ISSN 1011-8888

INSTITUTE OF ECONOMIC STUDIES

WORKING PAPER SERIES

W00:19 December 2000

Computations of Bayesian Estimators in ARMA Models

Helgi Tomasson

Address: Department of Economics
University of Iceland
Oddi at Sturlugata
101 Reykjavik
Iceland

Email: helgito@hi.is

Abstract

The computational problem of the likelihood and predictor of stationary ARMA models is discussed, from both a numerical and statistical point of view. The Bayesian estimator is derived by naive Monte Carlo integration, importance sampling Monte Carlo integration, Gaussian quadrature and Laplace's approximation of the integral. In a simulation study the maximum-likelihood estimator is compared with a Bayesian estimator, based on a non-informative prior. Results indicate that although for a given number of replications, importance sampling gives promising results for some parts of the parameter space, there are other parts where considerably more replications would be needed.

1 Introduction

The recent development in computer technology has made many computationally demanding tasks possible which were earlier only hypothetical. In this paper some of the computational and statistical aspects of Bayesian analysis of normal stationary and invertible ARMA(Auto-Regressive-Moving-Average) models are analyzed. The analysis focuses on estimation of the parameters of an ARMA(p,q). In such an analysis, several computational problems have to be solved. First, the computation of the likelihood is a non-trivial matter, and several authors have contributed algorithms to compute the likelihood (Ansley, 1979; Pearlman, 1980; Kohn and Ansley, 1985; Melard, 1984) Second, the computation of the best predictor as a linear combination of past observations in models with moving average terms when sample size is finite is not straightforward. Available methods are the Durbin-Levinson algorithm or the innovation algorithm (see for example, Brockwell and Davis (1991)or a modification of Ansley (1979), an algorithm for calculating the maximum-likelihood as in Tomasson (1986). The Kalman-filter algorithm also gives a computational way of calculating the likelihood and the predictor (see, for example, Harvey (1989) or Harvey (1993). Third, limiting the parameter space to the stationary-invertible region imposes complicated restrictions on the parameters. Implementing a numerical maximum-likelihood estimator requires an algorithm for efficient enforcement of this restriction. In the present paper, this is solved by using the recursive formulas usually used for calculating the partial autocorrelations. Fourth, to calculate the Bayesian estimator, as we do not have a conjugate prior, it is necessary to perform numerical integration which is difficult in high dimensions.

To measure the performance of an estimator, one has to define a criterion and a baseline estimator. Here, the constrained maximum-likelihood estimator is chosen as the baseline. The performance criteria analyzed are the mean-square-error and the mean-square-error of prediction. Ansley's (1979) algorithm was used to calculate the likelihood, and the modification of Ansley's algorithm in Tomasson (1986) was used to calculate the optimal predictor. The enforcement of the stationarity-invertibility constraint is done with a transformation described in Monahan (1984); Tomasson (1986). In high dimensions (say, higher than 3), exact numerical integration has not been an easy task up to now. Therefore, several Monte Carlo techniques have been suggested, such as importance sampling or Gibbs sampling (Geweke, 1992; Ripley, 1987). In this paper only one variant of importance sampling is tried, and Laplace approximation (Tierney et al., 1989) is applied.

2 The model and the likelihood

The process X_t follows a normal ARMA(p,q) model if it is defined by:

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \cdots - \theta_{t-q}\varepsilon_{t-q} \tag{1}$$

where

$$\epsilon_t \sim N(0, \sigma^2), \text{ and } E(\epsilon_t \epsilon_s) = 0 \text{ when t} \neq \text{s}$$

The process is stationary if the roots of the polynomial

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \tag{2}$$

are outside the unit circle, and similarly the model is said to be invertible if the roots of the polynomial:

$$1 - \theta_1 z - \theta_2 z^2 - \cdots - \theta_q z^q \tag{3}$$

are outside the unit circle. It is assumed that the data are in the form of one realization of a time series, x_1, x_2, ... , x_T. The likelihood is of the form:

$$L(\underline{\phi}, \underline{\theta}, \sigma | \underline{x}) = (\frac{1}{2\pi})^{T/2} |\Omega|^{-1/2} \exp(-\underline{x}\prime \Omega^{-1} \underline{x}/2)$$

where

$$\underline{x} = (x_1, \ldots, x_T)\prime, \quad \underline{\phi} = (\phi_1, \ldots \phi_p)\prime \quad \text{and} \quad \underline{\theta} = (\theta_1, \ldots, \theta_q)\prime \tag{4}$$

The variance-covariance matrix Ω is a complicated $T \times T$ function of the parameters, ϕ_i 's, θ_i 's and σ. It is therefore not computationally feasible to calculate the likelihood directly from equation (4).

3 The parameter space: the stationarity-invertibility region

The ARMA(p,q) model is stationary if the roots of the polynomial, (2), are outside the unit circle. Similarly, the model is invertible if the roots of the polynomial, (3), are outside the unit circle. In textbooks these restrictions are usually given for the easy cases when p(or q)=1, and p(or q)=2. When lags are larger these restrictions are more complex and difficult to visualize. Samuelson (1941) shows results on restrictions on the parameters, ensuring that the roots lie outside the unit circle. An algorithm for these conditions is based on calculating the partial autocorrelation for the AR parameters and then using the same algorithm for the MA parameters (Durbin, 1960). That is, the set:

$$C_p = \{\underline{\phi} \text{ such that the roots of (2) are outside the unit circle}\}$$

is given by the recursive relations:

$\underline{\phi} = (\phi_1(p), \dots, \phi_p(p)) \in C_p$ if and only if $|\phi_p(p)| < 1$ and $(\phi_1(p-1), \dots, \phi_p(p-1)) \in C_{p-1}$

where

$$\phi_p(p-1) = (\phi_p(p) - \phi_p(p-i)/(1-\phi_p(p)^2) \text{ for } i = 1, ..., p-1. \tag{5}$$

Defining $r_k = \phi_k(k)$, the reverse relation is given by:

$$\phi_i(k) = \phi_i(k-1) - r_k \phi_{k-i}(k-1) \text{ for } i = 1, ..., k-1. \tag{6}$$

