MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS TR 208
Robustness in time series and estimating ARMA models

Chin-Hui Lee and R. Douglas Martin

Department of Statistics, GN-22
University of Washington
Seattle, WA 98195

ONR Code N63374
1107 N.E. 45th Street
Seattle, WA 98105

7/84

35

Approved for public release: distribution unlimited.

M-estimates, ARMA models

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 (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.
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 \( \mu \) 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

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

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 \( \psi \) is a robustifying psi-function. With \( \psi = \rho' \), this estimating equation characterizes a stationary point of the minimization problem

\[
\min_{\mu} \sum_{i=1}^{n} \rho \left( \frac{Y_i - \mu}{c \cdot \hat{s}_y} \right)
\]

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 comparisons 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 \( \mu \) is a location parameter and that the observations are

\[
Y_t = \mu + \epsilon_t, \quad t = 1, 2, \ldots, n
\]  

(2.1)

where \( \epsilon_t \) is an ARMA\((p,q)\) model

\[
\epsilon_t = \phi_1 \epsilon_{t-1} + \cdots + \phi_p \epsilon_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}
\]

(2.2)

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

\[
Y_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} = \gamma + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}
\]

(2.3)

where the intercept is

\[
\gamma = \mu(1 + \Sigma \phi_t)
\]

(2.4)

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

\[
\min_{\alpha} \sum_{t=1}^{n} \rho \left[ \frac{\tau_t(\alpha')}{c \cdot s_\epsilon} \right]
\]

(2.5)

where \( \rho \) is a robustifying loss function. The constant \( c \) is a tuning constant and \( s_\epsilon \) is a robust estimate of the innovations scale \( s_\epsilon \).

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

\[
\hat{\mu} = \frac{\gamma}{1 + \Sigma \hat{\phi}_t}
\]

(2.6)
Consistency and asymptotic normality of $\hat{\alpha}$ and $\hat{\mu}$ have been established by Lee and Martin (1982a).

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

First consider an ordinary location $M$-estimate $\mu_{OM}$ computed from observations $Y_1, \ldots, Y_n$ in (2.1) which have a common marginal distribution $F(y) = F_0((y-\mu)/s_Y)$. Under regularity conditions (see for example Portnoy, 1977) $\mu_{OM}$ is consistent and asymptotically normal, with asymptotic variance given by

$$
V_{OM} = \frac{C(0)+2\sum_{i=1}^{\infty} C(i)}{E_F \psi_c(Y_1)}
$$

(3.1)

where

$$
C(i) = s_Y^2 E_{F_0} \psi_c(Y_1 \psi_c(Y_{1+i})).
$$

(3.2)

Here for $l = 0$, $F_0$ is the standardized marginal distribution $F_0$ of the $Y_t$, while for $l \geq 1$ $F_0$ is the bivariate distribution for $(Y_1, Y_{1+l})$ 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 $\psi_c$. In the special case of independent $Y_t$, $F_0 = G_0$ and $V_{OM}$ reduces to

$$
V_{OM} = s_Y^2 \frac{E_F \psi_c^2(Y_1)}{E_F \psi_c(Y_1)} = s_Y^2 V_{loc}(\psi_c,F_0)
$$

(3.3)

where $V_{loc} = V_{loc}(\psi,F_0)$, 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

$$
V = \frac{(1+\Sigma \theta_1)^2}{(1+\Sigma \varphi_1)^2} s_\varepsilon^2 V_{loc}(\psi_c,G_0)
$$

(3.4)

The quantity $s_\varepsilon^2(1+\Sigma \theta_1)^2/(1+\Sigma \varphi_1)^2$ differs by only a constant factor from the value at zero frequency of the spectrum of the process $Y_t$. When $\psi$ is the
identity function so that \( \tilde{\mu}_{GM} = \tilde{\mu}_{LS} = \bar{Y} \), and \( s_e \) 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 \( V_{OM} \) is quite attractive, particularly with regard to the relative ease of studentizing the estimate \( \tilde{\mu} \) for the purpose of constructing confidence intervals. Estimation of \( V \) from the data for this purpose may be quite manageable, whereas estimation of \( V_{OM} \) seems rather impractical when many \( C(i) \) are non-zero. In this regard the situation is particularly bad when an autoregression component is present, since then the \( C(i) \) only vanish asymptotically.

