Robust Estimators of Ar-Models: A Comparison

G.S. Donatos* and S.G. Meintanis**

Abstract

Many regression-estimation techniques have been extended to cover the case of dependent observations. The majority of such techniques are developed from the classical least squares, M and GM approaches and their properties have been investigated both on theoretical and empirical grounds. However, the behavior of some alternative methods- with satisfactory performance in the regression case- has not received equal attention in the context of time series. A simulation study of four robust estimators for autoregressive models containing innovation or additive outliers is presented. The robustness and efficiency properties of the methods are exhibited, some finite-sample results are discussed in combination with theoretical properties and the relative merits of the estimators are viewed in connection with the outlier-generating scheme.

Key Words: Robust estimation; Simulation; Innovation and additive outliers; Least median of squares; Functional least squares

1. 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. These errors are the most frequent reasons for outliers, namely data which are far away from the pattern set by the majority of the data. Traditionally two distinct kinds of outlier-schemes are considered. In the innovation-outlier (IO) scheme, the AR process is perfectly observed and gross errors are incorporated in the error distribution which is assumed to be non-normal. When the additive-outlier (AO) scheme is preferable, the error is normally distributed. However, the observations contain an additive effect not related to the AR process. In both outlier configurations, a heavy-tailed distribution is involved frequently assumed to have infinite variance. Although under the IO-model the least squares (LS) estimator is known to be consistent (Kanter and Steiger (1974), Hannan and Kanter (1977), 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 Bustos (1982). M-estimates in the case of innovations with finite variance are proposed by Beran (1976). Under less restrictive assumptions, the asymptotic properties of cer-

*Professor Dept. of Economics University of Athens

**Lecturer Dept. of Engineering Science University of Patras

tain M-estimators are studied by Campbell (1982) and Davis et al. (1992). Alternative robust approaches include the least absolute deviation method (Gross and Steiger (1979)), the S-estimates for AR-models (Martin and Yohai (1991)), the minimum distance estimators of Dhar (1993) and the recursive estimators considered by Cipra et al. (1993) and Sejling et al. (1994). For robust estimation methods in more general time-series models the reader is referred to Martin and Yohai (1985), Bustos and Yohai (1986), Chang et al. (1988), Li and McLeod (1988), Allende and Heiler (1992), McDougall (1994), Mikosch et al. (1995) and Salau (1995).

In most of the work cited above, finite-sample results are reported. However, similar results for the least median of squares (LMS) and the reweighted least squares (RLS) of Rousseeuw and Leroy (1987) and the functional least squares (FLS) estimator of Heathcote and Welsh (1983) are very limited. These estimators were initially introduced in the context of regression by Rousseeuw (1984) and Chambers and Heathcote (1981), respectively and were shown to have attractive theoretical properties (see also Maronna et al. (1993) for the LMS and the RLS and Csorgo (1983) for the FLS-estimator). In this paper we present a comparative simulation study for the LMS, the RLS and the FLS AR-model estimators in order to assess their performance in the presence of outliers under diverse sampling situations. Results corresponding to the trimmed least squares (TLS)-estimator of Ruppert and Carroll (1980) and the LS-estimator are also reported.

In the next section the AR-model, the outlier-schemes and the estimation methods are described. Numerical details are provided in Section 3 whereas, in the following section the Monte-Carlo experiment is specified. The simulation results are presented in Section 5 along with a comparative discussion and we conclude by summarizing our findings in the last section.

2. Ar-Models with Outliers and Robust Estimators

Consider the autoregressive-model AR(p) of order p

$$\boldsymbol{y}_t = \sum_{j=1}^{p} \gamma_j \boldsymbol{y}_{t-j} + \boldsymbol{e}_t \tag{1}$$

where $\gamma = (\gamma_1, \ldots, \gamma_p)'$ is a parameter-vector and the errors $\mathbf{e}_t$ are independent and identically distributed random variables. Assume that the stationarity condition

$$\sum_{j=1}^{p} \gamma_j \omega^j \neq 1$$

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

The three types of simulated noise acting on the observed process $\boldsymbol{z}_t$ are described as follows:

i) No Outliers. The observed process is $z_t = y_t$ and the errors e_t follow the standard normal-N(0,1)- distribution.

ii) IO-scheme. The AR-process is again perfectly observed. However, the errors now follow a
heavy-tailed distribution.

iii) AO-scheme. The observed process is $z_t = y_t + v_t$, where y_t as in the no-outliers case and independent of v_t and
the additive effects v_t are independent and identically distributed with

$$v_t \sim (1-\varepsilon)\delta_0 + \varepsilon D, \tag{2}$$

where $0 < \varepsilon < 1$ is the proportion of contamination, δ_0 denotes a distribution which is degenerate at zero and, the additive-effect distribution denoted by D is heavy-tailed.

Our aim is to estimate the parameter-vector $\gamma \in \Re^p$ based on the data z_t (t=1,2,...,T) and the corresponding AR-residuals

$$r_t = r_t(\gamma) = z_t - \sum_{j=1}^{p} \gamma_j z_{t-j}, \quad t = p+1, \ldots, T.$$

The estimators included in the study are:

i) The LMS which minimizes the criterion

$$\underset{t}{med}\; r_t^2 \,,$$

where med denotes the median.

ii) The RLS with corresponding objective function

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

where r_t are the LMS-residuals. Their scale estimate c is used to determine the weights

$$w_t = 1, \;\; if \left| \frac{r_t}{\hat{c}} \right| \le 2.5$$

$$= 0, \text{ otherwise.}$$

In Rousseeuw and Leroy (1987), the LMS and the RLS-method are employed in order to fit AR(1) and AR(2) models to the data RESX. For comparison with GM-estimates applied on this data set see Lee and Hui (1993).

iii) The FLS-estimator which is found by minimizing the function

$$L_T(\gamma;s) = -\frac{1}{s^2} log\left[U_T^2(\gamma;s) + V_T^2(\gamma;s)\right],\ s \in \Im, \qquad (3)$$

where $\Im$, is a compact set of the real line and

$$U_T(\gamma;s) = \frac{1}{T-p}\sum_{t=p+1}^{T} cos(s r_t(\gamma))$$

and

$$V_T(\gamma;s) = \frac{1}{T-p}\sum_{t=p+1}^{T} sin(s r_t(\gamma))$$

are (under the IO-scheme) the sample estimates for the real and the imaginary part of the error characteristic function $\varphi(s) = u(s) + iv(s),\ i = \sqrt{-1}$, respectively.

