta}_2\right)$ is a nonnegative definite $(n.n.d)$ matrix. Another criterion measure of goodness of an estimator is

$$smse\left(\tilde{\beta}\right) = tr\left(V\left(\tilde{\beta}\right)\right) + \left[Bias(\tilde{\beta})\right]'\left[Bias(\tilde{\beta})\right],$$

which is called as the scalar mean squared error $(smse)$ value of $\tilde{\beta}$.

If $MSE\left(\tilde{\beta}_1\right) - MSE\left(\tilde{\beta}_2\right)$ is a $n.n.d.$, then $smse\left(\tilde{\beta}_1\right) - smse\left(\tilde{\beta}_2\right) \geq 0$. The converse is not generally true (Theobald, [17]).

4. Superiority of the biased estimators under the MSE criterion

Almon-modified ridge and Almon-modified Liu estimators are biased alternatives to the Almon estimator in the presence of multicollinearity. In the following five subsections we compare Almon-modified ridge estimator with the Almon-ridge and Almon estimators. Also, Almon-modified Liu estimator is compared to the Almon-Liu and Almon estimators. In addition to these, Almon-modified ridge and Almon-Liu estimators are compared under

the MSE criterion. Canonical form of the estimators will be discussed in order to make these comparisons.

Model (1.7) can be written in canonical form

$$(4.1) \quad y = W\gamma + u, \qquad u \sim N\left(0, \sigma_u^2\right)$$

where $W = ZQ$, $\gamma = Q'\alpha$ and Q is the orthogonal matrix whose columns constitute the eigenvectors of $Z'Z$. Then

$$(4.2) \quad W'W = Q'Z'ZQ = \Lambda = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_{r+1}\right)$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{r+1} > 0$ are ordered eigenvalues of $Z'Z$. For model (4.1), we get the following representations.

Almon estimator is,

$$(4.3) \quad \widehat{\gamma}_A = \Lambda^{-1}W'y = C_1 y.$$

Almon-ridge estimator is,

$$(4.4) \quad \widehat{\gamma}_k = (\Lambda + kI)^{-1} W'y$$
$$= G_k W'y = C_2 y$$

where $G_k = (\Lambda + kI)^{-1}$. Here G_k is the diagonal and symmetric matrix.

Almon-modified ridge estimator is,

$$(4.5) \quad \widehat{\gamma}_m(k) = (\Lambda + kI)^{-1}\left(W'y + k\widehat{\gamma}_k\right)$$
$$= (\Lambda + kI)^{-1}\Lambda\widehat{\gamma}_A + k(\Lambda + kI)^{-1}\widehat{\gamma}_k$$
$$= \left[(\Lambda + kI)^{-1} + k(\Lambda + kI)^{-2}\right]W'y$$
$$= \left[G_k + kG_k^2\right]W'y = C_3 y.$$

Almon-Liu estimator is,

$$(4.6) \quad \widehat{\gamma}_d = (\Lambda + I)^{-1}(\Lambda + dI)\Lambda^{-1}W'y$$
$$= L_d \Lambda^{-1}W'y = C_4 y$$

where $L_d = (\Lambda + I)^{-1}(\Lambda + dI)$. Here L_d is diagonal and symmetric matrix.

Almon-modified Liu estimator is,

$$(4.7) \quad \widehat{\gamma}_m(d) = \left[(\Lambda + I)^{-1}(\Lambda + dI)\right]\widehat{\gamma}_A + \left[I - (\Lambda + I)^{-1}(\Lambda + dI)\right]\widehat{\gamma}_d$$
$$= L_d\widehat{\gamma}_A + (I - L_d)\widehat{\gamma}_d$$
$$= \left(2L_d - L_d^2\right)\widehat{\gamma}_A$$
$$= \left(2L_d - L_d^2\right)\Lambda^{-1}W'y = C_5 y.$$

It is evident that the above mentioned estimators are homogeneous linear. For the sake of convenience, we have an important Lemma needed in the following comparisons.

Lemma.(Trenkler, [18]). Let $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ be two homogeneous linear estimators of β such that $D = V\left(\widetilde{\beta}_1\right) - V\left(\widetilde{\beta}_2\right)$ is positive definite (p.d.).

If $Bias\left(\widetilde{\beta}_2\right)' D^{-1} Bias\left(\widetilde{\beta}_2\right) < \sigma^2$ then $MSE\left(\widetilde{\beta}_1\right) - MSE\left(\widetilde{\beta}_2\right)$ is p.d..

4.1. The comparison of Almon-modified ridge estimator and Almon estimator. In this section, we will discuss the superiority of Almon-modified ridge estimator over the Almon estimator by the MSE criterion. Also, we want to show that for any $k > 0$, we can always find k so that Almon-modified ridge estimator has less MSE as compared with Almon estimator.

As regards the performance by the variance-covariance matrix, we have the following theorem.

4.1. Theorem. *Let k be fixed and $k > 0$.*

If $b_1' D_1^{-1} b_1 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_A\right) - MSE\left(\widehat{\gamma}_m\left(k\right)\right)$ is p.d.,

where $D_1 = C_1 C_1' - C_3 C_3'$, $C_1 = \Lambda^{-1} W'$, $C_3 = \left[G_k + kG_k^2\right] W'$ and

$b_1 = Bias\left(\widehat{\gamma}_m\left(k\right)\right) = -k^2 G_k^2 \gamma$.

