



The Performance of Robust Modification of Breusch-Godfrey Test in the Presence of Outliers

Lim, H. A.^{1,2*} and Midi, H.¹

¹Mathematics Department, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

²School of Applied Science and Foundation Studies, Kuala Lumpur Infrastructure University College, Jalan Ikram-Uniten, 43000 Kajang, Malaysia

ABSTRACT

Autocorrelation problem causes unduly effects on the variance of Ordinary Least Squares (OLS) estimates. Hence, it is very essential to detect the autocorrelation problem so that appropriate remedial measures can be taken. The Breusch-Godfrey (BG) test is the most popular and commonly used test for the detection of autocorrelation. Since this test is based on the OLS estimates, which are not robust, it is easily affected by outliers. In this paper, we propose a robust Breusch-Godfrey (MBG) test which is not easily affected by outliers. The results of the study indicate that the MBG test is more powerful than the BG test in the detection of autocorrelation problem.

Keywords: Autocorrelation, outliers, robust Breusch-Godfrey test

INTRODUCTION

Many statisticians employ the Ordinary Least Squares (OLS) method to estimate the parameters of a linear model because of ease of computation. In many occasions, the assumptions of random and uncorrelated errors are taken for granted by statisticians without any rigorous check. These assumptions

may not be true most of the time. The residuals may be correlated with the previous errors, which means that $E(u_i, u_j) \neq 0$ or $\text{cov}(u_i, u_j) \neq 0$ for $i \neq j$. Many statistics practitioners are not aware of the consequences of the autocorrelation problem. In specific, it ruins the important properties of OLS (Grassian & Boer, 1980; White & Brisbon, 1980). The OLS estimators are no longer the Best Linear Unbiased Estimators (BLUE) in the sense that the residual variance $\hat{\sigma}^2$ is likely to be underestimated, the true σ^2 . Hence, less efficient estimates are obtained as a result of employing an incorrect model based on the erroneous assumption.

Article history:

Received: 30 May 2011

Accepted: 20 January 2012

E-mail addresses:

lim_ha@yahoo.com (Lim, H. A.),

habshahmidi@gmail.com (Midi, H.)

*Corresponding Author

Additionally, the usual t and F tests of significance are no longer persuasive. These tests tend to be statistically significant when in fact they are not. The coefficient of determination, R^2 , becomes inflated. As such, the estimator will look more accurate as compared to its actual value. All these problems contribute to the failure of the hypothesis testing. Hence, the autocorrelation problem will most likely give misleading conclusions about the statistical significance of the estimated regression coefficients (Gujarati & Porter, 2009). Therefore, it is very important to detect the presence of autocorrelation.

Many graphical methods have been developed and they are now available in the literature for detecting autocorrelation (Davidson & MacKinnon, 1998; Gujarati & Porter, 2009; Mirer, 1995; Murray, 2006). However, due to the fact that diagnostic plots can be very subjective, it is necessary to have some statistical methods to detect the problem of autocorrelation. Rigorous procedures for testing the autocorrelation of data have also been suggested in the literature (see Breusch, 1978; Durbin & Watson, 1951; Godfrey, 1978; Hosking, 1980; Hosking, 1981; Mirer, 1995; Murray, 2006). Most of these techniques are based on the OLS estimation.

The Breusch-Godfrey (BG) test is the most commonly used method to detect the presence of autocorrelation. It was developed by Breusch (1978) and Godfrey (1978). This test has many practical points than other existing tests of autocorrelation such as Durbin-Watson Test, Runs Test, and Portmanteau Test. First, it allows for nonstochastic regressors. Secondly, the regressors included in the regression may contain lagged values of the regressand Y , that is Y_{t-1} , Y_{t-2} , etc. These lagged values may also appear as explanatory variables in the model. Thirdly, it allows the lagged values of the regressand to follow higher-order autoregressive scheme such as AR(1), AR(2), etc. Other existing tests are not applicable in these circumstances (Breusch, 1978; Godfrey, 1978; Gujarati & Porter, 2009; Mirer, 1995; Murray, 2006).

Suppose

$$Y_t = X_t\beta + u_t, \quad (1)$$

if the error term u_t follows the p th-order autoregressive, AR(p), scheme

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t \quad (2)$$

where ε_t is a white noise error term that satisfies all the classical assumptions.