Furthermore, the effect of the tuning constant \( c \) on the asymptotic efficiency of \( \tilde{\mu} \) 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 \( \varphi_i \) and \( \theta_i \), efficiencies can be controlled through \( c \) without regard to the values of these parameters. This is not the case with regard to \( V_{OM} \), as can be seen in the following equivalent form of (3.1):

\[
V_{OM} = \left[ 1 + 2 \sum_{l=1}^{\infty} \rho_{1,l+1} \right] s_y^2 V_{loc}(\psi_c, F_0)
\]  

(3.5)

where \( \rho_{1,l+1} \) is the correlation coefficient for the random variables \( \psi_c(Y_{1}) \) and \( \psi_c(Y_{1+l}) \) when \( (Y_{1}, Y_{1+l}) \sim F_0 \). Here the effects of \( c \) appear not only in \( V_{loc} \), but also in the correlation coefficients \( \rho_{1,l+1} \), and the latter depend on the ARMA model parameters \( \varphi_i \) and \( \theta_i \). 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 \( \tilde{\mu}_{OM} \) 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 \( \varphi \). In this case \( V \) may be expressed in the following form, which facilitates comparison with (3.5):

\[
V = \left[ 1 - 2 \frac{\varphi}{1 + \varphi} \right] \frac{s_\epsilon^2}{1 - \varphi^2} V_{loc}(\psi, G_o)
\]

(3.6)

It is easy to check that the factors in square brackets in (3.5) and (3.6) are identical when \( \psi \) 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 \( \psi \) 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 \( V_{loc}(\psi, F_o) \), \( V_{loc}(\psi, G_o) \), \( s_{\psi}^2 \), and \( s_\epsilon^2/(1 - \varphi^2) \).

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 \( V_{loc}(\psi, F_o) \) will be larger than \( V_{loc}(\psi, G_o) \) when \( \varphi \neq 0 \). This is because \( Y_t \) is a weighted sum of the \( \epsilon_t \), and the convolutions which produce \( F_o \) from non-Gaussian \( C_o \) will often result in an \( F_o \) having heavier tails than \( C_o \). At the same time \( s_{\psi}^2 \) and \( s_\epsilon^2/(1 - \varphi^2) \) 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 \( F_o = C_o \), and then the two \( V_{loc} \)'s will be identical. However, in such a case \( s_{\psi}^2 \) and \( s_\epsilon^2/(1 - \varphi^2) \) will no longer be identical (except in the Gaussian case). For example, when \( G_o \) is a symmetric stable distribution with index \( \eta \), \( F_o \) is also a symmetric stable distribution, and it is easy
to check that (see Feller, 1966)

\[ R = \frac{\frac{\tilde{y}^2}{\varphi^2}}{s_x^2/(1-\varphi^2)} = \frac{1-\varphi^2}{(1-|\varphi|^\eta)^{2/\eta}} \tag{3.7} \]

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 \( V_{GW} \) 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 \( \tilde{\mu}_{GW} \).
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 \( \psi_v \), 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 
\[
\int \left( \frac{g'_o}{g_o} \right)^2 g_o \ 	ext{for the standardized innovations density } g_o \ (\text{Martin, 1982}).
\]

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

\[
AREFF = \frac{V_{OM}(\psi_v, C_o, \alpha)}{V(\psi_v, C_o, \alpha)}
\]

(4.1)

for various \( \psi_v, C_o \) and \( \alpha \).

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 \( C_o \) is non-Gaussian (symmetric stable \( C_o \) 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 \( E_{(0)} \psi_v \).

We treat here the \( MA(1) \) case with parameter \( \theta \), where (i) \( \varepsilon_1 \) has a contaminated normal distribution.
\[ CN(\gamma, \sigma^2) = (1-\delta)N(0,1) + \delta N(0,\sigma^2) \quad (4.2) \]

and (ii) \( \psi \) has either the normal distribution shape

\[ \psi(t) = \sqrt{2\pi} (\Phi(t) - \frac{1}{2}) \quad (4.3) \]

or the shape of the derivative of the normal density,

\[ \psi_{ND}(t) = t \cdot e^{-t^2/2} \quad (4.4) \]

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 \( \psi \) and \( \psi_{ND} \) 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 \( \psi_{ND} \) and \( \psi \), \( \hat{\mu}_{OM} \) and \( \hat{\mu} \) have matched asymptotic efficiencies of .90 for independent Gaussian observations (\( \theta = 0 \)).