Proof. Using the estimators $\widehat{\gamma}_A$ and $\widehat{\gamma}_m\left(k\right)$ in (4.3) and (4.5), the variance-covariance matrix of unbiased $\widehat{\gamma}_A$ is

(4.8) $\qquad V\left(\widehat{\gamma}_A\right) = \sigma_u^2 \Lambda^{-1}$

and the variance-covariance matrix and bias of $\widehat{\gamma}_m\left(k\right)$ are respectively,

$$V\left(\widehat{\gamma}_m\left(k\right)\right) = \sigma_u^2 \left(G_k + kG_k^2\right) \Lambda \left(G_k + kG_k^2\right)$$

(4.9) $\qquad\qquad\qquad = \sigma_u^2 G_k \left(I - kG_k\right)\left(I + kG_k\right)^2,$

(4.10) $\quad Bias\left(\widehat{\gamma}_m\left(k\right)\right) = -k^2 G_k^2 \gamma$

obtained. Then using (4.9) and (4.10), MSE matrix of $\widehat{\gamma}_m\left(k\right)$ is,

(4.11) $\quad MSE\left(\widehat{\gamma}_m\left(k\right)\right) = \sigma_u^2 G_k \left(I - kG_k\right)\left(I + kG_k\right)^2 + k^4 G_k^2 \gamma\gamma' G_k^2.$

Considering the following difference from (4.8) and (4.9), we obtain

$$\Delta_1 = V\left(\widehat{\gamma}_A\right) - V\left(\widehat{\gamma}_m\left(k\right)\right) = \sigma_u^2 \left(C_1 C_1' - C_3 C_3'\right)$$

(4.12) $\qquad = \sigma_u^2 k^2 G_k \left[G_k + \Lambda^{-1} + kG_k^2\right] G_k.$

Since $\left[G_k + \Lambda^{-1} + kG_k^2\right] > 0$, $\Delta_1 > 0$, namely D_1 will be p.d. for $k > 0$. By the Lemma, the proof is completed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.2. The comparison of Almon-modified ridge estimator and Almon-ridge estimator. We have already seen in the previous section that Almon-modified ridge estimator is superior to the Almon estimator. Now, the aim is to compare the performance of Almon-modified ridge to the Almon-ridge estimator according to the MSE criterion .

In the following theorem, we have obtained sufficient condition for the Almon-modified ridge estimator to outperform the Almon-ridge estimator in terms of MSE criterion.

4.2. Theorem. *Let k be fixed and $k > 0$.*

If $b_1' D_2^{-1} b_1 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_k\right) - MSE\left(\widehat{\gamma}_m\left(k\right)\right)$ is p.d.,
where $D_2 = C_2 C_2' - C_3 C_3'$, $C_2 = G_k W'$.

Proof. Using the estimator $\widehat{\gamma}_k$ in (4.4), the variance-covariance matrix of this estimator is,

$$V\left(\widehat{\gamma}_k\right) = \sigma_u^2 G_k \Lambda G_k'$$

$$(4.13) \qquad = \sigma_u^2 \left(I - k G_k\right) G_k$$

and bias is,

$$(4.14) \quad Bias\left(\widehat{\gamma}_k\right) = -k G_k \gamma.$$

Then using (4.13) and (4.14), MSE matrix of $\widehat{\gamma}_k$ is,

$$(4.15) \quad MSE\left(\widehat{\gamma}_k\right) = \sigma_u^2 \left(I - k G_k\right) G_k + k^2 G_k \gamma \gamma' G_k'$$

obtained. Then considering the following difference from (4.13) and (4.9) we obtain

$$\Delta_2 = V\left(\widehat{\gamma}_k\right) - V\left(\widehat{\gamma}_m\left(k\right)\right) = \sigma_u^2 \left(C_2 C_2' - C_3 C_3'\right)$$

$$(4.16) \qquad = \sigma_u^2 G_k \Lambda G_k \left(2 k G_k + k^2 G_k^2\right).$$

Since $\left[2 k G_k + k^2 G_k^2\right] > 0$, $\Delta_2 > 0$. Then D_2 will be *p.d.* for $k > 0$. By the Lemma, the proof is completed. $\square$

4.3. The comparison of Almon-modified Liu estimator and Almon estimator. Li and Yang [11] compared the modified Liu estimator with OLS, Liu, ridge and modified ridge estimators according to the MSE criterion in linear regression model. Now, our goal is to compare the Almon-modified Liu estimator that we have proposed here, with the Almon estimator for the distributed lag model.

Here we show that Almon-modified Liu estimator outperform to the Almon estimator in terms of MSE criterion by the following theorem.

4.3. Theorem. *Let d be fixed and $0 < d < 1$.*

If $b_2' D_3^{-1} b_2 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_A\right) - MSE\left(\widehat{\gamma}_m\left(d\right)\right)$ is p.d.

where $D_3 = C_1 C_1' - C_5 C_5'$, $C_5 = \left(2L_d - L_d^2\right)\Lambda^{-1}W'$ and

$b_2 = Bias\left(\widehat{\gamma}_m\left(d\right)\right) = -\left(1-d\right)^2\left(\Lambda + I\right)^{-2}\gamma.$

Proof. Using the estimator $\widehat{\gamma}_m\left(d\right)$ in (4.7), the variance-covariance matrix of this estimator is,

$$(4.17) \quad V\left(\widehat{\gamma}_m\left(d\right)\right) = \sigma_u^2\left[2L_d - L_d^2\right]\Lambda^{-1}\left[2L_d - L_d^2\right]$$

and bias is,

$$(4.18) \quad Bias\left(\widehat{\gamma}_m\left(d\right)\right) = -\left(1-d\right)^2\left(\Lambda + I\right)^{-2}\gamma$$

Then using (4.17) and (4.18), MSE matrix of $\widehat{\gamma}_m\left(d\right)$ is,

$$(4.19) \quad MSE\left(\widehat{\gamma}_m\left(d\right)\right) = \sigma_u^2\left[2L_d - L_d^2\right]\Lambda^{-1}\left[2L_d - L_d^2\right] + \left(1-d\right)^4\left(\Lambda + I\right)^{-2}\gamma\gamma'\left(\Lambda + I\right)^{-2}$$

