

Liu Estimates and Influence Analysis in Regression Models with Stochastic Linear Restrictions and AR (1) Errors

H. Mohammadi and A. R. Rasekh*

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

Received: 17 February 2019 / Revised: 27 April 2019 / Accepted: 26 May 2019

Abstract

In the linear regression models with AR (1) error structure when collinearity exists, stochastic linear restrictions or modifications of biased estimators (including Liu estimators) can be used to reduce the estimated variance of the regression coefficients estimates. In this paper, the combination of the biased Liu estimator and stochastic linear restrictions estimator is considered to overcome the effect of collinearity on the estimated coefficients. In addition, the deletion formulas for the detection of influential observations are presented for the proposed estimator. Finally, a simulation study and numerical example have been conducted to show the superiority of the proposed procedures.

Keywords: Liu estimator; Linear stochastic restrictions; Collinearity; Autocorrelated error; Influence analysis.

Introduction

The problem of collinearity in regression models refers to the situation where the explanatory variables have the near-linear dependency. By considering independent and identically distributed errors (homoscedasticity), it is well known that the ordinary least squares estimators (OLSE) are unbiased and have minimum variance in the class of linear unbiased estimators. However, in the presence of collinearity, they are no more reliable estimators, [1,2]. Collinearity causes the variance of the estimates to be large, and so the estimates of parameters will be unstable. To overcome this problem, one way is to make use of prior knowledge of observations and to deal with this information as linear stochastic restrictions. In such a case, the mixed estimators proposed by Theil and

Goldberger [3] and Theil [4] can be used, which is gained by unifying the sample and the prior information.

The other remedy to combat the collinearity problem is to use biased estimators such as the popular ridge estimator [5]. In 1993, Liu [6] introduced a new biased estimator called the Liu estimator by combining the Stein estimator [7] and the ridge estimator and showed that it can be superior over each of them in the MSEM sense. He also pointed out that it is easier to choose the Liu biasing parameter than to choose that of the ridge estimator since the Liu estimator is a linear function of its biasing parameter, but the ridge estimator is a complicated function of it.

As by combining different estimators, one can inherit the advantages of each of them. Ozkale [8] introduced the stochastic restricted ridge estimator under the

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

assumption of homoscedasticity by combining the mixed estimator and the ridge estimator. Also, Hubert and Wijekoon [9] introduced the stochastic restricted Liu estimator under the same assumption and Yang and Xu [10] introduced another stochastic restricted Liu estimator by combining them in an alternative way.

In practice, the assumption of independent and identically distributed errors (homoscedasticity) does not always hold. Sometimes the data are collected over time and so cause the errors to be correlated. A commonly occurring case is when the errors follow a 1st order autoregressive process (AR(1)). In such cases fitting an inappropriate model can have deleterious effects. A classic example can be found in Box and Newbold [11] who commented on a paper by Coen et al. [12]. The latter seemingly showed that car sales seven quarters earlier could be used to predict stock prices. But they failed to examine residuals and there was strong evidence that the errors were correlated. After fitting an appropriate model, Box and Newbold showed that there was no significant relationship between the two variable [13]. In fact, in practical applications, the neglect of such correlation in the errors may lead to inefficient parameter estimates and misleading inferences from hypothesis tests and inefficient predictions. The reason is that the OLSEs fail to achieve minimum variance estimates and the usual estimator of the variance-covariance matrix will be biased (see Griffith et al. [14]). To overcome these effects of autocorrelation, Aitken [15] proposed the generalized least squares estimator (GLSE). However, the presence of collinearity in the regression models with AR(1) error, will also result in unreliable GLS estimates, because of the large total variance. So these two problems should be examined simultaneously in order to achieve an appropriate estimation procedure for regression coefficients. The interested reader may refer to [16-20] for more details.

On the other hand, all the observations do not have the same impact on the estimated regression coefficients or on the resulting fitted values, therefore after fitting a regression model and using an estimation procedure to understand the relationship between variables, it is important to detect influential observations and (or) outliers in the framework of influence analysis. There are different statistical measures to detect these potentially influential observations, some of which are based on deleting cases. For example, the influence function for i th observation can be obtained as differences between the parameter estimated with and without the i th observation. The limitation of this approach is that it cannot be easily generalized to the linear regression model with 1st order autoregressive

errors, in which the dependency structure of the autoregressive model will not be valid after deleting a single observation from the data except for the last observation. But here with suitable modification as in Roy and Guria [21], we have kept intact the inherent autocorrelation structure. Detecting outliers is another problem, which has been extensively studied in the influence analysis of linear regression models, [1,2,22]. The method of mean-shift outlier model is one of the most important approaches for detecting discordant outliers in the regression analysis. Since the existence of outliers and influential observations are complicated by the presence of collinearity, it seems reasonable that after reducing the effects of collinearity by an appropriate estimator, the methods of influence analysis adequately be modified. Ullah et al. [23] and Jahufer [24] used the procedure of case deletion and derived influence measures for the Liu estimator in regression models with independent and identically distributed errors and without stochastic restrictions on parameters. Ullah et al. [23] also investigated the mean shift outlier model for the above-mentioned Liu regression. Zaherzadeh et al. [20] extended the method of mean-shift outlier model for detecting outliers in case of the ridge regression model with stochastic linear restrictions when the errors follow the AR(1) process. They also derived extensions of measures for diagnosing influential observations based on case deletion methods.

In the present paper, we have considered an estimator, which is a combination of the Liu estimator, and the mixed estimator in regression models with stochastic linear restrictions, in the case of correlated errors and especially when they are 1st order autoregressive. Also, a method has been given for choosing the Liu biasing parameter. Furthermore, some diagnostic measures are studied to identify influential observations or outliers that may be involved in the data modeled by the proposed method of estimation. In the preliminary section, the proposed model is introduced and the estimators are derived. In the results section, the case deletion diagnostics *DFBETAS* and *DFFITs*, are developed for the detection of influential observations. Furthermore, an outlier detection procedure based on the mean-shift outlier model is presented. Finally, an example to illustrate our results and a simulation study to show the performance of the achievements are given.

Preliminaries

Consider the following linear regression model

$$y = X\beta + \varepsilon \quad (1)$$

where y is an $n \times 1$ vector of observations on the dependent variable, X is an $n \times p$ matrix of

observations on the explanatory variables, β is a $p \times 1$ vector of unknown regression coefficients and ε is an $n \times 1$ vector of error terms. Also, we assume that $E(\varepsilon) = 0$ and $var(\varepsilon) = \sigma^2 V$, in which V is an $n \times n$ known p.d. matrix which can be decomposed as $V = TT'$; where T is a nonsingular matrix. Under these assumptions, by pre-multiplying the linear regression equation (1) with T^{-1} , we would have the following transformed model:

$$y^* = X^* \beta + \varepsilon^* \tag{2}$$

where $y^* = T^{-1}y$, $X^* = T^{-1}X$ and $\varepsilon^* = T^{-1}\varepsilon$. Therefore, the generalized least squares (GLS) estimate of β is $\tilde{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y$.

When the collinearity problem is presented, the matrix $X'V^{-1}X$ will be near singular. In such a case, the generalization of the Liu estimator can be used to reduce the effect of collinearity on the parameters estimates [18]. The biased generalized Liu estimator $\tilde{\beta}_d$ is expressed as:

$$\tilde{\beta}_d = \tilde{F}_d \tilde{\beta} \tag{3}$$

where $\tilde{F}_d = (X'V^{-1}X + I)^{-1}(X'V^{-1}X + dI)$,

$0 < d < 1$ is the Liu biasing parameter, and $\tilde{\beta}$ is the (GLS) estimator of β . In addition, we have

$var(\tilde{\beta}_d) = \sigma^2 \tilde{F}_d (X'V^{-1}X)^{-1} \tilde{F}_d'$. Also $\tilde{\beta}_d$ can be considered as the GLS estimator in the augmented model

$$\begin{pmatrix} y \\ d\tilde{\beta} \end{pmatrix} = \begin{pmatrix} X \\ I_p \end{pmatrix} \beta + \begin{pmatrix} \varepsilon \\ \dot{\varepsilon} \end{pmatrix} \quad \text{with } \tilde{n} (= n + p)$$

observations, where ε satisfies the conditions in the model (1) and where $\dot{\varepsilon}$ is considered to be a random vector of errors such that $E(\dot{\varepsilon}) = 0$ and $var(\dot{\varepsilon}) = \sigma^2 I_p$.

Besides using biased estimators, one can use the mixed estimators to overcome the collinearity problem, when there is some prior knowledge of the data in the form of linear stochastic restrictions. Suppose that historical observations related to the linear regression model (1) are available, which can be written as linear stochastic restrictions of the following form:

$$r = R\beta + \phi \tag{4}$$

where r is a known $m \times 1$ random vector, R is a known $m \times p$ matrix of prior information of

rank $m(\leq p)$ and ϕ is a random vector independent of ε with $E(\phi) = 0$ and $var(\phi) = \sigma^2 W$, where W is an $m \times m$ known p.d. matrix.

