CentER for Economic Research

Discussion paper

CIN00965

CBM
R
8414
1989
2

[redacted]

TILBURG UNIVERSITY

No. 8902

GENERALIZED LEAST SQUARES ESTIMATION OF LINEAR MODELS CONTAINING RATIONAL FUTURE EXPECTATIONS

by Theo Nijman and Franz Palm

January, 1989

GENERALIZED LEAST SQUARES ESTIMATION OF LINEAR MODELS CONTAINING RATIONAL FUTURE EXPECTATIONS

Theo Nijman *

Franz Palm **

ABSTRACT

We discuss the choice of approximations for unobserved expectations underlying consistent estimators in linear RE models with future expectations. We show how estimators which are more efficient than the commonly used GMM estimators can be obtained if it is assumed that the future expectation depends on a finite number of variables only. Numerical results for a simple model illustrate the relative efficiency of various estimators.

* Dept. of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

** Dept. of Economics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

1. INTRODUCTION

Models with expectational variables are widely used in empirical econometric research. Various estimators have been put forward to estimate the parameters in these models (see e.g. Pesaran (1987)). Many of these estimators belong to the general class of GMM estimators proposed by Hansen (1982). If a GMM estimator is used, the unobserved expectations are approximated by the corresponding realization following a suggestion of McCallum (1976). For static models an alternative referred to as the substitution approach by Wickens (1982) consists in fitting an auxiliary regression and approximating the unobserved expectation by the projection from the auxiliary equation. In this note we show how to approximate the future expectation by the projection from an auxiliary regression and obtain a generalized least squares estimator (GLS) that is at least as efficient as the GMM estimator based on future realizations as proxies for the future expectations, provided one is willing to assume that the future expectation depends on a finite number of variables only.
Although both the GLS estimator defined in this way and the GMM estimators are not fully efficient, in general they have many advantages over the efficient maximum likelihood (ML) estimator. For instance, they do not require a fully specified model and are therefore expected to be more robust with respect to specification uncertainty than full information methods. Moreover they are often computationally much more attractive than ML estimation.

The paper is organized as follows. In section 2 we introduce the GLS estimator and compare it with GMM estimators. In section 3 numerical results on the relative efficiency of the various estimators illustrate the argument. Finally section 4 contains some concluding remarks.

2. THE GLS ESTIMATOR

Consider the following linear model

$$y_t = \rho\, E[y_{t+1}|I_t] + \delta' x_t + \epsilon_t, \quad \epsilon_t \sim IN(0,\sigma_\epsilon^2), \tag{1}$$

$$x_t = \sum_{i=1}^{p} \Gamma_i\, x_{t-i} + v_t, \qquad v_t \sim IN(0,\Omega), \tag{2}$$

assume that ϵ_t and v_s are independent for all t and s, that v_t is independent of x_{t-1}, x_{t-2}, ... and define $I_t = \{y_t, x_t, y_{t-1}, x_{t-1}, \ldots\}$. Equation (1) describes a standard RE model with future expectation while equation (2) states that the k-dimensional vector of exogenous variables is generated by a vector autoregression, possibly with restrictions on the parameter matrices Γ_i. If x_t is stationary and $|\rho| < 1$ the model has a unique stationary solution which can be characterized by

$$y_{t+1} = \sum_{i=0}^{p} \psi_i\, x_{t-i} + u_t, \tag{3}$$

where u_t is independent of x_t, x_{t-1}, etc. and ψ_i is often a highly nonlinear function of the structural parameters in (1) and (2). By construction the error term u_t satisfies

$$E\, u_t u_{t-k} = 0 \quad \text{if } k \neq 0; \qquad E\, u_t^2 = \sigma_u^2\ ;$$

$$E\, u_t \epsilon_{t-k} = 0 \quad \text{if } k \neq 1; \qquad E\, u_t \epsilon_{t-1} = \sigma_\epsilon^2. \tag{4}$$

As (3) implies $E[y_{t+1}|I_t] = \sum_{i=0}^{p} \psi_i\, x_{t-i}$, ML estimation comes down to joint estimation of (2) and

$$y_t = (\rho\psi_0 + \delta')x_t + \rho \sum_{i=1}^{p} \psi_i \, x_{t-i} + \epsilon_t \qquad (5)$$

imposing all the restrictions on the Π's which can be computationally very demanding. Even if these restrictions are neglected simultaneous estimation of (5) and (2) may not be computationally very attractive.

