ISSN 1440-771X

MONASH University

Australia

Department of Econometrics
and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Parameter Estimation in Semi-Linear Models Using a Maximal Invariant Likelihood Function

Jahar L. Bhowmik and Maxwell L. King

July 2005

Working Paper 18/05

Parameter Estimation in Semi-Linear Models Using a Maximal Invariant Likelihood Function

Jahar L. Bhowmik and **Maxwell L. King***

Department of Econometrics and Business Statistics

Monash University, Victoria 3800, Australia

Abstract

In this paper, we consider the problem of estimation of semi-linear regression models. Using invariance arguments, Bhowmik and King (2001) have derived the probability density functions of the maximal invariant statistic for the nonlinear component of these models. Using these density functions as likelihood functions allows us to estimate these models in a two-step process. First the nonlinear component parameters are estimated by maximising the maximal invariant likelihood function. Then the nonlinear component, with the parameter values replaced by estimates, is treated as a regressor and ordinary least squares is used to estimate the remaining parameters. We report the results of a simulation study conducted to compare the accuracy of this approach with full maximum likelihood estimation. We find maximising the maximal invariant likelihood function typically results in less biased and lower variance estimates than those from full maximum likelihood.

Key words: Maximum likelihood estimation, nonlinear modelling, simulation experiment, two-step estimation.

JEL CLASSIFICATION: C2, C12

*Corresponding author. Tel.: +61-3-9905-2449; fax:+61-3-9905-8039

E-mail address: Max.King@BusEco.monash.edu.au (M.L. King)

1 Introduction

A major difficulty with full maximum likelihood estimation of multiparameter models is that it can result in poor estimates in some circumstances. There is a problem of potentially biased estimates arising from the joint estimation of multiple parameters. A good example is the estimate of the variance of the disturbances in the classical linear regression model. In this case, the maximum likelihood estimator is known to be biased and a simple correction is needed to make it unbiased in small samples. This is because the regression coefficients are nuisance parameters when it comes to estimating the variance. For further discussion of the problems of joint estimation of multiple parameters, see Neyman and Scott (1948), Anderson (1970) and Cox and Hinkely (1974). There is a vast amount of literature on the satisfactory handling of nuisance parameters, see for example, Fraser (1967), Kalbfleisch and Sprott (1970, 1973), Bellhouse (1978), King (1983), Barndorff-Nielsen (1983), Lehmann (1986), Cox and Reid (1987), Tunnicliffe Wilson (1989), McCullagh and Tibshirani (1990), Ara and King (1993, 1995), Ara (1995), and Laskar and King (1998, 2001).

One approach that has received a good deal of attention in the literature is the concept of the marginal likelihood which was first introduced by Fraser (1967), and further developed by Kalbfleisch and Sprott (1970). The main idea is to transform the data vector to another random vector, a subvector of which has a likelihood (marginal likelihood) that only involves the parameters of interest and the remainder of which contains no information about those parameters. There is a lot of evidence in the literature that the use of marginal likelihood methods can produce more accurate estimates and, in particular, less biased estimates. See for example Cooper and

Thompson (1977), Kitanidis and Vomvoris (1983), Kitanidis (1983, 1987), Hoeksema and Kitanidis (1985) and Kitanidis and Lane (1985), Cordus (1986), Tunnicliffe Wilson (1989), Bellhouse (1991), Shephard (1993), Ara (1995), Ara and King (1993, 1995), Laskar and King (1997) and Rahman and King (1998).

The use of invariance arguments has been a useful method for dealing with some of the problems caused by nuisance parameters, particularly for hypothesis testing. The approach involves noting that the testing problem is invariant to a certain class of transformations on the observed data vector and then requiring the chosen test to also be invariant to such transformations. A key device for test construction is the maximal invariant statistic. It is a vector function of the data vector that takes the same value for data vectors that can be connected by a transformation and different values for those data vectors that cannot be connected by a transformation. Thus the class of all invariant test statistics corresponds to the class of functions of the maximal invariant. This allows us to treat the maximal invariant as the observed data when designing a new test. The density function of the maximal invariant can be treated as a likelihood for this purpose. This function is known as the maximal invariant likelihood (MIL) function.

Ara (1995) showed that the marginal likelihood function and the likelihood of the maximal invariant statistic are equivalent in the case of nonspherical disturbances in the linear regression model. In the context of a linear regression model with a non-linear additive component, Bhowmik and King (2001) derived a MIL function for the non-linear component. The purpose of this paper is to compare maximum MIL (MMIL) and full maximum likelihood approaches to the estimation of parameters in the non-linear component. Using the MIL function to estimate these parameters results in a two-step process. First the non-linear component parameters are estimated by maximising the MIL. Then the non-linear component, with the parameter values replaced by estimates, is treated as a regressor and ordinary least squares is used to estimate the remaining parameters. Alternatively, the full likelihood of the complete model can be maximised to obtain the standard maximum likelihood estimates. The MMIL estimator might be expected to be superior to the full likelihood estimator given the evidence in the literature outlined above. We derive two MIL functions for two different models and these functions will be denoted as MIL1 and MIL2, where MIL1 stands for the linear model with a general non-linear component and MIL2 for a linear model with a regressor which is a non-linear function of unknown parameter(s).

The plan of this paper is as follows. In Section 2 we derive the likelihood functions (full likelihood and MIL) for the different non-linear models. Monte Carlo experiments to investigate the performance of different ML estimators in the context of non-linear parameters are outlined and reported in Section 3. Finally, concluding remarks are made in Section 4.

2 Theory

Our interest is in the following semi-linear model

$$y = X_1 \beta_1 + g(X_2, \beta_2) + u,\ u \sim N(0, \sigma^2 I_n) \qquad (2.1)$$

where y is an $n \times 1$ vector, X_1 is an $n \times q$ nonstochastic matrix, X_2 is an $n \times p$ nonstochastic matrix of n observations on p variables and $g(X_2, \beta_2)$ is a non-linear function of the $r \times 1$ parameter vector β_2 and X_2. Note r and p are different for flexibility. Bhowmik and King (2001) derived the density function of a maximal invariant statistic for the non-linear component of (2.1). This can be treated as a likelihood function for the parameter vector β_2 in order to construct the MIL1 function. The MIL1 function is

$$f(w) = \frac{1}{2}\Gamma(\frac{m}{2})\pi^{-m/2}\exp\{c(w,\beta_2)\}\ \{{}_1F_1[\frac{m}{2},\frac{1}{2},\frac{a^2(w,\beta_2)}{2}] +$$

$$\sqrt{2}\,a(w,\beta_2)\ \eta\ {}_1F_1[\frac{1+m}{2},\frac{3}{2},\frac{a^2(w,\beta_2)}{2}]\} \qquad (2.2)$$

where

$$a(w,\beta_2) = w'Pg^*(X_2,\beta_2)\,, \qquad (2.3)$$

$$g^*(X_2,\beta_2) = \frac{g(X_2,\beta_2)}{\sigma}\,, \qquad (2.4)$$

$$c(w,\beta_2) = b(w,\beta_2) - \frac{a^2(w,\beta_2)}{2} = -\frac{1}{2}g^{*\prime}(X_2,\beta_2)M_1g^*(X_2,\beta_2)\,, \qquad (2.5)$$

$$w = z/(z'z)^{1/2}\,, \qquad (2.6)$$

$$z = Py\,, \qquad (2.7)$$

$$\eta = \frac{\Gamma(\frac{1+m}{2})}{\Gamma(\frac{m}{2})}\,, \qquad (2.8)$$

${}_1F_1[.,.,.]$ is the confluent hypergeometric function, which has the form

$${}_1F_1[c,d,\delta] = 1 + \frac{c\delta}{d} + \frac{c(c+1)}{d(d+1)}\frac{\delta^2}{2!} + \ldots = \sum_{k=0}^{\infty}\frac{(c)_k}{(d)_k}\frac{\delta^k}{k!}\,, \qquad (2.9)$$

P is an $m \times n$ matrix such that $PP' = I_m$, $P'P = M_1$ and $m = n - q$.

