

SPECIAL ISSUE

STRUCTURAL BREAK INFERENCE USING INFORMATION CRITERIA
IN MODELS ESTIMATED BY TWO-STAGE LEAST SQUARES

ALASTAIR R. HALL^{a,*} DENISE R. OSBORN^{a,b} AND NIKOLAOS SAKKAS^c

^a *University of Manchester, Manchester, UK*

^b *University of Tasmania, Hobart, Tasmania, Australia*

^c *University of Bath, Bath, UK*

This paper makes two contributions in relation to the use of information criteria for inference on structural breaks when the coefficients of a linear model with endogenous regressors may experience multiple changes. First, we show that suitably defined information criteria yield consistent estimators of the number of breaks, when employed in the second stage of a two-stage least squares (2SLS) procedure with breaks in the reduced form taken into account in the first stage. Second, a Monte Carlo analysis investigates the finite sample performance of a range of criteria based on Bayesian information criterion (BIC), Hannan–Quinn information criterion (HQIC) and Akaike information criterion (AIC) for equations estimated by 2SLS. Versions of the consistent criteria BIC and HQIC perform well overall when the penalty term weights estimation of each break point more heavily than estimation of each coefficient, while AIC is inconsistent and badly over-estimates the number of true breaks.

Received 12 December 2013; Revised 23 June 2014; Accepted 10 October 2014

Keywords: structural breaks; information criteria; instrumental variables estimation
JEL. C13; C26.

1. INTRODUCTION

Information criteria (IC) are routinely employed in many contexts to select a model from a range of time-invariant linear specifications, such as selecting predictors (Pesaran and Timmermann, 1995) or specifying dynamics (Shibata, 1976; Ng and Perron, 2001). These methods are attractive to practitioners because they typically perform well, while the penalty functions on which they are based provide an intuitively attractive interpretation as a trade-off between goodness-of-fit and the dimension of the model, both defined appropriately. Hence, they are often preferred to the use of hypothesis tests for model specification.

Such criteria are, however, not widely used for estimation of the number of structural breaks in linear economic models, where, following the seminal studies of Andrews (1993) and Bai and Perron (1998), the predominant approach is based on the sequential application of hypothesis tests. One disadvantage of such a procedure is that the resulting estimator of the number of breaks is not consistent when the tests are performed at a fixed significance level. This is due to the probability of type one errors inherent in the decision rules for the tests, so that the estimator has a zero probability of under-fitting but a non-zero probability of over-fitting in the limit. On the other hand, appropriately defined IC yield consistent estimators for the number of breaks. Yao (1988) develops a version of the criterion of Schwarz (1978) [referred to as Bayesian information criterion (BIC)] for structural break inference.¹

* Correspondence to: Alastair R. Hall, Economics, School of Social Sciences, University of Manchester, Manchester M13 9PL, UK. E-mail: alastair.hall@manchester.ac.uk

¹ Also see Liu *et al.* (1997) and Zhang and Siegmund (2007).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Although the arguments of Bai (2000) indicate that this criterion, together with a wider range of penalty functions, will deliver a consistent estimator of the number of breaks,¹ the Monte Carlo study of Bai and Perron (2006) finds the Yao (1988) version of BIC to perform poorly in finite samples. The Akaike information criterion (AIC) of Ninomiya (2005), based on Akaike (1974), does not meet the Bai (2000) conditions and is not consistent for the number of breaks.

Nevertheless, consistent methods may not perform well in practice, and an important conclusion from the recent Monte Carlo analysis of Hall *et al.* (2013b) is that the penalty function of Yao (1988) may not adequately reflect the trade-off between the estimation of breaks and the residual sum of squares in finite samples. More specifically, based on analytical results obtained by Ninomiya (2005) for a simple mean shift model, Hall *et al.* (2013b) propose a modified penalty function that attaches a weight of three to break date estimation relative to estimation of an individual coefficient; this contrasts with the weight of one in the penalty used by Yao (1988). When modified by employing this higher weight, penalty functions based on BIC and the Hannan and Quinn (1979) criterion [Hannan and Quinn information criterion (HQIC)] perform well for estimating the number of breaks.

The studies referred to above relate to models with lagged dependent or exogenous regressors, but economic models often include endogenous regressors, rendering ordinary least square (OLS)-based techniques inappropriate. Although Hall *et al.* (2012) and Boldea *et al.* (2012) have very recently extended the OLS approach of Bai and Perron (1998) to develop a sequential hypothesis testing methodology for structural break inference in the two-stage least squares (2SLS) context, the IC approach remains unexplored. The purpose of this article is to study the usefulness of IC for structural break estimation in linear models estimated by 2SLS, both analytically and through a finite sample Monte Carlo analysis.