The class of GMM estimators is based on the substitution of (3) into (1) which yields

$$y_t = \rho \, y_{t+1} + \delta' x_t + \epsilon_t - \rho u_t \qquad (6)$$

from which ρ and δ can be consistently estimated using IV methods because x_t, x_{t-1}, ... are orthogonal to ϵ_t and u_t. As $\epsilon_t - \rho u_t$ is autocorrelated (see (4)), the standard IV estimators can be improved upon by taking into account the properties of this error term as proposed by Cumby, Huizinga and Obstfeld (1983) and Hayashi and Sims (1983) for the present linear model and by Hansen (1982) for the general case. Note that all GMM estimators are based on the fact that x_{t-k} $(k \geq 0)$ is orthogonal to $y_{t+1} - E[y_{t+1}|I_t]$ but do not use the restrictions on (3).

An alternative class of estimators starts by accounting for the exclusion restrictions on the regression in (3) by writing

$$y_{t+1} = z_t'\pi + u_t, \qquad (7)$$

where z_t is the vector of all elements of $(x_t', \ldots, x_{t-p}')$ with nonzero coefficient in (3) and π is a corresponding vector of parameters. As $E[y_{t+1}|I_t] = z_t'\pi$, a natural approximation for the unobserved expectation of y_{t+1} which can be used instead of the approximation y_{t+1} used in (6) is given by $z'\hat{\pi}_t$, where $\hat{\pi}$ is the OLS estimator in (7). This approximation yields

$$y_t = \rho\, z_t'\hat{\pi} + \delta' x_t + \epsilon_t + \rho z_t'(\pi - \hat{\pi}). \tag{8}$$

If ρ and δ are estimated from (8) using OLS the resulting estimator obviously coincides with the IV estimator from (6) when z_t is the vector of instruments because of the two stage least squares interpretation of IV estimators. The parameters in (8) however can alternatively be estimated by GLS which yields a more efficient estimator. In obvious vector notation (8) reads as

$$y = [Z(Z'Z)^{-1}Z'y_+ \mid X] \begin{bmatrix} \rho \\ \delta \end{bmatrix} + \epsilon - \rho\, Z(Z'Z)^{-1}Z'u. \tag{9}$$

In the appendix we present expressions for the inverse of the covariance matrix of $\epsilon - \rho\, Z(Z'Z)^{-1}Z'u$ which can be used to obtain the GLS estimator. Of course unknown parameters in the covariance matrix have to be replaced by consistent estimates, but these can easily be obtained and do not affect the limiting distribution.

An important point to note is that the GLS estimator is at least as efficient as any GMM estimator which is based on the orthogonality conditions in (6) only. This is true because the estimator proposed by Cumby, Huizinga and Obstfeld (1983) can be interpreted as an estimator based on premultiplication of (9) by Z' and therefore cannot be more efficient than GLS on (9). Moreover it is well known (see e.g. Hansen (1982)) that a GMM estimator based on (6) only, cannot be more efficient than the estimator proposed by Cumby et.al. if the number of instruments used in that estimator tends to infinity. The GLS estimator can of course be more efficient than GMM estimators because the zero restrictions in (3) are exploited. A GMM estimator which simultaneously imposes the orthogonality restrictions in (3) and (6) in a bivariate model will probably be as efficient as the GLS estimator, but this estimator is no longer computationally attractive.

