Malaysian Journal of Mathematical Sciences 14(3): 505–517 (2020)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

On the Robust Parameter Estimation Method for Linear Model with Autocorrelated Errors in the Presence of High Leverage Points and Outliers in the Y-Direction

Ali, D. A. *[1] and Midi, H. [2,3]

[1]*Department of Statistics and Planning, Faculty of Economics, Simad University, Mogadishu, Somalia*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia*
[3]*Laboratory of Computational Statistics and Operation Research, Institute for Mathematical Research, Universiti Putra Malaysia, Malaysia*

E-mail: daahirxy@gmail.com
* *Corresponding author*

Received: 10 May 2019
Accepted: 12 May 2020

ABSTRACT

In the existence of autocorrelation problem, the Ordinary Least Squares (OLS) estimates become incompetent. The Cochrane - Orcutt Prais - Winsten iterative method (COPW) is the most widely used remedial measure to rectify this problem. However, this iterative procedure is based on the OLS estimates, which is not resistant and easily influenced by high leverage points (outliers in the x-direction) and outliers in the y-direction. The COPW based on MM estimator is developed to remedy both problems of autocorrelation and high leverage points. Nevertheless, MM estimator does not perform well in the presence of bad leverage points. In this paper, we propose to improvise the Cochrane-Orcutt

Prais-Winsten iterative method based on GM6 estimator so that autocorrelated errors and high leverage points can be rectified. The performance of the COPW-GM6 is scrutinized widely by Monte Carlo simulation and real example. The results of this study show that the COPW-GM6 is more efficient than the COPW and COPW-MM.

Keywords: Autocorrelation, bad leverage points, Cochrane-Orcutt Prais-Winsten iterative method (COPW), good leverage points, high leverage points (HLPs), outliers.

1. Introduction

The most commonly used technique for estimating the parameters of multiple linear regression model is the Ordinary Least Square method (OLS) due to its general acceptance, neat statistical properties and computational simplicity. The OLS estimates has many attractive properties under normality assumptions of regression errors. Nevertheless, in many occasions, the statistical practitioners just apply it without any rigorous checking. One of the assumptions that is not usually fulfilled is the random and uncorrelated errors. White and Brisbin (1980) stated that the OLS estimators lose their Best Linear Unbiased Estimators (BLUE) properties when the error term of the current observation is correlated with the error term of the previous observation i.e. $E(\epsilon_{i,}\epsilon_j) \neq 0$ or $cov(\epsilon_i, \epsilon_j) \neq 0$ for $i \neq j$. The true population variance (σ^2) is likely to be underestimated by the sample variance. As a result, the OLS coefficients estimates become less efficient in the presence of autocorrelated errors in the sense that the usual t and F tests of significance are no longer trustworthy. These tests may become statistically significance when in fact they are not. In addition to that, the coefficient of multiple determination, R^2 gets inflated which shows that the fitting is good but indeed it is not. Therefore, autocorrelated error terms may lead to misleading conclusions about the statistical significance of the estimated regression coefficients (see Gujarati and Porter (2009)). Hence, suitable remedial measures should be considered after detecting the existence of autocorrelation problem.

To correct the problem of autocorrelation , there are several remedial measures such as Cochrane-Orcutt iterative method and Cochrane-Orcutt Prais-Winsten two-step or iterative procedures, Durbin two-step procedure, and the Hildreth-Lu, scanning or search procedure which are based on a specifically estimated correlation coefficient see (Green, 2008) and (Gujarati and Porter, 2009)). Among these methods, Cochrane-Orcutt Prais-Winsten iterative method (COPW) is the most commonly used measure in econometrics to get OLS estimators with the BLUE properties. However, this method (COPW) is based on the OLS estimates, which is not robust and therefore easily influenced by high leverage points (outliers in the x-direction) (see Habshah (1999), Rana et al. (2012) and Sani et al. (2019)). Many statistics practitioners are not aware of the unduly effects of high leverage points on the OLS estimates. The COPW based on MM estimator is developed to remedy both problems of autocorrelation and high leverage points in Habshah et al. (2013). Nevertheless, MM estimator does not perform well in the presence of bad leverage points. In this paper, we propose to improvise the Cochrane-Orcutt Prais-Winsten iterative method based on GM6 estimator so that autocorrelated errors and high leverage points can be rectified simultaneously. This method is developed by integrating high asymptotic efficiency, high breakdown and bounded influence property of GM6 estimator in the Cochrane-Orcutt Prais-Winsten iterative method. We name this new method as Cochrane-Orcutt Prais-Winsten iterative method based on GM6 estimator (COPW-GM6).