Figure 1 shows AREFF's based on \( \psi_{ND} \) for various \( \theta \) values, where \( \epsilon_t \sim CN(\delta, \sigma^2) \) with \( \delta = 0.1, 1.5 \sigma \leq 10 \). Although the AREFF's can be quite low for negative \( \theta \), they are quite high for a wide range of positive \( \theta \).

In Figure 2 we display AREFF's based on \( \psi_{ND} \) for the same values of \( \gamma, \sigma^2 \) and \( \theta \), except that \( c_{OM} \) has been adjusted to obtain matching asymptotic efficiencies of .90 for each value of \( \theta \) and Gaussian \( \epsilon_t \). The values of tuning constants \( c_{OM} = c_{OM}(\theta) \) needed to achieve various efficiencies are given in Table 1 for \( \psi \) and in Table 2 for \( \psi_{ND} \). While marked improvement in the relative performance
of $\hat{\mu}_{CM}$ is achieved at $\theta = -0.5$ and -0.9 at small values of $\sigma$, the improvement at large values of $\sigma$ is negligible. Thus even "proper" adjustment of $c$ using typically unavailable prior information on $\theta$ will not salvage $\hat{\mu}_{CM}$ for MA(1) models with negative $\theta$.

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

TABLE 1

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

which yield various efficiencies for $\psi_{\phi}$

<table>
<thead>
<tr>
<th>$\theta$ \ EFF</th>
<th>0.95</th>
<th>0.90</th>
<th>0.85</th>
<th>0.80</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.9</td>
<td>4.450</td>
<td>3.651</td>
<td>3.222</td>
<td>2.927</td>
</tr>
<tr>
<td>-0.7</td>
<td>2.343</td>
<td>1.870</td>
<td>1.611</td>
<td>1.431</td>
</tr>
<tr>
<td>-0.5</td>
<td>1.669</td>
<td>1.287</td>
<td>1.074</td>
<td>0.923</td>
</tr>
<tr>
<td>-0.3</td>
<td>1.299</td>
<td>0.959</td>
<td>0.765</td>
<td>0.625</td>
</tr>
<tr>
<td>-0.1</td>
<td>1.049</td>
<td>0.731</td>
<td>0.546</td>
<td>0.409</td>
</tr>
<tr>
<td>0.0</td>
<td>0.952</td>
<td>0.642</td>
<td>0.460</td>
<td>0.324</td>
</tr>
<tr>
<td>0.1</td>
<td>0.876</td>
<td>0.571</td>
<td>0.390</td>
<td>0.255</td>
</tr>
<tr>
<td>0.3</td>
<td>0.781</td>
<td>0.480</td>
<td>0.298</td>
<td>0.161</td>
</tr>
<tr>
<td>0.5</td>
<td>0.747</td>
<td>0.443</td>
<td>0.257</td>
<td>0.115</td>
</tr>
<tr>
<td>0.7</td>
<td>0.741</td>
<td>0.433</td>
<td>0.243</td>
<td>0.097</td>
</tr>
<tr>
<td>0.9</td>
<td>0.741</td>
<td>0.431</td>
<td>0.239</td>
<td>0.092</td>
</tr>
</tbody>
</table>
TABLE 2

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

which yield various efficiencies for $\psi_{HD}$

<table>
<thead>
<tr>
<th>$\theta$ \ EFF</th>
<th>0.95</th>
<th>0.90</th>
<th>0.85</th>
<th>0.80</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.9</td>
<td>7.839</td>
<td>6.478</td>
<td>5.752</td>
<td>5.258</td>
</tr>
<tr>
<td>-0.7</td>
<td>4.287</td>
<td>3.517</td>
<td>3.104</td>
<td>2.821</td>
</tr>
<tr>
<td>-0.5</td>
<td>3.195</td>
<td>2.600</td>
<td>2.281</td>
<td>2.061</td>
</tr>
<tr>
<td>-0.3</td>
<td>2.622</td>
<td>2.118</td>
<td>1.846</td>
<td>1.659</td>
</tr>
<tr>
<td>-0.1</td>
<td>2.249</td>
<td>1.802</td>
<td>1.561</td>
<td>1.397</td>
</tr>
<tr>
<td>0.0</td>
<td>2.110</td>
<td>1.685</td>
<td>1.456</td>
<td>1.300</td>
</tr>
<tr>
<td>0.1</td>
<td>2.002</td>
<td>1.594</td>
<td>1.374</td>
<td>1.225</td>
</tr>
<tr>
<td>0.3</td>
<td>1.874</td>
<td>1.485</td>
<td>1.276</td>
<td>1.134</td>
</tr>
<tr>
<td>0.5</td>
<td>1.832</td>
<td>1.448</td>
<td>1.242</td>
<td>1.101</td>
</tr>
<tr>
<td>0.7</td>
<td>1.827</td>
<td>1.443</td>
<td>1.235</td>
<td>1.094</td>
</tr>
<tr>
<td>0.9</td>
<td>1.829</td>
<td>1.443</td>
<td>1.235</td>
<td>1.094</td>
</tr>
</tbody>
</table>
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 $\psi_{\phi}$, followed by one iteration using $\psi_{ND}$. 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 $\psi_{ND}$ 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 $\psi_{ND}$ are given in Figure 4. Here REFF's can be quite low for positive $\varphi$ 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 $\sigma$, 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 $\varphi$, which contrasts sharply with the MA(1) results of Figure 3.

