

Department of Economics

A Note on an Iterative Least Squares Estimation Method for ARMA and VARMA Models

George Kapetanios

Working Paper No. 467

November 2002

ISSN 1473-0278



Queen Mary
University of London

A note on an iterative least squares estimation method for ARMA and VARMA models

George Kapetanios*
Queen Mary University of London

November 2002

Abstract

In this note we suggest a new iterative least squares method for estimating scalar and vector ARMA models. A Monte Carlo study shows that the method has better small sample properties than existing least squares methods and compares favourably with maximum likelihood estimation as well.

Keywords: ARMA Models
JEL Classification: C13, C22

*Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS. Phone: 00-44-20-78825097. Email: G.Kapetanios@qmul.ac.uk

1 Introduction

The class of univariate and multivariate ARMA models is a flexible and powerful modelling tool applicable in a variety of situations that has appealing theoretical properties. Nevertheless, its use in empirical work has been limited because of severe practical estimation problems. Maximum likelihood has received most of the attention as an estimation method for ARMA models. Problems with the maximum likelihood method include slow convergence and non-robustness with respect to initial conditions. An alternative class of methods that has received little attention in the literature are least squares methods. Exceptions include Hannan and Rissanen (1982), Koreisha and Pukkila (1990a), Koreisha and Pukkila (1990), Koreisha and Yoshimoto (1991) and Choudhury and Power (1998).

This note provides a easily implementable iterative least squares estimation method for ARMA and VARMA models with intuitive theoretical properties and good small sample properties as we show in a Monte Carlo study. This procedure has not been suggested anywhere else in the literature to the best of the author's knowledge. Section 2 discusses the new method and its theoretical properties. Section 3 presents the Monte Carlo results. Section 4 concludes.

2 The Method

The model we consider is of the form

$$\Phi(L)y_t = \Theta(L)\epsilon_t, t = 1, \dots, T$$

where y_t is an m dimensional series $m \geq 1$, $\Phi(L)$ and $\Theta(L)$ are matrices of lag polynomials of the form $I - \Phi_1 L - \dots - \Phi_p L^p$ and $I + \Theta_1 L + \dots + \Theta_q L^q$ respectively and ϵ_t is an m dimensional i.i.d. unobserved error sequence with finite fourth moment. We will not distinguish between the scalar case ($m = 1$) and the vector case ($m > 1$) to reduce the notational burden. Whenever

the distinction matters we will draw attention to that fact. We assume that the model is stable, invertible and identified. This implies that there are no common factors in the lag polynomials in the scalar case and that, in the vector case, the model is in a form which is sufficient for identification, i.e. in a final equations form or in an echelon form (for details see Lutkepohl (1993, pp. 246-248)).

Most of the previous work on least squares estimation methods suggest the construction of an initial consistent estimate for the error sequence and its use in a least squares procedure for the estimation of Φ_i , $i = 1, \dots, p$ and Θ_i , $i = 1 \dots, q$. Our suggestion is to iterate this least squares operation using the new estimate of the error sequence until the estimate of the error sequence converges. The consistency of the estimates follows straightforwardly from the consistency of the initial estimate of the error sequence.

To formalise our approach we have the following: The initial estimate of the error sequence denoted by $\hat{\epsilon}^0 = (\hat{\epsilon}_1^0, \dots, \hat{\epsilon}_T^0)$ may be obtained from the residuals of an autoregression of the form $C(L)y_t = v_t$ where the order of the (matrix of) lag polynomials tends to infinity asymptotically. Discussion of the conditions sufficient for consistency of $\hat{\epsilon}^0$ may be found in a number of papers (see e.g. Ng and Perron (1995)). It suffices to say that the rate should be equal to or larger than $c \ln(T)$ for some positive constant c . Then, the parameters are estimated either from the regression

$$y_t = \sum_{i=1}^p \Phi_i y_{t-i} + \sum_{i=1}^q \Theta_i \hat{\epsilon}_{t-i}^0 + \eta_t \quad (1)$$

or

$$y_t - \hat{\epsilon}_t^0 = \sum_{i=1}^p \Phi_i y_{t-i} + \sum_{i=1}^q \Theta_i \hat{\epsilon}_{t-i}^0 + \eta_t$$

