Journal of Sciences, Islamic Republic of Iran 29(1): 67 - 78 (2018) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Diagnostic Measures in Ridge Regression Model with AR(1) Errors under the Stochastic Linear Restrictions

A. Zaherzadeh, A. R. Rasekh* and B. Babadi

Department of Statistics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Islamic Republic of Iran

Received: 9 February 2017 / Revised: 10 October 2017 / Accepted: 20 November 2017

Abstract

Outliers and influential observations have important effects on the regression analysis. The goal of this paper is to extend the mean-shift model for detecting outliers in case of ridge regression model in the presence of stochastic linear restrictions when the error terms follow by an autoregressive AR(1) process. Furthermore, extensions of measures for diagnosing influential observations are derived. A numerical example of a real data set is used to illustrate the findings. Finally, a simulation study is conducted to evaluate the performance of the proposed procedure and measures. Results of this study show the efficiency of the proposed mean-shift outlier model for the proposed model. Also, the study resulted in some findings about the behavior of suggested measures for the specified model. In fact, these measures are affected by the degree of collinearity and the size of autocorrelation.

Keywords: Ridge regression; Stochastic linear rRestrictions; Autocorrelated error terms; Influential analysis; Mean-shift outlier model.

Introduction

Departures from underlying assumptions in the regression analysis can result in difficulties for ordinary least squares (OLS) estimates of the model parameters. The first departure is collinearity that occurs when two or more regressors are almost linearly dependent. Collinearity can make parameters estimates to have large variances so the confidence intervals of regression coefficients get wider and the p-values would be misleading; consequently, we have doubt about the necessity of a variable to enter the model. A number of remedies have been suggested to overcome these problems. Hoerl and Kennard [1] introduced the ridge regression estimator, as a biased estimator to improve the mean square error (MSE) of the parameter estimators against the OLS approach (See Belsley et al. [2] for more details). Belsley et al. [2] and Rao et al. [3] considered using prior information about parameters in the form of exact or stochastic linear restrictions that leads to decrease the effect of collinearity and improve the MSE of the estimator. In fact, the exact linear restriction refers to a deterministic relation between parameters while the stochastic linear restriction is referred to stochastic situation.

Sarkar [4] and Özkale [5] combined the ridge approach with exact linear restrictions and stochastic linear restrictions, respectively, to obtain restricted ridge regression estimator and stochastic restricted ridge regression estimator in order to gain advantages of the two approaches.

The second departure from underlying assumptions

* Corresponding author: Tel: +986133331043; Fax: +986133331043; Email: Rasekh_a@scu.ac.ir

occurs when the error terms are autocorrelated such as in the form of AR(1). Bayhan and Bayhan [6] considered the estimation of parameters in the linear regression model with both, autocorrelated errors and the existence of collinearity between the explanatory variables and Alkhamisi [7] concentrated on the ridge parameters in the same model.

The approach of using prior information for decreasing the effects of collinearity in case of autocorrelated errors is also of interest. Özkale [5] introduced a generalized least squares (GLS) estimator by using the idea of restricted ridge regression estimator proposed by Groβ [8] in the presence of exact linear restrictions. Alheety and Golam Kibria [9] proposed a generalized estimator as an alternative to the stochastic restricted ridge estimator when the error terms are not independent.

On the other hand, regression analysis can suffer from anomalous data including outliers or influential observations. Belsley et al. [2], and Chatterjee and Hadi [10] considered this subject for models with uncorrelated errors under constant variance error terms.

The case deletion method is the most common approach for influential analysis which compares the estimates of coefficients and the model fits for full model and reduced model (*DFBETA* and *DFFIT*).

Roy and Curia [11] studied the case deletion effect of individual observation in the autocorrelated errors model. Özkale and Açar [12] formulated *DFBETA* in the GLS estimators when the error terms are of the form AR(1) and AR(2). Then, Açar and Özkale [13] introduced influence measures for autocorrelated ridge regression model. Ghapani et al. [14] discussed the detection of outliers and influential observations in linear ridge measurement error models with stochastic linear restrictions.

Wang [15] and Rao et al. [3] discussed the method of mean-shift outlier model. This method is based on adding an ancillary parameter for each suspicious observation to the model and handling a hypothesis testing for significance of the parameter. Wen and Wing [16] discussed the subject of mean-shift outlier model in the general weighted regression and Pan and Xiong [17] extended the subject to the case of presence of collinearity and Ghapani et al. [18] studied this method in case of linear measurement error models with stochastic linear restrictions.

This paper concentrates on the outliers and influence measures through the mean-shift outlier model and case deletion methods using stochastic linear restrictions in the ridge regression model when the error terms are autocorrelated (of the type AR(1)). The above mentioned issues are organized as follows.

In the preliminary section the restricted ridge regression model and AR(1) error terms are introduced. In the main results section the method of mean-shift outlier model and the case deletion method for the ridge regression model using stochastic linear restrictions are discussed when the error terms are AR(1). Furthermore, the performance of the proposed methods will be illustrated using a real dataset and a simulation study through some tables and figures.

Preliminaries

Consider the linear regression model

$$y = X \quad + u \tag{1}$$

where *y* is the $n \times 1$ vector of responses, *X* is $n \times p$ matrix of explanatory variables, is $p \times 1$ vector of regression parameters and *u* is $n \times 1$ error vector with $E(u) = 0_{n\times 1}$ and $Var(u) = \pm^2 \Omega$, where Ω is an $n \times n$ positive definite (p.d) matrix.

In case of presence of collinearity, an approach for the estimation of parameters is to combine the method of stochastic linear restrictions with ridge regression approach (Özkale [5]). This approach is conducted through an extension of mixed estimation method introduced by Theil and Goldberger [19].

They assumed both ridge restrictions proposed by Troskie et al. [20] as $0_{p\times 1} = \sqrt{k} I_p \quad + ¦$ and stochastic linear restrictions of the form

$$r = R \quad + \mu \tag{2}$$

simultaneously. In these restrictions, $¦ \sim N_p(0, \pm^2 I)$ is a random vector, independent of u, *r* is an $m \times 1$ vector, *R* is an $m \times p$ prior information matrix of rank $m < p$ and $\mu \sim N_m(0, \pm^2 W)$ is a random vector, independent of u and $¦$, where *W* is an $n \times n$ known p.d matrix.

The mixed model is defined as follows:

$$\begin{pmatrix} y \\ 0 \\ r \end{pmatrix} = \begin{pmatrix} X \\ \sqrt{k} I_p \\ R \end{pmatrix} \quad + \begin{pmatrix} u \\ ¦ \\ \mu \end{pmatrix}$$

or

$$y_m = X_m \quad + £_m \tag{3}$$

where $Var(£_m) = \pm^2 \Lambda$ and $\Lambda = diag(\Omega, I, W)$ is a p.d matrix. There exist nonsingular matrices P and T so that $\Omega^{-1} = P'P$ and $W^{-1} = T'T$ (see Seber [23], p.461). Consequently a nonsingular matrix $L = diag(P, I, \mathrm{T})$ will be available such that $\Lambda^{-1} = L'L$.

