Methodology Branch

Time Series Research and Analysis
Division

Direction de la méthodologie

Division de la recherche
et de l'analyse des chroniques

Canada

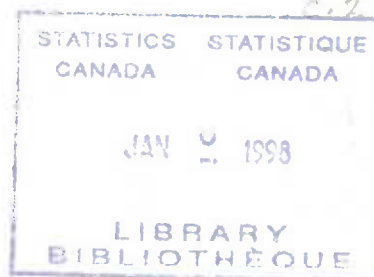# A MONTE-CARLO STUDY OF THE FIRST-ORDER MONTHLY
SEASONAL MOVING AVERAGE PROCESS

by

Normand J.D. Laniel

This is a preliminary version. Do not quote without author's permission.
Comments are welcome.

SUMMARY

A résumé of previous studies on estimators for auto regressive moving average models is given. The maximum likelihood, unconditional and conditional least squares estimators are described. Monte-Carlo results are presented for the $MA(1)_{12}$ seasonal model estimated by the three estimators. Conclusions about the choice of estimator are drawn from these results.

TABLE OF CONTENTS

1. Introduction

There exist several algorithms of procedures for estimation of autoregressive moving average models. The most widely used procedures are Least Squares [Box and Jenkins (1970)] and Maximum Likelihood [Newbold (1974), Anderson (1975), Osborn (1976), Dent (1977), Ali (1977), Phadke and Kedem (1978), Ansley (1979), Hillmer and Tiao (1979), Ljung and Box (1979), Gardner et al. (1980) McLeod and Sales (1983), Mélard (1984)]. The two Least Squares procedures proposed by Box and Jenkins, Conditional (CLS) and Unconditional (ULS), are approximations of the Maximum Likelihood procedure (ML). In ULS, the presample residuals are approximated by back-forecasting and in CLS they are set to zero.

Box and Jenkins suggested using ULS instead of CLS when the model is seasonal or when the series is short. ML procedures are efficient for seasonal or non-seasonal models and for short or moderately long series.

The properties of CLS, ULS and ML for autoregressive moving average models with criteria functions and Monte-Carlo experiments have been studied by many authors.

Nelson's (1974) Monte-Carlo experiments on the first order moving average model have shown that for a series of length N=30 the CLS estimator has a smaller mean square error than the ULS estimator if $|\theta|$ is near zero and the reverse situation if $|\theta|$ is near one. They also have shown for N=100 that the

two estimators are equally efficient. Nelson concluded that the analyst would do well to choose a final estimator on the basis of the apparent magnitude of the moving average parameter θ.

Dent and Min (1978) did a Monte Carlo study of a variety of estimators for six simple ARMA models with series of leng N=100 only. They found for the AR(1), AR(2), AR(3), MA(1), MA(2) and ARMA(1,1) models that CLS, ULS and ML are (except in a few AR(3) cases) equally efficient methods to estimate the parameters in regard of bias, variance and mean square error. The small differences in efficiency that Dent and Min found between the three methods are certainly due to the fact that their series were long in comparison with the orders of the six models. They would have found different results with shorter series.

Monte-Carlo experiments have been done by Ansley and Newbold (1980) for nonseasonal and seasonal autoregressive moving average models. Their study reveals that for the ARMA(1,1) and sample size 50 the CLS estimator has unacceptably large mean squared errors for large values of $|\theta|$ or when ϕ and θ are close to each other. On the other hand, ULS and ML are equally efficient. They also analyzed some quarterly seasonal models for sample size 50. For the $AR(1)(1)_4$ model they did not find large differences between the three estimators. However for the $MA(1)(1)_4$ model they found that for large values of the parameters CLS has a bias toward zero with large mean squared errors. ULS has smaller mean squared errors than ML and for these two estimators have much smaller MSE than CLS. For the same model and small values of the parameters ULS has a bias toward one and CLS has smaller mean squared errors than

ML but they are still of moderate size. The authors also presented results for some monthly seasonal models for sample sizes 50 and 100. These show that for the nonseasonal parameter of the $AR(1)(1)_{12}$ model ULS and ML perform well and CLS does poorly. On the other hand, for the seasonal parameter CLS and ML are almost equally efficient but, ULS estimates are biased towards the boundary values ± 1. For the $MA(1)(1)_{12}$ model the results are pretty well the same as for the quarterly model.

Davidson (1981) presents an analysis of the MA(1) model small sample properties. The results are essentially the same as Nelson (1974) but, the author explains the properties with criteria functions.

A study based on criteria functions has been done by Osborn (1982) for the MA(1) model. It gave results compatible with other authors and showed that CLS has a tendency to overestimate the residual variance, ULS to underestimate it and ML to give an unbiased estimate of it.

Most of the monthly seasonal Canadian socio-economic time series are fitted weel, after suitable differencing, by an ARMA $(p,q) \times (0,1)_{12}$ multiplicative model. In some applications the estimation of the parameters for that model are crucial. For example, in seasonal adjustment based on ARIMA models the seasonal parameter will determine how the seasonal component will be extracted from the original series, because it indirectly defines the type of seasonality (moving or stable).

Practitioners know that the estimated value for a MA seasonal parameter obtained from CLS, ULS, and ML are often very different. They also know that with moderate or large sample size the three estimators lead to the same value for nonseasonal parameters (Dent and Min (1978)).

The present study concentrates on the $MA(1)_{12}$ seasonal model, since the results could be extended to most of the $ARMA(p,q) \times (0,1)_{12}$ models. It consists of Monte-Carlo experiments for sample sizes N=20, 40, 80 and 160 and ten parameter values ranging from 0.1 to 1.0 which correspond to the range where estimates are found most of the time. In section 2, the two least squared and the maximum likelihood estimators are presented. The design of the Monte-Carlo experiments is described in detail and the results are presented in the next section. Finally, section 4 discusses the results and some conclusions are drawn about the choice of estimator.

2. The ML estimator and approximations

An autoregressive/moving average process of order (p,q) is defined by

$$\phi(B)w_t = \theta(B)a_t \tag{2.1}$$

where the a_t's are independent normally distributed random variables with mean zero and variance σ^2, $\phi(B) = 1-\phi_1 B- \ldots -\phi_p B^p$, $\theta(B) = 1-\theta_1 B- \ldots - \theta_q B^q$ and B is the backshift operator, such that $B^k w_t = w_{t-k}$. The model is stationary and invertible if the equations $\phi(B) = 0$ and $\theta(B) = 0$ have all roots outside the unit circle.

Since the a_t's and hence the w_t's have a normal distribution, the likelihood function for the parameters $\phi = (\phi_1, \ldots, \phi_p)'$, $\theta = (\theta_1, \ldots, \theta_q)'$ and σ^2 is $L(\phi, \theta, \sigma^2|w) = (2\pi\sigma^2)^{-\frac{1}{2}n} |\Sigma|^{-\frac{1}{2}} \exp\{-(2\sigma^2)^{-1} w'\Sigma^{-1} w\}$ (2.2) the propability density function of $w = (w_1, \ldots, w_n)'$, where $\sigma^2\Sigma$ is the nxn variance covariance matrix of w. The mean of the w_t's is assumed to be zero.

It is easy to show that the maximum likelihood estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{w'\hat{\Sigma}^{-1}w}{n} \tag{2.3}$$

Using equation (2.2) in (2.1) and taking the $(-2/n)$ power one obtains the concentrated likelihood

$$L_o(\phi, \theta|w) = (w'\Sigma^{-1}w) |\Sigma|^{1/n} \tag{2.4}$$

wich minimization yields the ML estimates of ϕ and θ. In our Monte-Carlo experiments the algorithm of Ansley (1979) was used to minimize (2.4).

For the general model (2.1) Box and Jenkins (1970) show that

$$w' \Sigma^{-1} w = S(\phi, \theta) = \sum_{t=-\infty}^{n} \hat{a}_t^2 \tag{2.5}$$

where $\hat{a}_t = E[a_t|w]$. If the process is MA(q) (p=o) then (2.5) reduces to

$$S(\phi, \theta) = \sum_{t=1-q}^{n} \hat{a}_t^2. \tag{2.6}$$

The first approximation to ML, namely ULS, is obtained by neglecting $|\Sigma|^{1/n}$ in (2.4), which tends to 1 for large n, and minimizing (2.5). The pre-sample residuals are computed by back-forecasting as described in Box and Jenkins (1970).