Then, the null hypothesis, H_0 , to be tested is:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0 \quad (3)$$

that is, there is no serial correlation between $u_t, u_{t-1}, \dots, u_{t-p}$ of any order.

The procedures of the BG test are as follows:

Step 1: Estimate the coefficients of Eq. 1 by the OLS and obtain the estimated residual, \hat{u}_t .

Step 2: Regress \hat{u}_t on the original X_t and lagged values of the estimated residuals in Step (1). In summary, the following auxiliary regression is carried out:

$$\hat{u}_t = X_t \alpha + \hat{\rho}_1 \hat{u}_{t-1} + \hat{\rho}_2 \hat{u}_{t-2} + \dots + \hat{\rho}_p \hat{u}_{t-p} + \varepsilon_t \tag{4}$$

where α is the regression coefficients of matrix X .

Step 3: Obtain R^2 from the above auxiliary regression. R^2 is given by:

$$R^2 = \frac{SSR}{SST}, \tag{5}$$

where SSR is the sum of the squared regression and SST is the sum of the squared total of the auxiliary regression.

When the sample size is large, the statistic $(n - p)R^2$ is asymptotically following the Chi-squared distribution with a degree of freedom of p , that is $(n - p)R^2 \sim \chi_p^2$. The null hypothesis is rejected if the statistic $(n - p)R^2$ exceeds the Chi-square value at the level of significant, which means at least one ρ_i in Eq. 2 is statistically significantly different from zero.

In this article, a simple linear regression with autocorrelated errors are considered, as follows:

$$Y_t = \beta_1 + \beta_2 X_t + \mu_t, \tag{6}$$

and the error term is set to follow the first-order autoregressive AR(1) scheme,

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad -1 < \rho < 1. \tag{7}$$

The auxiliary regression to be examined is therefore simplified to:

$$\hat{u}_t = X_t \alpha + \hat{\rho} \hat{u}_{t-1} + \varepsilon_t \tag{8}$$

Since this test is based on the OLS estimates, it is suspected to be easily affected by the outliers. It is now evident that the outlier(s) have an unduly effect on the OLS estimates (Midi, 1999; Habshah *et al.*, 2009; Rana *et al.*, 2008; Riazosham *et al.*, 2010).

In this paper, an attempt was made to robustify the Breusch-Godfrey test by incorporating the high efficient and high breakdown MM-estimator (Yohai, 1987) in the formulation of the new robust test for the identification of autocorrelation problem. We called this new test the Modified Breusch-Godfrey test (MBG). Real data and simulation experiments show that the proposed MBG outperforms the classical BG test in detecting autocorrelation in the presence of outliers.

MATERIALS AND METHODS

We have briefly discussed the Breusch-Godfrey (BG) test for autocorrelation detection. The BG test uses the OLS to estimate the regression coefficient, so we expect it to suffer a huge setback when outliers are present in the data. Therefore, we propose a test which is robust against outliers. Here, we propose a new test which is a modification of Breusch-Godfrey test. We first identify the components of the BG test that are affected by the outliers and then

replace them with robust alternative. From the preceding procedures, we can see that the BG test requires two times of minimizing the sum of squares residuals to get the estimated coefficients. Firstly, we regress the original regression and then regress the auxiliary regression. Edgeworth (1887) has proven that squaring of the residual causes the least square to become extremely vulnerable to the presence of outliers. Therefore, the coefficients obtained are easily affected by the outlier. The MM-estimators introduced by Yohai (1987), which combined high-breakdown point and a high efficiency, are incorporated into the BG test. The robustified BG test is proven to minimize the impact of outliers on the regression model. This test is called the Modified Breusch-Godfrey test, or in short, MBG.

The proposed MBG test is summarized as follows:

Step 1: Unlike the classical BG test, we estimate the coefficients of the two variables regression by MM-estimator and get the residuals, \hat{u}_t .

Step 2: Regress X_i on the original X_t and \hat{u}_{t-1} or run the auxiliary regression stated in equation (8) by the MM-estimator.

Step 3: Find R^2 from the auxiliary regression in Step 2. R^2 for MBG test is defined as:

$$R^2 = \frac{SSR}{(SSE + SSR)}, \quad (9)$$

where, SSE is the sum of squared errors and SSR is the sum of squared regression of the auxiliary regression.

