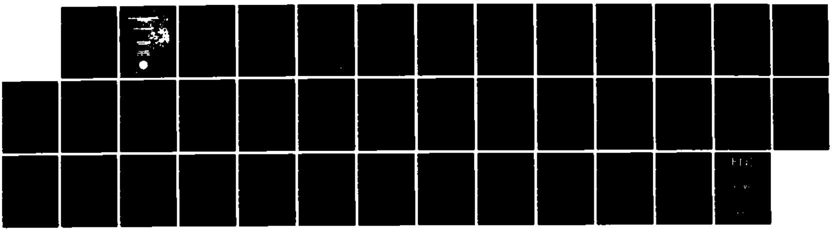
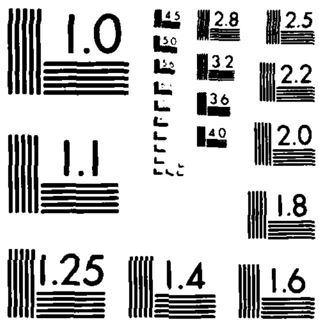
AD-A149 705 ORDINARY AND PROPER LOCATION M-ESTIMATES FOR ARMA MODELS REVISED(U) WASHINGTON UNIV SEATTLE DEPT OF STATISTICS C H LEE ET AL. JUL 84 TR-29 1/1

UNCLASSIFIED N00014-84-C-0169 F/G 12/1 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

AD-A149 705

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 29	2. GOVT ACCESSION NO. AD A149 905	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Robustness in time series and estimating ARMA models		5. TYPE OF REPORT & PERIOD COVERED TR 1/1/82 - 11/30/83
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Chin-Hui Lee and R. Douglas Martin		8. CONTRACT OR GRANT NUMBER(s) N00014-84-C-0169
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics, GN-22 University of Washington Seattle, WA 98195		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-608-002
11. CONTROLLING OFFICE NAME AND ADDRESS ONR Code N63374 1107 N.E. 45th Street Seattle, WA 98105		12. REPORT DATE 7/84
		13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

M-estimates, ARMA models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Proper location M-estimates for a model with non-Gaussian autoregressive-moving average type errors are genuine maximum likelihood type estimates, whereas ordinary location M-estimates are those introduced by P. Huber for independent and idetically distributed errors. The relative behavior of ordinary location M-estimates and proper location M-estimates is studied for situations with dependent errors of purely autoregressive and purely moving average type. It is shown through asymptotic calculations and (continued)

20. ABSTRACT (continued)

finite-sample Monte Carlo studies that although ordinary location M-estimates are adequate for weak dependency structure, they can be quite inefficient compared with proper M-estimates of location when the non-Gaussian errors have a moderate to strong dependency structure.

4

Ordinary and Proper Location M-Estimates for ARMA Models

Chin-Hui Lee

R. Douglas Martin

Department of Statistics
University of Washington
Seattle, Washington

ABSTRACT

Proper location M-estimates for a model with non-Gaussian autoregressive-moving average type errors are genuine maximum likelihood type estimates, whereas ordinary location M-estimates are those introduced by P. Huber for independent and identically distributed errors. The relative behavior of ordinary location M-estimates and proper location M-estimates is studied for situations with dependent errors of purely autoregressive and purely moving average type. It is shown through asymptotic calculations and finite-sample size Monte Carlo studies that although ordinary location M-estimates are adequate for weak dependency structure, they can be quite inefficient compared with proper M-estimates of location when the non-Gaussian errors have a moderate to strong dependency structure.

Revised July 1984

Ordinary and Proper Location M-Estimates for ARMA Models

Chin-Hui Lee
R. Douglas Martin
Department of Statistics
University of Washington
Seattle, Washington

1. INTRODUCTION

By now, P. Huber's (1964) M-estimates of location are well known. These estimates were introduced in the context of obtaining robust estimates of location μ for independent and identically distributed observations $Y_1, Y_2, \ldots, Y_n$. For reasons which become clear in the next section we refer to Huber's estimates as *ordinary* location M-estimates, and label them $\hat{\mu}_{OM}$. An ordinary location M-estimate is obtained by solving

$$\sum_{i=1}^{n} \psi\left[\frac{Y_i - \hat{\mu}_{OM}}{c \cdot \hat{s}_y}\right] = 0 \qquad (1.1)$$

with a good algorithm, where $\hat{s}_y$ is a consistent robust estimate of the scale s_y of the Y_t, c is a tuning constant and ψ is a robustifying psi-function. With $\psi = \rho'$, this estimating equation characterizes a stationary point of the minimization problem

$$\min_{\mu'} \sum_{1}^{n} \rho\left[\frac{Y_i - \mu'}{c \cdot \hat{s}_y}\right] .$$

Bounded and continuous psi-functions result in qualitative robustness for ordinary location M-estimates at certain distributions, including the normal distribution. This is true not only when the Y_t are independent and identically distributed (Hampel, 1971), but also when the Y_t are dependent (Papantoni-

Kazakos and Gray, 1979; Cox, 1981; Boente, Fraiman and Yohai, 1982).

The asymptotic and finite-sample size efficiency robustness of ordinary location M-estimates have been extensively studied under the independent and identically distributed observations setup. The issue of efficiency robustness where the distribution for the data is both dependent and possibly has a heavy-tailed non-Gaussian has received relatively little attention. Notable exceptions include the theoretical work of Portnoy (1977), and the Monte Carlo study of Wegman and Carrol (1977).

The essence of Portnoy's results are that for moving-average type non-Gaussian errors with *weak* correlation structure, ordinary location M-estimates do well in terms of efficiency relative to the asymptotic Cramer-Rao lower bound. In addition, through use of a small correlation expansion, Portnoy was able to obtain approximate asymptotic min-max results which involved a *redescending* psi-function.

Portnoy's work left unanswered the question of how ordinary location M-estimates would fare with moderate to large correlation structures and a heavy-tailed distribution. This paper partially answers the question through efficiency comparisons at perfectly-observed non-Gaussian first-order autoregressive and moving-average models. Efficiencies are obtained by some exact asymptotic variance calculations, and by Monte Carlo. The results show that ordinary location M-estimates can be seriously lacking of efficiency robustness in such situations. On the other hand, as expected, proper M-estimates have high efficiency robustness.

The next section briefly introduces proper M-estimates, while Section 3 gives the asymptotic variance expressions for both ordinary and proper M-estimates. These expressions reveal almost immediately some substantially negative aspects of ordinary location M-estimates in dependent process situations. Section 4 gives exact asymptotic comparisions for first-order moving average models, while Section 5 gives finite-sample Monte Carlo results for both first-order moving average and first-order autoregressive models.

2. PROPER M-ESTIMATES OF LOCATION

Suppose that μ is a location parameter and that the observations are

$$Y_t = \mu + V_t, \quad t = 1,2, \ldots, n \tag{2.1}$$

where V_t is an ARMA(p,q) model

$$V_t + \varphi_1 V_{t-1} + \cdots + \varphi_p V_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{2.2}$$

with the ε_t being independent and having a common symmetric distribution $G(\varepsilon) = G_0(\varepsilon / s_\varepsilon)$, s_ε being a scale parameter for the innovations. The ε_t are often called the *innovations* process. This yields the equivalent ARMA(p,q) model

$$Y_t + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} = \gamma + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{2.3}$$

where the intercept is

$$\gamma = \mu(1 + \Sigma \varphi_i) \quad . \tag{2.4}$$

Let $\underline{\alpha}' = (\gamma', \underline{\varphi}', \underline{\theta}')$ represent arbitrary parameter values in the region of stationarity and invertiblity for the ARMA process, and let $\underline{\alpha} = (\gamma, \underline{\varphi}, \underline{\theta})$ represent the true parameter values. Denote by $r_t(\underline{\alpha}')$ the residuals computed from an observed sample $Y_1, \cdots, Y_n$ by one of the usual variants with regard to initial conditions (see for example, Box and Jenkins, 1976). An M-estimate of $\underline{\alpha}$ is a solution of the minimization problem

$$\min_{\alpha'} \sum_{t=1}^{n} \rho \left[\frac{r_t(\underline{\alpha}')}{c \cdot \hat{s}_\varepsilon} \right] \tag{2.5}$$

where ρ is a robustifying loss function. The constant c is a tuning constant and $\hat{s}_\varepsilon$ is a robust estimate of the innovations scale s_ε.