Lemma 1: For the *mixed stochastic restricted ridge regression (MSR)* model (3), the *MSR* estimator of is given by:

$$\hat{}_{MSR}(k) = \left(X'\Omega^{-1}X + R'W^{-1}R + kI_p\right)^{-1}\left(X'\Omega^{-1}\tilde{y} + R'W^{-1}r\right)$$

Proof: Pre-multiplying (3) by L and defining $\tilde{y}_m = Ly_m$, $\tilde{X}_m = LX_m$ and $\tilde{£}_m = L£_m$, the model changes to $\tilde{y}_m = \tilde{X}_m \quad + \tilde{£}_m$, where $Var(\tilde{£}_m) = ±^2 I$. Hence the OLS estimator of is readily derived as follows:

$$\begin{aligned}
\hat{}_{MSR} &= \left(\tilde{X}'_m \tilde{X}_m\right)^{-1} \tilde{X}'_m \tilde{y}_m \\
&= \left(X'_m \Lambda^{-1} X_m\right)^{-1} X'_m \Lambda^{-1} y_m \\
&= \left(\begin{bmatrix} X' & \sqrt{k} I_p & R' \end{bmatrix} \begin{bmatrix} \Omega^{-1} & 0 & 0 \\ 0 & I_p & 0 \\ 0 & 0 & W^{-1} \end{bmatrix} \begin{bmatrix} X \\ \sqrt{k} I_p \\ R \end{bmatrix} \right)^{-1} \times \\
&\qquad \begin{bmatrix} X' & \sqrt{k} I_p & R' \end{bmatrix} \begin{bmatrix} \Omega^{-1} & 0 & 0 \\ 0 & I_p & 0 \\ 0 & 0 & W^{-1} \end{bmatrix} \begin{bmatrix} y \\ 0 \\ r \end{bmatrix} \\
&= \left(X'\Omega^{-1}X + R'W^{-1}R + kI_p\right)^{-1}\left(X'\Omega^{-1}y + R'W^{-1}r\right)
\end{aligned}$$

Replacing Ω^{-1} with $P'P$ in $\hat{}_{MSR}$ and letting $\tilde{X} = PX$ and $\tilde{y} = Py$ we have

$$\hat{}_{MSR}(k) = \left(\tilde{X}'\tilde{X} + R'W^{-1}R + kI_p\right)^{-1}\left(\tilde{X}'\tilde{y} + R'W^{-1}r\right). \quad (4)$$

This estimator is the same as one given by Özkale [5].

In the model (1) with the autocorrelated error terms, the p.d matrix Ω is given by (See Firinguetti [21])

$$\Omega = \frac{1}{1-°^2}\begin{bmatrix} 1 & ° & \dots & °^{n-1} \\ ° & 1 & \dots & °^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ °^{n-1} & °^{n-2} & \dots & 1 \end{bmatrix}. \quad (5)$$

It should be noticed that for each element of vector y in (1) denoted by $y_t = x_t \quad + u_t \quad t = 1, ..., n$, the error term u_t follows an AR(1) process, so $u_t = ° u_{t-1} + £_t$ where $|°| < 1$, $E(£_t) = 0$ and $E(£_t^2) = ±^2$ for each t and $E(£_i £_j) = 0 \quad \forall i \neq j$. The inverse matrix of Ω is

$$\Omega^{-1} = \begin{bmatrix} 1 & -° & 0 & \cdots & 0 & 0 \\ -° & 1+°^2 & -° & \cdots & 0 & 0 \\ & & \cdots & & & \\ 0 & 0 & 0 & \cdots & 1+°^2 & -° \\ 0 & 0 & 0 & \cdots & -° & 1 \end{bmatrix}$$

and there exists an $n \times n$ nonsingular matrix P, such that $\Omega^{-1} = P'P$ (Judge et al. [22]) as follows

$$P = \begin{bmatrix} \sqrt{1-°^2} & 0 & 0 & \cdots & 0 & 0 \\ -° & 1 & 0 & \cdots & 0 & 0 \\ 0 & -° & 1 & \cdots & 0 & 0 \\ & & \cdots & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -° & 1 \end{bmatrix}.$$

At this point we assume $°$ is known, but in the unknown case letting $\hat{u}_t = y_t - x_t \hat{}$ which are OLS residuals, $°$ can be estimated from $\hat{°} = \sum_{t=1}^{n-1} \hat{u}_t \hat{u}_{t+1} \Big/ \sum_{t=1}^{n} \hat{u}_t^2$.

Results

Mean-shift outlier method

Suppose that the ith observation is suspicious as an outlier. It is common to use mean-shift outlier model in order to test whether this observation is an outlier. Based on this method the data are arranged so that the ith observation is moved to the end of data set. The mean-shift outlier model with the rearranged data for model (1) is written as $y = X \quad + Z¥ + £$ where $Z = (0, ..., 0, 1)'$ is an $n \times 1$ vector and $¥$ is an unknown shift parameter. To check the significance of $¥$ we test the hypothesis $H_\circ : ¥ = 0$ versus $H_1 : ¥ \neq 0$ (see Seber [23] and Rao et al. [3]).

However, in case of autocorrelated error model, displacement of any observation changes the autocorrelated structure of errors into a disorder state. Therefore, first, the mean-shift outlier model is constructed without any changes in the order of data and $\overline{}_i$ is introduced as a column vector of one in the ith element and zeros elsewhere, so the model is

$$\begin{pmatrix} y \\ 0 \\ r \end{pmatrix} = \begin{pmatrix} X & \overline{}_i \\ \sqrt{k} I_p & 0 \\ R & 0 \end{pmatrix} \begin{pmatrix} \ \\ ¥ \end{pmatrix} + \begin{pmatrix} u \\ ¦ \\ µ \end{pmatrix}$$

or

$$y_m = B_m ¢ + £_m . \qquad (6)$$

Pre-multiplying (6) by matrix L gets $\tilde{y}_m = \tilde{B}_m ¢ + \tilde{£}_m$ where $\tilde{B}_m = LX_m$ and $Var(\tilde{£}_m) = ±^2 I$. The changes of the model elements are such that ith and ($i+1$)th transformed observations involve the shift parameter ¥ . We proposed to displace these two elements to the end of matrix B_m . Therefore, without any changes in the model structure the final mean-shift outlier model can be written as follows:

$$\begin{pmatrix} \tilde{y}_{(i,i+1)} \\ 0 \\ \tilde{r} \\ \tilde{y}_{i,i+1} \end{pmatrix} = \begin{pmatrix} \tilde{X}_{(i,i+1)} & 0 \\ \sqrt{k} I_p & 0 \\ \tilde{R} & 0 \\ \tilde{X}_{i,i+1} & I_2 \end{pmatrix} \begin{pmatrix} ¥ \\ -°¥ \end{pmatrix} + \begin{pmatrix} \tilde{u}_{(i,i+1)} \\ ¦ \\ \tilde{µ} \\ \tilde{u}_{i,i+1} \end{pmatrix},$$