The null hypothesis of no serial correlation between μ_t and μ_{t-1} will be rejected if the statistic $(n-1)R^2$ exceeds the Chi-square value at 0.05 significant level.

RESULTS AND DISCUSSION

In this section, a few real world examples and a simulation study are presented to demonstrate the advantage of using the proposed Breusch-Godfrey test over the classical Breusch-Godfrey test in detecting serial autocorrelation problems.

Indexes of Real Compensation and Productivity Data

The first example is the Indexes of Real Compensation and Productivity data by Gujarati and Porter (2009). The data set contains 46 observations that give the Index of Output (X) and the Index of Real Compensation per hour (Y) in U.S from 1960 to 2005. The data are shown in Table 1.

In this study, the performances of the classical BG test and the MBG test in the original data and contaminated data sets were examined. Three types of contaminated data sets were studied. The first type of the contaminated data is the data with one outlier in the x direction. An observation in X is replaced with an outlier; there will be a point that is in the far lower right corner. The second type of contaminated data is the data with one outlier in the y direction.

One observation in Y is replaced with an outlier; there will be a point that it is in the far upper left corner. The third type of the contaminated data is the data with a point that is in the far upper and far right corner, and the outlier is in both the x and y directions. For this case, a good observation is randomly replaced with an outlier. There are many definitions of outlier. In this study, outliers are considered as the values that lay outside the 3 deviation scopes from its mean. Fig.1 shows a scatter plot of the original data and the contaminated data.

Fig.2 shows the scatter plot of the current residuals (Res1) versus lagged residuals (Res(-1)) for the original data. From the plot, it is clearly seen that there is a positive serial correlation problem in the data.

