Journal of Econometrics 158 (2010) 37–50

Contents lists available at ScienceDirect

Journal of Econometrics

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

Model-based asymptotic inference on the effect of infrequent large shocks on cointegrated variables

Iliyan Georgiev *

Faculdade de Economia, Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal

ARTICLE INFO

Article history:
Available online 11 March 2010

JEL classification:
C32

Keywords:
Cointegration
Vector autoregression
Rare events
Impulse response

ABSTRACT

Quasi-maximum-likelihood (QML) estimation of a model combining cointegration in the conditional mean and rare large shocks (outliers) with a factor structure in the innovations is studied. The goal is not only to robustify inference on the conditional-mean parameters, but also to find regularities and conduct inference on the instantaneous and long-run effect of the large shocks. Given the cointegration rank and the factor order, χ^2 asymptotic inference is obtained for the cointegration vectors, the short-run parameters, and the direction of each column of both the factor loading matrix and the matrix of long-run impacts of the large shocks. Large shocks, whose location is assumed unknown a priori, can be detected and classified consistently into the factor components.

1. Introduction

The presence of occasional large shocks (outliers) in time series data poses two tasks. The traditional one is to develop valid procedures for inference about conditional-mean parameters. A second task is to uncover the regularities, if such exist, in the generation of large shocks and in their effect on the data. From this perspective outlying observations are not a nuisance, but rather, an object of interest. In the analysis of economic data the interest in outliers could be due to their possible relation with institutional and policy interventions, among other reasons.

In this paper, the two inferential tasks formulated above are addressed jointly in the framework of a cointegrated vector autoregressive (VAR) model with Gaussian underlying innovations, occasionally perturbed by influential large shocks obeying a factor structure. For this model the Gaussian QML (equivalently, reduced rank regression) estimators of the cointegration vectors and the short-run parameters do not have the asymptotic distributions derived in Johansen (1996) under the assumption of independent identically distributed (IID) errors with finite variance. In particular, χ^2 asymptotic inference does not apply to these estimators, and the task of robustification is relevant.

The factor structure of the large shocks is introduced as a context to address the second inferential task. As usual in factor models, shocks in the same factor component are thought to be similar. For example, if variables from different countries are studied, one component could contain world shocks, i.e., shocks affecting all countries, a second component could consist of euro-area shocks, while other components could contain country-specific shocks.

In a related paper on cointegration and outliers, Franses and Lucas (1998) also formulate two econometric tasks—one concerning the conditional-mean parameters, and the other one—the outliers. Their formulation carries a different emphasis, however. Most of the time Franses and Lucas do not commit themselves to a statistical model and view their procedure as an auxiliary diagnostic device for the influentiability and the location of outliers. The choice here, to study a class of models, makes it possible to learn about aspects of the outliers beyond their location, and to obtain analytical results that are of interest in applications where a model in the studied class provides a good description of the data. When the latter is not the case, one is (in the best case) back to the diagnostic device.

The main resemblance with Franses and Lucas is that non-Gaussian quasi-likelihood functions are used to define the parameter estimators. It is not required to take a decision in which periods outliers have occurred. Instead, a set of weights reflecting how likely it is outliers to have occurred is assigned implicitly to each period. If desired, these weights can be used to detect the periods with outliers, but this does not affect parameter estimation. On the other hand, a substantial difference from Franses and Lucas is in the scope and the setup of the asymptotic theory. Franses and Lucas do not derive asymptotics and refer to the asymptotic distribution of a quasi-likelihood ratio test on the cointegration rank from Lucas (1997, 1998), whereas here the focus is on asymptotic theory. The results of the two papers are not directly complementary, however, as different asymptotic schemes are used.

* Tel.: +351 213 801 643; fax: +351 213 870 933.
E-mail address: igeorgiev@fe.unl.pt.

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.03.005

Georgiev (2007), hereafter Gv07, gives results that we build upon in the study of factor estimation and inference. That paper, although admitting the possibility of estimating conditional-mean and factor parameters jointly, concentrates on a two-step procedure where inference on the outliers is based on first-step residuals obtained in an *ad hoc* fashion. In contrast, the present paper is entirely set up in the QML framework, which is more appropriate when joint inference on conditional-mean and factor parameters is undertaken. Furthermore, it is more widely applicable: the two-step results in Gv07 apply to the cointegrated VAR model only if outliers are separated by sufficiently large time spans, whereas QML estimators retain many of their asymptotic properties even if outliers occur at consecutive dates. For instance, the QML estimators of the cointegration vectors and of the short-run parameters have mixed Gaussian asymptotic distribution and χ^2 asymptotic inference is regained, as in the Gaussian cointegrated model. The time location of the outliers can be estimated consistently, and in the limit they can be correctly classified into the factor components. Thanks to the large size of the outliers, the direction of each column of the factor loading matrix is estimated consistently, but since outliers are rare, the estimated length of the columns remains random in the limit.

Regarding inference on the loading columns, we go beyond Gv07 and provide χ^2 asymptotic results for hypotheses not only on the instantaneous, but also on the long-run effect of large shocks. This relies on the property of cointegrated models that innovational outliers can cumulate to level shifts in the non-stationary directions.

The paper starts with a formulation of the model in Section 2, where some properties of the generated stochastic process are also addressed. Section 3 discusses the quasi-likelihood function used to define the estimators. The probability analysis in Section 4 is organized according to the standard sequence of existence, consistency and asymptotic distribution of the QML estimator of the conditional-mean parameters, followed by a subsection on the estimator of the factor loading matrix. Proofs are collected in the Appendix.