Existing methods estimate these regressions either by OLS or GLS. Let us denote these parameter estimates by $\hat{\Phi}_i^0$, $i = 1, \dots, p$ and $\hat{\Theta}_i^0$, $i = 1, \dots, q$. We simply suggest that the estimated residuals from (1) be reused iteratively in regressions of the same form as (1) but with $\hat{\epsilon}^j$ in the place of $\hat{\epsilon}^0$ where $\hat{\epsilon}^j$ denotes the residuals of the j -th iteration. If the iterative procedure converges

such that $\|\hat{\epsilon}^{j+1} - \hat{\epsilon}^j\| < \varepsilon$ for some suitably defined constant ε then the adopted estimates are $\hat{\Phi}_i^{j+1}$, $i = 1, \dots, p$ and $\hat{\Theta}_i^{j+1}$, $i = 1, \dots, q$. Otherwise, after some prespecified number of iterations the procedure is abandoned and the initial estimates are adopted as the final estimates. The procedure is consistent and equivalent to other least squares methods asymptotically, since the residuals from the long autoregression will converge to the true error terms. Clearly, there is no obvious reason why the procedure will converge in small samples, although the simulation results show that the procedure converges most of the time for reasonably large samples. Nevertheless, we can provide a simple numerical test for determining whether the procedure is likely to converge in small samples.

We restrict the analysis to the scalar case. Extension to the vector case is straightforward. Let us denote the estimated regressors at the j -th iteration by $X(\hat{\epsilon}^{j-1})$ and the vector of y_t , $t = 1, \dots, T$ by Y . Then, $\hat{\epsilon}^j = [I - X(\hat{\epsilon}^{j-1})(X(\hat{\epsilon}^{j-1})'X(\hat{\epsilon}^{j-1}))^{-1}X(\hat{\epsilon}^{j-1})]Y = f(\hat{\epsilon}^{j-1}; Y)$. The iterative procedure will converge if the above mapping is a contraction mapping¹. A sufficient condition for this is that the L^∞ norm of the Jacobian of the above mapping is less than unity in absolute value for all $\hat{\epsilon}^{j-1}$ (see Judd (1998, pp. 167)). In small samples little can be said about this mapping. We examine the asymptotic structure of the Jacobian as a guide to its small sample properties. The probability limit of the Jacobian is a strictly lower triangular matrix whose non zero elements are the true coefficients of the MA component of the model. To show that, in a simplified framework, consider an ARMA(1,1) model. Then,

$$\hat{\epsilon}_t^j = y_t - \hat{\phi}_1^j y_{t-1} - \hat{\theta}_1^j \hat{\epsilon}_{t-1}^{j-1}$$

Whereas in small samples a change in $\hat{\epsilon}_{t-1}^{j-1}$ will affect the whole of $\hat{\epsilon}_t^j$ because it will change the estimated parameters, asymptotically, the parameters will not be affected by a change in $\hat{\epsilon}_{t-1}^{j-1}$ only. The only term that will be affected

¹A contraction mapping, f , defined on some subset, \mathcal{D} , of a metric space (\mathcal{X}, ρ) is an operator such that $\rho(f(x), f(x')) \leq \rho(x, x')$, $\forall x, x' \in \mathcal{D}$.

is $\hat{\epsilon}_t^j$. The partial derivative will be equal to $\hat{\theta}_1^j$. From this follows the fact that the probability limit of the Jacobian is a banded strictly lower triangular matrix whose non zero band is made up of the true coefficients of the MA component.