The formulas (5) and (6) give a one-to-one relation between $(r_1, \dots, r_p)$, the vector of partial auto-correlations in the AR case, and $(\phi_1, \dots, \phi_p)$. As the partial auto-correlations can take any value in the interval $(-1, 1)$, this translates the complicated set C_p into the p-dimensional cube. The Jacobian of the transformation F:

$$F : (-1,1)^p \to C_p \tag{7}$$

is found by the chain rule and induction on p

$$J_F = \prod_{j=1}^{\lfloor p-1 \rfloor /2} (1 - r_{2j-1}^2) \prod_{j=1}^{\lfloor p/2 \rfloor} (1 - r_{2j}) \tag{8}$$

This transformation makes it easy to enforce the stationary-invertibility restriction of the parameter space (Samuelson, 1941; Monahan, 1984). This transformation and NAG subroutines were used to get a constrained maximum-likelihood estimate.

4 Bayesian estimation and numerical integration.

One type of Bayes estimator of a parameter β is defined as the expected value of the posterior distribution.

$$\hat{\underline{\beta}}_{BAYES} = \int \underline{\beta}\pi(\underline{\beta}|data)d\underline{\beta} \text{ where } \pi(\underline{\beta}|data) \text{ is the posteriori density for } \underline{\beta} \tag{9}$$

The posteriori distribution, $\pi(\underline{\beta}|data)$, is defined as the scaled likelihood function:

$$\pi(\underline{\beta}|data) = \frac{\pi(\underline{\beta})likelihood(\underline{\beta}|data)}{\int \pi(\underline{\beta})likelhood(\underline{\beta}|data)d\underline{\beta}} \tag{10}$$

where $\pi(\underline{\beta})$ is the a priori distribution of the unknown parameter vector, $\underline{\beta}$. In the cases where there exists a conjugate prior, the above integration can be performed analytically. In other cases, numerical integration is necessary. In this paper, naive Monte Carlo integration is used as well as numerical quadrature integration. In a naive Monte Carlo estimation of the mean of a Bayes estimator, points are sampled uniformly over the parameter space and then averaged. The transformation of the previous section turns the parameter space into a cube, so the regular shape of the transformed parameter spaces makes this particularly easy. In the case of ARMA models, one can, by the transformations in (5) and (6) and the corresponding Jacobian sample from the p+q dimensional cube, calculate the average value of the function weighted with the Jacobian. A more intelligent way of sampling, importance sampling (Ripley, 1987), a concept closely related to stratified sampling is also used. The idea is to sample from a distribution having the same mode as the likelihood but a slightly thicker tails. The integral:

$$\int_{C_p\times C_q} f(\underline{y}|\underline{\beta})\pi(\underline{\beta})d\underline{\beta} = \int_{C_p\times C_q} f(\underline{y}|\underline{\beta})(g(\underline{\beta})/I(\underline{\beta}))I(\underline{\beta})d\underline{\beta} \tag{11}$$

is approximated with

$$\Sigma w_i f(\underline{y}|\underline{\beta}_i)/\Sigma w_i \quad \text{where} \quad w_i = g(\underline{\beta}_i)/I(\underline{\beta}_i) \tag{12}$$

Samples are then drawn from the importance function I. The implementation used here is to approximate the likelihood (the parameter space has been transformed into a p+q dimensional cube) with a beta distribution on $(-1, 1)$ in each dimension. The parameters of the beta distribution are such that the mode is the same as that of the likelihood function, and the variance is proportional to 1/T, the order of the asymptotic variance of the maximum-likelihood estimator. When doing Monte Carlo numerical integration, the relative error has a variance order:

$$constant/n$$

where n is the number of replications in the simulation (Ripley, 1987). The constant in the numerator is a function of many different properties of the function to be integrated and the design of the Monte Carlo experiment. As an alternative to Monte Carlo integration and Gaussian quadrature, one can use some approximation. An interesting approximation is based on the Laplace method. The integral:

$$\int \exp(-mh(\underline{\beta}))d\underline{\beta} \tag{13}$$

is approximated with:

$$\sigma_h \sqrt{\frac{2\pi}{m}} \exp(-mh(\hat{\underline{\beta}}))$$

where σ_h is the square root of the inverse Hessian of h calculated at the mode, $\hat{\underline{\beta}}$, of $exp(-mh(\underline{\beta}))$. The error is of the order $1/m$. Tierney et al. (1989) show that when using Laplace's approximation for the ratio of moments, as in estimation of the posterior mean, the error is of the order $1/m^2$. This is particularly appealing since the likelihood consists partly of an exponent with a sum of T terms. This result is used in the present paper. The log-likelihood will consist of a sum of T-squared terms, so taking $m = T$ is a natural choice, which will result in the error of the numerical approximation being of an order less than the asymptotic variance of the maximum-likelihood estimator.

5 Calculation of the log-likelihood and its derivative

The calculation of the likelihood for an ARMA model is a non-trivial matter. Several algorithms have been developed, e.g., Gailbraith and Gailbraith (1974), Newbold (1974), Dent (1977), Ljung and Box (1979). Here, the method by Ansley (1979) was used. The idea is to avoid inverting an $n \times n$ matrix by transforming the series to one that is mostly a moving average model.

$$Z_t = \begin{cases} X_t & \text{for t } \leq \max(p,q)+1 \\ \phi(B)X_t & \text{else} \end{cases} \tag{14}$$

define the vector $\underline{Z}$ as:

$$\underline{Z} = (Z_1, ..., Z_T)' \tag{15}$$

Then the covariance matrix $E(\underline{Z}\underline{Z}') = \sigma^2\text{M}$ is a TxT band matrix. Now, find a Choleski decomposition LL' for M. The matrix L is a lower triangular bandmatrix. Define:

$$\underline{e} = L^{-1}\underline{Z} \tag{16}$$

The vector $\underline{e}$ is normally distributed with mean zero and its covariance matrix is:

$$E(\underline{e}\underline{e}') = E(L^{-1}\underline{Z}\underline{Z}'L^{-1'}) = L^{-1}E(\underline{Z}\underline{Z}')L^{-1'} = L^{-1}LL'L^{-1'}\sigma^2 = I\sigma^2$$