Let us also consider the following slightly more specific semi-linear model,

$$y = X_1\beta_1 + \beta_2 g(X_2, \beta_3) + u,\ u \sim N(0, \sigma^2 I_n), \tag{2.10}$$

where X_1 is an $n \times q$ nonstochastic matrix, X_2 is an $n \times p$ nonstochastic matrix and $g(X_2, \beta_3)$ is a non-linear function of β_3 and X_2. Bhowmik and King (2002) derived the density function of a maximal invariant statistic for the non-linear component of this model. The MIL2 function is

$$\begin{aligned} f(d) &= \left(2\pi\sigma^2\right)^{-\frac{n-q-1}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(d'd\right)\right] \\ &= \left(2\pi\sigma^2\right)^{-\frac{n-q-1}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(z'M_g(\beta_3)z\right)\right], \end{aligned} \tag{2.11}$$

where

$$d = Q(X_2, \beta_3)Py \sim N(0, \sigma^2 I_{n-q-1}), \tag{2.12}$$

$Q(X_2, \beta_3)$ is an $(n-q-1) \times m$ matrix such that $Q(X_2, \beta_3)Q(X_2, \beta_3)' = I_{n-q-1}$, $Q(X_2, \beta_3)'Q(X_2, \beta_3) = M_g(\beta_3)$,

and

$$M_g(\beta_3) = I - g(X_2, \beta_3)\{g(X_2, \beta_3)'g(X_2, \beta_3)\}^{-1} g(X_2, \beta_3)'. \tag{2.13}$$

Our aim is to use these two likelihood functions for the estimation of the non-linear parameters. A maximal invariant is a random vector and therefore the use of its density as a likelihood means that resultant estimators will have the usual asymptotic properties that have been demonstrated for the classical likelihood (see Lehmann 1983, Stuart and Ord 1991 and Ara 1995).

As mentioned earlier, the two-step estimation process involves estimating the non-linear component parameters by maximising the MIL function or equivalently the log of MIL. For the MIL2 function (2.11), the log likelihood function is

$$L_2 = -\frac{n-q-1}{2}\log(2\pi) - \frac{n-q-1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left(z'M_g(\beta_3)z\right).$$

Now

$$\frac{\partial L_2}{\partial \sigma^2} = -\frac{n-q-1}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\left(z'M_g(\beta_3)z\right),$$

and $\frac{\partial L_2}{\partial \sigma^2} = 0$ implies $\hat{\sigma}^2 = \frac{z'M_g(\beta_3)z}{n-q-1}$. Thus, replacing $\hat{\sigma}^2$ in the log likelihood, we have

$$L_2 = -\frac{n-q-1}{2}\log(2\pi) - \frac{n-q-1}{2}\log\left(\frac{z'M_g(\beta_3)z}{n-q-1}\right) - \frac{n-q-1}{2}.$$

Therefore, maximising L_2 is equivalent to minimising $z'M_g(\beta_3)z$ with respect to β_3.

3 Empirical comparisons

In order to compare the small sample performance of the three estimators, we conducted a simulation study outlined below.

3.1 Experimental design

We evaluated the different estimation methods on the following three semi-linear models, namely

$$K_t = \gamma V_t + \frac{\beta}{R_t - \alpha} + u_t ,\ u_t \sim IN(0,\sigma^2) , \tag{3.1}$$

$$Y_t = f(X_t,\theta) = \theta_1 X_{1t} + \theta_2 X_{2t} + \theta_4 \exp(\theta_3 X_{3t}) + u_t ,\ u_t \sim IN(0,\sigma^2) , \tag{3.2}$$

$$C_t = \alpha U_t + \beta W_t^{\gamma} + u_t ,\ u_t \sim IN(0,\sigma^2) ,\ t = 1,2,\ldots,n. \tag{3.3}$$

Model (3.1) is a non-linear money demand function used by Konstas and Khouja (1969), where

K_t = quantity of money demanded,

V_t = national income,

R_t = rate of interest,

γ , β and α are three unknown parameters such that $0 < \alpha < \infty$, $\beta > 0$ and $\gamma > 0$. Model (3.2) was given by Gallant (1975), where X_{1t} , X_{2t} and X_{3t} are three input variables, $Y_t = f(X_t,\theta)$ is the output variable, and θ_1 , θ_2 , θ_3 and θ_4 are unknown parameters. Model (3.3) is a modified model of the general consumption function from Greene (1997), where

W_t = aggregate income,

C_t = consumption,

U_t = regressor of independent random variables from $N(0,1)$,

α , β and γ are three unknown parameters such that $\alpha > 0$, $0 < \beta < 1$ and $\gamma > 0$.

In Greene's model, U_t is a vector of ones but in our case U_t is an $n \times 1$ vector of independent random variables from $N(0,1)$. We made this modification to avoid problems caused by an ill-conditioned non-linear model. It is worth noting that ill-conditioning will occur in the models for near zero values of β in (3.1), θ_4 in (3.2) and β in (3.3).

For model (3.1), we used generated data to construct the design matrix with V_t and R_t being independent observations from the [0,1] uniform distribution. The estimates based on (i) full maximum likelihood (FML), (ii) maximum MIL1 (MMIL1) and maximum MIL2 (MMIL2) when

$\left(\gamma, \beta, \alpha, \sigma^2\right) =$ (0.5, 0.1, 0.05, 0.05), (0.3, 0.1, 0.03, 0.05), (0.3 0.2, 0.015, 0.25), (0.7, 0.5, 0.15, 0.05), (0.3, 0.1, 0.03, 0.75), (0.5, 0.1, 0.5, 0.75), (0.5, 0.05, 0.01, 0.25), (0.5, 0.1, 0.5, 0.05), (1.5, 0.15, 0.2, 0.25), (0.25, 0.05, 0.01, 0.25)

were used for comparison. At an early stage in our simulations, we identified a problem with local maxima of the various likelihood functions and so used a range of starting values for γ, β, α and σ^2 in the respective optimisation procedures. For model (3.1) we used five sets of starting values for γ, β, α and σ^2 and these were (0.05, 0.03, 0.02, 0.05), (0.2, 0.05, 0.05, 0.2), (0.5, 0.1, 0.07, 0.5), (1, 0.65, 0.15, 0.75) and (1.5, 0.85, 0.25, 0.95) for FML, (β, α) = (0.1, 0.05), (0.3, 0.15), (0.5, 0.25), (0.7, 0.35) and (1, 0.5) for MMIL1 and α = 0.05, 0.1, 0.3, 0.4, and 0.5 for MMIL2.

