Asian Journal of Mathematics & Statistics

Methods of Estimation for Autoregressive Models with Outliers

Moawad El-Fallah Abd El-Sallam
Department of Statistics and Mathematics and Insurance, Faculty of Commerce, Zagazig University, Egypt

ABSTRACT

A common problem in autoregressive regression models is outliers, which produces undesirable effects on the least squares estimators. Many regression estimation techniques have been suggested to deal with this problem. The majority of such techniques are developed from the classical least squares. Some other robust approaches have been investigated in the regression case both on theoretical and empirical grounds. However, the behavior of these alternative methods has not received more attention in the context of time series. In this study, we present a comparative simulation study for three robust alternatives to the least squares estimator under innovation and additives outliers. The empirical study indicated that, the Reweighted Least Squares estimator (RLS) seems to be very reasonable because of having smaller "total" root normalized absolute bias and "total" root normalized variance. The RLS estimator was recommended as worthy robust alternative to the least squares estimator in the case of autoregressive models with the two kinds of outliers schemes.

Key words: Robust estimation, autoregressive models, innovation and additive outliers, least median of squares, trimmed least squares, reweighted least squares

INTRODUCTION

When estimating the parameters of an autoregressive (AR) model, a desirable quality of the underlying algorithm is to provide protection against gross errors occurring during the collection or recording of the data. Two distinct kinds of outliers-schemes are considered in regression literature; namely the innovation outliers and the additive outliers. In the Innovation Outliers (IO) scheme, the AR process is perfectly observed and gross errors are incorporated in the error distribution which is assumed to be non-normal. While, when the Additive Outliers (AO) scheme is preferable, the errors are normally distributed. However, the observations contain an additive effect not related to the AR process. In both outlier configurations, a heavy-tailed distribution, frequently assumed to have infinite variance, is involved. Although under the IO-model, the Least Squares (LS) estimator is known to be consistent (Yohai and Maronna, 1977), its Mean Squared Error (MSE) performance in the presence of outliers has been questioned. As a result, robust GM-estimators are developed by Denby and Martin (1979) and Simpson *et al.* (1992). In addition, M-estimates in the case of innovations with finite variance are proposed and studied with its asymptotic properties by Campbell (1982). Other alternative robust approaches such as the Functional Least Squares (FLS) estimator (Chambers and Heathcote, 1981), the Least Median of Squares (LMS) (Rousseeuw, 1984) and the Reweighted Least Squares (RLS) of Rousseeuw and Leroy (1987) are considered in regression literature. Also, finite-sample properties for most estimators cited above are reported in regression to the presence of heavy tails error distributions.

However, similar results for these robust estimators in the context of time series are very limited. Hence, a finite-sample comparison of the performance of some robust estimators in time series is of interest.

In this study, we present a comparative simulation study for the LMS and RLS in the case of AR-models with outliers and in order to assess their performance, diverse sampling situations are considered. In addition, results corresponding to the Trimmed Least Squares (TLS) estimator of Ruppert and Carroll (1980) and the LS-estimator are also reported.

ROBUST ESTIMATORS FOR AR-MODELS WITH OUTLIERS

Consider the autoregressive model AR (P) of order P:

$$y_t = \sum_{j=1}^{p} \alpha_j y_{t-1} + e_t \tag{1}$$

where, α $(\alpha_1, \alpha_2, ..., \alpha_p)'$ is a parameter vector and the errors e_t are independent and identically distributed random variables.

Assumed that the stationary condition:

$$\sum_{j=1}^{p} \alpha_j w^j \neq 1 \tag{2}$$

for every complex w, with $|w| \leq 1$ holds.