or

$$y_f = X_f ¢_f + £_f = A \quad + Z_t ¥_t + £_f . \qquad (7)$$

where $A' = \left(\tilde{X}'_{(i,i+1)} \quad \sqrt{k} I_p \quad \tilde{R}' \quad \tilde{x}_i \quad \tilde{x}_{i+1}\right)$, $Z_t = \begin{pmatrix} 0_{q\times 2} \\ I_2 \end{pmatrix}$, and $¥_t = \begin{pmatrix} ¥ \\ -°¥ \end{pmatrix}$ with $q = (n+p+m-2)$. Also, $\tilde{u} = Pu$, $\tilde{r} = Tr$, $\tilde{R} = TR$, $\tilde{µ} = Tµ$, $\tilde{y}_{i,i+1} = \begin{pmatrix} \tilde{y}_i \\ \tilde{y}_{i+1} \end{pmatrix}$, $\tilde{X}_{i,i+1} = \begin{pmatrix} \tilde{x}'_i \\ \tilde{x}'_{i+1} \end{pmatrix}$, $\tilde{u}_{i,i+1} = \begin{pmatrix} \tilde{u}_i \\ \tilde{u}_{i+1} \end{pmatrix}$, in which $\tilde{x}'_i$, $\tilde{y}_i$ and $\tilde{u}_i$ are the ith rows of $\tilde{X}$, $\tilde{y}$ and $\tilde{u}$, respectively. Furthermore, $\tilde{y}_{(i,i+1)}$, $\tilde{X}_{(i,i+1)}$ and $\tilde{u}_{(i,i+1)}$ are $\tilde{y}$, $\tilde{X}$ and $\tilde{u}$ with deleting the ith and (i+1)th rows, respectively. Throughout the rest of paper subscript (i) indicates that the ith row is deleted.

Remark 1: Testing the hypothesis of $H_{\circ 2} : ¥_t = \begin{pmatrix} ¥ \\ -°¥ \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for model (7) is equivalent to the hypothesis $H_{\circ 1} : ¥ = 0$ for the original mean-shift model (6).

Proof: If $H_{\circ 2}$ is not rejected, then ¥ is absolutely equal to zero and $H_{\circ 1}$ is not rejected too. If $H_{\circ 2}$ is rejected, then $¥_t = \begin{pmatrix} ¥ \\ -°¥ \end{pmatrix} = \begin{pmatrix} 1 \\ -° \end{pmatrix} ¥ \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, but from the basic definitions $° \neq 0$ so that $¥ \neq 0$ and $H_{\circ 1}$ is rejected too.

Applying the method used by Seber [23] to test $H_{\circ 2}$, under the null hypothesis of $¥_t = 0$, let $\tilde{H}$ be the hat matrix of model

$$y_f = A \quad + £_f$$
(8)

and then it is partitioned as

$$\tilde{H} = A(A'A)^{-1}A' = \begin{pmatrix} \underset{q\times q}{\tilde{H}_{11}} & \underset{q\times 2}{\tilde{H}_{12}} \\ \underset{2\times q}{\tilde{H}_{21}} & \underset{2\times 2}{\tilde{H}_{22}} \end{pmatrix}.$$

Then the residuals of model (8) can be written conformably with $\tilde{H}$ as follows:

$$\tilde{e} = (I_{q+2} - \tilde{H}) y_f = G y_f = \begin{pmatrix} (I_q - \tilde{H}_{11}) y_{f1} - \tilde{H}_{12} \tilde{y}_{i,i+1} \\ (I_2 - \tilde{H}_{22}) \tilde{y}_{i,i+1} - \tilde{H}_{21} y_{f1} \end{pmatrix} = \begin{pmatrix} \tilde{e}_1 \\ \tilde{e}_2 \end{pmatrix}$$
,

in which $y'_{f1} = \left(\tilde{y}'_{(i,i+1)} \quad 0'_{1\times p} \quad \tilde{r}'\right)$. The OLS estimate of $¥_t$ is

$$\hat{¥}_t = (Z'_t G Z_t)^{-1} Z'_t G y_f = \left(I_2 - \tilde{H}_{22}\right)^{-1} \tilde{e}_2 .$$

For testing $H_{\circ 2} : ¥_t = 0$, firstly the numerator of the F-test is calculated (Seber [23]–Theorem 3.6) as $\hat{¥}'_t Z'_t G y_f = \tilde{e}'_2 \left[I_2 - \tilde{H}_{22} \right]^{-1} \tilde{e}_2$ and the denominator is derived as $y'_f G y_f - \hat{¥}'_t Z'_t G y_f$ that equals to $RSS - \tilde{e}_2' \left[I_2 - \tilde{H}_{22} \right]^{-1} \tilde{e}_2$, where *RSS* is residual sum of squares of (8). Then the *F* statistic is as follows:

$$F = \frac{\tilde{e}_2' \left[I_2 - \tilde{H}_{22} \right]^{-1} \tilde{e}_2 \big/ 2}{\left(RSS - \tilde{e}_2' \left[I_2 - \tilde{H}_{22} \right]^{-1} \tilde{e}_2 \right) \big/ n - p - 2},$$

which is distributed as $F_{2,n-p-2}$.

***Remark* 2**: It should be noted that for any dataset with autocorrelated errors, an observation can be recognized as an outlier not only for a large vertical distance from the bulk of the data, but also for inconsistency with the autoregressive structure. The existence of such an outlier observation may also causes the previous or subsequent observations to be recognized as outliers. This occurs when the autoregressive connection between error terms does not exist anymore and these observations are considered as outliers in accordance with the autoregressive structure of the errors. More details are discussed in the numerical example section.

Influential diagnostics through the case deletion method

Following Belsley [2] consider the *i*th observation as an influential observation candidate. The case deletion method is based on determining changes in the estimated parameters and fitted values, when the *i*th observation is deleted, by measuring *DFBETA* and *DFFIT*, respectively. Acar and Özkale [13] found the measures in case of autocorrelated ridge regression model using leverage values. We intend to find these measures when the stochastic restrictions are added to the model, but through using the elements e_i^* and v_i^* defined by Roy and Curia [11] for a simpler calculation.

We start with some adjustments in the previous details and formulas according to deletion of the *i*th observation as follows:

$$\begin{pmatrix} y_{(i)} \\ 0 \\ r \end{pmatrix} = \begin{pmatrix} X_{(i)} \\ \sqrt{k} I_p \\ R \end{pmatrix} \quad + \begin{pmatrix} u_{(i)} \\ ¦ \\ μ \end{pmatrix}$$

or

$$y_{m(i)} = X_{m(i)} \quad + £_{m(i)} \qquad (9)$$

where $Var(£_{m(i)}) = ±^2 \Lambda_{(i)}$ and $\Lambda_{(i)} = diag(\Omega_{(i)}, I, W)$ is a p.d matrix. The matrix $\Omega_{(i)}$ is also a p.d matrix, so a nonsingular matrix $P_{(i)}$ is available through the Result 3.1 in Roy and Curia [11] such that $\Omega_{(i)}^{-1} = P_{(i)}' P_{(i)}$. Then There exists a nonsingular matrix $L_{(i)} = diag(P_{(i)}, I, T)$ such that $\Lambda_{(i)}^{-1} = L_{(i)}' L_{(i)}$