The variance-covariance matrix of $\widehat{\gamma}_m\left(d\right)$ can be rewrite in the following:

$$(4.20) \quad V\left(\widehat{\gamma}_m\left(d\right)\right) = \left[2L_d - L_d^2\right]^2 V\left(\widehat{\gamma}_A\right).$$

Here matrix $\left[2L_d - L_d^2\right]$ is the diagonal and symmetric matrix. Let B defined as

$$(4.21) \quad B = \left[2L_d - L_d^2\right]^2 = diag\left(b_1, b_2, \ldots, b_p\right).$$

We can see that $V\left(\widehat{\gamma}_m\left(d\right)\right)$ is decreasing due to the factor B in equation (4.20). The $i-th$ element of matrix B in (4.21) is

$$(4.22) \quad b_i = \left[\frac{\lambda_i^2 + 2\lambda_i + 2d - d^2}{\left(\lambda_i + 1\right)^2}\right]^2.$$

From (4.22), we have the conclusions that $\lambda_i^2 + 2\lambda_i + 2d - d^2 > 0$ and $\frac{\lambda_i^2 + 2\lambda_i + 2d - d^2}{\left(\lambda_i + 1\right)^2} < 1$ for $0 < d < 1$. Therefore, $0 < b_i < 1$ is ensured for the $i-th$ element of matrix B. Consequently, we obtain $V\left(\widehat{\gamma}_A\right) - V\left(\widehat{\gamma}_m\left(d\right)\right) > 0$, namely, D_3 is *p.d.* for $0 < d < 1$. By the Lemma, the proof is completed. $\square$

4.4. The comparison of Almon-modified Liu estimator and Almon-Liu estimator. Modified Liu estimator has smaller estimated MSE values than Liu, ridge and modified ridge estimators, respectively, in Liu and Yang [11]. In this section, we show that Almon-Liu estimator is better than Almon-modified Liu estimator according to the MSE criterion.

In the following theorem, we have obtained a sufficient condition for the Almon-Liu estimator to be superior to the Almon-modified Liu estimator in terms of MSE criterion.

4.4. Theorem. *Let d be fixed and $0 < d < 1$.*

If $b_3' D_4^{-1} b_3 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_m\left(d\right)\right) - MSE\left(\widehat{\gamma}_d\right)$ is p.d.,

where $D_4 = C_5 C_5' - C_4 C_4'$, $C_4 = L_d \Lambda^{-1} W'$, $L_d = \left(\Lambda + I\right)^{-1} \left(\Lambda + dI\right)$ and

$b_3 = Bias\left(\widehat{\gamma}_d\right) = -\left(1 - d\right)\left(\Lambda + I\right)^{-1} \gamma.$

Proof. Using the estimator $\widehat{\gamma}_d$ in (4.6), the variance-covariance matrix and the bias of this estimator are obtained respectively in the following:

$$(4.23) \quad V\left(\widehat{\gamma}_d\right) = \sigma_u^2 L_d \Lambda^{-1} L_d$$

$$(4.24) \quad Bias\left(\widehat{\gamma}_d\right) = -\left(1 - d\right)\left(\Lambda + I\right)^{-1} \gamma.$$

Then using (4.23) and (4.24), MSE matrix of $\widehat{\gamma}_d$ is,

$$(4.25) \quad MSE\left(\widehat{\gamma}_d\right) = \sigma_u^2 L_d \Lambda^{-1} L_d + \left(1 - d\right)^2 \left(\Lambda + I\right)^{-1} \gamma\gamma' \left(\Lambda + I\right)^{-1}.$$

Considering the following difference from (4.17) and (4.23), we obtain

$$\Delta_3 = V\left(\widehat{\gamma}_m\left(d\right)\right) - V\left(\widehat{\gamma}_d\right) = \sigma_u^2 \left(C_5 C_5' - C_4 C_4'\right)$$

$$= \sigma_u^2 L_d \left[\left(I + \left(1 - d\right)\left(\Lambda + I\right)^{-1}\right) \Lambda^{-1} \left(I + \left(1 - d\right)\left(\Lambda + I\right)^{-1}\right) - \Lambda^{-1}\right] L_d$$

$$(4.26) \quad = \sigma_u^2 L_d \left[2\left(1 - d\right)\Lambda^{-1}\left(\Lambda + I\right)^{-1} + \left(1 - d\right)^2 \left(\Lambda + I\right)^{-1}\Lambda^{-1}\left(\Lambda + I\right)^{-1}\right] L_d.$$

Since the last equation in (4.26) is *p.d.* for $0 < d < 1$, $V\left(\widehat{\gamma}_m\left(d\right)\right) - V\left(\widehat{\gamma}_d\right) > 0$. Therefore, $D_4 = C_5 C_5' - C_4 C_4'$ will be *p.d.* for $0 < d < 1$. By the Lemma, the proof is completed. $\square$

4.5. The comparison of Almon-modified ridge estimator and Almon-Liu estimator. Now, we compare the second order moment matrices of Almon-modified ridge and Almon-Liu estimators. Let now d be fixed for the moment, we may state the following theorem.