The three types of simulated noise used on the observed process Z_t are described as follows:

- **No. outliers:** The observed process is $Z_t = y_t$ and the errors e_t follows the standard normal N (0, 1) distribution
- **IO-scheme:** The AR-process is again perfectly observed. However, the errors now follow a heavy-tailed distribution
- **AO-scheme:** The observed process is $Z_t = y_t + u_t$ where y_t as in the no-outliners case and independent of u_t and the additive effects u_t are independent and identically distributed with:

$$u_t \sim (1-\varepsilon)F_0 + \varepsilon D \tag{3}$$

where, $0 < \varepsilon < 1$ is the proportion of contamination, F_0 denotes a distribution which is generated at zero and the additive-effect distribution, denoted the additive effect distribution denoted by D is heavy-tailed.

The aim in this study, is to obtain the estimate $\hat{\alpha}$ $(\hat{\alpha}_1, \hat{\alpha}_2, ..., \hat{\alpha}_p)'$ of the parameter vector $\alpha \in \mathbb{R}^p$ based on the data Z_t (t = 1, ..., T) and the corresponding AR-residuals are:

$$r_t = r_t(\alpha) = Z_t - \sum_{j=1}^{p} \alpha_j y_{t-1}, t = p+1, ..., T$$

The robust estimators included in the study are:

- The LMS which minimizes the criterion:

$$\text{med}\, r_t^2 \tag{4}$$

where "med" denotes the median.

- The RLS with corresponding objective function:

$$\sum_{t=p+1}^{T} w_t r_t^2 \tag{5}$$

where the final weights:

$$w_t = \begin{cases} 1 & \text{if } \left|\dfrac{r_t}{\hat{C}}\right| \leq 2.5 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

are computed from the LMS-residuals r_t and their final scale estimate is:

$$\hat{C} = \sqrt{\dfrac{\displaystyle\sum_{t=p+1}^{T} w_t^{\alpha} r_t^2}{\displaystyle\sum_{t=p+1}^{T} w_t^{\alpha} - p}} \tag{7}$$

For the definition of the initial weights w_t^0, refer to Rousseeuw and Leroy (1987).

It is of interest in regression case to note that, despite the fact that both estimators share the same (slow) rate of convergence to their asymptotic distribution, the RLS- estimator is far more efficient than the LMS in samples of size as large as T = 200 (Lucas, 1997). The robustness properties of the LMS and the RLS estimator in the context of regression can be found by Rousseeuw and Leroy (1987). The same authors provide an indication on the robustness of these estimators in the context of time series by employing the LMS and RLS methods in order to fit AR (1) and AR (2) models. Lucas (1997) showed that, the influence function of the LMS-estimator under the AO- scheme is bounded, but that the corresponding breakdown point may fall below on half and confirmed its dependence on the order of the AR-model.

- The q% TLS-estimator with corresponding objective function:

$$\sum_{t=1}^{h} r_{(t)}^2 \tag{8}$$

where, $r_{(t)}^2$ denotes the ordered squared LS-residuals and the value of h must be determined based on trimming the data values. For example, the 20% trimmed LTS estimator is defined to minimize the function $\sum r_{(t)}^2$. Chen (2002) defined h in the range:

$$\sum_{t=1}^{4T/5}[(T/2)+1] \leq h \leq \left[(3T+P+1)/4\right]$$

and recommended that the breakdown value for the TLS-estimator is T-h/T, when h = [(3T+P+1)/4]. The calculation of TLS estimator is based on a preliminary LS-fit. Then the data yielding the $100_q\%$ largest squared residuals are removed. Thus the TLS-estimator is defined to be the LS- estimator which corresponds to the reduced data set. In the simulation study, we use q = 0.20, whereas the ordinary LS-estimator corresponds to the limiting case q = 0.

Monte carlo simulation: The model and the sampling situations included in simulation study are chosen so as to ensure a feasible comparison of the estimators. Specifically, we have incorporated models of different orders and a wide range of shapes for the IO and AO simulated noise. The random number generation is implemented using the IMSL subroutines computer package.