Now given an M-estimate $\hat{\underline{\alpha}}$ of $\underline{\alpha} = (\gamma, \underline{\varphi}, \underline{\theta})$, the relation (2.4) leads to the *proper* M-estimate of location

$$\hat{\mu} = \frac{\hat{\gamma}}{1 + \Sigma \hat{\varphi}_i} \quad . \tag{2.6}$$

Consistency and asymptotic normality of $\hat{\underline{\alpha}}$ and $\hat{\mu}$ have been established by Lee and Martin (1982a).

In the special case where $\rho(t) = -\log g_o(t)$, with g_o the density for G_o, $\hat{\underline{\alpha}}$ and $\hat{\mu}$ are conditional maximum-likelihood estimates of $\underline{\alpha}$ and μ, where the conditioning involves fixing not only $Y_1, \cdots Y_p$, but also estimates $\hat{\varepsilon}_1, \cdots, \hat{\varepsilon}_q$ of $\varepsilon_1, \cdots, \varepsilon_q$. These conditional maximum-likelihood estimates are of course asymptotically efficient under regularity conditions.

3. ASYMPTOTIC CONSIDERATIONS

First consider an ordinary location M-estimate $\hat{\mu}_{0M}$ computed from observations $Y_1, \ldots, Y_n$ in (2.1) which have a common marginal distribution $F(y) = F_o((y-\mu)/s_y)$. Under regularity conditions (see for example Portnoy, 1977) $\hat{\mu}_{0M}$ is consistent and asymptotically normal, with asymptotic variance given by

$$V_{0M} = \frac{C(0)+2\sum_{l=1}^{\infty} C(l)}{E^2_{F_o}\psi'_c(Y_1)} \tag{3.1}$$

where

$$C(l) = s_y^2 E_{F_{0l}}\psi_c(Y_1)\psi_c(Y_{1+l}), \quad l = 0,1,2,\cdots \quad . \tag{3.2}$$

Here for $l = 0$, F_{0l} is the standardized marginal distribution F_0 of the Y_t, while for $l \geq 1$ F_{0l} is the bivariate distribution for (Y_1, Y_{l+l}) obtained when $\mu=0$ and $s_y=1$. The tuning constant c appearing in (1.1) is now (and henceforth) absorbed in the definition of ψ_c. In the special case of independent Y_t, $F_0 = G_0$ and V_{0M} reduces to

$$V_{0M} = s_y^2 \frac{E_{F_o}\psi_c^2(Y_1)}{E^2_{F_o}\psi'_c(Y_1)} = s_y^2 V_{loc}(\psi_c, F_o) \tag{3.3}$$

where $V_{loc} = V_{loc}(\psi, F_o)$, defined by the right-hand equality above, is P. Huber's (1964) well-known expression for the asymptotic variance of ordinary location M-estimates.

Now for the case of a *proper* location M-estimate $\hat{\mu}$, it can be shown (Lee and Martin, 1982a), that the asymptotic variance expression is

$$V = \frac{(1+\Sigma\theta_i)^2}{(1+\Sigma\varphi_i)^2} s_\varepsilon^2 \; V_{loc}(\psi_c, G_o) \; . \tag{3.4}$$

The quantity $s_\varepsilon^2(1+\Sigma\theta_i)^2/(1+\Sigma\varphi_i)^2$ differs by only a constant factor from the value at zero frequency of the spectrum of the process Y_t. When ψ is the identity function so that $\hat{\mu}_{0M} = \hat{\mu}_{LS} = \overline{Y}$, and s_ε is the standard deviation, (3.4) yields the well-known result that the asymptotic variance of the sample mean is given by the spectrum of the process evaluated at zero frequency (Grenander, 1954, 1981).

The simplicity of the expression for V relative to that of V_{0M} is quite attractive, particularly with regard to the relative ease of studentizing the estimate $\hat{\mu}$ for the purpose of constructing confidence intervals. Estimation of V from the data for this purpose may be quite manageable, whereas estimation of V_{0M} seems rather impractical when many $C(l)$ are non-zero. In this regard the situation is particularly bad when an autoregression component is present, since then the $C(l)$ only vanish asymptotically.

Furthermore, the effect of the tuning constant c on the asymptotic efficiency of $\hat{\mu}$ shows up only in the V_{loc} factor of the expression for V. Since V_{loc} is not affected by the dependency structure for Y_t, as specified by the parameters φ_i and θ_i, efficiencies can be controlled through c without regard to the values of these parameters. This is not the case with regard to V_{0M}, as can be seen in the following equivalent form of (3.1):

$$V_{0M} = \left[1+2\sum_{l=1}^{\infty}\rho_{1,1+l}\right] s_y^2 \cdot V_{loc}(\psi_c, F_o) \tag{3.5}$$

where $\rho_{1,1+l}$ is the correlation coefficient for the random variables $\psi_c(Y_1)$ and $\psi_c(Y_{1+l})$ when $(Y_1, Y_{1+l}) \sim F_{0l}$. Here the effects of c appear not only in V_{loc}, but also in the correlation coefficients $\rho_{1,1+l}$, and the latter depend on the ARMA model parameters φ_i and θ_i. This makes the adjustment of c to obtain desired Gaussian process efficiencies quite onerous, if not impractical.

In lieu of a better scheme, one would probably choose c for $\hat{\mu}_{0M}$ such that a desired efficiency is obtained for independent and identically distributed Gaussian data. It should be noted that such a value of c yields the same efficiency for $\hat{\mu}$ at *any* Gaussian ARMA process (see first paragraph of Section 4 in this regard).

In order to gain some insight into why $\hat{\mu}$ might be significantly more efficient than $\hat{\mu}_{OM}$ at highly correlated non-Gaussian ARMA situations, consider the case where Y_t is a first-order autoregression with parameter φ. In this case V may be expressed in the following form, which facilitates comparison with (3.5):

$$V = \left[1-2\frac{\varphi}{1+\varphi}\right]\frac{s_\varepsilon^2}{1-\varphi^2}V_{loc}(\psi_c,G_o) \tag{3.6}$$

It is easy to check that the factors in square brackets in (3.5) and (3.6) are identical when ψ is the identity function. We conjecture that these factors do not differ by too much for either Gaussian or non-Gaussian processes Y_t when ψ is one of the popular psi-functions. Assuming that this is the case, the behavior of V_{OM} relative to V will be determined by the relative values of $V_{loc}(\psi_c,F_o)$, $V_{loc}(\psi_c,G_o)$, s_y^2 and $s_\varepsilon^2/(1-\varphi^2)$.

Suppose that the same value of tuning constant c is used for both the ordinary and proper location M-estimates (in view of our previous comments, this is not an unlikely scenario). Then we can expect that in many non-Gaussian situations $V_{loc}(\psi_c,F_o)$ will be larger than $V_{loc}(\psi_c,G_o)$ when $\varphi\neq 0$. This is because Y_t is a weighted sum of the ε_t, and the convolutions which produce F_o from non-Gaussian G_o will often result in an F_o having heavier tails than G_o. At the same time s_y^2 and $s_\varepsilon^2/(1-\varphi^2)$ will be identical in finite-variance non-Gaussian situations, and then we may expect that V_{OM} is larger than V.

Of course for stable G_o we will have $F_o = G_o$, and then the two V_{loc}'s will be identical. However, in such a case s_y^2 and $s_\varepsilon^2/(1-\varphi^2)$ will no longer be identical (except in The Gaussian case). For example, when G_o is a symmetric stable distribution with index η, F_o is also a symmetric stable distribution, and it is easy to check that (see Feller, 1966)

$$R \equiv \frac{s_y^2}{s_\varepsilon^2/(1-\varphi^2)} = \frac{1-\varphi^2}{(1-|\varphi|^\eta)^{2/\eta}} \quad . \tag{3.7}$$

The Cauchy distribution is obtained when $\eta=1$, and in this case we have $R=3$ and 19 when $\varphi=0.5$ and 0.9, respectively. If we assume that the expressions (3.5) and (3.6) hold for infinite-variance situations, and that the square-bracketed factors in (3.5) and (3.6) are not too different, then V_{0M} may be much larger than V.

In the concluding comments section of the paper, a more direct heuristic argument is also offered in explanation of the relative inefficiency of $\hat{\mu}_{0M}$.

4. EXACT ASYMPTOTIC RELATIVE EFFICIENCY RESULTS

The asymptotic absolute efficiencies of a proper M-estimate at various distributions are the same as those of an ordinary location M-estimate based on matching ψ_c, with independent observations. This follows from the fact that the asymptotic lower bound on variance is given by (3.4) with V_{loc} replaced by the reciprocal of the Fisher information $i(g_o) = \int (g'_o/g_o)^2 g_o$ for the standardized innovations density g_o (Martin, 1982).

