UvA ECONOMETRICS

Discussion Paper: 2013/05

Bias correcting adjustment coefficients in a cointegrated VAR with known cointegrating vectors

K.J. van Garderen and H.P. Boswijk

www.ase.uva.nl/uva-econometrics

Amsterdam School of Economics
Department of Economics & Econometrics
Valckenierstraat 65-67
1018 XE AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Bias Correcting Adjustment Coefficients in a Cointegrated VAR with Known Cointegrating Vectors*

Kees Jan van Garderen
University of Amsterdam

H. Peter Boswijk†
Tinbergen Institute and University of Amsterdam

June 4, 2013

Abstract

The maximum likelihood estimator of the adjustment coefficient in a cointegrated vector autoregressive model (CVAR) is generally biased. For the case where the cointegrating vector is known in a first-order CVAR with no intercept, we derive a condition for the unbiasedness of the maximum likelihood estimator of the adjustment coefficients, and provide a simple characterization of the bias in case this condition is violated. A feasible bias correction method is shown to virtually eliminate the bias over a large part of the parameter space.

1 Introduction

Consider an m-dimensional first-order vector autoregressive (VAR) model in error correction representation

$$\Delta Y_t = \Pi Y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

where ε_t are $(m \times 1)$ mean zero independently normally distributed disturbances with contemporaneous covariance matrix Ω, independent of the observed starting value Y_0. The process is stable when the eigenvalues of the $m \times m$ matrix $(I_m + \Pi)$ are inside the unit circle. If exactly $m - r$ eigenvalues are unity, the matrix Π is of reduced rank r and we write $\Pi = \alpha\beta'$, where α and β are $(m \times r)$-dimensional matrices. If all eigenvalues of $I_r + \beta'\alpha$ are inside the unit circle (so that $\beta'\alpha$ is non-singular), then Y_t is an $I(1)$ process and the model becomes a cointegrated VAR (CVAR). The column vectors of β are cointegrating vectors with the property that for each $j = 1, \ldots, r$, $\beta_j' Y_t$ is a stable process which defines an equilibrium relationship between the variables in Y_t. The equilibrium space is an $(m - r)$-dimensional space orthogonal to β called the attractor set. The components α_{ij} of the adjustment matrix α describe the reaction of variable i to last period's disequilibrium $\beta_j' Y_{t-1}$.

*Helpful comments from Noud van Giersbergen are gratefully acknowledged.

†Corresponding author. Address: Amsterdam School of Economics, Valckenierstraat 65-67, 1018 XE Amsterdam, The Netherlands. Email: h.p.boswijk@uva.nl

We are interested in the bias when α is estimated by maximum likelihood. Even though the asymptotic distribution of $\widehat{\alpha}$ is centered around α (e.g. Johansen (1996), Theorem 13.3), there can be considerable bias in $\widehat{\alpha}$ in small samples, especially when $\beta'\alpha$ is small. We consider the case where β is known, which occurs, e.g., under the Purchasing Power Parity or Forward Rate Unbiasedness hypotheses. In this case there is a simple connection between the bias of $\widehat{\alpha}$ and the bias for the autoregressive parameter in the AR(1) model. This is obvious since pre-multiplication of (1) by β' gives:

$$\beta' Y_t = \rho \beta' Y_{t-1} + \beta' \varepsilon_t. \tag{2}$$

where $\rho = I_r + \beta'\alpha$ is a matrix which describes the memory of the disequilibrium process. If there is only one cointegrating vector then ρ is the scalar autoregressive parameter in an AR(1) of the univariate process $\beta' Y_t$. We could estimate ρ as $\widehat{\rho} = I_r + \beta'\widehat{\alpha}$ and the bias in both estimators is obviously related. The dimension of α is $m \times r$, however, and larger than the dimension of ρ which is $r \times r$, since $m > r$.

When $\rho = 0$ in the univariate AR(1) model without regressors, the OLS estimator for ρ is unbiased, which can be proved using an invariance argument. We can invoke the same argument here to prove the analogous result for the unbiasedness of $\widehat{\alpha}$. In the present context, $\rho = 0$ means that any deviation from equilibrium has no persistence and the expected value of the process in the next period, given the current value, always lies in the equilibrium set for every period t. The process is therefore symmetrically distributed around the equilibrium set and as a consequence the estimator for the adjustment coefficient is unbiased as we shall prove in Section 2. When this condition for unbiasedness is violated, i.e., when $\rho \neq 0$, we show that the bias in $\widehat{\alpha}$ can be expressed in terms of the bias in $\widehat{\rho}$, which leads to a simple bias correction method, illustrated in Section 3.