More details are provided as follows:

- **Model and parameter choices:** The AR-model (1) with p = 2, 3 and α_j $(-1)^{j+1}$ 0.8_j, j = 1, 2, α_j $(-1)^{j+1}$ 0.6^j, j = 1, 2, 3, are, respectively utilized

- **Simulated distributions:** The specific shapes of simulated noise under the IO-scheme correspond to the following innovation-distribution:

  - The symmetric stable distribution with shape parameter, a, denoted by S (a). The values of, a, are limited to the interval [0, 2]. The smaller the value of a, the thicker the tails of the distribution. For a = 1, the Cauchy distribution results, whereas, the normal (a = 2) is the only member of the family with finite variance. In any other case, S (a) has finite moments of order only less than a
  - The λ-contaminated normal distribution-CN (λ, μ, σ)-where:

$$e_t \sim (1-\lambda) \, N \, (0, \, 1) + \lambda N \, (\mu, \, \sigma^2) \tag{9}$$

with $0 < \lambda < 1$ and $-\infty < \mu < +\infty$, $\sigma^2 > 1$. This is a standard model for thick-tailed distributions possessing a N (0, 1). The distribution is symmetric for all λ and σ when $\mu = 0$ but asymmetric otherwise.

- The Student's -t_v distribution with v degrees of freedom

The three families of distributions considered above have the advantage of containing the standard normal distribution as a limiting case (a→2, λ→0, v→). This enables to study the behavior of the estimators as we depart from the N (0,1) model and approach alternative (heavier-tailed) innovation distributions. When simulating the AO-scheme, the distribution D in Eq. 3 is taken to belong to any of the three categories mentioned above (in the case of CN (λ, μ, σ), λ = 1 is used).

Also, the AR-values y_t are adjusted to have variance equal to one. In this case, the necessary tuning constants are determined by Hamilton (1994) and correspond to the asymptotic variance of the AR-values, which for p = 3 is given by:

$$\frac{\alpha_2(K_1K_4 + K_2) + K_1K_5 - \alpha_3K_2}{(K_1\alpha_1 + K_3\alpha_3)(K_1K_4 + K_2) - (K_1\alpha_2 - K_3)(K_1K_5 - K_2\alpha_3)} \tag{10}$$

Where:

$$K_1 = \alpha_2 + \alpha_1\alpha_3$$

$$K_2 = \alpha_1 + \alpha_2\alpha_3$$

$$K_3 = 1 - \alpha_3^2$$

$$K_4 = \alpha_4 + \alpha_3$$

$$K_5 = \alpha_2 - 1$$

Also, if we set $\alpha 1 = 0$, $(\alpha_2 = \alpha_3 = 0)$ in (10), the corresponding variance for p = 0 (p = 1) results.

- **Random number generation:** The subroutine DRNSTA is used for stable random numbers whereas, when N (0, 1) variates the subroutine DRNNOR is called. In the CN (λ, μ, σ) case, the standard normal variates are transformed to σ with probability λ simulated by the subroutine DRNUN which produces random numbers in (0, 1). The same subroutine is used to simulate the proportion ε in Eq. 3. For t_ν sampling situations, we have utilized the normal/chi-squared method, with chi-squared generated by calling DRNGAM-as a special case of the gamma distribution

In order to reduce dependence of the results on initial process values, a block of 100 variates is generated according to Eq. 1 and the last p are used as starting values for calculating the observations involved in the estimation process.

RESULTS OF COMPARISONS

The results presented in this section correspond to 1000 replications. With each replication, the methods are applied to 100 standardized observations $Z_t | \text{med} | Z_t |$. For selected cases, we have generated samples (up to 2000) and our conclusions is that the discussion reported below would not be changed if a higher number of replications was used. In the left entries of Table (1-3) the "total" root normalized absolute bias is reported, whereas in the right entries, the "total" root normalized variance is reported, respectively defined by:

$$\sum_{j=1}^{p} \sqrt{100|BS_j|}$$

Where:

$$BS_j = \frac{1}{1000}\sum_{m=1}^{1000}\hat{\alpha}_{mj} - \alpha_j$$

and:

Table 1: IO-Scheme: Total root normalized absolute bias (left entry) and total root normalized variance (right entry) of AR (2)-estimators for standard normal, symmetric stable distributions with shape a, λ-contaminated normal distributions with mean $\mu = 0, 2, 5, 10$ and standard deviation σ and student's t_v distribution with v degrees of freedom