The above line of argument suggests a conceptually simple numerical test that could be carried out at each iteration to determine whether the iterative procedure is likely to converge. We note that the Jacobian will, in small samples, almost surely be a solid matrix with full rank and linearly independent eigenvectors. It then follows that the mapping is a contraction mapping if the eigenvalues of the Jacobian are less than one in absolute value for all $\hat{\epsilon}^{j-1}$. Therefore, the eigenvalues of the Jacobian can be calculated and checked at each iteration. Our experience suggests that even if at one iteration one or more eigenvalues exceed unity in absolute value the iterative procedure is unlikely to converge. The asymptotic analysis of the Jacobian above suggests that, since both the trace and determinant of the Jacobian is asymptotically zero, the eigenvalues are likely to be small in absolute value, guaranteeing convergence, even in small samples. The results that we report in the next section support this since convergence occurs most of the time.

The new estimation method we propose provides consistent estimates of the parameters. It is intuitively appealing because it corrects the inherent incongruity in existing methods whereby the ex post estimated residual is not equal to the ex ante residual used to obtain the estimates. In more formal terms and assuming convergence, the superiority of this method to, say, the GLS method suggested by Koreisha and Pukkila (1990a) can be seen by noting that this method eliminates the random noise component underlying the justification of the GLS application. Further, the method is easy to implement as it requires just a series of OLS estimations. Note that the procedure can also be combined with the existing GLS method, which would be used in the event of no convergence instead of the simple LS method, to provide a hybrid estimation method.

3 Monte Carlo study

In this section we present a Monte Carlo study of the new method mainly in an ARMA framework but with one experiment on a VARMA model. For the ARMA experiments we consider 5 different estimation methods. We consider 8 ARMA experiments and one VARMA experiment. The ARMA experiment parameters and lag orders are given together with the results in Tables 1 to 3. The experiments have been chosen to conform with the format of previous studies in the area, see e.g. Koreisha and Pukkila (1990). The VARMA model is a VARMA(1,1) model and is given by a final equations form. The autoregressive parameter is 0.2 and the MA parameters are $vec(\Theta_1) = (0.25, 0.15, -0.2, -0.1)$. The data have been generated using pseudo random standard normal numbers generated within GAUSS. To minimise the effect of initial conditions, which are set to 0, the first 10 observations for each sample have been dropped. We consider samples of 50, 100, 200 and 400 observations. 1000 Monte Carlo replications have been undertaken. For the ARMA models the following estimation methods have been used: (i) the iterative OLS method (IOLS), (ii) the Hannan-Rissanen (1982) OLS method (OLS), (iii) the Koreisha and Pukkila (1990a) GLS method (KP), (iv) the conditional maximum likelihood method with the true parameters as initial estimates (ML 1) and (v) the conditional maximum likelihood method with the OLS estimates as initial estimates (ML 2). For the VARMA model only estimation methods (i), (ii) and (v) have been used. The least squares method is in this case restricted OLS corresponding to the restrictions imposed by the final equations form of the model. The maximum number of iterations for IOLS has been set to 500. If the ML estimation methods do not converge after 1000 iterations the OLS estimates are adopted as the ML estimates. We present the mean estimate of the parameter and the standard deviation of the parameter estimate over the replications. We also present the number of replications where the IOLS method did not converge (DNC in the Tables). The results, presented in Tables 1 to 4, make interesting reading.

For the ARMA models we see that both OLS and KP methods are biased for some experiments, even for samples of 400 observations, whereas this is never the case for IOLS. Further, the standard deviation of the parameter estimates over the Monte Carlo replications are smaller for IOLS than all the other LS methods in general and in particular the KP method. They also favourably compare with the ML methods. In a number of cases they are considerably smaller than those of the ML methods. Moving to the VARMA experiment all methods perform reasonably well, but again the IOLS method has lower standard deviation than the other methods. Overall, the two highest percentages of non convergent replications over the 1000 replications, out of the nine experiments, for the IOLS method are 9.6% and 2.4 % for 100 observations, 2.3% and 0.3 % for 200 observations and 0.2% and 0.1 % for 400 observations.