2 Bias Expressions

For known β, the maximum likelihood estimator of the adjustment parameter matrix α, based on the conditional likelihood (treating the starting value Y_0 as fixed) is given by the least-squares estimator

$$\widehat{\alpha} = \sum_{t=1}^{T} \Delta Y_t Y_{t-1}' \beta \left(\sum_{t=1}^{T} \beta' Y_{t-1} Y_{t-1}' \beta \right)^{-1}. \tag{3}$$

Proposition 1 *The maximum likelihood estimator $\widehat{\alpha}$ is unbiased when $\beta'\alpha = -I_r$.*

Proof. We use a simple invariance argument as highlighted by Kakwani (1967), and used in a slightly different context by Abadir *et al.* (1999). First, substitution of $\Delta Y_t = \alpha\beta' Y_{t-1} + \varepsilon_t$ in (3) gives

$$\widehat{\alpha} = \alpha + \sum_{t=1}^{T} \varepsilon_t Y_{t-1}' \beta \left(\sum_{t=1}^{T} \beta' Y_{t-1} Y_{t-1}' \beta \right)^{-1}.$$

When $\beta'\alpha = -I_r$ so that $\rho = 0$, then $\beta' Y_t = \beta'\varepsilon_t$ for $t = 1, \ldots, T$. Therefore, defining $\varepsilon_0 = Y_0$,

$$\begin{aligned}\widehat{\alpha}(\varepsilon) - \alpha &= \sum_{t=1}^{T} \varepsilon_t \varepsilon_{t-1}' \beta \left(\beta' \sum_{t=1}^{T} \varepsilon_{t-1}\varepsilon_{t-1}' \beta \right)^{-1} \\ &= \varepsilon' A \varepsilon \beta \left(\beta' \varepsilon' B \varepsilon \beta \right)^{-1},\end{aligned}$$

where $\varepsilon = (Y_0, \varepsilon_1, \ldots, \varepsilon_T)'$, a $(T+1) \times m$ matrix, and A and B are $(T+1) \times (T+1)$ matrices:

$$A = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}.$$

Next, define a $(T+1)\times(T+1)$ orthogonal matrix $H = \text{diag}(1, -1, 1, -1, \ldots)$ and let $\tilde{\varepsilon} = H\varepsilon$, such that $\tilde{\varepsilon}$ and ε will have the same distribution whenever the distribution of $\{\varepsilon_t\}_{t=1}^{T}$ is symmetric. (The first row of both ε and $\tilde{\varepsilon}$ is Y_0'.) It is easily checked that $H'AH = -A$ and $H'BH = B$, so

$$\begin{aligned}\widehat{\alpha}(\tilde{\varepsilon}) - \alpha &= \varepsilon' H' A H \varepsilon \beta \left(\beta' \varepsilon' H' B H \varepsilon \beta \right)^{-1} \\ &= -\varepsilon' A \varepsilon \beta \left(\beta' \varepsilon' B \varepsilon \beta \right)^{-1} \\ &= -\left(\widehat{\alpha}(\varepsilon) - \alpha \right).\end{aligned}$$

Since ε and $\tilde{\varepsilon}$ have the same distribution, $\widehat{\alpha}(\varepsilon) - \alpha$ and $-\left(\widehat{\alpha}(\varepsilon) - \alpha\right)$ will also have identical distributions, symmetric around 0. This distribution has finite mean, as follows from the criteria derived by Magnus (1986) for the existence of moments of ratios of quadratic forms in normal vectors. Therefore, $E[\widehat{\alpha}(\varepsilon) - \alpha] = -E[\widehat{\alpha}(\varepsilon) - \alpha]$, which implies $E[\widehat{\alpha} - \alpha] = 0$. $\square$

When $\beta'\alpha \neq -I_r$, and hence $\rho \neq 0_r$, then $\widehat{\alpha}$ is not unbiased. The bias in $\widehat{\alpha}$ is naturally related to the bias in

$$\begin{aligned}\widehat{\rho} &= I_r + \beta'\widehat{\alpha} \\ &= \sum_{t=1}^{T} Z_t Z_{t-1}' \left(\sum_{t=1}^{T} Z_{t-1} Z_{t-1}' \right)^{-1},\end{aligned}$$

where $Z_t = \beta' Y_t$. The question now becomes how to exploit knowledge concerning the bias in $\widehat{\rho}$ for obtaining bias expressions for $\widehat{\alpha}$.