More explicitly, we establish generic conditions under which IC methods yield consistent estimation of the number of breaks in the 2SLS structural equation. These conditions cover penalty functions that behave as a function of the sample size like either BIC or HQIC. In line with other results relating to model specification, including Shibata (1976), methods based on AIC are not consistent and may asymptotically over-estimate the number of true breaks. Our extensive Monte Carlo analysis examines two versions of BIC, HQIC and AIC for 2SLS structural break inference, namely counting an estimated break as effectively equivalent to one individual coefficient and a 'structural break' version that employs a weight of three. In line with our OLS analysis in Hall *et al.* (2013b), we find that BIC and HQIC perform well when combined with the higher relative weight for break estimation, and this applies in cases with both i.i.d. and positively autocorrelated disturbances.

The outline of the paper is as follows. Section 2 discusses the assumptions on the structural equation of interest to the researcher, with the consistency of the IC approach in the 2SLS context established in Section 3. The results of our Monte Carlo study are detailed in Section 4, with conclusions drawn in Section 5. Throughout the paper, \xrightarrow{p} denotes limit in probability.

2. THE STRUCTURAL EQUATION

Consider the case in which the equation of interest is a structural relationship from a simultaneous system, with this equation exhibiting m breaks, such that

$$y_t = x_t' \beta_{x,i}^0 + z_{1,t}' \beta_{z_1,i}^0 + u_t, \quad i = 1, \dots, m+1, \quad t = T_{i-1}^0 + 1, \dots, T_i^0 \quad (1)$$

where $T_0^0 = 0$ and $T_{m+1}^0 = T$, and T is the total sample size. Thus, y_t is the dependent variable, while x_t is a $p_1 \times 1$ vector of endogenous explanatory variables, $z_{1,t}$ is a $p_2 \times 1$ vector of exogenous variables including the intercept and u_t is a mean zero error. We define $p = p_1 + p_2$. As usual in the literature, we require the break points to be asymptotically distinct.

¹ Bai's (2000) analysis is in the context of vector autoregressions.

Assumption 1. $T_i^0 = [T\lambda_i^0]$, where $0 < \lambda_1^0 < \dots < \lambda_m^0 < 1$.

Here $[\cdot]$ denotes the integer part of the quantity in the brackets.²

As a structural equation, we allow the explanatory variables, x_t , to be correlated with the errors u_t , and the first stage of 2SLS requires a reduced form (RF) representation of x_t to be estimated using appropriate instruments. Furthermore, we allow for this RF to be subject to discrete shifts in the sample period,

$$x'_t = z'_t \Delta_0^{(i)} + v'_t, \quad i = 1, 2, \dots, h + 1, \quad t = T_{i-1}^* + 1, \dots, T_i^* \tag{2}$$

where $T_0^* = 0$ and $T_{h+1}^* = T$. The vector $z_t = (z'_{1,t}, z'_{2,t})'$ is $q \times 1$ and contains variables that are uncorrelated with both u_t and v_t and are appropriate instruments for x_t in the first stage of the 2SLS estimation. The parameter matrices are $\Delta_0^{(i)} = (\delta_{1,0}^{(i)}, \delta_{2,0}^{(i)}, \dots, \delta_{p_1,0}^{(i)})$, each with dimension $q \times p_1$, and each $\delta_{j,0}^{(i)}$ is dimension $q \times 1$, for $j = 1, \dots, p_1$. The points $\{T_i^*\}$ are assumed to be generated analogously to the structural form (SF) breaks as follows.

Assumption 2. $T_i^* = [T\pi_i^0]$, where $0 < \pi_1^0 < \dots < \pi_h^0 < 1$.

Note that the break fractions in the RF, $\pi^0 = [\pi_1^0, \pi_2^0, \dots, \pi_h^0]'$, may or may not coincide with the breaks in the structural equation, $\lambda^0 = [\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0]'$. Also note that (2) can be re-written as follows:

$$x_t(\pi^0)' = \tilde{z}_t(\pi^0)' \Theta_0 + v'_t, \quad t = 1, 2, \dots, T \tag{3}$$

where $\Theta_0 = [\Delta_0^{(1)'}, \Delta_0^{(2)'}, \dots, \Delta_0^{(h+1)'}]'$. $\tilde{z}_t(\pi^0) = \iota(t, T) \otimes z_t$, $\iota(t, T)$ is a $(h + 1) \times 1$ vector with first element $\mathcal{I}\{t/T \in (0, \pi_1^0]\}$, $h + 1$ th element $\mathcal{I}\{t/T \in (\pi_h^0, 1]\}$, k th element $\mathcal{I}\{t/T \in (\pi_{k-1}^0, \pi_k^0]\}$ for $k = 1, 2, \dots, h$ and $\mathcal{I}\{\cdot\}$ is an indicator variable that takes the value one if the event in the curly brackets occurs.

Let $\hat{\pi} = [\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_h]'$ denote estimators of π^0 . It is assumed these estimators satisfy the following condition.

Assumption 3. $\hat{\pi} = \pi^0 + O_p(T^{-1})$.