CLS is the second approximation to ML and is the result of minimizing

$$S^*(\phi, \theta) = \sum_{t=p+1}^{n} \hat{a}^2 t. \tag{2.7}$$

The residuals are calculated recursively by letting $\hat{a}_t = 0$ for $t < p + 1$ (see Box and Jenkins, 1970).

3. Monte-Carlo experiments

This section reports Monte-Carlo results obtained for the $MA(1)_{12}$ model
$w_t = a_t - \theta\, a_{t-12}$, $a_t \sim NI(0,1)$. For each parameter value and sample size
(listed in the introduction), 50 realizations of the process were generated as
described below. Then, the parameters θ and σ^2 were estimated with CLS, ULS
and ML (Ansley's procedure, 1979). Finally the biases, variances and mean
squared errors were computed for the three estimators.

All computations were performed on an IBM 360 computer at Statistics
Canada. An in-house pseudo-random number generator was used to generate inde-
pendent uniform [0,1] deviates, from which independent normal variates were
generated using the exact polar method of Box and Muller (1958).

To generate an ARMA (p,q) time series given by (2.1) we require p+q star-
ting values. Suitable starting values were obtained with the Waterloo Simula-
tion Procedure 2 of Mc Leod and Hipel (1978).

The numerical function minimizations were done with the non-linear least
squares algorithm of Marquardt (1963) using the ARIMA procedure of SAS/ETS
(1982). The convergence criterion was to stop iterating when the absolute
change in the parameter estimate $\hat\theta$ was smaller then 10^{-4}. For the three esti-
mators the variance of the white noise was estimated by dividing the sum of
squares error by the number of residuals minus one.

The results of the experiments (MSE, bias and variance curves) are shown in figures 1 to 6 (see appendix). It is seen that the efficiency of the three estimators depends upon the true value of θ and the length of the simulated series (this is not surprising considering their asymptotic properties).

Generally, when θ is estimated (see figures 1, 2 and 3), CLS has the smallest mean squared error for small values(*) of θ but often the largest for large values of θ. This is mainly due to its bias wich has a similar pattern while its variance is often small. ULS behaves in a complementary way. For small θ it has the largest MSE but the smallest for large θ. This is due to both its bias and variance which follow the same pattern. ML is somewhat different. Except for N=20, it has often nearly the smallest MSE for all θ but it usually does not have the smallest error. This is explained by a relatively small bias and large variance for all θ.

It is interesting observe that the following inequality holds almost all the time $\hat{\theta}_{ULS} > \hat{\theta}_{ML} > \hat{\theta}_{CLS}$.

For the estimation of σ^2 (see figures 4, 5 and 6), CLS has the smallest MSE for small θ and the largest for large θ. This is due to both its bias and variance which become larger as θ increases. ULS has a large MSE for small θ but the smallest for large θ. This is due to its variance which follows the same pattern while its bias is relatively large for all θ. ML has often MSE nearly the smallest or the smallest for all θ. This is explained by it small bias (often the smallest) while its variance is, most of the time, the largest for all θ.

(*) The meaning of small (or large) value of θ varies with the sample size. For N=20 it means $\theta < 0.3$, for N=40 $\theta < 0.6$ and for N=80 and 160 $\theta < 0.8$.

Another observation one can draw from the figures is that CLS has, for all Θ, a larger estimated value for σ^2 than ULS and ML. Furthermore, ULS gives a smaller value for $\hat{\sigma}^2$ than ML when Θ is small while the reverse is true for large Θ.

One can verify that all these results agree with other studies.

4. Conclusion

This study shows that for an ARIMA model with an $IMA(1,1)_{12}$ seasonal part the best estimator for small values of θ is CLS and the best for large θ is ULS. It also shows that Ansley's ML procedure generally does well for the bias as an estimator for θ but its variance is large.

The investigation and the development of other estimators than the ones considered here should also be considered.

Since CLS and ULS are the best estimators for small and large values of θ respectively, the use of a criterion should be considered to choose between CLS and ULS (For example a lack of fit test). One has to note that the estimated variance of the residuals cannot be used since $\hat{\sigma}^2_{CLS} > \hat{\sigma}^2_{ULS}$, as shown in the study.

BIBLIOGRAPHY

Ali, M.M. (1977). Analysis of autoregressive – moving average models:
    Estimation and prediction. Biometrika 64,535-45.

Anderson (1975). Maximum likelihood estimation of parameters of
    autoregressive processes with moving average residuals and other
    covariance matrices with linear structure. Ann. Statist. 3.1283-1304

Ansley, C.F. (1979). An algorithm for the exact likelihood of a mixed
    autoregressive moving average process. Bionetrika 66,59-65.

Ansley, C.F. and Newbold, P. (1980). Finite Sample Properties of Estimators
    for Autoregressive Moving Average Models. Journal of Econometrics,
    13,159-183

Box, G.E.P., and Jenkins, G. M. (1970). Time Series Analysis, Forecasting
    and Control, San Francisco: Holden-Day.

Box, G.E.P. and Muller, M. E. (1958). A note on the generation of random
    normal deviates, Annals of Mathematical Statistics 29,610-611.

Davidson, J.E.H. (1981). Problems with the Estimation of Moving Average
    Processes. Journal of Econometrics, 16,295-310

Dent, W.T. (1977). Computation of the exact likelihood function of an ARIMA
    process, Journal of Statistical Computation and Simulation 5,193-206.

Dent, W.T. and Min, A.S. (1978). A Monte Carlo Study of Autoregressive
        Integrated Moving Average Processes. Journal of Econometrics,
        7,23-55.

Gardner, G., Harvey, A.C. and Phillips, G.D.A. (1980). An Algorithm for Exact
        Maximum Likelihood Estimation of Autoregressive - Moving Average
        Models by Means of Kalman Filtering, Applied Statistics, 29,311-322.

Ljung, G.M. and Box, G.E.P. (1979). The Likelihood Function of Stationary
        Autoregressive - Moving Average Models, Biometrika, 66,265-270.

Marquardt, D.W. (1963). An algorithm for least squares estimation of
        nonlinear parameters, Journal of the Society of Industrial Applied
        Mathematics, 2,431-441.

McLeod, A.I. and Hipel, K.W. (1978). Simulation Procedures for Box - Jenkins
        Models. Water Resources Research, 14,969-975.

McLeod, A.I. and Sales, P.R.H. (1983). An Algorithm for Approximate
        Likelihood Calculation of ARMA and Seasonal ARMA Models, Applied
        Statistics, 32,211-223.

Mélard, G. (1984). A Fast Algorithm for the Exact Likelihood of
        Autoregressive - moving Average Models, Applied Statistics,
        33,104-114.

Nelson, C.R. (1974). The First-Order Moving Average Process, Journal of
    Econometrics, 2,121-141.

Newbold, P. (1974). The exact likelihood function for a mixed autoregressive
    - moving average process, Biometrika, 61,423-426.

Osborn, D.R. (1976). Maximum Likelihood Estimation of Moving Average
    Processes, Annals of Economic and Social Measurement, 5,75-87.

Osborn, D.R. (1982). On the Criteria Functions Used for the Estimation of
    Moving Average Processes, Journal of the American Statistical
    Association, 77,388-392.

Phadke, M.S. and Kedem, G. (1978). Computation of the exact likelihood
    function of multivariate moving average models, Biometrika,
    65,511-519.