The transformation $\underline{X}$ to $\underline{Z}$ has unit Jacobian so the log-likelihood can be written as

$$log(likelihood) = \text{-T}/2log(2\pi) - T/2log(\sigma^2) - \underline{e}'\underline{e}/2\sigma^2 - log|L|$$

If σ is concentrated out of the likelihood by substituting the maximum-likelihood estimate $\hat{\sigma}^2 = e'e/T$, the concentrated likelihood L^* has the form

$$log(L^*) = -T/2log(2\pi) - T/2 - T/2log(\underline{e}'\underline{e}/T) - log|L| = constant - T/2log(\underline{e}'\underline{e}) - log(|L|)$$

Calculating the derivative of the log-likelihood is a straightforward, but tedious, algebraic exercise. For further reading, see Kohn and Ansley (1985).

6 Calculating the optimal predictor

The optimal one step-ahead predictor is given by:

$$f_{1,T} = \underline{\gamma}_1{}'\Omega^{-1}\underline{X} \tag{17}$$

and

$$\underline{\gamma}_1{}' = (\gamma(1), \gamma(2), ..., \gamma(T))$$

where γ(k) is the auto-covariance function. This is computationally a very difficult expression. Efficient computation is given by either a Choleski decomposition of the same type as when the likelihood is computed, or by use of the Kalman-filter algorithm. If the $\underline{Z}$-vector of the previous section is calculated, one step-ahead prediction of $\underline{Z}$ is equivalent to one step-ahead prediction of $\underline{X}$, and then the vector $\underline{\gamma}_1{}'$ only has q non-zero values, so calculations can be saved (Tomasson, 1986).

7 Measures of performance of an estimator

The mean-square-error of the estimator is estimated. It is measured by simulation by calculating:

$$\frac{1}{n}\sum_{i=1}^{n}(\underline{\beta} - \hat{\underline{\beta}}_i)^2 \tag{18}$$

where $\hat{\underline{\beta}}_i$ is the estimate in replication number i. The mean-square-error keeps both the variance and the bias of the estimator under control. In a time series model, another measure is natural. The mean-square-error-of-prediction:

$$MSEP = E(X_{T+1} - f_{1,T}(\hat{\underline{\beta}})^2)$$

where $f_{1,T}(\hat{\underline{\beta}})$ is the predictor for time point $T + 1$, given information at point T for the estimated parameter. Calculating

$$\frac{1}{n}\sum_{i=1}^{n}(X_{T+1} - f_{1,T}(\hat{\underline{\beta}}))$$

directly by simulation is time-consuming. The simulation procedure can be improved by decomposing $MSEP$ in two parts:

$$MSEP(\hat{\underline{\beta}}) = E(X_{T+1} - f_{1,T}(\underline{\beta}))^2 + E(f_{1,T}(\underline{\beta}) - f_{1,T}(\hat{\underline{\beta}}))^2 = \\ MSEP(\underline{\beta}) + MSEP^*(\hat{\underline{\beta}}) \quad (19)$$

The first part is the $MSEP$ when the true parameter is used, and this can be analytically calculated. It is therefore only necessary to estimate the second term in (19) by simulation. The excess mean-square-error of prediction, $MSEP^*$, is due to using estimated parameters instead of the true ones. The task of the simulation is then limited to estimate $MSEP^*$ and no simulation power is wasted on the natural variance in the forecast.

8 Simulation results

The simulation analysis consists of generating a time series from the true model. The model is then estimated with maximum-likelihood (ML) and Bayesian methods based on a flat prior. In the Bayesian Monte Carlo analysis, 2000 points were sampled. That number was chosen for practical reasons: running the programs takes reasonably long time that way. Sometimes 2000 is more than enough; sometimes more points are needed. The set of parameter values was chosen to be partly comparable to Monahan (1983) and Dent and Min (1978). In Tables 1 to 5, some simulation results are shown. Expected values, mean-square-error of parameter estimates and the part of the mean-square-error-of-prediction due to estimation are tabulated ($MSEP*$, the second term in (19)). Standard errors of the estimates are shown in parenthesis. It is clear from Tables 1 and 2 that in most cases there is very little difference between ML and Bayesian importance sampling for AR models. It is only for one point (ϕ_1=1.9, ϕ_2=-0.95) that the importance sampling method seems to break down. The Bayes estimates based on Laplace approximations were virtually the same as the maximum-likelihood (results not shown in tables). The naive, uniform Bayes estimator (not shown in tables) gave results that were clearly inferior to the importance sampling results. When using Gaussian quadrature, the Bayes estimator based on a non-informative prior is similar to that of the maximum-likelihood estimator. For MA models, Tables 3 and 4 indicate that the ML-estimator is better. This contracicts with the results from Tomasson(1986). The maximum-likelihood estimator has a distribution with a point mass on the boundary of the parameter space. The small sample properties of the Bayes estimator should therefore be better than those of the maximum-likelihood, as is confirmed in Tomasson(1986). A possible explanation is that the numerical deviation is probably too large.

$$\underline{\hat{\beta}}_{MC-BAYES} = \underline{\hat{\beta}}_{BAYES} + \text{Monte Carlo deviation}$$

This explanation is also suggested by Table 5 where results for Bayesian estimation with non-stochastic Gaussian quadrature rules are shown. Only models with two parameters were generated because of the computer time required to obtain good numerical accuracy in high-dimensional integrals. Other simulations performed were Bayesian analyses, based on sampling uniformly from the parameter space and performing Laplace approximations to the integral. The naive uniform method was inferior to the importance sampling method in most cases, though least inferior in models with little correlation structure. The Laplace approximation worked well in AR models and was practically indistinguishable from the ML method (as it should be with a flat prior). For MA models, it is usual to get an ML-estimate on the boundary of the parameter space, and the Laplace approximation assumes that the mode is in the interior of the parameter space. The Laplace approximation therefore breaks down in such cases. Results for mixed models indicate problems similar to those with the MA models at perhaps even more serious degree.