Since the literature abounds with asymptotic efficiency computations for ordinary location M-estimates based on various ψ_c and independent Y_t, our main interest is in comparing $\hat{\mu}_{OM}$ with $\hat{\mu}$ for the model (2.1) - (2.2). Thus we wish to compute the *asymptotic relative efficiencies*

$$AREFF = AREFF(\psi_c, G_o, \underline{\alpha}) = \frac{V_{OM}(\psi_c, G_o, \underline{\alpha})}{V(\psi_c, G_o, \underline{\alpha})} \tag{4.1}$$

for various ψ_c, G_o and $\underline{\alpha}$.

This task is made difficult mainly because of the relatively complex structure of V_{OM}. For example, to compute (3.1) in the case of first-order autoregressions, both the stationary distribution F_o, and the bivariate distributions F_{0l}, $l = 1, 2, \ldots$, are required. Unfortunately, we can seldom specify F_o and F_{0l}, $l = 1, 2, \ldots$, in closed form when G_o is non-Gaussian (symmetric stable G_o is the main exception). Thus we study the case of a first-order autoregression solely via Monte Carlo in the next section.

On the other hand for moving-average processes of order q, the summation in (3.1) contains only a finite number of non-zero terms, and for small q we can sometimes find closed form expressions for the $C(l)$, $l = 0, 1, \ldots, q$, and $E_{F_0}^2 \psi'_c$.

We treat here the $MA(1)$ case with parameter θ, where (i) ε_1 has a contaminated normal distribution,

$$CN(\gamma,\sigma^2) = (1-\delta)N(0,1) + \delta N(0,\sigma^2) \tag{4.2}$$

and (ii) ψ has either the normal distribution shape

$$\psi_\Phi(t) = \sqrt{2\pi}\,(\Phi(t)-\tfrac{1}{2}) \tag{4.3}$$

or the shape of the derivative of the normal density,

$$\psi_{ND}(t) = t \cdot e^{-t^2/2} \quad . \tag{4.4}$$

For either of the combinations (4.2) - (4.3) or (4.2) - (4.4), a closed form expression for V_{OM} (and also for V) is obtained in a straightforward but tedious manner. These rather ugly expressions are developed in the Appendix.

It should be kept in mind that ψ_Φ and ψ_{ND} are used here only because: (i) they facilitate an exact calculation, and (ii) at the same time yield comparable efficiency robustness to that obtainable with Huber's (1964) favorite psi-function $\psi_H(t) = \max(-1,\min(1,t))$, and Tukey's bisquare psi-function (see Mosteller and Tukey, 1977), respectively. Point (ii) was verified through Monte Carlo results not reported here.

Except for the second set of results in this section, the tuning constants c_{OM} and c for the ordinary and proper M-estimates are adjusted so that for both ψ_{ND} and ψ_Φ, $\hat{\mu}_{OM}$ and $\hat{\mu}$ have matched asymptotic efficiencies of .90 for independent Gaussian observations (θ=0).

Figure 1 shows AREFF's based on ψ_{ND} for various θ values, where $\varepsilon_t \sim CN(\delta,\sigma^2)$ with δ = 0.1, $1 \leq \sigma \leq 10$. Although the AREFF's can be quite low for negative θ, they are quite high for a wide range of positive θ.

In Figure 2 we display AREFF's based on ψ_{ND} for the same values of γ,σ^2 and θ, except that c_{OM} has been adjusted to obtain matching asymptotic efficiencies of .90 for *each* value of θ and Gaussian ε_t. The values of tuning constants $c_{OM} = c_{OM}(\theta)$ needed to achieve various efficiencies are given in Table 1 for ψ_Φ, and in Table 2 for ψ_{ND}. While marked improvement in the relative performance of $\hat{\mu}_{OM}$ is achieved at θ = -.5 and -.9 at small values of σ, the improvement at large values of σ is negligible. Thus even "*proper*" adjustment of c using typically unavailable prior information on θ will not salvage $\hat{\mu}_{OM}$ for MA(1) models with negative θ.

Figures 1a and 2a give corresponding AREFF's based on ψ_ϕ. Although ψ_{ND} has the edge over ψ_ϕ at some θ values, the results are not overall too different from those in Figures 1 and 2.

TABLE 1

Tuning constants $c_{OM} = c_{OM}(\theta)$

which yield various efficiencies for ψ_ϕ

θ\EFF	0.95	0.90	0.85	0.80
-0.9	4.450	3.651	3.222	2.927
-0.7	2.343	1.870	1.611	1.431
-0.5	1.669	1.287	1.074	0.923
-0.3	1.299	0.959	0.765	0.625
-0.1	1.049	0.731	0.546	0.409
0.0	0.952	0.642	0.460	0.324
0.1	0.876	0.571	0.390	0.255
0.3	0.781	0.480	0.298	0.161
0.5	0.747	0.443	0.257	0.115
0.7	0.741	0.433	0.243	0.097
0.9	0.741	0.431	0.239	0.092

TABLE 2

Tuning constants $c_{OM} = c_{OM}(\theta)$

which yield various efficiencies for ψ_{ND}

θ\EFF	0.95	0.90	0.85	0.80
-0.9	7.839	6.478	5.752	5.258
-0.7	4.287	3.517	3.104	2.821
-0.5	3.195	2.600	2.281	2.061
-0.3	2.622	2.118	1.846	1.659
-0.1	2.249	1.802	1.561	1.397
0.0	2.110	1.685	1.456	1.300
0.1	2.002	1.594	1.374	1.225
0.3	1.874	1.485	1.276	1.134
0.5	1.832	1.448	1.242	1.101
0.7	1.827	1.443	1.235	1.094
0.9	1.829	1.443	1.235	1.094

5. MONTE CARLO RELATIVE EFFICIENCIES

In order to check both the finite-sample size *relative efficiencies* (REFF's) of $\hat{\mu}_{OM}$ and $\hat{\mu}$ for both MA(1) models as used for Figure 1, and AR(1) models, some Monte Carlo computations were carried out using 500 replications at sample size 100. Tuning constants c_{OM} were adjusted for asymptotic efficiencies of 0.9 at independent Gaussian Y_t, as described in the previous section.

The ordinary location M-estimates were computed using the median as a starting point, followed by 4 iterations of iterated-weighted least-squares using ψ_{Φ}, followed by one iteration using ψ_{ND}. The proper M-estimates were computed using 10 iterations of a nonlinear optimization algorithm for solving (2.5), which is described in Lee and Martin (1982b), followed by computing $\hat{\mu}$ from (2.6).

The results for the MA(1) case using ψ_{ND} are shown in Figure 3. The REFF's are in quite good agreement with the asymptotic REFF's of Figure 1, except for $\sigma=1$ (the Gaussian case).

Results for the AR(1) case using ψ_{ND} are given in Figure 4. Here REFF's can be quite low for positive φ as well as negative, the former case being the more commonly encountered one in practice. Furthermore, $\varphi = \pm 0.5$ can already result in REFF's as low as 70% for large σ, and for larger $|\varphi|$ the relative loss in efficiency associated with $\hat{\mu}_{OM}$ may become quite intolerable. Also, the REFF's are roughly symmetric in φ, which contrasts sharply with the MA(1) results of Figure 3.

Figures 3a and 4a give corresponding results using ψ_{Φ}. Again, ψ_{ND} tends to dominate ψ_{Φ} somewhat, but the differences are not overwhelming.

As a check on the "absolute" efficiencies of $\hat{\mu}$ at MA(1) and AR(1) models, we provide Figures 5 and 6 for ψ_{ND}, and Figures 5a and 5b for ψ_{Φ}. By "absolute" efficiencies (EFF's) we mean the asymptotic Cramer-Rao lower bound divided by the Monte Carlo variance. Except for the case $\theta = -0.9$ which requires large sample sizes to achieve high absolute efficiencies, $\hat{\mu}$ is very efficient for almost all other cases at a sample size of 100. With regard to the case $\theta = -.9$, one should keep in mind that $\theta = -1$ is a distinguished point of superefficiency (see for example, Chapter 4.4 of Grenander, 1981).