2. The Proposed Robust COPW-GM6 Estimator

2.1 The COPW-GM6

It is important to mention that in each step of COPW iterative method, the ordinary least square method is employed to obtain the parameter estimates of the multiple linear regression model. In addition to that, the transformation to correct for autocorrelation problem does not remove the influence of HLPs on OLS method. Hence, COPW is anticipated to show inconsistent results in the existence of HLPs. COPW is modified so that it is robust and not affected by HLPs. The robust COPW is formulated by incorporating GM-estimator which has a breakdown point close to 0.5, a bounded influence function and a high asymptotic efficiency for the normal model (see Coakley and Hettmansperger (1993)) in the COPW procedure. We name this regression estimator as Cochrane-Orcutt Prais-Winsten iterative method based on GM6 estimator (COPW-GM6). In the COPW-GM6 estimator, the OLS estimator are substituted with GM6 estimator. COPW-GM6 is expected to be more consistent and precise relative to COPW, COPW-MM in the presence of bad leverage points.

Consider the following multiple linear regression model:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} +, \dots, + \beta_k x_{kt} + u_t \tag{1}$$

where u_t follows the first order autoregressive model:

$$u_t = \rho u_{t-1} + \epsilon_t, -1 \leq \rho \leq 1. \tag{2}$$

The proposed COPW-GM6 algorithm can be summarized as follows:

Step 1: Estimate the coefficients of the multiple linear regression in Equation (1) by using GM6-estimator and get the residuals, u_t.

Step 2: Using the residuals obtained in Step 1, regress the following equation by using GM6-estimator and get the ρ.

$$u_t = \rho u_{t-1} + \epsilon_t. \tag{3}$$

Step 3: Using the estimate of ρ found in Step 2, estimate the following generalized difference equation by using the GM6-estimator.

$$y_t^* = \beta_0^* + \beta_1^* x_{1t}^* + \beta_2^* x_{2t}^* + \ldots + \beta_k^* x_{kt}^* + \epsilon_t \tag{4}$$

where $y_t^* = (y_t - \hat{\rho} y_{t-1})$, $x_{jt}^* = (x_{jt} - \hat{\rho} x_{jt-1})$, $\beta_0^* = \beta_0(1 - \hat{\rho})$ and $\beta_j^* = \beta_j$ for $j = 1, 2, \ldots, k$.

Step 4: Substitute the values of $\hat{\beta}_0^*$ and $\hat{\beta}_j^*$ for $j = 1, 2, \ldots, k$ found in Equation (4) and hence obtain the new residuals.

Step 5: Repeat Step 2 to Step 4 until the successive estimates of ρ differ by less than 0.00001 or maximum of 20 iterations if the convergence criteria is not met.

In the differencing procedure the first observation is lost as it has no antecedent. To prevent this loss of one observation, the first observation on y and x_j is transformed by the following formula:

$$y_1^* = y_{1\sqrt{1-\hat{\rho}^2}} \text{ and } \; x_{j1}^* = x_{j1\sqrt{1-\hat{\rho}^2}}.$$

2.2 The MM-Estimator

Yohai (1987) introduced MM estimator which employed more than one M-estimation process to find the final estimates. The procedure of MM-estimator is summarized as follows (see Rousseeuw and Leroy (2003)):

Step 1: By employing a high breakdown estimator such as S-estimator with Huber or bisquare weight function, compute the initial estimates of the coefficients and the corresponding residuals r_i, i=1,2,. . . , n.