Lemma 2: For the *mixed stochastic restricted autocorrelated ridge regression* (*MSAR*) model (9), the *MSAR* estimator of is given by:

$$\hat{}_{MSAR(i)}(k) = \left(X_{(i)}' \Omega_{(i)}^{-1} X_{(i)} + R'W^{-1}R + kI_p \right)^{-1} \left(X_{(i)}' \Omega_{(i)}^{-1} y_{(i)} + R'W^{-1}r \right)$$

Proof: Pre-multiplying model (9) by $L_{(i)}$ and defining $\tilde{y}_{m(i)} = L_{(i)} y_{m(i)}$, $\tilde{X}_{m(i)} = L_{(i)} X_{m(i)}$ and $£_{m(i)} = L_{(i)} £_{m(i)}$, the model changes to $\tilde{y}_{m(i)} = \tilde{X}_{m(i)} \quad + £_{m(i)}$, where $Var(£_{m(i)}) = ±^2 I$. Hence the OLS estimator of is readily derived as follows:

$$\hat{}_{MSAR(i)} = \left(\tilde{X}_{m(i)}' \tilde{X}_{m(i)} \right)^{-1} \tilde{X}_{m(i)}' \tilde{y}_{m(i)} = \left(X_{m(i)}' \Lambda_{(i)}^{-1} X_{m(i)} \right)^{-1} X_{m(i)}' \Lambda_{(i)}^{-1} y_{m(i)}$$

$$= \left(\begin{bmatrix} X_{(i)}' & \sqrt{k} I_p & R' \end{bmatrix} \begin{bmatrix} \Omega_{(i)}^{-1} & 0 & 0 \\ 0 & I_p & 0 \\ 0 & 0 & W^{-1} \end{bmatrix} \begin{bmatrix} X_{(i)} \\ \sqrt{k} I_p \\ R \end{bmatrix} \right)^{-1} \times$$

$$\begin{bmatrix} X_{(i)}' & \sqrt{k} I_p & R' \end{bmatrix} \begin{bmatrix} \Omega^{-1} & 0 & 0 \\ 0 & I_p & 0 \\ 0 & 0 & W^{-1} \end{bmatrix} \begin{bmatrix} y_{(i)} \\ 0 \\ r \end{bmatrix}$$

$$= \left(X_{(i)}' \Omega_{(i)}^{-1} X_{(i)} + R'W^{-1}R + kI_p \right)^{-1} \left(X_{(i)}' \Omega_{(i)}^{-1} y_{(i)} + R'W^{-1}r \right)$$

Replacing $\Omega_{(i)}^{-1}$ with $P_{(i)}' P_{(i)}$ in $\hat{}_{MSAR(i)}$ and letting $\tilde{X}_{(i)} = P_{(i)} X_{(i)}$ and $\tilde{y}_{(i)} = P_{(i)} y_{(i)}$ we have

$$\hat{}_{MSAR(i)}(k) = \left(\tilde{X}_{(i)}' \tilde{X}_{(i)} + R'W^{-1}R + kI_p \right)^{-1} \left(\tilde{X}_{(i)}' \tilde{y}_{(i)} + R'W^{-1}r \right)$$

.

Taking into account the Result 3.2 of Roy and Curia [11], for *i*=*2*,…,(*n*-1), we have $\tilde{X}_{(i)}' \tilde{X}_{(i)} = \tilde{X}'\tilde{X} - e_i^* e_i^{*\prime}$ and $\tilde{X}_{(i)}' \tilde{y}_{(i)} = \tilde{X}'\tilde{y} - e_i^* v_i^*$ where $e_i^* = (1 + °^2)^{-1/2} (° \tilde{x}_{i+1} - \tilde{x}_i)$ and

$v_i^* = (1+\phi^2)^{-1/2}(\phi \tilde{y}_{i+1} - \tilde{y}_i)$. So it follows that

$$\hat{\beta}_{MSAR(i)}(k) = \left[\left(\tilde{X}'\tilde{X} + R'W^{-1}R + kI_p\right) - e_i^* e_i^{*\prime}\right]^{-1}\left[\left(\tilde{X}'\tilde{y} + R'W^{-1}r\right) - e_i^* v_i^*\right]$$

Letting $C = \tilde{X}'\tilde{X} + R'W^{-1}R + kI_p$ and $D = \tilde{X}'\tilde{y} + R'W^{-1}r$ then

$$\hat{\beta}_{MSAR(i)}(k) = \left(C - e_i^* e_i^{*\prime}\right)^{-1}\left(D - e_i^* v_i^*\right) = \left(C^{-1} + \frac{C^{-1}e_i^* e_i^{*\prime}C^{-1}}{1 - e_i^{*\prime}C^{-1}e_i^*}\right)\left(D - e_i^* v_i^*\right).$$

Applying Sherman-Morrison-Woodbury formula (Seber [23], p.467) for the second equality leads to

$$\hat{\beta}_{MSAR(i)}(k) = \hat{\beta}_{MSAR}(k) - C^{-1}e_i^*\left(\frac{v_i^* - e_i^{*\prime}\hat{\beta}_{MSAR}(k)}{1 - e_i^{*\prime}C^{-1}e_i^*}\right) = \hat{\beta}_{MSAR}(k) - \frac{(C^{-1}e_i^*)\eta_i}{1 - e_i^{*\prime}C^{-1}e_i^*}$$

where $\hat{\beta}_{MSAR}(k)$ is the same as $\hat{\beta}_{MSR}(k)$ with Ω of the form (5) and $\eta_i = (1+\phi^2)^{-1/2}(\phi\tilde{\varepsilon}_{i+1} - \tilde{\varepsilon}_i)$ in which $\tilde{\varepsilon}_i = \tilde{y}_i - \hat{\beta}'_{MSAR}(k)\tilde{x}_i$. Therefore, it follows that

$$DFBETA_{MSAR(i)} = \frac{(C^{-1}e_i^*)\eta_i}{1 - e_i^{*\prime}C^{-1}e_i^*} \quad i = 2, \ldots, (n-1).$$

In order to standardize the $DFBETA_{MSAR(i)}$ for the jth parameter, we note that

$$Var(\hat{\beta}_{MSAR}(k)) = Var\left[C^{-1}\left(\tilde{X}'\tilde{X} + R'W^{-1}r\right)\right] = \sigma^2\left[C^{-1}\left(\tilde{X}'\tilde{X} + R'W^{-1}R\right)C^{-1}\right]$$

so the estimate of standard error of $\hat{\beta}_{MSAR_j}(k)$ is

$$SE(\hat{\beta}_{MSAR_j}(k)) = S(i)\left[C^{-1}\left(\tilde{X}'\tilde{X} + R'W^{-1}R\right)C^{-1}\right]_{j,j}^{1/2}$$

where $S(i)$ is the mixed restricted autocorrelated estimate of σ after deleting the ith case. Then

$$DFBETAS_{MSAR(i)j} = \frac{\left(C^{-1}e_i^*\right)_j \eta_i}{\left(1 - e_i^{*\prime}C^{-1}e_i^*\right)SE(\hat{\beta}_{MSAR_j}(k))}.$$

