Rochester Center for Economic Research

Inference in Cointegrated Models using VAR Prewhitening to Estimate Shortrun Dynamics

Joon Y. Park   Masao Ogaki

Working Paper No. 281
June 1991

University of Rochester

Inference in Cointegrated Models Using VAR Prewhitening to Estimate Shortrun Dynamics

Joon Y. Park and Masao Ogaki

Rochester Center for Economic Research
Working Paper No. 281

June 1991

Inference in Cointegrated Models Using VAR Prewhitening to Estimate Shortrun Dynamics

Joon Y. Park
Cornell University
and
Masao Ogaki
University of Rochester

and

The Rochester Center for Economic Research
Working Paper No. 281

Abstract

This paper introduces an improved method of inference in cointegrated models, which uses the VAR prewhitening procedure to estimate shortrun dynamics. The prewhitening procedure provides a very flexible framework to incorporate the knowledge of shortrun dynamics, to efficiently estimate the longrun parameters in cointegrated systems. It can be used for the commonly used nonparametric methods of inference in cointegrated models. When the shortrun dynamics is given explicitly, as in ECM's, the nonparametric methods can be implemented in a *parametric form* with the proposed VAR prewhitening procedure. Therefore, they become conformable with the ECM-based methods. Unlike the ECM-based methods, however, they can also be made valid quite easily for misspecified models through the analysis of the spectrum for the prewhitened errors. The effect of the VAR prewhitening and other important issues on the use of shortrun information in estimating cointegrated models are investigated through an extensive Monte Carlo simulation.

First Draft: March 1991

We are grateful to A. Pagan and H. Pesaran for helpful discussions.

1. Introduction

The cointegrating models have increasingly been more popular in applied research, since the publication of an influential paper by Engle and Granger (1987). The issue of efficient estimation of cointegrated systems now seems to be largely settled down, at least in the theoretical domain. The usual least squares estimator is super-consistent, but known to be inefficient and biased asymptotically. The asymptotic sub-optimality of course well predicts unsatisfactory finite sample performance of the least squares estimate, as has been documented in the simulation studies by various authors.

There are two strands in the literature on the theory of efficient estimation of cointegrated systems: one nonparametric, and the other parametric. The exact ML approaches by Johansen (1988, 1989) and Park (1990b), for instance, are based on vector-autoregression (VAR) of known order, represented as an error correction model (ECM). They require a parametric specification of the shortrun transient dynamics, as well as the longrun static equilibrium relationships. In contrast, the approaches by Phillips (1988, 1989), Park (1990a) and Phillips and Hansen (1990) do not presume any specific transient dynamics. In their approaches, only the longrun equilibrium relationships are modelled in parametric forms. The shortrun dynamics are estimated nonparametrically to efficiently estimate the longrun parameters. It is shown in Park (1990b) that the two approaches are asymptotically equivalent.

We consider in this paper the VAR prewhitening method to estimate the shortrun dynamics of a cointegrated model. The method has recently been used by Andrews and Monahan (1990) to get improved asymptotic variance estimators for heteroskedastic and autocorrelated time series. The finite sample performance of the aforementioned nonparametric methods is, needless to say, heavily dependent upon the quality of estimates for the shortrun dynamics, which are effectively concentrated on the spectrum (especially, at the origin) of the stationary process driving a cointegrated model. The VAR prewhitening procedure therefore offers an obvious potential to improve the efficiency of the longrun parameter estimates in finite samples.

The VAR prewhitening method seems very attractive especially in the context of estimating cointegrated models. It provides a very flexible framework within which we may conveniently incorporate the knowledge (or the assumption, more plausibly) on the shortrun dynamics of a cointegrated model, to more efficiently estimate the longrun parameters in finite samples. Often, one may wish to consider a shortrun ECM, jointly with a longrun cointegrated model, as an essential ingredient of his empirical model or for the purpose of forecasting. This line of research has been taken by many – too many to enumerate – authors. For the specification of an ECM, the assumption of a finite order VAR structure for the underlying data generatng process (DGP) is unavoidable.

When the true DGP for a cointegrating model is given by a finite VAR of known order, the VAR prewhitening procedure for the stationary components of the model yields pure white noise residuals, whose spectrum is flat over the entire range. The *complete* prewhitening is possible in this case, and no more dynamic structure to be analyzed is left over in the residuals. As a result, any of the aforementioned 'nonparametric' methods of inference for cointegrated models can be implemented in a *parametric form*, since the VAR coefficients in the prewhitening procedure fully represent the shortrun dynamics of the model. The 'nonparametric' methods can therefore be made conformable with the parametric procedures by Johansen (1988, 1989) and Park (1990b) based on the ECM.

This is obviously an extreme case. One may hope at best that a postulated VAR closely approximate the true model. First, the order of the underlying VAR is typically unknown, even when it is justifiable that the true DGP has a VAR structure of finite order. Second, the underlying DGP may deviate from the standard finite order VAR in various directions. For instance, the errors may have MA and/or heteroskedastic, unconditional or conditional (such as ARCH), components. Economic models generated by optimizing behavior often suggest linear cointegrating relationships, but typically with the shortrun dynamics much more complicated than a simple finite order VAR structure. For the concrete examples of such models, the reader is referred to Cooley and Ogaki (1990), Gregory, Pagan and Smith (1990) and Ogaki and Park (1990).

The VAR prewhitening for the stationary components of a cointegrated model is therefore rarely expected in practice to produce pure white noise residuals. The motivation for the *pseudo*-VAR prewhitening is to obtain residuals whose spectrum is flatter at the locality of the origin, and easier to precisely estimate in finite samples. The more closely a postulated VAR approximates the true model, the less dynamics would be left over in the prewhitened residuals. Here we simply use a parametric specification of a VAR model, for presumably more complicated shortrun dynamics, to get better estimates for the spectrum of the stationary components of the model. Of course, it is allowed that the prewhitened residuals are serially correlated and/or heteroskedastic. The ML methods by Johansen (1988, 1989) and Park (1990b) are not asymptotically efficient, unless the complete prewhitening is possible.

We perform an extensive Monte Carlo experiment to evaluate the effect of VAR prewhitening, and to study some of the important issues on the use of the shortrun information to efficiently estimate the longrun parameters in finite samples. In the simulation, the VAR prewhitening method is applied to the canonical cointegrating regression (CCR) estimator developed by Park (1990a). The resulting estimator performs truly well in finite samples. Especially when the complete prewhitening is allowed, it in many cases yields virtually zero bias and mean squared error (MSE) really close to the theoretical asymptotic variance even for the samples of size 100. In terms of MSE, the CCR estimator with the VAR prewhitening outperforms the exact ML method in small samples, unmbiguously and often very substantially. The VAR prewhitening method appears to be quite effective in estimating the shortrun dynamics of cointegrated models.

Several other important issues on the use of shortrun information in estimating cointegrated models are also examined in our simulation. As we mentioned above, the specification and estimation of the shortrun dynamics becomes unimportant in asymptotics. Clearly, it is a finite sample issue. The method of VAR prewhitening makes any of the 'nonparametric' methods mentioned above applicable in both parametric and nonparametric form. This versatility makes it by far easier and more straightforward to see how important for the efficiency of the longrun parameter estimators it is to use the information on the shortrun dynamics. The question has been raised by several authors, including Gozalo (1989) and Inder (1990), but answered only indirectly by comparing *different* estimators.

We found by comparing the *same* CCR estimator implemented in parametric and nonparametric form that the precise information on the structure of shortrun dynamics greatly improves the efficiency of the longrun parameter estimates. The CCR estimator performs substantially better in finite samples, when the information on the shortrun dynamics is utilized. The information on the shortrun dynamics, however, provides only a *potential* to improve the longrun parameter estimators. In particular, it seems that estimators using the exact specification of the shortrun dynamics do not necessarily perform better than any other estimators not relying on such specification. Our simulation results indeed show that the exact ML method is in no sense better in small samples than the other 'nonparametric' methods which do not use any information on the shortrun dynamics. It very often yields completely nonsensical, and unacceptable, estimates in small samples.