By combining the linear regression model (1) with the restrictions (4), we would have the augmented model, $\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} \varepsilon \\ \phi \end{pmatrix}$ or

$$\tilde{y} = \tilde{X} \beta + \tilde{\varepsilon}$$

(5)

where $E(\tilde{\varepsilon}) = 0$, $var(\tilde{\varepsilon}) = \sigma^2 \tilde{W}$ and

$\tilde{W} = \begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix}$ is a p.d. matrix. The generalized mixed

least squares estimator of β in (5) is:

$$\begin{aligned} \tilde{\beta}_m &= (\tilde{X}'\tilde{W}^{-1}\tilde{X})^{-1}\tilde{X}'\tilde{W}^{-1}\tilde{y} \\ &= (X'V^{-1}X + R'W^{-1}R)^{-1}(X'V^{-1}y + R'W^{-1}r). \end{aligned} \tag{6}$$

The mixed estimator $\tilde{\beta}_m$ is unbiased and

$$var(\tilde{\beta}_m) = \sigma^2 \tilde{A} \tag{7}$$

where $\tilde{A} = (X'V^{-1}X + R'W^{-1}R)^{-1}$.

Hubert and Wijekoon [9] improved the Liu estimator in the ordinary linear regression model by considering simultaneously the two approaches followed in obtaining the mixed estimator $\tilde{\beta}_m$ with $V = I$ and the Liu estimator and proposed a new biased estimator of β called stochastic restricted Liu estimator. We consider such an estimator in the case of unequal variance and (or) correlated errors and combine the Liu estimator $\tilde{\beta}_d$ and the mixed estimator $\tilde{\beta}_m$. By substituting $\tilde{\beta}$ in equation (3) with $\tilde{\beta}_m$, we have

$$\tilde{\beta}_{srd} = \tilde{F}_d \tilde{\beta}_m$$

(8)

Since the mixed estimator $\tilde{\beta}_m$ can be written as $\tilde{\beta}_m = \tilde{\beta} + \tilde{S}^{-1}R'(W + R\tilde{S}^{-1}R')^{-1}(r - R\tilde{\beta})$ it will be easily seen that $\tilde{\beta}_{srd}$ can also be considered as the GLS estimator in the following augmented model with $\tilde{n} (= n + p)$ observations:

$$\begin{pmatrix} y \\ d\tilde{\beta}_m + \tilde{S}\tilde{g} \end{pmatrix} = \begin{pmatrix} X \\ I_p \end{pmatrix} \beta + \begin{pmatrix} \varepsilon \\ \varepsilon^d \end{pmatrix} \tag{9}$$

where $\tilde{S} = X'V^{-1}X$,

$\tilde{g} = \tilde{S}^{-1}R'(W + R\tilde{S}^{-1}R')^{-1}(r - R\tilde{\beta})$, and ε satisfies the conditions $E(\varepsilon) = 0$ and $var(\varepsilon) = \sigma^2V$, and ε^d is considered to be a random vector of errors independent of ε such that $E(\varepsilon^d) = 0$ and $var(\varepsilon^d) = \sigma^2I_p$. We refer to the augmented regression model (9) as stochastic restricted Liu regression under unequal-variance or correlated errors.

The expectation and variance of $\tilde{\beta}_{srd}$, are obtained by the unbiasedness of $\tilde{\beta}_m$ and its variance in (7), as

$$E(\tilde{\beta}_{srd}) = \tilde{F}_d\beta, \quad var(\tilde{\beta}_{srd}) = \sigma^2\tilde{F}_d\tilde{A}\tilde{F}_d' \quad (10)$$

The conditions for the superiority of $\tilde{\beta}_{srd}$ over $\tilde{\beta}_d$ and $\tilde{\beta}_m$ can be obtained by a simple modification of the proofs given in Hubert and Wijekoon [9] which are for the case of $V = I$.

In what follows, we consider a special structure of V , in which the data are collected over time and the error terms follow a 1st order autoregressive process (AR (1)), that is

$$\varepsilon_i = \rho\varepsilon_{i-1} + u_i \quad \text{where} \quad E(u_i) = 0 \quad \text{and} \\ var(u_i) = \sigma^2 \quad \text{for } i = 2, \dots, n$$

and where $|\rho| < 1$. In this case, as it is well known, the matrix V is expressed as

$$V = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix} \quad (11)$$

and its inverse is given by

$$V^{-1} = \begin{pmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix} \quad (12)$$

which can be decomposed as $V^{-1} = P'P$, where

$$P = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix}. \quad (13)$$

In practice, the matrix V in (11) is generally

unknown. By writing V as a function of the unknown parameter ρ , say $V(\rho)$, an estimator of V can be defined by replacing unknown ρ by an estimator $\hat{\rho}$, which can be derived, using the OLS technique as [1]

$$\hat{\rho} = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sum_{i=1}^n e_i^2} \quad (14)$$

where e_i is the i th element of the residual vector by ordinary least squares estimator. By denoting an estimator of V by $V(\hat{\rho})$, the estimator $\tilde{\beta}_{srd}$ in (8) can be expressed in the following expansion form

$$\tilde{\beta}_{srd} = [X'V^{-1}(\hat{\rho})X + I]^{-1} [X'V^{-1}(\hat{\rho})X + dI] \\ \times [X'V^{-1}(\hat{\rho})X + R'W^{-1}R]^{-1} [X'V^{-1}(\hat{\rho})y + R'W^{-1}r]. \quad (15)$$

Selection of d

It should be noted that Liu [6] has given different estimates of the Liu parameter d , in the linear model (1) with the assumption of independent and identically distributed errors. The given estimates are extended to the case of unequal variance or correlated errors by Alheety and Kibria [25]. They gave the optimal value of d by minimizing the *MSE* of the Liu estimator in the canonical form of the transformed regression model (2).

By symmetry of $(X^*X^*)X'V^{-1}X$, there exists an orthogonal matrix E containing normalized eigenvectors of $X'V^{-1}X$ such that $E'X'V^{-1}XE = \Gamma$, where $\Gamma = \text{diag}\{\gamma_1, \dots, \gamma_p\}$ is a diagonal matrix and the γ_i , $i = 1, \dots, p$, are eigenvalues of $X'V^{-1}X$. Since the matrix E is orthonormal, the transformed model (2) can be written in the following canonical form:

$$y^* = \tilde{Z}\alpha + \varepsilon^*, \quad \text{where } \tilde{Z} = X^*E \quad \text{and} \quad \alpha = E'\beta. \quad (16)$$

The least squares estimate and the Liu estimate of α are $\tilde{\alpha} = (\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'y^* = \Gamma^{-1}E'X'V^{-1}y$ and $\tilde{\alpha}_d = (\tilde{Z}'\tilde{Z} + I)^{-1}(\tilde{Z}'\tilde{Z} + dI)\tilde{\alpha} = (\Gamma + I)^{-1}(\Gamma + dI)\tilde{\alpha}$ respectively, where

$$E(\tilde{\alpha}) = \alpha \quad \text{and} \quad var(\tilde{\alpha}) = \sigma^2(\tilde{Z}'\tilde{Z})^{-1} = \sigma^2\Gamma^{-1} \quad (17)$$

and

$$E(\tilde{\alpha}_d) = (\Gamma + I)^{-1}(\Gamma + dI)\alpha \quad \text{and} \\ var(\tilde{\alpha}_d) = \sigma^2(\Gamma + I)^{-1}(\Gamma + dI)\Gamma^{-1}(\Gamma + dI)(\Gamma + I)^{-1}. \quad (18)$$

In order to find the optimal value of d , the *MSE* of

$\tilde{\alpha}_d$ is used which is defined as $var(\tilde{\alpha}_d) + B(\tilde{\alpha}_d)B(\tilde{\alpha}_d)'$, where $B(\tilde{\alpha}_d)$ is the bias of $\tilde{\alpha}_d$ in estimating α . Since from (18), $B(\tilde{\alpha}_d) = [(\Gamma + I)^{-1}(\Gamma + dI) - I]\alpha$, the trace of MSE of $\tilde{\alpha}_d$ is obtained as

$$TMSE(\tilde{\alpha}_d) = \sigma^2 \sum_{i=1}^p \frac{(\gamma_i + d)^2}{(\gamma_i + 1)^2 \gamma_i} + (d-1)^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\gamma_i + 1)^2}$$

where γ_i 's are the diagonal elements of the matrix Γ , and α_i ($i=1, \dots, p$) is the i th element of the regression coefficients vector α in the canonical model (16). By differentiating the $TMSE(\tilde{\alpha}_d)$ with respect to d , and equating it by zero, an estimate of d is found which minimizes the $TMSE$ criteria. After some simplifications, it is obtained

$$d = \left[\sum_{i=1}^p \frac{\alpha_i^2 - \sigma^2}{(\gamma_i + 1)^2} \right] \left[\sum_{j=1}^p \frac{\sigma^2 + \gamma_j \alpha_j^2}{(\gamma_j + 1)^2 \gamma_j} \right]^{-1}. \text{ Now by}$$

substituting unknown parameters σ^2 and α_i^2 with their unbiased estimates, an estimate of d can be obtained. Since the expectation and the variance of $\tilde{\alpha}_i$, the i th element of the estimated regression coefficients $\tilde{\alpha}$, are as given in (17), we have

$$E(\tilde{\alpha}_i^2) = var(\tilde{\alpha}_i) + E^2(\tilde{\alpha}_i) = \frac{\sigma^2}{\gamma_i} + \alpha_i^2. \text{ So an}$$

unbiased estimator of α_i^2 is $\tilde{\alpha}_i^2 - \frac{\hat{\sigma}^2}{\gamma_i}$, where in turn,

$$\hat{\sigma}^2 = \frac{(\tilde{y} - \tilde{X} \tilde{\beta}_m)' \tilde{W}^{-1} (\tilde{y} - \tilde{X} \tilde{\beta}_m)}{n + m - p}$$

is an estimator of σ^2 based on the mixed model (5).