Innovation distribution		LS		LMS		RLS		TLS	
S (a)	a								
	1.0	0.92	2.24	0.83	0.57	0.65	0.38	0.90	0.53
	1.3	1.51	1.50	1.24	1.42	1.19	0.82	1.20	0.86
	1.6	1.86	1.43	1.80	2.36	1.83	1.23	1.70	1.27
	1.9	2.14	1.51	2.33	3.50	2.32	1.63	2.21	1.66
N (0,1)		1.98	1.55	2.06	3.39	2.01	1.81	1.86	1.84
CN (λ, μ, σ)	σ								
$\lambda = 0.05$	2	2.03	1.75	0.93	3.79	1.67	1.54	1.90	1.76
$\lambda = 0.10$	3	2.13	1.53	1.85	3.10	1.45	1.45	1.83	1.52
$\lambda = 0.20$	5	1.99	1.46	1.22	1.81	0.93	0.96	1.53	0.99
$\lambda = 0.30$	10	1.80	1.45	0.98	0.75	0.76	0.45	0.82	0.59
CN (λ, μ, σ)	σ								
$\lambda = 0.05$	2	1.68	1.74	1.23	3.53	1.57	1.65	1.67	1.66
$\lambda = 0.10$	3	1.57	1.51	1.62	2.86	1.45	1.40	1.55	1.41
$\lambda = 0.20$	5	1.81	1.47	1.26	1.78	1.22	0.92	1.20	0.94
$\lambda = 0.30$	10	1.57	1.46	0.98	0.75	0.78	0.44	0.69	0.59
CN (λ, μ, σ)	σ								
$\lambda = 0.05$	2	2.04	1.52	0.47	2.81	0.87	1.32	1.50	1.36
$\lambda = 0.10$	3	2.94	1.51	1.12	2.10	1.09	1.07	2.02	1.13
$\lambda = 0.20$	5	3.67	1.48	1.23	1.37	0.77	0.72	2.06	0.73
$\lambda = 0.30$	10	2.97	1.50	0.96	0.68	0.76	0.40	1.39	0.61
CN (λ, μ, σ)	σ								
$\lambda = 0.05$	2	2.46	1.46	0.77	2.00	0.75	0.91	1.69	1.00
$\lambda = 0.10$	3	3.50	1.51	0.63	1.32	0.67	0.63	1.99	0.77
$\lambda = 0.20$	5	4.76	1.49	0.87	0.84	0.76	0.45	2.22	0.60
$\lambda = 0.30$	10	4.70	1.52	0.84	0.54	0.57	0.32	2.45	0.68
t_v	V								
	2	1.86	1.39	0.01	1.91	1.32	1.10	1.55	1.02
	3	1.63	1.49	1.29	2.46	1.23	1.36	1.33	1.33
	6	1.98	1.66	2.27	3.25	1.88	1.53	1.92	1.60
	12	2.17	1.68	2.15	3.48	2.00	1.55	2.08	1.69

$$\sum_{j=1}^{p} \sqrt{100 \left| VR_j \right|}$$

Where:

$$VR_j = \frac{1}{1000} \sum_{m=1}^{1000} \hat{\alpha}_{mj}^2 - \bar{\alpha}^2$$

with $\hat{\alpha}_{mj}$ denoting the estimate of $\hat{\alpha}_j$ (j = 1, 2,... , p) in the mth replication, m = 1, 2,. .., 1000. The "total" root normalized mean-squared error is defined by:

$$TRMSE = \sum_{j=1}^{p} \sqrt{100 MSE_j}$$

Table 2: AO-Scheme: Total root normalized absolute bias (left entry) and total root normalized variance (right entry) of AR (3)-estimators with proportion of contamination $\varepsilon = 0.01$ symmetric stable distributions with shape a, normal distributions with mean $\mu = 0$, 2, 5, 10 and standard deviation σ and student's t_v distribution with v degrees of freedom