6. CONCLUDING COMMENTS

The following simple heuristic argument indicates why $\hat{\mu}$ should generally be more precise than $\hat{\mu}_{OM}$, particularly in the case of autoregressions with moderate to large correlation. Suppose one is using $\hat{\mu}_{OM}$ with robust scale estimate $\hat{s}_y$, and that the series contains just one huge isolated outlier in the ε_t at time t_0 say, after which the sample path will decay roughly like the homogenous solution to (2.2). The first part of this decay will produce residuals $r_t = Y_t - \hat{\mu}$, $t \geq t_0$, which exceed s_y in magnitude and will thus be down-weighted. Unfortunately, it is *only* the initial residual r_{t_0} that deserves downweighting, and this results in loss of information. Because the residuals in (2.5) are based on the regression with intercept form (2.3), only the residual at time t_0 will be heavily downweighted, and information in the immediately succeeding observations will be utilized.

This argument can also be cast in terms of the scatter plot of Y_t versus Y_{t-1}, say for an AR(1) process, in the spirit of Cox's (1966) comments with regard to the null distribution of the serial correlation coefficient. The pair (Y_{t_0-1}, Y_{t_0}) will be far removed from the regression line with slope φ and intercept γ, but the pairs (Y_{t-1}, Y_t), $t = t_0+1, \cdots$, constitute good *leverage* points (i.e., points which will lie close to the regression line and are large in magnitude) for estimating γ and φ -- the latter with ultra precision when ε_t has a heavy-tailed distribution (Martin, 1982). The ordinary location M-estimate would downweight such points.

The asymptotic and finite-sample efficiencies of $\hat{\mu}_{OM}$ relative to $\hat{\mu}_M$, along with awkwardness and impracticality of assessing the variability of $\hat{\mu}_{OM}$, suggest that it should be used only when one is certain that the correlation structure of the errors is quite weak. For situations where the non-Gaussian ARMA model (2.1)-(2.2) is a good approximation to reality, the proper M-estimate $\hat{\mu}$ is preferred.

When (2.1)-(2.2) does not provide a good model for non-Gaussian time series with outliers, e.g., when Y_t is corrupted with *additive outliers*, then the proper M-estimate $\hat{\mu}$ will no longer be advisable since it is not robust toward such deviations from a nominal Gaussian ARMA model (see Martin and Yohai, 1984). More generally, $\hat{\mu}$ is not robust over a full neighborhood of the nominal Gaussian model. An alternative proposal for estimating μ is mentioned in Section VIII of Martin (1981). A detailed study of this alternative, among others, is called for.

Appendix

ASYMPTOTIC VARIANCE EXPRESSIONS

As was mentioned in Section 4, one can obtain closed form expressions for V_{OM} in (3.1) as well as V in (3.4) for the special case where $\varepsilon_t \sim CN(\delta,\sigma^2)$ and either $\psi_1=\psi_\Phi$ or $\psi_1=\psi_{ND}$ (see equations (4.2)-(4.4)). The keys to this are the following relationships:

$$\int_{-\infty}^{\infty} \Phi(\alpha x+\beta)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx = \Phi\left[\frac{\beta}{\sqrt{1+\alpha^2}}\right] \tag{A.1}$$

$$\int_{-\infty}^{\infty} \Phi(\alpha x)\Phi(\beta x)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx = \frac{1}{4}+\frac{1}{2\pi}\tan^{-1}\frac{\alpha\beta}{\sqrt{1+\alpha^2+\beta^2}} \tag{A.2}$$

A.1 was given by Gupta, S.S. and Pillai, K.C.S. (1965), and a proof of A.2 may be found in Jong (1977, Lemma 16).

The Cumulative Normal Psi-Function

Since $\varepsilon_t \sim G = CN(\delta,\sigma^2) = (1-\delta)N(0,1) + \delta N(0,\sigma^2)$, the MA(1) process $\{V_t\}$ has the four-component normal mixture distribution $F = NM(\theta,\delta,\sigma^2)$ $= (1-\delta)^2 N(0,(1+\theta^2)) + \delta(1-\delta)N(0,\sigma^2+\theta^2) + \delta(1-\delta)N(0,1+\theta^2\sigma^2) + \delta^2 N(0,(1+\theta^2)\sigma^2)$. Let $\psi_c(\varepsilon)$ denote ψ_Φ, scaled for the error process ε_t:

$$\psi_c(\varepsilon) = c \cdot s_\varepsilon\, \psi_{CN}\left[\frac{\varepsilon}{c s_\varepsilon}\right] = c s_\varepsilon \sqrt{2\pi}\left[\Phi(\frac{\varepsilon}{c\ s_\varepsilon}) - \frac{1}{2}\right] \tag{A.3}$$

where c is the tuning constant. This we use in computing V. Similarly, in computing V_{OM} we use

$$\psi_{c_{OM}}(y) = c_{OM}\, s_y \sqrt{2\pi}\left[\Phi\left[\frac{y}{c_{OM}\, s_y}\right] - \frac{1}{2}\right] \tag{A.4}$$

which is ψ_{ϕ} scaled for the Y_t process, with tuning constant c_{OM}.

First we get the expression for V with $k = cs_{\varepsilon}$ and $G = CN(\delta,\sigma^2)$, A.2 and A.3:

$$E_G\psi_c^2(\varepsilon) = k^2\left[(1-\delta)\tan^{-1}\frac{(1/k^2)}{\sqrt{1+2/k^2}} + \delta\tan^{-1}\frac{(\sigma^2/k^2)}{\sqrt{1+2\sigma^2/k^2}}\right] . \quad (A.5)$$

Also

$$E_G\psi'_c(\varepsilon) = (1-\delta)\sqrt{k^2/(1+k^2)} + \delta\sqrt{k^2/(\sigma^2+k^2)} . \quad (A.6)$$

Thus

$$V = (1+\theta)^2 E_G\psi^2(\varepsilon)/E_G^2\psi'_c(\varepsilon)$$

$$= \frac{(1+\theta)^2k^2\left[(1-\delta)\tan^{-1}\frac{(1/k^2)}{\sqrt{1+2/k^2}} + \delta\tan^{-1}\frac{(\sigma^2/k^2)}{\sqrt{1+2\sigma^2/k^2}}\right]}{\left[(1-\delta)\sqrt{k^2/(1+k^2)} + \delta\sqrt{k^2/(\sigma^2+k^2)}\right]^2} . \quad (A.7)$$

As for V_{OM}, we need to evaluate $C(0)$ and $C(1)$ using $\psi_{c_{OM}}$. With $k_y = c_{OM}s_y$ and $F = NM(\theta,\delta,\sigma^2)$, A.2 and A.4 give

$$E_F\psi'_{c_{OM}}(Y_1) = (1-\delta)^2\sqrt{k_y^2/(1+\theta^2+k_y^2)} + \delta(1-\delta)\sqrt{k_y^2/(1+\theta^2\sigma^2+k_y^2)}$$

$$+ \delta(1-\delta)\sqrt{k_y^2/(\theta^2+\sigma^2+k_y^2)}$$

$$+ \delta^2\sqrt{k_y^2/[(1+\theta^2)\sigma^2+k_y^2]} . \quad (A.8)$$

$$C(0) = E_F\psi_{c_{OM}}^2(Y_1) = k_y^2\left\{(1-\delta)^2\tan^{-1}\left[(1+\theta^2)\cdot k_y^{-2}/\sqrt{1+2\theta^2k_y^{-2}}\right]\right.$$

$$+ \delta(1-\delta)\tan^{-1}\left[(1+\theta^2+\sigma^2)\cdot k_y^{-2}/\sqrt{1+2(1+\theta^2+\sigma^2)k_y^{-2}}\right]$$

$$+ \delta(1-\delta)\tan^{-1}\left[(1+\theta^2\sigma^2)\cdot k_y^{-2}/\sqrt{1+2(1+\theta^2\sigma^2)k_y^{-2}}\right]$$

$$\left. + \delta^2\tan^{-1}\left[(1+\theta^2)\sigma^2k_y^{-2}/\sqrt{1+2(1+\theta^2)\sigma^2k_y^{-2}}\right]\right\} . \quad (A.9)$$

Now for $C(1)$, first note that

$$\psi_{c_{OM}}(Y_t)\psi_{c_{OM}}(Y_{t+1}) = 2\pi k_y^2\left[\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1})\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t)\right.$$

$$- \frac{1}{2}\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1}) - \frac{1}{2}\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t) + \frac{1}{4}\Big] \quad . \tag{A.10}$$