The FLS-method produces a family of estimators for $s \in \Im$, with the limiting case $s = 0$ corresponding to the LS-estimator. In the regression case, Csorgo (1983) provided the theoretical background for an efficient version of the method in which s is selected by minimizing the asymptotic variance function of the FLS-estimator

$$\sigma^2(s) = \frac{u^2(s)(1-u(2s)) - 2u(s)v(s)v(2s) + v^2(s)(1+u(2s))}{2s^2(u^2(s)+v^2(s))^2},$$

over $\Im$. Dhar (1990) proves the weak convergence and the asymptotic unbiasedness of the FLS-estimator in the AR(1)-model with additive outliers. However, he imposes restrictive conditions pertaining to the existence of moments for the additive-effect distribution and to the asymptotic behavior of a non-constant proportion of contamination. Some encouraging simulation results on the FLS, when applied to AR(1) models, can be found in Heathcote and Welsh (1983) and Dhar (1993).

iv) The TLS-estimator with corresponding objective function

$$\sum_{t=p+1}^{T} w_t r_{(t)}^2$$

with

$$w_t = 1, \text{ if } t \leq [T - q \times (T - p)]$$
$$= 0, \text{ otherwise}$$

where $q \in (0,1)$, and $[\cdot]$, $r^2_{(\cdot)}$, denote the integer-part function and the ordered squared LS-residuals, respectively. Naturally the LS-estimator corresponds to the limiting case $q=0$.

3. Specification of Algorithms

Calculations are performed in double precision arithmetic on a HP Vectra VL2, 4/100 with a

Professional Fortran compiler. The extensive version of the routine PROGRESS (Rousseeuw and Leroy (1987)) is employed for the calculation of the LMS and the RLS-estimator. This routine- kindly provided by Professor P.J. Rousseeuw- is modified in the data-generation part where the AR-process and the outliers are simulated.

In order to obtain the FLS-estimator, we use equation (3) and the resulting estimating equations

$$l_T(\gamma, s) = 0, \tag{4}$$

where l_T is a vector with jth-element (j=1,...,p)

$$l_{Tj}(\gamma; s) = \frac{\partial}{\partial \gamma_j} L_T(\gamma; s)$$

$$= \frac{1}{(T-p)} \frac{1}{s} \sum_{t=p+1}^{T} z_{t-j} (U_T(\gamma; s) sin\{ sr_t(\gamma)\} - V_T(\gamma; s) cos\{ sr_t(\gamma)\}).$$

Equation (4) produces the iterative scheme

$$\hat{\gamma}^{(m+1)} = \hat{\gamma}^{(m)} + \left[U_T^2(\hat{\gamma}^{(m)}; s) + V_T^2(\hat{\gamma}^{(m)}; s) \right]^{-1} (ZZ)^{-1} l_T(\hat{\gamma}^{(m)}; s) \tag{5}$$

where ZZ is a $p \times p$ matrix with (i,j) element

$$\frac{1}{T-p} \sum_{t=p+1}^{T} z_{t-i} z_{t-j} - \frac{1}{T-p} \sum_{t=p+1}^{T} z_{t-i} \; \frac{1}{T-p} \sum_{t=p+1}^{T} z_{t-j}$$

and $\hat{\gamma}^{(0)}$ is the 20% TLS-estimator $(q=0.20)$.

The specific value of s used in (5) is determined by employing the IMSL(1987) routine DUVMGS to minimize the current sample estimate of the variance function $\sigma^2(s)$ over $\Im = [10^{-3}, 1.0]$. Iteration (5) stops when the overall absolute change in the values of the estimates falls below 10^{-9}. A few convergence problems- due to the periodic nature of the trigonometric functions involved- are resolved by reducing the upper boundary of in problematic cases.

4. Description of the Monte Carlo Study

The model and the sampling situations included are chosen so as to ensure a credible comparison of the estimators. Specifically, we have incorporated models of different orders and different parameter-values and a wide range of shapes for the IO and AO simulated noise. The random number generation is implemented by calling the appropriate routines of the IMSL (1987) computer package. More details are provided below.

i) Model and parameter choices. The AR-model (1) with p=2,3 and

$$\gamma_j = (-1)^{j+1} 0.80^j,\ j = 1,2,\quad \gamma_j = (-1)^{j+1} 0.60^j,\ j = 1,2,3,$$

respectively, is utilized. The scheme $\gamma_j = (-1)^{j+1} a^j$, $a < 1$ (due to H. Wold) is adopted from Anderson (1971).

ii) Simulated distributions. The specific shapes of simulated noise under the IO-scheme correspond to the following innovation-distributions:

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

b) The contaminated normal distribution- CN(λ,σ)- where

$$e_t \sim ((1-\lambda)N(0,1) + \lambda N(0,\sigma^2)),$$

with $0 < \lambda < 1$ and $\sigma^2 > 1$.

This is a standard model for thick-tailed distributions possessing a N(0,1) core.

c) The Student's-t_ν distribution with ν degrees of freedom.

The three families of distributions considered above have the advantage of containing the standard normal distribution as a limiting case ($\alpha \to 2$, $\lambda \to 0$, $v \to \infty$).

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 (2) is taken to belong to any of the three categories mentioned above (in the case of CN(λ,σ), λ=1 is used). Also, the AR-values y_t are adjusted to have variance equal to one and the determination of the necessary tuning constants is based on equations [3.4.4] and [3.4.36] in Hamilton (1994).

iii) Random number generation. The routine DRNSTA is used for stable random numbers whereas, when N(0,1) variates are required the routine DRNNOR is called. In the CN(λ,σ) case, the standard normal variates are multiplied by σ with probability λ simulated by the routine DRNUN which produces random numbers in (0,1). The same routine is used to simulate the proportion ε in (2). For t_v 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 equation (1) and the last *p* are used as starting values for calculating the observations involved in the estimation process.

5. Simulation Results

The results presented in this section correspond to 1000 replications. With each replication, the methods are applied to 100 standardized observations $z_t / med|z_t|$. For selected cases, we have generated more samples (up to 3000) and our conclusion is that the discussion reported below would not be altered if a higher number of replications was used. In the left entries of Table 1, the sums of the root normalized (times 100) absolute biases over all estimated coefficients are reported for the IO-case. The right entries show the corresponding figures for the variances. The same statistics under the AO-scheme are reported in Tables 2,3,4 for ε= 0.01,0.05,0.10, respectively. Based on the simulation results the following observations can be made:

i) IO-scheme

The RLS-estimator has (overall) the smallest bias followed by the LMS whereas, the TLS, the FLS and the LS are less reliable methods under this criterion. 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 consistency rates related to $(T\text{-}p)^{1/\alpha}$ (see, for example, Davis et al. (1992)). Turning to variability we observe that the FLS followed by the TLS and the RLS are the most preferable methods whereas, the LS and (by far) the LMS are less efficient. 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 (see, for example, 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. For the LS-method, the exhibited flatness in cases of errors with finite variance is consistent with the fact that the asymptotic variance of the estimator is independent of the innovation-distribution. In order to assist final comparison we have graphed the sum of the root normalized MSE's (TRMSE) of the parameter-estimators in Figures 1-4. Apparently the FLS is the best and the LMS is the worst among the estimators considered in terms of estimated TRMSE. However, when the underlying error-distribution is extreme-tailed, the last estimator is on the whole more reliable than the LS-estimator which, naturally, is very efficient at (or near) the standard normal model. Finally, the TLS and the RLS-estimator exhibit a considerable agreement in their values of estimated TRMSE in the majority of cases.