Additive effect distribution		LS		LMS		RLS		TLS	
S (a)	a								
	1.0	6.90	4.03	3.56	7.54	4.36	3.62	6.56	4.20
	1.3	5.77	3.59	3.88	7.57	4.06	3.55	5.40	3.82
	1.6	5.02	3.54	3.51	7.57	3.86	3.32	4.77	3.67
	1.9	4.50	3.10	3.73	7.53	3.81	3.49	4.77	3.67
N (0, σ^2)	σ								
	2	5.64	3.54	2.85	7.55	4.47	3.16	5.02	3.53
	3	7.48	3.66	4.19	7.52	5.57	3.46	6.58	3.62
	5	10.00	4.00	4.79	7.50	6.52	3.84	8.88	3.92
	10	13.20	4.49	5.04	7.62	7.37	4.15	11.90	4.36
N (2, σ^2)	σ								
	2	7.29	3.36	4.07	7.55	5.60	3.62	6.44	3.57
	3	8.48	3.60	4.22	7.62	6.03	3.71	7.46	3.68
	5	10.50	4.02	4.89	7.61	6.82	3.83	9.28	3.94
	10	13.40	4.36	5.32	7.65	7.47	4.17	12.00	4.28
N (5, σ^2)	σ								
	2	11.5	3.55	5.62	7.60	7.48	3.91	10.0	3.66
	3	11.6	3.88	5.50	7.55	7.15	3.95	10.1	3.88
	5	12.1	4.21	5.29	7.53	7.13	4.03	10.7	4.10
	10	13.8	4.33	5.21	7.60	7.37	4.23	12.3	4.26
N (10, σ^2)	σ								
	2	15.6	2.72	5.51	7.51	8.13	4.43	43.9	3.24
	3	15.4	3.09	5.48	7.46	8.04	4.38	13.7	3.44
	5	14.9	3.84	5.17	7.15	7.71	4.32	13.3	3.88
	10	14.8	4.34	5.38	7.50	7.42	4.31	13.4	4.26
t_v	V								
	2	5.23	3.55	3.88	7.50	3.88	3.42	5.03	3.72
	3	4.57	3.54	3.82	7.66	3.80	3.21	4.24	3.62
	6	3.73	3.49	2.46	7.61	3.21	3.12	3.52	3.49
	12	3.43	3.44	2.49	7.61	2.99	3.03	3.37	3.55

Where:

$$MSE_j = BS_j^2 + VR_j, \; j = 1, ..., p$$

The following results based on the simulation results.

IO-Scheme (Table 1): The RLS estimator has (overall) the smallest bias, whereas, the LS is the least reliable method under this criterion, particularly in the presence of considerable skewness (when a = 1.9 or near to 2). In such skewed situations, the LMS-estimator should be preferred but under symmetric (or nearly symmetric) errors, the LMS and TLS-estimators are comparable with respect to bias. Also, within the family of symmetric-stable error-distributions, the bias of the estimators decreases, as the tail-length of the underlying distribution increases. This phenomenon is in agreement with theoretical results-pertaining to the LS and M-estimators- which indicate

Table 3: AO-Scheme: Total root normalized absolute bias (left entry) and total root normalized variance (right entry) of AR (3)-estimators with proportion of contamination $\varepsilon = 0.05$ for symmetric stable distributions with shape a, normal distributions with mean $\mu = 0, 2, 5, 10$ and standard deviation σ and student's t_v distribution with v degrees of freedom