Furthermore, the *DFFIT* measure can be derived as

$$DFFIT_{MSAR(i)} = \tilde{x}_i'\left[\hat{\beta}_{MSAR}(k) - \hat{\beta}_{MSAR(i)}(k)\right] = \frac{\tilde{x}_i'\left(C^{-1}e_i^*\right)\eta_i}{1 - e_i^{*\prime}C^{-1}e_i^*}.$$

The estimated standard error of the ith fitted value is given by

$$SE(\tilde{x}_i'\hat{\beta}_{MSAR}(k)) = S(i)\left[\tilde{x}_i'C^{-1}\left(\tilde{X}'\tilde{X} + R'W^{-1}R\right)C^{-1}\tilde{x}_i\right]^{1/2}$$

so that the standardized *DFFIT* of the ith fitted value, is as follows:

$$DFFITS_{MSAR(i)} = \frac{\tilde{x}_i'\left(C^{-1}e_i^*\right)\eta_i}{\left(1 - e_i^{*\prime}C^{-1}e_i^*\right)SE\left(\tilde{x}_i'\hat{\beta}_{MSAR}(k)\right)}.$$

Numerical Example

In order to illustrate diagnostic measures discussed in the preceding sections, we use an example previously employed by Bayhan and Bayhan [6] and Özkale [5]. This dataset belongs to a Turkish shampoo and soap firm in which the purpose is to estimate the demand of future weekly sales based on the weekly variation of shampoo sales. 75 records of weekly observations of sales collected in a period of time with a high and irregular inflation are used. The first 60 observations in Table 1 are assumed as historical data and the last 15 observations in Table 2 as fresh data. Two considered explanatory variables and a dependent variable are as follows: weekly list prices (averages from selected

Table 1. Historical data for weekly sales of shampoos and prices

Row	y	X_1	X_2	Row	y	X_1	X_2	Row	y	X_1	X_2
1	28.445	49	12.5	21	30.441	72.4	18	41	32.441	88.5	22.3
2	28.547	49	12.5	22	30.549	72.4	21	42	32.545	68.7	22.9
3	28.644	51.2	13	23	30.641	80	21	43	32.643	68.7	22.9
4	28.746	51.2	13	24	30.739	72	18.3	44	32.748	91.3	22.9
5	28.849	40.3	13	25	30.845	72	18.3	45	32.842	91.3	22.9
6	28.940	52	13	26	30.949	55	19	46	32.950	91.3	23
7	29.045	52.3	13.8	27	31.051	48	19	47	33.039	92.8	23
8	29.142	58	14.4	28	31.148	80.1	19.4	48	33.144	92.8	23
9	29.248	58	14.4	29	31.245	80.1	21.2	49	33.249	76	25.4
10	29.250	58	14.4	30	31.342	84.4	21.2	50	33.347	76	26
11	29.443	62	16	31	31.446	84.4	21.2	51	33.442	93.4	24.1
12	29.545	62	16	32	31.549	85	21.2	52	33.543	93.4	24.1
13	29.644	62	16	33	31.641	85	21.2	53	33.647	93.4	24.1
14	29.747	52	17.1	34	31.743	78	20.1	54	33.746	96.3	24.1
15	29.841	67.2	17.1	35	31.848	78	20.1	55	33.849	96.3	24.3
16	29.045	67.2	17.1	36	31.940	81.3	20.1	56	33.940	97.2	24.3
17	30.046	67.2	18	37	32.043	83.1	21	57	34.041	97.2	24.3
18	30.142	67.2	18	38	32.146	83	21	58	34.143	75.2	25.1
19	30.245	72.4	18	39	32.250	88.5	22.3	59	34.248	100	25.1
20	30.348	72.4	18	40	32.344	88.5	22.3	60	34.345	101.5	25.4

Table 2. Fresh data for weekly sales of shampoos and prices

Row	y	X_1	X_2	Row	y	X_1	X_2	Row	y	X_1	X_2
1	34.481	101.3	25.3	6	34.308	104.9	26.2	11	34.780	108.5	27.1
2	34.369	102	25.5	7	34.402	105.6	26.4	12	34.875	109.1	27.3
3	34.268	102.7	25.7	8	34.479	106.9	26.6	13	34.963	109.9	27.5
4	34.160	103.5	25.9	9	34.580	107	26.8	14	35.540	110.6	27.7
5	34.215	104.2	26.1	10	34.682	107.7	27	15	35.173	111.3	27.8

Table 3. Summary of descriptive statistics for shampoos and prices

Data	n	y		X_1		X_2	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Historical	60	31.35	1.79	75.05	15.94	19.86	3.9
Fresh	15	34.59	0.32	106.35	3.19	26.59	0.8

supermarkets) of the firm's shampoos (X_1), weekly list prices of a certain brand of soap, substituted for shampoos (X_2) and weekly quantities of shampoos sold (bottle average for the selected supermarkets) (*Y*). Prices are in Turkish Liras. Table 3 represents a summary of descriptive statistics for all variables in both data series.

Our computations are carried out by R software of version 3.30. First, each variable is centered and scaled by the unit normal scaling technique such that

$$w_j = \frac{y_j - \bar{y}}{S_{yy}}, \quad z_{j1} = \frac{x_{j1} - \bar{x}_1}{S_1}, \quad z_{j2} = \frac{x_{j2} - \bar{x}_2}{S_2}$$

where

$$S_{yy}^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(y_j - \bar{y}\right)^2, \qquad S_1^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{j1} - \bar{x}_1\right)^2,$$

$S_2^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{j2} - \bar{x}_2\right)^2$ (see Montgomery and Peck[24], p.113). Then following Bayhan and Bayhan [6] and Özkale [5], the Durbin-Watson (DW) statistic of fresh and historical data are calculated as 0.3533 and 0.562, respectively, which mean that error terms of the both datasets are of the form AR(1) at 0.05 significance level (for more details about the autocorrelated structure of this data set, see Özkale [5]). The estimated value of ρ using historical data is 0.7072. Substituting $\hat{\rho}$ in Ω the condition number of $X'\hat{\Omega}^{-1}X$ for fresh data is 126.3 which indicates that the fresh dataset has strong collinearity problem.

We employ the observations 48 and 49 from transformed historical data as prior information in the form of stochastic linear restrictions with $r = \begin{bmatrix} 0.1303 \\ 0.1380 \end{bmatrix}$

$R = \begin{bmatrix} 0.1450 & 0.1049 \\ 0.0077 & 0.1850 \end{bmatrix}$, and $W = \begin{bmatrix} 1 & 0.7072 \\ 0.7072 & 1 \end{bmatrix}$. Following Firinguetti [21], the estimate of ridge parameter estimate is $\hat{k} = 0.356$. So we have, $\hat{\beta}_{MSAR}(k) = \begin{pmatrix} 0.2835 \\ 0.4383 \end{pmatrix}$.