Step 2: By using M-estimator, compute M-estimate of scale, $\hat{s_0}$ by minimizing the function

$$min \sum_{i=1}^{n} \rho_0 \left\{ \frac{y_i - x_i^T \hat{\beta}_0}{\hat{s_0}} \right\}.$$

***Step 3*:** Based on the appropriate re-descending function again use the M estimator using the iterative procedure starting at $\hat{\beta}_0$. That is, compute the MM-estimate, $\hat{\beta}$ which is obtained from the solution of

$$min \sum_{i=1}^{n} \rho_1 \left\{ \frac{y_i - x_i^T \hat{\beta}}{\hat{s}} \right\}.$$

Yohai (1987) noted that $\rho_0(r)$ and $\rho_1(r)$ can be taken as $\rho_0\left(\frac{r}{k_0}\right)$ and $\rho_1\left(\frac{r}{k_1}\right)$, respectively. Selecting $k_0 = 0.212$ and $k_1 = 0.9014$ will guarantee a high breakdown estimate and will result in 95% efficiency at normal errors, respectively. Generally, the MM estimate is resistant estimator except for the case of high leverage points (see Hekimoglu and Erenoglu (2013)).

2.3 The GM6-Estimator

Coakley and Hettmansperger (1993) proposed GM6 estimator which is defined as a solution of normal equations as follows:

$$\sum_{i=1}^{n} \pi_i \psi \left\{ \frac{y_i - x_i^t \hat{\beta}}{\hat{\sigma} \pi_i} \right\} x_i = 0$$

where $\psi = \rho^{'}$ is a derivative of redescending function (weight function) and $\pi_i, i = 1, 2, \ldots, n$ is the i^{th} initial weight element of the diagonal matrix W, $\hat{\sigma}$ is the scale estimate, and $\hat{\beta}$ is the vector of parameters estimates.

***Step 1*:** Calculate the residuals (r_i) based on Least Trimmed of Squares (LTS) estimator.

***Step 2*:** Compute the estimated scale (σ) of the residuals, $s = (1.4826)(1 + \frac{5}{(n-p-1)})($ median ($|r_i|$), where r_i is obtained from Step 1.

***Step 3*:** Compute the standardized residuals (e_i), where, $e_i = r_i/s$.

***Step 4*:** Calculate the initial weight, denoted as π_i, where $\pi_i = min\left\{1, \frac{\chi^2_{0.95,k}}{RMD(MVE)}\right\}$, where RMD, MVE and k are Robust Mahalanobis Distance, Minimum Volume Ellipsoid and the number of predictors in the model, respectively.

***Step 5*:**: Compute the bounded influence function, $t_i = e_i/\pi_i$.

***Step 6*:**: Compute one-step Newton Raphson to obtain the GM6 estimates.

3. Results and Discussion

3.1 Monte Carlo Simulation Study

A Monte Carlo simulation study is performed to examine the performance of our new proposed method (COPW-GM6) with some existing methods (OLS, COPW and COPW-MM) in the simultaneous presence of autocorrelation problem, high leverage points(in this case bad leverage points) and outliers in the y-direction. In this study, a multiple linear regression model with two predictor variables and different sample sizes, that is n = 20, 40, 60 and 80 are considered. For each sample size n = 20, 40, 60 and 80, the clean data are generated using the following relationship (see Habshah et al. (2013)):

$$y_t = 1 + 2x_{1t} + 3x_{2t} + u_t \tag{5}$$

where x_1 and x_2 are generated from uniform distribution, U(0,1). In order to make sure that the simulated data have autocorrelated errors, the error term u_t is generated by using the following first order autoregressive scheme:

$$u_t = 0.9u_{t-1} + \epsilon_t \tag{6}$$

with an initial value of u_1 is generated from Normal Distribution, N(0,4) and the white noise,ϵ_t is generated from Normal Distribution, N(0,1).

In order to create 5% and 10% HLPs in both x_1 and x_2 and in y-directions, certain clean observations are replaced by contaminated observations in each sample size. The contaminated observations in both x_1 and x_2 directions as well as in y-directions are generated from U(4.9,5), U(9.9,10) and U(29.9,30) respectively. To compare the performance of COPW-GM6 with COPW, COPW-MM and OLS, the Absolute Bias, standard errors and RMSE of parameter estimates are obtained based on the average of R simulation runs where is R is the number of replications. The results of clean and contaminated data are displayed in Tables 1, 2 and 3 respectively. It can be clearly seen from Table 1 that OLS performs poorly and has inflated standard errors in the presence of only autocorrelation problem. On the other hand, the COPW, COPW-MM and COPW-GM6 are quite similar as the sample size increases. However, the COPW is slightly better than COPW-MM and COPW-GM6 when there is only autocorrelation problem in the data.