3. NUMERICAL EXAMPLE

In order to illustrate the argument in the previous section we present numerical results on the relative asymptotic efficiency of the various estimators for the very simple case where $p = 2$ and $k = 1$, with k being the number of exogenous variables in (1). It can be easily checked that in this case

$$\begin{aligned} \pi_0 &= \delta(\gamma_1 + \rho\gamma_2)\{1 - \rho\gamma_1 - \rho^2\gamma_2\}^{-1} \\ \pi_1 &= \delta\gamma_2\{1 - \rho\gamma_1 - \rho^2\gamma_2\}^{-1} \end{aligned} \tag{10}$$

and

$$u_t = \epsilon_{t+1} + (\rho\pi_0 + \delta)\, v_{t+1},$$

where lower case letters γ and π indicate that we consider a scalar case. Using (6) it is straightforward to evaluate the asymptotic variances of the estimators. Moreover it can be shown that the relative efficiency does not depend on all six parameters in the model but on ρ, γ_1, γ_2 and $R^2 = E(y_t-\epsilon_t)^2/Ey_t^2$ only.

In table 1 the asymptotic efficiency of four estimators of ρ and δ is presented. The efficiency is defined as the ratio of the large sample variance of an estimator over that of the ML estimator. The first estimator considered is the IV estimator when x_t and x_{t-1} are used as instruments in (6). It is evident that increasing the number of instruments does not change the asymptotic variance of the IV estimator. The second estimator applied to (6) has been proposed by Cumby et.al. (1983), denoted by CHO, with x_t, x_{t-1} and x_{t-2} being the instruments. In this case increasing the number of instruments would lower the asymptotic variance. Irrespective of the number of instruments, in this model, the CHO estimator will be less efficient than the third estimator that we consider which has been proposed by Hayashi and Sims (1983), HS. It is the efficient estimator within the class of GMM estimators based on the orthogonality restrictions in (6) only. Finally in table 1 we present the relative efficiency of the GLS estimator proposed in section 2 when it is a priori assumed that $p = 3$. The GLS estimator based on the more restrictive assumption that $p = 2$ (which coincides with the data generating process) is in this special case fully efficient and therefore the relative efficiency of this estimator is not presented in the table.

Table 1 : Relative efficiency of the maximum likelihood estimator compared with alternative estimators for ρ and δ if $k=1$, $p=2$ and $\rho = 0.9$.

			Rel.Eff. $\hat{\rho}$				Rel.Eff. $\hat{\delta}$			
R^2	γ_1	γ_2	IV	CHO	HS	GLS	IV	CHO	HS	GLS
.5	1.2	-.35	1.87	1.45	1.32	1.19	2.01	1.53	1.37	1.22
.5	1.4	-.45	2.30	1.62	1.30	1.13	2.43	1.69	1.33	1.15
.5	1.5	-.56	2.29	1.63	1.34	1.17	2.45	1.71	1.39	1.19
.5	1.7	-.72	2.91	1.89	1.32	1.12	3.06	1.96	1.34	1.13
.9	1.2	-.35	1.23	1.18	1.18	1.17	1.23	1.18	1.17	1.16
.9	1.4	-.45	1.46	1.30	1.28	1.23	1.46	1.30	1.28	1.23
.9	1.5	-.56	1.35	1.24	1.22	1.20	1.33	1.22	1.21	1.19
.9	1.7	-.72	1.54	1.31	1.27	1.20	1.51	1.30	1.25	1.19

efficiency between the various estimators can be considerable. In this example, the asymptotic variance of the CHO estimator with x_{t-2} included as additional instrument is substantially smaller than that of the standard IV estimator. Moreover it pays either to use more instruments in the computation of the CHO estimator or to use the HS estimator. Finally it is clear from table 1 the GLS estimator proposed in section 2 can in turn be substantially more efficient than the HS estimator.

7. CONCLUSIONS

In this note we showed how the assumption that a future expectation depends on a finite number of variables only can be exploited to increase the efficiency of simple consistent estimators. In the example, the ML estimator appeared to be only about 20 percent more efficient asymptotically than the GLS estimator which is computationally more easy to implement. This finding suggests that in empirical work, it is more appropriate to approximate the unobserved future expectations by a conditional expectation based on a finite number of past observations and then apply GLS than to substitute the future observation and apply some IV method. The result can be extended to more general models where for example lagged endogenous variables are included in (2).

Cumby, R.E., Huizinga, J. and M. Obstfeld (1983), "Two-step two-stage least squares estimation in models with rational expectations", Journal of Econometrics, 21, p. 333-356.

Hansen, L.P. (1982), "Large sample properties of generalized methods of moments estimators", Econometrica, 50, p. 1029-1054.