This condition is sufficient to ensure that the RF coefficients are consistently estimated, which is required for consistent inference on the SF. Since this is a standard OLS-based problem, the assumption could be satisfied by consistently estimating the RF break dates equation by equation, applying either the Bai and Perron's (1998) methodology or an appropriate IC and then pooling the estimates of the break fractions. Let $\hat{x}_t(\hat{\pi})$ denote the resulting fitted values, that is,

$$\hat{x}_t(\hat{\pi})' = \tilde{z}_t(\hat{\pi})' \hat{\Theta}_T(\hat{\pi}) = \tilde{z}_t(\hat{\pi})' \left(\sum_{t=1}^T \tilde{z}_t(\hat{\pi}) \tilde{z}_t(\hat{\pi})' \right)^{-1} \sum_{t=1}^T \tilde{z}_t(\hat{\pi}) x'_t \tag{4}$$

where $\tilde{z}_t(\hat{\pi})$ is defined analogously to $\tilde{z}_t(\pi^0)$ based on the estimator of the true break points in the RF.

² This assumption fixes the break fraction λ_i and hence implies that the break dates T_i^0 change with the sample size T . Although fixed break dates may appear more realistic, by their nature, they do not lead to asymptotically large regimes. Assumption 1 is widely used in the theoretical structural break literature because it implies asymptotically large regimes leading to tractable analytical results on the limiting behaviour of the estimators that can be used to approximate finite sample behaviour. Dufour *et al.* (1994) and Andrews (2003) develop tests for structural instability in short regimes, such as at the end of the sample. For these methods, the critical value is calculated by bootstrap (Andrews, 2003) or bounded using moment inequalities (Dufour *et al.*, 1994).

To facilitate our analysis, we also impose the following assumptions:

Assumption 4. (i) $h_t = (u_t, v_t)' \otimes z_t$ is an array of real valued $(p+1)q \times 1$ random vectors defined on the probability space (Ω, \mathcal{F}, P) , $V_T = \text{Var} \left[\sum_{t=1}^T h_t \right]$ is such that $\text{diag} \left[\xi_{T,1}^{-1}, \dots, \xi_{T,(p+1)q}^{-1} \right] = \Xi_T^{-1}$ is $O(T^{-1})$, where Ξ_T is the $(p+1)q \times (p+1)q$ diagonal matrix with the eigenvalues $(\xi_{T,1}, \dots, \xi_{T,(p+1)q})$ of V_T along the diagonal; (ii) for each element $h_{t,i}$ of h_t , $E[h_{t,i}] = 0$ and, for some $d > 2$, $\|h_{t,i}\|_d < \Gamma < \infty$ for $t = 1, 2, \dots$, where $\|\cdot\|_d$ denotes the $L_d(P)$ norm; (iii) $\{h_{t,i}\}$ is near epoch dependent with respect to $\{g_t\}$ such that $\|h_t - E[h_t | \mathcal{G}_{t-\zeta}^{t+\zeta}]\|_2 \leq v_\zeta$ with $v_\zeta = O(\zeta^{-1/2})$, where $\mathcal{G}_{t-\zeta}^{t+\zeta}$ is a sigma-algebra based on $(g_{t-\zeta}, \dots, g_{t+\zeta})$; (iv) $\{g_t\}$ is either ϕ -mixing of size $\zeta^{-d/(2(d-1))}$ or α -mixing of size $\zeta^{-d/(d-2)}$; (v) $V_T(r) = \text{Var} \left[T^{-1/2} \sum_{t=1}^{[Tr]} h_t \right]$ satisfies $V_T(r) \rightarrow rV$ uniformly in $r \in [0, 1]$, where V is a pd matrix.

Assumption 5. $\text{Var}[u_t] = \sigma_u^2$, $\text{Cov}[u_t, v_t] = \Sigma_{uv}$ and $\text{Var}[v_t] = \Sigma_v$, for all t .

Assumption 6. $\text{Rank} \{ \Upsilon_i^0 \} = p$, where $\Upsilon_i^0 = \begin{bmatrix} \Delta_0^{(i)} & \Pi \end{bmatrix}$, for $i = 1, 2, \dots, h+1$, where $\Pi' = [I_{p_2}, 0_{p_2 \times (q-p_2)}]$, I_a denotes the $a \times a$ identity matrix and $0_{a \times b}$ is the $a \times b$ null matrix.

Assumption 7. For $\# = 0, *$, there exists an $l_\# > 0$ such that for all $l > l_\#$, the minimum eigenvalues of $A_{il} = (1/l) \sum_{t=T_i^\#}^{T_i^\#+l} z_t z_t'$ and of $\bar{A}_{il} = (1/l) \sum_{t=T_i^\#-l}^{T_i^\#} z_t z_t'$ are bounded away from zero for all $i = 1, \dots, v^\# + 1$, where $v^0 = m$ and $v^* = h$.

Assumption 8. $T^{-1} \sum_{t=1}^{[Tr]} z_t z_t' \xrightarrow{P} Q_{ZZ}(r)$ uniformly in $r \in [0, 1]$; $Q_{ZZ}(r)$ is positive definite for any $r > 0$ and strictly increasing in r , while $Q_{ZZ}(r) - Q_{ZZ}(s)$ is positive definite for any $r > s$.