Table 1. IO-scheme: Total root normalized bias (left entry) and total root normalized variance (right entry) of estimatorsAND FIGURES 1-4 APPROXIMATELY HERE

Innovation Distribution		*LS*		*LMS*		*RLS*		*FLS*		*TLS*	
						p=2					
S(α)	α										
	1.0	0.92	2.24	0.83	0.57	0.65	0.38	0.51	0.43	0.90	0.53
	1.3	1.51	1.50	1.24	1.42	1.19	0.86	0.92	0.79	1.20	0.82
	1.6	1.83	1.43	1.80	2.36	1.86	1.27	1.69	1.13	1.70	1.23
	1.9	2.14	1.51	2.33	3.50	2.32	1.63	2.22	1.44	2.21	1.66
N(0,1)		1.98	1.55	2.06	3.39	2.01	1.81	2.02	1.57	1.86	1.84
CN(λ,σ)	σ										
λ=0.05	2	2.03	1.54	0.93	3.79	1.67	1.76	1.89	1.52	1.90	1.75
λ=0.10	3	2.13	1.53	1.85	3.10	1.45	1.45	1.58	1.35	1.83	1.52
λ=0.20	5	1.99	1.46	1.22	1.81	0.93	0.96	1.09	0.94	1.53	0.99
λ=0.30	10	1.80	1.45	0.98	0.75	0.76	0.45	0.83	0.52	0.82	0.59
t_V	V										
	2	1.86	1.39	1.01	1.91	1.32	1.10	1.03	1.09	1.55	1.02
	3	1.63	1.49	1.29	2.46	1.23	1.36	1.49	1.28	1.33	1.33
	6	1.98	1.53	2.27	3.25	1.88	1.66	1.84	1.48	1.92	1.60
	12	2.17	1.55	2.15	3.48	2.00	1.68	2.18	1.54	2.08	1.69
						p=3					
S(α)	α										
	1.0	2.03	7.70	1.11	1.64	0.96	1.08	1.59	1.18	1.59	1.19
	1.3	2.30	3.07	1.41	3.25	1.36	1.93	1.87	1.75	2.19	1.71
	1.6	2.62	2.80	3.16	5.15	2.54	2.69	2.90	2.33	2.72	2.38
	1.9	2.98	2.93	4.07	7.14	3.26	3.30	3.74	2.84	3.23	3.23
N(0,1)		2.22	3.02	1.76	7.60	2.18	3.45	3.63	3.04	2.24	3.52
CN(λ,σ)	σ										
λ=0.05	2	2.08	3.01	3.43	7.16	2.13	3.39	2.89	3.01	2.41	3.39
λ=0.10	3	2.11	2.93	0.97	6.10	1.43	2.92	2.63	2.62	1.92	2.89
λ=0.20	5	1.98	2.83	0.68	3.65	1.34	1.95	1.79	1.93	1.65	1.92
λ=0.30	10	1.87	2.89	1.07	1.49	0.44	0.87	0.49	1.06	1.15	1.22
t_V	V										
	2	2.91	2.76	1.56	4.02	1.30	2.33	1.67	2.20	2.30	2.08
	3	2.57	2.91	2.42	5.08	2.71	2.81	3.02	2.56	2.45	2.58
	6	2.07	3.02	1.62	6.30	1.13	3.22	2.93	2.92	1.67	3.11
	12	2.54	2.96	3.01	6.79	2.50	3.28	3.36	2.97	2.86	3.23