For model (3.2), Gallant (1975), used simulated data for X_{1t}, X_{2t} and X_{3t}. In our study,

X_{1t} was independently generated from $N(0,1)$,

X_{2t} was independently generated from $N(0,1)$,

X_{3t} was independently generated from the [0,1] uniform distribution.

Data was generated for nine different sets of values for $\theta_1, \theta_2, \theta_3$ θ_4 and σ^2, namely

(0.5, 0.25, 0.01, 0.25, 0.25), (0.75, 0.5, 0.1, 0.5, 0.9), (0.03, 1.02, -1.1, -0.5, 0.25), (0.15, 0.5, 0.15, 0.35, 0.5), (-0.05, 1.02, -0.95, -0.5, 0.25), (-0.05, -1.1, 0.5, -0.5, 0.95), (-0.025, 1.1, -1.1, -0.5, 0.25), (-0.5, -1.1, 0.5, -0.5, 0.5), (-0.75, -0.5, 0.5, -0.75, 0.25).

Again we identified a problem with local maxima and used five different sets of starting values in the respective optimisation procedures to overcome this problem. The five sets of starting values for θ_1, θ_2, θ_3 θ_4 and σ^2 were (-0.85, -1, -1.5, -1.2, 0.05), (-0.5, -0.75, -0.85, -0.9, 0.25), (0.5, -0.25, -0.5, -0.4, 0.5), (0.75, 0.5, 0.5, 0.5, 0.75) and (1.5, 1, 1.5, 1.2, 0.9) for FML, (θ_3, θ_4) = (-0.5, -0.5), (-0.15, -0.2), (0.15, 0.15), (0.25,0.25) and (0.75, 0.75) for MMIL1 and $\theta_3 = -1, -0.5, 0.5, 1$ and 1.5 for MMIL2.

For model (3.3), U_t was independently generated from $N(0,1)$ and W_t was generated from the [0,1] uniform distribution. Data for C_t was generated for seven different sets of values for α, β, γ and σ^2 namely

(3.5, 0.5, 1.15, 0.05), (1.5, 0.25, 1.155, 0.05), (0.5, 0.25, 1.5, 0.05), (1.5, 0.25, 1.1, 0.05), (1.5, 0.25, 1.1, 0.01), (0.5, 0.25, 0.25, 0.25), (1.5, 0.75, 0.5, 0.25).

To overcome problems with local maxima, we used five different sets of starting values for α, β, γ and σ^2 and these were (0.25, 0.05, 0.25, 0.05), (0.75, 0.25, 0.5, 0.25), (1.2, 0.5, 0.95, 0.5), (1.75, 0.75, 1.25, 0.75) and (3.5, 0.9, 1.75, 0.95) for FML, (β, γ) = (0.1, 0.2), (0.3, 0.5), (0.5, 0.8), (0.7, 1) and (0.95, 1.5) for MMIL1 and γ = 0.25, 0.5, 0.75, 1 and 1.5 for MMIL2.

For each case, 2000 iterations were used to simulate the distributions of the estimators. We used two sample sizes, $n = 30$ and $n = 60$. In order to maximise the likelihood functions, the Gauss (see Aptech 1995) Co-optimisation routine was used.

From the simulations, we recorded estimated bias, standard deviation, mean squared error and quantiles (5%, 50% and 95%) of the three different estimators (FML, MMIL1 and MMIL2) of the non-linear parameters of the three different models.

3.2 The question of existence of second-order moments of estimators

There is an issue of whether the second-order moments of the estimators exist. If they do not exist then our estimates of SD and MSE are meaningless because they will be finite estimates of infinity. The possibility of the estimator having an infinite variance can be revealed by running the simulations for a range of different numbers of iterations. An infinite variance would be reflected in the estimate of SD increasing with the number of iterations. We examined this by running simulations for different numbers of iterations for each of the models, namely 500, 2000, 5000, 7000, 10,000 and 15,000. In this simulation experiment, we used only one set of values for the parameters for each of the models. These were $\gamma = 0.5$, $\beta = 0.1$, $\alpha = 0.05$, $\sigma^2 = 0.05$ for model (3.1); $\theta_1 = 0.5$, $\theta_2 = 0.25$, $\theta_3 = 0.01$, $\theta_4 = 0.25$, $\sigma^2 = 0.25$ for model (3.2) and $\alpha = 0.5$, $\beta = 0.25$, $\gamma = 1.5$, $\sigma^2 = 0.05$ for model (3.3). Sample sizes of $n = 30$ and $n = 60$ were used.

The resultant estimates of the SD are presented in Table 1. They show that SDs for model (3.1) for each of the methods, (FML, MMIL1 and MMIL2) are stable for different numbers of iterations and for both sample sizes ($n = 30$ and $n = 60$). For model (3.2), we notice that from the use of the FML, MMIL1 and MMIL2 methods, estimates of SD for the non-linear parameter decrease slightly when the number of iterations is increased, especially for the larger numbers of iterations (7000, 10,000 and 15,000). Similarly for model (3.3), we observe a small decrease in SD estimates for both of the sample sizes when the number of iterations is increased. Therefore, the SD results in Table 1 confirm that the second-order moments of the estimators exist, at least for the models we considered.

3.3 Simulation results

Estimated bias, standard deviation, mean squared error (MSE), quantiles (5%, 50%, and 95%) of the three different estimators (FML, MMIL1 and MMIL2) of the non-linear parameters of models (3.1), (3.2) and (3.3), for selected parameter combinations are presented in Tables 2-4. The following is a discussion of the full set of results.

For model (3.1), the results show that both FML and MMIL (MMIL1 and MMIL2) estimators have little bias. In cases where the FML estimator has a small bias, it is reduced by the use of the MMIL1 and MMIL2 estimators, especially for the parameters in the non-linear component and for $n = 30$. Using an MMIL estimator in place of the FML estimator can reduce bias by up to 99.8%.

The results also confirm that the MMIL estimators (MMIL1 and MMIL2) have smaller SDs than the FML estimator, particularly for the parameters in the non-linear component. In some cases, the MMIL2 estimator has a smaller SD than the MMIL1 estimator, particularly when $n = 60$. For $n = 30$, the MMIL1 estimator is typically better than the MMIL2 estimator.

When MSE is considered, we see in general the MMIL estimators are better than the FML estimator, especially for the parameters in the non-linear component and when $n = 30$. Given that the bias and SD both decrease when an MMIL estimator is used in place of the FML estimator, it is no surprise to see that the MSE also decreases. Sometimes for the parameters in the non-linear component, we observe up to a 99.9% reduction in MSE when an MMIL estimator is used. The MMIL2 estimator is better than the MMIL1 estimator, particularly when $n = 60$. On the other hand, for $n = 30$ in most cases, the MMIL1 estimator is better than the MMIL2 estimator with respect to MSE.

An analysis of the quantile results reveals that the differences from the median (50th percentile) to 5th percentile and 95th percentile for the MMIL estimators are less than for those of the FML estimator, especially for the parameters in the non-linear component and for $n = 30$. The $100(1-\alpha)\%$ percentile range (PR) of an estimator is calculated as $Q_{1-\alpha/2} - Q_{\alpha/2}$, where $Q_p =$ pth quantile. We observe that the 90% PRs for the MMIL estimators are less than those for the FML estimator. The PR values show that sometimes, the middle 90% parameter estimates are up to 98.4% more tightly distributed for the two-step MMIL method compared to FML estimates, especially for the parameters in the non-linear component and for $n = 30$. If we observe the difference between the 50th percentile (median) and the true value of the parameters, we see that the resulting difference is generally less for the MMIL estimators than for the FML estimator.