Additive effect distribution		LS		LMS		RLS		TLS	
S (a)	a								
	1.0	12.10	4.72	6.52	7.81	8.21	4.09	11.50	4.74
	1.3	10.40	4.85	6.96	7.57	7.88	3.30	9.88	4.37
	1.6	9.09	3.82	6.71	7.57	7.53	3.78	8.63	3.98
	1.9	8.01	3.35	6.62	7.49	7.14	3.61	7.66	3.68
N $(0, \sigma^2)$	σ								
	2	9.94	3.73	7.58	7.44	8.69	3.32	9.29	3.61
	3	12.50	3.43	9.01	7.34	10.3	3.94	11.70	3.64
	5	15.20	4.20	9.81	7.37	11.7	3.22	14.40	3.48
	10	17.30	4.61	10.7	7.15	12.5	2.68	16.90	2.63
N $(2, \sigma^2)$	σ								
	2	12.10	3.37	9.28	7.30	10.3	3.82	11.4	3.55
	3	13.50	3.39	9.88	7.33	11.0	3.94	12.8	3.59
	5	15.50	4.18	10.4	7.14	12.0	3.14	14.8	3.43
	10	17.30	4.62	11.0	7.15	12.8	2.32	16.9	2.60
N $(5, \sigma^2)$	σ								
	2	15.9	2.65	11.4	7.12	13.0	3.99	15.4	2.81
	3	16.0	2.75	10.8	7.18	12.6	4.15	15.5	2.96
	5	16.5	2.75	10.6	7.25	12.5	4.39	16.0	2.99
	10	17.5	2.12	11.0	7.27	12.8	4.64	17.1	2.47
N $(10, \sigma^2)$	σ								
	2	17.6	2.65	12.1	6.92	13.5	4.62	17.5	1.81
	3	17.7	1.82	11.8	7.01	13.4	4.65	17.5	1.83
	5	17.7	1.85	11.5	7.11	13.2	4.67	17.5	2.00
	10	17.8	1.76	10.9	7.22	12.7	4.73	17.6	2.00
t_v	V								
	2	9.20	3.89	6.47	7.54	7.31	3.74	8.76	4.01
	3	7.96	3.65	6.19	7.51	6.78	3.50	7.49	3.82
	6	6.74	3.65	3.99	7.56	5.86	3.23	6.47	3.57
	12	6.23	3.47	4.68	7.55	5.79	3.12	6.02	3.56

consistency rates related to Campbell (1982). Turning to variability we observe that the RLS, also, is the most preferable method, whereas, the LS and the LMS are less efficient. However, the last estimator should be used instead of the LS when asymmetric heavy-tailed innovations are expected. With the exception of the LS-method, the estimated variance is a decreasing function of the tail length of the innovation distribution. Thus, previous results (Denby and Martin, 1979) indicating that heavy-tailed innovation distributions improve the accuracy of certain estimators can be possibly extended to the robust methods considered here.

To concluded, the presence of considerable skewness increases the bias of the LS-estimator but leaves the variance practically unaffected. On the contrary the variance (and in many cases the bias) of the LMS and RLS-estimators is reduced with increasing asymmetry. Furthermore, the LMS estimator is the worst among the estimators considered in terms of estimated TRMSE. However, when the underlying error-distributions is extreme-tailed, the last estimator is on the whole more

reliable than the LS-estimator which, naturally, is very efficient (or near) at the standard normal model. Finally, the TLS and the RLS-estimators are both fairly accurate estimators but the latter estimator should be preferred since it is more bias-robust in the presence of skewness.

AO-Scheme (Table 2-3): In this case, the effect of additive outliers on the estimated bias increases with properties of contamination ε and it seems impossible to remove. This drastic effect of additive outliers is also obtained by Sejling *et al.* (1994) in their simulation study of recursive robust estimators. Interesting enough, comparison of biases favors the LMS and the RLS robust estimators over all additive-effect distributions. For the former estimator, this behavior is obviously connected with the specific type of outliers as it does not appear neither in the regression context nor in the IO- case. Among the other two less reliable estimators in terms of bias, the TLS has some advantage under symmetric (or nearly symmetric) additive effects for $\varepsilon = 0.01$. This advantage, however practically, disappears with increasing skewness and proportion of contamination. On the other hand, the bias of the LMS and RLS estimators is significantly affected by the order of the AR-model. Hence, we my conjecture that, the breakdown point of these methods is reduced compared to the regression setup and depends on the order p (Denby and Martin, 1979).