In the past many expressions have been derived for the bias in $\widehat{\rho}$ in the autoregressive model. Early contributions include Marriott and Pope (1954), Kendall (1954) and White (1961) but there are many others. In order to use these results we need the inverse of the bias relation $\beta' E[\widehat{\alpha} - \alpha] = E[\widehat{\rho} - \rho]$. The dimension of ρ is smaller than α and hence the equation $\beta' E[\widehat{\alpha} - \alpha] = E[\widehat{\rho} - \rho]$ has general solution (see e.g. Magnus and Neudecker (1988), p. 37):

$$E[\widehat{\alpha} - \alpha] = \beta \left(\beta'\beta \right)^{-1} E[\widehat{\rho} - \rho] + \beta_\perp \left(\beta_\perp' \beta_\perp \right)^{-1} \beta_\perp' q, \qquad q \in \mathbb{R}^{m \times r},$$

where q in general will depend on the unknown parameters (α, Ω) and the fixed β.

In order to resolve the indeterminacy in q, we write the model with known β as

$$\begin{aligned} Z_t &= \rho Z_{t-1} + u_{1t}, \\ W_t &= \gamma Z_{t-1} + u_{2t}, \end{aligned}$$

with $\gamma = \beta_\perp' \alpha$, $W_t = \beta_\perp' \Delta Y_t$, $u_{1t} = \beta' \varepsilon_t$ and $u_{2t} = \beta_\perp' \varepsilon_t$. Conditional on the initial values we can calculate the maximum likelihood estimates of ρ and γ by OLS since Z_{t-1} is common in both equations. Using the explicit expression for $\widehat{\alpha}$ we have the following relations

$$\begin{aligned} \beta' \widehat{\alpha} &= \widehat{\rho} - I_r, \\ \beta_\perp' \widehat{\alpha} &= \widehat{\gamma}, \end{aligned}$$

where $\widehat{\gamma} = \sum_{t=1}^T W_t Z_{t-1}' \left(\sum_{t=1}^T Z_{t-1} Z_{t-1}' \right)^{-1}$. This relation can be inverted to obtain

$$\widehat{\alpha} = \beta \left(\beta' \beta \right)^{-1} \left(\widehat{\rho} - I_r \right) + \beta_\perp \left(\beta_\perp' \beta_\perp \right)^{-1} \widehat{\gamma}.$$

This leads to the following proposition:

Proposition 2 $E[\widehat{\alpha} - \alpha] = \left(\beta \left(\beta' \beta \right)^{-1} + \beta_\perp \left(\beta_\perp' \beta_\perp \right)^{-1} \delta \right) E[\widehat{\rho} - \rho]$ *where* $\delta = \beta_\perp' \Omega \beta \left(\beta' \Omega \beta \right)^{-1}$.

When the covariance matrix is scalar the second term vanishes since $\beta_\perp' \Omega \beta = 0$ and we have:

Corollary 1 $E[\widehat{\alpha} - \alpha] = \beta \left(\beta' \beta \right)^{-1} E[\widehat{\rho} - \rho]$ *when* $\Omega = \sigma^2 I_m$.

Proof. Using $\widehat{\gamma} - \gamma = \sum_{t=1}^T u_{2t} Z_{t-1}' \left(\sum_{t=1}^T Z_{t-1} Z_{t-1}' \right)^{-1}$ and writing $Z_t = \rho^t Z_0 + \sum_{j=0}^{t-1} \rho^j u_{1,t-j}$ it follows that $\{u_{2t}\}_{t=1}^T$ is independent of $\{Z_{t-1}\}_{t=1}^T$, so that $E[\widehat{\gamma} - \gamma] = 0$, if u_{1t} is independent of u_{2t}. When ε_t is Gaussian with covariance matrix Ω, this happens if and only if $\text{Cov}[u_{1t}, u_{2t}] = \beta' \Omega \beta_\perp = 0$. This proves the corollary when $\Omega = \sigma^2 I_m$.
In other cases we have

$$u_{2t} = \delta u_{1t} + u_{2\cdot 1,t},$$

where $u_{2\cdot 1,t}$ is independent of u_{1t}. Using $(\widehat{\rho} - \rho) = \sum_{t=1}^T u_{1t} Z_{t-1}' \left(\sum_{t=1}^T Z_{t-1} Z_{t-1}' \right)^{-1}$ we have $\widehat{\gamma} - \gamma = \delta \left(\widehat{\rho} - \rho \right) + \sum_{t=1}^T u_{2\cdot 1,t} Z_{t-1}' \left(\sum_{t=1}^T Z_{t-1} Z_{t-1}' \right)^{-1}$, where the last term has expectation 0 because $\{u_{2\cdot 1,t}\}_{t=1}^T$ is independent of $\{Z_{t-1}\}_{t=1}^T$. This leads to the result of Proposition 2. $\square$