4 Conclusion

In this note we have suggested a new iterative least squares method for estimating scalar and vector ARMA models. We have provided a simple test for determining whether the iterative procedure is likely to converge in small samples. The method is easy to implement and requires no specialised programming routines. The Monte Carlo study showed that the method has better small sample properties than existing least squares methods and compares favourably with maximum likelihood estimation as well. Its use in the context of VARMA models is of particular relevance given the computational difficulty of ML estimation of these models.

References

CHOUDHURY, A. H., AND S. POWER (1998): "A Simplified GLS Estimator for ARMA Models," *Applied Economics Letters*, 5, 247–250.

- HANNAN, E. J., AND J. RISSANEN (1982): "Recursive Estimation of Mixed ARMA Order," *Biometrika*, 69, 81–94.
- JUDD, K. L. (1998): *Numerical Methods in Economics*. MIT Press.
- KOREISHA, S., AND T. PUKKILA (1990): "Linear Methods for Estimating ARMA and Regression Models with Serial Correlation," *Communications in Statistics - Simulation*, 19, 71–102.
- KOREISHA, S., AND T. PUKKILA (1990a): "A GLS Approach for Estimation of ARMA Models," *Journal of Time Series Analysis*, 11, 139–151.
- KOREISHA, S., AND G. YOSHIMOTO (1991): "A Comparison Between Identification Procedures for ARMA Models," *International Statistical Review*, 59, 37–57.
- LUTKEPOHL, H. (1993): *Introduction to Multiple Time Series Analysis*. Springer.
- NG, S., AND P. PERRON (1995): "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90, 268–281.

Table 1: Experiments 1-3

Exp.	Number of Observations									
	50					100				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.80 \ -0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.557	0.613	0.468	0.508	0.492	0.690	0.745	0.622	0.678	0.717
$\text{std}(\phi_1)$	0.608	0.660	0.389	0.426	0.526	0.380	0.431	0.265	0.223	0.428
θ_1	-0.298	-0.347	-0.201	-0.238	-0.221	-0.409	-0.448	-0.321	-0.389	-0.429
$\text{std}(\theta_1)$	0.630	0.680	0.390	0.431	0.528	0.405	0.464	0.268	0.243	0.457
DNC	85					11				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.80 \ -0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.750	0.779	0.721	0.751	0.767	0.775	0.789	0.770	0.760	0.768
$\text{std}(\phi_1)$	0.117	0.183	0.166	0.109	0.171	0.076	0.113	0.101	0.174	0.150
θ_1	-0.461	-0.470	-0.419	-0.457	-0.474	-0.478	-0.471	-0.463	-0.462	-0.470
$\text{std}(\theta_1)$	0.157	0.221	0.191	0.143	0.202	0.106	0.140	0.130	0.164	0.148
DNC	0					0				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ -0.80)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.465	0.499	0.088	0.236	0.314	0.443	0.508	0.245	0.359	0.516
$\text{std}(\phi_1)$	3.693	3.702	0.378	1.825	3.673	1.370	1.414	0.327	1.328	1.407
θ_1	-0.735	-0.745	-0.410	-0.523	-0.589	-0.730	-0.759	-0.585	-0.673	-0.806
$\text{std}(\theta_1)$	3.684	3.693	0.450	1.835	3.673	1.354	1.395	0.355	1.329	1.393
DNC	218					96				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ -0.80)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.448	0.505	0.402	0.291	0.574	0.471	0.511	0.495	0.442	0.571
$\text{std}(\phi_1)$	0.319	0.485	0.282	0.430	0.475	0.158	0.332	0.244	0.327	0.337
θ_1	-0.742	-0.756	-0.728	-0.582	-0.860	-0.766	-0.759	-0.799	-0.722	-0.847
$\text{std}(\theta_1)$	0.292	0.454	0.283	0.517	0.453	0.134	0.307	0.218	0.354	0.331
DNC	23					2				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ 0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.430	0.396	0.416	0.451	0.450	0.470	0.455	0.427	0.477	0.477
$\text{std}(\phi_1)$	0.179	0.222	0.370	0.158	0.167	0.120	0.141	0.386	0.113	0.113
θ_1	0.539	0.509	0.493	0.528	0.530	0.514	0.465	0.510	0.511	0.511
$\text{std}(\theta_1)$	0.186	0.209	0.251	0.171	0.177	0.130	0.138	0.217	0.116	0.116
DNC	2					0				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ 0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.484	0.477	0.401	0.487	0.487	0.490	0.487	0.398	0.492	0.492
$\text{std}(\phi_1)$	0.081	0.092	0.339	0.078 ⁹	0.078	0.060	0.067	0.379	0.057	0.057
θ_1	0.510	0.453	0.527	0.508	0.508	0.505	0.445	0.531	0.505	0.505
$\text{std}(\theta_1)$	0.089	0.092	0.164	0.078	0.078	0.063	0.065	0.154	0.054	0.054
DNC	0					0				