Throughout, $\mathbb{I}\{\cdot\}$ and $\lfloor\cdot\rfloor$ stand for the indicator function and the integer part function, respectively. The cardinality of a set is denoted by $\#(\cdot)$. A block diagonal matrix with diagonal blocks $A_1, \ldots, A_s$ is written as $diag(A_1, \ldots, A_s)$. $D_n[0, 1]$ denotes the space of càdlàg functions from $[0, 1]$ to $\mathbb{R}^n$, endowed with the Skorokhod topology.

2. The statistical model

The first subsection introduces the standard part of the model, the second one deals with the factor structure.

2.1. The conditional mean

The specification of the conditional mean given the past is in error-correction form:

$$\Delta X_t = \mu + \alpha\beta' X_{t-1} + \sum_{i=1}^{l} \Gamma_i \Delta X_{t-i} + e_{Tt}, \quad t = 1, \ldots, T, \tag{1}$$

where the $n \times 1$ vectors X_t and e_{Tt} are the observable variable and the innovation in period t, respectively, and T is the sample size. Initial values are specified in Section 2.3. The $n \times r$ matrices α and β $(0 < r < n)$, the $n \times n$ matrices Γ_i $(i = 1, \ldots, l)$ and the $n \times 1$ vector μ are assumed to satisfy the following conditions.

Assumption C. a. The rank of α and β is r.

b. The characteristic polynomial $|(1-z)I - \alpha\beta' z - \sum_{i=1}^{l} \Gamma_i (1-z) z^i|$ has all its roots at unity and outside the unit complex disk.

c. The matrix $\alpha_\perp' \Psi \beta_\perp$ is invertible (here $\Psi = I - \sum_{i=1}^{l} \Gamma_i$, and $\alpha_\perp$ and $\beta_\perp$ of dimension $n \times (n-r)$ are orthogonal complements of α and β).

d. $\alpha_\perp' \mu \neq 0$.

The implications of Assumption C for the properties of the process X_t will be discussed after the innovation term is fully specified. Now we note that Assumption C rules out seasonal unit roots and explosive roots (part b), as well as $I(2)$ stochastic trends (part c). Part (d) implies that the constant cumulates to a linear trend in the directions of $\beta_\perp$. This assumption is made for concreteness only; other specifications of the deterministics can be handled similarly by modifying, if necessary, the quasi-likelihood function introduced in Section 3.

The cointegration rank r and the lag length l are considered known. Issues related to their determination are addressed in the Concluding remarks section.

2.2. The innovations

The innovations are specified as $e_{Tt} = \varepsilon_t + T^{1/2} \sum_{i=1}^{k} M_i \pi_{T,ti} \eta_{ti}$, where $\{\varepsilon_t\}$ is an IID sequence, $M = (M_1, \ldots, M_k)$ is a fixed $n \times k$ matrix, $\pi_{T,ti}$ are binary random variables thought to be zero in most periods and η_{ti} are a.s. non-zero random variables. Thus, in most periods the innovation is drawn from the sequence $\{\varepsilon_t\}$, whereas in some periods large shocks are superimposed on this sequence. Specifically, when some $\pi_{T,ti}$ takes the value of one, an extra contribution to the innovation with random size $T^{1/2} M_i \eta_{ti}$ is generated. The normalization with $T^{1/2}$ renders the extra shocks asymptotically influential; it has been used, e.g., in Leybourne and Newbold (2000a,b) and Cavaliere and Georgiev (2007).

Some of the assumptions we make next are understood best from the perspective of factor models. Subsuming the dependence on T in the subscripts, the innovation in time t can be written compactly as $e_t = \varepsilon_t + T^{1/2} M \delta_t$, where $\delta_t = \Pi_t \eta_t$, $\Pi_t = diag(\pi_{t1}, \ldots, \pi_{tk})$, $\eta_t = (\eta_{t1}, \ldots, \eta_{tk})'$. This is the usual factor model notation, with M standing for the matrix of factor loadings. The meaning of the factor structure is that large shocks corresponding to unit values of the same indicator variable are proportional to each other and, under Assumption E, also the converse is true, with probability approaching one.

Assumption E. a. No column of M is proportional to any other column, and in particular, all columns of M are non-zero.

b. $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is an $n \times 1$ doubly infinite sequence of IID $N(0, \Omega)$ random vectors, with Ω strictly positive definite.

c. $\{\eta_t\}_{t=1}^{\infty}$ is a $k \times 1$ sequence of IID random vectors, where $\eta_t = (\eta_{t1}, \ldots, \eta_{tk})'$, and $P(\eta_{1i} = 0) = 0$ for $i = 1, \ldots, k$.

d. For every $i = 1, \ldots, k$, $\{\pi_{ti}\}_{t=1}^{T}$ is an array of independent Bernoulli random variables, and for some non-decreasing continuous function $F_i : [0, 1] \to [0, \infty)$, not identically equal to 0, it holds that $\sum_{t=1}^{\lfloor Tu \rfloor} P(\pi_{ti} = 1) \to F_i(u)$ as $T \to \infty$ for all $u \in [0, 1]$.

e. $\{\varepsilon_t\}_{t=-\infty}^{\infty}$, $\{\eta_t\}_{t=1}^{\infty}$ and $\{\pi_{ti}\}_{t=1}^{T}$ are jointly independent for all T and $i = 1, \ldots, k$.

Remark 2.1. Under Assumption E(e) the occurrence of large shocks of different kinds is independent. Independence between factor components is a usual ingredient of factor models, see, e.g., Harvey (1989, Ch. 8.5). Further, under Assumption E(d) the number of large shocks in factor component i (called in what follows large shocks of kind i, $i = 1, \ldots, k$) converges in distribution, and is, therefore, bounded in probability. Assumption E(d,e) imply that, with probability approaching one, no large shocks of different kinds occur simultaneously anywhere in the sample, and furthermore, the time distance between consecutive large shocks diverges