We see that the bias in $\widehat{\alpha}$ is proportional to the bias in $\widehat{\rho}$ in the direction of the cointegrating vector, orthogonal to the equilibrium set if the contemporaneous covariance matrix is scalar, and a second term that is governed by the non-orthogonality of β and $\beta_\perp$ in de metric defined by the contemporaneous covariance matrix Ω .

3 Bias Correction

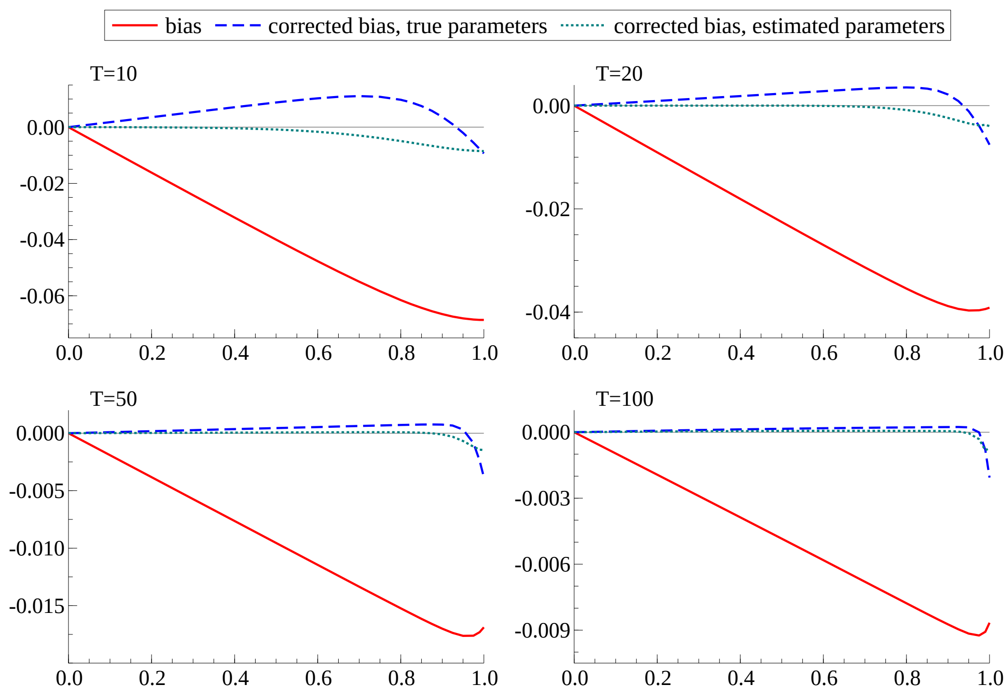
In order to illustrate the result and to show that we can successfully use bias expressions for autoregressive parameters to bias adjust the estimator $\widehat{\alpha}$, we consider a bivariate CVAR with one cointegrating vector $\beta = (1, -1)'$, inspired by e.g. the Forward Rate Unbiasedness hypothesis and present value models. We choose as adjustment vector $\alpha = \frac{1}{2}(\rho - 1)(1, -1)'$, for various values of ρ. The disturbance covariance matrix is taken as $\Omega = \frac{1}{2}\,\mathrm{diag}(1+\delta, 1-\delta)$ with $\delta \in (0, 1)$, such that $\beta'\Omega\beta = 1$ and $\beta_\perp'\Omega\beta = \delta$ (where we have taken $\beta_\perp = (1, 1)'$). The initial condition satisfies $\beta' Y_0 = 0$.