The rest of this paper is organized as follows. The models and estimators are given in Section 2. The parametric and nonparametric specifications of a cointegrated model are compared, and their implications on the structure of the shortrun dynamics are contrasted. The existing 'nonparametric' procedures for inference in cointegrated models are briefly discussed. The VAR prewhitening procedure is introduced in Section 3. The method of the VAR prewhitening procedure to estimate the critical shortrun parameters is proposed. It is also explained how to implement the procedure to do inference in a cointegrated model. Section 4 reports the simulation results for the effect of the VAR prewhitening on the finite sample efficiency of the CCR procedure. Several other issues on the use and importance of shortrun information on estimating cointegrated models are investigated there too. The finite sample performance of the CCR estimator with the VAR prewhitening is also compared with that of the exact ML estimator. Section 5 concludes the paper, and the mathematical proofs are given in Appendix.

2. The Models and Estimators

We consider time series $\{y_t\}$ and $\{x_t\}$, which are respectively ℓ and m-dimensional integrated processes of order one. Let $\{y_t\}$ and $\{x_t\}$ be cointegrated, and write

$$\mathrm{M}(a) \ : \qquad y_t = \Pi' x_t + u_t$$

where $\{u_t\}$ is stationary. It is assumed in M(a) that there is no cointegration in $\{x_t\}$, and Π is uniquely determined. The model M(a) represents only a static longrun equilibrium relationship. No specific dynamic structure is presumed in the model.

When it is desirable to specify the shortrun transient dynamics, as well as the longrun static equilibrium relationship, we may look at an ECM. To define it precisely, let

$$z_t = (y_t', x_t')' \tag{1}$$

be an r-dimensional, $r = \ell + m$, time series, and define an $r \times \ell$ matrix

$$B = (\,\mathrm{I}, -\Pi')' \tag{2}$$

The usual ECM for $\{z_t\}$ in (1) is given in the form

$$\mathrm{M}(b) \ : \qquad \Delta z_t = AB' z_{t-p} + \sum_{k=1}^{p-1} C_k \Delta z_{t-k} + \epsilon_t$$

where $\{\epsilon_t\}$ is assumed to be white noise. In M(b), A is $r \times \ell$ matrix of error correction coefficients. The error correction model M(b) is derived in Johansen (1988) from a p-th order VAR model for $\{z_t\}$ under the assumption of the presence of unit roots and cointegration that is implied by M(a).

In M(a), the process

$$w_t = (u_t', \Delta x_t')' \tag{3}$$

which drives the model is assumed to be a general stationary process, without any precise specification of its dynamic structure. Only the presence of the unit root in $\{x_t\}$ and cointegration between $\{y_t\}$ and $\{x_t\}$ to insure, respectively, that $\{\Delta x_t\}$ and $\{u_t\}$ are stationary.

The process is, however, specified in M(b) in an exact parametric form. To see this, define an $(\ell + m)$-dimensional matrix

$$H = \begin{pmatrix} \mathrm{I} & -\Pi' \\ 0 & \mathrm{I} \end{pmatrix}$$

and let I_a be an $r \times r$ matrix the $\ell \times \ell$ northwest block of which is an identity and zero elsewhere. Also, define A_a to be an $r \times r$ matrix which is obtained by augmenting $r \times m$ zeros to A. Then we have

Lemma 1 *Suppose* DGP *is given by* M(b). *Then* $\{w_t\}$ *defined in* (3) *follows* VAR *of order* p

$$w_t = \sum_{k=1}^{p} \Phi_k w_{t-k} + e_t$$

with $\Phi_k = HC_kH^{-1} - HC_{k-1}I_a$, *for* $k = 1, \ldots, p-1$, $\Phi_p = HA_a - HC_{p-1}I_a$, *and* $e_t = H\epsilon_t$.

The existing nonparametric estimators for Π in M(a), such as those developed in Phillips (1988, 1989), Park (1990a) and Phillips and Hansen (1990), do not rely on any specific dynamic structure of $\{w_t\}$. They simply assume that it satisfies an invariance principle. The invariance principle is known to hold for a very wide class of stationary and possibly heteroskedastic processes, including of course the process generated by M(b). The reader is referred to Phillips (1989) and the references cited there for the explicit conditions. On the other hand, the parametric methods of Johansen (1988, 1989) and Park (1990b) are based on the ECM M(b). Therefore, they use the exact parametric specification of the dynamic structure of $\{w_t\}$, as shown in Lemma 1. They also impose the normality of the error distribution to derive the exact ML estimators. The ML estimate for Π may be obtained either using the Johansen's (1988, 1989) method with the posterior normalization of B, or following the procedure by Park (1990b) with the a priori identification of B. The posterior normalization or a priori identification of B is, of course, given by (2).

We define

$$\Sigma = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E(w_t w_t') \tag{4}$$

$$\Gamma = \lim_{n\to\infty} \frac{1}{n} \sum_{t=2}^{n} \sum_{k=1}^{t-1} E(w_t w_{t-k}') \tag{5}$$

$$\Omega = \lim_{n\to\infty} \frac{1}{n} E\left(\sum_{t=1}^{n} w_t\right)\left(\sum_{t=1}^{n} w_t\right)' \tag{6}$$

Notice that $\Omega = \Sigma + \Gamma + \Gamma'$. Also, we let $\Lambda = \Sigma + \Gamma$. Partition Ω and Λ conformably with $\{w_t\}$ in (3) as

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \tag{7}$$

and define

$$\Omega_* = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21} \quad \text{and} \quad \Lambda_2 = (\Lambda_{12}', \Lambda_{22}')' \tag{8}$$

The aforementioned nonparametric methods of inference in cointegrated systems require consistent estimators of some of the parameters defined in (4) – (8). In the paper, we specifically look at the canonical cointegrating regression (CCR) method by Park (1990a). The procedure requires the transformation of $\{y_t\}$ and $\{x_t\}$, using the stationary components $\{w_t\}$ of the model. The transformed series $\{y_t^*\}$ and $\{x_t^*\}$ are given explicitly by

$$x_t^* = x_t - \left(\Sigma^{-1}\Lambda_2\right)' w_t \tag{9}$$

$$y_t^* = y_t - \left(\Sigma^{-1}\Lambda_2\Pi + \left(0, \Omega_{12}\Omega_{22}^{-1}\right)'\right)' w_t \tag{10}$$

The regression reformulated with these transformed series is called the CCR.

The efficient estimation of the parameter Π may now simply be based on the OLS in the CCR

$$y_t^* = \Pi' x_t^* + u_t^* \tag{11}$$

Note that the cointegrating relationship between $\{y_t\}$ and $\{x_t\}$ in M(a) is preserved in (11), since the transformations to obtain $\{y_t^*\}$ and $\{x_t^*\}$ only involve stationary terms. Notice that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E(x_t^* u_t^*) = 0$$

i.e., the usual orthogonality between the regressors and the regression errors holds in the CCR. Moreover, the CCR errors $\{u_t^*\}$ become

$$u_t^* = u_t - \Omega_{12}\Omega_{22}^{-1}\Delta x_t$$

which is asymptotically independent of $\{\Delta x_t\}$. This is the reason that the OLS in CCR (11) is asymptotically equivalent to the ML estimator. The longrun variance of $\{u_t^*\}$ is given by Ω_*, which is defined in (8).

In this paper, we are primarily concerned with purely stochastic models, where the individual series $\{y_t\}$ and $\{x_t\}$ do not have any deterministic trends. The CCR method for more general models containing deterministic trends, however, is essentially identical. It only requires some obvious modifications in the definition $\{w_t\}$ in (3) so that $\{w_t\}$ represents purely stochastic stationary components of the underlying model. Any deterministic regressors included in the regression do not need the transformation in (9). The reader is referred to Park (1990a) for more details.

The method by Phillips and Hansen (1990) is quite similar to the CCR procedure introduced above. It also requires the nonparametric estimation of the nuisance parameters Ω and Λ_2, defined respectively in (6) and (7). More precisely, their estimator is given for our model by

$$\hat{\Pi}^{+*} = \hat{\Pi}^{+} - n\left(\sum_{t=1}^{n} x_t x_t'\right)^{-1}\Lambda_2'(\mathrm{I}, -\Omega_{12}\Omega_{22}^{-1})'$$

where $\hat{\Pi}^{+}$ is the least squares estimate from the regression of $\{y_t^+\}$ on $\{x_t\}$, $y_t^+ = y_t - \Omega_{12}\Omega_{22}^{-1}\Delta x_t$. Their estimator therefore modifies $\{y_t\}$ in the first step, and then corrects in the second step the OLS estimate from the regression of the modified $\{y_t^+\}$ on $\{x_t\}$. In contrast, the CCR method modifies both $\{y_t\}$ and $\{x_t\}$ simultaneously.