By the above substitution, the optimal value of d after some simplifications is obtained as:

$$\hat{d} = 1 - \hat{\sigma}^2 \left[\sum_{j=1}^p \tilde{\alpha}_j^2 (\gamma_j + 1)^{-2} \right]^{-1} \left[\sum_{i=1}^p (\gamma_i (\gamma_i + 1))^{-1} \right]. \quad (19)$$

Results

Some diagnostic measures

After using a particular estimation procedure and fitting a regression model, one may be interested in the influence of individual observations on different aspects of the model including estimates of the parameters and

predicted values. Different measures of influence have been proposed, some of which are based on the deletion of cases.

Here we concentrate on the influential measures based on the estimator $\tilde{\beta}_{srd}$, which is presented in the preliminary section and especially we consider the case where the error terms of the linear model are 1st order autocorrelated. To determine the effect of the i th observation on the j th element of $\tilde{\beta}_{srd}$, we consider $DFBETAS$ criteria, which are based on the deletion of the i th case and are defined as:

$$DFBETAS_{(srd)j}(i) = \frac{\tilde{\beta}_{srd j} - \tilde{\beta}_{srd j}(i)}{S.E.(\tilde{\beta}_{srd j})} \quad (20)$$

where $\tilde{\beta}_{srd j}$ and $\tilde{\beta}_{srd j}(i)$ are the j th elements of $\tilde{\beta}_{srd}$ with and without the i th observation, respectively. In addition, $S.E.(\tilde{\beta}_{srd j})$ in the denominator of (20) is the standard error of j th regression coefficient estimated by $\tilde{\beta}_{srd}$, which is an estimate of the square root of the diagonal element of $var(\tilde{\beta}_{srd})$ in (10), namely, $S(i)(\tilde{F}_d \tilde{A} \tilde{F}')_{j,j}$. $S(i)$ is the estimate of σ based on the mixed model (5) and after deleting the i th case.

Although, Kim and Huggins [26] and Tsai and Wu [27] claim that the deletion approach is inappropriate in studying the diagnostics in a regression model with autocorrelated errors, but Roy and Guria [21] mentioned that this claim is relevant to the point that the deletion of an observation disrupt the autocorrelation structure. Roy and Guria [21] with suitable modification, found a transformation matrix for the case of AR(1) errors, that incorporates the deletion of an observation keeping intact the inherent autocorrelation structure.

For AR(1) errors, by using the matrix P given in (13) and putting $T^{-1} = P$ in the transformed model (2), we have $X^* = PX$, $y^* = Py$ and $\varepsilon^* = P\varepsilon$. In this case, the elements of y^* and X^* are as follow:

$$y_i^* = \begin{cases} \sqrt{1-\rho^2} y_1 & i=1 \\ -\rho y_{i-1} + y_i & i=2, \dots, n \end{cases}, \quad x_i^* = \begin{cases} \sqrt{1-\rho^2} x_1' & i=1 \\ -\rho x_{i-1}' + x_i' & i=2, \dots, n \end{cases} \quad (21)$$

where y_i^* and $x_i^{*'}$ denotes the i th element of y^* and the i th row of X^* , respectively, and where y_i and x_i' are the i th element of y and the i th row of X .

Now, following Roy and Guria [21], by deleting the i th observation, the i th row and the i th column of the variance matrix V also needs to be deleted. The impact

of this deletion on V^{-1} and P has been shown in Roy and Guria [21]. Suppose that $V_{(i)}$ is the matrix V after deleting its i th row and i th column and the matrices $V_{(i)}^{-1}$ and $P_{(i)}$ are obtained from $V_{(i)}$.

Result 1 (Roy and Guria [21]) For $i = 2, \dots, n-1$, $V_{(i)}^{-1}$ is obtained from V^{-1} in (12) by deleting its i th row and i th column, and replacing in V^{-1} the $(i-1, i-1)$ and $(i+1, i+1)$ elements by $(1+\rho^4)/(1+\rho^2)$, and the $(i-1, i+1)$ and $(i+1, i-1)$ elements by $(-\rho^2)/(1+\rho^2)$. The corresponding $P_{(i)}$ is obtained from P in (13) by deleting its i th row and i th column, and replacing the $(i+1, i-1)$ element of P by $(-\rho^2)/(1+\rho^2)^{1/2}$, and replacing the $(i+1, i+1)$ element of it by $1/(1+\rho^2)^{1/2}$.

Remark 1 It should be noted that by deleting the first row and the first column, or by deleting the last row and the last column of the matrix V , the overall structure of this matrix does not change. So for $i = 1$ and $i = n$, $V_{(i)}$ will be of the same form as V except for a single reduction in the dimension. That is the difference between V and $V_{(i)}$ will be that V is a square matrix of order n but $V_{(i)}$ is a square matrix of order $n-1$, where both are of the form given in (11). Hence the corresponding $V_{(i)}^{-1}$ and $P_{(i)}$ will also be the same as V^{-1} and P with one dimension less.

Let $X_{(i)}$ and $y_{(i)}$ be the matrix X and the vector y without the i th observation. Define $X_{(i)}^* = P_{(i)}X_{(i)}$ and $y_{(i)}^* = P_{(i)}y_{(i)}$, where $P_{(i)}$ is as obtained in result 1. It can be seen for $i = 2, \dots, n-1$ that (Roy and Guria [21]):

$$\begin{aligned} X_{(i)}^{*'} X_{(i)}^* &= X^{*'} X^* - u_i^* u_i^{*'} && \text{and} \\ X_{(i)}^{*'} y_{(i)}^* &= X^{*'} y^* - u_i^* \mathcal{G}_i^*, \quad i = 2, \dots, n-1 \end{aligned} \quad (22)$$

in which $u_i^* = (1+\rho^2)^{-1/2}(\rho x_{i+1}^* - x_i^*)$ and $\mathcal{G}_i^* = (1+\rho^2)^{-1/2}(\rho y_{i+1}^* - y_i^*)$, and where x_i^* and y_i^* are as defined in (21).

We found similarly for the cases of $i = 1$ and n .

First, we define $X_{(1)}^* = P_{(1)}X_{(1)}$ and $y_{(1)}^* = P_{(1)}y_{(1)}$, where $P_{(1)}$ is as mentioned in remark 1. It is seen that the first row of $X_{(1)}^*$ is equal to the $\sqrt{1-\rho^2} x_2'$, and for $j = 2, \dots, n-1$ the j th row of $X_{(1)}^*$ is equal to the $(j+1)$ th row of X^* , where the rows of X^* are as given in (21). So we have the following result:

$$\begin{aligned} X_{(1)}^{*'} X_{(1)}^* &= \sum_{i=3}^n x_i^* x_i^{*'} + (1-\rho^2)x_2 x_2' \\ &= X^{*'} X^* - x_1^* x_1^{*'} - x_2^* x_2^{*'} + (1-\rho^2)x_2 x_2' \\ &= X^{*'} X^* - w_1^* w_1^{*'} \end{aligned} \quad (23)$$

$$\text{in which } w_1^* = \sqrt{1-\rho^2} x_1^* - \rho x_2^*.$$

Also, it is seen that the first element of $y_{(1)}^*$ is equal to the $\sqrt{1-\rho^2} y_2$, and for $j = 2, \dots, n-1$ the j th element of $y_{(1)}^*$ is equal to the $(j+1)$ th element of y^* , where the elements of y^* are as given in (21). So we have

$$\begin{aligned} X_{(1)}^{*'} y_{(1)}^* &= \sum_{i=3}^n x_i^* y_i^* + (1-\rho^2)x_2 y_2 \\ &= X^{*'} y^* - x_1^* y_1^* - x_2^* y_2^* + (1-\rho^2)x_2 y_2 \\ &= X^{*'} y^* - v_1^* w_1^{*'} \end{aligned} \quad (24)$$

$$\text{where } v_1^* = \sqrt{1-\rho^2} y_1^* - \rho y_2^*.$$

Doing the same for the n th observation, and by defining $X_{(n)}^* = P_{(n)}X_{(n)}$ and $y_{(n)}^* = P_{(n)}y_{(n)}$, where $P_{(n)}$ is as mentioned in Remark 1, we also have:

$$\begin{aligned} X_{(n)}^{*'} X_{(n)}^* &= X^{*'} X^* - x_n^* x_n^{*'} && \text{and} \\ X_{(n)}^{*'} y_{(n)}^* &= X^{*'} y^* - y_n^* x_n^{*'} \end{aligned} \quad (25)$$