There are various bias expressions for $\widehat{\rho}$, but we use one based on the geometry of the AR(1) model, see van Garderen (1997, 1999) and calculated using general second order bias expression given in, e.g., Amari (1985). For the case where $Z_0 = 0$, this results in the explicit bias formula

$$E[\widehat{\rho} - \rho] = \frac{(1 - \rho^2)\left(4\rho^2 - 2T\rho^2 + 2T\rho^4 - 2T\rho^{2T} - 4\rho^{2+2T} + 2T\rho^{2+2T}\right)}{\rho(T - 1 - T\rho^2 + \rho^{2T})^2} + o(T^{-1}). \quad (4)$$

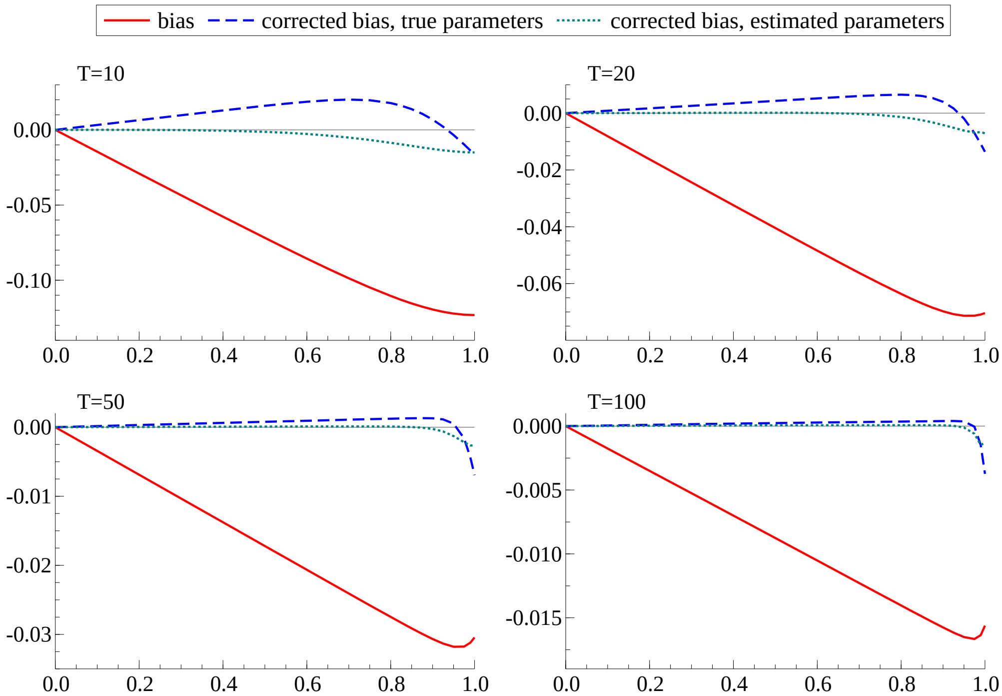
Figures 1–3 display the bias in $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ against $\rho \in [0, 1]$, with $T \in \{10, 20, 50, 100\}$ and $\delta \in \{0, 0.8\}$. When $\delta = 0$, then the distribution of $\widehat{\alpha}_2$ is the same as that of $-\widehat{\alpha}_1$, so this case is not displayed. For similar reasons of symmetry, we do not consider $\rho < 0$ or $\delta < 0$. In addition to the bias, we have calculated the remaining bias after correction using Proposition 2 in combination with (4), either using the true parameter values of ρ and δ, or their estimates, where we have imposed $\widehat{\rho} \leq 1$ by taking $\widehat{\rho} = \min\left\{1, 1 + \beta'\widehat{\alpha}\right\}$.

Figure 1: Bias and corrected bias in $\widehat{\alpha}_1$ against ρ, with $\delta = 0$



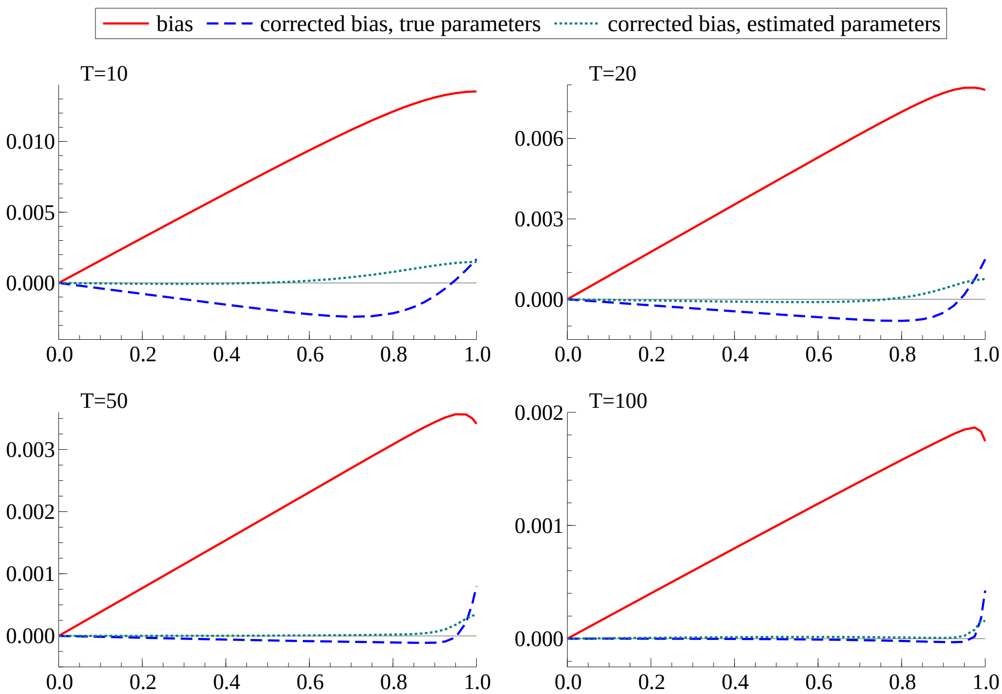
Note: All graphs have ρ on the horizontal axis, and bias on the vertical axis.