The Phillips' (1988, 1989) procedures are based on an ECM, just as those of Johansen (1988, 1989) and Park (1990b). He, however, uses an essentailly nonparametric ECM, and its asymptotic likelihood function, to derive the quasi-ML estimators. The procedure, in particular, does not presume the specification of the explicit shortrun dynamics; instead, it requires the nonparametric estimation of the longrun variance Ω in (6), similarly as Park (1990a) and Phillips and Hansen (1990).

3. VAR Prewhitening Method

We consider a p-th order VAR model

$$w_t = \sum_{k=1}^{p} \Phi_k w_{t-k} + e_t \tag{12}$$

for the process $\{w_t\}$ in (3) which generates the cointegrated model M(a) or M(b). When the model is generated by M(b), $\{w_t\}$ follows the VAR process exactly as given by (12) with white noise residual $\{e_t\}$. The VAR coefficients Φ_k's are defined, in a one-to-one fashion, from the coefficients C_k's in the ECM. This was shown in Lemma 1. The VAR model in (12) is, however, not meant to be a true model in general. It is to be understood here primarily as a *pseudo*-model, with the coefficients Φ_k's defined trivially as

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=k+1}^{n} E(w_{t-k}e_t') = 0 \tag{13}$$

for $k = 1, \ldots, p$.

The pseudo-VAR model has recently been considered by Andrews and Monahan (1990) to obtain an improved heteroskedasticity and autocorrelation consistent estimator of covariance matrix. Their method is directly applicable for the estimation of Ω defined in (6). The basic idea is to estimate the asymptotic variance of $\{w_t\}$, indirectly through fitting the VAR model (12) by the least squares and analyzing the spectrum of the prewhitened residual $\{e_t\}$. As they explain, we may easily obtain a consistent estimate for the asymptotic variance (or the longrun variance, in our terminology) of $\{w_t\}$ from that of the residual $\{e_t\}$ and the estimates of the VAR coefficients Φ_k for $k = 1, \ldots, p$, i.e., by 'recoloring' the spectrum of the prewhitened residual $\{e_t\}$.

The VAR model (12) is used here simply as a tool to get residuals which possess a decreased temporal dependence. In general, the prewhitening procedure leaves a process whose spectrum is flatter at the locality of the origin, which may be estimated with a smaller error. Clearly, this does not necessarily imply that we may more precisely estimate the spectrum of the original series, due to the errors involved in the estimation of the VAR coefficients used to the recoloring procedure. Nevertheless, Andrews and Monahan (1990) report a convincing evidence that the VAR prewhitening offers some significant improvement.

To employ the CCR procedure (and the method by Phillips and Hansen (1990)) introduced in the previous section, we need a consistent estimate of Γ in (5) (more precisely, Λ_2), as well as that of the asymptotic variance Ω of $\{w_t\}$ in (6). As one may well expect, the VAR prewhitening method can also be used to more efficiently estimate one-way spectrum Γ (and Λ) in (5). To show the relationship between Γ of $\{w_t\}$ and the corresponding parameter for the prewhitened series $\{e_t\}$ in (12), we let Γ^0 be defined for the residual $\{e_t\}$ similarly as Γ for $\{w_t\}$ in (5). Also, define

$$\Sigma_k = \lim_{n\to\infty} \frac{1}{n} \sum_{t=k+1}^{n} E(w_t w_{t-k}') \tag{14}$$

which can be consistently estimated by the corresponding product moments of $\{w_t\}$. Notice that $\Sigma_0 = \Sigma$ defined in (4).

It is rather straightforward to deduce

Proposition 2 *Let the notation be defined as above. Then we have*

$$\Gamma = \Phi(1)^{-1}\Gamma^0\Phi(1)^{-1\prime} + \Phi(1)^{-1}\sum_{i=0}^{p-1}\sum_{j=i+1}^{p} \Sigma_i'\Phi_j$$

where $\Phi(1) = \mathrm{I} - \sum_{k=1}^{p} \Phi_k$.

A consistent estimator of the one-way spectrum Γ of the original series $\{w_t\}$ can therefore be obtained essentially from that of the prewhitened residual $\{e_t\}$ through recoloring, using the estimated VAR coefficients Φ_k for $k = 1, \ldots, p$.

Once a consistent estimate for Γ is obtained, it is easy to estimate other parameters Ω and Λ in (7) consistently. In particular, Ω can be estimated from the relationship

$$\Omega = \Sigma + \Gamma + \Gamma' \tag{15}$$

It can also be estimated directly from the estimated longrun variance of the residual $\{e_t\}$, which we denote by Ω^0, using the relationship

$$\Omega = \Phi(1)^{-1}\Omega^0\Phi(1)^{-1\prime} \tag{16}$$

as suggested in Andrews and Monahan (1990).

We may now easily incorporate the VAR prewhitening procedure in the CCR method (or the method by Phillips and Hansen (1990)). First, obtain by the OLS on M(a) an estimate of Π and the fitted $\{w_t\}$. Estimate Σ_k's by the corresponding product moments of the fitted $\{w_t\}$. Second, fit the VAR model (12) to get the OLS estimates for Φ_k for $k = 1, \ldots, p$ and the fitted values of the residual $\{e_t\}$. Then estimate Γ^0 (or Γ^0 and Ω^0) applying any of the available methods for kernel estimation to the fitted residuals. Third, recover Γ and Ω using the relationships in Proposition 2 and (15) or (16). The corresponding estimates of Λ_2 and $\Omega_{12}\Omega_{22}^{-1}$ required in the CCR transformations (9) and (10) can easily be obtained.

The typical estimator of Γ^0 is of the form

$$\hat{\Gamma}^0 = \frac{1}{n} \sum_k c(k) \sum_{t=k+1}^{n} e_t e'_{t-k} \tag{17}$$

where $c(k)$ is a weight function, or a kernel. Usual kernels are truncated by the bandwidth parameter, l, say, so that $c(k) = 0$ for $k > l$. The bandwidth parameter l may be selected a priori, or left to be determined by some data-dependent scheme, as proposed by Andrews (1990). The reader is referred to Andrews (1989) and the references cited there for a detailed discussion on the estimation of Γ^0.

When the CCR and the other existing 'nonparametric' methods are employed in the nonparametric context as for M(a), the first order VAR may often be a reasonable choice for the model used for prewhitening. When the prewhitening is based on the first order VAR, the result for Γ in Proposition 2 is simplified. If we let Φ be the coefficient matrix for the first order VAR (i.e., $\Phi = \Phi_1$), then it follows that

$$\Gamma = (\mathrm{I} - \Phi)^{-1}\Gamma^0(\mathrm{I} - \Phi')^{-1} + (\mathrm{I} - \Phi)^{-1}\Phi\Sigma$$

High order VAR's may well be preferred, of course, if the DGP is given by M(b). In this case, $\{w_t\}$ is truly generated as a VAR of order p, and the complete prewhitening is achieved by fitting the VAR (12). As a result, the residual $\{e_t\}$ in (12) becomes white noise, and we have

$$\Gamma^0 = 0 \tag{18}$$

Moreover, the specification of a cointegrated model as the ECM in M(b) implies the restriction

$$\Phi_p = \begin{pmatrix} \Phi_{p1} & 0 \\ \Phi_{p2} & 0 \end{pmatrix} \tag{19}$$

on the VAR coefficient Φ_p. The restrictions (18) and (19) are precisely the additional informations provided by the parametric specification of the shortrun dynamics in M(b). These restrictions can of course be easily imposed, when we estimate the shortrun dynamics based on VAR (12).

The formula for Γ in Proposition 2 is simplified correspondingly for model M(b) to

$$\Gamma = \Phi(1)^{-1} \sum_{i=0}^{p-1} \sum_{j=i+1}^{p} \Sigma_i' \Phi_j$$

Also, we have

$$\Omega^0 = \Sigma^0$$

for the relationship (16), where Σ^0 is the usual variance of the residual $\{e_t\}$.

