Journal of Econometrics 164 (2011) 218–238

Contents lists available at ScienceDirect

Journal of Econometrics

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

Model selection criteria in multivariate models with multiple structural changes[☆]

Eiji Kurozumi[a,*], Purevdorj Tuvaandorj[b]

[a] Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan

[b] Department of Economics, McGill University, Canada

ARTICLE INFO

Article history:
Received 10 June 2010
Received in revised form
14 December 2010
Accepted 26 April 2011
Available online 6 May 2011

JEL classification:
C13
C32

Keywords:
Structural breaks
AIC
Mallows' Cp
BIC
Information criteria

ABSTRACT

This paper considers the issue of selecting the number of regressors and the number of structural breaks in multivariate regression models in the possible presence of multiple structural changes. We develop a modified Akaike information criterion (AIC), a modified Mallows' C_p criterion and a modified Bayesian information criterion (BIC). The penalty terms in these criteria are shown to be different from the usual terms. We prove that the modified BIC consistently selects the regressors and the number of breaks whereas the modified AIC and the modified C_p criterion tend to overfit with positive probability. The finite sample performance of these criteria is investigated through Monte Carlo simulations and it turns out that our modification is successful in comparison to the classical model selection criteria and the sequential testing procedure robust to heteroskedasticity and autocorrelation.

1. Introduction

This paper considers the selection of regressors and estimation of the number of structural changes in multivariate regression models in the possible presence of multiple structural changes. Many methods for the selection of regressors have been proposed in the econometric and statistical literature, and it is often the case in practical analyses that the regressors are selected using either testing procedures or model selection criteria. The former methods select the regressors by testing the significance of the coefficients of the regressors and deleting the insignificant coefficients from the models, while the model selection criteria choose the regressors that minimize the given risk functions. The representative model selection criteria in econometric analysis are the Akaike information criterion (AIC) by Akaike (1973), the C_p criterion by Mallows (1973) and the Bayesian information criterion (BIC) by Schwarz (1978) among others. See Burnham and Anderson (2002) and Konishi and Kitagawa (2008) for a general treatment of the model selection criteria.

In addition to the selection of the regressors, we need to consider the possibility of structural changes when we investigate data covering a relatively long sample period. In such a case we usually test for structural changes. Various tests for structural changes have been proposed in the literature, and the most commonly used tests in recent practical analyses are the sup-type test of Andrews (1993) and the exponential and average-type tests of Andrews et al. (1996) among others. These tests assume the null hypothesis of no changes against the alternative of (multiple) change(s), whereas Bai and Perron (1998) and Bai (1999) proposed tests for the null of ℓ breaks against the alternative of $\ell + 1$ breaks for univariate models. These tests are extended to multivariate models by Qu and Perron (2007), who give a comprehensive treatment on the issue of the estimation, inference, and computation in a system of equations with multiple structural changes. Their treatment is general enough in that less restrictive assumptions are placed on the error term and that models such as vector autoregressions (VAR), seemingly unrelated regressions (SURE) and panel data models are included in their setup as special cases. See Perron (2006) for a review of the testing and estimation of structural changes.

Once the evidence of structural breaks is found, the next step is to estimate the number of breaks. Bai (1997b, 1999), Bai and Perron (1998) and Qu and Perron (2007) proposed to implement tests for structural changes sequentially and proved that the estimated number of structural changes is consistent by letting

☆ We are grateful to Professors Takeshi Amemiya (Editor), Kohtaro Hitomi, Yoshihiko Nishiyama, Ryo Okui, Pierre Perron, Zhongjun Qu, Katsumi Shimotsu, Toshiaki Watanabe and anonymous referees for helpful comments. All errors are our responsibility.

* Corresponding author. Tel.: +81 42 580 8787.
E-mail address: kurozumi@stat.hit-u.ac.jp (E. Kurozumi).