APPENDIX

FIGURES 1 TO 6
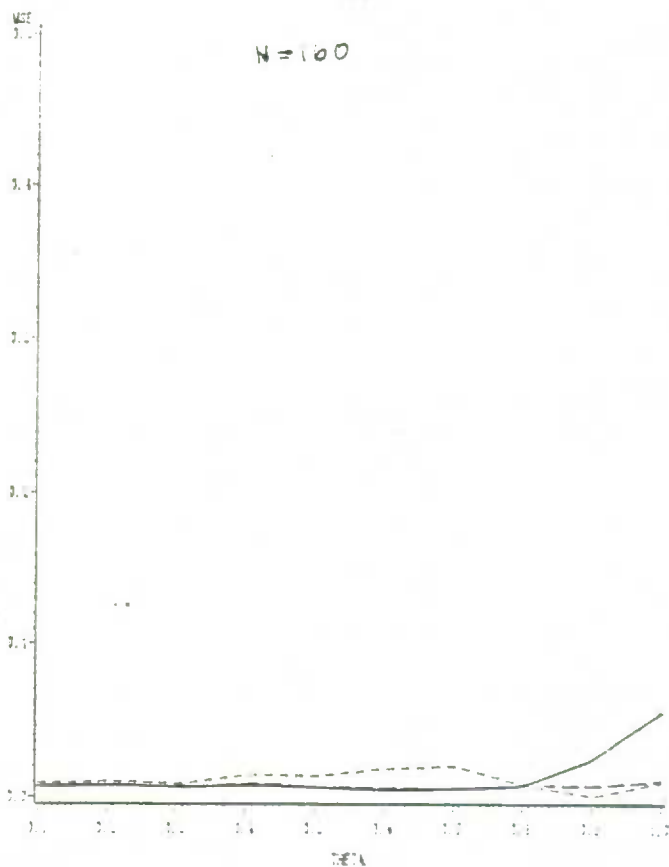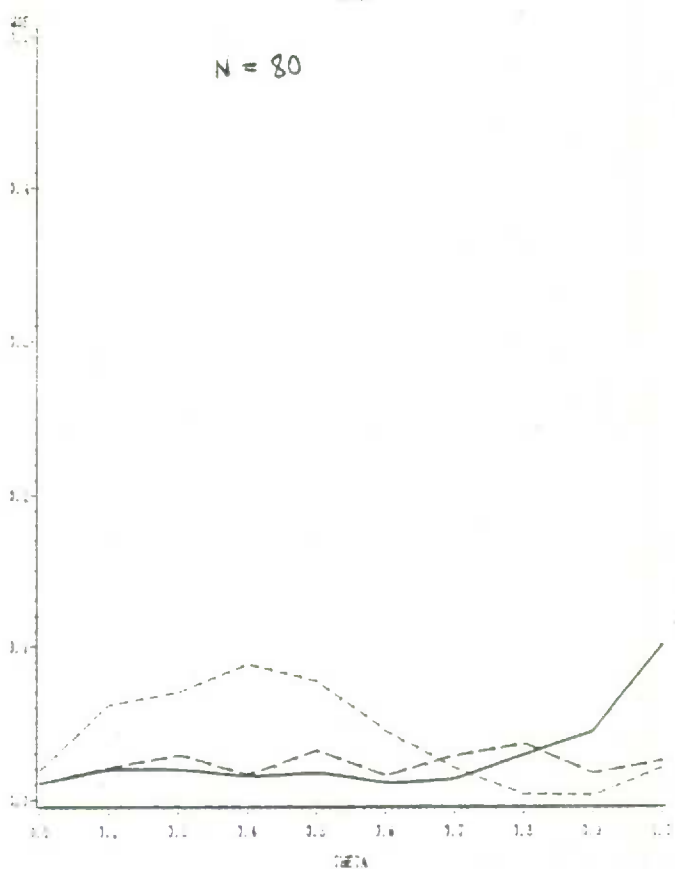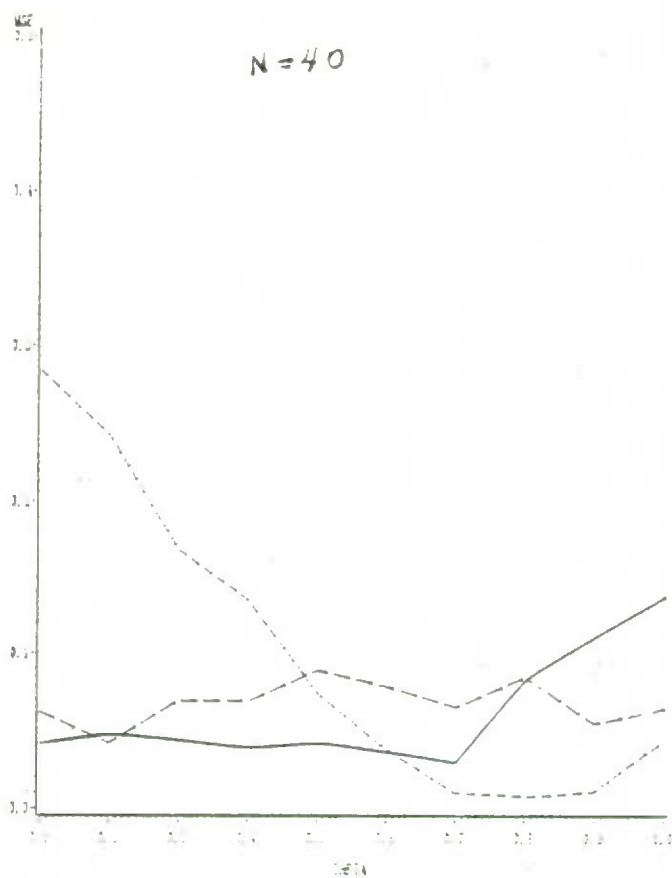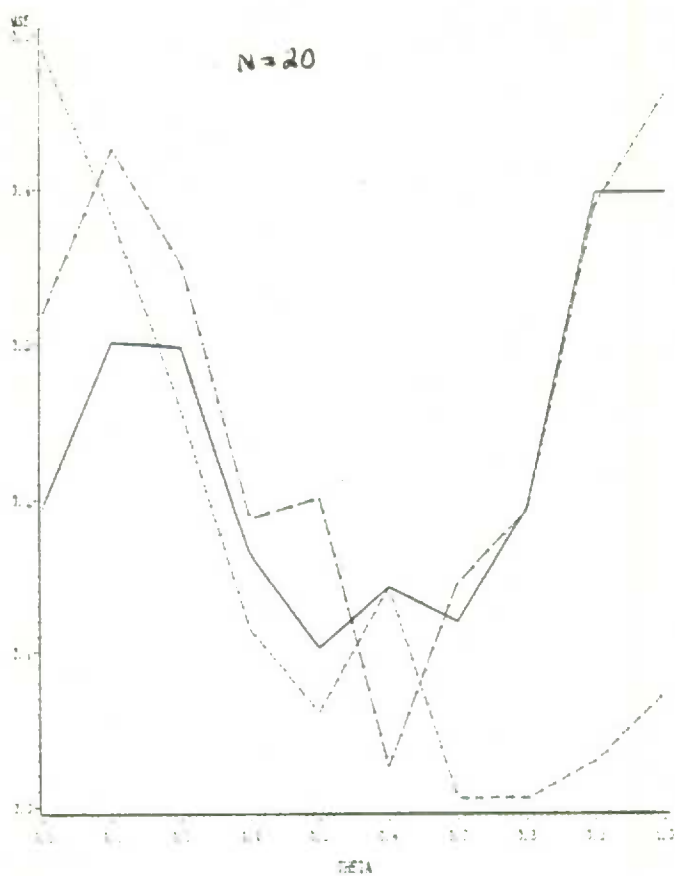
FIGURE 1. MSE of $\hat{M}$.   CLS ———— , ULS ‑‑‑‑‑‑‑ , ML ‑‑‑‑