Table 3 shows selected results for model (3.2). In this case, in general for the parameters in the non-linear component, bias is less for the MMIL estimators than for the FML estimator for $n = 30$. We observe up to an 82% reduction in bias from using an MMIL estimator in place of the FML estimator. The reduction is most noticeable for the parameters in the non-linear component and when $n = 30$. Notable exceptions for the MMIL1 estimator occur when θ_3 is very small. The MMIL2 estimator typically results in a higher reduction in bias compared to the FML estimator when $n = 30$. The MMIL2 estimator is often better than the MMIL1 estimator for the parameters in the non-linear component. For near zero values of θ_3 and large values of σ^2, we observe more biased estimates of the parameters in the non-linear component.

Results reported in Table 3 show that in most cases, an MMIL estimator (MMIL1 or MMIL2) has a lower SD than the FML estimator for both sample sizes, particularly for the parameters in the non-linear component. For the linear parameters, sometimes we have a reduction in SD from using an MMIL estimator in place of the FML estimator but in most cases, the SDs of the MMIL estimators and the FML estimator are almost the same.

We see that bias and SD results for the parameters in the non-linear component both decrease when the MMIL estimator is used and as a result, the MSE also decreases. The results in Table 3 show that for the parameters in the non-linear component and for $n = 30$, we obtain up to a 61.8% reduction in MSE from the use of an MMIL estimator in place of the FML estimator. When θ_3 and θ_4 both are positive large numbers or are both negative then for $n = 30$ with respect to MSE, the MMIL estimators are better than the FML estimator. The MMIL2 estimator is often better than the MMIL1 estimator in this regard, particularly for $n = 30$. For some values of θ_3 and θ_4 ($\theta_3 = -0.95$ and $\theta_4 = -0.5$), we do not get a reduction in MSE from using the MMIL2 estimator in place of the FML estimator.

The quantile results in Table 3 show that for the parameters in the non-linear component, the 90% PRs for the MMIL estimators are less than those of the FML estimator. The 90% PR values generally reveal that the middle 90% estimated values of the parameters are up to 40.6% more tightly distributed for the MMIL estimators compared to the FML estimator. The difference between the median and true value of the parameters is almost always less for the MMIL estimators than for the FML estimator, especially for the parameters in the non-linear component. The MMIL estimates of non-linear component parameters are more concentrated around their true value than are the FML estimates.

Table 4 shows selected simulation results for model (3.3). Again we have a reduction in bias for the MMIL estimators in comparison to the FML estimator, especially for the parameters in the non-linear component and for small sample sizes. The overall bias results show that there can be up to a 81.8% reduction in bias from the use of an MMIL estimator (MMIL1 or MMIL2) in place of the FML estimator. Extremes in bias reduction occur when the traditional FML estimator is more biased. However when

$n = 60$, we do not have a reduction in bias from using the MMIL estimators, particularly for the parameters in the non-linear component. For this model, the MMIL1 estimator is better than the MMIL2 estimator at reducing bias, particularly when $n = 30$.

When $n = 30$, we have a sizable reduction in SD from the use of the MMIL estimators in place of the FML estimator. However, for $n = 60$ in most cases, especially for the parameters in the non-linear component, we do not have a reduction in SD for the MMIL estimators compared to the FML estimator. Among the MMIL estimators, the MMIL1 estimator is better than the MMIL2 estimator when SD is considered.

For model (3.3) with $n = 30$, both bias and SD are almost always decreased when an MMIL estimator is used especially for the parameters in the non-linear component. As a result, we can obtain up to a 79.9% reduction in MSE for $n = 30$ from the use of an MMIL estimator (MMIL1 or MMIL2) in place of the FML estimator. In many cases, the MMIL1 estimator has a smaller MSE than the MMIL2 estimator for both sample sizes. For the α parameter and for $n = 60$, the MMIL estimators are typically better than the FML estimator with respect to MSE, but for the non-linear parameters, the differences for $n = 60$ are very small.

When $n = 30$, the quantile results in Table 4 show that the 90% PRs for the MMIL estimators are smaller than those of the FML estimator, particularly for the parameters in the non-linear component. The difference between these two percentiles show that the middle 90% estimated values of the parameters are up to 32.4% more tightly distributed for the MMIL estimators than for the FML estimator.

4 Concluding remarks

The results for the three models show that overall the MMIL estimators are less biased than the traditional FML estimator, particularly for the parameters in the non-linear component and for small sample sizes. The MMIL estimators (MMIL1 and MMIL2) typically have smaller SDs than the FML estimator. Similarly, with respect to MSE, the MMIL estimators are better than the FML estimator, especially for the parameters in the non-linear component and for small sample sizes. The quantile results of the estimators give us a closer view of the two methods. The estimated values from the MMIL estimators are more concentrated around their true parameter value than for the FML estimator. Therefore, we can conclude that when there is a measurable bias in the FML estimates, the MMIL estimators will help to reduce this bias, for small sample sizes and particularly for the parameters in the non-linear component. However, for all the models and for $n = 60$, sometimes the FML estimator and the MMIL estimators are nearly the same, especially for the parameters in the non-linear component.

When we compare the MMIL1 and MMIL2 estimators, their performance is more or less equal. However, MIL1 is a complicated mathematical function and, for our simulations, the MMIL1 estimator was more time consuming to apply. Therefore, we recommend the MMIL2 estimator ahead of the MMIL1 estimator because it is straightforward and more easily applied.

Acknowledgement

The research reported in this paper forms part of the first author’s Ph.D. thesis. This research was supported by the Monash Research Graduate School and in part by an

ARC grant. We are grateful to Don Poskitt, Rob J. Hyndman, Simone D. Grose and Baki Billah for helpful suggestions. An earlier version of this paper was presented at the 2002 *Econometric Society Australasian Meeting* held in Brisbane, Australia.

References

Anderson, E.B., 1970. Asymptotic properties of conditional maximum-likelihood estimators. J. Roy. Statist. Soc. B 32, 283-302.

Aptech System, 1995. Gauss: System and Graphics Manual. Vol. 1, Maple Vally, WA.

Ara, I., 1995. Marginal Likelihood Based Tests of Regression Disturbances. Ph.D. Thesis, Monash University, Clayton, Melbourne.

Ara, I., King, M.L., 1993. Marginal likelihood based tests of regression disturbances. Paper presented at the 1993 Australasian Meeting of the Econometric Society, Sydney.

Ara, I., King, M.L., 1995. Marginal likelihood based tests of a subvector of the parameter vector of linear regression disturbances. Proceedings of the Econometrics Conference at Monash University, C.S. Forbes, P. Kofman and T.R.L. Fry (eds.), Monash University, 69-106.

Barndorff-Nielsen, O.E., 1983. On a formula for the distribution of the maximum likelihood estimator. Biometrika 70, 343-365.

Bellhouse, D.R., 1978. Marginal likelihood methods for distributed lag models. Statist. Hefte 19, 2-14.

Bellhouse, D.R., 1991. Marginal and approximate conditional likelihoods for sampling on successive occasions. Survey Methodology 17, 69-78.