4.5. Theorem. *Let d be fixed and $0 < d < 1$.*
a. *If $b_3' \left(C_3 C_3' - C_4 C_4'\right)^{-1} b_3 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_m\left(k\right)\right) - MSE\left(\widehat{\gamma}_d\right)$ is p.d. for*
   $0 < k < k_j.$
b. *If $b_1' \left(C_4 C_4' - C_3 C_3'\right)^{-1} b_1 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_d\right) - MSE\left(\widehat{\gamma}_m\left(k\right)\right)$ is p.d. for*
   $0 < k_j < k$, *where $k_j = \frac{\lambda_j\left(1 - d\right)}{\lambda_j + d}$, $j = 1, 2, \ldots, r + 1$ $b_1 = Bias\left(\widehat{\gamma}_m\left(k\right)\right)$ and*
   $b_3 = Bias\left(\widehat{\gamma}_d\right).$

Proof. Using (4.9) and (4.23), we obtain

$$\Delta_3 = V\left(\widehat{\gamma}_m\left(k\right)\right) - V\left(\widehat{\gamma}_d\right) = \sigma_u^2 \left(C_3 C_3' - C_4 C_4'\right)$$

$$= \sigma_u^2 \left[\left(G_k + kG_k^2\right)\Lambda\left(G_k + kG_k^2\right) - L_d \Lambda^{-1} L_d\right].$$

Evidently, $C_3C_3' - C_4C_4'$ will be *p.d.* if and only if $\Psi_j > 0$, for all $j = 1, 2, \ldots, r+1$ where

$$\Psi_j = \frac{\lambda_j}{(\lambda_j + k)^2} - \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} + \frac{2k\lambda_j}{(\lambda_j + k)^3} + \frac{k^2\lambda_j}{(\lambda_j + k)^4}.$$

For $k > 0$, since $\frac{2k\lambda_j}{(\lambda_j + k)^3}$ and $\frac{k^2\lambda_j}{(\lambda_j + k)^4}$ are positive, a sufficient condition for $C_3C_3' - C_4C_4'$ being *p.d.* is

(4.27) $\quad \dfrac{\lambda_j}{(\lambda_j + k)^2} - \dfrac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2}$

greater than zero. So, this inequality requires than $C_3C_3' - C_4C_4'$ is *p.d.* for $0 < k < k_j$. Similarly, $C_4C_4' - C_3C_3'$ will be *p.d.* for $0 < k_j < k$ (see also Sakallioglu et al. [15]). By the Lemma, the proof is completed. $\qquad\square$

Let now k be fixed for the moment and let be $0 < k < 1$. Thus we have the following theorem.

4.6. Theorem. *Let k be fixed and $0 < k < 1$.*
a.*If $b_3'\left(C_3C_3' - C_4C_4'\right)^{-1} b_3 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_m(k)\right) - MSE\left(\widehat{\gamma}_d\right)$ is p.d. for*
$\quad 0 < d < d_j < 1$.
b.*If $b_1'\left(C_4C_4' - C_3C_3'\right)^{-1} b_1 < \sigma_u^2$, then $MSE\left(\widehat{\gamma}_d\right) - MSE\left(\widehat{\gamma}_m(k)\right)$ is p.d. for*
$\quad 0 < d_j < d < 1$ *where* $d_j = 1 - \frac{k(\lambda_j + 1)}{\lambda_j + k}, \quad j = 1, 2, \ldots, r+1.$

Proof. From the above theorem's proof, we know that $C_3C_3' - C_4C_4'$ will be *p.d.* if and only if $\Psi_j > 0$, for all $j = 1, 2, \ldots, r+1$. For fixed $k > 0$, (4.27) requires that $C_3C_3' - C_4C_4'$ is *p.d.* for $0 < d < d_j < 1$ and $C_4C_4' - C_3C_3'$ will be *p.d.* for $0 < d_j < d < 1$. By the Lemma, the proof is completed. $\qquad\square$

To illustrate our theoretical results, it is easy to use *smse* in practical applications. Therefore, the *smse* formulas for the $\widehat{\gamma}_A, \widehat{\gamma}_k, \widehat{\gamma}_m(k), \widehat{\gamma}_d$ and $\widehat{\gamma}_m(d)$ are given respectively:

(4.28) $\quad smse\left(\widehat{\gamma}_A\right) = \sigma_u^2 \sum_{i=1}^{r+1} \frac{1}{\lambda_i}$

(4.29) $\quad smse\left(\widehat{\gamma}_k\right) = \sigma_u^2 \sum_{i=1}^{r+1} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^{r+1} \frac{\gamma_i^2}{(\lambda_i + k)^2}$

(4.30) $\quad smse\left(\widehat{\gamma}_m(k)\right) = \sigma_u^2 \sum_{i=1}^{r+1} \frac{\lambda_i(\lambda_i + 2k)^2}{(\lambda_i + k)^4} + k^4 \sum_{i=1}^{r+1} \frac{\gamma_i^2}{(\lambda_i + k)^4}$

(4.31) $\quad smse\left(\widehat{\gamma}_d\right) = \sigma_u^2 \sum_{i=1}^{r+1} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2} + (1 - d)^2 \sum_{i=1}^{r+1} \frac{\gamma_i^2}{(\lambda_i + 1)^2}$

$$(4.32) \quad smse\left(\widehat{\gamma}_m\left(d\right)\right) = \sigma_u^2 \sum_{i=1}^{r+1} \frac{b_i}{\lambda_i} + (1-d)^4 \sum_{i=1}^{r+1} \frac{\gamma_i^2}{(\lambda_i+1)^4}$$

where b_i is defined in (4.22). A very important issue in the study of ridge regression is how to find an appropriate biasing parameter k. Hoerl and Kennard [6], [7], Hoerl, Kennard and Baldwin [8] and Lawless and Wang [10] suggested the following ridge parameters, that we can estimate for the model (4.1) respectively;

$$(4.33) \quad \widehat{k}_{HK} = \frac{\widehat{\sigma}_u^2}{\sum_{i=1}^{r+1} \widehat{\gamma}_i^2}$$