Table 1: SE, Absolute Bias and RMSE of estimates, autocorrelated errors, no outliers ($\alpha = 0\%$).

Method	Coef	$n = 20$			$n = 40$		
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	1.4404	0.0039	1.4404	1.2049	0.0387	1.2100
	$\hat{\beta}_2$	1.4600	0.0521	1.4609	1.1544	0.0043	1.1544
COPW	$\hat{\beta}_1$	0.1467	0.0440	0.1532	0.3778	0.0327	0.3792
	$\hat{\beta}_2$	0.1693	0.1426	0.2214	0.3898	0.0007	0.3898
COPW-MM	$\hat{\beta}_1$	0.7184	0.0158	0.7186	0.4736	0.0018	0.4736
	$\hat{\beta}_2$	0.7834	0.0204	0.7837	0.4539	0.0034	0.4539
COPW-GM6	$\hat{\beta}_1$	1.1906	0.0341	1.1910	0.4718	0.0114	0.4719
	$\hat{\beta}_2$	1.1626	0.0080	1.1626	0.4719	0.0043	0.4719
Method	Coef	$n = 60$			$n = 80$		
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	0.9785	0.0097	0.9786	0.8529	0.0043	0.8529
	$\hat{\beta}_2$	0.9526	0.0012	0.9527	0.8627	0.0355	0.8694
COPW	$\hat{\beta}_1$	0.3338	0.0097	0.3339	0.2957	0.0156	0.2961
	$\hat{\beta}_2$	0.3441	0.0004	0.3441	0.2983	0.0084	0.2994
COPW-MM	$\hat{\beta}_1$	0.3603	0.0063	0.3604	0.2995	0.0007	0.2995
	$\hat{\beta}_2$	0.3585	0.0178	0.3589	0.3142	0.0018	0.3142
COPW-GM6	$\hat{\beta}_1$	0.3706	0.0016	0.3706	0.3139	0.0052	0.3140
	$\hat{\beta}_2$	0.3675	0.0070	0.3675	0.3094	0.0005	0.3094

Let's now discuss the results when the generated data has autocorrelation problem, HLPs (bad leverage points) and outliers in the y-directions. Tables 2 and 3 show the results for regression estimates when the simulated data have autocorrelation, HLPs and outliers in the y-direction. We can clearly observe that the values of SE and RMSE for OLS and COPW are higher than the values of SE and RMSE for other robust regression estimates for all possible combinations of sample sizes n and α. On the other hand, we observe that in the presence of autocorrelation and different percentages of contamination points in x and y directions, COPW-GM6 and COPW-MM are the best regression methods relative to other methods. However, COPW-GM6 outperforms COPW-MM.

Table 2: SE, Absolute Bias and RMSE of estimates, autocorrelated errors, no outliers ($\alpha = 5\%$).

Method	Coef	$n = 20$			$n = 40$		
		SE	Bias	RMSE	SE	Bias	RMSE
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	4.9742	1.4905	5.1928	3.3549	1.5073	3.6779
	$\hat{\beta}_2$	2.3786	3.3274	4.0902	1.5999	3.3179	3.6835
COPW	$\hat{\beta}_1$	5.3465	1.6529	5.5961	1.6574	2.7366	3.7463
	$\hat{\beta}_2$	2.5446	3.2533	4.1303	2.5584	0.3945	1.2737
COPW-MM	$\hat{\beta}_1$	1.3666	9.8112	9.9059	0.6127	0.0913	0.6195
	$\hat{\beta}_2$	1.6830	1.4158	2.1994	0.8493	0.1855	0.8693
COPW-GM6	$\hat{\beta}_1$	1.0695	0.0423	1.0703	0.5066	0.0245	0.5072
	$\hat{\beta}_2$	1.0438	0.0606	1.0456	0.4946	0.0055	0.4946
Method	Coef	$n = 60$			$n = 80$		
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	2.7988	1.6814	3.2651	2.2577	1.6598	2.8022
	$\hat{\beta}_2$	1.3273	3.2342	3.4960	1.0938	3.2538	3.4328
COPW	$\hat{\beta}_1$	1.6952	1.5372	2.2884	1.1902	1.4679	1.8898
	$\hat{\beta}_2$	0.8084	3.0974	3.2012	0.5908	3.0444	3.1012
COPW-MM	$\hat{\beta}_1$	0.4163	0.0524	0.4196	0.3479	0.0794	0.3569
	$\hat{\beta}_2$	0.4595	0.0990	0.4700	0.3736	0.1095	0.3893
COPW-GM6	$\hat{\beta}_1$	0.4070	0.0087	0.4071	0.3438	0.0061	0.3438
	$\hat{\beta}_2$	0.4156	0.0209	0.4161	0.3815	0.0415	0.3639