4. Simulation Results

In this section, Monte Carlo methods are used to examine the finite sample effect of the VAR prewhitening procedure introduced in the previous section. The finite sample performance of the CCR method is evaluated under various VAR prewhitening schemes, and also compared with those of the usual OLS and the ML procedure based on the ECM by Johansen (1988, 1989). Since we only consider here cointegrating models without any restriction on the longrun coefficients, the ML method by Park (1990b) is not required. For the simulations, we use the bivariate model

$$\triangle z_t = \alpha\beta' z_{t-1} + C\triangle z_{t-1} + e_t \tag{20}$$

where $z_t = (y_t, x_t)'$ and $\beta = (1, -\pi)'$. The errors $\{e_t\}$ are generated by

$$e_t = \epsilon_t + D\epsilon_{t-1} \tag{21}$$

where $\{\epsilon_t\}$ are standard normals with covariance matrix Υ. Here and elsewhere in this section, we use the lower case letters to denote the parameters defined in previous sections, whenever they are vectors or scalars.

More explicitly, the models that we use for simulations are given by

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \tag{22}$$

for α in (20), where we let $\alpha_1 = -0.2$ and $\alpha_2 = 0, 0.2, 0.4$ or 0.6. We denote the DGP's corresponding each of these values of α_2 by (i), (ii), (iii) and (iv). We set $\pi = 1$. Also, the covariance matrix Υ of $\{\epsilon_t\}$ in (21) is specified as

$$\Upsilon = \begin{pmatrix} 1 & \upsilon \\ \upsilon & 1 \end{pmatrix} \tag{23}$$

where $\upsilon = -0.3, 0, 0.3$. Similarly as before, we denote the models for each of these values of υ by a, b and c. Taken (22) and (23) together, the DGP's are signified as (ia) and so on.

We consider three different types of models: VAR(1), VAR(2) and ARMA(1,1). For the VAR(1) model, we let $C = D = 0$ with other parameters given as above. The coefficient matrix C in the VAR(2) model is set

$$C = \begin{pmatrix} -0.4 & 0 \\ 0 & 0.2 \end{pmatrix}$$

with $D = 0$. The ARMA(1,1) model is specified by $C = 0$ and

$$D = \begin{pmatrix} 1 & 0.2 \\ -0.4 & 1 \end{pmatrix}$$

The parameters C and D in the VAR(2) and ARMA(1,1) models are chosen so that the theoretical asymptotic variances of the estimators in these models are largely the same as those in the VAR(1) model.

The asymptotic variances for the CCR and ML estimators for the DGP's given by (20) and (21) are easy to obtain from the results in Lemma 1 and (16). To derive the asymptotic variances more explicitly, let us define

$$\omega_*^2 = \frac{\omega_{11}}{\omega_{22}} - \frac{\omega_{12}^2}{\omega_{22}^2} \tag{24}$$

using the notation defined in (7). Moreover, denote by W_1 and W_2 two independent standard Brownian motions on the unit interval, and define $\overline{W}_2$ to be the demeaned

Brownian motion given by

$$\overline{W}_2 = W_2 - \int_0^1 W_2(t)\,dt$$

In our simulation, individual integrated series are demeaned, prior to analyses, to avoid the dependency of the results on the initial values. Clearly, this is equivalent to including a constant in cointegrating regressions. The asymptotic distributions of the estimators are consequently represented by $\overline{W}_2$, rather than W_2.

As shown in Park (1990b), the CCR and ML estimators are asymptotically equivalent, and have the same limiting distribution. If we let $\hat{\pi}$ be the CCR or ML estimator of the cointegrating coefficient π, then we have

$$n\,(\hat{\pi} - \pi) \xrightarrow{\mathcal{D}} \omega_* \frac{\int_0^1 \overline{W}_2\,dW_1}{\int_0^1 \overline{W}_2^2} \tag{25}$$

where n is the sample size and ω_*^2 is defined in (24). The asymptotic variance of the CCR or ML estimator can therefore be easily obtained for given n. We found through simulation the value of

$$\operatorname{var}\left(\frac{\int_0^1 \overline{W}_2\,dW_1}{\int_0^1 \overline{W}_2^2}\right) = \mathrm{E}\left(\int_0^1 \overline{W}_2^2\right)^{-1}$$

to be about 10.78, which we used to compute the reported asymptotic variances.

We consider the following estimators:

	Nonparametric Estimators
CCR	CCR estimator with fixed bandwidth
CCR_0	CCR estimator with automatic bandwidth
CCR_1	CCR estimator with pseudo-VAR(1) prewhitening and automatic bandwidth
	ECM-based Estimators
ML_p	ML estimator based on VAR(p) model
CCR_p^*	CCR estimator based on VAR(p) model

We use in the simulation the Parzen window for the CCR, CCR_0 and CCR_1 estimators. In the prewhitening procedure for the CCR_1 estimator, an adjustment for the estimated VAR coefficient (Φ) is made as suggested by Andrews and Monahan (1990), to avoid the near-singularity of the longrun impact matrix ($I-\Phi$). The actual threshold point used is 0.95. The reader is referred to their paper for the detailed explanation of this procedure. No such adjustment is made for the CCR_p^* estimator based on the VAR(p) model. One step iteration is done for the CCR_p^* estimators, as suggested in Park (1990a). Our simulation results are reported in Tables 1 – 6.

The Effect of VAR *Prewhitening*

The effect of the VAR prewhitening on the finite sample performance of the cointegrating coefficient estimator can be seen by comparing the results for the CCR_0 and CCR_1 estimators. In all three different models, i.e., VAR(1), VAR(2) and ARMA(1,1), and across various DGP's, the VAR prewhitening procedure offers an unambiguous and substantial improvement. This is so for both $n = 100$ and $n = 300$. It is expected that methods of estimating shortrun dynamics have a decreasing effect on the efficiency of the longrun estimators, as the sample size increases. In asymptotics, the theory suggests that they have no effect at all. Nevertheless, the use of a more efficient estimation method for the shortrun dynamics seems to be of great practical importance for the samples as moderately large as $n = 300$.

It appears in our simulation that the VAR prewhitening method works slightly better for the models with unknown AR components than those including MA part. The overall performance of the CCR_1 estimator is somewhat better in the VAR(2) model than in the ARMA(1,1) model, relative to the CCR_0 estimator. The residuals in the prewhitening VAR in these two models have, respectively, AR and MA components. This is perhaps because the automatic bandwidth selection procedure by Andrews (1990), used in our simulation, is designed to more effectively deal with unknown AR components. In the pseudo-VAR model, the coefficient estimates are not consistent for the true coefficient values of the AR components, and the actual prewhitened residuals in these models therefore have an ARMA structure in general.

The prewhitened residuals are, however, expected to have a smaller MA components for the VAR(2) models.

The improvement due to the use of the VAR prewhitening procedure is in most cases somewhat more drastic, if we compare biases. This is consistent with the finding by Andrews and Monahan (1990) that the VAR prewhitening yields spectrum estimates which are less biased and more concentrated around the true value. In contrast to their result, however, we observed that the prewhitening also decreases MSE's of the estimates of the longrun parameters substantially in virtually all the cases that we considered in the paper. They reported deteriorating effect of the prewhitening on MSE's of the spectrum estimates.

Shortrun Information: How Important in Finite Samples

Now we look at how important it is to use the shortrun information for the estimation of a cointegrated model. For this, we first compare the results for the estimator CCR_1 with those for the CCR_1^* and CCR_2^* estimators, respectively, for the VAR(1) and VAR(2) models. The CCR_1^* and CCR_2^* estimators use the exact specification of the shortrun dynamics, and the restrictions (18) and (19) are imposed. On the other hand, the CCR_1 estimator is nonparametric, and does not rely on any parametric specification of the shortrun dynamics. It only uses the pseudo-VAR(1) model for prewhitening.

It is clearly seen in our simulation results that the use of the information on shortrun dynamics, if available, is extremely important. In both the VAR(1) and VAR(2) models, the CCR procedures in parametric form yield significantly smaller biases and MSE's than those based on the pseudo-VAR prewhitening. The reduction in biases and MSE's is often bigger than fifty percent when the sample size $n = 100$. The importance of the use of shortrun information is expected to decrease, as the sample size increases. However, the relative performance of the CCR_1^* and CCR_2^* estimators respectively in the VAR(1) and VAR(2) models is significantly better than the nonparametric CCR_1 estimator even when $n = 300$.