$$(4.34) \quad \widehat{k}_{HKB} = \frac{(r+1)\widehat{\sigma}_u^2}{\sum_{i=1}^{r+1} \widehat{\gamma}_i^2}$$

$$(4.35) \quad \widehat{k}_{LW} = \frac{(r+1)\widehat{\sigma}_u^2}{\sum_{i=1}^{r+1} \lambda_i\widehat{\gamma}_i^2}$$

where $\widehat{\gamma}$ and $\widehat{\sigma}_u^2$ are the OLS estimates of γ and σ_u^2. On the other hand Liu [13] gave the some estimates of d by analogy with the estimate of k in ridge estimate. Two of these estimates are defined as for the model (4.1):

$$(4.36) \quad \widehat{d}_{mm} = 1 - \widehat{\sigma}_u^2 \left[\sum_{i=1}^{r+1} \frac{1}{\lambda_i\left(\lambda_i+1\right)} \Big/ \sum_{i=1}^{r+1} \frac{\widehat{\gamma}_i^2}{(\lambda_i+1)^2}\right]$$

$$(4.37) \quad \widehat{d}_{CL} = 1 - \widehat{\sigma}_u^2 \left[\sum_{i=1}^{r+1} \frac{1}{\lambda_i+1} \Big/ \sum_{i=1}^{r+1} \frac{\lambda_i\widehat{\gamma}_i^2}{(\lambda_i+1)^2}\right]$$

where $\widehat{\gamma}$ and $\widehat{\sigma}_u^2$ are the OLS estimates of γ and σ_u^2.

5. A numeric example with Almon data

To illustrate our theoretical results we now consider a dataset due to Almon [2]. These data was taken in the years 1953-1967 using quarterly data where independent variable is appropriations and dependent variable is expenditures. Consideration of these data, the following results were obtained. Firstly, the smallest value of SIC was obtained 12.75 if the length of lag is $p=8$ using "*Schwartz Information Criteria* (SIC) ". Starting from the assumption that the prior information on β_i is fifth degree ($r = 5$) polynomial in (1.5), after testing the significance of the coefficient then, the optimal polynomial degree ($r = 2$) is obtained. Here, in order to obtain the form (1.7), Z matrix is obtained by means of X matrix produced by multiplying matrix A defined earlier. The condition number of Z matrix is 63.5 which imply the existence of highly multicollinearity in the

data set. In this case, the results of Almon method that based on the OLS will not be appropriate.

Theoretical comparisons for the alternative estimators to the Almon estimator have been made in terms of the MSE criterion. Also, *smse* formulas have been given for these estimators. Using *smse* is generally the most convenient for applications or simulation studies. Then, we decided that which one is the best estimator for distributed lag models. For this data, we find the following results:

(a) The eigenvalues of $Z'Z$: (0.0007, 0.0634, 2.9359)

(b) The Almon estimates of

$$\alpha : (\widehat{\alpha}_A)' = (0.0962,\ 0.0320,\ -0.0052)$$

$$\widehat{\beta}_A = (A\widehat{\alpha}_A)' = (0.096,\ 0.123,\ 0.140,\ 0.146,\ 0.142,\ 0.127,\ 0.102,\ 0.067,\ 0.021)\,.$$

(c) The estimate of $\sigma^2 : \widehat{\sigma}_u^2 = 0.0164$

The 3×3 matrix Q is the matrix of normalized eigenvectors, Λ is a 3×3 diagonal matrix of eigenvalues of $Z'Z$ such that $Z'Z = Q\Lambda Q'$. Then, $W = ZQ$ and $\gamma = Q'\alpha$ so that, $y = Z\alpha + u = W\gamma + u$, where

$$Q = \begin{bmatrix} -0.2478 & -0.7818 & 0.5722 \\ 0.7934 & 0.1751 & 0.5829 \\ -0.5559 & 0.5985 & 0.5769 \end{bmatrix}$$

and

$$W'W = \Lambda = \begin{bmatrix} 0.0007 & 0 & 0 \\ 0 & 0.0634 & 0 \\ 0 & 0 & 2.9359 \end{bmatrix}$$

In orthogonal coordinates the OLS estimator of the regression coefficients is

$$\widehat{\gamma} = \Lambda^{-1}W'y = [1.2297,\ -1.0754,\ 0.5580]'$$

obtained. Using the equations in (4.33)-(4.35) estimators of k obtained for the evaluate the estimated *smse* values of Almon-ridge and Almon-modified ridge estimators. Then, for the practical purposes various values of k and the corresponding estimated *smse* values of the estimators are shown in Table 1. In Figure 1, the graph of estimated *smse* values of the Almon-ridge and Almon-modified ridge estimators is illustrated for the range of k values that performance of Almon-modified ridge estimator is better than Almon-ridge estimator.

Let us consider the Almon-Liu and Almon-modified Liu estimators various values of d and the corresponding estimated *smse* values of the estimators are shown in Table 2.

Also, the performances of Almon-Liu and Almon-modified Liu estimators are illustrated for the various values of d in Figure 2.

In Table 3, we compared the Almon-modified ridge, Almon-modified Liu and also Almon-Liu estimators and comparisons are shown on the graph for the common values of k and d in Figure 3.