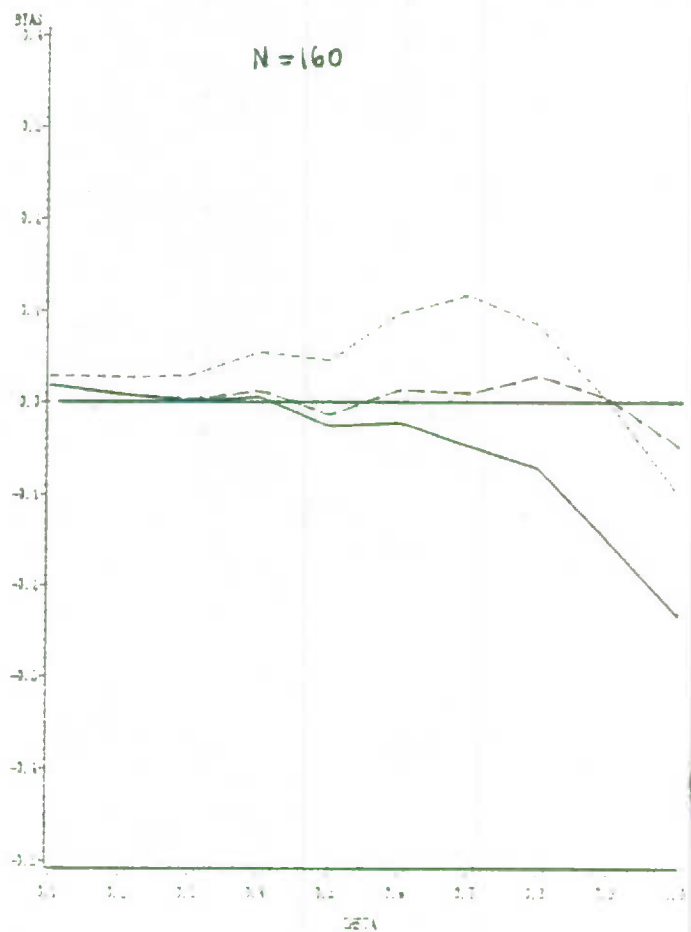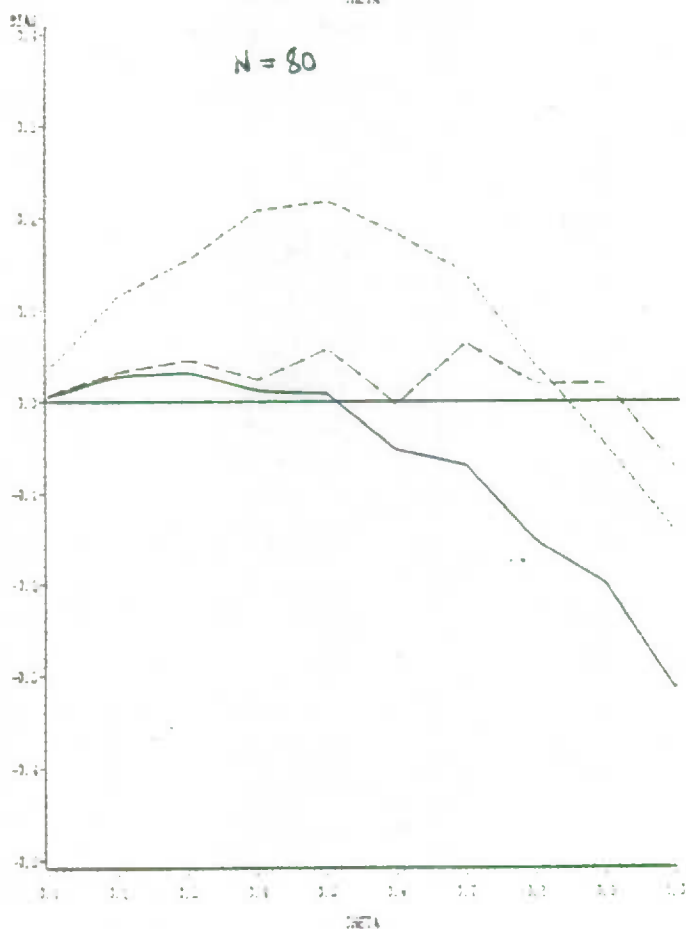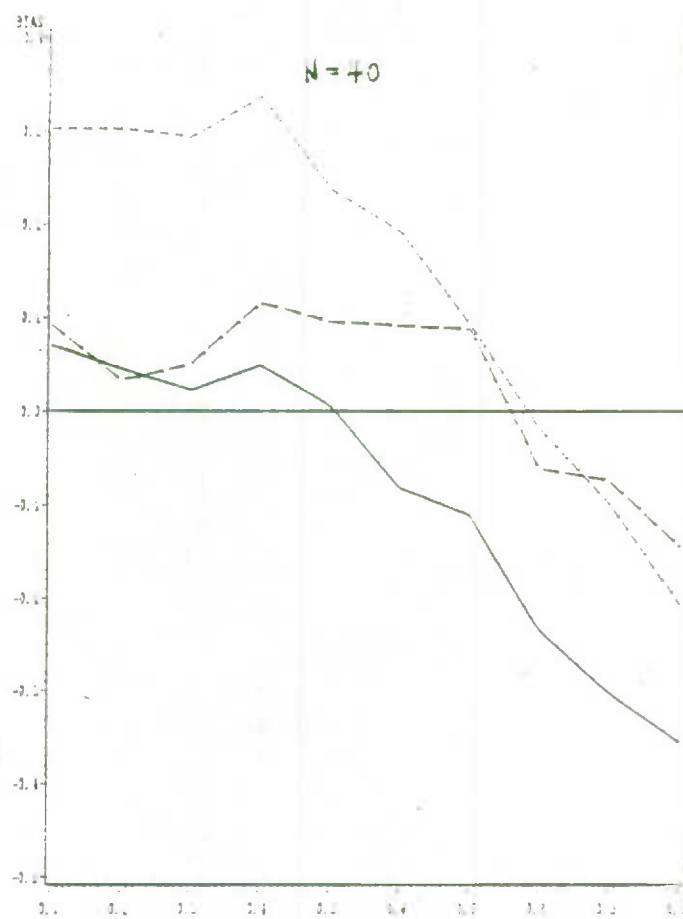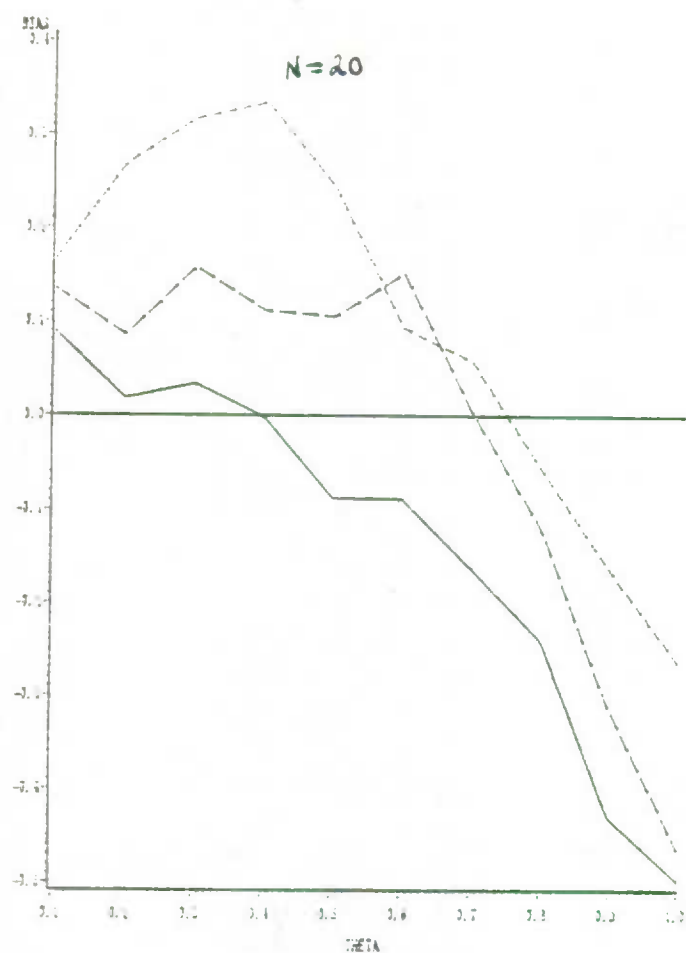
FIGURE 2. BIAS of $\hat{\Theta}$.   CLS ———— , ULS ······· , ML ————