Bhowmik, J.L., King M.L., 2001. Deriving tests of the regression model using the density function of a maximal invariant. Paper presented at the 2001 Australasian Meeting of the Econometric Society, Auckland, New Zealand.

Bhowmik, J.L., King M.L., 2002. Parameter estimation in linear models with a non-linear component using a maximal invariant likelihood function. Paper presented at the 2002 Australasian Meeting of the Econometric Society, Brisbane, Australia.

Cooper, D.R., Thompson, R., 1977. A note on the estimation of parameters of the autoregressive-moving average process. Biometrika 64, 625-628.

Corduas, M., 1986. The Use of Marginal Likelihood in Testing for Correlation in Time Series Regression. M. Phil. Thesis, University of Lancaster.

Cox, D.R., Hinkley, D.V., 1974. Theoretical Statistics. Chapman and Hall, London.

Cox, D.R., Reid, N., 1987. Parameter orthogonality and approximate conditional inference (with discussion). J. Roy. Statist. Soc. B 49, 1-39.

Fraser, D.A.S., 1967. Data transformations and the linear model. Ann. Math. Statist. 38, 1456-1465.

Gallant, A.R., 1975. Non-linear regression. Amer. Statist. 29, 75-81.

Greene, W.H., 1997. Econometric Analysis. Prentice Hall, New York.

Hoekseman, R.J., Kitanidis, P.K., 1985. Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution to the inverse problem. Water Resources Research 21, 825-836.

Kalbfleisch, J.D., Sprott, D.A., 1970. Application of likelihood methods to models involving large numbers of parameters. J. Roy. Statist. Soc. B 32, 175-208.

Kalbfleisch, J.D., Sprott, D.A., 1973. Marginal and conditional likelihoods. Sankhya A 35, 311-328.

King, M.L., 1983. Testing for autoregressive against moving average errors in the linear regression model. J. Econometrics 21, 35-51.

Kitanidis, P.K., 1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. Water Resources Research 19, 909-921.

Kitanidis, P.K., 1987. Parametric estimation of covariances of regionalized variables. Water Resources Bulletin, American Water Resources Association 23, 557-567.

Kitanidis, P.K., Lane, R.W., 1985. Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method. J. Hydrology 79, 59-71.

Kitanidis, P.K., Vomvoris, E.G., 1983. A geostatistical approach to the inverse problem in ground water modelling (steady state) and one-dimensional simulation. Water Resources Research 19, 677-690.

Konstas, P., Khouja, M.W., 1969. The Keynesian demand-for-money function: Another look and some additional evidence. J. Mon. Cred. Bank. 1, 765-777.

Laskar, M.R., King, M.L., 1998. Estimation and testing of regression disturbances based on modified likelihood functions. J. Statist. Plann. Inference 71, 75-92.

Laskar, M.R., King, M.L., 2001. Modified likelihood and related methods for handling nuisance parameters in the linear regression model. In A.K.M.E. Saleh ed. Data Analysis from Statistical Foundations, Nova Science Publisher, Inc. Huntington, New York, 119-142.

Lehmann, E.L., 1983. Theory of Point Estimation. Wiley, New York.

Lehmann, E.L., 1986. Testing Statistical Hypotheses. Second Edition, Wiley, New York.

McCullagh, P., Tibshirani, R., 1990. A simple method for the adjustment of profile likelihoods. J. Roy. Statist. Soc. B 52, 325-344.

Neyman, J., Scott, E.L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16, 1-32.

Rahman, S., King, M.L., 1998. Marginal likelihood score based tests of regression disturbances in the presence of nuisance parameters. J. Econometrics 82, 81-106.

Shephard, N., 1993. Maximum likelihood estimation of regression models with stochastic trend components. J. Amer. Statist. Assoc. 88, 590-595.

Stuart, A., Ord, K., 1991. Kendall's Advanced Theory of Statistics. Vol. 2, Oxford University Press, New York.

Tunnicliffe Wilson, G., 1989. On the use of marginal likelihood in time series model estimation. J. Roy. Statist. Soc. B 51, 15-27.

Table 1 Estimates of SD for the non-linear parameters of three different models for different numbers of iterations and for two different sample sizes

Sample size		Number of iterations						
		500	1000	2000	5000	7000	10000	15000
		Model 1: $M_t = \gamma Y_t + \frac{\beta}{R_t - \alpha} + u_t$; SD of estimates of α						
	FML	0.00049	0.00049	0.00049	0.00049	0.00049	0.00049	0.00049
30	MMIL1	0.00050	0.00061	0.00057	0.00055	0.00055	0.00050	0.00050
	MMIL2	0.00049	0.00050	0.00049	0.00049	0.00048	0.00048	0.00048
	FML	0.00009	0.00009	0.00009	0.00009	0.00009	0.00008	0.00008
60	MMIL1	0.00012	0.00012	0.00012	0.00012	0.00013	0.00012	0.00012
	MMIL2	0.00008	0.00009	0.00009	0.00009	0.00008	0.00008	0.00008
		Model 2: $Y_t = \theta_1 X_{1t} + \theta_2 X_{2t} + \theta_4 \exp(\theta_3 X_{3t}) + u_t$; SD of estimates of θ_3						
	FML	6.96426	6.93151	6.89750	7.06477	6.95384	6.71809	6.61518
30	MMIL1	7.10115	6.96940	6.81410	6.17519	6.17230	6.16213	6.16185
	MMIL2	7.81624	7.79325	8.07040	7.39501	7.35290	7.34501	7.34463
	FML	1.95792	1.95177	1.95035	1.88706	1.82699	1.83637	1.83559
60	MMIL1	1.74588	1.74496	1.74383	1.72274	1.71955	1.71015	1.71002
	MMIL2	1.86419	1.79879	1.88635	1.89355	1.88940	1.88637	1.86428
		Model 3: $C_t = \alpha U_t + \beta W_t^{\gamma} + u_t$; SD of estimates of γ						
	FML	5.23450	4.99587	4.98870	5.00981	4.89432	4.88754	4.78921
30	MMIL1	2.18325	2.28940	2.20283	2.40910	2.35460	2.34981	2.34672
	MMIL2	4.39880	4.35099	4.32676	4.31617	4.33603	4.32175	4.31568
	FML	2.10130	1.98715	1.91666	1.97758	1.93872	1.94116	1.93987
60	MMIL1	1.50103	1.49665	1.51905	1.51505	1.50655	1.50310	1.50265
	MMIL2	2.31661	2.30525	2.21115	2.20402	2.16357	2.06473	2.05153

Table 2 Estimated bias, standard deviation, mean squared error, and quantiles (5%, 50% and 95%) of estimators for model (3.1) based on FML and two step MMIL (MMIL1 and MMIL2) estimators