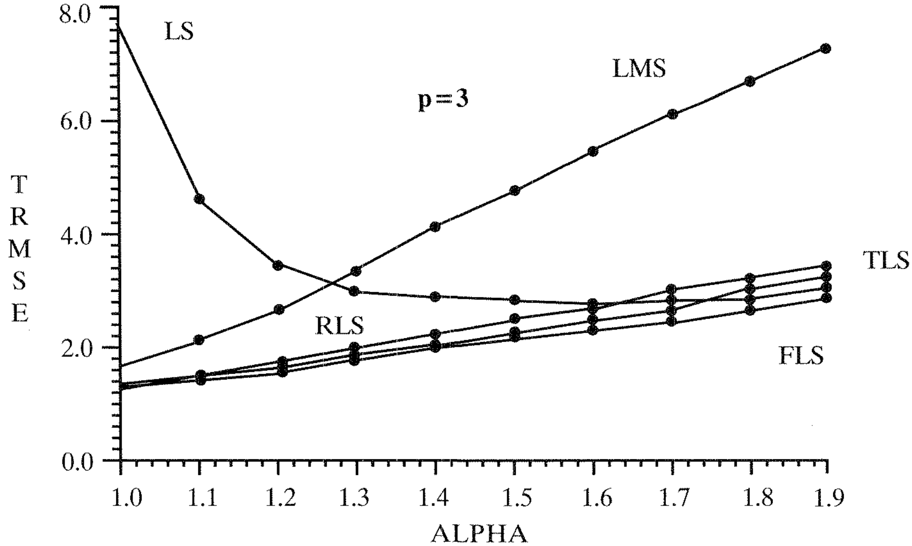

Figure 1. IO-scheme: Efficiency under symmetric-samle errors

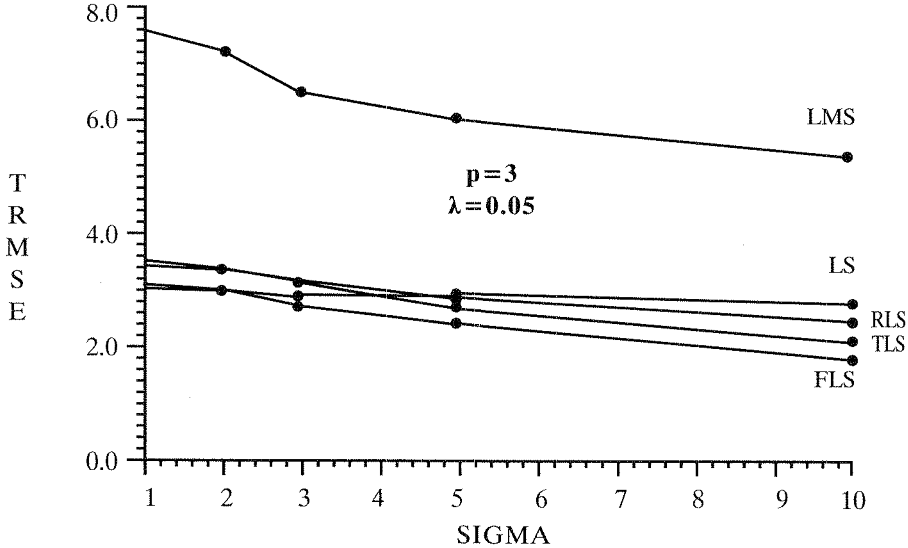

Figure 2. IO-scheme: Efficiency under contaminated-normal errors

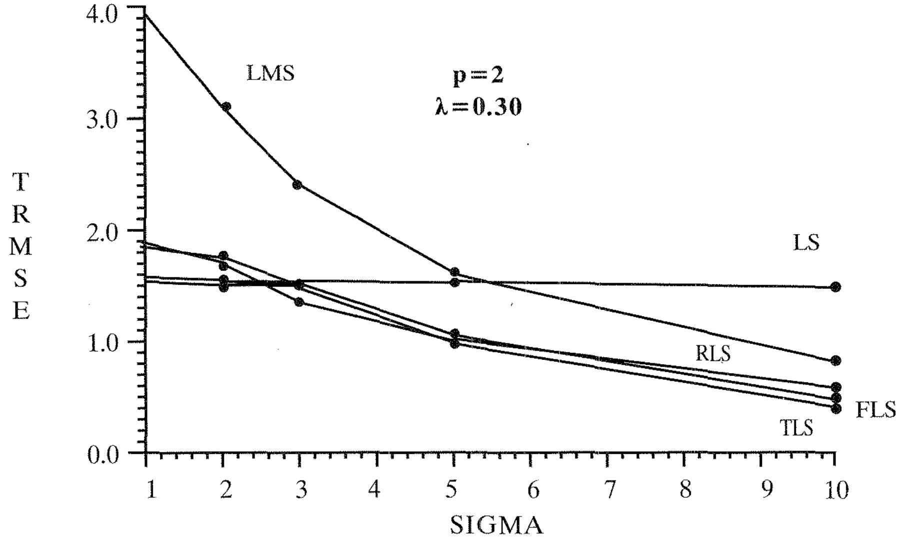

Figure 3. IO-scheme: Efficiency under contaminated-normal errors

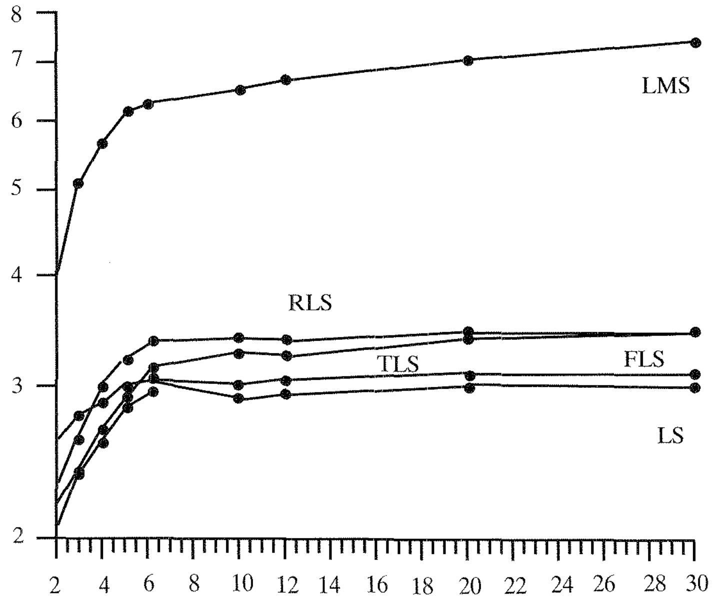

Figure 4. IO-scheme: Efficiency under Student's-t errors

ii) AO-scheme

In this case, the effect of additive outliers on the estimated bias increases with the proportion of contamination and seems impossible to remove it. This drastic effect of additive outliers is also observed by Sejling et al. (1994) in their simulation study of recursive robust estimators. Comparison of biases favors the LMS-estimator uniformly in and over all additive-effect distributions. The RLS follows whereas, the LS and, to a lesser extent, the TLS and the FLS are less reliable methods in terms of bias. For the last three estimators the estimated bias increases with the tail-length of the additive-effect distribution. This resembles the monotonicity-with respect to the variance of $\boldsymbol{D}$- property of the asymptotic bias of the LS-estimator in the AR(1) case (Denby and Martin (1979)). In addition, the bias of the LMS, the RLS and the FLS-estimator is significantly affected by the order of the AR-model. Hence, we may conjecture that the breakdown point of these methods is reduced compared to the regression setup and depends on the order $\boldsymbol{p}$, as its is the case for the GM and the S-estimator (see Martin and Yohai (1985) and (1991), respectively). Turning to efficiency, it can be seen that the calculated variance indicates the LMS to be the least appropriate method. All other methods appear to be (overall) comparable, since no estimator clearly dominates the others. Nevertheless, differences do exist in extreme-tailed situations particularly for the lower order model. In Figures 5-12 the TRMSE is shown for selected sizes and shapes of contamination. Apparently the performance of the LMS-estimator is not significantly affected by tail-changes within any ε-family of additive-effect distributions. This is also true, though to a lesser extent, for increases in the value of the contamination parameter ε. The flatness of the LMS in combination with the relative sensitiveness of the remaining estimators to changes in size and shape of contamination makes the LMS competitive for large values of heavy-tailed ε-contamination. Also the FLS and the TLS become indistinguishable for $\varepsilon \geq 0.10$ indicating that iteration (5) fails to produce a more efficient estimator in such cases, especially for the AR(3)-model. The estimator with the best performance overall is the RLS followed by the FLS and the TLS-estimator whereas, the LS is comparable to the best estimator for small percentage of medium-tailed contamination but looses MSE-efficiency with increasing and tail-heaviness.

Table 2. AO-scheme: Total root normalized bias (left entry) and total root normalized variance (right entry) of estimators