Table 2: Experiments 4-6

Exp.	Number of Observations									
	50					100				
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.80 \ 0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.718	0.707	0.672	0.724	0.725	0.761	0.756	0.698	0.764	0.764
std(ϕ_1)	0.125	0.140	0.343	0.121	0.120	0.080	0.087	0.341	0.076	0.076
θ_1	0.529	0.456	0.503	0.529	0.528	0.515	0.435	0.512	0.513	0.513
std(θ_1)	0.152	0.142	0.246	0.147	0.146	0.109	0.097	0.180	0.100	0.100
DNC	1					0				
200										
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.80 \ 0.50)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.782	0.780	0.666	0.783	0.783	0.792	0.791	0.701	0.793	0.793
std(ϕ_1)	0.053	0.056	0.405	0.051	0.051	0.035	0.037	0.422	0.034	0.034
θ_1	0.506	0.424	0.531	0.505	0.505	0.501	0.418	0.521	0.501	0.501
std(θ_1)	0.078	0.070	0.164	0.069	0.069	0.054	0.048	0.146	0.046	0.046
DNC	0					0				
50										
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ 0.80)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.435	0.410	0.473	0.470	0.469	0.469	0.454	0.470	0.485	0.485
std(ϕ_1)	0.180	0.213	0.374	0.152	0.152	0.110	0.132	0.322	0.096	0.098
θ_1	0.763	0.609	0.755	0.770	0.770	0.784	0.583	0.832	0.793	0.794
std(θ_1)	0.151	0.184	0.285	0.124	0.122	0.105	0.121	0.217	0.077	0.078
DNC	65					14				
200										
ARMA(1,1)										
$(\phi_1, \theta_1) = (0.50 \ 0.80)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.486	0.477	0.362	0.494	0.494	0.492	0.488	0.472	0.496	0.496
std(ϕ_1)	0.073	0.090	0.304	0.064	0.064	0.051	0.062	0.289	0.045	0.045
θ_1	0.794	0.571	0.828	0.798	0.798	0.796	0.563	0.795	0.799	0.799
std(θ_1)	0.076	0.085	0.214	0.050	0.049	0.054	0.060	0.149	0.034	0.034
DNC	1					0				
50										
ARMA(1,2)										
$(\phi_1, \theta_1, \theta_2) = (0.60 \ 0.00 \ 0.64)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.484	0.453	0.550	0.535	0.526	0.554	0.527	0.587	0.573	0.575
std(ϕ_1)	0.288	0.427	0.302	0.192	0.268	0.144	0.212	0.270	0.108	0.110
θ_1	0.080	0.132	0.015	0.024	0.035	0.033	0.087	-0.013	0.010	0.009
std(θ_1)	0.287	0.433	0.282	0.187	0.259	0.146	0.219	0.202	0.101	0.108
θ_2	0.599	0.487	0.482	0.666	0.669	0.624	0.489	0.514	0.648	0.648
std(θ_2)	0.166	0.208	0.228	0.180	0.184	0.108	0.137	0.160	0.098	0.100
DNC	124					24				
200										
ARMA(1,2)										
$(\phi_1, \theta_1, \theta_2) = (0.60 \ 0.00 \ 0.64)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.580	0.562	0.620	0.587	0.587	0.588	0.575	0.570	0.592	0.592
std(ϕ_1)	0.081	0.128	0.261	0.072	0.072	0.058	0.088	0.212	0.051	0.051
θ_1	0.014	0.062	-0.018	0.006	0.006	0.009	0.053	0.020	0.005	0.005
std(θ_1)	0.089	0.140	0.156	0.068	0.068	0.062	0.095	0.124	0.047	0.047
θ_2	0.633	0.487	0.583	0.642	0.642	0.638	0.486	0.660	0.642	0.642
std(θ_2)	0.076	0.094	0.128	0.062	0.062	0.052	0.065	0.108	0.043	0.043