9 Discussion

Despite great progress in computing speed, numerical Bayesian analysis of ARMA models is still a difficult task. The importance sampling technique is an improvement over naive uniform Monte Carlo methods. Indications of large numerical deviations are observed in areas close to the border of the parameter space in models with a complicated correlation structure. However, these are precisely the most interesting cases. The gain in $MSEP$ of a good estimator is large when the variance of the process is relatively high compared to the noise. This paper has not emphasised the Bayesian inference aspects. Results have been interpreted in a purely frequentistic manner. The choice of the maximum-likelihood estimator as a baseline for comparison can be interpreted as a search for an estimator with better small sample properties in a frequentistic interpretation. The results of the Laplace approximation for pure AR models are promising. Analyzing the small sample properties of the Laplace approximation for complex priors would require using a Gaussian quadrature integration as a baseline. Despite the progress in computers, this is still a highly demanding problem. For the MA and mixed models, the importance sampling Monte Carlo method did not give a satisfactory result, compared with maximum-likelihood, despite the fact that the maximum-likelihood estimator has been found to behave strangely in small samples. There seems to be room for improvement, and the Bayesian quadrature results in Tomasson (1986) seem to confirm this. A better numerical technique is needed for high dimensional Bayesian analysis of MA and mixed models. A fashionable simulation method not mentioned here is Gibb's sampling (Geweke, 1992). This was not tried here for the following reasons. First, it requires sampling one (or a set of) coordinate of the parameter vector from a likelihood that is of a non-standard form. Second, in mixed models there is a possibility of a thin edge in the likelihood. If Gibb's sampling was performed on such an edge one coordinate of the parameter vector at a time, it seems likely that one would just be jumping back and forth without (or with a slow) convergence. The longer the series, the more peaked the likelihood will be, and it will be more and more difficult to estimate the integral with a Monte Carlo experiment. On the other hand, the Laplace approximation becomes better and better as the sample size increases.

Acknowledgements

This paper has been presented at the 'Symposium on computer-aided time series analysis, Turku 1994', and seminars at the Stockholm School of Economics 1994 and the University of Iceland 1995. The author thanks the participants for useful comments.

References

Ansley, C. (1979). An algorithm for the exact likelihood of a mixed autoregressive moving average process. *Biometrika*, 77:59–65.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and method.* Springer, New York.

Dent, W. (1977). Computation of the exact likelihood function of an arima process. *Journal of Statistical Computation and Simulation*, 5:193–206.

Dent, W. and Min, A. (1978). A [m]onte [c]arlo study of autoregressive integrated moving average processes. *Journal of Econometrics*, 11:63–71.

Durbin, J. (1960). The fitting of time series model. *Rev. Inst. Int Stat*, 28:233–244.

Gailbraith, R. and Gailbraith, J. (1974). On the inverse of some patterned matrices arising in the theory of stationary time series. *Journal of Applied Probability*, 11:63–71.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moment. In Bernardo, J., Berger, J., Dawid, A., and Smit, A., editors, *Bayesian Statistics 4*. Clarendon Press Oxford.

Harvey, A. (1993). *Time Series Models.* Harvester Wheatsheaf, Hertfortshire.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Kohn, R. and Ansley, C. (1985). Computing the likelihood and its derivatives for a gaussian arma model. *Journal Statistical Compututation and Simulation.*

Ljung, G. and Box, G. (1979). The likelihood function of stationary auto-regressive moving average models. *Biometrika*, 66:265–270.

Melard, G. (1984). Algorithm as197: A fast algorithm for the exact likelihood of autoregressive-moving average model. *Applied Statistics*, 33:104–114.

Monahan, J. (1984). A note on enforcing stationarity ing autoregressive moving average models. *Biometrika*, 71:403–404.

Monahan, J. F. (1983). Fully bayesian analysis of arma time series models. *Journal of Econometrics*, 21:307–331.

Newbold, P. (1974). The exact likelihood function for a mixed auto-regressive moving-a average process. *Biometrika*, 61:323–336.

Pearlman, J. (1980). An algorithm for the exact likelihood of a high order auto-regressive moving average process. *Biometrika*, 67:232–233.

Ripley, B. (1987). *Stochastic Simulation*. John Wiley and Sons, New-York.

Samuelson, P. (1941). Conditions that the roots of a polynomial be less that unity in absolute value. *Annals of Mathematical Statistics*, 12:360–364.

Tierney, L., Kass, R., and Kadane, J. (1989). Fully exponential approximations to expectations and variances of non-positive functions. *Journal of the American Statistical Associations*, 84(407):711–716.

Tomasson, H. (1986). *Prediction and estimation in ARMA models*. Almqvist & Wiksell, Stockholm.