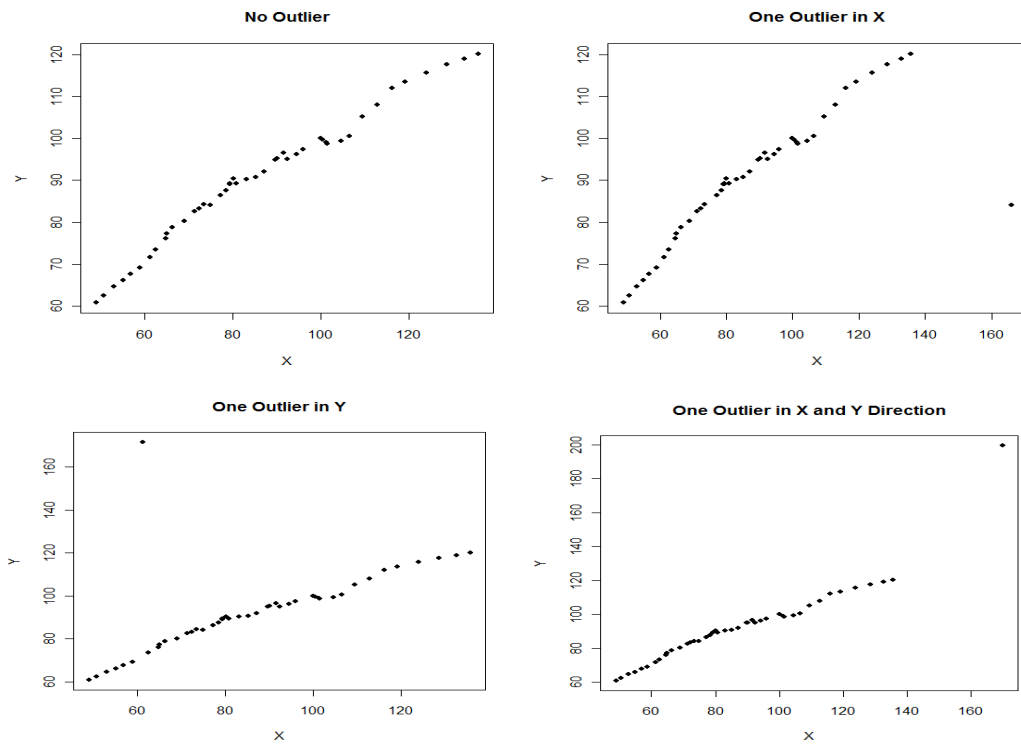


Fig.1: Scatter plot for the original and contaminated data

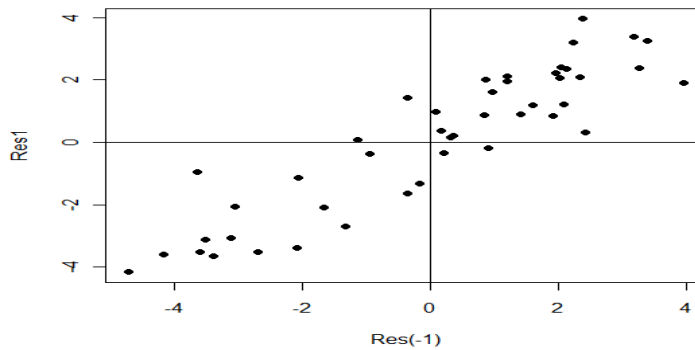


Fig.2: Current residuals (Res1) versus lagged residuals (Res(-1))

TABLE 1
Original and modified Real Compensation and Productivity Data, 1960-2005

No	X	Y	No	X	Y
1	48.9	60.8	24	83.0	90.3
2	50.6	62.5	25	85.2	90.7
3	52.9	64.6	26	87.1	92.0
4	55.0	66.1	27	89.7	94.95
5	6.8	67.7	28	90.1	95.2
6	58.8	69.1	29	91.5	96.5
7	61.2	71.7 [171.7]	30	92.4	95.0
8	62.5	73.5	31	94.4	96.2
9	64.7	76.2	32	95.9	97.4
10	65.0	77.3	33	100.0 (170)	100.0 (200)
11	66.3	78.8	34	100.4	99.7
12	69.0	80.2	35	101.3	99.0
13	71.2	82.6	36	101.5	98.7
14	73.4	84.3	37	104.5	99.4
15	72.3	83.3	38	106.5	100.5
16	74.8 {166}	84.1	39	109.5	105.2
17	77.1	86.4	40	112.8	108.0
18	78.5	87.6	41	116.1	112.0
19	79.3	89.1	42	119.1	113.5
20	79.3	89.3	43	124.0	115.7
21	79.2	89.1	44	128.7	117.7
22	80.8	89.3	45	132.7	119.0
23	80.1	90.4	46	135.7	120.2

Note: X = index of output
 Y = index of real compensation per hour
 { } = outlier in X
 [] = outlier in Y
 () = outlier in X and Y direction

The performances of the BG and MBG tests are evaluated based on the *p*-values and the results are presented in Table 2.

TABLE 2
Autocorrelation diagnostics for Real Compensation and Productivity

Test	No Outlier (<i>p</i> -value)	One Outlier in X (<i>p</i> -value)	One Outlier in Y (<i>p</i> -value)	One Outlier in X and Y Direction (<i>p</i> -value)
BG	7.667e-10	5.664e-02	5.650e-01	2.590e-01
MBG	5.703e-10	1.268e-04	1.571e-04	1.363e-04

We observe from this table that the classical BG test is able to detect autocorrelation at 0.05 significance level if there is no outlier in the data. However, it fails to detect the problem of autocorrelation when the outlier occurs in the data set. We now observe the results of the MBG test on the original and modified Indexes of Real Compensation and Productivity data. Unlike the BG test, the MBG test can successfully detect the autocorrelation in the presence of an outlier yielding a highly significant p -value.

Economic Report of the President 1982 Data

Our next example is the economic report of the president data given by Mirer (1995). These data contain 25 observations that show the relationship between personal consumption expenditures (CON) and disposable personal income (DPI). We deliberately replace a good observation with an outlier into the data set in order to get the modified data in vertical direction, horizontal direction, as well as both vertical and horizontal directions. This data set, together with the contaminated data, is presented in Table 3.

TABLE 3
Original and modified Economic Report of the President 1982 data

No	DPI(X)	CON(Y)	No	DPI(X)	CON(Y)
1	446.1	405.4	14	722.5	657.9
2	455.5	413.8	15	751.6	672.1
3	460.7	418.0	16	779.2	696.8
4	479.7	440.4	17	810.3	737.1
5	489.7	452.0	18	865.3	768.5
6	503.8(1570.0)	461.4(1461.4)	19	858.4	763.6
7	524.9	482.0	20	875.8	780.2
8	542.3	500.5	21	907.4	823.7
9	580.8{1400.0}	528.0	22	939.8	863.9
10	616.3	557.5[1557.5]	23	981.5	904.8
11	646.8	585.7	24	1011.5	930.9
12	673.5	602.7	25	1018.4	935.1
13	701.3	634.4			

Note: { } = outlier in X
[] = outlier in Y
() = outlier in X and Y directions

Fig.3 shows the scatter plot of the original and modified economic reports of the president 1982 data, while Fig.4 illustrates the scatter plot of the current residuals (Res1) versus lagged residuals (Res(-1)) for the original data. Most of the residuals are bunched, in the first and the third quadrants, suggesting a positive correlation in the data.