Now in order to find an expression for $\tilde{\beta}_{srd}(i)$ in terms of $\tilde{\beta}_{srd}$, we first consider the mixed estimator $\tilde{\beta}_m$ in equation (6). Suppose that $\tilde{\beta}_m(i)$ is the estimator $\tilde{\beta}_m$ after deleting the i th observation, which is computed as

$$\tilde{\beta}_m(i) = (X_{(i)}^{*'} X_{(i)}^* + R'W^{-1}R)^{-1}(X_{(i)}^{*'} y_{(i)}^* + R'W^{-1}r).$$

For $i = 2, \dots, n-1$, applying Sherman-Morrison-Woodbury formula ([1]) and using the expressions given in (22) and putting $M = X^{*'} X^* + R'W^{-1}R$,

$\tilde{\beta}_m(i)$ can be written as:

$$\begin{aligned} \tilde{\beta}_m(i) &= (X^* X^* - u_i^* u_i^{*'} + R^* W^{-1} R)^{-1} (X^* y^* - u_i^* \vartheta_i^* + R^* W^{-1} r) \\ &= \left(M^{-1} + \frac{M^{-1} u_i^* u_i^{*'} M^{-1}}{1 - u_i^{*'} M^{-1} u_i^*} \right) (X^* y^* + R^* W^{-1} r - u_i^* \vartheta_i^*). \end{aligned}$$

$i = 2, \dots, n-1$

Moreover, $\tilde{\beta}_m(i)$ can be simplified as

$$\begin{aligned} \tilde{\beta}_m(i) &= \tilde{\beta}_m - \frac{M^{-1} u_i^* (\vartheta_i^* - u_i^{*'} \tilde{\beta}_m)}{1 - u_i^{*'} M^{-1} u_i^*} \\ &= \tilde{\beta}_m - (1 - u_i^{*'} M^{-1} u_i^*)^{-1} (1 + \rho^2)^{-1/2} M^{-1} u_i^* (\rho \tilde{e}_{m_{i+1}} - \tilde{e}_{m_i}) \end{aligned}$$

$i = 2, \dots, n-1$ (26)

where \tilde{e}_{m_i} is the i th residual given by $\tilde{e}_{m_i} = y_i^* - x_i^{*'} \tilde{\beta}_m$. Doing similarly for $i = 1, n$ and using the expressions given in (23), (24) and (25), the estimators $\tilde{\beta}_m(1)$ and $\tilde{\beta}_m(n)$ are

$$\begin{aligned} \tilde{\beta}_m(1) &= \tilde{\beta}_m - (1 - w_1^{*'} M^{-1} w_1^*)^{-1} M^{-1} w_1^* ((1 - \rho^2)^{1/2} \tilde{e}_{m_1} - \rho \tilde{e}_{m_2}), \\ \tilde{\beta}_m(n) &= \tilde{\beta}_m - (1 - x_n^{*'} M^{-1} x_n^*)^{-1} M^{-1} x_n^* \tilde{e}_{m_n} \end{aligned}$$

where \tilde{e}_{m_i} ($i = 1, 2, n$) is as defined in (26).

Now we investigate the estimator $\tilde{\beta}_{srd}$ after deleting the i th case, which can be obtained as $\tilde{\beta}_{srd}(i) = (X_{(i)}^* X_{(i)}^* + I)^{-1} (X_{(i)}^* X_{(i)}^* + dI) \tilde{\beta}_m(i)$.

For $i = 2, \dots, n-1$, we have $\tilde{\beta}_{srd}(i) = (X^* X^* - u_i^* u_i^{*'} + I)^{-1} (X^* X^* - u_i^* u_i^{*'} + dI) \tilde{\beta}_m(i)$. By using Sherman-Morrison-Woodbury formula and after substituting $\tilde{\beta}_m(i)$ with the expression obtained in equation (26), and putting $h_i = u_i^{*'} M^{-1} u_i^*$ and $g_i = u_i^{*'} (X^* X^* + I)^{-1} u_i^*$, the estimator $\tilde{\beta}_{srd}(i)$ can be expressed as $\tilde{\beta}_{srd}(i) = \tilde{\beta}_{srd} - \zeta_i$, in which

$$\begin{aligned} \zeta_i &= \frac{(1 + \rho^2)^{-1/2}}{1 - g_i} (\rho \tilde{e}_{m_{i+1}} - \tilde{e}_{m_i}) (X^* X^* + I)^{-1} \\ &\quad \times \left\{ \left[(1 - g_i) (X^* X^* + dI) - u_i^* u_i^{*'} (I - \tilde{F}_d) \right] \frac{M^{-1}}{1 - h_i} - I_p \right\} u_i^* \\ &\quad + \frac{(1 + \rho^2)^{-1/2}}{1 - g_i} (\rho \tilde{e}_{srd_{i+1}} - \tilde{e}_{srd_i}) (X^* X^* + I)^{-1} u_i^* \end{aligned}$$

(27)

where \tilde{e}_{srd_i} is the i th residual given by $\tilde{e}_{srd_i} = y_i^* - x_i^{*'} \tilde{\beta}_{srd}$. For $i = 1, n$, by defining the

scalars $h_i = w_1^{*'} M^{-1} w_1^*$, $g_i = w_1^{*'} (X^* X^* + I)^{-1} w_1^*$,

$h_n = x_n^{*'} M^{-1} x_n^*$ and $g_n = x_n^{*'} (X^* X^* + I)^{-1} x_n^*$, we find:

$\tilde{\beta}_{srd}(i) = \tilde{\beta}_{srd} - \zeta_i$ ($i = 1, n$), in which ζ_1 and ζ_n are obtained as

$$\begin{aligned} \zeta_1 &= \frac{1}{1 - g_1} \left[(1 - \rho^2)^{1/2} \tilde{e}_{m_1} - \rho \tilde{e}_{m_2} \right] (X^* X^* + I)^{-1} \\ &\quad \times \left\{ \left[(1 - g_1) (X^* X^* + dI) - w_1^* w_1^{*'} (I - \tilde{F}_d) \right] \frac{M^{-1}}{1 - h_1} - I_p \right\} w_1^* \\ &\quad + \frac{1}{1 - g_1} ((1 - \rho^2)^{1/2} \tilde{e}_{srd_1} - \rho \tilde{e}_{srd_2}) (X^* X^* + I)^{-1} w_1^* \end{aligned}$$

(28)

$$\begin{aligned} \zeta_n &= \frac{\tilde{e}_{m_n}}{1 - g_n} (X^* X^* + I)^{-1} \\ &\quad \times \left\{ \left[(1 - g_n) (X^* X^* + dI) - x_n^* x_n^{*'} (I - \tilde{F}_d) \right] \frac{M^{-1}}{1 - h_n} - I_p \right\} x_n^* \\ &\quad + \frac{\tilde{e}_{srd_n}}{1 - g_n} (X^* X^* + I)^{-1} x_n^*. \end{aligned}$$

(29)

By defining ζ_{ij} ($i = 1, \dots, n, j = 1, \dots, p$) to be the j th element of the vector ζ_i , where ζ_i ($i = 2, \dots, n-1$), ζ_1 and ζ_n are respectively defined as in (27), (28) and (29). The *DFBETAS* criteria in (20) can be considered as below:

$$DFBETAS_{(srd)j}(i) = \frac{\zeta_{ij}}{S.E.(\tilde{\beta}_{srd_j})}$$

(30)

DFFITs criteria

DFFITs is another criterion, which can be used to derive the impact of each observation on the predicted values of the response variable by considering the predicted values before and after deleting the i th observation. Using the estimator $\tilde{\beta}_{srd}$, the *DFFITs* criteria for the i th observation is defined as

$$DFFITs_{(srd)}(i) = \frac{x_i^{*'} (\tilde{\beta}_{srd} - \tilde{\beta}_{srd}(i))}{S.E.(x_i^{*'} \tilde{\beta}_{srd})}$$

where $S.E.(x_i^{*'} \tilde{\beta}_{srd})$ is the estimated standard error of the fitted values by the estimator $\tilde{\beta}_{srd}$. This criterion can be written as

$$DFFITs_{(srd)}(i) = \frac{x_i^{*'} \zeta_i}{S.E.(x_i^{*'} \tilde{\beta}_{srd})}$$

(31)

where ζ_i ($i = 2, \dots, n-1$), ζ_1 and ζ_n are

respectively as defined in (27), (28) and (29). It should be noted that in the denominator of (31), using the equation (10), we have:

$$S.E.(x_i^* \tilde{\beta}_{srd}) = S(i)(x_i^* \tilde{F}_d \tilde{A} \tilde{F}_d' x_i^*)^{1/2} \text{ for } i = 1, \dots, n.$$

At last, it should be noted that regularly in ordinary least squares regression, the cutoff points for *DFBETAS* and *DFFITs* criteria are taken as $\frac{2}{\sqrt{n}}$ and

$$2\sqrt{\frac{p}{n-p}}.$$

But here they have to be modified according to equation (9). By taking into account the number of data in the augmented model and substituting n in the above expressions with \tilde{n} , the cutoff points are considered as $\frac{2}{\sqrt{n+p}}$ and $2\sqrt{\frac{p}{n}}$.