Since the ε_t are i.i.d. with distribution $CN(\delta,\sigma^2)$, we can condition on ε_t and apply (A.1) to get

$$E[\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1}) \mid \varepsilon_t]$$

$$= (1-\delta)\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2 k_y^{-2}}\right] + \delta\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2\sigma^2 k_y^{-2}}\right] \quad . \tag{A.11}$$

Similarly

$$E[\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t) \mid \varepsilon_t]$$

$$= (1-\delta)\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+k_y^{-2}}\right] + \delta\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+\sigma^2 k_y^{-2}}\right] \quad . \tag{A.12}$$

Taking expectation with respect to ε_t in (A.11) and (A.12), and using A.1 with $\beta=0$, gives

$$E_G\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1}) = \Phi(0) = \frac{1}{2} \tag{A.13}$$

$$E_G\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t) = \Phi(0) = \frac{1}{2} \quad . \tag{A.14}$$

For the expectation of the first term on the right-hand side of (A.10), we again use the results in (A.11) and (A.12)

$$E[\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1})\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t)]$$

$$= (1-\delta)^2 E_G\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2 k_y^{-2}}\right]\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+k_y^{-2}}\right]$$

$$+ \delta(1-\delta)E_G\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2 k_y^{-2}}\right]\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+\sigma^2 k_y^{-2}}\right]$$

$$+ \delta(1-\delta)E_G\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2\sigma^2 k_y^{-2}}\right]\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+k_y^{-2}}\right]$$

$$+ \delta^2 E_G\Phi\left[k_y^{-1}\varepsilon_t / \sqrt{1+\theta^2\sigma^2 k_y^{-2}}\right]\Phi\left[\theta k_y^{-1}\varepsilon_t / \sqrt{1+\sigma^2 k_y^{-2}}\right]$$

$$= (1-\delta)^2 A_1 + \delta(1-\delta)[A_2 + A_3] + \delta^2 A_4 \quad . \tag{A.15}$$

The expectations $A_1 - A_4$ in (A.15) can be obtained by applying (A.2) with constants appropriately adjusted:

$$A_1 = \frac{1}{4} + \frac{(1-\delta)}{2\pi}\tan^{-1}\left\{\theta k_y^{-1}\left[k_y^2(1+\theta^2 k_y^{-2})(1+k_y^{-2})+(1+k_y^{-2})+\theta^2(1+\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$+ \frac{\delta}{2\pi}\tan^{-1}\left\{\theta\sigma^2 k_y^{-1}\left[k_y^2(1+\theta^2 k_y^{-2})(1+k_y^{-2})+\sigma^2(1+k_y^{-2})+\theta^2\sigma^2(1+\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$= \frac{1}{4} + \frac{(1-\delta)}{2\pi} A_{11} + \frac{\delta}{2\pi} A_{12} \quad ; \qquad \text{(A.16)}$$

$$A_2 = \frac{1}{4} + \frac{(1-\delta)}{2\pi}\tan^{-1}\left\{\theta k_y^{-1}\left[k_y^2(1+\theta^2 k_y^{-2})(1+\sigma^2 k_y^{-2})+(1+\sigma^2 k_y^{-2})+\theta^2(1+\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$+ \frac{\delta}{2\pi}\tan^{-1}\left\{\theta\sigma^2 k_y^{-1}\left[k_y^2(1+\theta^2 k_y^{-2})(1+\sigma^2 k_y^{-2})+\sigma^2(1+\sigma^2 k_y^{-2})+\theta^2\sigma^2(1+\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$= \frac{1}{4} + \frac{(1-\delta)}{2\pi} A_{21} + \frac{\delta}{2\pi} A_{22} \quad ; \qquad \text{(A.17)}$$

$$A_3 = \frac{1}{4} + \frac{(1-\delta)}{2\pi}\tan^{-1}\left\{\theta k_y^{-1}\left[k_y^2(1+\sigma^2\theta^2 k_y^{-2})(1+k_y^{-2})+(1+k_y^{-2})+\theta^2(1+\sigma^2\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$+ \frac{\delta}{2\pi}\tan^{-1}\left\{\theta\sigma^2 k_y^{-1}\left[k_y^2(1+\sigma^2\theta^2 k_y^{-2})(1+k_y^{-2})+\sigma^2(1+k_y^{-2})+\sigma^2\theta^2(1+\sigma^2\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$= \frac{1}{4} + \frac{(1-\delta)}{2\pi} A_{31} + \frac{\delta}{2\pi} A_{32} \quad ; \qquad \text{(A.18)}$$

and finally,

$$A_4 = \frac{1}{4} + \frac{(1-\delta)}{2\pi}\tan^{-1}\left\{\theta k_y^{-1}\left[k_y^2(1+\sigma^2\theta^2 k_y^{-2})(1+\sigma^2 k_y^{-2})+(1+\sigma^2 k_y^{-2})+\theta^2(1+\sigma^2\theta^2 k_y^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$+ \frac{\delta}{2\pi} \tan^{-1}\left\{\theta\sigma^2 k_y^{-1}\left[k_y^2(1+\sigma^2\theta^2 k_y^{-2})(1+\sigma^2 k_y^{-2})+\sigma^2(1+\sigma^2 k_y^{-2})+\sigma^2\theta^2(1+\sigma^2\theta^2 k_v^{-2})\right]^{-\frac{1}{2}}\right\}$$

$$= \frac{1}{4} + \frac{(1-\delta)}{2\pi} A_{41} + \frac{\delta}{2\pi} A_{42} \quad . \tag{A.19}$$

Now applying (A.15) - (A.19), we have

$$C(1) = E_F[\psi_{c_{OM}}(Y_t)\psi_{c_{OM}}(Y_{t+1})] = k_y^2\Big[(1-\delta)^3 A_{11} + \delta(1-\delta)^2(A_{12}+A_{21}+A_{31})$$

$$+ \delta^2(1-\delta)(A_{22}+A_{32}+A_{41}) + \delta^3 A_{42}\Big] \quad . \tag{A.20}$$

Therefore, (A.8), (A.9) and (A.20) can be combined to get the closed form for V_{OM}:

$$V_{OM} = \frac{C(0)+2C(1)}{E_F^2 \psi'_{c_{OM}}(Y_1)} \quad . \tag{A.21}$$

Normal Derivative Psi-Function

Let ψ_c denote ψ_{ND} scaled for ε_t, with tuning constant c:

$$\psi_c(\varepsilon) = cs_\varepsilon \psi_{ND}(\varepsilon / cs_\varepsilon) = \varepsilon \exp[-\varepsilon^2/2c^2 s_\varepsilon^2] \quad . \tag{A.22}$$

Similarly, let $\psi_{c_{OM}}$ denote ψ_{ND} scaled for Y_t, with tuning constant c_{OM}:

$$\psi_{c_{OM}}(y) = y \exp[-y^2/2c_{OM}^2 s_y^2] \quad . \tag{A.23}$$

First we obtain the expression for V, with $k = cs_\varepsilon$ and $G = CN(\delta, \sigma^2)$. Direct evaluation gives

$$E_G \psi_c^2(\varepsilon) = (1-\delta)k^3/(2+k^2)^{\frac{3}{2}} + \delta\sigma^2 k^3/(2\sigma^2+k^2)^{\frac{3}{2}} \tag{A.24}$$

and

$$E_G \psi'_c(\varepsilon) = (1-\delta)\frac{k^3}{(1+k^2)^{\frac{3}{2}}} + \delta\frac{k^3}{(\sigma^2+k^2)^{\frac{3}{2}}} \quad . \tag{A.25}$$

Now $V = E_G\psi_c^2(\varepsilon)/E_G^2\psi'_c(\varepsilon)$ may be computed from A.24 and A.25.