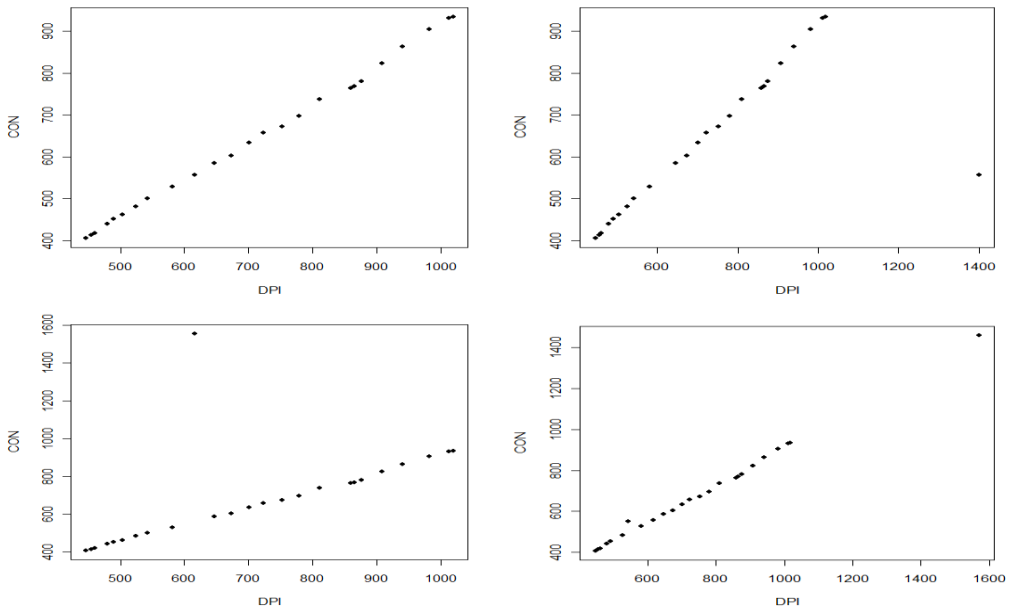


Fig.3: Scatter plots for the original and contaminated data for the Economic Report of the President data

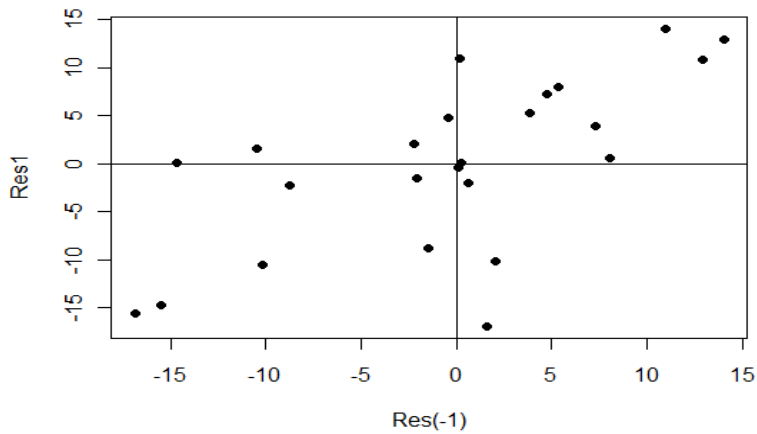


Fig.4: Current residuals (Res1) versus lagged residuals (Res(-1)) for the Economic Report of the President data

The results of the newly proposed MBG test and the classical BG test in detecting autocorrelation for the Economic Report of President data are presented in Table 4. Table 4 signifies that the classical BG test can only correctly identify the autocorrelation problem at 0.05 significance level, i.e. when the data are free from contamination although they give a false detection in the presence of outliers. The MBG test still successfully detects the presence of autocorrelation problem with and without the presence of outliers.