Table 3: SE, Absolute Bias and RMSE of estimates, autocorrelated errors, no outliers ($\alpha = 10\%$).

Method	Coef	$n = 20$			$n = 40$		
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	7.0178	1.8515	7.2579	4.5385	1.4434	4.7625
	$\hat{\beta}_2$	3.3197	3.3671	4.7284	2.1543	3.5418	4.1455
COPW	$\hat{\beta}_1$	5.7136	1.6530	5.9479	2.4775	1.5703	2.9333
	$\hat{\beta}_2$	2.7147	3.3100	4.2808	1.1784	3.1947	3.4051
COPW-MM	$\hat{\beta}_1$	1.7822	0.6085	1.8832	0.8025	0.3401	0.8716
	$\hat{\beta}_2$	1.8925	1.2590	2.2731	1.2311	0.7244	1.4285
COPW-GM6	$\hat{\beta}_1$	1.1088	0.0206	1.1090	0.6099	0.0335	0.6108
	$\hat{\beta}_2$	1.1050	0.0303	1.1054	0.6485	0.0677	0.6521
Method	Coef	$n = 60$			$n = 80$		
		SE	Bias	RMSE	SE	Bias	RMSE
OLS	$\hat{\beta}_1$	3.7730	1.5536	4.0804	2.6900	2.4047	3.6082
	$\hat{\beta}_2$	1.7922	3.4912	3.9244	2.2389	2.7624	3.5559
COPW	$\hat{\beta}_1$	1.6978	1.4571	2.2373	1.0750	2.2737	2.5151
	$\hat{\beta}_2$	0.8189	3.1372	3.2423	0.8938	2.7709	2.9115
COPW-MM	$\hat{\beta}_1$	0.5360	0.2591	0.5954	0.4863	0.2979	0.5703
	$\hat{\beta}_2$	0.8123	0.4300	0.9192	0.4886	0.4239	0.6469
COPW-GM6	$\hat{\beta}_1$	0.4795	0.0585	0.4831	0.3799	0.0115	0.3801
	$\hat{\beta}_2$	0.6041	0.1379	0.6196	0.3669	0.0079	0.3670

3.2 U.S Consumption Expenditure Data set

The real data set to be discussed here is the U.S consumption expenditure data set for the period 1947 through 2000 taken from Gujarati and Porter

(2009). The data set shows the real consumption expenditures in billions of dollars (y, real disposable income in billions of dollars (x_1) and real wealth in billions of dollars (x_2) for the years between 1947 to 2000). The existence of autocorrelation problem in the original U.S consumption expenditure data set is evident since most of the points cluster around the first and the third quadrants as shown in Figure 1. This is confirmed by Breusch (1978), Godfrey (1978) and Modified Breusch-Godfrey (MBG) of Lim and Midi (2014) as depicted in Table 4 since the p-values for both BG and MBG are significant at 0.05 significance level.