Assumption 4 allows substantial dependence and heterogeneity in $(u_t, v_t)' \otimes z_t$ but at the same time imposes sufficient restrictions to deduce a functional central limit theorem for $T^{-1/2} \sum_{t=1}^{[Tr]} h_t$; see Wooldridge and White (1988). This assumption also contains the restrictions that the implicit population moment condition in 2SLS is valid – that is, $E[z_t u_t] = 0$ – and the conditional mean of the RF is correctly specified. Assumption 5 restricts the unconditional variance and covariances of the structural equation and RF errors to be constant over time. Assumption 6 implies the standard rank condition for identification in IV estimation in the linear regression model³ because Assumptions 4(ii), 6 and 8 together imply that

$$T^{-1} \sum_{t=[sT]+1}^{[Tr]} z_t [x_t', z_{1,t}'] \xrightarrow{P} [Q_{ZZ}(r) - Q_{ZZ}(s)] \Upsilon_0 = Q_{Z,[X,Z_1]}(r,s) \text{ uniformly in } r > s + \epsilon, r, s \in [0, 1] \quad (5)$$

where $Q_{Z,[X,Z_1]}(r,s)$ has rank equal to p for any r, s (satisfying the above conditions). Note that this assumption implies $q \geq p$. Assumption 7 requires that there be enough observations near the true break points in either the structural equation or RF so that they can be identified and is analogous to the extension proposed in Bai and Perron (1998) to their Assumption A2.

As already noted, Hall *et al.* (2012) (HHB hereafter) develop a hypothesis testing approach for inference on breaks in (1). They are the first authors to study this context analytically, and we retain almost all of their assumptions. However, the theory underlying certain tests employed in HHB's methodology requires the standardized partial sum instrument cross-product matrix to be linear in the sampling fraction within the assumed

³ See, e.g. Hall (2005, p. 35).

regimes under the appropriate null, that is, $T^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+rT} z_t z_t' \xrightarrow{p} rQ_i$ uniformly in $r \in (0, \lambda_i^0 - \lambda_{i-1}^0]$, where Q_i is a positive definite matrix of constants. This assumption is more restrictive than Assumption 8 and rules out changes in the mean and variance of the instruments at different times from the changes in the structural parameters.

3. CONSISTENCY OF AN INFORMATION CRITERION

Suppose now that a researcher knows neither the number nor the location of the breaks in the structural equation. Consider the case where an arbitrary number n breaks are estimated at $\tau(n) = [\tau_1, \tau_2, \dots, \tau_n]'$ with $0 < \tau_1 < \tau_2 < \dots < \tau_n < 1$, $\tau_0 = 0$, and $\tau_{n+1} = 1$. Then, the second stage of 2SLS can begin with the estimation of (1) via OLS for each possible n -partition of the sample, that is,

$$y_t = \hat{x}_t(\hat{\pi})' \beta_{x,i}^* + z'_{1,t} \beta_{z1,i}^* + \tilde{u}_t(\hat{\pi}), \quad i = 1, \dots, n + 1; \quad t = T_{i-1} + 1, \dots, T_i \tag{6}$$

where $T_i = [\tau_i T]$, and the regressors x_t are estimated using the fitted values of the first stage of 2SLS, $\hat{x}_t(\hat{\pi})$ as in (4). We further assume the following.

Assumption 9. Equation (6) is estimated over all partitions (T_1, \dots, T_n) such that $T_i - T_{i-1} > \max\{q - 1, \epsilon T\}$ for some $\epsilon > 0$ and $\epsilon < \inf_i (\lambda_{i+1}^0 - \lambda_i^0)$ and $\epsilon < \inf_j (\pi_{j+1}^0 - \pi_j^0)$.

Assumption 9 requires that each segment considered in the estimation contains a positive fraction of the sample asymptotically; in practice, ϵ is chosen to be small in the hope that the last part of the assumption is valid. Letting $\beta_i^{*'} = (\beta_{x,i}^{*'}, \beta_{z1,i}^{*'})'$, for a given $(n + 1)$ -partition, the estimates of $\beta^* = (\beta_1^{*'}, \beta_2^{*'}, \dots, \beta_{n+1}^{*'})'$ are obtained by minimizing the sum of squared residuals

$$S_T(T_1, \dots, T_n; \beta) = \sum_{i=1}^{n+1} \sum_{t=T_{i-1}+1}^{T_i} \{y_t - \hat{x}_t(\hat{\pi})' \beta_{x,i} - z'_{1,t} \beta_{z1,i}\}^2$$

with respect to $\beta = (\beta_1', \beta_2', \dots, \beta_{n+1}')'$. We denote these estimators by $\hat{\beta}(\tau(n))$. The estimators of the break points, $(\hat{T}_1, \dots, \hat{T}_n)$, are then defined as

$$\hat{\tau}(n) = (\hat{T}_1, \dots, \hat{T}_n) = \arg \min_{T_1, \dots, T_n} S_T(T_1, \dots, T_n; \hat{\beta}(\tau(n))) \tag{7}$$

where the minimization is taken over all possible partitions implied by (T_1, \dots, T_n) . The 2SLS estimates of the regression parameters, $\hat{\beta}(\tau(n)) = (\hat{\beta}'_1, \hat{\beta}'_2, \dots, \hat{\beta}'_{n+1})'$, are the regression parameter estimates associated with each of the estimated partitions.