Hayashi, F. and C. Sims (1983), "Nearly efficient estimation of time series models with predetermined, but not exogenous, instruments", Econometrica, 51, p. 783-798.

McCallum, B.T. (1976), "Rational expectations and the estimation of econometric models : an alternative procedure", Econometrica, 17, p. 484-490.

Pesaran, M.H. (1987), "The Limits to Rational Expectations", Basil Blackwell, Oxford.

Wickens, M.R. (1982), "The efficient estimation of econometric models with rational expectations", Review of Economic Studies, 49, p. 55-67.

<u>APPENDIX</u> : Details on the computation of the GLS estimator.

In this appendix we show how to find a weighting matrix which can be used to compute the GLS estimator described in section 2. If (9) is rewritten as

$$y = W\,\theta + w, \tag{A1}$$

where $W = [Z(Z'Z)^{-1}Z'y_+ \,|X]$, $\theta' = (\rho,\delta')$ and $w = \epsilon - \rho Z(Z'Z)^{-1}Z'u$, the problem is to find a matrix $\hat{\Omega}^{-1}$ such that

$$\sqrt{T}\,(W'\hat{\Omega}^{-1}W)^{-1}\,W'\hat{\Omega}^{-1}w \sim N(0,\ \mathrm{plim}\ T(W'\hat{\Omega}^{-1}W)^{-1}), \tag{A2}$$

where T denotes the sample size. A matrix $\hat{\Omega}^{-1}$ which is appropriate in this respect is the inverse of the covariance matrix Ω of w assuming that Z is deterministic,

$$\Omega = \sigma_\epsilon^2 I_T - \rho\sigma_\epsilon^2\,\{Z(Z'Z)^{-1}Z'_{-1} + Z_{-1}(Z'Z)^{-1}Z'\} + \rho^2\sigma_u^2\,Z(Z'Z)^{-1}Z', \tag{A3}$$

if unknown parameters are replaced by consistent estimates. As Ω is a (TxT) matrix, it is computationally attractive not to invert Ω directly but to define

$$\begin{aligned}
\Omega &= \Omega_1 + \Omega_2\Omega_3^{-1}\Omega'_2 \\
\Omega_1 &= \sigma_\epsilon^2 I_T - \sigma_\epsilon^4\sigma_u^{-2}Z_{-1}(Z'Z)^{-1}Z'_{-1} \\
\Omega_2 &= Z - \rho^{-1}\sigma_\epsilon^2\sigma_u^{-2}Z_{-1} \\
\Omega_3 &= \rho^{-2}\sigma_u^{-2}Z'Z
\end{aligned} \tag{A4}$$

and to use the matrix inversion lemma

$$(\Omega_1 + \Omega_2\Omega_3^{-1}\Omega_2')^{-1} = \Omega_1^{-1} - \Omega_1^{-1}\Omega_2(\Omega_3^{-1} + \Omega_2'\Omega_1^{-1}\Omega_2)^{-1}\Omega_2'\Omega_1^{-1} \qquad \text{(A5)}$$

which only requires inversion of (k×k) dimensional matrices if the matrix inversion lemma is applied in a similar way to invert Ω_1.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

No.	Author(s)	Title
8801	Th. van de Klundert and F. van der Ploeg	Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidity
8802	J.R. Magnus and B. Pesaran	The Bias of Forecasts from a First-order Autoregression
8803	A.A. Weber	The Credibility of Monetary Policies, Policy-makers' Reputation and the EMS-Hypothesis: Empirical Evidence from 13 Countries
8804	F. van der Ploeg and A.J. de Zeeuw	Perfect Equilibrium in a Model of Competitive Arms Accumulation
8805	M.F.J. Steel	Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach
8806	Th. Ten Raa and E.N. Wolff	Secondary Products and the Measurement of Productivity Growth
8807	F. van der Ploeg	Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth
8901	Th. Ten Raa and P. Kop Jansen	The Choice of Model in the Construction of Input-Output Coefficient Matrices
8902	Th. Nijman and F. Palm	Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS
Bibliotheek K. U. Brabant
17 000 01117387 0