Figures 3a and 4a give corresponding results using $\psi_{\phi}$. Again, $\psi_{ND}$ tends to dominate $\psi_{\phi}$ 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 $\psi_{ND}$, and Figures 5a and 5b for $\psi_{\phi}$. 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 \( \hat{\mu}_{OL} \), particularly in the case of autoregressions with moderate to large correlation. Suppose one is using \( \hat{\mu}_{OL} \) with robust scale estimate \( \hat{\sigma}_R \), and that the series contains just one huge isolated outlier in the \( \epsilon_t \) 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 \( \sigma_R \) in magnitude and will thus be down-weighted. Unfortunately, it is only the initial residual \( r_t \) 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 \( (Y_{t-1}, Y_t) \) will be far removed from the regression line with slope \( \varphi \) and intercept \( \gamma \), but the pairs \( (Y_{t-1}, Y_t) \), \( t = t_0 + 1, \ldots \), constitute good leverage points (i.e., points which will lie close to the regression line and are large in magnitude) for estimating \( \gamma \) and \( \varphi \) -- the latter with ultra precision when \( \epsilon_t \) has a heavy-tailed distribution (Martin, 1982). The ordinary location M-estimate would downweight such points.

The asymptotic and finite-sample efficiencies of \( \hat{\mu}_{OL} \) relative to \( \hat{\mu}_M \), along with awkwardness and impracticality of assessing the variability of \( \hat{\mu}_{OL} \), 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 \( \mu \) 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 $\psi_1=\psi_4$ or $\psi_1=\psi_{ND}$ (see equations (4.2)-(4.4)). The keys to this are the following relationships:

\begin{align}
\int_{-\infty}^{\infty} \Phi(ax+\beta) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx &= \Phi\left(\frac{\beta}{\sqrt{1+\alpha^2}}\right) \quad (A.1) \\
\int_{-\infty}^{\infty} \Phi(ax)\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}} \quad (A.2)
\end{align}

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

\begin{align}
\psi_\varepsilon(\varepsilon) = c \cdot s_{\varepsilon} \psi_{CN}\left[ \frac{\varepsilon}{c s_{\varepsilon}} \right] = cs_{\varepsilon} \sqrt{2\pi} \left[ \Phi\left( \frac{\varepsilon}{c s_{\varepsilon}} \right) - \frac{1}{2} \right] \quad (A.3)
\end{align}

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

\begin{align}
\psi_{COM}(y) = c_{OM} s_y \sqrt{2\pi} \left[ \Phi\left( \frac{y}{c_{OM} s_y} \right) - \frac{1}{2} \right] \quad (A.4)
\end{align}
which is \( \psi \) scaled for the \( Y \) process, with tuning constant \( c_{W} \).