Next we evaluate $E_F\psi'(Y_1)$, $C(0)$ and $C(1)$, with $F = NM(\theta,\delta,\sigma^2)$ and $k_y = c_{OM}s_y$, in order to compute V_{OM}. First, we have

$$E_F\psi'_{c_{OM}}(Y_1) = (1-\delta)^2 k_y^3/(1+\theta^2+k_y^2)^{\frac{3}{2}} + \delta(1-\delta)k_y^3/(\sigma^2+\theta^2+k_y^2)^{\frac{3}{2}}$$

$$+ \delta(1-\delta)k_y^3/(1+\theta^2\sigma^2+k_y^2)^{\frac{3}{2}}$$

$$+ \delta^2 k_y^3/(\sigma^2+\theta^2\sigma^2+k_y^2)^{\frac{3}{2}} \quad . \qquad \text{(A.26)}$$

As for $C(0)$:

$$C(0) = E_F\psi^2_{c_{OM}}(Y_1) = (1-\delta)^2(1+\theta^2)k_y^3/[2(1+\theta^2)+k_y^2]^{\frac{3}{2}}$$

$$+ \delta(1-\delta)(\sigma^2+\theta^2)k_y^3/[2(\sigma^2+\theta^2)+k_y^2]^{\frac{3}{2}}$$

$$+ \delta(1-\delta)(1+\theta^2\sigma^2)k_y^3/[2(1+\theta^2\sigma^2)+k_y^2]^{\frac{3}{2}}$$

$$+ \delta^2(1+\theta^2)\sigma^2 k_y^3/[2\sigma^2(1+\theta^2)+k_y^2]^{\frac{3}{2}} \quad . \qquad \text{(A.27)}$$

As for $C(1)$, consider first the expectation conditioned on ε_t:

$$E[\psi(y_{t+1})\psi(y_t)\,|\,\varepsilon_t] = E_F\left\{[\varepsilon_{t+1}+\theta\varepsilon_t]exp[-(\varepsilon_{t+1}+\theta\varepsilon_t)^2/2k_y^2]\Big|\varepsilon_t\right\}$$

$$\cdot E_F\left\{[\varepsilon_t+\theta\varepsilon_{t-1}]\exp[-(\varepsilon_t+\theta\varepsilon_{t-1})^2/2k_y^2]\Big|\varepsilon_t\right\}$$

$$= K_1(\varepsilon_t)\cdot K_2(\varepsilon_t) \qquad \text{(A.28)}$$

where

$$K_1(\varepsilon_t) = \frac{(1-\delta)\theta k_y^3}{(1+k_y^2)^{\frac{3}{2}}}\varepsilon_t \cdot \exp[-\theta^2\varepsilon_t^2/2(1+k_y^2)]$$

$$+ \frac{\delta \theta k_y^3}{\sigma^2+k_y^2)^{\frac{3}{2}}} \varepsilon_t \cdot \exp[-\theta^2 \varepsilon_t^2 / 2(\sigma^2+k_y^2)]$$

and

$$K_2(\varepsilon_t) = \frac{(1-\delta)k_y^3}{(\theta^2+k_y^2)^{\frac{3}{2}}} \varepsilon_t \cdot \exp[-\varepsilon_t^2 / 2(\theta^2+k_y^2)]$$

$$+ \frac{\delta k_y^3}{(\theta^2\sigma^2+k_y^2)^{\frac{3}{2}}} \varepsilon_t \exp[-\varepsilon_t^2 / 2(\theta^2\sigma^2+k_y^2)] \quad .$$

Therefore,

$$C(1) = E_G[K_1(\varepsilon_t)K_2(\varepsilon_t)]$$

$$= \theta k_y^6 \Big\{ (1-\delta)^3 / [\theta^2(\theta^2+k_y^2)+(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)^2\delta / [\theta^2(\theta^2\sigma^2+k_y^2)+(1+k_y^2)+(1+k_y^2)(\theta^2\sigma^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)^2\delta / [\theta^2(\theta^2+k_y^2)+(\sigma^2+k_y^2)+(\sigma^2+k_y^2)(\theta^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)\delta^2 / [\theta^2(\theta^2\sigma^2+k_y^2)+(\sigma^2+k_y^2)+(\sigma^2+k_y^2)(\theta^2\sigma^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)^2\delta\sigma^2 / [\theta^2\sigma^2(\theta^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)\delta^2\sigma^2 / [\theta^2\sigma^2(\theta^2\sigma^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2\sigma^2+k_y^2)]^{\frac{3}{2}}$$

$$+ (1-\delta)\delta^2\sigma^2 / [\theta^2\sigma^2(\theta^2+k_y^2)+\sigma^2(\sigma^2k_y^2)+(\sigma^2+k_y^2)(\theta^2+k_y^2)]^{\frac{3}{2}}$$

$$+ \delta^3\sigma^2 / [\theta^2\sigma^2(\theta^2\sigma^2+k_y^2)+\sigma^2(\sigma^2+k_y^2)+(\sigma^2+k_y^2)(\theta^2\sigma^2+k_y^2)]^{\frac{3}{2}} \Big\} \quad .$$

(A.29)

Now (A.26), (A.27) and (A.29) can be combined to obtain the closed form for V_{0M}.

References

Boente, G., Fraiman, R. and Yohai, V. (1982). "Qualitative robustness for general stochastic processes," Tech Rep. 26, Dept. of Statistics, Univ. of Washington, Seattle, WA.

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

Cox, D.R. (1966). "The null distribution of the first serial correlation coefficient," *Biometrika* **53**, 623-626.

Cox, D. (1981). "Metrics on stochastic processes and qualitative robustness," Tech. Rep. 3, Dept. of Statistics, Univ. of Washington, Seattle, WA.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications, Volume II*, John Wiley, New York.

Grenander, U. (1954). "On the estimation of regression coefficients in the case of an autocorrelated disturbance," *Annals Math. Stat.* **25**, 252-272.

Grenander, U. (1981). *Abstract Inference*, John Wiley, New York.

Gupta, S.S. and Pillai, K.C.S. (1965). "On linear functions of ordered correlated normal random variables," *Biometrika* **52**, 367-379.

Hampel, F.R. (1971). "A general qualitative definition of robustness," *Ann. Math. Stat.* **42**, 1887-1895.

Huber, P.J. (1964). "Robust estimation of a location parameter," *Annals of Math. Stat.* **35**, 73-101.

Huber, P.J. (1981). *Robust Statistics*, Wiley, 1981.

Jong, J. (1977). "Robust generalized M-estimates for the autoregressive parameter," Ph.D. dissertation, University of Washington, Seattle, WA.

Lee, C.H. and Martin, R.D. (1982a). "M-estimates for ARMA processes," Tech. Rep. No. 23, Dept. of Statistics, Univ. of Washington, Seattle, WA.

Lee, C.H. and Martin, R.D. (1982b). "The information matrix and robust M-estimates for ARMA models," Tech. Rep. No. 24, Dept. of Statistics, Univ. of Washington, Seattle, WA.

Martin, R.D. (1981). "Robust methods for time series," in *Applied Time Series II*, edited by D.F. Findley, Academic Press, New York.

Martin, R.D. (1982). "The Cramer-Rao lower bound and robust M-estimates for autoregressions," *Biometrika* **69**, 437-442.

Martin, R.D. and Yohai, V.J. (1984). "Robustness in time series and estimating ARMA models," in *Handbook of Statistics, Vol. 4, Time Series: The Time*

Domain," edited by Brillinger and Krishnaiah.

Mosteller, F. and Tukey, J.W. (1977). *Data Analysis and Regression*, Addison Wesley, Reading, MA.

Papantoni-Kazakos, P. and Gray, R.M. (1979). "Robustness of estimators on stationary observations," *Ann. Probability* **7**, 6:989-1002.

Portnoy, S.L. (1977). "Robust estimation in dependent situations," *Ann. Statist.* **5**, 522-529.

Wegman, E.J. and Carroll, R.J. (1977). "A Monte Carlo study of robust estimators of location," *Comm. in Stat.–Theory and Methods* **A6**, 795-812.

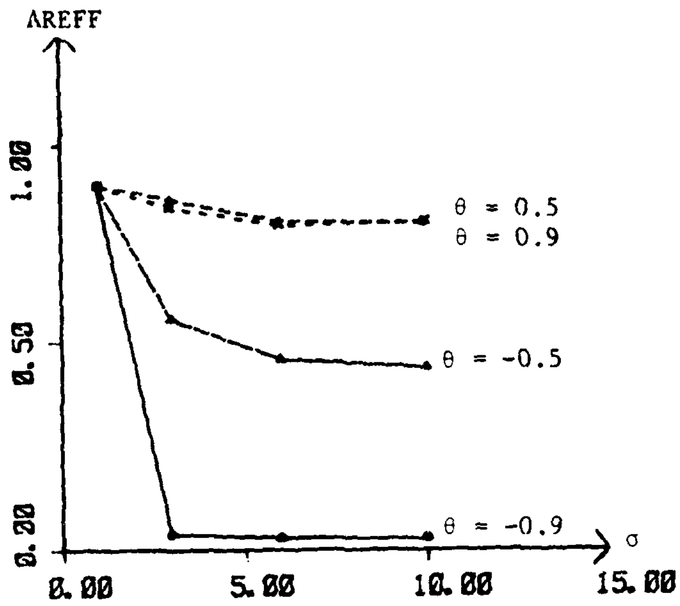


Figure 1. AREFF versus σ for MA(1) model using ψ_{ND} .