Model	Expected value	MSE	$MSEP^*$
ϕ_1=0.9	$0.839_{(0.014)}$	$0.019_{(0.005)}$	$0.046_{(0.021)}$
ϕ_1=0.3	$0.268_{(0.018)}$	$0.032_{(0.004)}$	$0.025_{(0.002)}$
ϕ_1=0.5	$0.500_{(0.018)}$	$0.034_{(0.003)}$	$0.060_{(0.010)}$
ϕ_2=0.4	$0.337_{(0.019)}$	$0.037_{(0.005)}$	
ϕ_1=1	$0.937_{(0.020)}$	$0.042_{(0.004)}$	$0.092_{(0.031)}$
ϕ_2=-0.5	$-0.452_{(0.017)}$	$0.030_{(0.007)}$	
ϕ_1=-0.8	$-0.787_{(0.012)}$	$0.014_{(0.002)}$	$0.080_{(0.027)}$
ϕ_2=-0.8	$-0.765_{(0.012)}$	$0.014_{(0.002)}$	
ϕ_1=0.3	$0.281_{(0.018)}$	$0.034_{(0.005)}$	$0.103_{(0.033)}$
ϕ_2=0.3	$0.230_{(0.021)}$	$0.044_{(0.006)}$	
ϕ_1=-0.5	$-0.473_{(0.019)}$	$0.036_{(0.002)}$	$0.093_{(0.037)}$
ϕ_2=-0.4	$-0.390_{(0.018)}$	$0.031_{(0.004)}$	
ϕ_1=1.9	$1.872_{(0.007)}$	$0.005_{(0.001)}$	$0.082_{(0.023)}$
ϕ_2=-0.95	$-0.926_{(0.007)}$	$0.004_{(0.001)}$	
ϕ_1=0.3	$0.305_{(0.019)}$	$0.037_{(0.004)}$	$0.115_{(0.049)}$
ϕ_2=0.5	$0.426_{(0.018)}$	$0.034_{(0.007)}$	
ϕ_3=-0.5	$-0.463_{(0.017)}$	$0.028_{(0.004)}$	
ϕ_1=0.75	$0.728_{(0.020)}$	$0.041_{(0.007)}$	$0.124_{(0.043)}$
ϕ_2=-0.6	$-0.591_{(0.020)}$	$0.042_{(0.006)}$	
ϕ_3=0.6	$0.515_{(0.019)}$	$0.038_{(0.008)}$	
ϕ_1=0.4	$0.419_{(0.020)}$	$0.042_{(0.010)}$	$0.148_{(0.101)}$
ϕ_2=0.25	$0.178_{(0.024)}$	$0.060_{(0.015)}$	
ϕ_3=0.3	$0.316_{(0.019)}$	$0.037_{(0.008)}$	
ϕ_4=0.5	$0.394_{(0.025)}$	$0.063_{(0.023)}$	
ϕ_5=-0.65	$-0.552_{(0.022)}$	$0.049_{(0.012)}$	
ϕ_1=0.45	$0.413_{(0.020)}$	$0.040_{(0.005)}$	$0.182_{(0.05)}$
ϕ_2=-0.2	$-0.193_{(0.023)}$	$0.053_{(0.006)}$	
ϕ_3=0.25	$0.232_{(0.020)}$	$0.039_{(0.005)}$	
ϕ_4=-0.35	$-0.349_{(0.022)}$	$0.047_{(0.007)}$	
ϕ_5=0.35	$0.286_{(0.020)}$	$0.039_{(0.006)}$	
ϕ_1=-0.85	$-0.808_{(0.023)}$	$0.052_{(0.007)}$	$0.227_{(0.092)}$
ϕ_2=-0.6	$-0.619_{(0.027)}$	$0.078_{(0.012)}$	
ϕ_3=0.4	$0.271_{(0.031)}$	$0.093_{(0.018)}$	
ϕ_4=0.6	$0.419_{(0.030)}$	$0.091_{(0.013)}$	
ϕ_5=0.5	$0.359_{(0.026)}$	$0.067_{(0.010)}$	
ϕ_1=0.25	$0.221_{(0.018)}$	$0.033_{(0.004)}$	$0.194_{(0.025)}$
ϕ_2=-0.85	$-0.825_{(0.020)}$	$0.039_{(0.010)}$	
ϕ_3=0.9	$0.772_{(0.023)}$	$0.055_{(0.010)}$	
ϕ_4=-0.2	$-0.212_{(0.020)}$	$0.041_{(0.005)}$	
ϕ_5=0.65	$0.526_{(0.023)}$	$0.052_{(0.009)}$	

Table 1: Simulation results for some AR models. 100 replications of length 25. Maximum-likelihood estimator. Standard deviations given in parentheses.

Model	Expected value	MSE	$MSEP^*$
ϕ_1=0.9	$0.813_{(0.017)}$	$0.029_{(0.006)}$	$0.083_{(0.017)}$
ϕ_1=0.3	$0.261_{(0.018)}$	$0.031_{(0.004)}$	$0.024_{(0.002)}$
ϕ_1=0.5	$0.500_{(0.019)}$	$0.034_{(0.005)}$	$0.098_{(0.028)}$
ϕ_2=0.4	$0.309_{(0.020)}$	$0.041_{(0.006)}$	
ϕ_1=1.0	$0.922_{(0.022)}$	$0.046_{(0.002)}$	$0.092_{(0.026)}$
ϕ_2=-0.5	$-0.438_{(0.017)}$	$0.029_{(0.005)}$	
ϕ_1=-0.8	$-0.768_{(0.013)}$	$0.016_{(0.002)}$	$0.100_{(0.040)}$
ϕ_2=-0.8	$-0.739_{(0.013)}$	$0.017_{(0.003)}$	
ϕ_1=0.3	$0.275_{(0.018)}$	$0.033_{(0.005)}$	$0.090_{(0.028)}$
ϕ_2 =0.3	$0.202_{(0.021)}$	$0.043_{(0.006)}$	
ϕ_1=-0.5	$-0.464_{(0.019)}$	$0.035_{(0.005)}$	$0.089_{(0.038)}$
ϕ_2=-0.4	$-0.380_{(0.017)}$	$0.029_{(0.004)}$	
ϕ_1=1.9	$1.993_{(0.083)}$	$0.696_{(0.159)}$	*******
ϕ_2=-0.95	$-0.921_{(0.047)}$	$0.224_{(0.073)}$	
ϕ_1=0.3	$0.299_{(0.020)}$	$0.039_{(0.005)}$	$0.111_{(0.023)}$
ϕ_2=0.5	$0.390_{(0.019)}$	$0.038_{(0.007)}$	
ϕ_3=-0.5	$-0.420_{(0.017)}$	$0.029_{(0.004)}$	
ϕ_1=0.75	$0.697_{(0.022)}$	$0.048_{(0.007)}$	$0.125_{(0.052)}$
ϕ_2=-0.6	$-0.558_{(0.020)}$	$0.040_{(0.006)}$	
ϕ_3=0.6	$0.459_{(0.021)}$	$0.045_{(0.008)}$	
ϕ_1=0.4	$0.419_{(0.021)}$	$0.045_{(0.010)}$	$0.220_{(0.046)}$
ϕ_2=0.25	$0.158_{(0.024)}$	$0.058_{(0.012)}$	
ϕ_3=0.3	$0.307_{(0.020)}$	$0.040_{(0.010)}$	
ϕ_4=0.5	$0.366_{(0.027)}$	$0.074_{(0.023)}$	
ϕ_5=-0.65	$-0.503_{(0.024)}$	$0.059_{(0.011)}$	
ϕ_1=0.45	$0.393_{(0.020)}$	$0.041_{(0.005)}$	$0.181_{(0.024)}$
ϕ_2=-0.2	$-0.190_{(0.023)}$	$0.052_{(0.006)}$	
ϕ_3=0.25	$0.213_{(0.018)}$	$0.034_{(0.004)}$	
ϕ_4=-0.35	$-0.340_{(0.021)}$	$0.043_{(0.006)}$	
ϕ_5=0.35	$0.247_{(0.020)}$	$0.038_{(0.005)}$	
ϕ_1=-0.85	$-0.805_{(0.024)}$	$0.060_{(0.008)}$	$0.309_{(0.071)}$
ϕ_2=-0.6	$-0.627_{(0.028)}$	$0.077_{(0.011)}$	
ϕ_3=0.4	$0.215_{(0.033)}$	$0.111_{(0.021)}$	
ϕ_4=0.6	$0.364_{(0.034)}$	$0.113_{(0.014)}$	
ϕ_5=0.5	$0.310_{(0.027)}$	$0.075_{(0.010)}$	
ϕ_1=0.25	$0.210_{(0.020)}$	$0.040_{(0.005)}$	$0.635_{(0.308)}$
ϕ_2=-0.85	$-0.827_{(0.030)}$	$0.090_{(0.027)}$	
ϕ_3=0.9	$0.716_{(0.034)}$	$0.118_{(0.013)}$	
ϕ_4=-0.2	$-0.221_{(0.021)}$	$0.044_{(0.006)}$	
ϕ_5=0.65	$0.464_{(0.028)}$	$0.080_{(0.012)}$	