Table 4: BG and MBG for Auction Price Data Set

Method	Test statistic	P-value
BG	8.2369	0.0041
MBG	11.4020	0.0007

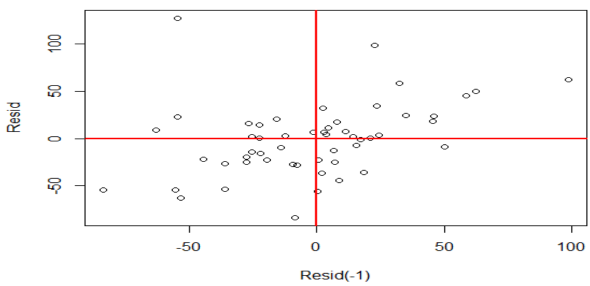

Figure 1: Current Residuals (Res1) Versus Lagged Residuals (Res (-1)) for U.S consumption expenditure

We assess the merit of OLS, COPW, COPW-MM and COPW-GM6 in both scenarios, original and modified U.S consumption expenditure data set. To create bad leverage points as well as outliers in the y-direction, the ith observation of the response variable and explanatory variables are replaced by an arbitrarily a large values. To make sure these large values are really bad leverage points, DRGP-MGt plot of Alguraibawi et al. (2015) is employed as displayed in Figure 2. This diagnostic plot classifies the observations into regular observations, vertical outliers, good and bad leverage points for the U.S consumption expenditure data set. Figure 3 displays the DRGP versus MGt plot for the modified U.S consumption expenditure data set. It can be clearly seen that case 17 is a bad leverage point while case 54 is a vertical outlier. Since it is confirmed that the modified U.S consumption expenditure data set has bad leverage point and an outlier in the y-directions, we now proceed to examine the performance of our proposed new method relative to some existing methods.

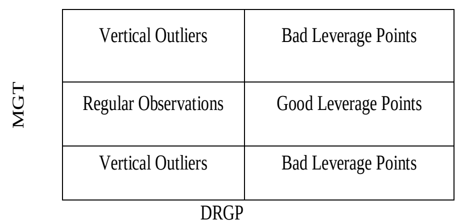

Figure 2: Plot of DRGP against MGt

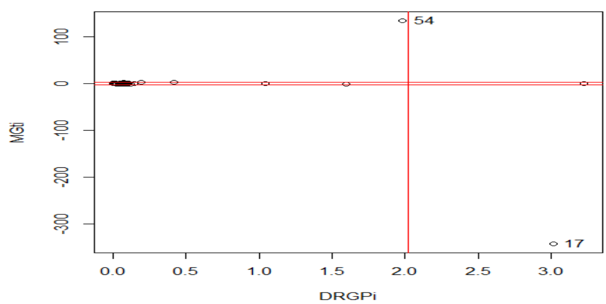

Figure 3: DRGP versus MGt for Modified U.S consumption expenditure Data Set

The parameter estimates and their standard errors of OLS, COPW, COPW-MM and COPW-GM6 regression estimators for original and modified U.S consumption expenditure data set are exhibited in Tables 5 and 6, respectively. From Table 5, as expected, the standard errors of the parameter estimates obtained by COPW, COPW-MM, and COPW-GM6 regression estimators are fairly close to each other in the presence of only autocorrelation problem and no outliers in the data.

Table 5: Estimates and standard deviation of original U.S consumption expenditure data set.

Method		β_1	β_2
OLS	Est.	0.7264	0.0364
	S.E	0.0139	0.0025
COPW	Est.	0.7417	0.0339
	S.E	0.0210	0.0038
COPW-MM	Est.	0.7422	0.0323
	S.E	0.0178	0.0032
COPW-GM6	Est.	0.7736	0.0238
	S.E	0.0339	0.0074

On the other hand, COPW performs poorly and have inflated standard errors in the presence of autocorrelation, bad leverage point and outliers in the y-direction, see Table 6. The parameter estimates of COPW regression estimator for modified data are completely different from the one found in the clean data set. It is interesting to see that not only the values but also the signs of the parameter estimates changed. The parameter estimate of β_1 obtained by COPW estimator in the original data set is 0.7417 while the parameter estimate of β_2 obtained by COPW estimator when the data set is modified in x_1, x_2 and the y- directions is -0.5860. Similarly, the estimated value of β_2 obtained by COPW estimator in the original dataset is 0.0339 whereas the parameter estimate of β_2 obtained by COPW estimator when the data set is modified in x_1, x_2 and the y-directions is 0.1237 which is not stable. On the contrary, the COPW-MM and COPW-GM6 estimators provide close parameter estimates of β_1 and β_2 to the parameter estimates obtained in the original data set when it is applied to the modified data set to estimate the regression coefficients. This indicates that the parameter estimates based on COPW-MM and COPW-GM6 estimators are stable. Furthermore, the standard errors of $\hat{\beta_1}$ and $\hat{\beta_2}$ based on COPW-GM6 estimator is less than the standard errors of $\hat{\beta_1}$ and $\hat{\beta_2}$ based on COPW-MM estimator in the modified data set. This exhibits that the parameter estimates based on COPW-GM6 estimator are more efficient than those obtained by COPW-MM estimator.