It may also be interesting to compare the results for the CCR and CCR_0 estimators. In a sense, the automatic bandwidth selection procedure employed in the CCR_0 estimator uses some shortrun information. The procedure, of course, does not rely on any parametric specification of the shortrun dynamics. It, however, determines the bandwidth using the shortrun information provided by the data. The CCR_0 estimator that we considered in the simulation, in this sense, is compared with the CCR estimator, for which the bandwidth is fixed *a priori*. The CCR_0 estimator performs better than the CCR estimator in most cases. It seems apparant that the use of the shortrun information in the way of the automatic bandwidth selection generally has a positive effect on estimating the longrun parameters.

Our strong evidence for the overall positive effect of the use of shortrun information, however, should not be used generally to argue favorably for a 'parametric' estimator against any other 'nonparametric' estimators. The exact ML estimator, which utilizes not only the structure of the shortrun dynamics but also the Gaussianity of the error distribution, does not seem to be effective in using shortrun information in small samples. The exact ML estimator based on the precise parametric specification of the shortrun dynamics performs in small samples in no sense better than the nonparametric CCR_1 estimator, for instance. The ML estimator indeed performs very poorly in small samples.

In particular, the ML estimator can be extremely unreliable for samples of relatively small sizes. Even the exact ML_1 and ML_2 estimators, respectively for the VAR(1) and VAR(2) models, often show rather irratic behavior. In our simulation for $n = 100$, they yield untolerably large MSE's in several cases. It is suspected that the ML estimator does not have finite variance in some of these cases. The ML estimator behaves badly especially when, though not exclusively, the underlying model implies large asymptotic variance for the estimator. Complexity in the structure of shortrun dynamics usually worsen the poor performance of the ML estimator in small samples. The problem becomes worse for samples of a smaller size. Overparametrization also has a severe adverse effect, as we will discuss subsequently in more detail.

Our results on the ML estimator for $n = 100$ are consistent with those in Gonzalo (1989). In his result for $n = 100$, the ML estimator has significantly larger MSE's than the OLS estimator, in all five cases that he considered. Stock and Watson (1991) also observed large finite sample variance of the ML estimator in their simulation. The finite sample problem of the ML estimator disappears, as is expected from the asymptotic theory, as the sample size increases. When $n = 300$, as Gonzalo (1989) reported, the ML estimator behaves well. Its performance is largely comparable to that of the CCR estimator based on the same model.

Misspecification and Overparametrization

The CCR_p^* and ML_p estimators are based on VAR of order p. In practice, however, the underlying DGP is typically unknown. It would therefore be interesting to see the effects of misspecification and overparametrization of the underlying VAR model on the longrun parameter estimates. The effect of misspecification can be looked at from the results for the CCR_1^* and ML_1 estimators in the VAR(2) and ARMA(1,1) models. The finite sample performance of the estimators in these two models show the potential adverse effects of the misspecifications of the AR and MA parts, respectively. The effect of overparametrization, on the other hand, can be inferred from the results for the CCR_4^* and ML_4 estimators in the VAR(1) and VAR(2) models.

In our simulation results, it is clearly shown that both the misspecification and overparametrization of the shortrun dynamics can have significant adverse effects on the longrun parameter estimates in finite samples. As is well expected, misspecification has generally a more significant adverse effect on the bias, than on the MSE. In sharp contrast, the magnitude of the adverse effect of overparametrization is much larger for the MSE than for the bias. These are so, for both the samples of size $n = 100$ and $n = 300$. The comparisons between the effects of the misspecification and overparametrization, however, diverge for the two different sample sizes.

When we consider the samples of size $n = 100$, the overall adverse effect of misspecification on the MSE's of the CCR and ML estimators does not appear to stand out. For the VAR(2) and ARMA(1,1) models, the ML_1 estimators based on the misspecified models yield smaller MSE's than the estimators on the true model, or on a model which better approximates the true model. Similarly, the CCR_1^* estimator performs mostly better than the CCR_4^* estimator in the ARMA(1,1) model. When the true model is generated as VAR(2), the former based on the misspecified VAR(1) model yielded in one case smaller MSE than the latter based on the overparametrized VAR(4) model. In small samples, it seems that the positive effect of parsimonious specification may well dominate the negative effect of misspecification.

This is especially so for the ML procedure. Using a parametric model that involves many unknown parameters dramatically increases the small sample variance of the ML estimator. The ML_4 estimator yields variances incomparably larger than the ML_1 estimator, in the VAR(1) model. For the VAR(2) model, the ML_1 estimator based on a model that is misspecified but contains less unknown parameters indeed often behaves much better than the exact ML_2 estimator, across various DGP's that we investigate. Moreover, in the ARMA(1,1) model, the ML_4 estimator is in virtually all cases outperformed by the ML_1 estimator, which is based on a model futher away from the true model.

The adverse effect of misspecification on the bias is more conspicuous in our simulation results. For the VAR(2) model, both the CCR_1^* and ML_1 estimators yield substantially larger finite sample biases than the corresponding estimators, CCR_2^* and ML_2, based on the true model. On the contrary, overparametrization does not generally have a significant adverse effect on the finite sample bias. The CCR_4^* and ML_4 estimators in the VAR(1) and VAR(2) models have biases not much greater, on the average, than the estimators based true models. They also have biases largely comparable to those of the CCR_1 and ML_1 estimators in the ARMA(1,1) model, respectively.

A quite different picture emerges, when we increase the sample size to $n = 300$. For the models and estimators that we consider in the simulation, overparametrization clearly seems less problematic than misspecification, even in terms of MSE's. The results largely coincide with the asymptotic theory. For the CCR procedure, the

CCR_4^* estimator based on the overparametrized model behave not much worse than the CCR_1^* and CCR_2^* estimators respectively in the VAR(1) and VAR(2) models. It is also seen that the CCR_4^* estimator, in the VAR(2) model, outperforms the CCR_1^* estimator. Largely the same results are observed for the ML procedure. The nonparametric CCR_1 estimator performs better in many cases than the other estimators, both the CCR and ML estimators, based on incorrectly specified models.

Our results for $n = 300$ should, however, be interpreted with caution in this context. What is observed is that overparametrization up to the order of VAR(4) looks fine, and not that overparametrization is generally not problematic, when the sample size is as big as 300. It seems obvious that overparametrization using higher order VAR's can still have a serious problem, even when the sample size is moderately large. It is indeed suggested on a theoretical ground that the order of VAR be increased at some rates of the sample size, to better approximate the unknown underlying DGP. The strategy is commonly employed in practical applications. Long VAR's are indeed frequently used. For long VAR's with the order increased proportionately to the sample size, our results for $n = 100$ may well be more relevant.

5. Conclusion

We have considered the VAR prewhitening method to estimate critical shortrun parameters in cointegrated models, which is required to implement commonly used nonparametric methods of inference. This paper was motivated by a recent study by Andrews and Monahan (1990) on an improved estimation of spectrum in stationary regressions. It has been shown in the paper that the nonparametric methods by Park (1990a), Phillips (1988, 1989) and Phillips and Hansen (1990) can easily be employed in a parametric form, through the complete prewhitening of the stationary components of cointegrating models given in parametric ECM form. The nonparametric methods are directly comparable in this case to the ECM-based methods by Johansen (1988, 1989) and Park (1990b). Yet, the VAR prewhitening procedure allows us to analyze the spectrum of the prewhitened residuals, and provides very simple and effective methods of inference in cointegrated models, which are valid against various potential misspecifications.

To see the effect of the VAR prewhitening, based on both the true and pseudo-VAR models, we performed a rather extensive Monte Carlo simulation. It was found that the VAR prewhitening in general provides substantial finite sample efficiency gains. The CCR estimator by Park (1990a) based on a pseudo-VAR model performs better, significantly in most cases, than the original estimator with no prewhitening procedure. This is so, in terms of both bias and MSE. When the true structure of the shortrun dynamics is assumed to be known, the CCR estimator with the complete VAR prewhitening behaves truly well. In particular, the resulting estimator outperforms the exact ML estimator by Johansen (1988, 1989) in small samples, clearly and often substantially, in terms of MSE.

