Journal of Econometrics 192 (2016) 64–85

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Inference on co-integration parameters in heteroskedastic vector autoregressions[☆]

H. Peter Boswijk[a], Giuseppe Cavaliere[b,*], Anders Rahbek[c,d], A.M. Robert Taylor[e]

[a] Amsterdam School of Economics and Tinbergen Institute, University of Amsterdam, Netherlands
[b] Department of Statistical Sciences, University of Bologna, Italy
[c] Department of Economics, University of Copenhagen, Denmark
[d] CREATES, Denmark
[e] Essex Business School, University of Essex, United Kingdom

ARTICLE INFO

Article history:
Received 15 November 2013
Received in revised form
24 January 2015
Accepted 30 July 2015
Available online 8 December 2015

JEL classification:
C30
C32

Keywords:
Co-integration
Adjustment coefficients
(un)Conditional heteroskedasticity
Heteroskedasticity-robust inference
Wild bootstrap

ABSTRACT

We consider estimation and hypothesis testing on the coefficients of the co-integrating relations and the adjustment coefficients in vector autoregressions driven by shocks which display both conditional and unconditional heteroskedasticity of a quite general and unknown form. We show that the conventional results in Johansen (1996) for the maximum likelihood estimators and associated likelihood ratio tests derived under homoskedasticity do not in general hold under heteroskedasticity. As a result, standard confidence intervals and hypothesis tests on these coefficients are potentially unreliable. Solutions based on Wald tests (using a "sandwich" estimator of the variance matrix) and on the use of the wild bootstrap are discussed. These do not require the practitioner to specify a parametric model for volatility. We establish the conditions under which these methods are asymptotically valid. A Monte Carlo simulation study demonstrates that significant improvements in finite sample size can be obtained by the bootstrap over the corresponding asymptotic tests in both heteroskedastic and homoskedastic environments. An application to the term structure of interest rates in the US illustrates the difference between standard and bootstrap inferences regarding hypotheses on the co-integrating vectors and adjustment coefficients.

1. Introduction

In this paper we focus on the problem of conducting inference (estimation and hypothesis testing) on the coefficients of the co-integrating relations and associated adjustment parameters, based around the likelihood-based methods of Johansen (1996), in vector autoregressive time series which display time-varying behaviour in the variance of the driving shocks. We allow for both unconditional heteroskedasticity (often referred to as non-stationary volatility in the literature) and conditional heteroskedasticity in our analysis. It is well known that the assumption of conditional homoskedasticity appears inconsistent with financial and macroeconomic data; see, for example, Gonçalves and Kilian (2004). A large body of recent applied work has grown suggesting that the assumption of constant *unconditional* volatility is also at odds with what is observed in the data, with a general decline in the unconditional volatility of the shocks driving macroeconomic series in the twenty years or so leading up to the recent financial crisis, the so-called "Great Moderation", commonly observed; see, for example, *inter alia*, Kim and Nelson (1999) and McConnell and Perez Quiros (2000) and the references therein.

These empirical findings have helped stimulate research into the impact of time-varying conditional and unconditional volatility on standard time series methods. Of most relevance to this paper, Cavaliere et al. (2010b) analyse the impact this has on the conventional co-integration rank pseudo likelihood ratio (PLR) tests of Johansen (1996). They demonstrate that the asymptotic

[☆] We are grateful to the Editor, Peter Robinson, and Associate Editor and four referees for their helpful and constructive comments on an earlier version of this paper. Cavaliere, Rahbek and Taylor thank the Danish Council for Independent Research, Sapere Aude | DFF Advanced Grant (Grant nr: 12-124980) for financial support. Cavaliere also thanks the Italian Ministry of Education, University and Research (PRIN project "Multivariate statistical models for risk assessment") for financial support.

* Correspondence to: Department of Statistical Sciences, Via Belle Arti 41, I-40126 Bologna, Italy. Tel.: +39 051 2098230; fax: +39 051 232153.