Mean shift outlier model

In fitting a regression model, some of the observations may arouse suspicions as they are discordant with other observations. Such observations are usually referred to as outliers and they may or may not have an effect on estimation and inference using the prescribed regression model [14]. The mean shift outlier model can be used to test if any observation is an outlier. For the regression model (1), it is defined as

$$y = X\beta + z_i\gamma + \varepsilon, \quad i = 1, \dots, n$$

where z_i is an n-dimensional vector with 1 at the i th position and zero elsewhere, and γ is the shift for the possibly outlying observation. In the above mean shift outlier model, $\gamma \neq 0$ indicates that the suspected observation (that is the i th case) is an outlier. For testing $H_0: \gamma = 0$ against $H_1: \gamma \neq 0$, the following F-statistic can be considered

$$F = \frac{RSS_0 - RSS_1}{RSS_1} \tag{32}$$

where RSS_0 and RSS_1 are the residual sum of squares under H_0 and H_1 , respectively. We consider the mean shift outlier model and define it in the stochastic restricted Liu regression (9) when the data have AR(1) errors where the variance of the error term in (9) is $\sigma^2 \begin{pmatrix} V & 0 \\ 0 & I_p \end{pmatrix}$, in which V is as defined in (11).

We continue by testing if the i th ($i = 1, 2, \dots, n$) observation is an outlier. For this purpose, we reorder the data such that the suspected observation to be located at the end of the data.

In the previous subsection, we mentioned that Roy and Guria [21] with a suitable modification found a transformation matrix that with a deletion of an observation the AR(1) structure do not be disrupted. In the same way, in what follows and after reordering the data, we will also find an appropriate covariance structure of the reordered error term such that the inherent autocorrelation structure be kept intact. and it's inverse will be used in the F-statistic (32).

Suppose that the $\tilde{n} \times 1$ vector of responses after reordering is $y^o = \begin{pmatrix} y_{(i)} \\ d\tilde{\beta}_m + \tilde{S}\tilde{g} \\ y_i \end{pmatrix}$, where $y_{(i)}$ is the

vector y without the i th case, and also the $\tilde{n} \times p$ reordered matrix of explanatory variables is

$$X^o = \begin{pmatrix} X_{(i)} \\ I_p \\ x_i' \end{pmatrix},$$

where $X_{(i)}$ is the matrix X without the i th case. Similarly, the $\tilde{n} \times p$ reordered error vector is

$$\varepsilon^o = \begin{pmatrix} \varepsilon_{(i)} \\ \varepsilon^d \\ \varepsilon_i \end{pmatrix}.$$

In other words, suppose that the regression model (9) is stated in the following reordered form

$$y^o = X^o\beta + \varepsilon^o$$

(33)

where

$$E(\varepsilon^o) = 0 \quad \text{and}$$

$$var(\varepsilon^o) = \sigma^2 V^o = \sigma^2 \begin{pmatrix} V_{(i)} & 0_{(n-1) \times p} & \mathcal{G}_i(i) \\ 0_{p \times (n-1)} & I_p & 0_{p \times 1} \\ \mathcal{G}_i(i) & 0_{1 \times p} & (1-\rho^2)^{-1} \end{pmatrix}, \quad \text{and}$$

where $V_{(i)}$ is the matrix V without the i th row and the i th column, and the vector $\mathcal{G}_i(i)'$ is the i th row of the matrix V which its i th case is omitted. In addition, for $i = 1, \dots, n$, the inverse of the covariance matrix of errors is $\sigma^{-2} V^{o-1}$, where V^{o-1} is found as

$$V^{o-1} = \begin{pmatrix} V^{-1(i)} & 0_{(n-1) \times p} & b \\ 0_{p \times (n-1)} & I_p & 0_{p \times 1} \\ b' & 0_{1 \times p} & 1 + \rho^2 \end{pmatrix} \tag{34}$$

where $V^{-1(i)}$ is the matrix V^{-1} in (12) without the i th row and the i th column, and b' is the i th row of the matrix V^{-1} without the i th case. Now for the regression model (33), the mean shift outlier model can be defined as

$$y^o = X^o \beta + z\gamma + \varepsilon^o \quad (35)$$

$$= X^{oo} \delta + \varepsilon^o$$

where $z = \begin{pmatrix} 0_{\tilde{n}-1} \\ 1 \end{pmatrix}$, $X^{oo} = (X^o \quad z)$ and $\delta = \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$. The

GLS estimate of β and the residual sum of squares under the hypothesis $H_0 : \gamma = 0$ are obtained as

$$\tilde{\beta}_0 = (X^{o'} V^{o-1} X^o)^{-1} X^{o'} V^{o-1} y^o \quad \text{and} \quad (36)$$

$$e^{o'} V^{o-1} e^o = y^{o'} (V^{o-1} - P_0) y^o$$

respectively, where $P_0 = V^{o-1} X^o (X^{o'} V^{o-1} X^o)^{-1} X^{o'} V^{o-1}$

and $\tilde{\beta}_0$ is the stochastic restricted Liu estimator $\tilde{\beta}_{srd}$ with AR(1) errors. Under the alternative hypothesis $H_1 : \gamma \neq 0$, the GLS estimates of β , γ and the corresponding residual vector are

$$\tilde{\delta} = \begin{pmatrix} \tilde{\beta}_1 \\ \tilde{\gamma}_1 \end{pmatrix} = (X^{oo'} V^{o-1} X^{oo})^{-1} X^{oo'} V^{o-1} y^o \quad \text{and}$$

$r^o = y^o - X^{oo} \tilde{\delta}$. Similarly, the residual sum of squares can be expressed as

$$r^{o'} V^{o-1} r^o = y^{o'} (V^{o-1} - P_1) y^o \quad (37)$$

where $P_1 = V^{o-1} X^{oo} (X^{oo'} V^{o-1} X^{oo})^{-1} X^{oo'} V^{o-1}$.

By the expressions obtained in equations (36) and (37), the difference between the residual sum of squares under H_0 and H_1 (that is the i th observation isn't an outlier or it is), equals

$$e^{o'} V^{o-1} e^o - r^{o'} V^{o-1} r^o = y^{o'} (P_1 - P_0) y^o$$

On the other hand, by the definition of X^{oo} and z in (35), we have

$$(X^{oo'} V^{o-1} X^{oo})^{-1} = \begin{pmatrix} X^{o'} V^{o-1} X^o & -\rho x_{i-1} - \rho x_{i+1} + (1 + \rho^2) x_i \\ -\rho x'_{i-1} - \rho x'_{i+1} + (1 + \rho^2) x'_i & 1 + \rho^2 \end{pmatrix}^{-1}$$

Using the theorem of the inverse of partitioned matrices [1], it is found that

$$X^{oo} (X^{oo'} V^{o-1} X^{oo})^{-1} X^{oo'} = X^o (X^{o'} V^{o-1} X^o)^{-1} X^{o'}$$

$$+ \frac{1}{1 + \rho^2 - p_{0i}} X^o (X^{o'} V^{o-1} X^o)^{-1} X^{o'} V^{o-1} z z' V^{o-1} X^o (X^{o'} V^{o-1} X^o)^{-1} X^{o'}$$

$$- \frac{1}{1 + \rho^2 - p_{0i}} z z' V^{o-1} X^o (X^{o'} V^{o-1} X^o)^{-1} X^{o'}$$

$$- \frac{1}{1 + \rho^2 - p_{0i}} (X^{o'} V^{o-1} X^o)^{-1} X^{o'} V^{o-1} z z' + \frac{z z'}{1 + \rho^2 - p_{0i}}$$

where p_{0i} is the last element of P_0 defined in (36).