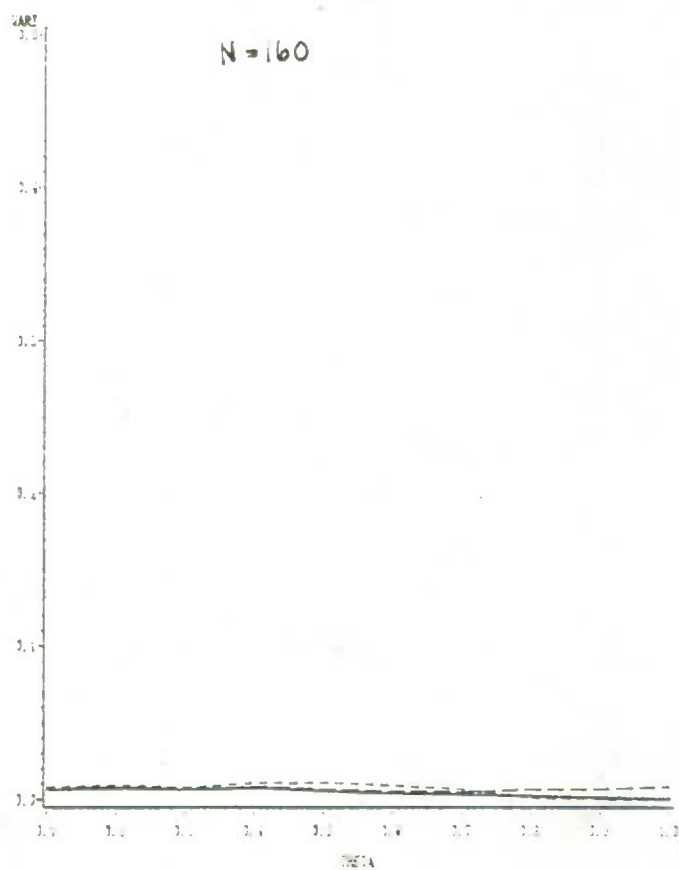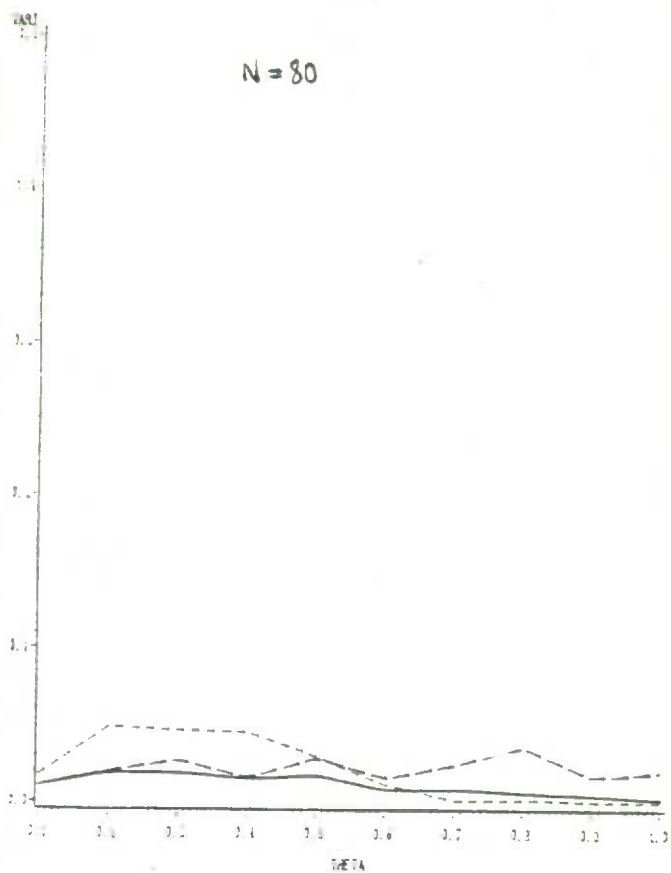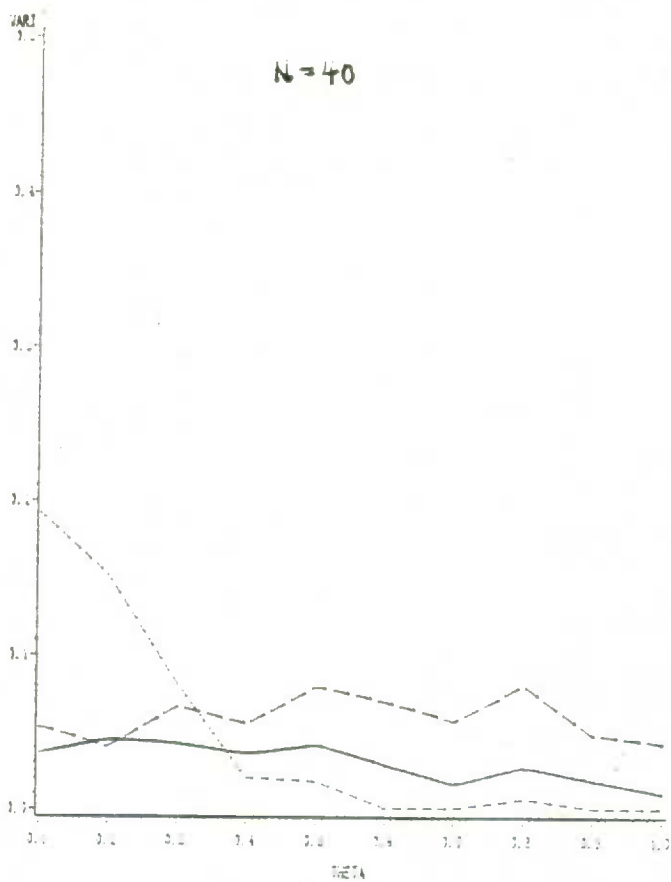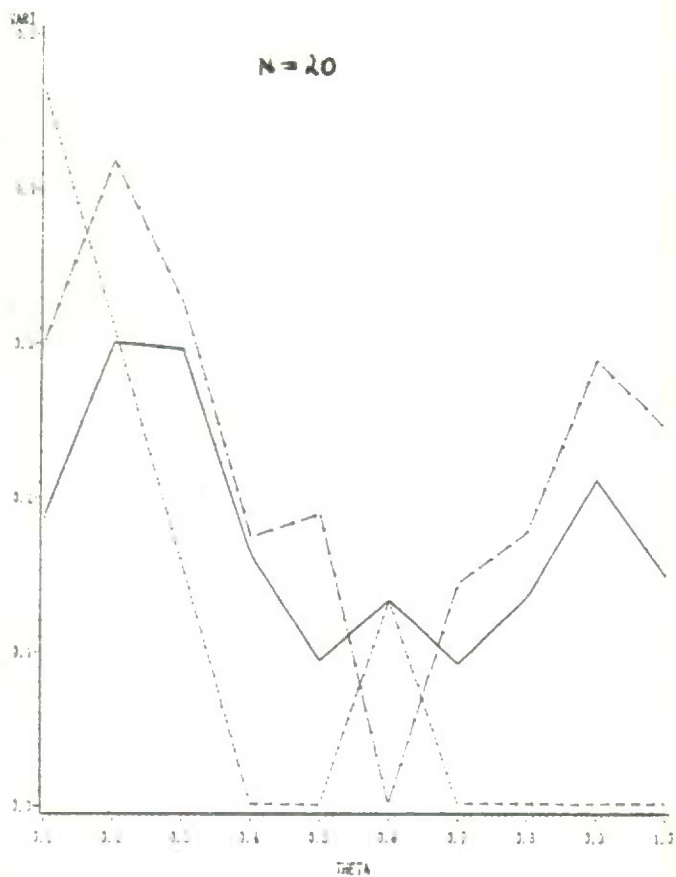
FIGURE 3. VARIANCE of $\hat{\Theta}$.   CLS——, ULS-----, ML---