Table 1. Estimated $smse$ values of Almon, Almon-ridge and Almon-modified ridge estimators

	$\widehat{smse}\left(\widehat{\gamma}_A\right)$	$\widehat{smse}\left(\widehat{\gamma}_k\right)$	$\widehat{smse}\left(\widehat{\gamma}_m\left(k\right)\right)$
$k=0$	23.6928	23.6928	23.6928
$k_{HK}=0.0055$	23.6928	1.7206	2.2609
$k=0.01$	23.6928	1.6411	1.7840
$k_{HKB}=0.0165$	23.6928	1.6481	1.6743
$k=0.02$	23.6928	1.6599	1.6605
$k=0.03$	23.6928	1.7002	1.6524
$k_{LW}=0.0498$	23.6928	1.7855	1.6649
$k=0.1$	23.6928	1.9700	1.7440
$k=0.2$	23.6928	2.1898	1.9286
$k=0.3$	23.6928	2.3086	2.0675
$k=0.4$	23.6928	2.3823	2.1662
$k=0.5$	23.6928	2.4328	2.2384
$k=0.6$	23.6928	2.4699	2.2931
$k=0.7$	23.6928	2.4985	2.3359
$k=0.8$	23.6928	2.5215	2.3702
$k=0.9$	23.6928	2.5406	2.3984
$k=1$	23.6928	2.5569	2.3310
$k=2$	23.6928	2.6510	2.4949

When we compare Almon, Almon-ridge and Almon-modified ridge estimators, we observe that as k increases, Almon-modified ridge estimator always gives better performance than the other estimators. On the other hand, the performance of Almon-ridge estimator is better than Almon estimator with in the wide range k values. The plot of $\widehat{smse}\left(\widehat{\gamma}_k\right)$ and $\widehat{smse}\left(\widehat{\gamma}_m\left(k\right)\right)$ vs. k in the interval [0,1] has been presented in Fig.1. This figure indicates that $\widehat{smse}\left(\widehat{\gamma}_k\right)$ and $\widehat{smse}\left(\widehat{\gamma}_m\left(k\right)\right)$ increase as k increases. The Almon-modified ridge estimator dominates Almon-ridge estimator when $k > 0.02$. These findings have supported the results in Section 4.1 and 4.2.

Considering the performance of the other alternative estimators we can see that Almon-modified Liu estimator outperforms to the Almon-Liu and Almon estimator for all values of d satisfying $0 < d < 1$. The plot of $\widehat{smse}\left(\widehat{\gamma}_d\right)$ and $\widehat{smse}\left(\widehat{\gamma}_m\left(d\right)\right)$ has been presented in Fig.2. This figure indicates that $\widehat{smse}\left(\widehat{\gamma}_d\right)$ and $\widehat{smse}\left(\widehat{\gamma}_m\left(d\right)\right)$ increase as
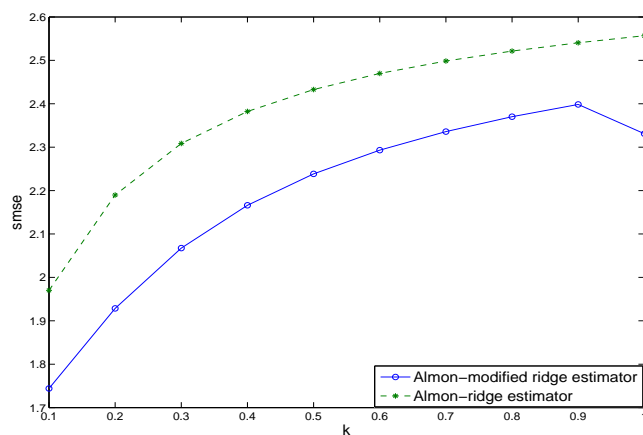
Figure 1. Estimated *smse* of Almon-ridge and Almon-modified ridge estimators versus k

Table 2. Estimated *smse* values of Almon, Almon-Liu and Almon-modified Liu estimators

	$\widehat{smse}(\widehat{\gamma}_A)$	$\widehat{smse}(\widehat{\gamma}_d)$	$\widehat{smse}(\widehat{\gamma}_m(d))$
$d = 0$	23.6928	0.1721	0.0161
$d = 0.001$	23.6928	0.1718	0.0160
$d = 0.01$	23.6928	0.1688	0.0156
$d = 0.1$	23.6928	0.1403	0.0119
$d = 0.2$	23.6928	0.1118	0.0089
$d = 0.3$	23.6928	0.0867	0.0069
$d = 0.4$	23.6928	0.0650	0.0056
$d = 0.5$	23.6928	0.0467	0.0049
$d = 0.6$	23.6928	0.0318	0.0045
$d_{CL} = 0.712$	23.6928	0.0192	0.0043
$d = 0.8$	23.6928	0.0122	0.0042
$d = 0.9$	23.6928	0.0076	0.0042

d increases and large value of d Almon-modified Liu estimator dominates the Almon-Liu estimator. On the other hand the increasing of $\widehat{smse}(\widehat{\gamma}_m(d))$ is slowly than the $\widehat{smse}(\widehat{\gamma}_d)$.

Finally, comparison of the three estimators is illustrated in Figure 3. It can be seen that not only Almon-modified Liu estimator but also Almon-Liu estimator outperforms Almon-modified ridge estimator in Figure 3.

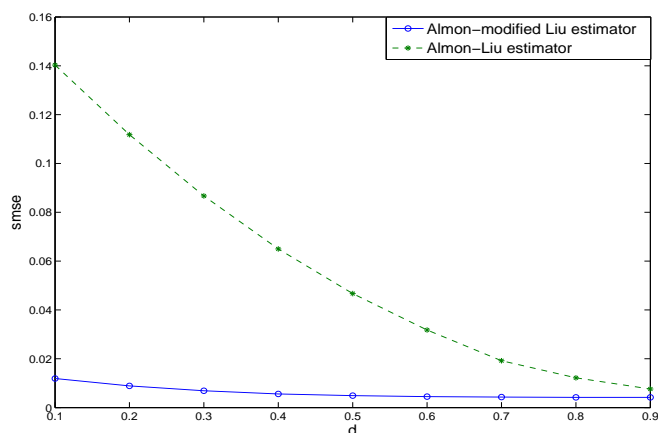From Table 3 and Figure 3, we can also obtain the following conclusions:

Figure 2. Estimated *smse* of Almon-Liu and Almon-modified Liu estimators versus d

Table 3. Comparisons between Almon-modified ridge, Almon Liu and Almon-modified Liu estimators in *smse* sense

$k = d$	$\widehat{smse}\,(\widehat{\gamma}_d)$	$\widehat{smse}\,(\widehat{\gamma}_m\,(k))$	$\widehat{smse}\,(\widehat{\gamma}_m\,(d))$
0.01	0.1688	1.7840	0.0156
0.1	0.1403	1.7440	0.0119
0.2	0.1118	1.9286	0.0089
0.3	0.0867	2.0675	0.0069
0.4	0.0650	2.1662	0.0056
0.5	0.0467	2.2384	0.0049
0.6	0.0318	2.2931	0.0045
0.7	0.0192	2.3359	0.0043
0.8	0.0122	2.3702	0.0042
0.9	0.0076	2.3984	0.0042

(i) Let $d = 0.1$ be fixed. We get values of k_j by using Theorem 4.5.

$$k_j : 0.8704, 0.3492, 0.0062.$$

Comparing $\widehat{smse}\,(\widehat{\gamma}_{d=0.1}) = 0.1403$ with $\widehat{smse}\,(\widehat{\gamma}_m\,(k = 0.005)) = 2.4019$ for $0 < k < 0.0062$, we see that $\widehat{\gamma}_d$ has a smaller estimated *smse* value than $\widehat{\gamma}_m\,(k)$ (see also Figure 3). Comparing $\widehat{smse}\,(\widehat{\gamma}_{d=0.1}) = 0.1403$ with $\widehat{smse}\,(\widehat{\gamma}_m\,(k = 0.9)) = 2.3984$ is obtained for $0 < 0.8704 < k$. Since the sufficient condition in Theorem 4.5.(*b*) is not satisfied, $\widehat{\gamma}_m\,(k)$ does not have estimated *smse* value than $\widehat{\gamma}_d$.
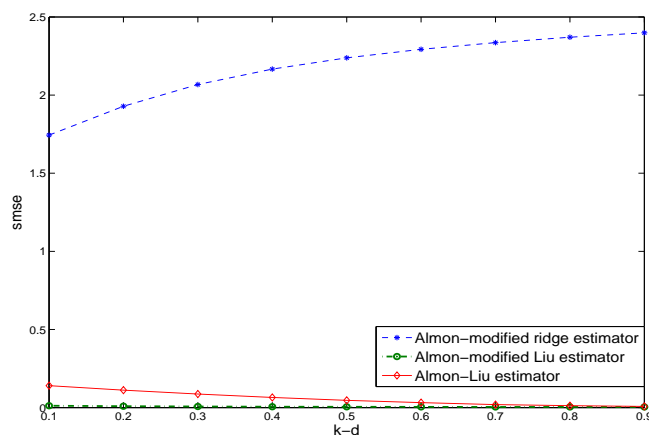
Figure 3. Estimated *smse* of Almon-modified ridge, Almon-Liu, Almon-modified Liu estimators versus $k - d$

(ii) Let $d = 0.9$ be fixed. By using Theorem 4.5 k_j values are obtained as

$k_j : 0.00765, 0.00658, 0.000077.$

Comparing $\widehat{smse}\left(\widehat{\gamma}_{d=0.9}\right) = 0.0076$ with $\widehat{smse}\left(\widehat{\gamma}_m\left(k = 0.00007\right)\right) = 23.235$ for $0 < k < 0.000077$, we see that $\widehat{\gamma}_d$ has a smaller estimated *smse* value than $\widehat{\gamma}_m\left(k\right)$ (see also Figure 3). Comparing $\widehat{smse}\left(\widehat{\gamma}_{d=0.9}\right) = 0.0076$ with $\widehat{smse}\left(\widehat{\gamma}_m\left(k = 0.008\right)\right) = 1.8979$ is obtained for $0 < 0.00765 < k$. Since the sufficient condition in Theorem 4.5.(b) is not satisfied, $\widehat{\gamma}_m\left(k\right)$ does not have smaller estimated *smse* value than $\widehat{\gamma}_d$.

(iii) Let $k = 0.2$ be fixed. We get values of d_j by using Theorem 4.6.

$d_j : 0.749, 0.1926, 0.0028.$

Comparing $\widehat{smse}\left(\widehat{\gamma}_m\left(k = 0.2\right)\right) = 1.9286$ with $\widehat{smse}\left(\widehat{\gamma}_{d=0.002}\right) = 0.1715$ for $0 < d < 0.0028 < 1$. So $\widehat{\gamma}_d$ has a smaller estimated *smse* value than $\widehat{\gamma}_m\left(k\right)$ as it is indicated in (a) part of the Theorem 4.6 (see also Figure 3). On the other hand, comparing $\widehat{smse}\left(\widehat{\gamma}_m\left(k = 0.2\right)\right) = 1.9286$ with $\widehat{smse}\left(\widehat{\gamma}_{d=0.8}\right) = 0.0122$ is obtained for $0 < 0.749 < d < 1$. Since the sufficient condition in Theorem 4.6.(b) is not satisfied, $\widehat{\gamma}_m\left(k\right)$ does not have smaller estimated *smse* value than $\widehat{\gamma}_d$.