Using the above expression and after some simplifications, $P_1 - P_0$ can be expressed as

$$P_1 - P_0 = \frac{1}{1 + \rho^2 - p_{0i}} (V^{o-1} - P_0) z z' (V^{o-1} - P_0). \quad (38)$$

Therefore, the difference between the residual sum of squares can be stated as

$$e^{o'} V^{o-1} e^o - r^{o'} V^{o-1} r^o = \frac{1}{1 + \rho^2 - p_{0i}} y^{o'} (V^{o-1} - P_0) z z' (V^{o-1} - P_0) y^o \quad (39)$$

By (32) and using the expressions given in (37) and (39), the following F -statistic can be used

$$F = \frac{(\tilde{n} - p - 1) y^{o'} (V^{o-1} - P_0) z z' (V^{o-1} - P_0) y^o}{(1 + \rho^2 - p_{0i}) y^{o'} (V^{o-1} - P_1) y^o} \quad (40)$$

which has F distribution with 1 and $\tilde{n} - p - 1$ degrees of freedom under H_0 , when the parameter ρ is known. In practice, the parameter ρ is unknown. By Judge et al. [28], an estimate of ρ ($\hat{\rho}$) can be used in place of the actual parameter value and the F -statistic given in (40) goes in distribution to $\chi^2_{(1)}$. Therefore, for finding the outliers using the mean shift outlier model (35), the $\chi^2_{(1)}$ value can be used as the critical value of the proposed F -statistic in (40).

Numerical Example

In this section, the dataset due to Bayhan and Bayhan [17] is considered to illustrate the theoretical results of the previous sections. There are 75 observations and the dataset consists of two independent variables: weekly list prices (averages from selected supermarkets) of the shampoos (X_1) and of a certain brand of soap, substituted for shampoos (X_2), and one response variable: the weekly quantities of shampoos sold. As the data are given for a high and irregular inflationary period, the sixty of observations are taken as historical data, and the fifteen observations of the last 15 weeks as fresh data. In what follows the rescaled data are used, and each variable is centered and scaled by the unit length scaling technique. Bayhan and Bayhan [17] have shown that the data are strongly collinear by finding the correlation matrix of regressor variables and the error terms have AR(1) structure by the Durbin-Watson statistic.

As mentioned before, for the case of AR(1) errors, the matrix V takes the form (11) which depends on the unknown parameter ρ . By using the historical data and the estimator given in (14), the parameter ρ is estimated as $\hat{\rho} = 0.707$. Then following Bayhan and Bayhan [17], we selected the observations 48 and 49 of

the transformed historical data, and constructed the stochastic linear restriction (4); that is we have chosen

$$R = \begin{pmatrix} 0.1450 & 0.1049 \\ 0.0077 & 0.1850 \end{pmatrix} \text{ and } r = \begin{pmatrix} 0.1303 \\ 0.1380 \end{pmatrix}.$$

By taking the first and the second elements of the first two rows and columns of the matrix V , the variance of the error term ϕ in stochastic linear restriction is determined as

$$\text{var}(\phi) = \sigma^2 \begin{pmatrix} (1-\rho^2)^{-1} & \rho(1-\rho^2)^{-1} \\ \rho(1-\rho^2)^{-1} & (1-\rho^2)^{-1} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1.999 & 1.413 \\ 1.413 & 1.999 \end{pmatrix}.$$

. Now in order to find the Liu biasing parameter d , we first consider the canonical model (16) and estimate the regression coefficient α as $\tilde{\alpha}' = (0.5, -2.8)$, then using the estimator of the parameter d given in (19), we have an estimate of d as $\hat{d} = 0.417$, where $\hat{\sigma}_2 = 0.006$.

Different estimators of β which were discussed in the preliminary section and the corresponding estimated standard errors are calculated and given in Table 1. For the estimator $\tilde{\beta}$, the estimated standard errors are obtained about 1.5 but using the Liu estimator ($\tilde{\beta}_d$), the estimated standard errors reduce to about 0.6. Furthermore, by taking into account additional observations and combining mixed estimator with the

Liu estimator ($\tilde{\beta}_{srd}$), the standard errors are about 0.1, which shows that lower standard errors of the estimated parameters are achieved. Also as mentioned in [8], these additional information corrects the wrong sign problem of the estimated coefficients.

For more investigation, the $DFBETAS$ values for the j th ($j=1,2$) element of $\tilde{\beta}_{srd}$ and the $DFFITs$ values are calculated and the results are presented in Figure 1. The Straight lines are the cutoff points. The cutoff points are chosen as $\frac{2}{\sqrt{n+p}}$ and $2\sqrt{\frac{p}{n}}$, which are 0.48

and 0.73, respectively, for this study. From Figure 1, one can find out that only the first observation is recognized as influential with respect to the $DFFITs$ criteria since the calculated value exceeds the cutoff point 0.73.

At last, we calculated the F -statistic in (40) for the fresh data. The results are given in Table 2. The first, the fourth and the fifteenth observations have the largest values of the F -statistic, which are 1.31, 1.25 and 2.17, respectively. For other observations, the F -statistic is obtained less than 1. Therefore, since $\chi^2_{(1)} = 3.84$ (for $\alpha = 0.05$), none of the observations can be considered as an outlier for this data set.

To reveal the efficiency of our methods, a shift is

Table 1. The estimated values of regression coefficients and their standard errors (in parenthesis) for $d = 0.417$

	β_1	β_2
$\tilde{\beta}$	2.38 (1.54)	-1.62 (1.54)
$\tilde{\beta}_d$	1.05 (0.65)	-0.61 (0.65)
$\tilde{\beta}_{srd}$	0.20 (0.16)	0.24 (0.15)

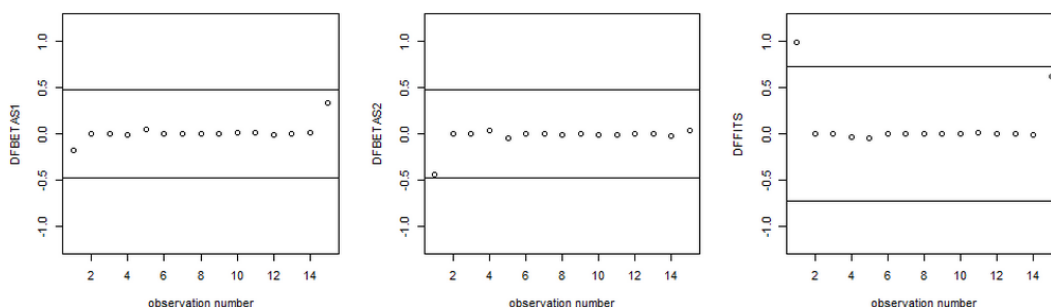


Figure 1. $DFBETAS$ for β_1 ($DFBETAS_1$) and $DFBETAS$ for β_2 ($DFBETAS_2$) and $DFFITs$ for fitted values by the estimator $\tilde{\beta}_{srd}$. The straight lines are cutoff points.

Table 2. F -statistic values by the estimator $\tilde{\beta}_{srd}$

Case	F -statistic	Case	F -statistic	Case	F -statistic
1	1.31	6	0.007	11	0.01
2	0.005	7	0.0001	12	0.02
3	0.002	8	0.03	13	0.01
4	1.25	9	0.001	14	0.0006
5	0.19	10	0.00001	15	2.17

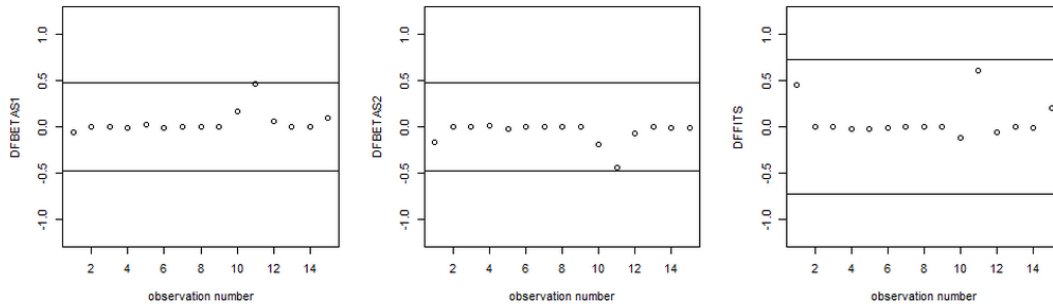


Figure 2. $DFBETAS$ for β_1 ($DFBETAS_1$) and $DFBETAS$ for β_2 ($DFBETAS_2$) and $DFFITS$ for fitted values by the estimator $\tilde{\beta}_{srd}$ after changing the 11th observation. The straight lines are cutoff points.

imposed on the dependent variable to investigate its impact on finding the outlier and on diagnosing influential observations. We have chosen a shift equal to 0.52 (which is about two standard deviation of y) and exerted on 11th observation. In fact, the chosen shift value is the largest possible value such that ensures us that the error terms of the rescaled shifted data are still of the form AR(1). For the new data set, the Liu biasing parameter is estimated as $d = 0.855$, where $\hat{\sigma}^2 = 0.018$. The corresponding calculated $DFBETAS$ and $DFFIT$ values are displayed in Figure 2.

As it is seen in Figure 2, the absolute value of the $DFBETAS$ and $DFFITS$ criteria for the 11th observation tends to be near the cutoff points but do not exceed the corresponding values 0.48 and 0.73. Also, the imposed shift on 11th observation has reduced the influence effect of the first observation on the estimated fitted values.