E-mail addresses: h.p.boswijk@uva.nl (H.P. Boswijk), giuseppe.cavaliere@unibo.it (G. Cavaliere), anders.rahbek@econ.ku.dk (A. Rahbek), rtaylor@essex.ac.uk (A.M.R. Taylor).

http://dx.doi.org/10.1016/j.jeconom.2015.07.005
0304-4076/

null distributions of the PLR statistics, which are constructed under the assumption that the innovations are i.i.d. and Gaussian, are non-pivotal in the presence of unconditional heteroskedasticity. Cavaliere et al. (2014) [CRT] show that wild bootstrap implementations of the PLR tests are, however, asymptotically valid.[1] Cavaliere et al. (2010a) provide a separate treatment for the case where the shocks are conditionally heteroskedastic but unconditionally homoskedastic. They show that the standard PLR tests (based on asymptotic critical values) are asymptotically valid, but that the corresponding wild bootstrap tests can deliver considerable finite sample improvements.

In this paper we make two distinct contributions to the literature. Utilising a very general set-up which combines the assumptions of Cavaliere et al. (2010a,b) into a unified framework, our first contribution is to examine the impact of time-varying volatility on the large sample properties of the standard likelihood-based methods of estimation and hypothesis testing on the coefficients of the long run relations and the associated adjustment coefficients (β and α, respectively, in standard notation) detailed in Johansen (1996). In particular, we analyse the pseudo maximum likelihood (PML) estimates of these parameters and the associated PLR test for linear restrictions on these parameters, both derived under the assumption of an i.i.d. Gaussian pseudo-likelihood. We also analyse the corresponding Wald statistic, based around a PML ("sandwich") variance matrix estimator. We demonstrate that although the PML estimates are consistent, standard confidence intervals and PLR test statistics based on the PML estimates of α and β will not be reliable in general, their form depending on nuisance parameters arising from any heteroskedasticity in the shocks. Where the shocks are unconditionally homoskedastic, however, inference on β alone is shown to be asymptotically pivotal. For this to hold for the PLR tests involving α, conditional heteroskedasticity must also be absent from the shocks. We show that asymptotically robust inference can be achieved on α, regardless of any heteroskedasticity present, by using the Wald statistic. This also holds when using the Wald statistic to test hypotheses involving β, provided the shocks are unconditionally homoskedastic, but in general is not true when non-stationary volatility is present. These results complement those given in Hansen (1992a) for the case of a single equation error-correction model (as in Engle and Granger, 1987), driven by an error term whose volatility follows a first-order integrated ($I(1)$) process.

Our second contribution is to develop wild bootstrap implementations of the standard PLR and Wald tests. Extant bootstrap methods for testing hypotheses on the co-integration parameters deal with tests on β only and are at most devised for the case of independent, identically distributed shocks; see Omtzigt and Fachin (2006), Cavaliere et al. (2015) and the references therein. In contrast, we derive the conditions under which wild bootstrap implementations of the PLR and Wald tests of hypotheses on both α and β can replicate the first order limiting null distributions of the corresponding standard test statistics. In such cases asymptotically valid bootstrap inference can be performed in the presence of time-varying volatility using the wild bootstrap versions of these tests. For the bootstrap PLR tests involving α this requires the assumption of a further moment condition and the assumption of the absence of asymmetric volatility clustering, as formally defined below after Assumption 2. For the PLR tests involving only β neither of these additional assumptions is required, while for the Wald tests, the additional assumption on the form of the volatility clustering is also not required. When testing joint hypotheses on α and β, statistical leverage effects (defined after Assumption 2) need to be ruled out for bootstrap inference based on PLR and Wald tests.