Figure 2: Bias and corrected bias in $\widehat{\alpha}_1$ against ρ, with $\delta = 0.8$



Note: All graphs have ρ on the horizontal axis, and bias on the vertical axis.

Figure 3: Bias and corrected bias in $\widehat{\alpha}_2$ against ρ, with $\delta = 0.8$



Note: All graphs have ρ on the horizontal axis, and bias on the vertical axis.

The results are based on $1,000,000$ replications. The same random numbers have been used for different values of ρ, and the result of Proposition 1 (zero bias at $\rho = 0$) has been enforced by taking "antithetic" variates in the spirit of the proof of Proposition 1.

In all three figures, we observe very similar features. The bias increases almost linearly in ρ for small values of ρ, but the function is curved and non-monotonous as ρ approaches 1. The (infeasible) bias correction based on the true parameters leads to an over-correction of the bias for smaller values of ρ and T, and an under-correction in the neighbourhood of $\rho = 1$. For a large part of the interval $\rho \in [0, 1]$, the correction based on estimated parameters leads to an almost unbiased estimator. This may be explained by the fact that the negative bias in $\widehat{\rho}$ reduces the over-correction caused by the approximation (4). As ρ approaches 1, the feasible correction method based on estimated parameters does not fully eliminate the bias, but still leads to a substantial bias reduction.

4 Concluding Remarks

We have shown that in the CVAR model with known β, the bias in $\widehat{\alpha}$ can be related to the bias in $\widehat{\rho}$ in pure (vector) autoregressive models. The bias can be very large relative to the true value of α, in particular for small values of α when return to the equilibrium set is slow and shocks are relatively persistent. Our feasible bias correction significantly reduces the bias of the adjustment estimator.

When the model is extended to include deterministics and lagged differences then the estimator $\widehat{\rho}$ is not unbiased, even when $\rho = 0$, which is well known. This means that Proposition 1 no longer applies; however, we conjecture that Proposition 2 can be extended to the case of deterministic components in the first-order model.

References

Abadir, K. M., K. Hadri and E. Tzavalis (1999), "The influence of VAR dimensions on estimator biases," *Econometrica* 67, 163–181.

Amari, S. (1985), *Differential-Geometrical Methods in Statistics.* Berlin: Springer-Verlag.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive models.* Oxford: Oxford University Press.

Kakwani, N. C. (1967), "The unbiasedness of Zellner's seemingly unrelated regression equations estimators," *Journal of the American Statistical Association* 62, 141–142.

Kendall, M. G. (1954), "Note on bias in the estimation of autocorrelation," *Biometrika* 41, 403–404.

Magnus, J. R. (1986), "The exact moments of a ratio of quadratic forms in normal variables," *Annales d'Économie et de Statistique* 4, 95–109.

Magnus, J. R. and H. Neudecker (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Chichester: John Wiley.

Marriott, F. H. C. and J. A. Pope (1954), "Bias in the estimation of autocorrelations," *Biometrika* 41, 390–402.

van Garderen, K. J. (1997), "Exact geometry of explosive autoregressive models," CORE Discussion Paper 9768.

van Garderen, K. J. (1999), "Exact geometry of first-order autoregressive models," *Journal of Time Series Analysis* 20, 1–21.

White, J. S. (1961), "Asymptotic expansions for the mean and variance of the serial correlation coefficient," *Biometrika* 48, 85–94.