In order to investigate outlier observations through the mean-shift outlier model, the *F*-statistic for each observation is given in Table 4. It can be seen that 1st and 3rd observations have larger *F* values than the table

Table 4. F-statistic for mean-shift outlier model

Obs.	F	p-value	Obs.	F	p-value
1	6.66	0.0085	8	0.11	0.8966
2	1.72	0.2126	9	0.01	0.9901
3	4.54	0.0287	10	0.23	0.7973
4	2.99	0.0808	11	0.39	0.6837
5	0.26	0.7745	12	0.43	0.6583
6	0.01	0.9901	13	0.60	0.5615
7	0.07	0.9327	14	1.83	0.1945

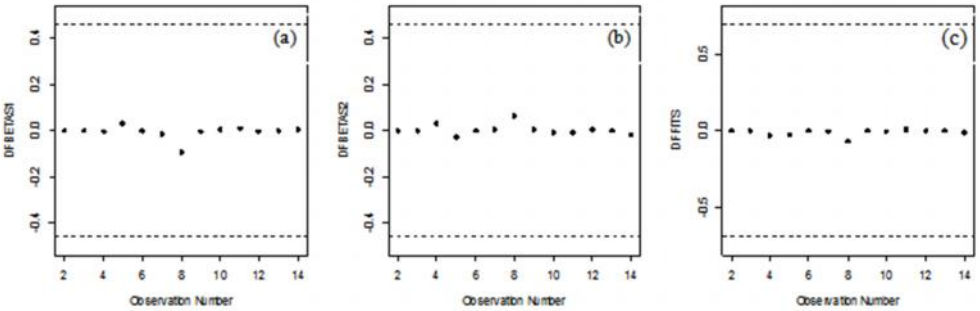

Figure 1. (a) *DFBETAS* for ${}_1$, (b) *DFBETAS* for ${}_2$ (c) *DFFITS* for fitted values after fitting Mixed Stochastic restricted Autocorrelated Ridge model to the shampoo data. The dashed lines are cutoff points.

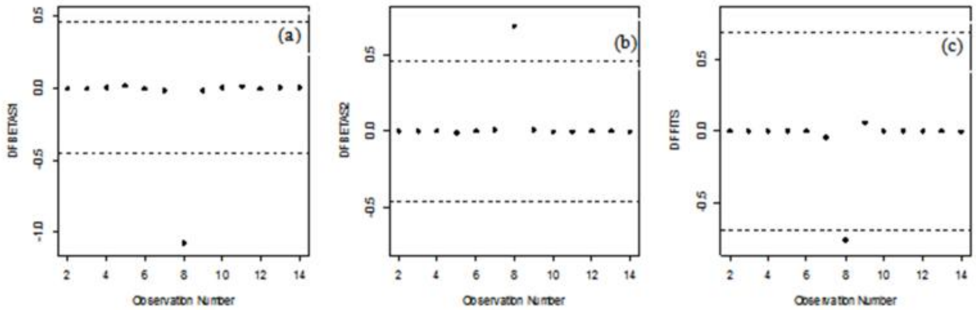

Figure 2. (a) DFBETAS for ^{1}, (b) DFBETAS for ^{2} (c) DFFITS for fitted values after fitting Mixed Stochastic Autocorrelated Ridge model to the shampoo data with a shift in 8th observation. The dashed lines are cutoff points.

$F_{2,n-p-2} = 3.68$ at $\sim = 0.05$, so these observations will be diagnosticated as outliers.

To investigate influential observations using the case deletion method, the related diagnostic measures are given in Figure 1. The cutoff points are $2/\sqrt{n} = 0.46$ for *DFBETAS* and $2\sqrt{p/n-p} = 0.69$ for *DFFITS* (see Seber [23], p.307). While the absolute values of *DFBETAS* and *DFFITS* for 8th observation are at least three times more than those for the other observations, none of these measures exceed the cutoff points and no influential observation is recognized (the above influence measures cannot be evaluated for 1st and 15th observations).

In order to identify the efficiency of our methods, a reduction shift for dependent variable, equal to 0.4 can be chosen (which is slightly more than one standard deviation of *y*) and exerted on 8th observation. We emphasize that this observation is more likely influential under reduction shift according to the position of 8th point in Figure 1. Table 5 shows that applying the mean-shift outlier model to the shifted dataset, the 8th observation is recognized as an outlier, but it is observed that 7th observation is also discerned as an outlier. According to the Remark 2, the reason is disbanding the autoregressive structure of data by shifting 8th observation implying that these observations are outliers in accordance with the autoregressive nature of data. Again, for more investigation, a proposed

Table 5. *F*-statistic for mean- shift outlier model after changing 8th observation

Obs.	F	p-value	Obs.	F	p-value
1	1.16	0.34	8	15.84	0.00
2	0.46	0.64	9	2.5	0.12
3	0.99	0.39	10	0.14	0.87
4	0.71	0.51	11	0.22	0.81
5	0.07	0.93	12	0.24	0.79
6	0.0002	0.99	13	0.33	0.72
7	5.67	0.01	14	0.79	0.47

reduction shift equal to $\delta \times SD(y)$ is exerted on 7th observation and the result of this relief action shows that, only 8th observation was recognized as an outlier. In other words, the shift exerted to 7th observation recovers the destroyed autocorrelated connection between the two observations.

By the reduction shift action for 8th observation, as expected, substantial changes occur in the outcome of diagnosing influential observations (see Figure 2). It can be seen that *DFBETAS* and *DFFITS* of 8th observation exceed the cutoff points. So, in the shifted dataset, 8th observation is an influential observation.

Simulation Study

In this section the performance of the proposed methods are investigated through a simulation study. We generate data with autocorrelated error terms taking into account different values of ° , sample size, degree of collinearity and shift values. Then the mean-shift outlier and the case deletion methods are conducted.

In order to produce autocollinear explanatory variables, the formula

$$x_{ij} = (1 - ª^2)^{1/2} d_{ij} + ª d_{i,p+1} \quad , \; i = 1,...,n \quad , \; j = 1,..., p$$

suggested by McDonald and Galarneau [25] is used where d_{ij} are independent standard normal pseudo-random numbers and ª is specified so that the correlation between any two explanatory variables is given by $ª^2$.

The dataset is generated based on the model $y_t = x_t \; + u_t \quad t = 1,...,n$ with $u_t = ° u_{t-1} + £_t$. For $n = 40, 100, 200$, define $ª = 0.5, 0.9$, $d_{ij} \sim N(0,5)$, $' = (\;_1, \;_2)' = (0.6, 0.4)'$, $° = 0.6, 0.9$ and $£_t$ is generated from $N(0,1)$.

Each data set is produced by a combination of model specification and generating the error terms $£_t$, then 40% of the sample sizes are assigned to historical data (i.e. 16, 40, 80) and 60% to fresh data (i.e. 24, 60, 120). Then the assigned data are centered and scaled by the unit normal scaling technique. The DW statistic for each set of data is calculated to assure that the autocorrelated property is still hold.

This procedure is replicated 1000 times by generating new error terms $£_t$, for each combination of model specification and using the assigned historical data, the estimates of Ω and *P* matrices are obtained. Then the mixed stochastic restricted ridge estimator (4) is calculated. Following Bayhan and Bayhan [6], a subset of $(n/2)^{th}$ and $(n/2+1)^{th}$ observations of transformed historical data are used as the stochastic linear restrictions (2), so $W = \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix}$. The ridge parameter estimate $\hat{k}$ is calculated as suggested by Firinguetti [21] from assigned fresh dataset.