Parameter		$n=30$						$n=60$					
		Bias*	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias*	SD	MSE	Q_5	Q_{50}	Q_{95}
$\gamma=.5$	FML	1.08	.1116	.01246	.316	.497	.685	.36	.0500	.0025	.417	.499	.583
	MMIL1	.070	.0189	.00036	.363	.500	.531	.18	.0289	.0024	.420	.500	.582
	MMIL2	1.40	.0850	.00722	.400	.500	.638	.04	.0508	.0026	.418	.499	.584
$\beta=.1$	FML	.070	.0054	2.9×10^{-5}	.091	.100	.109	.03	.0020	4.2×10^{-6}	.097	.100	.103
	MMIL1	.020	2.9×10^{-5}	1.2×10^{-9}	.092	.100	.102	.60	.0001	3.7×10^{-7}	.096	.100	.103
	MMIL2	.008	.0055	.00003	.091	.100	.109	.22	.0020	4.2×10^{-6}	.097	.100	.103
$\alpha=.05$	FML	.009	.0005	2.5×10^{-7}	.049	.050	.051	.002	.00009	7.8×10^{-9}	.049	.050	.050
	MMIL1	.005	.0005	2.5×10^{-7}	.049	.050	.051	.030	.00012	1.5×10^{-8}	.050	.050	.050
	MMIL2	.003	.0005	2.5×10^{-7}	.049	.050	.051	.008	8.7×10^{-5}	1.0×10^{-8}	.050	.050	.050
$\sigma^2=.05$	FML	.180	.0119	.00014	.026	.043	.065	2.07	.0089	.00008	.034	.048	.065
	MMIL1	.240	.0006	4.2×10^{-7}	.026	.043	.065	1.56	.0109	.0001	.033	.046	.062
	MMIL2	4.65	.0124	.00017	.027	.044	.067	1.99	.0090	.0001	.034	.048	.064
$\gamma=.5$	FML	66.22	.3558	.1309	.0001	.554	1.179	.03	.1911	.0365	.182	.502	.817
	MMIL1	1.89	.3317	.1099	-.047	.505	1.037	.65	.1917	.0367	.181	.502	.814
	MMIL2	4.81	.3309	.1095	-.052	.499	1.031	.02	.0958	.0092	.341	.501	.658
$\beta=.1$	FML	14.61	.0720	.0054	.0002	.077	.216	1.10	.0145	.0002	.074	.099	.122
	MMIL1	3.98	.0278	.0008	.041	.099	.138	3.44	.0213	.0005	.034	.098	.120
	MMIL2	.94	.0254	.0007	.057	.100	.139	.79	.0143	.0002	.074	.099	.122
$\alpha=.5$	FML	1.01	.1357	.0184	.243	.513	.771	.05	.0011	1.3×10^{-6}	.494	.500	.502
	MMIL1	.33	.0024	6.0×10^{-6}	.495	.499	.503	.35	.0015	2.4×10^{-6}	.498	.500	.505
	MMIL2	.07	.0022	5.0×10^{-6}	.496	.500	.504	.03	.0011	1.1×10^{-6}	.498	.500	.502
$\sigma^2=.75$	FML	77.36	.1833	.0396	.403	.658	1.005	34.34	.1349	.0194	.507	.708	.959
	MMIL1	62.14	.1960	.0423	.402	.673	1.034	17.71	.1526	.0236	.512	.715	1.01
	MMIL2	73.64	.1828	.0388	.399	.667	.992	34.47	.1345	.0193	.507	.708	.959

Q_5 = 5th percentile, Q_{50} = Median = 50th percentile and Q_{95} = 95th percentile

* These values have been multiplied by 1000.

Table 2 (continued)

Parameter		n = 30						n = 60					
		Bias*	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias*	SD	MSE	Q_5	Q_{50}	Q_{95}
	FML	3.43	.1843	.0334	.0001	.294	.617	1.53	.1142	.0130	.110	.296	.486
$\gamma = .3$	MMIL1	17.95	.1936	.0338	-.023	.319	.670	.33	.1146	.0131	.112	.298	.488
	MMIL2	1.37	.1960	.0384	-.031	.296	.617	.18	.1145	.0131	.110	.296	.486
$\beta = .2$	FML	.03	.0131	.0002	.179	.199	.222	.01	.0029	8.6×10^{-6}	.195	.200	.205
	MMIL1	.52	.0083	.0001	.196	.199	.220	.08	.0031	9.5×10^{-6}	.195	.200	.205
	MMIL2	.08	.0133	.0002	.179	.199	.222	.03	.0029	8.5×10^{-6}	.195	.200	.205
$\alpha = .015$	FML	.01	.0002	3.3×10^{-8}	.015	.015	.016	.001	.0001	3.9×10^{-9}	.015	.015	.015
	MMIL1	.007	.0001	1.1×10^{-8}	.015	.015	.015	.003	.0001	3.6×10^{-9}	.015	.015	.015
	MMIL2	.0006	.0002	3.0×10^{-8}	.015	.015	.015	.0002	.0001	3.6×10^{-9}	.015	.015	.015
$\sigma^2 = .25$	FML	23.5	.0617	.0044	.133	.221	.335	11.33	.0446	.0021	.170	.236	.320
	MMIL1	19.1	.0625	.0043	.135	.227	.338	10.51	.0448	.0021	.172	.237	.320
	MMIL2	23.73	.0616	.0043	.133	.221	.334	11.32	.0445	.0021	.170	.236	.320
$\gamma = .3$	FML	30.69	.2741	.0760	.0001	.302	.829	18.97	.1776	.0319	1.0×10^{-4}	.314	.607
	MMIL1	1.71	.3250	.1056	-.238	.302	.829	.30	.1975	.0390	-.030	.297	.621
	MMIL2	1.84	.3251	.1057	-.237	.302	.829	.69	.1975	.0389	-.032	.298	.620
$\beta = .1$	FML	0.36	.0074	.0001	.088	.100	.111	.37	.0048	.00002	.074	.096	.106
	MMIL1	0.38	.0073	.0001	.088	.100	.111	.07	.0045	.00002	.092	.099	.107
	MMIL2	0.39	.0074	.0001	.088	.100	.111	.03	.0045	.00002	.092	.099	.107
$\alpha = .03$	FML	0.02	.0009	7.6×10^{-7}	.029	.030	.031	6.05	.0080	.0001	.030	.030	.046
	MMIL1	0.01	.0008	7.3×10^{-7}	.029	.030	.031	.009	.0004	1.2×10^{-7}	.029	.030	.031
	MMIL2	.007	.0008	7.5×10^{-7}	.029	.030	.031	.003	.0004	1.2×10^{-7}	.029	.030	.031
$\sigma^2 = .75$	FML	69.82	.1847	.0389	.400	.669	1.004	44.58	.1586	.0271	.537	.830	1.001
	MMIL1	72.29	.1843	.0391	.400	.667	.998	33.97	.1334	.0189	.512	.708	.955
	MMIL2	72.32	.1843	.0391	.400	.667	.998	34.03	.1334	.0189	.512	.708	.955

Q_5 = 5th percentile, Q_{50} = 50th percentile and Q_{95} = 95th percentile

* These values have been multiplied by 1000.