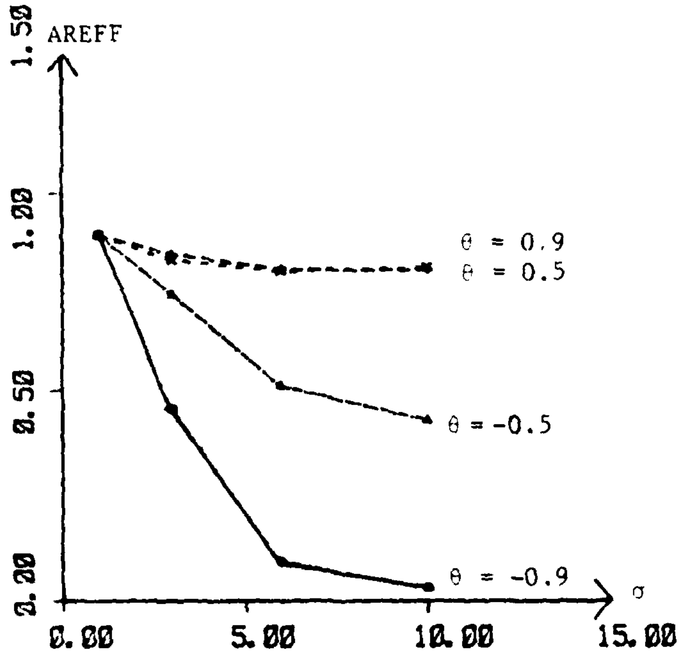


Figure 2. AREFF versus σ for MA(1) model using ψ_{ND} and $c_{OM} = c_{OM}(\theta)$.

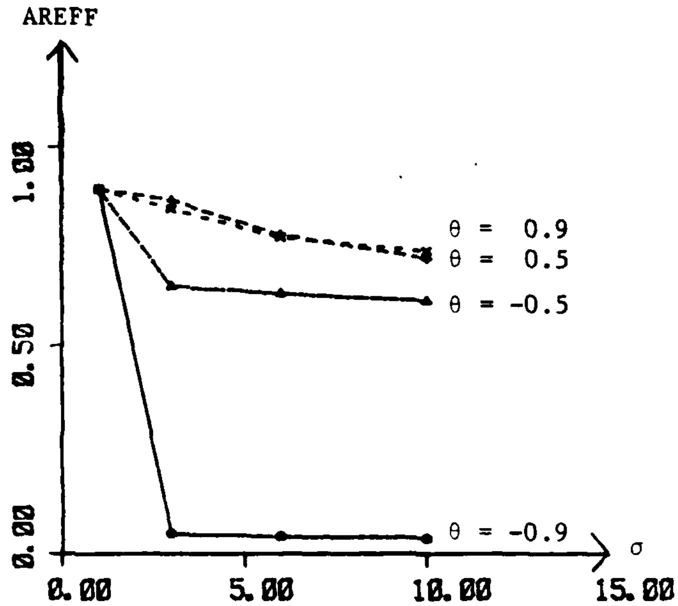


Figure 1a. AREFF versus σ for MA(1) model using ψ_ϕ .

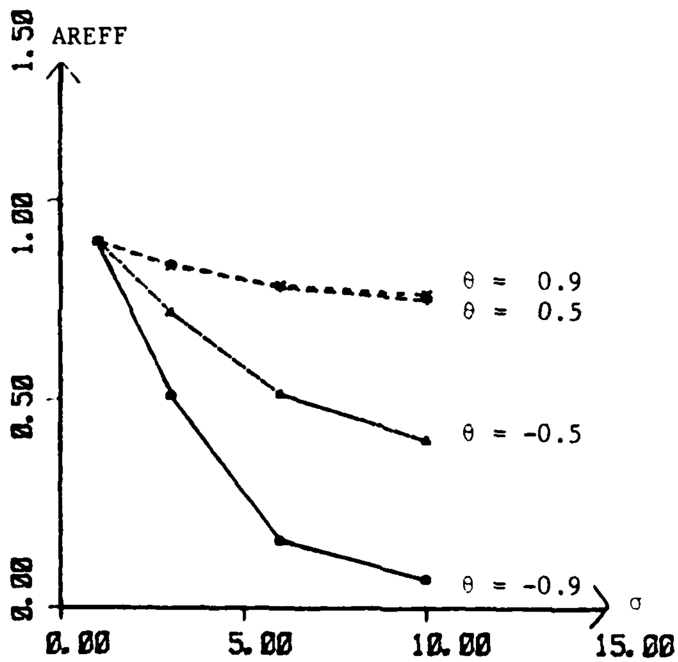


Figure 2a. AREFF versus σ for MA(1) model using ψ_ϕ and $c_{OM} = c_{OM}(\theta)$.

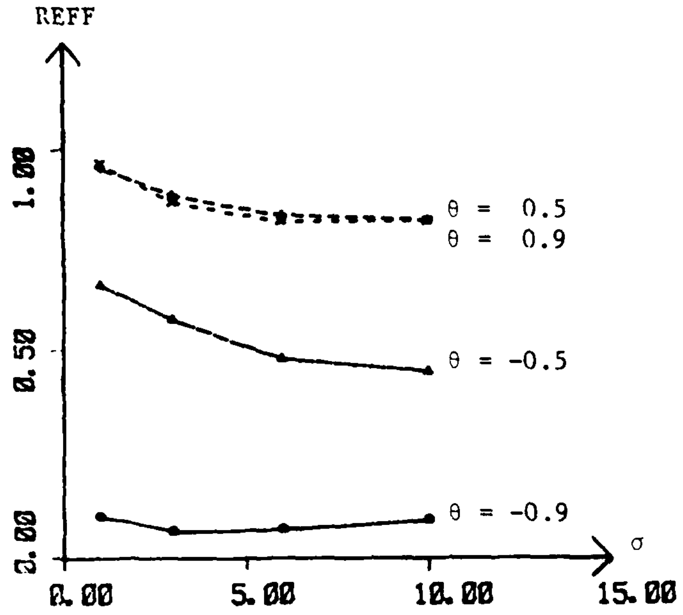


Figure 3. REFF versus σ for MA(1) model using ψ_{ND} .

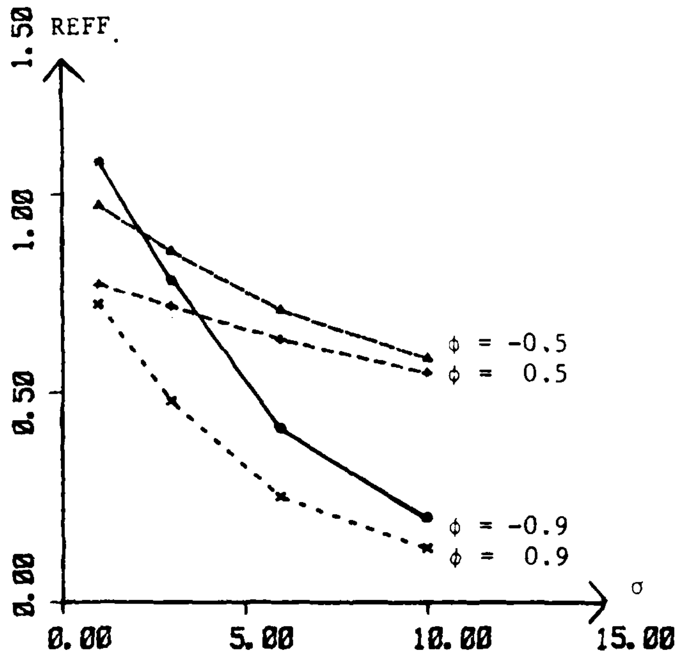


Figure 4. REFF versus σ for AR(1) model using ψ_{ND} .