Table 2: Simulation results for some AR models. 100 replications of length 25. Bayesian Monte-Carlo importance sampling estimator. Standard deviations given in parentheses.

Model	Expected value	MSE	$MSEP^*$
θ_1=0.9	$0.914_{(0.013)}$	$0.017_{(0.004)}$	$0.032_{(0.009)}$
θ_1=0.3	$0.347_{(0.027)}$	$0.075_{(0.012)}$	$0.064_{(0.014)}$
θ_1=0.5	$0.457_{(0.025)}$	$0.062_{(0.010)}$	$0.113_{(0.024)}$
θ_2=0.4	$0.398_{(0.029)}$	$0.089_{(0.017)}$	
θ_1=1	$1.069_{(0.027)}$	$0.072_{(0.012)}$	$0.150_{(0.031)}$
θ_2=-0.5	$-0.552_{(0.032)}$	$0.100_{(0.018)}$	
θ_1=-0.8	$-0.815_{(0.016)}$	$0.025_{(0.004)}$	$0.210_{(0.130)}$
θ_2=-0.8	$-0.827_{(0.024)}$	$0.055_{(0.014)}$	
θ_1=0.3	$0.289_{(0.025)}$	$0.064_{(0.009)}$	$0.131_{(0.026)}$
θ_2=0.3	$0.276_{(0.028)}$	$0.080_{(0.019)}$	
θ_1=-0.5	$-0.534_{(0.022)}$	$0.050_{(0.008)}$	$0.160_{(0.035)}$
θ_2=-0.4	$-0.493_{(0.031)}$	$0.094_{(0.012)}$	
θ_1=1.9	$1.903_{(0.013)}$	$0.017_{(0.005)}$	$0.080_{(0.023)}$
θ_2=-0.95	$-0.951_{(0.013)}$	$0.017_{(0.006)}$	
θ_1=0.3	$0.372_{(0.024)}$	$0.055_{(0.010)}$	$0.231_{(0.055)}$
θ_2=0.5	$0.499_{(0.025)}$	$0.063_{(0.010)}$	
θ_3=-0.5	$-0.555_{(0.034)}$	$0.116_{(0.023)}$	
θ_1=0.75	$0.773_{(0.022)}$	$0.051_{(0.009)}$	$0.268_{(0.129)}$
θ_2=-0.6	$-0.710_{(0.031)}$	$0.096_{(0.012)}$	
θ_3=0.6	$0.713_{(0.031)}$	$0.097_{(0.013)}$	
θ_1=0.4	$0.510_{(0.027)}$	$0.071_{(0.008)}$	$0.237_{(0.039)}$
θ_2=0.25	$0.236_{(0.027)}$	$0.073_{(0.014)}$	
θ_3=0.3	$0.323_{(0.025)}$	$0.061_{(0.010)}$	
θ_4=0.5	$0.471_{(0.027)}$	$0.074_{(0.016)}$	
θ_5=-0.65	$-0.718_{(0.035)}$	$0.120_{(0.023)}$	
θ_1=0.45	$0.470_{(0.026)}$	$0.070_{(0.014)}$	$0.312_{(0.061)}$
θ_2=-0.2	$-0.252_{(0.033)}$	$0.107_{(0.018)}$	
θ_3=0.25	$0.218_{(0.034)}$	$0.117_{(0.017)}$	
θ_4=-0.35	$-0.343_{(0.037)}$	$0.140_{(0.027)}$	
θ_5=0.35	$0.318_{(0.039)}$	$0.151_{(0.022)}$	
θ_1= -0.85	$-0.905_{(0.025)}$	$0.060_{(0.009)}$	$0.222_{(0.034)}$
θ_2=-0.6	$-0.630_{(0.032)}$	$0.101_{(0.019)}$	
θ_3=0.4	$0.484_{(0.033)}$	$0.109_{(0.016)}$	
θ_4=0.6	$0.690_{(0.036)}$	$0.129_{(0.040)}$	
θ_5=0.5	$0.603_{(0.034)}$	$0.116_{(0.016)}$	
θ_1=0.25	$0.237_{(0.022)}$	$0.048_{(0.008)}$	$0.335_{(0.051)}$
θ_2=-0.85	$-0.959_{(0.027)}$	$0.072_{(0.010)}$	
θ_3=0.9	$0.916_{(0.031)}$	$0.096_{(0.026)}$	
θ_4=-0.2	$-0.261_{(0.026)}$	$0.065_{(0.010)}$	
θ_5=0.65	$0.701_{(0.032)}$	$0.104_{(0.022)}$	

Table 3: Simulation results for some MA models. 100 replications of length 25. Maximum-likelihood estimator. Standard deviations in parentheses.