The ML method by Johansen (1988, 1989) indeed appears to have a serious small sample problem. When the underlying model implies large asymptotic variance for the estimator, the ML method very often produces nonsensical estimates. It seems extremely sensitive to outliers. The poor performance of the ML estimator in small samples rapidly gets worse, as the order of the underlying VAR increases. In contrast, the CCR procedure with VAR prewhitening performs very well also in small samples. Almost all the cases that we have considered (including those for which we did not report the details in the paper), the CCR estimator yields MSE's which are reasonably close to the theoretical asymptotic variances. Unknown shortrun dynamics seem to be well taken care of by the mixture of parametric and nonparametric adjustments that we have proposed in the paper, even for the samples of relatively small sizes.

Appendix: Mathematical Proofs

Proof of Lemma 1. Let

$$C(L) = I - C_1 L - \cdots - C_{p-1} L^{p-1}$$

and rewrite the ECM in M(b) as

$$HC(L)H^{-1} H\triangle z_t = HAu_{t-p} + H\epsilon_t$$

Notice that $H\triangle z_t = (\triangle u_t', \triangle x_t')'$, and therefore,

$$HC(L)H^{-1} H\triangle z_t = HC(L)H^{-1} G(L)w_t$$

where

$$G(L) = \begin{pmatrix} (1-L)I & 0 \\ 0 & I \end{pmatrix}$$

The result in Lemma 1 follows, after some trivial algebra, directly from this.

Proof of Proposition 2. Define

$$\Sigma_k^0 = \lim_{n\to\infty} \frac{1}{n} \sum_{t=k+1}^{n} E(e_t e_{t-k}') \tag{A1}$$

for $\{e_t\}$ in (12), similarly as Σ_k in (14) for $\{w_t\}$. Moreover, let

$$\Sigma_k^c = \lim_{n\to\infty} \frac{1}{n} \sum_{t=k+1}^{n} E(e_t w_{t-k}') \tag{A2}$$

In the notation used in the text, we have

$$\Gamma = \sum_{k=1}^{\infty} \Sigma_k \quad \text{and} \quad \Gamma^0 = \sum_{k=1}^{\infty} \Sigma_k^0 \tag{A3}$$

We also define

$$\Gamma^c = \sum_{k=1}^{\infty} \Sigma_k^c \tag{A4}$$

Notice that

$$\Sigma_1^c = \cdots = \Sigma_p^c = 0 \tag{A5}$$

by (13).

Post-multiplying the equation (12) by $\{w'_{t-k}\}$ and taking limits to the "average" expectations as in (13) and above definitions of Σ_k^0 and Σ_k^c in (A1) and (A2), we may easily deduce that

$$\Sigma_k = \Phi_1 \Sigma_{k-1} + \cdots + \Phi_p \Sigma_{k-p} + \Sigma_k^c \tag{A6}$$

It follows from the definitions of Γ in (A3) and Γ^c in (A4) that

$$\Gamma = \Phi(1)^{-1} \sum_{i=0}^{p-1} \sum_{j=i+1}^{p} \Sigma_i' \Phi_j + \Phi(1)^{-1} \Gamma^c \tag{A7}$$

by taking summation both sides of (A6), term by term, with respect to k from 1 to ∞, and rearranging terms.

Moreover, if we premultiply $\{e_t\}$ to the equation

$$w'_{t-k} = w'_{t-k-1}\Phi'_1 + \cdots + w'_{t-k-p}\Phi'_p + e'_{t-k}$$

and take the limits to the average expectations term by term as above, it can be easily deduced that

$$\Sigma_k^c = \Sigma_{k+1}^c \Phi'_1 + \cdots + \Sigma_{k+p}^c \Phi'_p + \Sigma_k^0 \tag{A8}$$

Now summing up each term in (A8) with respect to k from 1 to ∞, we have from (A5)

$$\Gamma^c = \Gamma^0 \Phi(1)^{-1\prime} \tag{A9}$$

The stated result in part (b) is now immediate from (A7) and (A9).

References

[1] Andrews, D.W.K. (1990). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," Cowles Foundation Discussion Paper No. 877R, Yale University.

[2] Andrews, D.W.K. and J.C. Monahan (1990). "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator," Cowles Foundation Discussion Paper No. 942, Yale University.

[3] Cooley, T.F. and M. Ogaki (1990). "Estimating a labor supply model using cointegration restriction," mimeograph, University of Rochester.

[4] Engle, R.F. and C.W.J. Granger (1987). "Co-integration and error correction: Representation, estimation and testing," *Econometrica*, 55, 251-276.

[5] Gonzalo, J. (1989). "Comparison of five alternative methods of estimating long run equilibrium relationships," UCSD Discussion Paper No. 89-55.

[6] Gregory, A.W., A.R. Pagan and G.W. Smith (1990). "Estimating linear quadratic models with integrated processes," Rochester Center for Economic Research Working Paper No. 247, University of Rochester.

[7] Inder, B (1990). "Estimating long run relationships in economics: A comparison of different approaches," mimeograph, Monash University.

[8] Johansen, S. (1988). "Statistical analysis of cointegrating vectors," *Journal of Economic Dynamics and Control*, 12, 231-254

[9] Johansen, S. (1989). "Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models," University of Copenhagen, Institute of Mathematical Statistics Preprint No. 3.

[10] Ogaki, M. and J.Y. Park (1990) "A cointegrating approach to estimating preference parameter," Rochester Center for Economic Research Working Paper No. 236, University of Rochester.

[11] Park, J.Y. (1990a). "Canonical cointegrating regressions," CAE Working Paper No. 88-29R, Cornell University, forthcoming in *Econometrica.*

[12] Park, J.Y. (1990b). "Maximum likelihood estimation of simultaneous cointegrated models," mimeograph, University of Aarhus.

[13] Phillips, P.C.B. (1988). "Spectral regression for cointegrated time series," Cowles Foundation Discussion Paper No. 872, Yale University.

[14] Phillips, P.C.B. (1989). "Optimal inference in cointegrated systems," Cowles Foundation Discussion Paper No. 866R, Yale University, forthcoming in *Econometrica.*

[15] Phillips, P.C.B. and B.E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies*, 57, 99-126.

[16] Stock, J.H. and M.W. Watson (1991). "A simple estimator of cointegrating vectors in higher order integrated systems," mimeograph, University of California, Berkeley and Northwestern University.

Table 1: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 100, DGP: VAR(1)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4
(ia)	−.2822	−.2031	−.1874	−.1474	−.0636	−.0857	.0114	.0562
(iia)	−.4124	−.2836	−.2417	−.1427	−.0295	−.0838	.0332	.0435
(iiia)	−.3151	−.1785	−.1461	−.0588	−.0014	−.0406	.0118	.0223
(iva)	−.2600	−.1180	−.1009	−.0311	−.0017	−.0302	.0052	.0059
(ib)	−.2179	−.1561	−.1409	−.1041	−.0404	−.0604	.0136	.0035
(iib)	−.2536	−.1658	−.1352	−.0750	−.0122	−.0439	.0141	.0349
(iiib)	−.2034	−.1084	−.0858	−.0358	−.0008	−.0242	.0062	.0147
(ivb)	−.1775	−.0750	−.0634	−.0212	−.0011	−.0183	.0035	.0045
(ic)	−.1532	−.1095	−.0970	−.0679	−.0239	−.0398	.0126	−.0042
(iic)	−.1485	−.0937	−.0738	−.0396	−.0058	−.0233	.0060	.0015
(iiic)	−.1232	−.0627	−.0485	−.0211	−.0008	−.0140	.0029	.0036
(ivc)	−.1120	−.0446	−.0379	−.0136	−.0004	−.0109	.0019	.0033

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4	Avar
(ia)	.1307	.0810	.0831	.0634	.0387	.0486	.0551	n.a.	.0246
(iia)	.2366	.1397	.1259	.0738	.0535	.0745	.1426	n.a.	.0501
(iiia)	.1397	.0598	.0509	.0223	.0175	.0267	.0253	.1289	.0153
(iva)	.0956	.0282	.0247	.0085	.0068	.0127	.0071	.0120	.0058
(ib)	.0857	.0572	.0584	.0461	.0365	.0448	.2029	n.a.	.0270
(iib)	.1003	.0589	.0529	.0353	.0309	.0394	.0468	n.a.	.0270
(iiib)	.0634	.0268	.0229	.0130	.0112	.0158	.0124	.0640	.0097
(ivb)	.0474	.0136	.0118	.0058	.0049	.0077	.0052	.0085	.0043
(ic)	.0497	.0369	.0382	.0331	.0323	.0383	.6209	n.a.	.0246
(iic)	.0398	.0248	.0229	.0177	.0172	.0214	.0226	.6310	.0145
(iiic)	.0259	.0118	.0104	.0073	.0066	.0089	.0071	.0727	.0058
(ivc)	.0204	.0063	.0056	.0036	.0031	.0046	.0033	.0062	.0028

Note: The MSE's of the ML_4 estimator reported as *n.a.* are unacceptably large. In these and some of other cases with large reported values, the MSE's do not appear to converge. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.