ε =0.01											
Additive Effect Distribution		*LS*		*LMS*		*RLS*		*FLS*		*TLS*	
						p=2					
S(α)	α										
	1.0	5.98	3.51	2.47	3.94	2.77	1.83	4.10	2.67	4.93	2.96
	1.3	4.93	2.74	2.68	3.90	2.80	1.82	3.66	2.16	4.09	2.42
	1.6	4.26	2.23	2.82	3.89	2.82	1.82	3.26	1.89	3.60	2.10
	1.9	3.78	1.82	2.66	3.89	2.84	1.82	3.21	1.72	3.30	1.92
$N(0,\sigma^2)$	σ										
	2	4.80	2.00	2.29	4.00	3.15	1.96	3.82	1.81	3.93	1.96
	3	6.33	2.57	1.95	4.02	3.17	2.04	3.79	2.05	4.83	2.10
	5	8.59	3.53	2.22	4.00	3.12	2.06	3.11	2.45	6.26	2.47
	10	11.6	4.49	2.06	4.03	2.89	2.09	3.01	3.56	8.65	3.63
t_v	*v*										
	2	4.47	2.43	2.75	3.81	2.68	1.80	3.13	1.92	3.76	2.18
	3	3.88	1.89	2.09	4.02	2.61	1.86	3.00	1.68	3.08	1.90
	6	3.11	1.72	1.78	3.87	2.33	1.84	2.71	1.68	2.69	1.85
	12	2.93	1.64	1.78	3.91	2.57	1.82	2.76	1.61	2.63	1.83
						p=3					
S(α)	α										
	1.0	6.90	4.03	3.56	7.54	4.36	3.62	6.81	3.86	6.56	4.20
	1.3	5.77	3.59	3.88	7.57	4.06	3.55	5.85	3.44	5.40	3.82
	1.6	5.02	3.32	3.51	7.57	3.86	3.54	5.21	3.31	4.77	3.67
	1.9	4.50	3.10	3.73	7.53	3.81	3.49	4.97	3.13	4.35	3.57
$N(0,\sigma^2)$	σ										
	2	5.64	3.16	2.85	7.55	4.47	3.54	6.00	3.15	5.02	3.53
	3	7.48	3.46	4.19	7.52	5.57	3.66	7.24	3.46	6.58	3.62
	5	10.0	4.00	4.79	7.50	6.52	3.84	9.22	3.93	8.88	3.93
	10	13.2	4.49	5.04	7.62	7.37	4.15	11.6	4.59	11.9	4.36
t_v	*v*										
	2	5.23	3.42	3.88	7.50	3.88	3.55	5.44	3.36	5.03	3.72
	3	4.57	3.21	3.82	7.66	3.80	3.54	5.07	3.17	4.24	3.62
	6	3.73	3.12	2.46	7.61	3.12	3.49	4.24	3.13	3.52	3.49
	12	3.43	3.03	2.49	7.61	2.99	3.44	4.10	3.06	3.37	3.55

Table 3. AO-scheme: Total root normalized bias (left entry) and total root normalized variance (right entry) of estimators

ε=0.05

Additive Effect Distribution		*LS*		*LMS*		*RLS*		*FLS*		*TLS*	
						p=2					
S(α)	α										
	1.0	10.6	4.85	3.94	4.16	4.65	2.22	7.80	4.48	9.06	4.43
	1.3	8.97	4.07	4.06	4.08	4.78	2.16	6.69	3.46	7.59	3.56
	1.6	7.71	3.24	4.17	4.06	4.81	2.06	5.95	2.57	6.50	2.80
	1.9	6.68	2.35	4.00	4.03	4.79	2.05	5.71	2.12	5.71	2.24
$N(0,\sigma^2)$	σ										
	2	8.32	2.40	4.02	4.26	5.48	2.39	6.87	2.50	6.96	2.28
	3	10.6	2.82	4.05	4.33	5.67	2.63	7.18	3.53	8.89	2.78
	5	13.3	2.85	3.68	4.37	5.51	2.86	8.88	5.08	11.6	3.37
	10	15.6	1.98	3.38	4.17	4.78	2.77	13.8	5.08	14.8	2.91
t_ν	ν										
	2	7.83	3.31	3.63	4.08	4.52	2.11	5.53	2.47	6.48	2.69
	3	6.64	2.55	3.67	4.21	4.39	2.11	5.07	2.10	5.48	2.19
	6	5.55	2.02	3.33	4.07	4.19	2.02	4.88	1.90	4.78	2.01
	12	5.06	1.86	3.52	4.03	4.20	2.00	4.78	1.81	4.49	1.98
						p=3					
S(α)	α										
	1.0	12.1	4.72	6.52	7.81	8.21	4.09	11.6	4.54	11.5	4.74
	1.3	10.4	4.30	6.96	7.57	7.88	3.85	10.2	4.10	9.88	4.37
	1.6	9.09	3.82	6.71	7.57	7.53	3.78	9.03	3.65	8.63	3.98
	1.9	8.01	3.35	6.62	7.49	7.14	3.61	8.23	3.29	7.66	3.68
$N(0,\sigma^2)$	σ										
	2	9.94	3.32	7.58	7.44	8.69	3.73	10.0	3.28	9.29	3.61
	3	12.5	3.43	9.01	7.34	10.3	3.94	12.1	3.56	11.7	3.64
	5	15.2	3.22	9.81	7.37	11.7	4.20	14.5	3.65	14.4	3.48
	10	17.3	2.31	10.7	7.15	12.5	4.61	16.8	2.83	16.9	2.63
t_ν	ν										
	2	9.20	3.89	6.47	7.54	7.31	3.74	9.01	3.72	8.76	4.01
	3	7.96	3.50	6.19	7.51	6.78	3.65	8.06	3.40	7.49	3.82
	6	6.74	3.23	3.99	7.56	5.86	3.65	7.06	3.22	6.47	3.57
	12	6.23	3.12	4.68	7.55	5.79	3.47	6.70	3.13	6.02	3.56

Table 4. AO-scheme: Total root normalized bias (left entry) and total root normalized variance (right entry) of estimators

$\varepsilon = 0.10$

Additive Effect Distribution		*LS*		*LMS*		*RLS*		*FLS*		*TLS*	
						p=2					
S(α)	α										
	1.0	14.1	3.93	5.81	4.48	7.31	2.90	12.3	5.14	13.1	4.35
	1.3	12.4	3.90	6.10	4.44	7.46	2.73	10.4	4.36	11.2	3.97
	1.6	10.8	3.37	6.20	4.45	7.57	2.66	9.19	3.33	9.63	3.19
	1.9	9.45	2.51	6.23	4.46	7.52	2.49	8.61	2.50	8.54	2.49
$N(0,\sigma^2)$	σ										
	2	11.2	2.45	6.58	4.72	8.49	2.92	10.3	2.90	10.2	2.63
	3	13.4	2.41	7.18	5.08	9.27	3.56	12.2	3.69	12.6	2.80
	5	15.2	1.98	7.19	5.42	9.28	4.33	14.6	3.42	14.8	2.30
	10	16.4	1.46	7.06	5.78	8.60	4.98	16.3	1.63	16.3	1.44
t_v	*V*										
	2	10.8	3.57	5.63	4.35	6.95	2.61	8.59	3.37	9.42	3.31
	3	9.36	2.86	5.59	4.44	6.76	2.52	7.97	2.75	8.22	2.62
	6	8.00	2.33	5.20	4.27	6.40	2.35	7.32	2.25	7.12	2.33
	12	7.31	2.09	5.47	4.38	6.30	2.30	6.99	2.06	6.68	2.20
						p=3					
S(α)	α										
	1.0	15.9	3.70	11.0	7.20	12.3	4.13	15.3	3.87	15.4	3.85
	1.3	14.2	3.88	10.6	7.34	11.6	4.00	13.8	3.85	13.6	3.97
	1.6	12.6	3.69	10.0	7.31	11.0	3.75	12.5	3.55	12.1	3.84
	1.9	11.2	3.30	9.91	7.19	10.4	3.60	11.3	3.26	10.9	3.62
$N(0,\sigma^2)$	σ										
	2	13.1	3.20	10.9	7.32	12.0	3.68	13.1	3.16	12.7	3.39
	3	15.2	3.07	12.8	6.90	13.9	3.67	15.1	2.96	14.9	3.13
	5	16.8	2.79	14.5	6.09	15.3	3.59	16.8	2.52	16.7	2.57
	10	17.9	2.43	15.5	5.50	16.2	3.64	17.9	1.72	17.8	1.80
t_v	*V*										
	2	12.5	3.89	9.77	7.27	10.6	3.92	12.2	3.77	12.0	3.94
	3	11.0	3.58	9.07	7.46	9.85	3.75	11.0	3.47	10.6	3.77
	6	9.53	3.34	8.03	7.52	8.70	3.76	9.72	3.32	9.23	3.59
	12	8.84	3.23	7.45	7.49	8.34	3.62	9.16	3.23	8.70	3.61