Table 6: Estimates and standard deviation of modified U.S consumption expenditure data set.

Method		β_1	β_2
OLS	Est.	-0.9969	0.2540
	S.E	0.5447	0.1053
COPW	Est.	-0.5860	0.1237
	S.E	0.6970	0.1386
COPW-MM	Est.	0.7434	0.0320
	S.E	0.0160	0.0032
COPW-GM6	Est.	0.7617	0.0265
	S.E	0.0073	0.0021

4. Concluding Remarks

In the presence of both autocorrelation and outliers (bad leverage points and vertical outliers), it is known that the multiple linear regression model using the OLS and COPW methods are severely affected. The COPW can only handle the problem of autocorrelation but not outliers. The COPW-MM can remedy both problems of autocorrelation and vertical outliers but not resistant to high leverage points. Hence, the gist of this study was to develop robust Cochrane-Orcutt Prais-Winsten (COPW) iterative method for multiple linear regression model with autocorrelated errors in the presence of high leverage points (i.e. bad leverage points) and outliers in the y-direction. In this respect, we proposed robust Cochrane-Orcutt Prais-Winsten (COPW) iterative method based on GM6-estimator, namely (COPW-GM6). Through Monte Carlo simulation study conducted followed by a real data set example, we showed that COPW performs better than some existing methods in the presence of autocorrelated errors, vertical outliers and bad leverage points.

References

Alguraibawi, M., Midi, H., and Imon, A. H. M. R. (2015). A new robust diagnostic plot for classifying good and bad high leverage points in a multiple linear regression model. *Mathematical Problems in Engineering*, Volume 2015. Article ID 279472, 12 pages, https://doi.org/10.1155/2015/279472.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355.

Coakley, C. W. and Hettmansperger, T. P. (1993). A bounded influence, high breakdown, efficient regression estimator. *Journal of the American Statistical Association*, 88(423):872–880.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 48(6):1293–1301.

Green, W. H. (2008). *Econometric Analysis.* New Jersey: Pearson Education Inc.

Gujarati, D. N. and Porter, D. C. (2009). *Basic Econometrics.* New York: McGraw-Hill Education.

Habshah, M. (1999). Preliminary estimators for robust non-linear regression estimation. *Journal of Applied Statistics*, 26(5):591–600.

Habshah, M., Lim, H. A., and Rana, S. (2013). On the robust parameter estimation for linear model with autocorrelated errors. *Advanced Science Letters*, 19(8):2494–2496.

Hekimoglu, S. and Erenoglu, R. (2013). A new gm-estimate with high breakdown point. *Acta Geodaetica et Geophysica*, 48:419–437. https://doi.org/10.1007/s40328-013-0029-1.

Lim, H. A. and Midi, H. (2014). The performance of robust modification of breusch-godfrey test in the presence of outliers. *Science and Technology*, 22(1):81–93.

Rana, S., Midi, H., and Imon, A. (2012). Robust wild bootstrap for stabilizing the variance of parameter estimates in heteroscedastic regression models in the presence of outliers. *Mathematical Problems in Engineering*, 12(7):1–14. https://doi.org/10.1155/2012/730328.

Rousseeuw, P. J. and Leroy, A. M. (2003). *Robust regression and outlier detection*. New York: John Wiley and Sons.

Sani, M., Midi, H., and Arasan, J. (2019). The performance of robust heteroscedasticity consistent covariance matrix estimator. *Malaysian Journal of Mathematical Sciences*, 13(April):71–88.

White, G. and Brisbin, I. (1980). Estimation and comparison of parameters in stochastic growth models for barn owls. *Growth*, 44(2):97–111.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, 15(20):642–656.