Table 3: Experiments 7-8

Exp.	Number of Observations									
	50					100				
ARMA(2,1)										
$(\phi_1, \phi_2, \theta_1) = (1.00 \ -0.64 \ -0.60)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.944	0.962	0.901	0.928	0.907	0.967	0.974	0.959	0.985	0.997
$\text{std}(\phi_1)$	0.280	0.328	0.249	0.317	0.343	0.137	0.190	0.168	0.126	0.149
ϕ_2	-0.616	-0.623	-0.619	-0.618	-0.610	-0.627	-0.630	-0.623	-0.636	-0.639
$\text{std}(\phi_2)$	0.132	0.143	0.169	0.158	0.176	0.080	0.088	0.149	0.077	0.079
θ_1	-0.558	-0.513	-0.490	-0.539	-0.514	-0.574	-0.517	-0.564	-0.603	-0.613
$\text{std}(\theta_1)$	0.321	0.355	0.291	0.352	0.388	0.170	0.211	0.220	0.152	0.174
DNC	51					2				
200										
ARMA(2,1)										
$(\phi_1, \phi_2, \theta_1) = (1.00 \ -0.64 \ -0.60)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	0.983	0.987	0.992	0.994	0.999	0.991	0.991	0.994	0.996	0.993
$\text{std}(\phi_1)$	0.091	0.135	0.128	0.082	0.108	0.062	0.095	0.120	0.072	0.114
ϕ_2	-0.634	-0.636	-0.614	-0.640	-0.641	-0.638	-0.638	-0.574	-0.639	-0.637
$\text{std}(\phi_2)$	0.053	0.060	0.148	0.051	0.059	0.037	0.043	0.213	0.049	0.074
θ_1	-0.584	-0.527	-0.619	-0.602	-0.606	-0.589	-0.530	-0.656	-0.599	-0.596
$\text{std}(\theta_1)$	0.118	0.156	0.191	0.100	0.122	0.079	0.109	0.200	0.082	0.121
DNC	0					0				
50										
ARMA(2,1)										
$(\phi_1, \phi_2, \theta_1) = (0.00 \ -0.64 \ -0.60)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	-0.018	-0.011	-0.008	-0.013	-0.017	-0.007	-0.002	0.009	-0.001	-0.002
$\text{std}(\phi_1)$	0.150	0.178	0.224	0.146	0.156	0.105	0.126	0.228	0.100	0.106
ϕ_2	-0.630	-0.628	-0.647	-0.633	-0.632	-0.631	-0.629	-0.611	-0.631	-0.630
$\text{std}(\phi_2)$	0.120	0.124	0.228	0.118	0.118	0.085	0.087	0.270	0.082	0.083
θ_1	-0.581	-0.486	-0.564	-0.598	-0.591	-0.596	-0.498	-0.607	-0.617	-0.616
$\text{std}(\theta_1)$	0.194	0.199	0.240	0.183	0.205	0.133	0.141	0.187	0.117	0.129
DNC	25					0				
200										
ARMA(2,1)										
$(\phi_1, \phi_2, \theta_1) = (0.00 \ -0.64 \ -0.60)$	IOLS	OLS	KP	ML 1	ML 2	IOLS	OLS	KP	ML 1	ML 2
ϕ_1	-0.003	0.001	0.036	0.000	0.001	-0.003	-0.001	-0.003	-0.002	-0.001
$\text{std}(\phi_1)$	0.071	0.085	0.223	0.067	0.068	0.050	0.060	0.244	0.059	0.047
ϕ_2	-0.636	-0.635	-0.564	-0.636	-0.636	-0.638	-0.638	-0.492	-0.638	-0.638
$\text{std}(\phi_2)$	0.059	0.061	0.317	0.057	0.057	0.041	0.042	0.416	0.042	0.040
θ_1	-0.598	-0.500	-0.639	-0.609	-0.609	-0.598	-0.500	-0.651	-0.601	-0.603
$\text{std}(\theta_1)$	0.091	0.096	0.162	0.077	0.077	0.061	0.067	0.161	0.072	0.051
DNC	0					0				