Table 2: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 100, DGP: VAR(2)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_2^*	CCR_4^*	ML_1	ML_2	ML_4
(ia)	−.3019	−.2472	−.2175	−.2017	−.1739	−.1153	−.1260	−.1484	.0271	.0421
(iia)	−.4986	−.3817	−.3270	−.2881	−.2428	−.1007	−.1400	−.2176	.1977	−.1612
(iiia)	−.4210	−.2709	−.2304	−.1706	−.1482	−.0324	−.0744	−.1387	.0279	.0340
(iva)	−.3696	−.1926	−.1770	−.1047	−.0946	−.0141	−.0529	−.0907	.0081	−.0001
(ib)	−.2578	−.2126	−.1833	−.1704	−.1579	−.0899	−.1010	−.1445	.0333	.0188
(iib)	−.3460	−.2570	−.2130	−.1888	−.1740	−.0563	−.0843	−.1628	.0657	.0115
(iiib)	−.3006	−.1856	−.1568	−.1223	−.1144	−.0215	−.0494	−.1093	.0131	.0232
(ivb)	−.2743	−.1365	−.1297	−.0839	−.0786	−.0109	−.0374	−.0761	.0065	.0086
(ic)	−.2139	−.1786	−.1509	−.1398	−.1455	−.0699	−.0805	−.1425	.0333	.0353
(iic)	−.2409	−.1759	−.1422	−.1273	−.1350	−.0343	−.0535	−.1305	.0519	.0448
(iiic)	−.2117	−.1271	−.1079	−.0871	−.0931	−.0145	−.0331	−.0906	.0135	.0173
(ivc)	−.1978	−.0950	−.0958	−.0641	−.0676	−.0081	−.0261	−.0662	.0047	.0064

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_2^*	CCR_4^*	ML_1	ML_2	ML_4	Avar
(ia)	.1388	.0989	.0907	.0805	.0634	.0486	.0545	.0519	.2420	n.a.	.0157
(iia)	.3113	.2056	.1740	.1480	.1161	.0752	.0990	.1085	n.a.	n.a.	.0607
(iiia)	.2272	.1132	.0934	.0626	.0485	.0268	.0417	.0452	.1315	.4714	.0220
(iva)	.1775	.0618	.0547	.0272	.0212	.0113	.0221	.0200	.0131	.4743	.0091
(ib)	.1043	.0768	.0692	.0619	.0554	.0389	.0451	.0535	.5629	n.a.	.0173
(iib)	.1606	.1033	.0858	.0733	.0662	.0400	.0518	.0638	n.a.	n.a.	.0327
(iiib)	.1223	.0584	.0485	.0351	.0306	.0168	.0242	.0292	.0382	.8565	.0140
(ivb)	.1020	.0340	.0317	.0181	.0152	.0080	.0137	.0145	.0099	.0151	.0067
(ic)	.0742	.0569	.0506	.0453	.0484	.0314	.0362	.0522	.6880	n.a.	.0157
(iic)	.0826	.0527	.0432	.0372	.0403	.0217	.0274	.0399	n.a.	n.a.	.0176
(iiic)	.0635	.0297	.0252	.0191	.0199	.0098	.0136	.0193	.1314	.1449	.0083
(ivc)	.0552	.0179	.0182	.0110	.0108	.0052	.0082	.0105	.0065	.0190	.0044

Note: The MSE's of the ML_2 and ML_4 estimators reported as *n.a.* are unacceptably large. In these and some of other cases with large reported values, the MSE's do not appear to converge. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.

Table 3: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 100, DGP: ARMA(1,1)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4
(ia)	−.2577	−.1980	−.1995	−.1403	−.1199	−.0896	.0133	.0050
(iia)	−.3980	−.3050	−.2748	−.0893	−.0622	−.0940	.0570	.1867
(iiia)	−.2846	−.1960	−.1606	−.0171	.0055	−.0481	.0242	.0139
(iva)	−.2023	−.1214	−.0925	−.0023	.0101	−.0309	.0118	.0171
(ib)	−.2100	−.1606	−.1602	−.1004	−.0830	−.0731	.0112	.0520
(iib)	−.2411	−.1783	−.1529	−.0234	−.0163	−.0555	.0262	.3601
(iiib)	−.1701	−.1124	−.0867	.0013	.0053	−.0314	.0101	.0096
(ivb)	−.1258	−.0726	−.0521	.0051	.0069	−.0214	.0082	−.0016
(ic)	−.1500	−.1141	−.1126	−.0632	−.0469	−.0591	.0116	.0385
(iic)	−.1319	−.0957	−.0776	−.0030	−.0013	−.0349	−.0081	−.0376
(iiic)	−.0927	−.0604	−.0433	.0053	.0046	−.0209	.0103	.0428
(ivc)	−.0711	−.0405	−.0272	.0059	.0048	−.0152	.0058	−.0035

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4	Avar
(ia)	.1126	.0744	.0886	.0656	.0516	.0460	.0392	n.a.	.0190
(iia)	.2269	.1573	.1558	.0764	.0742	.0874	.2932	n.a.	.0585
(iiia)	.1207	.0713	.0644	.0259	.0265	.0339	.0444	.1797	.0201
(iva)	.0630	.0301	.0253	.0102	.0111	.0144	.0114	.9091	.0076
(ib)	.0842	.0596	.0696	.0566	.0449	.0507	.1196	n.a.	.0260
(iib)	.0977	.0687	.0674	.0434	.0435	.0514	.1149	n.a.	.0328
(iiib)	.0497	.0298	.0266	.0152	.0149	.0196	.1277	.0962	.0115
(ivb)	.0274	.0135	.0113	.0065	.0066	.0088	.0070	.0176	.0049
(ic)	.0548	.0433	.0494	.0501	.0433	.0532	.2236	n.a.	.0302
(iic)	.0379	.0292	.0289	.0256	.0252	.0294	n.a.	n.a.	.0181
(iiic)	.0184	.0122	.0110	.0086	.0081	.0109	.0167	n.a.	.0064
(ivc)	.0105	.0059	.0052	.0040	.0038	.0053	.0041	.0100	.0030

Note: The MSE's of the ML_1, ML_2 and ML_4 estimators reported as *n.a.* are unacceptably large. In these and some of other cases with large reported values, the MSE's do not appear to converge. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.