Based on efficiency, it is evident however that, the LMS is the least efficient method. The other three methods appear to be (overall) comparable, since no estimator clearly dominates the others, with the exception of the RLS which is less efficient with increasing skewness for $\varepsilon \geq 0.05$. Although, the effect of increasing skewness on the bias of the estimators is of considerable size for all methods (particularly for small σ), the LMS and RLS-estimators are comparatively more robust in this respect. The same effect on the efficiency of all four estimators is not equally significant. It appears however that, the presence of skewness considerably enhance the efficiency of the LS, the TLS-estimator for $\varepsilon \geq 0.05$ but at the same time, as it has already been mentioned, exerts a somewhat deteriorating effect on the variance of the RLS-estimator.

CONCLUSION

Robust autoregressive-model estimators occupy a particularly important area in time - series methodology. In this paper, we present a simulation study of three robust alternatives to the LS estimator under innovation and additive outliers. The simulation results indicate that, under the IO-scheme, the reweighted least squares and the trimmed least squares estimators perform well and the general conclusion is that in most cases, the estimators are not seriously affected by this type of outliers. Also, the results suggest that:

- The consistency rates of the LMS and RLS estimators under stable innovations are a decreasing function of the shape parameter a of the distribution (possibly of order related to a^{-1})
- The presence of heavy tails improves the asymptotic efficiency of all robust estimators considered, regardless of whether or not the variance of the innovations exists

Turning to the results under the AO-scheme, we have observed that, the effect of additive outliers on all four estimators considered is significant and with the exception of the least median of squares, of increasing importance as the proportion and the shape of contamination changes. The least median of squares estimator is competitive in the presence of considerable heavy-tail contamination, especially, when the additive-effect distribution is asymmetric. Overall, however, the reweighted least squares appears to be the most efficient method. In addition, the least squares

estimators under both outliers schemes and although it is quite efficient near the standard normal model, exhibits less satisfactory performance on the whole. Finally, the findings of this study indicate that, in the case of auto regression with outliers, the RLS method should be considered as worthy robust alternative to the least squares method.

REFERENCES

Campbell, K., 1982. Recursive computation of M-estimates for the parameters of a finite autoregressive process. Ann. Statist., 10: 442-453.

Chambers, R.L. and C.R. Heathcote, 1981. On the estimation of slope and the identification of outliers in linear regression. Biometrika, 68: 21-23.

Chen, C., 2002. Robust regression and outlier detection with the ROBUSTREG procedure. SAS Institute: Cary, NC. SUGI Paper, 265-270.

Denby, L. and R.D. Martin, 1979. Robust estimation of the first-order autoregressive parameter. J. Am. Statistical Assoc., 74: 140-146.

Hamilton, J.D., 1994. Time Series Analysis. 1st Edn., Princeton University Press, New Jersey, ISBN: 0-691-04289-6.

Lucas, A., 1997. Asymptotic robustness of least median of squares for autoregressions with additive outliers. Comm. Statist. Theory Methods, 26: 2363-2380.

Rousseeuw, P.J. and A.M. Leroy, 1987. Robust Regression and Outliers Detection. Wiley, New York.

Rousseeuw, P.J., 1984. Least median of squares regression. J. Am. Statist. Assoc., 79: 871-880.

Ruppert, D. and R.J. Carroll, 1980. Trimmed least squares in the linear model. J. Amer. Statist. Assoc., 75: 828-838.

Sejling, K., H. Madsen, J. Holst, U. Holst and J.E. Englund, 1994. Methods for recursive robust estimation of AR parameters. Comput. Statist. Data Anal., 17: 509-536.

Simpson, D.G., D. Ruppert and R.J. Carroll, 1992. On one-step GM estimates and stability of inferences in linear regression. J. Amer. Statist. Assoc., 87: 439-450.

Yohai, V.J. and R.A. Maronna, 1977. Asymptotic behavior of least squares estimates for autoregressive process with infinite variance. Ann. Statist., 5: 554-560.