For each of the generated dataset the method of mean-shift outlier model for testing $H_{02}: ¥_t = 0$ is applied and the percentage of times that the *F*-statistic is greater than the corresponding critical *F* value for

Table 6. The probability of type I error ($\sim = 0.05$) and power of F test for the mean-shift outlier model with a combination of parameters *n*, ª , ° for $(\;_1, \;_2) = (0.6, 0.1)$ and shift=3.5 and 4

***n*-total**	***n*** Historical	Fresh	ª	°	**Sig level**	**Power** Shift=3.5	Shift=4
40	16	24	0.5	0.6	0.100	0.821	0.876
				0.9	0.065	0.826	0.872
			0.9	0.6	0.104	0.843	0.894
				0.9	0.063	0.830	0.879
100	40	60	0.5	0.6	0.058	0.950	0.975
				0.9	0.048	0.958	0.984
			0.9	0.6	0.050	0.944	0.986
				0.9	0.051	0.960	0.991
200	80	120	0.5	0.6	0.059	0.964	0.989
				0.9	0.044	0.982	0.999
			0.9	0.6	0.050	0.966	0.989
				0.9	0.030	0.990	0.999

~ $= 0.05$ is calculated.

For each generated data set, the 4th observation is considered as an outlier by exerting two shift values 3.5 and 4, which are around the mean of standard deviation of generated dependent observations, to its original dependent values. The autocorrelated structure of errors is tested again by calculating the DW statistic to assure this property still hold. The power of test is calculated by the foregoing method and the results of this simulation study for different combinations of model specifications are given in Table 6.

Table 6 indicates that except for simultaneous small values of ° and *n*, in other combinations of *n*, a and ° the significance level remains around ~ $= 0.05$. Moreover, it shows that the power of the test is increasing continuously with increase of the shift values.

In case-deletion method, the mean of absolutes of *DFBETAS*, *DFFITS* and the proportion of replications which exceed the related cutoff points in 1000 replicates are used as judgment tools. The results of different shift values are shown in Tables 7 and 8.

Some related results from this study are presented as follows:

- It can be seen that the mean of $|DFFITS|$ and $|DFBETAS|$ and their related proportions for the mixed stochastic autocorrelated ridge regression model are smaller than the measures for autocorrelated ridge regression model. It may be due to improved accuracy of the advanced model and its reduced sensitivity to the observations which do not have a large displacement from the bulk of data.
- Increment in ° does not have a significant effect on the mean of $|DFFITS|$ and the related proportion in the small sample sizes for the both models, whereas for large sample sizes, it mainly causes a growth in the measures. These results are somehow valid for the mean of $|DFBETAS|$ and related proportions.
- Increment in degree of collinearity which is delegated by a has mainly a reduction effect on the mean of $|DFFITS|$ and the related proportion for both models. These results are somehow valid for the mean of $|DFBETAS|$ and related proportions.

Table 7. Mean and proportion (%) of $|DFFITS|$ and $|DFBETAS|$ of 4th observation for different values of *n*, a and ° when Shift=3.5

***n*-total**	***n***		a	°	**Mean of *DFFITS* (% of Influential)**		**Mean of *DFFBETAS* (% of Influential)**			
							1		2	
	Historical	Fresh			With Restriction	Without Restriction	With Restriction	Without Restriction	With Restriction	Without Restriction
40	16	24	0.5	0.6	0.9429	1.0395	0.6017	0.6624	0.6681	0.7280
					(0.7)	(0.73)	(0.594)	(0.609)	(0.629)	(0.646)
				0.9	0.9115	1.0175	0.5902	0.6629	0.6298	0.6970
					(0.668)	(0.7)	(0.584)	(0.613)	(0.604)	(0.625)
			0.9	0.6	0.8940	0.9901	0.6046	0.6697	0.5981	0.6615
					(0.671)	(0.689)	(0.574)	(0.59)	(0.569	(0.604)
				0.9	0.9111	1.0267	0.6375	0.7132	0.6161	0.6917
					(0.672)	(0.704)	(0.6)	(0.62)	(0.582)	(0.616)
100	40	60	0.5	0.6	0.6787	0.7114	0.4299	0.4509	0.4886	0.5123
					(0.752)	(0.759)	(0.634)	(0.646)	(0.654)	(0.672)
				0.9	0.7271	0.7879	0.5052	0.5471	0.4955	0.5368
					(0.767)	(0.788)	(0.691)	(0.703)	(0.678)	(0.698)
			0.9	0.6	0.6772	0.7079	0.4667	0.4869	0.4887	0.511
					(0.746)	(0.758)	(0.659)	(0.662)	(0.659)	(0.676)
				0.9	0.6967	0.7617	0.4845	0.5311	0.4901	0.5357
					(0.722)	(0.751)	(0.656)	(0.679)	(0.654)	(0.677)
200	80	120	0.5	0.6	0.5127	0.5241	0.3380	0.3454	0.3549	0.3628
					(0.775)	(0.781)	(0.652)	(0.658)	(0.66)	(0.662)
				0.9	0.5344	0.5662	0.3763	0.3982	0.3717	0.3931
					(0.764)	(0.782)	(0.695)	(0.708)	(0.678)	(0.694)
			0.9	0.6	0.4905	0.5023	0.3412	0.3497	0.3466	0.3544
					(0.732)	(0.734)	(0.646)	(0.655)	(0.653)	(0.65)
				0.9	0.5313	0.5649	0.3793	0.4043	0.3830	0.4072
					(0.763)	(0.774)	(0.694)	(0.699)	(0.688)	(0.7)

Table 8. Mean and proportion (%) of $|DFFITS|$ and $|DFBETAS|$ of 4th observation for different values of *n*, a and ° when Shift=4

***n*-total**	***n***		a	°	**Mean of *DFFITS* (% of Influential)**		**Mean of *DFFBETAS* (% of Influential)**			
							1		2	
	Historical	Fresh			With Restriction	Without Restriction	With Restriction	Without Restriction	With Restriction	Without Restriction
40	16	24	0.5	0.6	1.0807	1.2044	0.7119	0.7945	0.7058	0.7807
					(0.783)	(0.795)	(0.657)	(0.675	(0.661)	(0.677)
				0.9	1.0302	1.1526	0.7117	0.7927	0.7081	0.7894
					(0.714)	(0.743)	(0.654)	(0.675)	(0.638)	(0.659)
			0.9	0.6	1.0468	1.1576	0.6860	0.7544	0.6957	0.7698
					(0.753)	(0.772)	(0.638)	(0.645)	(0.644)	(0.658)
				0.9	0.9728	1.0933	0.6631	0.7383	0.6638	0.7420
					(0.717)	(0.733)	(0.632)	(0.656)	(0.604)	(0.631)
100	40	60	0.5	0.6	0.7851	0.8223	0.5266	0.5512	0.5378	0.5613
					(0.791)	(0.802)	(0.708)	(0.71)	(0.711)	(0.72)
				0.9	0.8117	0.8834	0.5585	0.6046	0.5677	0.6183
					(0.794)	(0.807)	(0.729)	(0.74)	(0.7)	(0.72)
			0.9	0.6	0.7674	0.8023	0.5110	0.5341	0.5265	0.5482
					(0.8)	(0.813)	(0.688)	(0.698)	(0.681)	(0.69)
				0.9	0.8025	0.8761	0.5612	0.6141	0.5690	0.6197
					(0.788)	(0.811)	(0.694)	(0.732)	(0.71)	(0.724)
200	80	120	0.5	0.6	0.5802	0.5943	0.3883	0.3975	0.3963	0.4060
					(0.814)	(0.817)	(0.708)	(0.709)	(0.712)	(0.713)
				0.9	0.6580	0.6991	0.4447	0.4734	0.4667	0.4944
					(0.807)	(0.816)	(0.742)	(0757)	(0.734)	(0.753)
			0.9	0.6	0.5744	0.5892	0.3884	0.3982	0.3977	0.4081
					(0.81)	(0.818)	(0.713)	(0.717)	(0.71)	(0.716)
				0.9	0.6061	0.6417	0.4256	0.4486	0.4315	0.4558
					(0.788)	(0.795)	(0.724)	(0.728)	(0.736)	(0.747)