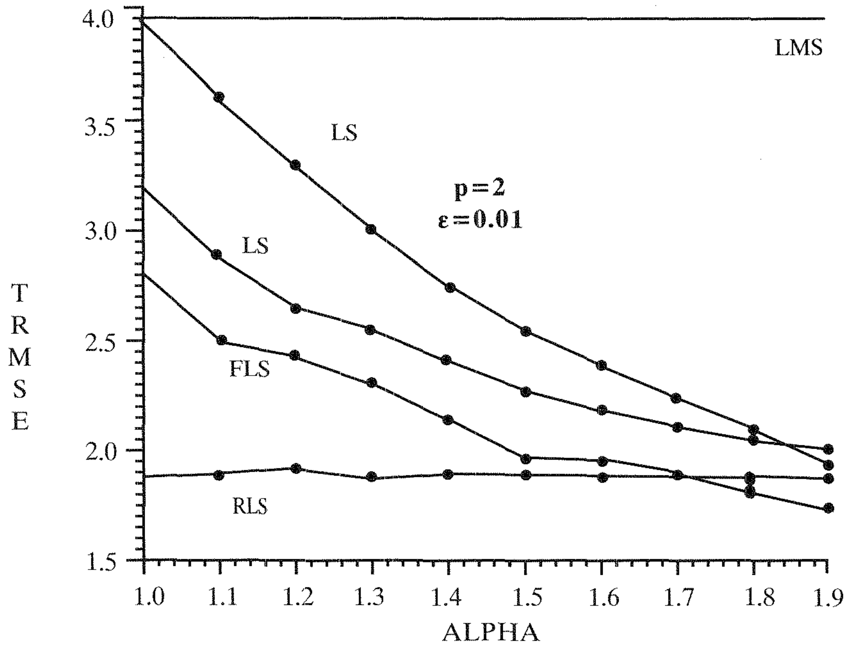

Figure 5. AO-scheme: Efficiency under symmetric-stable additive effects

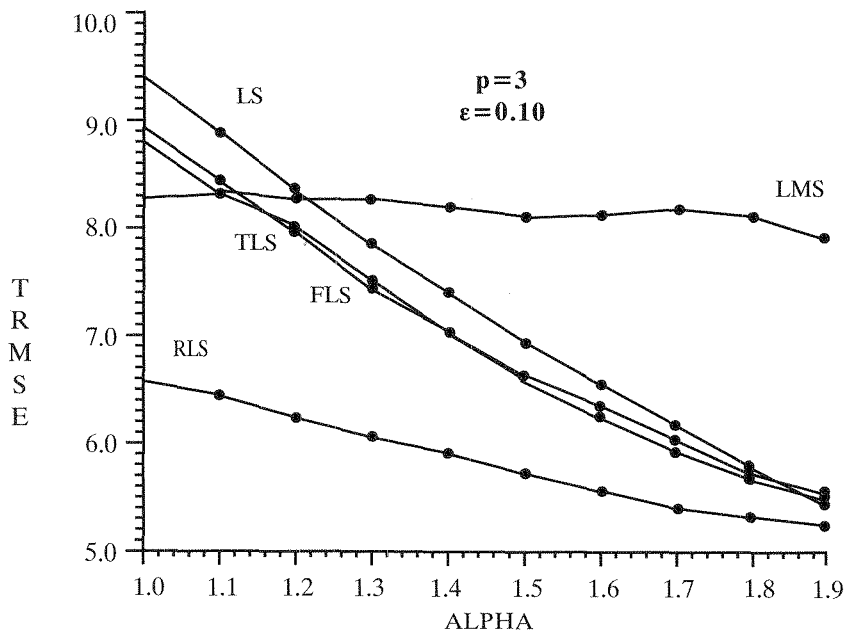

Figure 6. AO-scheme: Efficiency under symmetric-stable additive effects

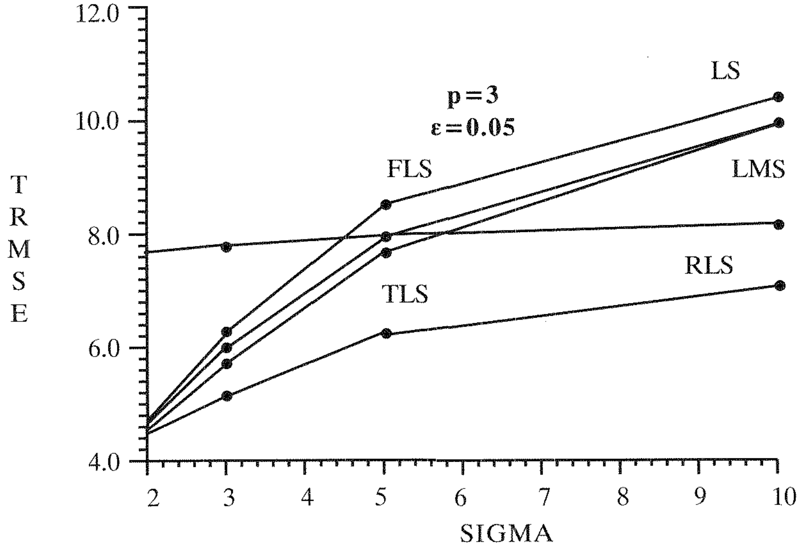

Figure 7. AO-scheme: Efficiency under normal additive effects

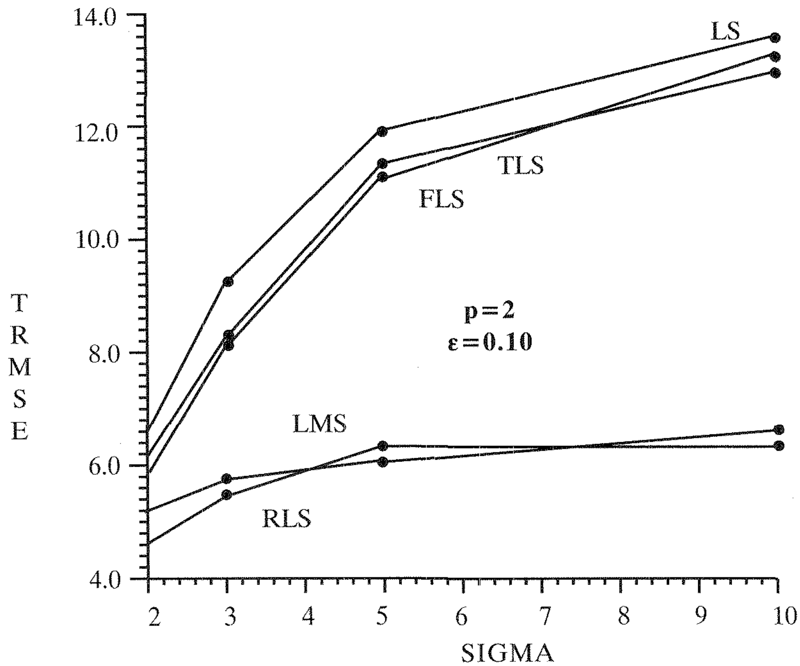

Figure 8. AO-scheme: Efficiency under normal additive effects

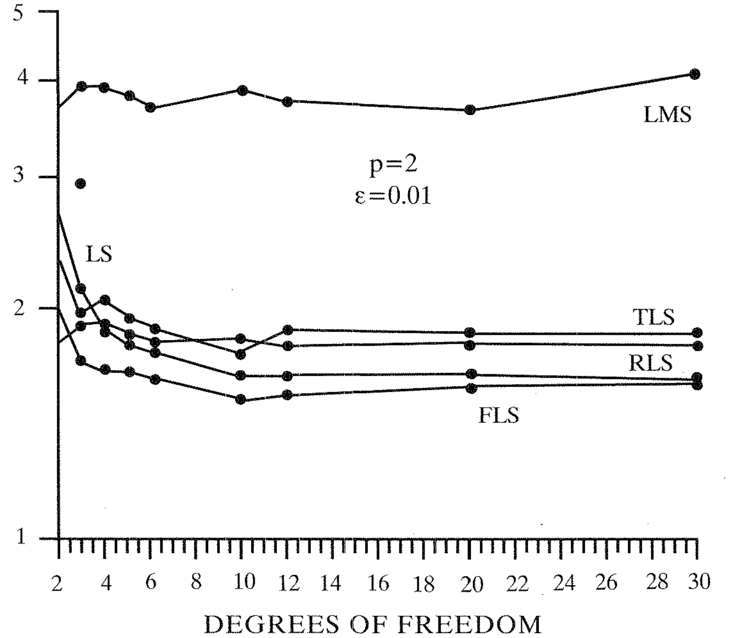

Figure 9. IO-scheme: Efficiency under Student's-t additive effects

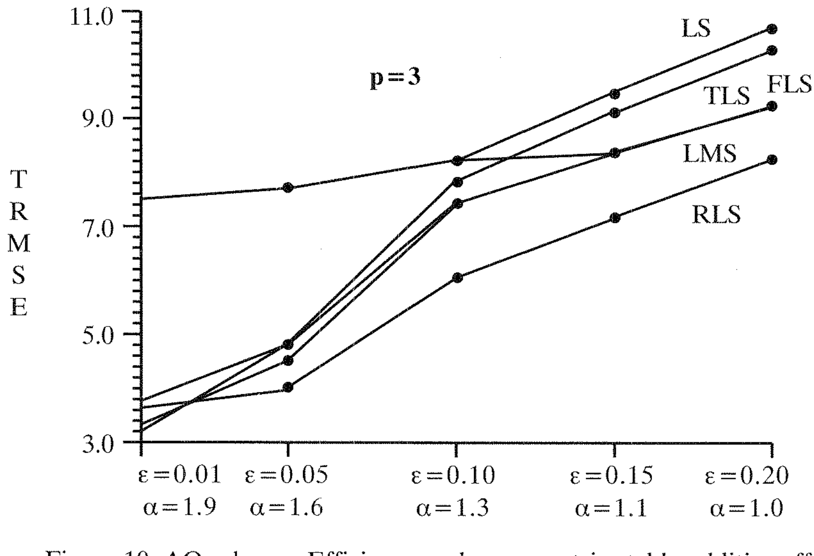

Figure 10. AO-scheme: Efficiency under symmetric-stable additive effects

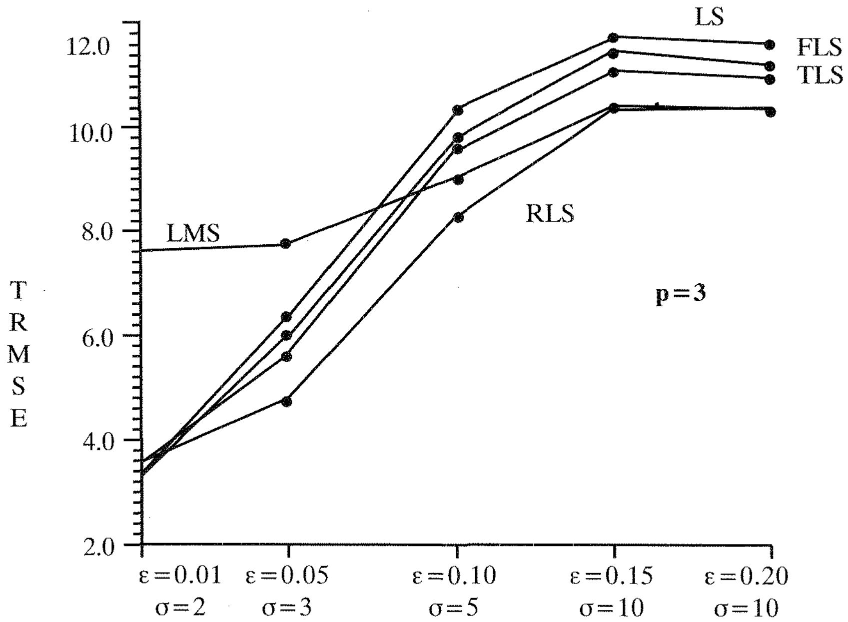

Figure 11. AO-scheme: Efficiency under normal additive effects

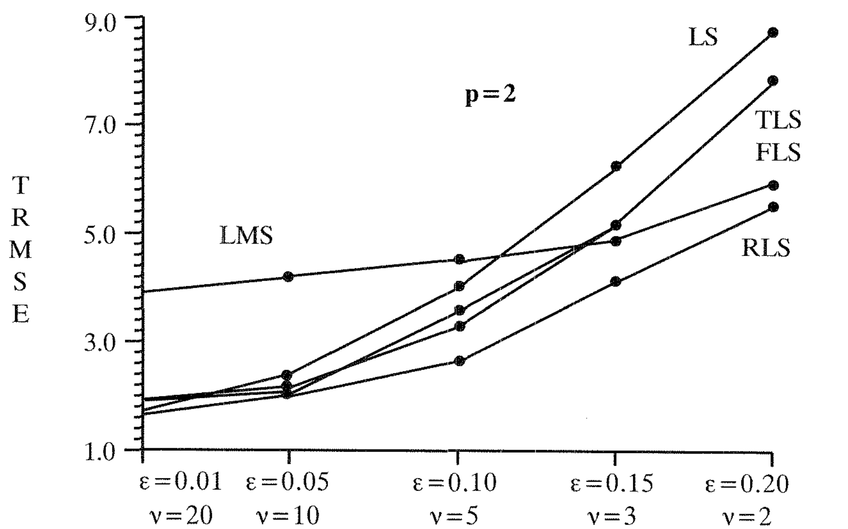

Figure 12. AO-scheme: Efficiency under studet's-t additive effects

6. Conclusion

Robust autoregressive-model estimators occupy a particularly important place in time-series methodology. In this article we present a simulation study of four robust alternatives to the LS- estimator under innovation and additive outliers. The results indicate that under the IO- scheme, the functional least squares is the estimator on which innovation outliers have the relatively mildest impact. However, the reweighted least squares and the trimmed least squares also perform well and the general conclusion is that in most cases the estimators are not seriously affected by this type of outliers. On the contrary, the effect of additive outliers on all five estimators considered is significant and for most of them, of increasing importance as the proportion and the shape of contamination changes. Nevertheless, the reweighted least squares appears to be the most efficient method. The other two estimators mentioned in the IO-case above also deserve some merit under the AO-scheme. Finally, on the whole and under both outlier schemes the least squares and the least median of squares are the estimators with less satisfactory performance, although the former exhibits considerable efficiency near the standard normal model while the latter outperforms almost all estimators in terms of estimated bias.

References