TABLE 4
Autocorrelation diagnostics for the real compensation and productivity data

Test	No Outlier (<i>p</i> -value)	One Outlier in X (<i>p</i> -value)	One Outlier in Y (<i>p</i> -value)	One Outlier in X and Y Direction (<i>p</i> -value)
BG	1.217e-03	6.284e-02	7.760e-01	8.752e-02
MBG	8.765e-04	1.495e-02	1.496e-02	3.711e-03

Inventories and Sales in U.S. Manufacturing, 1950 – 1991 data

For the last numerical example, we consider inventories and sales taken from Gujarati and Porter (2009). Once again, we randomly replace a good observation in the sales and inventories with the outliers and replace a coordinate paired with a contaminated pair in the sales and inventories direction. The original and contaminated data are shown in Table 5, and the scatter plot of each data set is shown in Fig.5. It can be seen by looking at the residual plot in Figure 6 that the data have positive autocorrelation problem.

TABLE 5
Original and contaminated Inventories and Sales data

No	Sales(X)	Inventories(Y)	No	Sales(X)	Inventories(Y)
1	46486	84646	22	224619	369374
2	50229	90560	23	236698	391212
3	53501	98145	24	242686	405073
4	52805	101599	25	239847	390950
5	55906	102567[802567]	26	250394	382510
6	63027	108121	27	242002	378762
7	72931	124499	28	251708	379706
8	84790	157625	29	269843	399970
9	86589	159708	30	289973	424843
10	98797	174636	31	299766	430518
11	113201	188378	32	319558	443622
12	126905	211691	33	324984	449083
13	143936	242157	34	335991	463563
14	154391	265215	35	350715	481633
15	168129{579000}	283413	36	330875	428108
16	163351(547551)	311852(1900000)	37	326227	423082
17	172547	312379	38	334616	408226
18	190682	339516	39	359081	439821
19	194538	334749	40	394615	479106
20	194657	322654	41	411663	509902
21	206326	338109			

Note: { } =outlier in X
[] =outlier in Y
() =outlier in X and Y directions

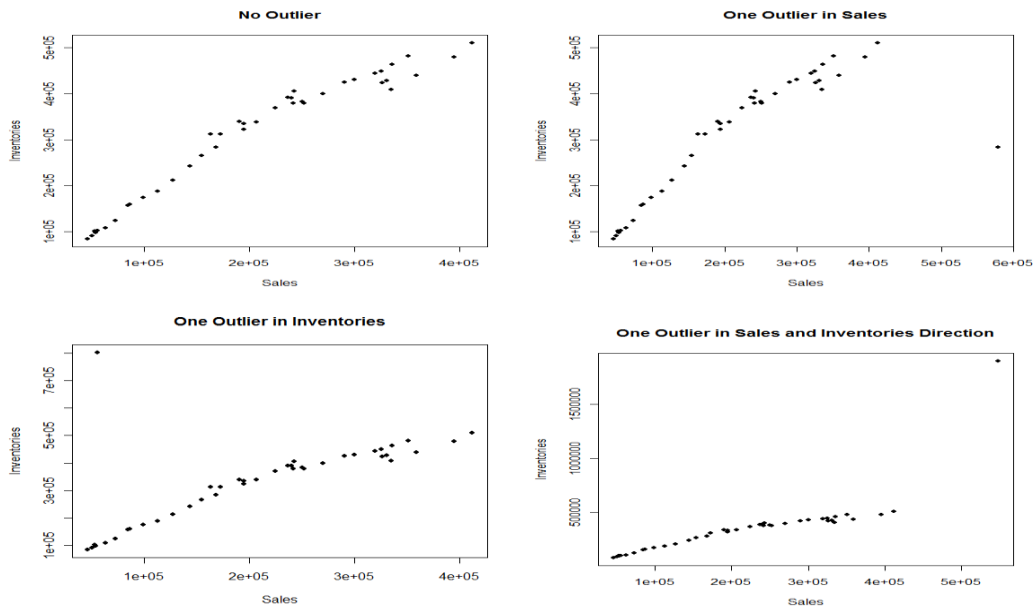


Fig.5: Scatter plot for the original and contaminated data for the Inventories and Sales data