The estimators $\hat{\tau}(n)$ and $\hat{\beta}(\tau(n))$ are calculated conditional on n . While the above considers arbitrary n , we seek an estimator for the true number of structural breaks m , which is typically unknown *a priori*. HHB propose a method for estimation of m based on the sequential application of certain test statistics for parameter variation. Although this methodology provides researchers with techniques that are (asymptotically) valid for 2SLS, nevertheless, it has the practical disadvantage that it involves dividing the sample into sub-samples over which the RF is judged stable. With the moderate sample sizes often available to practitioners, this sample splitting can lead to segments that have relatively few observations over which testing can be conducted for the SF equation. Further, as pointed out in Section 1, such a sequential procedure may not be consistent for the number of breaks.

In the light of the widespread use of IC in preference to hypothesis testing in other model specification contexts, we consider the application of IC for inference on the number of breaks in the structural equation (1) estimated by 2SLS. This involves minimizing

$$\text{IC}(\tau(n); n, \hat{\pi}) = \ln [\hat{\sigma}^2(\tau(n); n, \hat{\pi})] + K(n, T) \quad (8)$$

where

$$\hat{\sigma}^2(\tau(n); n, \hat{\pi}) = (T - p)^{-1} \text{RSS}(\tau(n); n, \hat{\pi}) \quad (9)$$

$$\text{RSS}(\tau(n); n, \hat{\pi}) = \sum_{j=1}^n \text{RSS}_j(\tau(n); n, \hat{\pi}) \quad (10)$$

$$\text{RSS}_j(\tau(n); n, \hat{\pi}) = \sum_{t=[\tau_{j-1}T]+1}^{[\tau_j T]} \left\{ y_t - \hat{x}_t(\hat{\pi})' \hat{\beta}_{x,i} - z'_{1,t} \hat{\beta}_{z_1,i} \right\}^2 \quad (11)$$

and $K(n, T)$ is a deterministic penalty term governed by the following assumption.

Assumption 10. $K(n, T) = o(1)$ as $T \rightarrow \infty$, it is a strictly increasing function of n , and $TK(n, T) \rightarrow \infty$ as $T \rightarrow \infty$.

The estimated number of breaks, denoted \hat{n} , is the value that minimizes the IC, that is,

$$\hat{n} = \underset{n \in \mathcal{N}}{\text{argmin}} \text{IC}(\tau(n); n, \hat{\pi}) \quad (12)$$

where $\mathcal{N} = \{0, 1, \dots, N\}$. The associated estimators of the break locations are $\hat{\tau}(\hat{n})$. N is the maximum number of breaks considered, and we assume this is large enough to ensure $m \in \mathcal{N}$.

Assumption 11. $N \geq m$.

The proof of consistency of our method (see the Appendix) rests on the limiting properties of $\text{RSS}(\hat{\tau}(\hat{n}); \hat{n}, \hat{\pi})$. For any partition of the sample with no neglected breaks, $T^{-1} \text{RSS}(\cdot)$ converges to $\Gamma = \sum_{i=1}^n \Gamma_i$, where

$$\Gamma_i = \sigma_u^2 + 2 \Sigma_{uv} \beta_{x,i}^0 + \beta_{x,i}^{0'} \Sigma_v \beta_{x,i}^0 \quad (13)$$

is the variance of the composite disturbance $u_t + v_t' \beta_{x,i}^0$ that applies for segment i in the second stage regression (1) when x_t' is replaced by the true RF model (2). On the other hand, for any partition with at least one neglected break, $T^{-1} \text{RSS}(\cdot)$ converges to the larger value $\Gamma + \xi$, $\xi > 0$. The consequent behaviour of (8), combined with Assumptions 10 and 11, implies the consistency of \hat{n} for m , with HHB (Theorem 1) then yielding the consistency of $\hat{\tau}(\hat{n})$.

Our result is stated formally in the following theorem.

Theorem 1. Under Assumptions 1–11,

$$[\hat{n}, \hat{\tau}(\hat{n})] \xrightarrow{P} [m, \lambda^0]$$

where $\lambda^0 = [\lambda_1^0, \dots, \lambda_m^0]'$ is the collection of the true break fractions in (1).

To implement the estimation procedure, it is necessary to pick a penalty term that satisfies Assumption 10. A natural choice that leads to a consistent IC is

$$K(n, T) = [(n + 1)p + kn] \ln(T)/T \quad (14)$$

which is associated with BIC, because this choice has been found to work well in other settings. In effect, the dimension of the model (6) is equal to the number of coefficients estimated, namely $(n + 1)p$, plus the number of break dates estimated, with the latter having a relative weight of k . Applied in this 2SLS context, the proposal of Yao (1988) sets $k = 1$, and this penalty gives the criterion we refer to simply as BIC. However, Hall *et al.* (2013a) show that although the analytical results of Ninomiya (2005) are obtained in the context of a simple mean shift model, the weight of $k = 3$ represents the asymptotic impact of break date estimation on the residual sum of squares in a more general linear model. This provides the justification for the higher weight proposed in Hall *et al.* (2013b) for models estimated by OLS. We apply this also in the 2SLS case and refer to the criterion with $k = 3$ as SBBIC, indicating structural break SIC. Following Hall *et al.* (2013b), we also employ two versions of HQIC, with

$$K(n, T) = 2[(n + 1)p + kn] \ln[\ln(T)]/T \quad (15)$$

for $k = 1$ (referred to as HQIC) and $k = 3$ (SBHQIC). These criteria using the penalty (15) also satisfy Assumption 10.