Table 4: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 300, DGP: VAR(1)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4
(ia)	−.3196	−.2157	−.1409	−.0879	−.0132	−.0231	.0043	.0069
(iia)	−.5226	−.3096	−.1812	−.0760	.0054	−.0088	.0149	.0179
(iiia)	−.3771	−.1736	−.0964	−.0245	.0045	−.0027	.0058	.0062
(iva)	−.3038	−.1047	−.0664	−.0122	.0021	−.0031	.0024	.0023
(ib)	−.2461	−.1658	−.1045	−.0582	−.0065	−.0145	.0039	.0062
(iib)	−.2975	−.1689	−.0891	−.0311	.0057	−.0012	.0092	.0111
(iiib)	−.2293	−.1004	−.0517	−.0120	.0037	−.0001	.0044	.0050
(ivb)	−.1977	−.0642	−.0396	−.0069	.0020	−.0009	.0022	.0022
(ic)	−.1721	−.1156	−.0702	−.0344	−.0020	−.0077	.0034	.0056
(iic)	−.1649	−.0909	−.0441	−.0130	.0044	.0011	.0058	.0071
(iiic)	−.1326	−.0558	−.0270	−.0057	.0028	.0009	.0032	.0037
(ivc)	−.1197	−.0370	−.0225	−.0037	.0017	.0001	.0018	.0020

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4	Avar
(ia)	.1839	.0989	.0698	.0461	.0291	.0326	.0295	.0330	.0246
(iia)	.4209	.1864	.1132	.0615	.0520	.0556	.0549	.0627	.0501
(iiia)	.2199	.0641	.0361	.0172	.0159	.0175	.0160	.0179	.0153
(iva)	.1432	.0257	.0157	.0066	.0061	.0069	.0060	.0067	.0058
(ib)	.1208	.0709	.0541	.0383	.0315	.0345	.0326	.0374	.0270
(iib)	.1506	.0681	.0436	.0296	.0282	.0301	.0292	.0333	.0270
(iiib)	.0869	.0260	.0159	.0104	.0100	.0108	.0101	.0113	.0097
(ivb)	.0633	.0117	.0078	.0046	.0044	.0049	.0044	.0049	.0043
(ic)	.0697	.0458	.0389	.0305	.0288	.0314	.0300	.0346	.0246
(iic)	.0529	.0263	.0191	.0154	.0152	.0163	.0156	.0178	.0145
(iiic)	.0318	.0107	.0074	.0060	.0059	.0063	.0059	.0067	.0058
(ivc)	.0247	.0053	.0039	.0030	.0029	.0031	.0029	.0032	.0028

Note: The biases and MSE's reported here are three and nine multiples, respectively, of the actual numbers. The theoretical asymptotic variances are also multiplied by nine, and are the same as those given for $n = 100$. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.

Table 5: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 300, DGP: VAR(2)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_2^*	CCR_4^*	ML_1	ML_2	ML_4
(ia)	-.3666	-.2906	-.1931	-.1743	-.1603	-.0431	-.0497	-.1497	.0072	.0094
(iia)	-.7255	-.4974	-.3245	-.2749	-.2723	-.0290	-.0460	-.2637	.0160	.0181
(iiia)	-.5685	-.3030	-.1958	-.1335	-.1585	-.0038	-.0150	-.1568	.0058	.0060
(iva)	-.4781	-.1893	-.1487	-.0750	-.0989	-.0003	-.0084	-.0987	.0027	.0025
(ib)	-.3146	-.2513	-.1624	-.1530	-.1529	-.0311	-.0373	-.1478	.0073	.0095
(iib)	-.4675	-.3121	-.1890	-.1713	-.1909	-.0114	-.0219	-.1881	.0090	.0101
(iiib)	-.3842	-.1975	-.1232	-.0959	-.1215	-.0018	-.0090	-.1209	.0040	.0041
(ivb)	-.3402	-.1290	-.1055	-.0625	-.0823	-.0003	-.0058	-.0824	.0019	.0018
(ic)	-.2628	-.2124	-.1340	-.1301	-.1484	-.0231	-.0286	-.1470	.0064	.0084
(iic)	-.3097	-.2041	-.1174	-.1140	-.1467	-.0054	-.0121	-.1460	.0052	.0058
(iiic)	-.2600	-.1307	-.0802	-.0693	-.0985	-.0008	-.0056	-.0984	.0026	.0026
(ivc)	-.2374	-.0872	-.0776	-.0495	-.0707	-.0002	-.0040	-.0709	.0013	.0012

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_2^*	CCR_4^*	ML_1	ML_2	ML_4	Avar
(ia)	.2235	.1468	.0906	.0752	.0603	.0281	.0303	.0526	.0208	.0229	.0157
(iia)	.7355	.3913	.2290	.1791	.1618	.0619	.0685	.1542	.0713	.0815	.0607
(iiia)	.4596	.1588	.0897	.0537	.0587	.0233	.0256	.0578	.0249	.0277	.0220
(iva)	.3281	.0679	.0470	.0200	.0241	.0098	.0108	.0239	.0099	.0108	.0091
(ib)	.1686	.1143	.0704	.0614	.0567	.0259	.0278	.0533	.0227	.0250	.0173
(iib)	.3247	.1678	.0939	.0799	.0846	.0338	.0366	.0829	.0374	.0425	.0327
(iiib)	.2194	.0733	.0417	.0302	.0358	.0147	.0158	.0356	.0157	.0175	.0140
(ivb)	.1715	.0343	.0258	.0140	.0171	.0071	.0077	.0171	.0073	.0080	.0067
(ic)	.1208	.0850	.0523	.0474	.0527	.0220	.0237	.0518	.0208	.0235	.0157
(iic)	.1495	.0769	.0424	.0388	.0495	.0184	.0199	.0492	.0199	.0225	.0176
(iiic)	.1042	.0344	.0201	.0166	.0228	.0087	.0093	.0228	.0092	.0102	.0083
(ivc)	.0862	.0172	.0148	.0089	.0122	.0046	.0050	.0122	.0048	.0053	.0044

Note: The biases and MSE's reported here are three and nine multiples, respectively, of the actual numbers. The theoretical asymptotic variances are also multiplied by nine, and are the same as those given for $n = 100$. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.

Table 6: *Finite Sample Bias, MSE and Asymptotic Variance*

Size: 300, DGP: ARMA(1,1)

Bias

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4
(ia)	−.2948	−.2161	−.1592	−.0625	−.0554	−.0299	.0047	−.0038
(iia)	−.5118	−.3573	−.2252	−.0045	−.0093	−.0320	.0142	−.0017
(iiia)	−.3397	−.2076	−.1086	.0217	.0068	−.0156	.0078	−.0051
(iva)	−.2307	−.1209	−.0570	.0167	.0054	−.0100	.0039	−.0048
(ib)	−.2412	−.1763	−.1280	−.0256	−.0333	−.0260	.0051	−.0062
(iib)	−.2876	−.1965	−.1125	.0347	.0025	−.0210	.0081	−.0060
(iiib)	−.1907	−.1144	−.0544	.0271	.0045	−.0120	.0049	−.0060
(ivb)	−.1369	−.0708	−.0308	.0185	.0036	−.0083	.0029	−.0050
(ic)	−.1726	−.1255	−.0900	−.0010	−.0124	−.0250	.0057	−.0107
(iic)	−.1491	−.1016	−.0537	.0333	.0043	−.0168	.0058	−.0092
(iiic)	−.0997	−.0603	−.0265	.0220	.0034	−.0104	.0037	−.0070
(ivc)	−.0746	−.0392	−.0158	.0157	.0028	−.0075	.0025	−.0055

MSE

DGP	OLS	CCR	CCR_0	CCR_1	CCR_1^*	CCR_4^*	ML_1	ML_4	Avar
(ia)	.1555	.0922	.0764	.0374	.0327	.0271	.0234	.0269	.0190
(iia)	.4194	.2413	.1667	.0589	.0672	.0673	.0745	.0742	.0585
(iiia)	.1889	.0876	.0501	.0217	.0229	.0235	.0234	.0239	.0201
(iva)	.0881	.0318	.0165	.0085	.0091	.0090	.0089	.0091	.0076
(ib)	.1162	.0739	.0652	.0379	.0342	.0345	.0324	.0370	.0260
(iib)	.1516	.0902	.0646	.0373	.0374	.0388	.0401	.0410	.0328
(iiib)	.0669	.0328	.0204	.0136	.0128	.0136	.0131	.0137	.0115
(ivb)	.0342	.0135	.0079	.0059	.0057	.0059	.0057	.0060	.0049
(ic)	.0746	.0532	.0521	.0404	.0365	.0398	.0387	.0450	.0302
(iic)	.0508	.0338	.0271	.0226	.0203	.0218	.0215	.0227	.0181
(iiic)	.0222	.0126	.0091	.0078	.0070	.0077	.0072	.0077	.0064
(ivc)	.0120	.0057	.0039	.0036	.0033	.0036	.0034	.0036	.0030

Note: The biases and MSE's reported here are three and nine multiples, respectively, of the actual numbers. The theoretical asymptotic variances are also multiplied by nine, and are the same as those given for $n = 100$. The simulations were based on the samples generated by the random number generator built in the GAUSS-386 program. All the computations were done using programs written in GAUSS. The number of iteration is 5000.