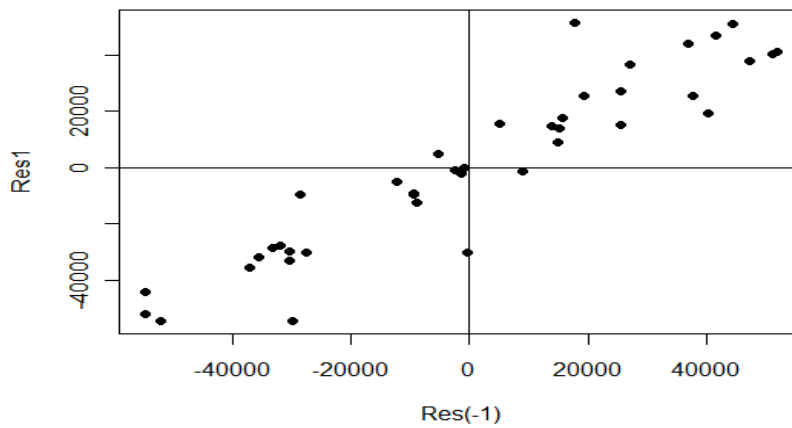


Fig.6: Current residuals (Res1) versus lagged residuals (Res(-1)) for the Inventories and Sales data

We employ the classical BG and MBG tests to the sales and inventories data. The test results are exhibited in Table 6. Similar results are obtained as in the previous examples. The power of detection of the classical BG test becomes poor when the outliers are present in the data. The MBG test is reliable in detecting the serial correlation irrespective of the presence of outliers at 0.05 significance level.

TABLE 6
Autocorrelation diagnostics for the real compensation and productivity data

Test	No Outlier (<i>p</i> -value)	One Outlier in X (<i>p</i> -value)	One Outlier in Y (<i>p</i> -value)	One Outlier in X and Y Direction (<i>p</i> -value)
BG	2.789e-09	5.043e-02	4.937e-01	5.047e-02
MBG	3.097e-09	3.996e-04	4.640e-04	2.831e-04

Simulation Study

We have seen the performance of the MBG test in the real world data. Now, we want to verify the results by checking a Monte Carlo simulation experiment. In this study, we considered three different samples sizes, $n = 20, 60$ and 100 , to represent the small, medium and large samples. For each sample, n ‘good’ data are generated according to the following relation:

$$Y = 2 + 4X + u \quad (10)$$

where, all the values of X are generated from Uniform Distribution, $U(0,10)$. The error term u_t is generated by the first-order autoregressive scheme, as follows:

$$u_t = 0.9u_{t-1} + \varepsilon_t \quad (11)$$

with an initial value of u_1 equals to 2. The white noise, ε_t is generated from the Normal distribution, with mean 0 and standard deviation 0.1. This autoregressive scheme is repeated for every 10 observations. Based on our experiences, the value of 0.9 is chosen in Eq. 11 to ensure the existence of a high autocorrelation problem.

We would like to compare the performance of the BG and MBG tests with 5% and 10% outlier in x , y and both x and y directions. For each sample size, outliers are generated by deleting the ‘good’ observations and substituting them with ‘bad’ data points. The outliers in x are represented by a uniform distributed variate x_i from Uniform Distribution $U(15,20)$, with y_i being randomly selected Y values which are less than 15. Similarly, the outliers in the y direction are represented by generating the y_i variate from a Uniform Distribution $U(50,60)$, with x_i being randomly chosen X values which are less than 4.

Finally, the data sets with the outliers in both x and y directions are created by randomly replacing good observations with x_i from $U(15,20)$ and y_i from $U(50,60)$. In this study, we set the significance level to 0.05 and in each simulation run, there are 10,000 simulations.

Table 7 exhibits the classical BG and MBG tests. The classical BG test performs very poorly in the simulation. Throughout the simulation, the classical BG tests show inconsistency in detecting autocorrelation. In fact, the BG tests fail when there are outliers in the data set for all the three sample sizes. Nonetheless, the MBG test performs superbly throughout. This test is robust when the data are contaminated with the outliers. The MBG test also has higher power of detection with the increase of sample sizes. Thus, the MBG test outperforms the classical BG test in every respect of contamination.