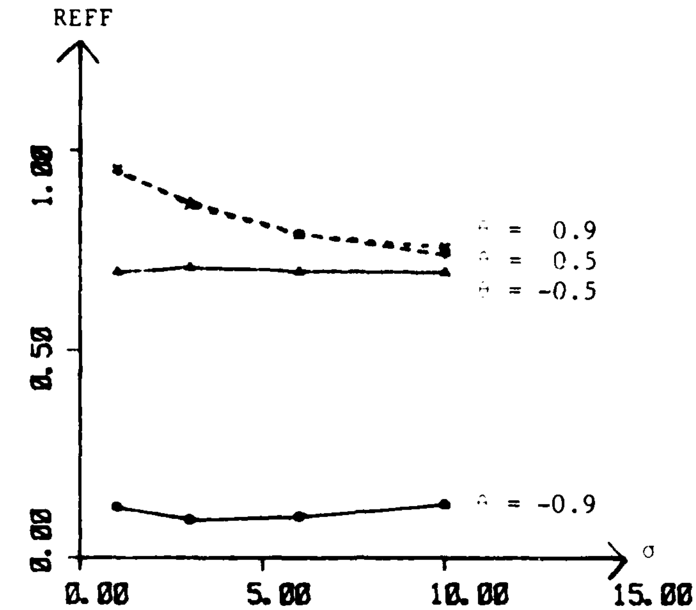


Figure 3a. REFF versus σ for MA(1) model using ψ_θ .

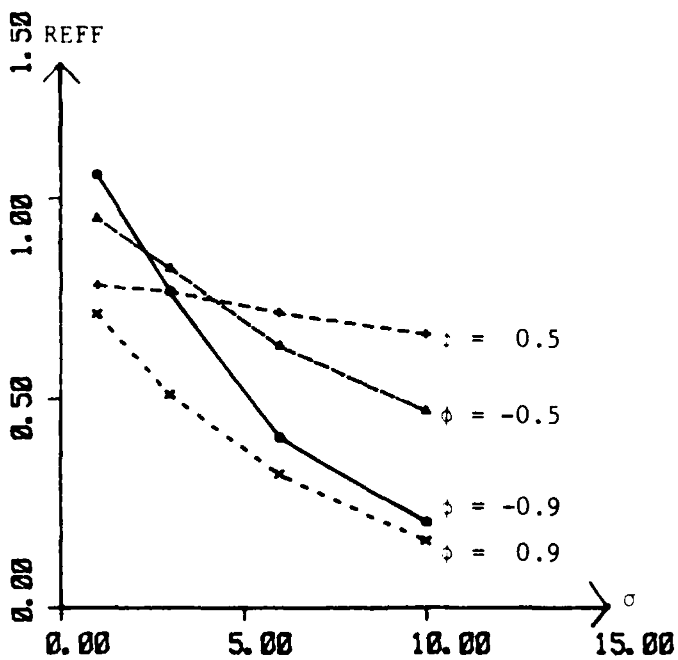


Figure 4a. REFF versus σ for AR(1) model using ψ_ϕ .

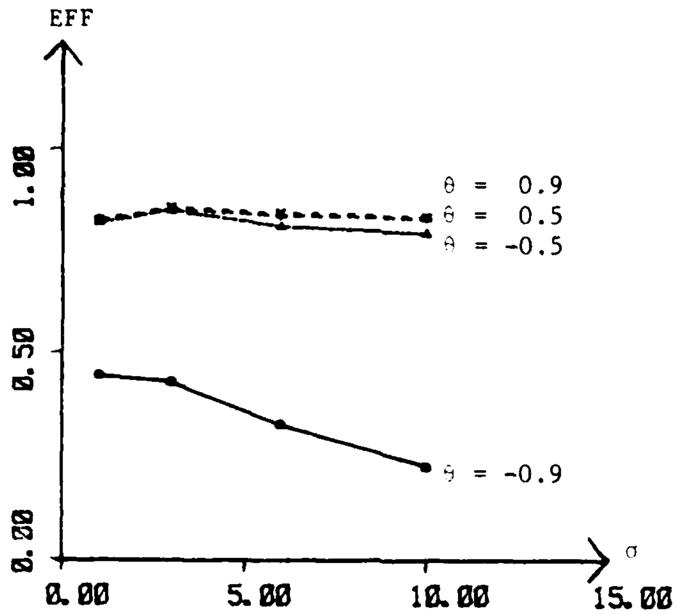


Figure 5. EFF versus σ for MA(1) model using ψ_{ND} .

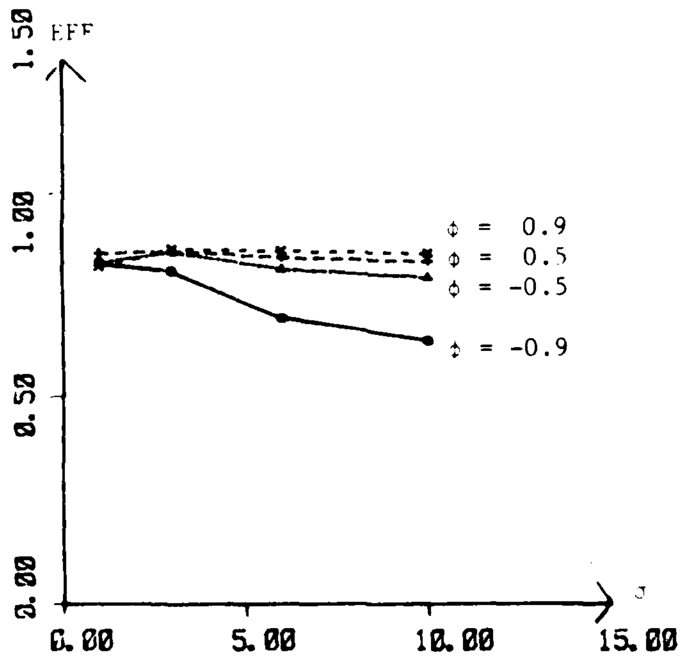


Figure 6. EFF versus σ for AR(1) model using ψ_{ND} .

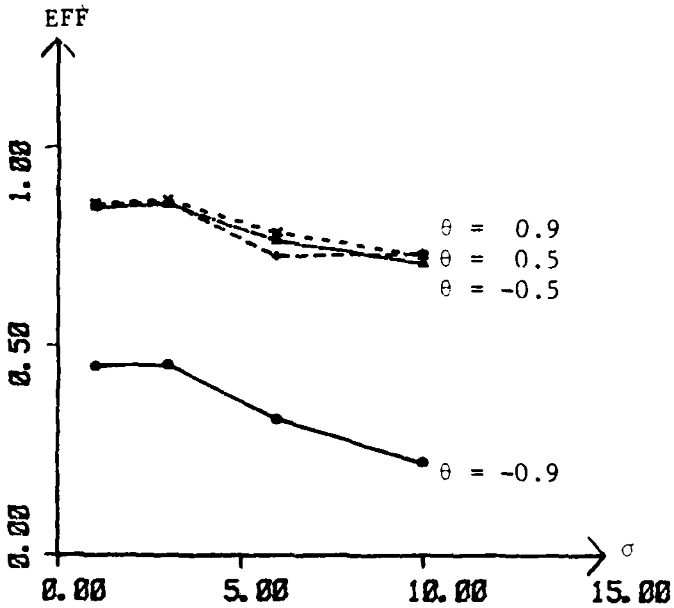


Figure 5a. EFF versus σ for MA(1) model using ψ_{ϕ} .

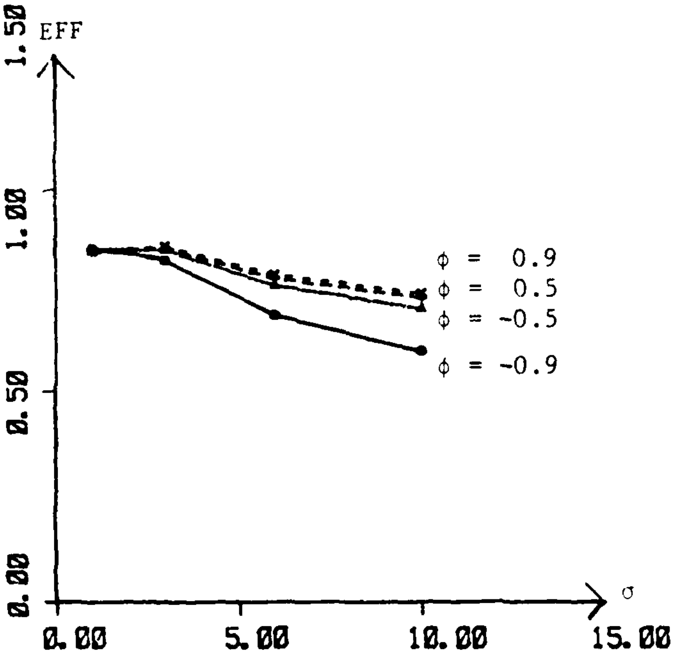


Figure 6a. EFF versus σ for AR(1) model using ψ_{ϕ} .

END

FILMED

2-85

DTIC