Table 3 Estimated bias, standard deviation, mean squared error, and quantiles (5%, 50% and 95%) of estimators for model (3.2) based on FML and two step MMIL (MMIL1 and MMIL2) estimators

		$n=30$						$n=60$					
Parameter		Bias	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias	SD	MSE	Q_5	Q_{50}	Q_{95}
$\theta_1=.5$	FML	.0010	.1087	.0118	.322	.498	.675	.0001	.0732	.0054	.383	.499	.619
	MMIL1	.0012	.1086	.0118	.322	.498	.675	.0001	.0732	.0054	.383	.499	.619
	MMIL2	.0009	.1089	.0119	.322	.499	.677	.0001	.0732	.0054	.383	.499	.619
$\theta_2=.25$	FML	.0001	.1037	.0108	.082	.248	.419	.0009	.0788	.0062	.125	.250	.381
	MMIL1	.0001	.1038	.0108	.083	.248	.419	.0010	.0773	.0059	.125	.250	.382
	MMIL2	.0003	.1040	.0108	.080	.247	.419	.0008	.0789	.0062	.125	.250	.382
$\theta_3=.01$	FML	.4520	6.898	47.76	-3.30	.035	3.13	.0678	1.95	3.809	-1.73	.017	2.14
	MMIL1	.4573	6.814	46.62	-3.01	.037	3.10	.0362	1.74	3.041	-1.73	.016	2.04
	MMIL2	.0820	8.070	65.11	-3.38	.035	3.10	.1029	1.89	3.568	-1.76	.023	1.98
$\theta_4=.25$	FML	.1037	2.46	6.072	.048	.252	.570	.0455	.3197	.1043	.017	.253	.624
	MMIL1	.0571	1.18	1.392	.048	.252	.570	.0457	.3078	.0968	.020	.252	.623
	MMIL2	.0282	.186	.0354	.048	.252	.570	.0236	.1849	.0347	.015	.253	.611
$\sigma^2=.25$	FML	.0356	.0591	.0048	.129	.210	.317	.0183	.0429	.0022	.166	.230	.304
	MMIL1	.0355	.0592	.0048	.129	.210	.317	.0182	.0429	.0022	.166	0.230	.304
	MMIL2	.0358	.0590	.0048	.128	.209	.317	.0183	.0429	.0022	.166	0.230	.304
$\theta_1=.75$	FML	.0017	.2064	.0426	.412	.747	1.09	.0003	.1390	.0193	.527	.749	.977
	MMIL1	.0023	.2061	.0425	.412	.746	1.08	.0003	.1389	.0193	.527	.749	.977
	MMIL2	.0014	.2066	.0427	.412	.748	1.09	.0003	.1390	.0193	.527	.749	.977
$\theta_2=.5$	FML	.0004	.1965	.0386	.188	.495	.822	.0018	.1495	.0224	.264	.500	.750
	MMIL1	.0001	.1964	.0385	.187	.495	.820	.0017	.1495	.0223	.264	.500	.751
	MMIL2	.0003	.1967	.0387	.188	.495	.820	.0018	.1497	.0224	.264	.500	.751
$\theta_3=.1$	FML	.2551	5.39	29.13	-2.44	.119	2.67	.0667	1.544	2.387	-1.406	.106	1.91
	MMIL1	.2430	5.20	27.13	-2.30	.124	2.34	.0847	1.158	1.348	-1.398	.105	1.91
	MMIL2	.0918	5.33	28.44	-2.36	.119	2.47	.0796	1.416	2.011	-1.398	.105	1.91
$\theta_4=.5$	FML	.0892	2.244	5.039	.123	.504	1.07	.0674	.4868	.2414	.065	.502	1.16
	MMIL1	.0396	.2953	.0887	.123	.504	1.07	.0663	.4620	.2178	.078	.502	1.15
	MMIL2	.0502	.4929	.2454	.123	.504	1.07	.0444	.3879	.1524	.061	.502	1.15
$\sigma^2=.9$	FML	.1253	.2133	.0612	.465	.758	1.14	.0633	.1549	.0280	.601	.829	1.10
	MMIL1	.1239	.2139	.0611	.465	.758	1.14	.0632	.1549	.0279	.600	.829	1.10
	MMIL2	.1257	.2133	.0613	.463	.758	1.14	.0633	.1549	.0279	.601	.829	1.10

Table 3 (continued)

Parameter		$n = 30$						$n = 60$					
		Bias	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias	SD	MSE	Q_5	Q_{50}	Q_{95}
$\theta_1 = .03$	FML	.0011	.1087	.0118	-.147	.031	.213	.0010	.0732	.0054	-.085	.033	.153
	MMIL1	.0011	.1087	.0118	-.147	.031	.213	.0010	.0732	.0054	-.085	.033	.153
	MMIL2	.0010	.1089	.0118	-.147	.031	.213	.0010	.0732	.0054	-.085	.033	.153
$\theta_2 = 1.02$	FML	.0018	.1028	.0106	.857	1.02	1.19	.0015	.0783	.0061	.899	1.022	1.15
	MMIL1	.0018	.1027	.0106	.857	1.02	1.19	.0015	.0784	.0061	.899	1.022	1.15
	MMIL2	.0018	.1028	.0106	.856	1.02	1.19	.0016	.0784	.0061	.899	1.022	1.15
$\theta_3 = -1.1$	FML	.5091	5.004	25.29	-3.79	-1.09	.466	.0978	1.655	2.748	-2.90	-1.087	.325
	MMIL1	.4713	4.488	20.35	-2.10	-1.08	.431	.0746	1.105	1.226	-2.90	-1.087	.325
	MMIL2	.3867	3.678	13.67	-3.80	-1.09	.443	.0299	1.938	3.754	-2.90	-1.087	.325
$\theta_4 = -.5$	FML	.0911	2.237	5.011	-.958	-.510	-.220	.0485	.2906	.0868	-.952	-.520	-.210
	MMIL1	.0476	.2873	.0848	-.948	-.520	-.210	.0413	.2436	.0610	-.958	-.510	-.220
	MMIL2	.0457	.2587	.0690	-.952	-.520	-.209	.0413	.2437	.0610	-.958	-.510	-.220
$\sigma^2 = .25$	FML	.0350	.0589	.0047	.130	.210	.317	.0181	.0430	.0022	.166	.230	.304
	MMIL1	.0350	.0589	.0047	.130	.210	.317	.0181	.0430	.0022	.166	.230	.304
	MMIL2	.0351	.0589	.0047	.130	.210	.317	.0181	.0430	.0022	.166	.230	.304
$\theta_1 = -.025$	FML	.0011	.1087	.0118	-.202	-.024	.158	.0011	.0732	.0054	-.140	-.022	.098
	MMIL1	.0011	.1087	.0118	-.202	-.0237	.158	.0012	.0732	.0054	-.140	-.022	.098
	MMIL2	.0011	.1088	.0118	-.2018	-.0236	.158	.0011	.0732	.0054	-.140	-.022	.098
$\theta_2 = 1.1$	FML	.0017	.1028	.0106	.937	1.10	1.27	.0015	.0784	.0061	.979	1.10	1.23
	MMIL1	.0018	.1028	.0106	.937	1.10	1.27	.0014	.0784	.0061	.978	1.10	1.23
	MMIL2	.0018	.1029	.0106	.936	1.10	1.27	.0015	.0784	.0061	.979	1.10	1.23
$\theta_3 = -1.1$	FML	.5178	5.21	27.41	-3.82	-1.09	.466	.0738	1.114	1.246	-2.89	-1.09	.325
	MMIL1	.4788	5.11	26.36	-3.78	-1.09	.466	.0872	1.140	1.307	-2.90	-1.10	.325
	MMIL2	.5008	5.03	25.54	-3.79	-1.09	.465	.0694	1.178	1.392	-2.89	-1.09	.325
$\theta_4 = -0.5$	FML	.0505	.3181	.1037	-.954	-.520	-.209	.0413	.2436	.0610	-.958	-.509	-.220
	MMIL1	.0486	.3162	.1023	-.948	-.519	-.209	.0443	.2511	.0650	-.960	-.510	-.220
	MMIL2	.0496	.3038	.0947	-.948	-.520	-.209	.0413	.2436	.0610	-.958	-.510	-.220
$\sigma^2 = 0.25$	FML	.0350	.0589	.0047	.130	.210	.317	.0181	.0430	.0022	.166	.230	.304
	MMIL1	.0350	.0589	.0047	.130	.210	.317	.0180	.0430	.0022	.166	.230	.304
	MMIL2	.0350	.0589	.0047	.130	.210	.317	.0181	.0430	.0022	.166	.230	.304