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

\[
E_{C} \psi_{c}^{2}(e) = k^{2} \left( 1 - \delta \right) \tan^{-1} \left( \frac{\sigma^{2}/k^{2}}{\sigma^{2}/k^{2}} \right) + \delta \tan^{-1} \left( \frac{\sigma^{2}/k^{2}}{\sigma^{2}/k^{2}} \right). \tag{A.5}
\]

Also

\[
E_{C} \psi_{e}^{2}(e) = (1 - \delta) \sqrt{k^{2}/(1 + k^{2})} + \delta \sqrt{k^{2}/(\sigma^{2} + k^{2})}. \tag{A.6}
\]

Thus

\[
V = (1 + \theta)^{2} E_{C} \psi_{e}^{2}(e)/E_{C} \psi_{c}^{2}(e)
\]

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

As for \( V_{c,W} \), we need to evaluate \( C(0) \) and \( C(1) \) using \( \psi_{c_{W}} \). With \( k_{y} = c_{W} s_{y} \) and \( F = NM(\theta, \delta, \sigma^{2}) \), A.2 and A.4 give

\[
E_{F} \psi_{c_{W}}^{2}(Y_{1}) = (1 - \delta)^{2} \sqrt{k_{y}^{2}/(1 + \theta^{2} + k_{y}^{2})} + \delta \sqrt{k_{y}^{2}/(1 + \theta^{2} + k_{y}^{2})}
\]

\[+ \delta \sqrt{k_{y}^{2}/(1 + \theta^{2} + k_{y}^{2})}
\]

\[+ \delta \sqrt{k_{y}^{2}/(1 + \theta^{2} + k_{y}^{2})}. \tag{A.8}
\]

\[C(0) = E_{F} \psi_{c_{W}}^{2}(Y_{1}) = k_{y}^{2} \left( 1 - \delta \right) \tan^{-1} \left( 1 + \theta^{2} \right) - k_{y}^{2}/\sqrt{1 + 2 \theta^{2} k_{y}^{2}}
\]

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

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

\[+ \delta \tan^{-1} \left( 1 + \theta^{2} \cdot k_{y}^{2}/\sqrt{1 + 2(1 + \theta^{2}) k_{y}^{2}} \right). \tag{A.9}
\]

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

\[
\psi_{c_{W}}(Y_{1}) \psi_{c_{W}}(Y_{t+1}) = 2\pi k_{y}^{2} \left[ \Phi(k_{y}^{-1} \varepsilon_{i} + \theta k_{y}^{-1} \varepsilon_{i-1}) \Phi(k_{y}^{-1} \varepsilon_{i+1} + \theta k_{y}^{-1} \varepsilon_{i}) \right]
\]
Since the $\varepsilon_t$ are i.i.d. with distribution $CN(0,\sigma^2)$, we can condition on $\varepsilon_t$ and apply (A.1) to get

$$E[\Phi(k_y^{-1}\varepsilon_t + \theta k_y^{-1}\varepsilon_{t-1})|\varepsilon_t] = (1-\delta)\Phi\left(k_y^{-1}\varepsilon_t/\sqrt{1+\theta^2k_y^{-2}}\right) + \delta\Phi\left(\theta k_y^{-1}\varepsilon_t/\sqrt{1+\sigma^2k_y^{-2}}\right).$$

(A.11)

Similarly

$$E[\Phi(k_y^{-1}\varepsilon_{t+1} + \theta k_y^{-1}\varepsilon_t)|\varepsilon_t] = (1-\delta)\Phi\left(\theta k_y^{-1}\varepsilon_t/\sqrt{1+\sigma^2k_y^{-2}}\right) + \delta\Phi\left(\theta k_y^{-1}\varepsilon_t/\sqrt{1+\sigma^2k_y^{-2}}\right).$$

(A.12)

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

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

(A.13)

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

(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)

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

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

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

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

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

$$= (1-\delta)^2A_1 + \delta(1-\delta)(A_2+A_3) + \delta^2A_4$$

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

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

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

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

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

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

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

and finally,

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

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

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

\[ C(1) = E_F \left[ \psi_{cOM}(Y_t) \psi_{cOM}(Y_{t+1}) \right] = k_y^3 \left[ (1-\delta)^3 A_{11} + \delta (1-\delta)^2 (A_{12}+A_{21}+A_{31}) + \delta^2 (1-\delta) (A_{22}+A_{32}+A_{41}) + \delta^3 A_{42} \right] \quad \text{(A.20)} \]

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

\[ V_{OM} = \frac{C(0)+2C(1)}{E_F \psi_{cOM}(Y_t)} \quad \text{(A.21)} \]

**Normal Derivative \( \psi \)-Function**

Let \( \psi_c \) denote \( \psi_{ND} \) scaled for \( \epsilon_t \), with tuning constant \( c \):

\[ \psi_c(\epsilon) = c \epsilon \psi_{ND}(\epsilon / c \epsilon_t) = \epsilon \exp \left[ -\epsilon^2 / 2c^2 \sigma^2 \right] \quad \text{(A.22)} \]