Table 4: VARMA Experiment

Exp.	Number of Observations					
	50			100		
VARMA(1,1)						
	IOLS	OLS	ML	IOLS	OLS	ML
$\phi_1 = 0.20$	0.155	0.163	0.197	0.181	0.187	0.196
$\text{std}(\phi_1)$	0.251	0.296	0.285	0.185	0.206	0.217
$\theta_{1,1} = 0.25$	0.269	0.252	0.232	0.252	0.242	0.248
$\text{std}(\theta_{1,1})$	0.276	0.317	0.349	0.192	0.213	0.230
$\theta_{1,2} = -0.20$	-0.204	-0.203	-0.193	-0.205	-0.205	-0.204
$\text{std}(\theta_{1,2})$	0.156	0.155	0.179	0.104	0.103	0.104
$\theta_{2,1} = 0.15$	0.153	0.149	0.151	0.150	0.149	0.154
$\text{std}(\theta_{2,1})$	0.155	0.154	0.179	0.104	0.104	0.108
$\theta_{2,2} = -0.10$	-0.075	-0.097	-0.091	-0.089	-0.101	-0.095
$\text{std}(\theta_{2,2})$	0.300	0.329	0.368	0.213	0.230	0.250
DNC	5			0		
	200			400		
	IOLS	OLS	ML	IOLS	OLS	ML
$\phi_1 = 0.20$	0.187	0.189	0.198	0.196	0.197	0.202
$\text{std}(\phi_1)$	0.140	0.153	0.149	0.100	0.109	0.101
$\theta_{1,1} = 0.25$	0.255	0.250	0.250	0.251	0.248	0.248
$\text{std}(\theta_{1,1})$	0.144	0.158	0.158	0.103	0.113	0.105
$\theta_{1,2} = -0.20$	-0.203	-0.203	-0.203	-0.202	-0.202	-0.203
$\text{std}(\theta_{1,2})$	0.069	0.069	0.068	0.050	0.050	0.049
$\theta_{2,1} = 0.15$	0.152	0.151	0.154	0.150	0.149	0.152
$\text{std}(\theta_{2,1})$	0.072	0.072	0.073	0.051	0.051	0.052
$\theta_{2,2} = -0.10$	-0.094	-0.100	-0.100	-0.100	-0.102	-0.103
$\text{std}(\theta_{2,2})$	0.157	0.168	0.168	0.109	0.117	0.111
DNC	0			0		

**This working paper has been produced by
the Department of Economics at
Queen Mary, University of London**

**Copyright © 2002 George Kapetanios
All rights reserved.**

**Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580
Email: j.conner@qmul.ac.uk
Website: www.econ.qmul.ac.uk/papers/wp.htm**