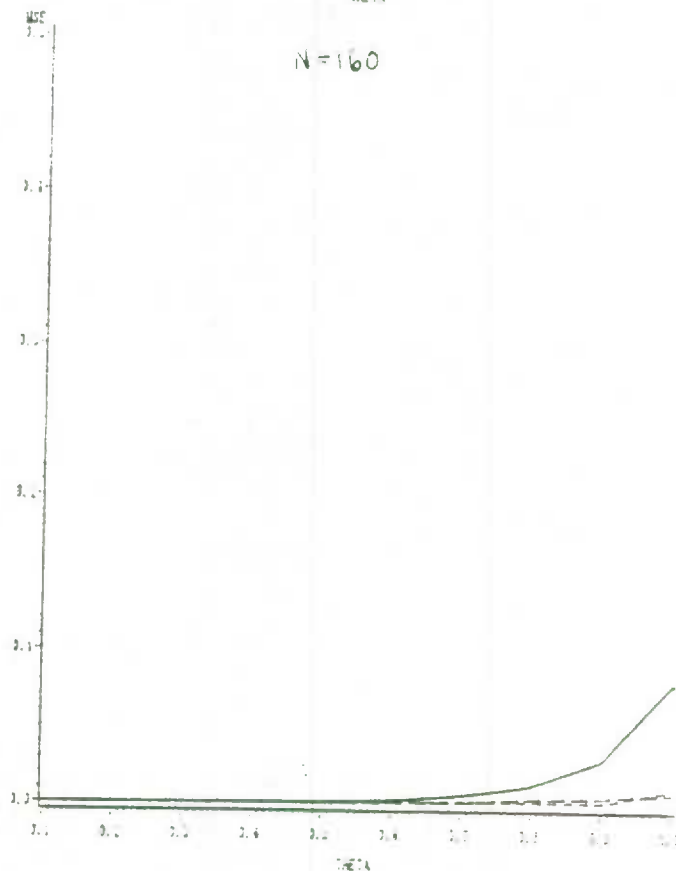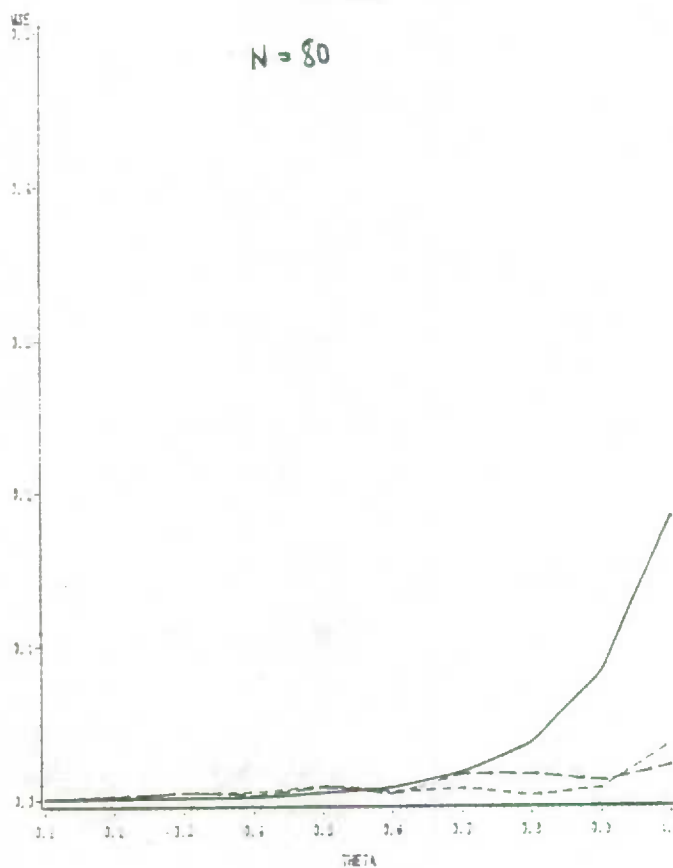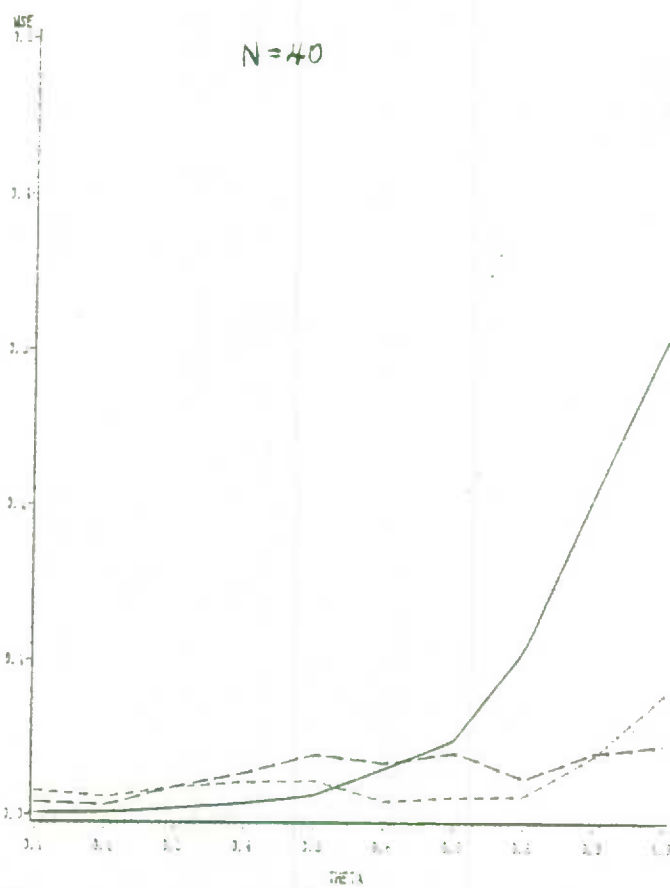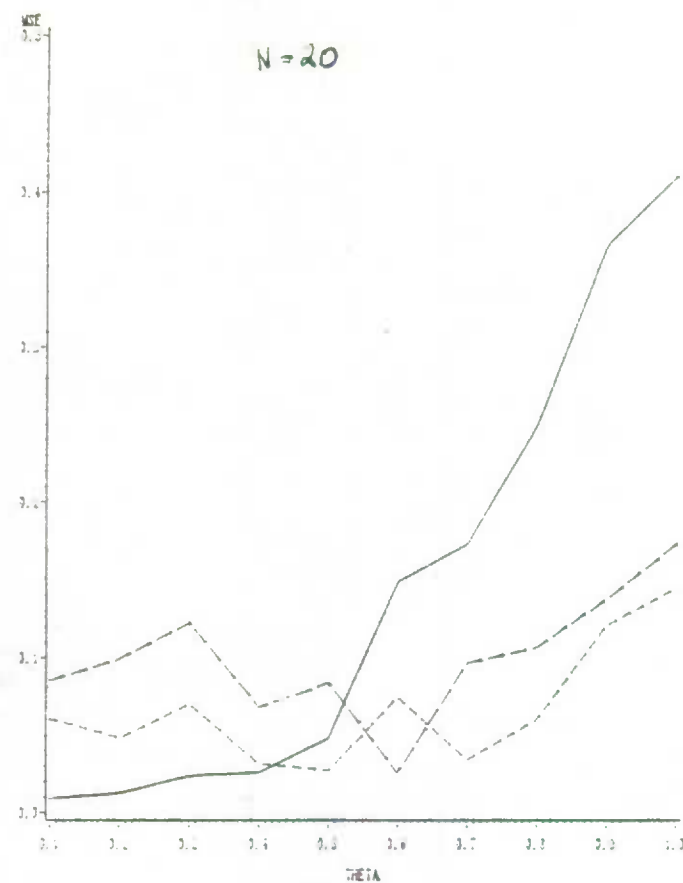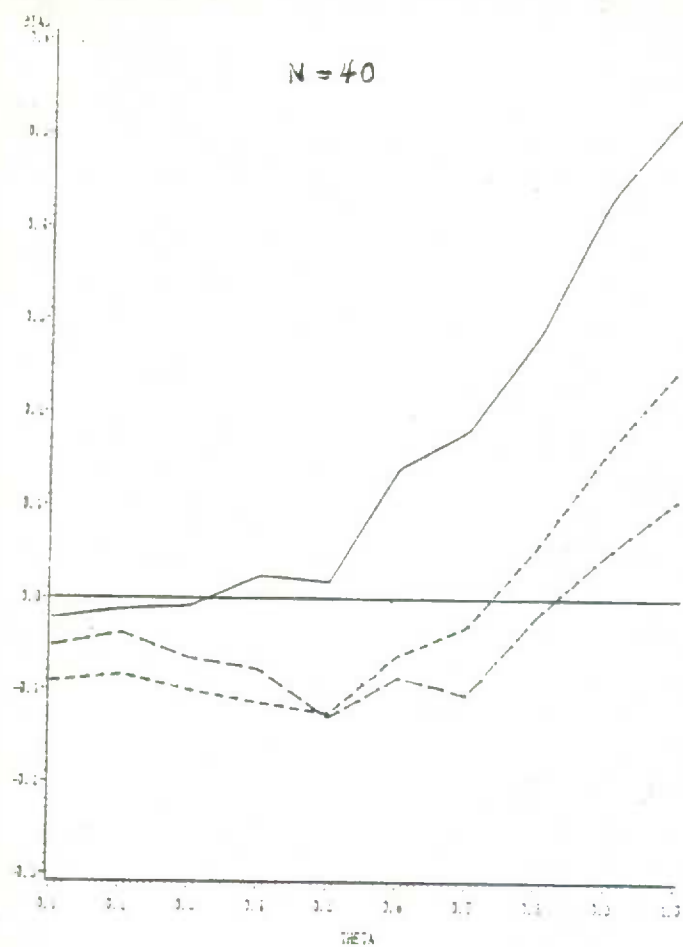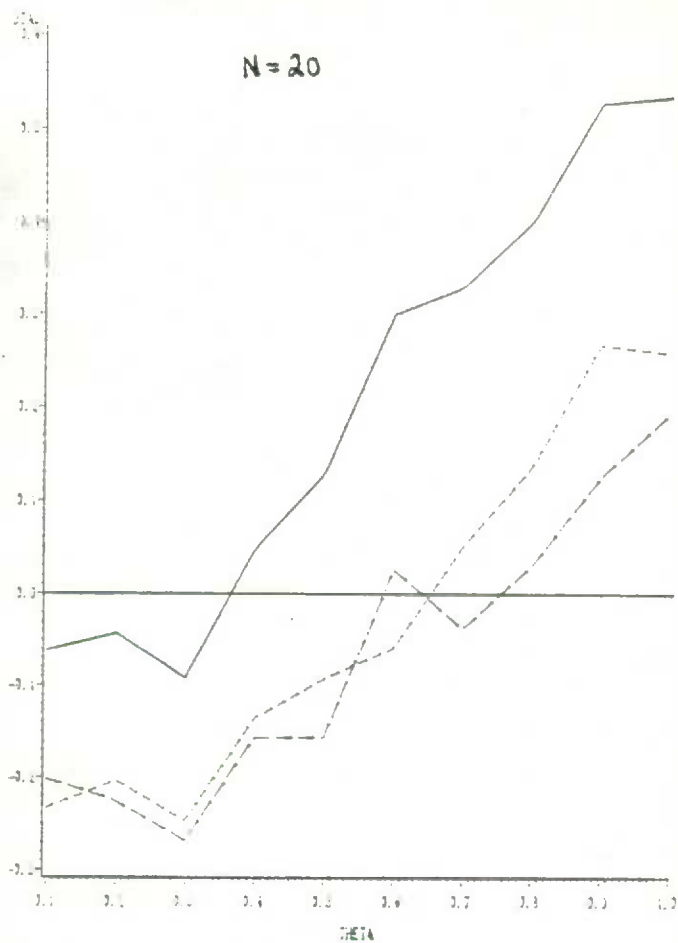