1. Allende, H. and Heiler, S.(1992). "Recursive generalized M estimates for atore gressive moving-average models", J.Time Ser.Anal., 13, 1-18.
2. Anderson, T.W.(1971). The Statistical Analysis of Time Series, John Wiley, New York.
3. Beran, R.(1976). "Adaptive estimates for autoregressive processes", Ann.Inst.Statist.Math., 28, 77-89.
4. Bustos, O.H.(1982). "General M-estimates for contaminated pth-order autoregressive processes: Consistency and asymptotic normality", Z .Wahrsch.verw.Gebiete, 59, 491-504.
5. Bustos, H.O. and Yohai, V.J.(1986). "Robust estimates for ARMA models", J.Amer.Statist.Assoc., 81, 155-168.
6. Campbell, K.(1982). "Recursive computation of M-estimates for the parameters of a finite autoregressive process", Ann.Statist., 10, 442-453.
7. Chambers, R.L. and Heathcote, C.R.(1981). "On the estimation of slope and the identification of outliers in linear regression", Biometrika, 68, 21-33.
8. Chang, I., Tiao, G.C. and Chen, C.(1988). "Estimation of time series parameters in the presence of outliers", Technometrics, 30, 193-204.
9. Cipra, T., Rubio, A. and Canal, J.L.(1993). "Recursive estimation in autoregressive models with additive outliers", Kybernetika, 29, 62-72.
10. Csorgo, S.(1983). "The theory of functional least squares", J. Austral. Math.Soc.Ser.A, 34, 336-355.
11. Davis, R.A., Knight, K. and Liu, J.(1992). "M-estimation for autoregressions

with infinite variance", Stoc.Proc. and their Appl., 40, 145-180.

12. Denby, L.and Martin, R.D.(1979). "Robust estimation of the first-order autoregressive parameter", *J.Amer.Statist.Assoc.*, 74, 140-146.
13. Dhar, S.K.(1990). "Functional least squares estimators in an additive effects outliers model", *J.Austral.Math.Soc.Ser.A*, 48, 299-319.
14. Dhar, S.K.(1993). "Computation of certain minimum distance estimators in AR[k] model", *J.Amer.Statist.Assoc.*, 88, 278-283.
15. Gross, S. and Steiger, W.L.(1979). "Least absolute deviation estimates in autoregression with infinite variance", *J.Appl.Prob.*, 16, 104-116.
16. Hamilton, J.D.(1994). *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
17. Hannan, E.J. and Kanter, M.(1977). "Autoregressive processes with infinite variance", *J.Appl.Prob.*,14, 411-415