(iv) Let $k = 0.8$ be fixed. By using Theorem 4.6 d_j values are obtained as

$d_j : 0.1572, 0.0147, 0.0002.$

Comparing $\widehat{smse}\left(\widehat{\gamma}_m\left(k=0.8\right)\right) = 2.3702$ with $\widehat{smse}\left(\widehat{\gamma}_{d=0.0001}\right) = 0.1721$ for $0 < d < 0.0002 < 1$. So $\widehat{\gamma}_d$ has a smaller estimated $smse$ value than $\widehat{\gamma}_m\left(k\right)$ as it is indicated in (a) part of the Theorem 4.6 (see also Figure 3). Beside this, comparing $\widehat{smse}\left(\widehat{\gamma}_m\left(k=0.8\right)\right) = 2.3712$ with $\widehat{smse}\left(\widehat{\gamma}_{d=0.2}\right) = 0.1118$ is obtained for $0 < 0.1572 < d < 1$. Since the sufficient condition in Theorem 4.6.(b) is not satisfied, $\widehat{\gamma}_m\left(k\right)$ does not have estimated $smse$ value than $\widehat{\gamma}_d$.

6. Conclusions

In this study, we have compared theoretical performances of Almon-ridge $(\widehat{\gamma}_k)$, Almon-modified ridge $(\widehat{\gamma}_m\left(k\right))$, Almon-Liu $(\widehat{\gamma}_d)$, Almon-modified Liu $(\widehat{\gamma}_m\left(d\right))$ estimators to the Almon $(\widehat{\gamma}_A)$ estimator according to the MSE criterion with using some theorems. These alternative estimators showed quite good performance to the Almon estimator. Also, some of the alternative estimators compared with each other. The performances of the estimators depends on biasing parameters k and d. To see more detailed results of the comparisons we plotted estimated $smse$ values of these estimators using k and d values in Figure 1-3.

Liu and Yang [11] showed with the increasing of the levels of multicollinearity, the $smse$ values of ridge, Liu, modified ridge and modified Liu estimators are decreasing in general for the linear regression model. Moreover, they showed that the $smse$ values of these estimators outperformed to the OLS estimator for all cases. Also, for most cases, modified Liu estimator has smaller $smse$ values than those of the Liu, ridge, and modified ridge estimator, respectively. In this study, we find similar results for the distributed lag models. Theoretical results suggested that, for an appropriate value of k and d Almon-modified ridge and Almon-modified Liu estimator give better estimates than the other alternative estimators in terms of MSE criterion for the distributed lag models.

The theoretical section is supported by a numerical example based on widely analyzed Almon [2] dataset. Almon-modified Liu estimator has been showed as the best estimator in distributed lag models.

Acknowledgement

References

[1] Akdeniz, F. and Kaçıranlar, S. *On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE*, Communications in Statistics Theory and Methods **24** (7), 1789-1797, 1995.

[2] Almon, S. *The distributed lag between capital appropriations and expenditures*, Econometrica, 178-196, 1965.

[3] Chanda, A.K. and Maddala, G.S. *Ridge estimators for distributed lag models*, Communications in Statistics Theory and Methods **13** (2), 217-225, 1984.

[4] Fisher, I. *Income in theory and income taxation practice*, Econometrica, **5**, 1-55, 1937.

[5] Gujarati, D. N. *Basic Econometrics*, New York:McGraw-Hill, 1999.

[6] Hoerl, A.E. and Kennard, R.W. *Ridge regression:biased estimation for nonorthogonal problems*, Technometrics, **12**, 55-67, 1970a.

[7] Hoerl, A.E. and Kennard, R.W. *Ridge regression:application for nonorthogonal problems*, Tecnometrics, **12**, 69-82, 1970b.

[8] Hoerl, A.E., Kennard, R.W., Baldwin, K. F. *Ridge regression: some simulation*, Communication in Statistics, **4** (2), 105-123, 1975.

[9] Kaçıranlar, S. *Alternative estimators for distributed lag models*, Journal of Statistical Research, **9** (1), 1-10, 2010.

[10] Lawless, J.F. and Wang, P.A. *Simulation study of ridge and other regression estimators*, Communication in Statistics Theory and Methods, **A5** (4), 307-323, 1976.

[11] Li Y. and Yang H. *A new Liu-type estimator in linear regression model*, Statistical Papers **53** (2), 427-437, 2012.

[12] Lindley, D. V. and Smith, A. F. M. *Bayes estimates for the linear model*, Journal of the Royal Statistical Society, B Series, 1-41, 1972.

[13] Liu, K. *A new class of biased estimate in linear regression*, Communication in Statistics Theory and Methods, **22** (2), 393-402, 1993.

[14] Maddala, G. S. *Ridge estimator for distributed lag models*, NBER Working Paper Series, **69**, 1974.

[15] Sakallıoglu, S., Kaçıranlar S., Akdeniz F. *Mean square error comparisons of some biased regression estimators*, Communication in Statistics Theory and Methods, **30** (2), 347-361, 2001.

[16] Swindel, B.F. *Good ridge estimators based on prior information*, Communication in Statistics Theory and Methods, **A5** (11), 1065-1075, 1976.

[17] Theobald, C.M. *Generalizations of mean square error applied to ridge regression*, J.Roy.Stat.Soc.Ser. **B** (36), 103-106, 1974.

[18] Trenkler G. *Generalized mean squared error comparisions of biased regression estimators*, Communication in Statistics Theory and Methods, **A9** (12), 1247-1259, 1980.

[19] Vinod H. D. and Ullah A. *Recent advances in regression methods*, Marcel Dekker, Inc, 226-238. 1981.

[20] Yeo, S. J. and Trivedi, P. K. *On using ridge-type estimators for a distributed lag model*, Oxford Bulletin of Economics and Statistics, **51** (1), 85-90, 1989.

[21] Yurdakul, F. *The econometrics analysis of relationships between of cotton production and prices: Koyck -Almon Approach*, Journal of Faculty of Economics and Administrative Sciences of Cukurova University, **8** (1), 341-353, 1998.