TABLE 7
Simulation results of autocorrelation

Sample sizes	Tests	No Outlier (<i>p</i> -value)	5% of Outliers (<i>p</i> -value)			10% of Outliers (<i>p</i> -value)		
			X	Y	Both X and Y	X	Y	Both X and Y
n = 20	BG	1.643e-02	4.702e-01	6.635e-01	4.669e-01	4.362e-01	4.975e-01	4.491e-01
	MBG	4.214e-03	4.701e-02	4.399e-02	4.729e-02	3.584e-02	3.553e-02	3.758e-02
n = 60	BG	5.906e-04	4.813e-01	5.947e-01	4.957e-01	4.740e-01	4.787e-01	4.781e-01
	MBG	5.099e-07	5.594e-05	5.181e-05	6.205e-05	6.211e-05	6.828e-05	7.495e-05
n = 100	BG	1.759e-05	4.870e-01	5.290e-01	5.032e-01	4.808e-01	4.887e-01	4.815e-01
	MBG	6.958e-11	1.294e-07	1.331e-07	1.429e-07	2.163e-07	1.797e-07	1.697e-07

CONCLUSION

In this research, the commonly used test for detecting autocorrelation has been shown to fail when outliers are present in any respect of the data. Hence, we formulate a simple but robust modification of the Breusch-Godfrey test to overcome the problem. Meanwhile, the comparison using the real data and Monte Carlo simulation experiments proved that the proposed Breusch-Godfrey test is consistent and reliable in offering substantial improvements over the classical Breusch-Godfrey test and also performs excellently in the detection of autocorrelation in the presence of outliers.

REFERENCES

- Breusch, T. S. (1978). Testing for autocorrelation in dynamic model. *Australian Economic Papers*, 17, 334-355.
- Davidson, R. and MacKinnon, J. G. (1998). Graphical methods for investigating the size and power of test statistics. *The Manchester School*, 66, 1-26.
- Durbin, J., & Watson G. S. (1951). Testing for serial correlation in least squares regression II. *Biometrika*, 38, 159-178.
- Edgeworth, F. Y. (1887). On Observations Relating to Several Quantities. *Hermathena*, 6, 279-285.
- Grassian, A., & Boer, E. S. (1980). Some methods of growth curve fitting. *Math. Scientist*, 5, 91-103.
- Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, 46, 1303-1310.
- Gujarati, D. N., & Porter, D. C. (2009). Autocorrelation: What Happens If the Error Terms Are Correlated? In M. A. Grove, M. Ito, H. Kim, P. V. Wunnava, & A. Paizis (Eds), *Basic Econometrics* (pp. 412-466). New York: McGraw-Hill.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P.J., & Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley & Son, Inc.
- Habshah, M., Norazan, M. R., & Imon, A. H. M. R. (2009). The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression. *Journal of Applied Statistics*, 36(5), 62-99.

- Hosking, J. R. M. (1980). The multivariate portmanteau statistics. *Journal of American Statistical Association*, 75, 602-608.
- Hosking, J. R. M. (1981). Lagrange-multiplier tests of multivariate time series model. *Journal of the Royal Statistical Society*, B43, 261-262.
- Midi, H. (1999). Preliminary estimators for robust non-linear regression estimation. *Journal of Applied Statistics*, 26(5), 591-600.
- Mirer, T. W. (1995). Heteroscedasticity and Autocorrelation. *Economic Statistics and Econometrics*. In P. V. Wunnava, K. Lahiri, B. Bechdolt, & G. Chowdhury-Bose (Eds). United States of America: Prentice-Hall.
- Murray, M. (2006). Autoregressive Disturbances. *Econometrics A Modern Introduction* (pp. 436-464). D. Clinton (Eds). United States of America: Pearson Addison-Wesley.
- Rana, M. S., Midi, H., & Imon, A. A. H. M. R. (2008). A robust modification of the Goldfeld-Quandt test for the detection of heteroscedasticity in the presence of outliers. *Journal of Mathematics and Statistics*, 4(4), 277-283.
- Riazoshams, H., Midi, H., & Sharipov, O. (2010). The performance of robust two-stage estimator in nonlinear regression with autocorrelated error. *Communications in Statistics. Simulation and Computation*, 39(6), 1236-1253.
- Yohai, V. (1987). High Breakdown-point and High Efficiency Estimates for Regression. *The Annals of Statistics*, 15, 642-665.
- White, G. C., & Brisbon, I. L. (1980). Estimation and comparison of parameters in stochastic growth model for barn owls. *Growth*, 44, 97-111.