Similarly, let \( \psi_{cOM} \) denote \( \psi_{ND} \) scaled for \( Y_t \), with tuning constant \( c_{OM} \):

\[ \psi_{cOM}(y) = y \exp \left[ -y^2 / 2c_{OM}^2 \sigma_y^2 \right] \quad \text{(A.23)} \]

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

\[ E_G \psi^2(\epsilon) = (1-\delta) k^3 / (2+k^3)^{\frac{3}{2}} + \delta \sigma^2 k^3 / (2\sigma^2+k^2)^{\frac{3}{2}} \quad \text{(A.24)} \]

and

\[ E_G \psi_\epsilon(\epsilon) = (1-\delta) \frac{k^3}{(1+k^2)^{\frac{3}{2}}} + \delta \frac{k^3}{(\sigma^2+k^2)^{\frac{3}{2}}} \quad \text{(A.25)} \]
Now $V = E_G \psi^2_e(\varepsilon)/E_G^2 \psi_e(\varepsilon)$ 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

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

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

$$+ \delta^2 k_y^3/(\sigma^2+\theta^2\sigma^2+k_y^2)^2.$$  \hspace{1cm} (A.26)

As for $C(0)$:

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

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

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

$$+ \delta^2(1+\theta^2)k_y^3/[2\sigma^2(1+\theta^2+k_y^2)]^2.$$  \hspace{1cm} (A.27)

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

$$E[\psi(Y_{t+1})\psi(Y_t) | \varepsilon_i] = E_F\left[ (\varepsilon_{t+1} + \theta \varepsilon_i) \exp[-(\varepsilon_{t+1} + \theta \varepsilon_i)^2/2k_y^2] \right]^2$$

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

$$= K_1(\varepsilon_i) \cdot K_2(\varepsilon_i).$$  \hspace{1cm} (A.28)

where

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

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

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

Therefore,

\[ C(1) = E_C[K_1(\epsilon_t)K_2(\epsilon_t)] \]

\[ = \theta k_y^3 \left\{ (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \right\} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + (1-\delta)^2 \left[ \theta^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + \delta \sigma_0^2 \left[ \theta^2\sigma_0^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + \delta \sigma_0^2 \left[ \theta^2\sigma_0^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + \delta \sigma_0^2 \left[ \theta^2\sigma_0^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ + \delta \sigma_0^2 \left[ \theta^2\sigma_0^2(\sigma_0^2+k_y^2)+\sigma^2(1+k_y^2)+(1+k_y^2)(\theta^2+k_y^2) \right] \frac{3}{2} \]

\[ \left( A.29 \right) \]
Now (A.26), (A.27) and (A.29) can be combined to obtain the closed form for $V_{OM}$. 
References


Domain," edited by Brillinger and Krishnaiah.


Figure 1. AREFF versus $\sigma$ for MA(1) model using $\psi_{ND}$.

Figure 2. AREFF versus $\sigma$ for MA(1) model using $\psi_{ND}$ and $c_{OM} = c_{OM}(\theta)$.
Figure 1a. AREFF versus $\sigma$ for MA(1) model using $\psi_\phi$.

Figure 2a. AREFF versus $\sigma$ for MA(1) model using $\psi_\phi$ and $c_{OM} = c_{OM}(\theta)$. 
$\text{REFF} = 0.5$
$\text{REFF} = 0.9$
$\text{REFF} = -0.5$
$\text{REFF} = -0.9$

Figure 3. REFF versus $\sigma$ for MA(1) model using $\psi_{\text{ND}}$.

$\phi = -0.5$
$\phi = 0.5$
$\phi = -0.9$
$\phi = 0.9$

Figure 4. REFF versus $\sigma$ for AR(1) model using $\psi_{\text{ND}}$. 
Figure 3a. REFF versus \( \sigma \) for MA(1) model using \( \psi_{x} \).

Figure 4a. REFF versus \( \sigma \) for AR(1) model using \( \psi_{\phi} \).
Figure 5. EFF versus $\sigma$ for MA(1) model using $\psi_{ND}$.

Figure 6. EFF versus $\sigma$ for AR(1) model using $\psi_{ND}$. 

Figure 5a. EFF versus $\sigma$ for MA(1) model using $\psi_\phi$.

Figure 6a. EFF versus $\sigma$ for AR(1) model using $\psi_\phi$. 