FIGURE 4. MSE of $\hat{\sigma}^2$.   CLS ——, ULS -----, ML ----
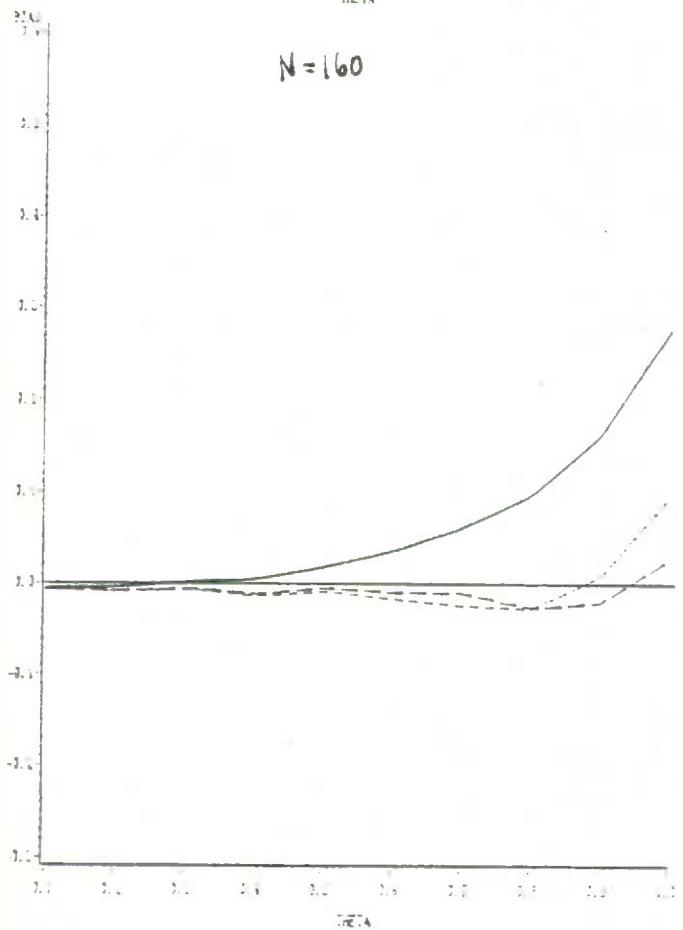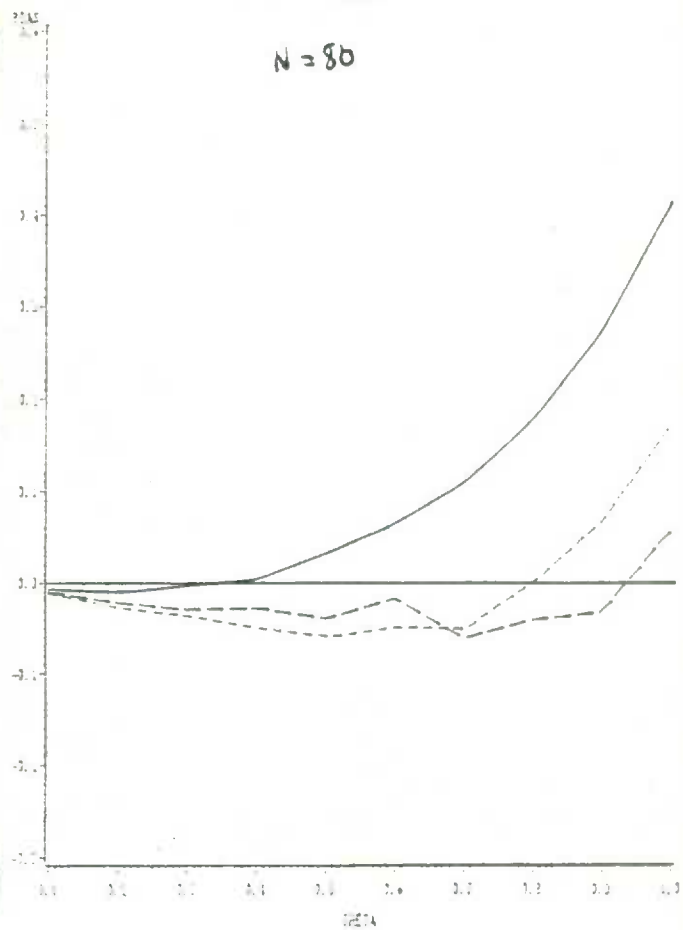
FIGURE 5. BIAS OF $\hat{\sigma}^2$.   CLS ——, ULS ......, ML ————