However, the choice associated with AIC (Akaike (1974)), where

$$K(n, T) = 2[(n + 1)p + kn]/T \quad (16)$$

does not satisfy Assumption 10 and yields an estimator that has a zero probability of choosing too few breaks but a non-zero probability of choosing too many breaks in the limit. Nevertheless, we include the penalty (16) in the Monte Carlo study of the next section order to examine the finite sample implications of this lack of consistency. Once again, we use $k = 1$ (labelled as AIC in the results) and $k = 3$ (SBAIC).

4. SIMULATION EVIDENCE

The first section outlines the set-up employed for our Monte Carlo analysis, with results discussed in the second section.

Table I. Simulation cases

Case	h	m	π_1	π_2	λ_1	λ_2
1	0	0	—	—	—	—
2	0	1	—	—	0.5	—
3	0	2	—	—	0.3	0.6
4	1	0	0.5	—	—	—
5	1	1	0.5	—	0.5	—
6	1	2	0.5	—	0.3	0.6
7	1	1	0.3	—	0.6	—
8	2	1	0.3	0.6	0.5	—
9	2	2	0.3	0.6	0.2	0.4

Note: The data generating processes are described in Section 4.

h , number of breaks in the reduced form; m , number of breaks in the structural form; π_1, π_2 , locations of reduced form breaks (as fractions of sample size); λ_1, λ_2 , locations of structural form breaks.

4.1. Methodology

We assess the performance of the aforementioned IC in a variety of cases with different numbers and locations of breaks in both the SF and RF equations of (1) and (2). These nine cases are given in Table I. We specify models with no breaks in one equation and zero to two breaks in the other, together with models involving breaks in both the RF and the SF, including a contrast between a coincidental break in both equations (case 5) and non-coincidental breaks (case 7). Since these nine cases cover realistic scenarios that might be encountered in applied settings, we anticipate that the results will be informative for practitioners even when more than two breaks may be present.

For each case, we investigate the effect of sample size ($T = 120$ and 240), break magnitude, autocorrelation and the effect of explanatory power in first stage (RF) on the second stage (SF) breaks estimation. We examine the realistic scenario where the number and locations of breaks are unknown in both equations, and the same IC is applied for structural break inference in each of these. Tables II–IX present the empirical probability (as a percentage) of each IC selecting 0, 1, 2 or ≥ 3 breaks in the RF and SF for each case, based on 10 000 replications of the data generating processes (DGPs) discussed below. The tables present RF results once for all cases with the same RF, since by keeping the same seed for the pseudo-random number generator, these estimations give identical results. The cases with a common RF are separated using horizontal lines in the tables.

The SF equation includes a constant and one endogenous variable, so that (1) becomes

$$y_t = \beta_{1,i} + \beta_{2,i}x_t + u_t \quad i = 1, \dots, m + 1.$$

All cases of no breaks use $\beta_1 = 0.5$, with $\beta_2 = 0.1$ in Tables II–V and $\beta_2 = 1$ in Tables VI–IX. When breaks exist in the SF, we use the same coefficient values as in the no breaks case but alternate the signs of both coefficients between segments. These coefficient values were chosen so as to present meaningful and comparable results, where the IC neither pick the true number of breaks 100% of the time nor have effectively zero power. Thus, for example, for two breaks, we set

$$y_t = \begin{cases} \beta_1 + \beta_2 x_t + u_t & \text{if } t \leq [\lambda_1 T] \\ -\beta_1 - \beta_2 x_t + u_t & \text{if } [\lambda_1 T] < t \leq [\lambda_2 T] \\ \beta_1 + \beta_2 x_t + u_t & \text{if } t > [\lambda_2 T] \end{cases}$$

The simulated RF equations based on (2) are

$$x_t = \delta_{1,j} + \delta_j^* \sum_{a=2}^q z_{a,t} + v_t \quad j = 1, \dots, h + 1$$

so that the $(q - 1)$ instruments, other than the intercept, take the common coefficient value δ_j^* in RF segment j . The intercept is $\delta_{1,j} = \pm 0.5$ as in the SF and $q = 5$. With the variances of $z_{2,t}, \dots, z_{5,t}$ and v_t set to unity, δ_j^* is determined to yield theoretical $R^2 = 0.3$ or $R^2 = 0.5$ by using $\delta_j^* = \sqrt{R^2 / (q - 1)(1 - R^2)}$, as in Hahn and Inoue (2002). Performance of the IC under these different levels of explanatory power is presented as separate rows of results. Across segments, the signs of the RF coefficients alternate as for the SF equations.