18. Heathcote, C.R. and Welsh, A.H.(1983). "The robust estimation of autoregressive processes by functional least squares", *J.Appl.Prob.*, 20, 737-753. IMSL (1987). *IMSL User's Manual Math and Stat/Library*, IMSL Inc., Houston, Texas.
19. Kanter, M. and Steiger, W.L.(1974). "Regression and autoregression with infinite variance", *Adv.Appl.Prob.*, 6, 768-783.
20. Lee, A. and Hui, Y.V.(1993). "Outlier detection in time series", *J.Statist.Comput.Simul.*, 45, 77-95. Li, W.K. and McLeod, A.I.(1988). "ARMA modelling with non-Gaussian innovations", *J. Time Ser.Anal.*, 9, 155-168.
21. Maronna, R.A., Yohai, V.J. and Zamar, R.J.(1993). "Bias-robust regression estimation: A partial survey", in
22. Morgenthaler, Ronchetti, Stahel (Eds.) *New Directions in Statistical Data Analysis and Robustness*, 157-176.
23. Martin, R.D. and Yohai, V.J.(1985). "Robustness in time series and estimating ARMA models", in Hannan, Krishnaiah, Rao (Eds.) *Handbook of Statistics*, Vol. 5, 119-155.
24. Martin, R.D. and Yohai, V.J.(1991). "Bias robust estimation of autoregressive parameters", in Stahel, Weisberg (Eds.) *Directions in Robust Statistics and Diagnostics*, Part I, 233-246.
25. McDougall, A.J.(1994). "Robust methods for recursive autoregressive moving average estimation", *J.R.Statist.Soc.B*, 56, 189-207.
26. Mikosch, T., Gadrich, T., Kluppelberg, C. and Adler, R.J.(1995). "Parameter estimation for ARMA models with infinite variance innovations", *Ann.Statist.*, 23, 305-326.
27. Rousseeuw, P.J.(1984). "Least median of squares regression", *J.Amer.Statist.Assoc.*, 79, 871-880.
28. Rousseeuw, P.J. and Leroy, A.M.(1987). *Robust Regression and Outlier Detection*, John Wiley , New York.
29. Ruppert, D. and Carroll, R.J.(1980). "Trimmed least squares in the linear model", *J.Amer.Statist.Assoc.*, 75, 828-838.

30. Salau, M.O.(1995). "On the numerical implementation of the generalized least squares procedure for ARMA estimation", *Commun.Statist.-Simul.Comput.*, 24, 111-130.
31. Sejling, K., Madsen, H., Holst, J. ,Holst, U. and Englund, J.E.(1994). "Methods for recursive robust estimation of AR parameters", *Comput.Statist. and Dat.Anal.*, 17, 509-536.
32. Yohai, V.J. and Maronna, R.A.(1977). "Asymptotic behavior of least-squares estimates for autoregressive processes with infinite variance", *Ann.Statist.*, 5, 554-560.