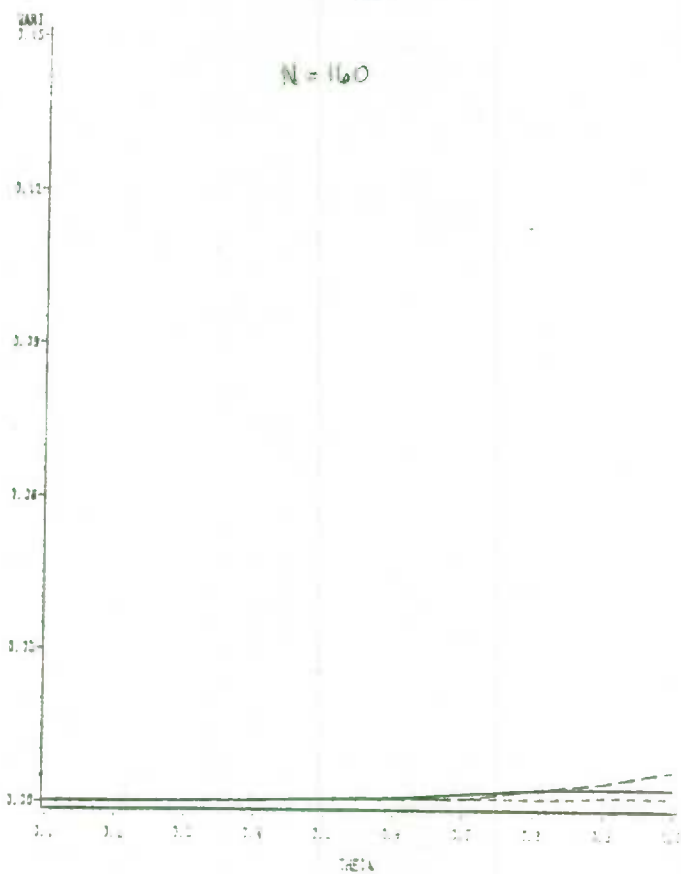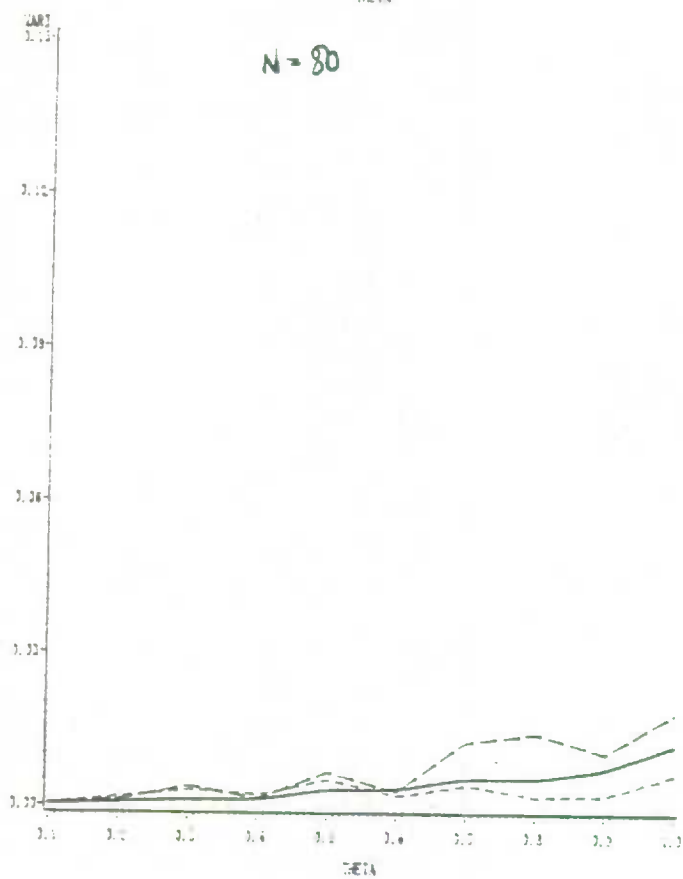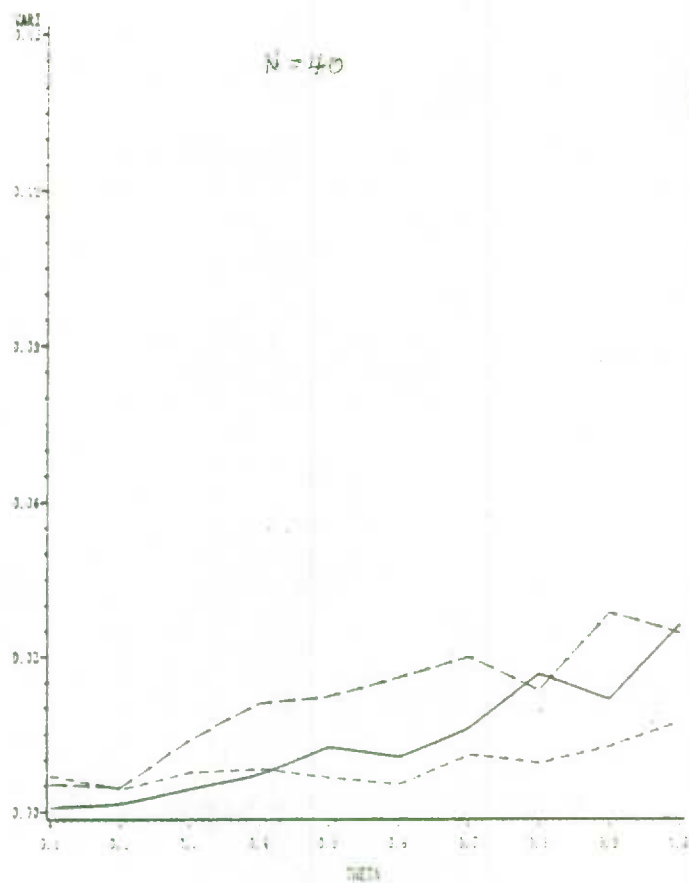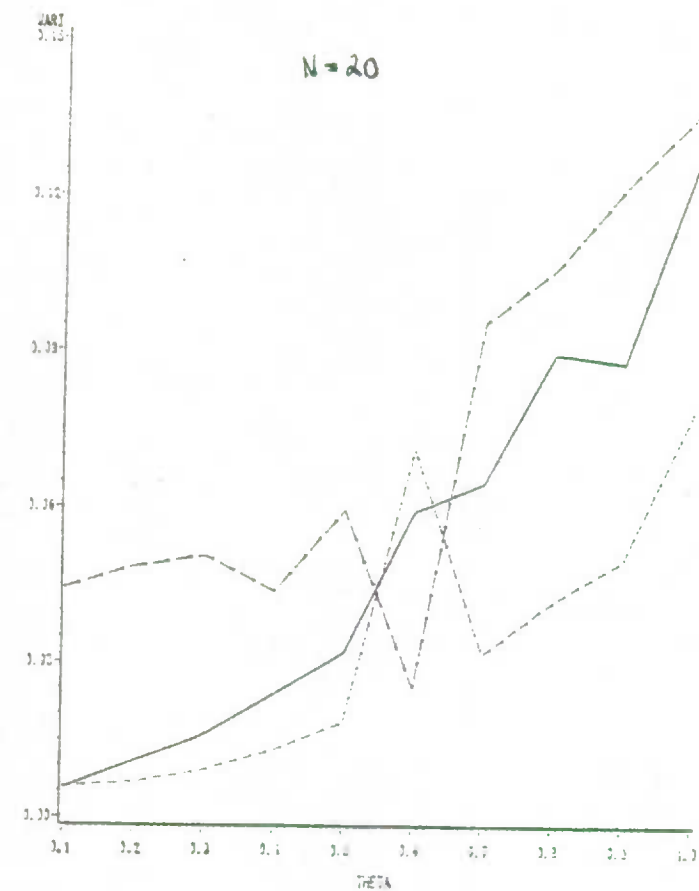
FIGURE 6. VARIANCE of $\hat{\sigma}^2$.   CLS ——— , ULS ······ , ML ———