We use two different dynamic structures, each presented in different tables, to generate the u_t , v_t and the instruments z_t . In the case of i.i.d. errors, we draw from a multi-variate (six-dimensional) standard normal distribution, with $\text{Cov}[u_t, v_t] = 0.5$ and uncorrelated with the instruments, while the instruments have $\text{Cov}[z_{it}, z_{jt}] = 0 \quad \forall i \neq j$. To explore the effect of autocorrelation in the behaviour of the IC, we also simulate each case with AR(1) processes for both the SF errors $u_t = \phi_u u_{t-1} + \varepsilon_t$, and the instruments $z_{a,t} = \phi_z z_{a,t-1} + \varepsilon_{a,t}$ ($a = 2, \dots, 5$). We set the autoregressive parameter to 0.5 for both, and to ensure that $\text{Var}[u_t] = 1$, we set $\text{Var}[\varepsilon_t] = (1 - \phi_u^2) \text{Var}[u_t] = 0.75$, while to retain $\text{Cov}[u_t, v_t] = \text{Cov}[(1 - \phi_u L)^{-1} \varepsilon_t, v_t] = 0.5$, we set $\text{Cov}[\varepsilon_t, v_t] = 0.5$. Similar considerations for the AR(1) in the instruments imply setting $\text{Var}[\varepsilon_{it}] = (1 - \phi_z^2) = 0.75$. Finally, when searching for the break locations, we allow for a maximum of five breaks, set the trimming parameter (ϵ in Assumption 9) to 0.10,

that is, the minimum length of a segment can be 10% of the sample size and use the efficient search algorithm developed in Bai and Perron (2003).

The presentation of the results is as follows. Tables II and III give the results for the small sample size ($T = 120$) and ‘small’ breaks ($\beta_2 = 0.1$) for the two different dynamic structures, i.i.d. and AR(1) respectively. Tables IV and V change to the larger sample size ($T = 240$), and Tables VI–IX repeat the models of Tables II–V but for the larger magnitude of breaks given by $\beta_2 = 1$. To aid interpretation within these tables, the highest empirical probability of detecting the true number of breaks is shown in bold for each case considered.

4.2. Results

When no breaks occur in the DGP for either the SF or RF equations ($h = 0, m = 0$), case 1 of Tables II, IV, VI and VIII show a good performance of BIC when the disturbances are i.i.d. More explicitly, the BIC criterion, which employs $k = 1$ in (14), performs very well in not detecting spurious breaks in the RF, with good results consequently also seen in the SF when there is no autocorrelation. Even with the smaller sample size of $T = 120$, spurious breaks are infrequently detected by BIC (Tables II and VI). However, any such spurious breaks are almost always removed by the use of the criterion SBBIC, which applies a higher weight of $k = 3$ to break date estimation. The use of HQIC leads to the estimation of some spurious breaks in both the RF and SF equations, with this feature being more marked for the SF. An increase in the sample size from $T = 120$ to $T = 240$ reduces spurious break detection in the SF from around 15% to about 11% (compare Table II with Table IV, and Table VI with Table VIII). Use of the modified criterion SBHQIC, however, eliminates the vast majority of these, resulting in 1% or fewer spurious SF breaks for case 1 across Tables II, IV, VI and VIII.

Compared with the performances of these criteria, the use of the inconsistent AIC yields poor inference on the number of breaks when none apply in the DGP. This is particularly marked when the penalty term (16) employs $k = 1$, which effectively counts each break date estimated as equivalent to a single coefficient and leads to three or more spurious breaks being detected in the clear majority of replications with i.i.d. disturbances. While the number of spurious breaks is reduced by the use of SBAIC, these nevertheless occur in a substantial percentage of replications, standing at around 18% in the most favourable scenario of Table VIII.

Turning to the DGPs with autocorrelation, notice first that autocorrelation in the regressors with $h = 0$ in Tables III, V, VII and IX leads to very similar break detection results for the RF compared with when the regressors are i.i.d. However, when breaks occur ($h = 1$ or 2) and the sample size is relatively small, autocorrelation reduces the accuracy of RF break detection by the BIC-based and HQIC-based methods in Tables III and VII compared with Tables II and VI respectively. The AIC-based methods are always poor, and autocorrelated regressors have little effect on their RF performance.

Although the BIC-based and HQIC-based criteria remain consistent in the presence of stationary autocorrelation, it is clear that the positively autocorrelated AR(1) disturbances lead to a deterioration of performance for all criteria applied to the SF when this experiences no breaks.⁴ However, allocating the heavier weight to break date estimation in SBBIC and SBHQIC alleviates this feature. For example, for case 1 in Table III, BIC yields spurious breaks in more than 40% of the replications, which is reduced to less than 7% by SBBIC, with the corresponding percentages for HQIC and SBHQIC being 83% and 33% respectively, with these performances improving marginally with $T = 240$ in Table V. Not surprisingly, a stronger role for the SF regressors ($\beta_2 = 1$) also leads to improved performances for these consistent criteria in Tables VII and IX, with the marked improvements shown by BIC, HQIC and SBHQIC particularly noteworthy. AIC and SBAIC also show an increased tendency to detect spurious breaks with AR(1) rather than i.i.d. disturbances, but they remain poor in comparison with the consistent criteria. Indeed, this is always the case irrespective of the number of true breaks, and hence, we do not explicitly discuss these criteria further.