Table 4 Estimated bias, standard deviation, mean squared error, and quantiles (5%, 50% and 95%) of estimators for model (3.3) based on FML and two step MMIL (MMIL1 and MMIL2) estimators

		$n=30$						$n=60$					
Parameter		Bias	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias	SD	MSE	Q_5	Q_{50}	Q_{95}
$\alpha=1.5$	FML	.0017	.0427	.0018	1.47	1.50	1.57	.0008	.0282	.0008	1.48	1.50	1.55
	MMIL1	.0030	.0446	.0020	1.43	1.50	1.58	.0002	.0249	.0006	1.46	1.50	1.54
	MMIL2	.0030	.0446	.0020	1.43	1.50	1.58	.0002	.0249	.0006	1.46	1.50	1.54
$\beta=.25$	FML	.0272	.1191	.0149	.194	.257	.498	.0097	.0833	.0070	.202	.253	.400
	MMIL1	.0216	.1030	.0111	.200	.256	.491	.0200	.0881	.0082	.147	.261	.435
	MMIL2	.0210	.1162	.0139	.197	.256	.499	.0198	.0882	.0082	.147	.261	.435
$\gamma=1.155$	FML	.5505	2.500	6.55	.557	1.10	4.85	.1171	1.007	1.028	.734	1.13	2.60
	MMIL1	.4113	1.683	3.00	.573	1.14	4.57	.2244	1.043	1.138	.255	1.15	3.12
	MMIL2	.4591	1.776	3.36	.583	1.14	4.49	.2202	1.031	1.111	.255	1.15	3.12
$\sigma^2=.05$	FML	.0049	.0124	.0002	.036	.044	.067	.0022	.0091	.0001	.041	.047	.064
	MMIL1	.0050	.0125	.0002	.027	.044	.067	.0022	.0092	.0001	.034	.047	.064
	MMIL2	.0050	.0125	.0002	.027	.044	.067	.0022	.0092	.0001	.034	.047	.064
$\alpha=.5$	FML	.0021	.0428	.0018	.473	.502	.574	.0008	.0282	.0008	.480	.499	.545
	MMIL1	.0029	.0446	.0020	.431	.502	.579	3.0×10^{-6}	.0249	.0006	.458	.500	.540
	MMIL2	.0028	.0447	.0020	.431	.502	.578	.00001	.0249	.0006	.458	.500	.540
$\beta=.25$	FML	.0289	.1298	.0177	.188	.256	.521	.0122	.0951	.0092	.198	.255	.422
	MMIL1	.0279	.1211	.0154	.193	.253	.518	.0230	.1000	.0105	.134	.261	.455
	MMIL2	.0293	.1264	.0168	.192	.254	.520	.0243	.1156	.0139	.135	.261	.455
$\gamma=1.5$	FML	.8890	4.99	25.67	.698	1.41	6.25	.2312	1.917	3.73	.969	1.48	3.50
	MMIL1	.5512	2.20	5.15	.750	1.46	5.89	.2944	1.519	2.40	.310	1.48	4.07
	MMIL2	.6133	4.33	19.10	.732	1.49	6.20	.2454	2.211	4.95	.327	1.49	4.09
$\sigma^2=.05$	FML	.0050	.0124	.0002	.036	.044	.067	.0022	.0091	.0001	.041	.047	.064
	MMIL1	.0050	.0125	.0002	.027	.044	.067	.0022	.0092	.0001	.034	.047	.064
	MMIL2	.0050	.0124	.0002	.027	.044	.067	.0022	.0092	.0001	.034	.047	.064

Table 4 (continued)

Parameter		$n=30$						$n=60$					
		Bias	SD	MSE	Q_5	Q_{50}	Q_{95}	Bias	SD	MSE	Q_5	Q_{50}	Q_{95}
$\alpha=1.5$	FML	.0017	.0426	.0018	1.43	1.50	1.57	.0008	.0282	.00080	1.45	1.50	1.55
	MMIL1	.0030	.0446	.0020	1.43	1.50	1.58	.0008	.0249	.00062	1.46	1.50	1.54
	MMIL2	.0030	.0447	.0020	1.43	1.50	1.58	.0008	.0249	.00062	1.46	1.50	1.54
$\beta=0.25$	FML	.0269	.1174	.0145	.127	.257	.495	.0095	.0820	.00682	.140	.254	0.397
	MMIL1	.0208	.1108	.0123	.127	.253	.494	.0195	.0864	.00783	.149	.261	0.431
	MMIL2	.0244	.1164	.0142	.127	.254	.494	.0196	.0864	.00784	.149	.261	0.431
$\gamma=1.1$	FML	.5008	2.21	5.13	.071	1.06	4.61	.1082	.9594	.93172	.162	1.08	2.46
	MMIL1	.3920	1.62	2.79	.081	1.09	4.26	.2126	.9881	1.0210	.242	1.09	2.98
	MMIL2	.4442	1.83	3.55	.077	1.09	4.24	.2138	.9884	1.0222	.242	1.09	2.98
$\sigma^2=0.05$	FML	.0049	.0125	.0002	.027	.044	.067	.0022	.0091	.00009	.034	.047	.064
	MMIL1	.0050	.0125	.0002	.027	.044	.067	.0022	.0092	.00009	.034	.047	.064
	MMIL2	.0050	.0125	.0002	.027	.044	.067	.0022	.0092	.00009	.034	.047	.064
$\alpha=.5$	FML	.0066	.0942	.0089	.445	.505	.669	.0008	.0630	.0040	.457	.498	.600
	MMIL1	.0054	.0988	.0098	.343	.505	.670	.0002	.0555	.0031	.405	.501	.590
	MMIL2	.0054	.0990	.0098	.343	.505	.671	.0002	.0555	.0031	.405	.501	.590
$\beta=.25$	FML	.0718	.1950	.0432	.192	.280	.737	.0453	.1339	.0200	.206	.270	.535
	MMIL1	.0600	.1593	.0290	.193	.281	.735	.0603	.1023	.0141	.136	.278	.596
	MMIL2	.0711	.1623	.0314	.191	.282	.737	.0589	.1539	.0272	.138	.278	.601
$\gamma=.25$	FML	.9922	5.75	34.06	.001	.238	4.26	.3305	1.495	2.34	.001	.244	1.86
	MMIL1	.8378	4.37	19.81	.001	.223	3.76	.4901	2.474	6.36	.001	.274	2.53
	MMIL2	.7670	3.33	11.67	.001	.223	3.77	.4673	1.917	3.89	.001	.276	2.63
$\sigma^2=.25$	FML	.0236	.0625	.0045	.182	.221	.335	.0110	.0453	.0022	.208	.237	.319
	MMIL1	.0234	.0625	.0045	.133	.222	.337	.0112	.0458	.0022	.168	.236	.321
	MMIL2	.0235	.0624	.0044	.133	.222	.336	.0112	.0458	.0022	.168	.236	.321