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.04.003

the significance level go to zero. Alternatively, in the statistical literature, the model selection criteria have been proposed to select the number of breaks. For independent normal random variables with mean shifts, Yao (1988) and Zhang and Siegmund (2007) derived the modified BIC and Ninomiya (2005) proposed to modify the AIC, while Liu et al. (1997) considered the modified BIC in regression models with i.i.d. regressors. According to these works, the penalty terms of these new criteria are different from those of the corresponding classical criteria because of the irregularity in the change points. Although these results are of interest from a statistical point of view, they cannot be directly applied to economic data because while economic time series variables are typically serially correlated, assumptions such as i.i.d. observations and regressors are made in the above papers. Exceptions are Ninomiya (2006), Davis et al. (2006) and Hansen (2009). Ninomiya (2006) considered the modified AIC in finite order autoregressive models, but derived under the assumption of known variance, while Davis et al. (2006) proposed a model selection criterion based on the minimum description length principle, but only the consistency of the break fraction estimators was discussed. Hansen (2009) established the modified Mallows' C_p criterion but with only a single break being allowed.

In this paper we develop the model selection criteria in multivariate models allowing lagged dependent variables as regressors in the possible presence of multiple structural changes in both the coefficients and the variance matrices. Our criteria have an advantage over the existing ones in that (i) multivariate models are considered, (ii) serial correlation is taken into account in models by allowing serially correlated regressors, including lagged dependent variables, (iii) structural changes in the variance matrices are allowed. We theoretically derive the AIC, C_p criterion and BIC in models with structural changes and show that the penalty terms should be modified compared with those of the corresponding classical ones. We confirm by Monte Carlo simulations that this modification of the penalty terms is very important to correctly select the regressors and the number of structural changes in finite samples.

The rest of this paper is organized as follows. We explain the model and assumptions in Section 2. Section 3 establishes the modified AIC, the modified C_p criterion and the modified BIC with multiple structural breaks and discusses the consistency of these criteria. In Section 4, we investigate the finite sample performance of our model selection criteria via simulations. Concluding remarks are provided in Section 5.

2. Model and assumptions

Let us consider the following n-dimensional regression model with m structural changes ($m+1$ regimes):

$$y_t = \Phi_j x_{jt} + \varepsilon_t \quad (j = 1, \ldots, m+1$$
$$\text{and} \quad t = T_{j-1}+1, \ldots, T_j) \tag{1}$$

where y_t and x_{jt} are $n \times 1$ and $p_{x_j} \times 1$ vectors of observations, respectively, ε_t is an error term and Φ_j is an $n \times p_{x_j}$ unknown coefficient matrix in the jth regime. Typically, the regressor x_{jt} includes a constant but trending regressors are not allowed in our model. We use the term $p_{\phi_j} = np_{x_j}$ to denote the number of unknown coefficients in each regime, so that the total number of coefficients is given by $p_\phi^{all} = \sum_{j=1}^{m+1} p_{\phi_j} = \sum_{j=1}^{m+1} np_{x_j}$. Similarly, by allowing structural changes in the variance matrix of ε_t, the number of unknown variance components in each regime is $p_\sigma = n(n+1)/2$ and that in all regimes is given by $p_\sigma^{all} = (m+1)p_\sigma = (m+1)n(n+1)/2$. We set $T_0 = 0$ and $T_{m+1} = T$, so that the total number of observations is T. In model (1) there are m structural changes ($m+1$ regimes) with change points given by $T_1, \ldots, T_m$. We allow the lagged dependent variables as regressors and in that case, the initial observations of y_t for $t \le 0$ are assumed to be given. Thus, model (1) includes a VAR model as a special case. Note that the different regressors and the different orders of the lagged dependent variables are allowed depending on the regimes. The main purpose of this paper is to derive the model selection criteria to choose the regressors among the $\bar{p}_x$ candidates for regressors and to estimate the number of structural changes m. In what follows, while we will continue using "choose the number p_x among the $\bar{p}_x$ regressors," we, however, imply "choose the regressors $x_{1t}, x_{2t}, \ldots, x_{m+1t}$ for all the regimes among the $\bar{p}_x$ candidates for regressors."