Also, the values of F -statistic for the shifted data is given in Table 3. It can be seen from table 3, the calculated F -statistic for the 11th observation is 4.73 and

for other observations, it is less than 3. So, since $7.70 > \chi^2_{(1)} = 3.84$, the 11th observation is correctly recognized as an outlier.

Simulation study

In what follows a simulation study is carried out to investigate the performance of the proposed mean-shift outlier model and to conduct a survey on changes in influential measures in the presence of influential observation. Following McDonald and Galarneau [29], two collinear explanatory variables are generated from the following equation:

$$x_{ij} = \left(\frac{1-\lambda}{\lambda}\right)^{1/2} z_{ij} + z_{i3}, \quad i = 1, \dots, t, \quad j = 1, 2$$

where z_{i1} , z_{i2} and z_{i3} are independent standard normal pseudo-random numbers, and λ represents the correlation between two explanatory variables x_1 and

Table 3. F -statistic value by the estimator $\tilde{\beta}_{srd}$ after changing the 11th observation

Case	F -statistic	Case	F -statistic	Case	F -statistic
1	0.65	6	0.003	11	4.73
2	0.0001	7	0.0004	12	2.59
3	0.00003	8	0.009	13	0.000004
4	0.36	9	0.002	14	0.003
5	0.04	10	2.92	15	0.21

Table 4. The probability of type I error and Power of the F test for the mean-shift outlier model for $(\beta_1, \beta_2) = (1, -0.5)$

n	λ	ρ	Sig. level	Power	
				Shift=3.5	Shift=4
15	0.9	0.6	0.025	0.898	0.946
		0.9	0.054	0.948	0.976
	0.99	0.6	0.034	0.919	0.947
		0.9	0.076	0.943	0.974
30	0.9	0.6	0.034	0.988	0.998
		0.9	0.057	0.995	0.999
	0.99	0.6	0.036	0.994	0.999
		0.9	0.070	0.992	0.998
70	0.9	0.6	0.050	0.997	1.000
		0.9	0.039	1.000	1.000
	0.99	0.6	0.066	0.998	1.000
		0.9	0.062	0.997	1.000
100	0.9	0.6	0.050	1.000	1.000
		0.9	0.044	0.996	1.000
	0.99	0.6	0.055	0.999	1.000
		0.9	0.054	0.997	1.000

x_2 , which is set to be 0.9 and 0.99 in this study. Also, the response variable is generated based on the model $y_i = \sum_{j=1}^2 \beta_j x_{ij} + \varepsilon_i, i = 1, \dots, t$, with $\varepsilon_i = \rho \varepsilon_{i-1} + u_i$, where $u_i \sim NID(0,1)$ and the initial value ε_0 is sampled from $N(0, (1 - \rho^2)^{-1})$. The vector of coefficients is set to $\beta = (1, -0.5)'$. With 0.6 and 0.9 values of ρ , 75, 90, 130 and 160 units of samples are produced for 1000 times. Each time the first 60 observations are taken as historical data, and after being rescaled by the unit length scaling technique, they are used for estimating the ρ parameter. The remaining (i.e. $n = 15, 30, 70$ and 100) are taken as fresh data. Also for the fresh data, the x_1, x_2 and y values are rescaled. Furthermore, the 59th and 60th historical data (which are the nearest observations to the fresh data) are used as the stochastic restrictions. In each iteration, the optimal d Liu parameter is found for the stochastic restricted Liu regression under AR(1). For different values of λ, ρ and n , the probability of type I error ($\alpha = 0.05$) is calculated using the method of mean-shift outlier model. For this purpose, the percentage of the time that the F-statistic for testing $H_0 : \gamma = 0$ is greater than the corresponding critical value is calculated and the results are shown in Table 4. A glance in the results of this table indicates that the significance level remains around 0.05.

Also, the power of the test for the mean-shift outlier model is investigated in Table 4 after shifting the

generated data set. At this point, in each dataset, the 4th fresh observation is taken as an outlier by exerting two shift values 3.5 and 4 to the original value of the response. These values are more than one time the mean of standard deviation of generated dependent observations and less than twice of it. Then the power of the test is calculated for the shifted data by the foregoing method. We remind that also here the shifted data are rescaled and before being used in the calculations, they're AR(1) structure is checked by the Durbin-Watson statistic. It is shown in Table 4 that for each combination of n, λ and ρ , the power of the test increases with the increase of the value of the shift. Also, it increases by the increase in the sample size.

In order to study the impact of the shifted 4th observation on the estimated regression coefficients β_1, β_2 the predicted values of the response variable, the mean of absolute values of the *DFBETAS* and *DFFITs* in 1000 replications, with and without stochastic linear restrictions, are computed. The results are shown in Tables 5 and 6 for the shift values 3.5 and 4. The proportion of times that the absolute values of calculated *DFBETAS* and *DFFITs* values exceed

$$\frac{2}{\sqrt{n+p}} \text{ and } 2\sqrt{\frac{p}{n}}$$

(in which $p = 2$ in this study) are reported in parenthesis.

From Tables 5 and 6, it can be seen that for larger sample sizes, the mean of $|DFBETAS|, |DFFITs|$ and their related proportions for the stochastic linear restrictions cases are smaller than the

Table 5. Mean of $|DFBETAS|$ and $|DFFITS|$ of the 4th observation for $shift = 3.5$ and in parenthesis the proportion of times the calculated criteria exceed the cutoff points.

n	λ	ρ	Mean of $ DFBETAS_1 $ (%)		Mean of $ DFBETAS_2 $ (%)		Mean of $ DFFITS $ (%)		
			of influential)	(% of influential)	of influential)	(% of influential)	of influential)	(% of influential)	
15	0.9	0.6	With	Without	With	Without	With	Without	
			Restriction	Restriction	Restriction	Restriction	Restriction	Restriction	
			0.884	0.860	0.852	0.821	1.325	1.291	
			(0.688)	(0.671)	(0.655)	(0.649)	(0.773)	(0.763)	
		0.9	0.926	0.897	0.903	0.882	1.339	1.302	
			(0.677)	(0.658)	(0.668)	(0.659)	(0.767)	(0.761)	
30	0.99	0.6	0.908	0.888	0.907	0.883	1.344	1.309	
			(0.694)	(0.695)	(0.694)	(0.684)	(0.777)	(0.768)	
		0.9	1.018	0.998	0.993	0.981	1.422	1.365	
			(0.715)	(0.715)	(0.699)	(0.710)	(0.775)	(0.764)	
	0.9	0.6	0.563	0.566	0.560	0.556	0.854	0.851	
			(0.6)	(0.603)	(0.617)	(0.613)	(0.709)	(0.704)	
		0.9	0.627	0.622	0.614	0.609	0.911	0.903	
			(0.636)	(0.632)	(0.644)	(0.637)	(0.726)	(0.722)	
	0.99	0.6	0.583	0.582	0.571	0.576	0.869	0.865	
(0.617)			(0.615)	(0.596)	(0.606)	(0.730)	(0.718)		
		0.9	0.623	0.642	0.623	0.635	0.912	0.935	
			(0.629)	(0.642)	(0.624)	(0.627)	(0.722)	(0.733)	
70	0.99	0.6	0.424	0.428	0.434	0.439	0.641	0.652	
			(0.648)	(0.652)	(0.662)	(0.659)	(0.754)	(0.763)	
		0.9	0.455	0.463	0.462	0.471	0.674	0.685	
			(0.665)	(0.669)	(0.680)	(0.687)	(0.750)	(0.765)	
		0.99	0.6	0.410	0.417	0.413	0.420	0.623	0.633
	(0.618)			(0.620)	(0.630)	(0.639)	(0.747)	(0.749)	
		0.9	0.426	0.440	0.436	0.445	0.637	0.654	
			(0.632)	(0.644)	(0.643)	(0.649)	(0.729)	(0.745)	
100	0.99	0.6	0.364	0.372	0.366	0.373	0.541	0.552	
			(0.637)	(0.647)	(0.65)	(0.655)	(0.77)	(0.765)	
		0.9	0.392	0.404	0.392	0.404	0.566	0.583	
			(0.662)	(0.672)	(0.676)	(0.68)	(0.744)	(0.754)	
		0.99	0.6	0.335	0.345	0.338	0.344	0.510	0.519
	(0.628)			(0.630)	(0.634)	(0.648)	(0.755)	(0.767)	
		0.9	0.358	0.371	0.362	0.373	0.539	0.551	
			(0.646)	(0.655)	(0.651)	(0.665)	(0.735)	(0.740)	

measures for the cases with no stochastic linear restrictions. This result may be due to the achievement of improved accuracy for the advanced model when the sample size is enough large, and its reduced sensitivity to the observation which do not have a large displacement from the bulk of data.