The remainder of the paper is organised as follows. Section 2 defines the heteroskedastic model, discussing in detail the type of time-varying volatility that we consider. We then characterise the asymptotic behaviour of the common trends in the process. Next, we introduce a class of hypotheses on the co-integrating vectors and error correction coefficients. Section 3 derives the asymptotic null distributions of the PLR and Wald test statistics for the class of hypotheses we consider. The wild bootstrap approach, based on a sieve-type procedure using the PML coefficient matrix estimates from the co-integrated VAR model, is outlined in Section 4. Here the conditions under which the wild bootstrap tests deliver asymptotically valid inference are also detailed. In Section 5 we use Monte Carlo simulation evidence to compare the small sample size properties of the standard (asymptotic) tests and their bootstrap analogues for a variety of heteroskedastic co-integrated VAR models. An empirical application of the proposed methods to the term structure of interest rates in the US is presented in Section 6. Section 7 concludes. All proofs are contained in the Appendix.

In the following '$\xrightarrow{w}$' denotes weak convergence and '$\xrightarrow{p}$' convergence in probability, in each case as the sample size, T, diverges; $\mathbb{I}(\cdot)$ denotes the indicator function and '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x); $\lfloor \cdot \rfloor$ denotes the integer part of its argument. The notation $\mathcal{C}_{\mathbb{R}^{m\times n}}[0,1]$ is used to denote the space of $m \times n$ matrices of continuous functions on $[0, 1]$; $\mathcal{D}_{\mathbb{R}^{m\times n}}[0,1]$ denotes the space of $m \times n$ matrices of càdlàg functions on $[0, 1]$, equipped with the Skorohod metric. The space spanned by the columns of any $m \times n$ matrix A is denoted as $\text{col}(A)$; if A is of full column rank $n < m$, then $A_\perp$ denotes an $m \times (m-n)$ matrix of full column rank satisfying $A_\perp' A = 0$. For any square matrix, A, $|A|$ is used to denote the determinant of A, $\|A\|$ the norm $\|A\|^2 := \text{tr}\left\{A'A\right\}$, and $\rho\,(A)$ its spectral radius (that is, the maximal modulus of the eigenvalues of A). For any vector, x, $\|x\|$ denotes the usual Euclidean norm, $\|x\| := \left(x'x\right)^{1/2}$. Finally, $\otimes$ denotes the Kronecker product.

2. The heteroskedastic VAR model and hypotheses

We consider the following VAR(k) model in error-correction format:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j} + \alpha\rho_1' D_{1t} + \mu_2 D_{2t} + \varepsilon_t,$$
$$t = 1, \ldots, T, \tag{1}$$

where X_t is a p-variate vector process, with initial values $(X_{1-k}, \ldots, X_0)$, which are known and taken to be fixed in the statistical analysis, and D_{1t} and D_{2t} are vectors of deterministic terms, such as a constant or linear trend, of dimensions d_1 and d_2, respectively. The disturbance ε_t is assumed to be a p-variate vector martingale difference sequence relative to some filtration $\mathcal{F}_t$, with finite and positive definite conditional variance matrix. Further conditions on ε_t are discussed below. The parameter matrices α and β, which are our key focus in this paper, are of dimension $p \times r$, with $0 < r < p$, and $\{\Gamma_j\}_{j=1}^{k-1}$ are $p \times p$ lag coefficient matrices. The co-integration rank, r, is assumed to be known in what follows; in practice this would first be determined using the wild bootstrap co-integration rank tests of CRT. The parameter matrices ρ_1 and μ_2 are of dimension $d_1 \times r$ and $p \times d_2$, respectively; note that D_{1t} enters the

[1] The algorithm proposed in CRT generates bootstrap samples using estimates all of which are obtained under the rank restriction imposed by the null, as is also done in Cavaliere et al. (2012), who use an i.i.d., rather than wild, re-sampling scheme. Cavaliere et al. (2010a,b) also propose an alternative algorithm, along the lines of that considered in Swensen (2006) using restricted estimates only for the long run parameters of the model. Cavaliere et al. (2012) and CRT demonstrate that the algorithms they propose are preferable to those proposed in Cavaliere et al. (2010a,b).