Model (1) can be rewritten as $y_t = (x_{jt}' \otimes I_n)\varphi_j + \varepsilon_t$ for the jth regime where $\varphi_j = \text{vec}(\Phi_j)$, or equivalently, $y_t = (x_t' \otimes I_n)\phi_j + \varepsilon_t$ where x_t consists of all the regressors, which are collected from the element of x_{jt}'s without overlap, while ϕ_j may contain 0's if the different regressors are allowed in each regime. We denote the true value of a parameter with superscript 0. For example, ϕ_j^0 and T_j^0 denote the true value of ϕ_j in the jth regime and the true jth break point, respectively. Hence, the data generating process is given by

$$y_t = (x_{jt}' \otimes I_n)\varphi_j^0 + \varepsilon_t = (x_t' \otimes I_n)\phi_j^0 + \varepsilon_t.$$

The following assumptions are made mainly for the derivation of the modified AIC and the modified C_p criterion.

Assumption A1. (a) There exists a positive integer $l_0 > 0$ such that for all $l > l_0$, the minimum eigenvalues of $(1/l)\sum_{t=T_{j-1}^0+1}^{T_{j-1}^0+l} x_t x_t'$ and $(1/l)\sum_{t=T_j^0-l}^{T_j^0} x_t x_t'$ are bounded away from zero ($j = 1, \ldots, m^0+1$). (b) $\sum_{t=k}^{l} x_t x_t'$ is invertible for $l - k > k_0$ for some $0 < k_0 < \infty$. (c) $\sup_t E\|x_t\|^{4+\delta} < \infty$ for some $\delta > 0$.

Assumption A2. When the lagged dependent variables are allowed as regressors, all the characteristic roots associated with the lag polynomials are inside the unit circle.

Assumption A3. (a) $\varepsilon_t = (\Sigma_j^0)^{1/2}\eta_t$ for $T_{j-1}^0 + 1 \le t \le T_j^0$ ($j = 1, \ldots, m^0+1$), where Σ_j^0 is a symmetric and positive definite unknown matrix and $\{\eta_t\}$ is a martingale difference sequence with respect to $\mathcal{F}_t = \sigma\{\eta_t, \eta_{t-1}, \ldots, x_{t+1}, x_t, \ldots,\}$ with $E[\eta_t\eta_t'|\mathcal{F}_{t-1}] = I_n$ for all t. (b) $\sup_t E\|\eta_t\|^{4+\delta} < \infty$ for some $\delta > 0$. (c) $E[\eta_{it}\eta_{jt}\eta_{kt}] = 0$ ($i, j, k = 1, \ldots, n$). (d) $(1/\Delta T_j^0)\text{tr}\left\{\left[\sum_{t=T_{j-1}^0+1}^{T_j^0}(\eta_t\eta_t' - I_n)\right]^2\right\} \xrightarrow{p} \kappa_{4j}$, where κ_4 is some positive number and $\Delta T_j^0 = T_j^0 - T_{j-1}^0$ ($j = 1, \ldots, m^0+1$).

Assumption A4. $\phi_{j+1}^0 - \phi_j^0 = \upsilon_T\delta_j$ and $\Sigma_{j+1}^0 - \Sigma_j^0 = \upsilon_T\Psi_j$, where $(\delta_j, \Psi_j) \neq 0$ ($j = 1, \ldots, m^0$), $\Sigma_j^0 \to \Sigma^0$ as $T \to \infty$ for all j and υ_T is a sequence of positive numbers such that $\upsilon_T \to 0$ and $\sqrt{T}\upsilon_T/(\log T)^2 \to \infty$.

Assumption A5. $0 = \lambda_0 < \lambda_1^0 < \cdots < \lambda_{m^0}^0 < \lambda_{m^0+1} = 1$, where $T_j^0 = [T\lambda_j^0]$ ($j = 0, \ldots, m^0+1$).

Assumption A6. The following weak law of large numbers and the functional central limit theorems hold ($j = 1, \ldots, m^0$):

$$\frac{1}{\Delta T_j^0}\sum_{t=T_{j-1}^0+1}^{T_j^0}\left(x_t x_t' \otimes (\Sigma_j^0)^{-1}\right) \xrightarrow{p} Q_{1j},$$