For fix values of λ and ρ , the mean value of $|DFBETAS|$ and $|DFFITS|$ decreases as n increases for both models. But not any significant effect on the proportion of detection of influential observation can be shown. This may happen because cutoff points have an inverse relationship with the sample size and they decrease by the increase of n . So although we have a decrease in the mean of absolutes of measures, but simultaneously a decrease in cutoff points happens and causes the proportions to remain almost unchanged. It can also be seen from Tables 5 and 6 that with the

increase in the degree of collinearity (λ from 0.9 to 0.99), the mean value of the measures increases for small samples while it decreases for large samples. Also, it is seen in all cases that increasing of ρ (the autocorrelation parameter) causes a growth in the mean value of $|DFBETAS|$, $|DFFITS|$ and related proportions of both models. At last, increasing the size of the shift values causes the measures and the related proportions to increase in all cases.

Discussion

In this article, a stochastic restricted Liu estimator is presented to reduce the effect of collinearity when there are also additional stochastic linear restrictions on the vector of parameters and when the data are correlated with AR(1) errors. The AR(1) error structure is the one

Table 6. Mean of $|DFBETAS|$ and $|DFFITs|$ of the 4th observation for $shift = 4$ and in parenthesis the proportion of times the calculated criteria exceed the cutoff points.

n	λ	ρ	Mean of $ DFBETAS_1 $ (% of influential)		Mean of $ DFBETAS_2 $ (% of influential)		Mean of $ DFFITs $ (% of influential)	
			With Restriction	Without Restriction	With Restriction	Without Restriction	With Restriction	Without Restriction
15	0.9	0.6	1.001 (0.729)	0.971 (0.720)	0.955 (0.680)	0.934 (0.679)	1.496 (0.840)	1.445 (0.821)
		0.9	1.021 (0.714)	1.016 (0.717)	0.989 (0.682)	0.973 (0.685)	1.476 (0.801)	1.448 (0.800)
	0.99	0.6	1.046 (0.754)	1.041 (0.744)	1.030 (0.726)	1.022 (0.727)	1.524 (0.829)	1.490 (0.818)
		0.9	1.162 (0.769)	1.130 (0.766)	1.151 (0.752)	1.119 (0.759)	1.590 (0.812)	1.537 (0.811)
30	0.9	0.6	0.619 (0.649)	0.618 (0.642)	0.602 (0.647)	0.601 (0.644)	0.936 (0.758)	0.931 (0.755)
		0.9	0.674 (0.669)	0.666 (0.656)	0.658 (0.660)	0.645 (0.652)	0.982 (0.753)	0.969 (0.745)
	0.99	0.6	0.661 (0.680)	0.654 (0.665)	0.644 (0.648)	0.645 (0.648)	0.970 (0.776)	0.952 (0.762)
		0.9	0.697 (0.675)	0.720 (0.678)	0.689 (0.656)	0.714 (0.676)	1.001 (0.748)	1.027 (0.765)
70	0.9	0.6	0.463 (0.675)	0.467 (0.682)	0.472 (0.677)	0.478 (0.684)	0.701 (0.781)	0.712 (0.787)
		0.9	0.499 (0.687)	0.508 (0.694)	0.558 (0.702)	0.516 (0.715)	0.740 (0.794)	0.753 (0.801)
	0.99	0.6	0.445 (0.642)	0.457 (0.650)	0.450 (0.659)	0.460 (0.668)	0.676 (0.772)	0.692 (0.776)
		0.9	0.467 (0.668)	0.481 (0.669)	0.467 (0.667)	0.484 (0.686)	0.695 (0.760)	0.718 (0.770)
100	0.9	0.6	0.401 (0.666)	0.410 (0.673)	0.403 (0.682)	0.411 (0.680)	0.596 (0.791)	0.609 (0.796)
		0.9	0.434 (0.694)	0.447 (0.702)	0.434 (0.700)	0.447 (0.711)	0.627 (0.779)	0.645 (0.785)
	0.99	0.6	0.368 (0.65)	0.376 (0.663)	0.369 (0.664)	0.376 (0.670)	0.563 (0.786)	0.572 (0.793)
		0.9	0.396 (0.665)	0.407 (0.681)	0.399 (0.687)	0.412 (0.692)	0.592 (0.760)	0.603 (0.762)

widely used in regressions with autocorrelated errors. By an example it is shown that using the proposed biased estimator, more computational accuracy will be gained, i.e. there can be found values of the biasing parameter d for which the estimator has smaller MSE than the GLS estimator and the Liu estimator. Measuring the influence should be done after controlling the collinearity, as suggested by Belsely et al. [2]. Therefore, we derived influence measures and mean-shift outlier method in the case of stochastic restricted Liu estimator with AR(1) errors. The performance of the results is shown and approved by a simulation study. It is shown that the sample size affects the power of the proposed test for finding outlier using the mean-shift outlier method. A larger sample size gives the test more power to detect an outlier. Furthermore, when the sample size is large with the

increase in the degree of collinearity, outliers seem to be less likely to be influential and have less influence on the regression fit or on the regression parameter estimates. The result is inversed for small samples. Also generally when the error terms have the higher autocorrelated parameter, outliers have more influence on the influence measures.

Acknowledgments

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. Seber G. and Lee A. *Linear Regression Analysis*. Wiley, New York, (2003).

2. Belsley D., Kuh E., and Welsch R. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Wiley, New York, (1980).
3. Theil H. and Goldberger A. On pure and mixed estimation in economics. *Int. Econ. Rev.*, **2**: 65-77 (1961).
4. Theil H. On the use of incomplete prior information in regression analysis. *J. Am. Stat. Assoc.*, **58**: 401-414 (1963).
5. Hoerl A. and Kennard R. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **12**: 55-67 (1970).
6. Liu K. A new class of biased estimate in linear regression. *Commun. Stat. Theory Methods*, **22**: 393-402 (1993).
7. Stein C. Inadmissibility of usual estimator for the mean of a multivariate normal distribution. *Proceedings of the third Berkley symposium on mathematical statistics and probability*, **1**: 197-206 (1956).
8. Ozkale M. A stochastic restricted ridge regression estimator. *J. Multivar. Anal.*, **100**: 1706-1716 (2009).
9. Hubert M. and Wijekoon P. Improvement of the Liu estimator in linear regression model. *Stat. Pap.*, **47**: 471-479 (2006).
10. Yang H. and Xu J. An alternative stochastic restricted Liu estimator in linear regression. *Stat. Pap.*, **50**: 639-647 (2009).
11. Box G. and Newbold P. Some comments on a paper by Coen, Gomme, and Kendall. *J. R. Stat. Soc. Ser. A*, **134**: 229-240 (1971).
12. Coen P., Gomme E., and Kendall M. Lagged relationships in economic forecasting. *J. R. Stat. Soc. Ser. A*, **132**: 133-152 (1969).
13. Ryan T. *Modern Regression Methods*. Wiley, New York, (1997).
14. Griffiths W., Hill R., and Judge G. *Learning and Practicing Econometrics*. Wiley, New York, (1993).
15. Aitken A. On least squares and linear combinations of observations. *Proceedings of Royal Statistical Society, Edinburgh*, **55**: 42-48 (1935).
16. Trenkler K. On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors. *J. Econom.* **25**: 179-190 (1984).
17. Bayhan G. and Bayhan M. Forecasting using autocorrelated errors and multicollinear predictor variables. *Comp. Ind. Eng.*, **34**: 413-421 (1998).
18. Kaciranlar S. Liu estimator in the general linear regression model. *J. Appl. Statist. Sci.*, **13**: 229-234 (2003).
19. Alkhamisi M. Ridge estimation in linear models with autocorrelated errors. *Commun. Stat. Theory Methods*, **39**: 2630-2644 (2010).
20. Zaherzadeh A., Rasekh A., and Babadi B. Diagnostic measures in ridge regression model with AR(1) errors under the stochastic linear restrictions. *J. Sci. I. R. Iran*, **29**: 67-78 (2018).
21. Roy S. and Guria S. Regression diagnostics in an autocorrelated model. *Braz. J. Probab. Stat.*, **18**: 103-112 (2004).
22. Cook R. and Weisberg S. *Residuals and Influence in Regression*. Chapman and Hall, (1882).
23. Ullah M., Pasha G., and Aslam M. Assessing influence on the Liu estimates in linear regression models. *Commun. Stat. Theory Methods*, **42**: 3100-3116 (2013).
24. Jahufer A. Detecting global influential observations in Liu regression model. *Open J. Stat.*, **3**: 5-11 (2013).
25. Alheety M. and Kibria B. On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors. *Surveys in Mathematics and its Applications*, **4**: 155-167 (2009).
26. Kim S. and Huggins R. Diagnostics for autocorrelated regression models. *Aust. N. Z. J. Stat.*, **40**: 65-71 (1998).
27. Tsai C. and Wu X. Assessing local influence in linear regression models with 1st order autoregressive or heteroscedastic error structure. *Stat. Probab. Lett.*, **14**:

247-252 (1992).

28. Judge G., Griffiths W., Hill R., Lutkepohl H., and Lee T. *The Theory and Practice of Econometrics*. Wiley, New York, (1985).
29. Mcdonald G. and Galarneau D. A monte carlo evaluation of some ridge-type estimators. *J. Am. Stat. assoc.*, **70**: 407-416 (1975).