Model	Expected value	MSE	$MSEP*$
θ_1=0.9	$0.772_{(0.017)}$	$0.027_{(0.004)}$	$0.055_{(0.012)}$
θ_1= 0.3	$0.308_{(0.023)}$	$0.051_{(0.007)}$	$0.050_{(0.012)}$
θ_1=0.5	$0.374_{(0.025)}$	$0.063_{(0.012)}$	$0.135_{(0.027)}$
θ_2=0.4	$0.281_{(0.026)}$	$0.068_{(0.014)}$	
θ_1=1	$1.008_{(0.026)}$	$0.068_{(0.013)}$	$0.102_{(0.040)}$
θ_2=-0.5	$-0.511_{(0.023)}$	$0.052_{(0.009)}$	
θ_1=-0.8	$-0.750_{(0.015)}$	$0.024_{(0.003)}$	$0.081_{(0.025)}$
θ_2=-0.8	$-0.712_{(0.017)}$	$0.027_{(0.004)}$	
θ_1=0.3	$0.268_{(0.022)}$	$0.047_{(0.006)}$	$0.111_{(0.024)}$
θ_2=0.3	$0.203_{(0.022)}$	$0.050_{(0.007)}$	
θ_1=1-0.5	$-0.496_{(0.020)}$	$0.039_{(0.006)}$	$0.126_{(0.028)}$
θ_2=-0.4	$-0.446_{(0.024)}$	$0.058_{(0.008)}$	
θ_1=1.9	$1.767_{(0.061)}$	$0.374_{(0.055)}$	$0.595_{(0.115)}$
θ_2=-0.95	$-0.780_{(0.034)}$	$0.119_{(0.026)}$	
θ_1=0.3	$0.322_{(0.022)}$	$0.049_{(0.011)}$	$0.139_{(0.028)}$
θ_2=0.5	$0.395_{(0.021)}$	$0.046_{(0.009)}$	
θ_3=-0.5	$-0.443_{(0.024)}$	$0.055_{(0.012)}$	
θ_1=0.75	$0.697_{(0.023)}$	$0.051_{(0.007)}$	$0.239_{(0.099)}$
θ_2=-0.6	$-0.595_{(0.025)}$	$0.062_{(0.009)}$	
θ_3=0.6	$0.507_{(0.020)}$	$0.041_{(0.008)}$	
θ_1= 0.4	$0.423_{(0.026)}$	$0.067_{(0.009)}$	$0.226_{(0.040)}$
θ_2=0.25	$0.148_{(0.024)}$	$0.057_{(0.015)}$	
θ_3=0.3	$0.251_{(0.021)}$	$0.043_{(0.008)}$	
θ_4=0.5	$0.309_{(0.028)}$	$0.081_{(0.015)}$	
θ_5=-0.65	$-0.518_{(0.026)}$	$0.067_{(0.015)}$	
θ_1=0.45	$0.410_{(0.024)}$	$0.062_{(0.009)}$	$0.275_{(0.060)}$
θ_2=-0.2	$-0.263_{(0.029)}$	$0.087_{(0.015)}$	
θ_3=0.25	$0.196_{(0.028)}$	$0.076_{(0.011)}$	
θ_4=-0.35	$-0.327_{(0.029)}$	$0.084_{(0.013)}$	
θ_5=0.35	$0.228_{(0.030)}$	$0.092_{(0.016)}$	
θ_1=-0.85	$-0.819_{(0.025)}$	$0.065_{(0.009)}$	$0.300_{(0.043)}$
θ_2=-0.6	$-0.600_{(0.032)}$	$0.104_{(0.017)}$	
θ_3=0.4	$0.339_{(0.028)}$	$0.080_{(0.015)}$	
θ_4=0.6	$0.475_{(0.029)}$	$0.084_{(0.017)}$	
θ_5=0.5	$0.409_{(0.023)}$	$0.051_{(0.014)}$	
θ_1=0.25	$0.200_{(0.020)}$	$0.040_{(0.007)}$	$0.332_{(0.047)}$
θ_2=-0.85	$-0.857_{(0.029)}$	$0.083_{(0.012)}$	
θ_3=0.9	$0.698_{(0.033)}$	$0.107_{(0.023)}$	
θ_4=-0.2	$-0.235_{(0.020)}$	$0.040_{(0.008)}$	
θ_5=0.65	$0.486_{(0.029)}$	$0.086_{(0.018)}$	

Table 4: Simulation results for some MA models. 100 replications of length 25. Bayesian Monte-Carlo importance sampling estimation. Standard deviations given in parentheses.

Model	Expected value	MSE	$MSEP$*
θ_1=0.5	$0.462_{(0.032)}$	$0.050_{(0.009)}$	$0.108_{(0.027)}$
θ_2=0.4	$0.359_{(0.033)}$	$0.053_{(0.014)}$	
θ_1=1	$1.075_{(0.027)}$	$0.037_{(0.008)}$	$0.200_{(0.051)}$
θ_2=-0.5	$-0.528_{(0.037)}$	$0.069_{(0.012)}$	
θ_1=-0.8	$-0.824_{(0.026)}$	$0.033_{(0.006)}$	$0.050_{(0.011)}$
θ_2=-0.8	$-0.812_{(0.019)}$	$0.018_{(0.004)}$	
θ_1=0.3	$0.346_{(0.037)}$	$0.068_{(0.020)}$	$0.083_{(0.025)}$
θ_2=0.3	$0.229_{(0.038)}$	$0.073_{(0.021)}$	
θ_1=-0.5	$-0.571_{(0.031)}$	$0.049_{(0.008)}$	$0.112_{(0.030)}$
θ_2=-0.4	$-0.525_{(0.038)}$	$0.071_{(0.010)}$	
θ_1=1.9	$1.843_{(0.012)}$	$0.008_{(0.003)}$	$0.051_{(0.014)}$
θ_2=-0.95	$-0.885_{(0.013)}$	$0.008_{(0.003)}$	

Table 5: Simulation results for some MA models. 50 replications of length 25. Bayesian estimation with normal quadrature. Standard deviations given in parentheses.