- Increment in the sample size *n* has similar effects on the mean of $|DFFITS|$, the mean of $|DFBETAS|$ and the related proportions as follows. Almost for all of the cases it causes a significant reduction in both $|DFFITS|$ and $|DFBETAS|$, whereas it usually causes the related proportions increases, since the cutoff points are dependent on *n*.
- Increment in the size of shift values causes the measures and related proportions increase in all cases.

Discussion

In this article we extended the mean-shift outlier model, *DFFITS* and *DFBETAS* measures to the case of autocorrelated ridge regression model under stochastic linear restrictions. We applied our results to a real dataset with AR(1) error terms and stochastic linear restrictions. We observed that the proposed mean-shift outlier model is efficient for detecting observations which do not conform to the nature of data which are known as outliers. Also, the derived measures are suitable for diagnosing influential observations. In addition, a simulation investigation conducted to study the performance of mean-shift outlier method showed that if shift values increase, the power of the mean-shift *F* test will also increase. Also, a simulation was carried out to study the behavior of *DFFITS* and *DFBETAS* for the cases of mixed stochastic autocorrelated ridge regression model and autocorrelated ridge model. This simulation study demonstrated that, in general, the influence measures in the first model are smaller than those in the second model. A reduction occurs mainly when the degree of collinearity increases, whereas increasing autocorrelation coefficient for large sample sizes causes mainly a growth in the measures and, increasing sample size almost always causes a reduction in the measures.

Finally, in this paper we concentrated on the case deletion methods in restricted autocorrelated ridge models. This work can be extended in different directions. One direction is to study the local influence of observations in the above mentioned model. The other direction is using different biased estimators including Liu estimators and Lasso estimators.

Regarding the possible limitations, the main obstacle was that the authors had no access to the suitable real data sets from national research studies.

Acknowledgement

The authors would like to thank the editor and anonymous referees for several helpful comments and suggestions, which resulted in a significant improvement in the presentation of this paper.

References

1. Hoerl A.E. and Kennard R.W. Ridge regression: biased estimation for non-orthogonal problems. *Technometrics*, **12**: 69–82 (1970).
2. Belsley D.A., Kuh E. and Welsch R.E. *Regression diagnostics: identifying data and sources of collinearity*. John Wiley & Sons, New York, (2004).
3. Rao C.R., Toutenburg H., Shalabh and Heumann C. *Linear models and generalizations, Least squares and alternatives*. Springer, Berlin, (2008).
4. Sarkar N. A new estimator combining the ridge regression and the restricted least squares methods of estimation. *Commun. Stat.-Theor M*, **21**: 1987–2000 (1992).
5. Özkale M.R. A stochastic restricted ridge regression estimator. *J. Multivar Anal.*, **100**: 1706–1716 (2009).
6. Bayhan G.M. and Bayhan M. Forcasting using autocorrelated errors and multicollinear predictor variables. *Comp. ind. Eng.*, **34**(2): 413-421 (1998).
7. Alkhamisi M.A. Ridge estimation in linear models with autocorrelated errors. *Commun. Stat.-Theor M*, **39**: 2630–2644 (2010).
8. Groβ J. Restricted ridge estimation. *Stat. Probabil. Lett.*, **65**: 57-64 (2003).
9. Alheety M.I. and Golam Kibria B.M. A Generalized stochastic restricted ridge regression estimator. *Commun. Stat.-Theor M*, **43**: 4415–4427 (2014).
10. Chatterjee S. and Hadi A.S. Influential observations, high leverage points, and outliers in linear regression. *Stat. Science*, **1**(3): 379–416 (1986).
11. Roy S.S. and Guria S. Regression diagnostics in an autocorrelated model. *Braz. J. Probab.* Statist., **18**: 103–112 (2004).
12. Özkale M.R. and Acar T.S. Leverages and influential observations in a regression model with autocorrelated errors. *Commun. Stat.-Theor M*, **44**: 2267–2290 (2015).
13. Acar T.S and Özkale M.R. Influence measures in ridge regression when the error terms follow an AR(1) process. *Comput. Stat.*, **31**(3): 879–898 (2016).
14. Ghapani F., Rasekh A.R., Akhoond M.R. and Babadi B. Detection of outliers and influential observations in linear ridge measurement error models with stochastic linear restrictions. *J. Sci. I. R. Iran*, **26**(4): 355 - 366 (2015).
15. Wang J. *Statistical diagnosis of linear regression model with the random constraints and Bayes method.* Nanjing University of Science and Technology, Nanjing, (2007).
16. Wen H.W. and Wing K.F. The mean-shift outlier model in general weighted regression and its applications. *Comp. Stat. data an.*, **33**(4): 429-441 (1999).
17. Pan J. and Xiong H. Outliers and influential observations in a ridge mean shift regression, *System science and Mathematical science*, **9**(1): 12-26 (1995).
18. Ghapani F., Rasekh A.R. and Babadi B. Mean shift and influence measures in linear measurement error models with stochastic linear restrictions. *Comm. Stat-Simulat.C.*, **46**(6): 4499-4512 (2017).
19. Theil H. and Goldberger A.S. On pure and mixed statistical estimation in economics. *Int. Econ. Rev.*, **2** : 65-78 (1961).
20 Troskie C.G., Chalton D.O., Stewart T.J. and Jacobs M. Detection of outliers and influential observations in regression analysis using stochastic prior information, *Commun. Stat.-Theor M*, **23**(12): 3453–3476 (1994).
21. Firinguetti L. A simulation study of ridge regression estimators with autocorrelated errors. *Commun. Stat. Simulat.*, **18**(2): 673-701 (1989).
22. Judge G.C., Hill R.C., Griffiths W.E., Lütkepohl H. and Lee T.C. *Introduction to the Theory and Practice of Econometrics.* John Wiley & Sons, New York, (1988).
23. Seber G.A.F. and Lee A. *Linear Regression Analysis*. John Wiley & Sons, New Jersey, (2003).
24. Montgomery D.C., Peck E.A., and Vining G.G. *Linear Regression Analysis*,5th Ed., John Wiley & Sons, New Jersey, (2012).
26. McDonald G.C. and Galarneau D.I. A monte carlo evaluation of some ridge type estimators. *J. Am. Stat. Assoc.*, **70**: 407–416 (1975).