INSTITUTE OF ECONOMIC STUDIES WORKING PAPERS 1987-2000
Formerly Iceland Economic Papers Series

Editor Tryggvi Thor Herbertsson

A complete list of IoES Working Papers and Reprints can be accessed on our World Wide Web site at http://www.ioes.hi.is

W98:06 Thorarinn G. Petursson: Explaining the Term Stucture: The Expectations Hypothesis and Time Varying Term Premia

W98:07 Thorarinn G. Petursson: The Representative Houshold's Demand for Money in a Cointegrated VAR Model

W98:08 Tryggvi Thor Herbertsson and Gylfi Zoega: Trade Surpluses and Life-cycle Saving Behaviour

W98:09 Thorvaldur Gylfason: Privatization, Efficiency, and Economic Growth

W98:10 Tor Einarsson and Milton Marquis: Uncertain Labor Quality and Asset Pricing

W98:11 J. Michael Orzag: Actuarial Discretion vs. Rules in Regulation of Pension Funds

W98:12 J. Michael Orzag and Dennis Snower: From Unemployment Benefits to Unemployment Support Accounts

W98:13 Sveinn Agnarsson, Axel Hall, Tryggvi Thor Herbertsson, Sigurdur Ingolfsson, Gylfi Magnusson, and Gylfi Zoega: EMU and the Icelandic Labor Market

W98:14 Thorvaldur Gylfason: Exports, Inflation, and Growth

W99:01 Alison L. Booth, Yu-Fu Chen, and Gylfi Zoega: Hiring and Firing: A Tale of Two Thresholds

W99:02 Martin Paldam and Gert Tinggaard Svendsen: Is Social Capital an Effective Smoke Condenser?

W99:03 J. Michael Orszag and Dennis J. Snower: Anatomy of Policy Complementarities

W99:04 Thorvaldur Gylfason and Tryggvi Thor Herbertsson: Does Inflation Matter for Growth?

W99:05 Marco Biamchi, Bjorn R. Gudmundsson, and Gylfi Zoega: Iceland's Natural Experiment in Supply-side Economics

W99:06 Haukur C. Benediktsson, Tryggvi Thor Herbertsson, Gylfi Magnússon, and Marta G. Skúladóttir: Generational Accounts for Iceland

W99:07 Axel Hall and Jon Thor Sturluson: Testing a CGE Model: The Tax Free Year in Iceland as a Natural Experiment

W99:08 Edmund S. Phelps: Equilibrium and Disequilibrium in 20th Century 'Macro': With Attention to the Share Price Boom of the 1990s

W99:09 Kenneth F. Wallis: Macroeconometric Modelling

W99:10 Tryggvi Thor Herbertsson, Marta G. Skúladóttir, and Gylfi Zoega: Three Symptoms and a Cure: A Contribution to the Economics of the Dutch Disease

W99:11 Tryggvi Thor Herbertsson and J. Michael Orszag: Issues in European Pension Reforms: Supplementary Pensions

W99:12 Tryggvi Thor Herbertsson, J. Michael Orszag and Peter R. Orszag: Population Dynamics and Convergence in Fertility Rates

W99:13 Ragnar Arnason: Costs of Fisheries Management: Theoretical and Practical Implications

W99:14 Ragnar Arnason: Economic Instruments to Achieve Ecosystem Objectives in Fisheries Management

W99:15 Ragnar Arnason: Property Rights as a Means of Economic Organization

W00:01 Jerry Coakley, Ana-Maria Fuertes and Gylfi Zoega: Testing the Persistence and Structuralist Theories of Unemployment

W00:02 Thrainn Eggertsson: Norms in Economics – With Special Reference to Economic Development

W00:03 Thorvaldur Gylfason: Growing Apart

W00:04 Jon Danielsson: The Emperor has no Clothes: Limits to Risk Modelling

W00:05 Thorolfur Matthiasson: The Icelandic Debate on the Case for a Fishing Fee: A Non-Technical Introduction

W00:06 Willem H. Buiter: Is Iceland an Optimal Currency Area?

W00:07 Alison L. Booth and Gylfi Zoega: Why do Firms Invest in General Training? 'Good' Firms and 'Bad' Firms as a Source of Monopsony Power

W00:08 Eduard Hochreiter: "Exchange rate regimes and capital mobility: Issues and some lessons from central and eastern European applicant countries"

W00:09 Thorvaldur Gylfason: Fix or Flex? Alternative Exchange Rate Regimes in an Era of Global Capital Mobility

W00:10 Thorvaldur Gylfason: Natural Resources, Education and Economic Development

W00:11 Helgi Tomasson: Signal-noise Decomposition in Financial Markets: An Empirical Stochastic Process Analysis for Infrequent Trading

W00:12 Thorolfur Matthiasson: Changing Rules for Regulation of Icelandic Fisheries

W00:13 E. Tumusiime-Mutebile: Economic Reforms and their Impact in Uganda

W00:14 Sveinn Agnarsson: Productivity in Icelandic Fish Processing Industry 1985 – 1995: A Comparison of Methods

W00:15 Sveinn Agnarsson: Development of Efficiency in Icelandic Fish Processing Firms: A DEA Approach

W00:16 Jon Danielsson, Bjorn N. Jorgensen and Casper G. de Vries: Risk Management and Regulation in Incomplete Markets

W00:17 Ragnar Arnason, Gylfi Magnusson and Sveinn Agnarsson: The Norwegian Spring Spawning Herring Fishery: A Stylised Game Model

W00:18 Helgi Tomasson: Estimation of Correlations in Financial Markets when Trading is Infrequent

W00:19 Helgi Tomasson: Computations of Bayesian Estimators in ARMA Models