⁴ Since there is generally only modest deterioration in the detection of RF breaks with autocorrelated regressors, the deterioration in performance in the SF can be attributed primarily to autocorrelation in the second stage model itself.

Table II. Empirical distributions of estimated numbers of breaks, i.i.d. errors, $T = 120, \beta_2 = 0.1$

Case	R^2	BIC			SBBIC			HQ			SBHQIC			AIC			SBAIC									
		0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3					
1	RF	0.3	99.81	0.19	0.00	0.00	100.00	0.00	0.00	93.54	5.24	1.00	0.22	99.39	0.58	0.03	0.00	0.01	0.04	0.13	99.82	0.94	1.02	2.99	95.05	
	($h = 0$)	0.5	99.81	0.19	0.00	0.00	100.00	0.00	0.00	93.54	5.24	1.00	0.22	99.39	0.58	0.03	0.00	0.01	0.04	0.13	99.82	0.94	1.02	2.99	95.05	
	SF	0.3	98.20	1.68	0.10	0.02	99.98	0.02	0.00	84.71	10.15	4.04	1.10	98.97	1.02	0.01	0.00	1.03	1.68	6.55	90.74	30.73	12.94	20.38	35.95	
	($h = 0, m = 0$)	0.5	98.13	1.76	0.09	0.02	99.98	0.02	0.00	84.64	10.15	4.09	1.12	99.01	0.99	0.00	0.00	0.97	1.43	6.42	91.18	28.46	12.95	21.03	37.56	
2	SF	0.3	0.72	96.82	2.31	0.15	10.51	89.45	0.04	0.00	0.03	79.99	15.59	4.39	1.17	97.28	1.49	0.06	0.00	1.07	4.24	94.69	0.01	22.02	23.91	54.06
	($h = 0, m = 1$)	0.5	0.58	96.94	2.33	0.15	9.79	90.17	0.04	0.00	0.05	79.61	15.84	4.50	1.04	97.42	1.48	0.06	0.00	1.11	4.09	94.80	0.00	20.91	24.33	54.76
3	SF	0.3	25.81	4.69	67.39	2.11	89.55	1.72	8.72	0.01	2.46	1.78	77.12	18.64	36.06	4.96	57.81	1.17	0.00	0.00	2.45	97.55	0.00	0.01	26.08	73.91
	($h = 0, m = 2$)	0.5	24.69	4.61	68.50	2.20	88.81	1.85	9.33	0.01	2.28	1.61	77.22	18.89	34.66	4.91	59.18	1.25	0.00	0.00	2.15	97.85	0.00	0.01	25.40	74.59
4	RF	0.3	0.07	99.56	0.37	0.00	0.96	99.04	0.00	0.00	88.35	9.59	2.06	0.00	98.73	1.22	0.05	0.00	0.02	0.12	99.86	0.00	0.82	2.21	96.97	
	($h = 1$)	0.5	0.00	99.65	0.35	0.00	0.00	99.99	0.01	0.00	88.79	9.09	2.12	0.00	98.74	1.20	0.06	0.00	0.03	0.10	99.87	0.00	0.99	2.36	96.65	
	SF	0.3	98.56	1.29	0.15	0.00	100.00	0.00	0.00	84.34	10.78	3.88	1.00	99.20	0.75	0.05	0.00	0.74	1.63	6.25	91.38	21.90	16.82	22.72	38.56	
	($h = 1, m = 0$)	0.5	98.56	1.34	0.10	0.00	100.00	0.00	0.00	85.33	9.63	4.02	1.02	99.22	0.74	0.04	0.00	0.84	1.50	5.90	91.76	24.73	14.10	21.97	39.20	
5	SF	0.3	19.54	78.19	2.16	0.11	69.92	30.06	0.02	0.00	3.81	76.92	15.20	4.07	25.88	72.77	1.29	0.06	0.00	1.08	4.13	94.79	0.63	21.60	24.96	52.81
	($h = 1, m = 1$)	0.5	6.85	90.75	2.28	0.12	43.31	56.64	0.05	0.00	0.78	79.29	15.53	4.40	10.11	88.46	1.38	0.05	0.00	0.90	3.89	95.21	0.12	20.33	24.33	55.22
6	SF	0.3	34.51	11.06	52.57	1.86	91.79	3.46	4.75	0.00	4.22	5.21	72.90	17.67	45.76	10.62	42.79	0.83	0.00	0.00	2.23	97.77	0.00	0.06	25.40	74.54
	($h = 1, m = 2$)	0.5	34.98	7.72	55.55	1.75	92.28	2.39	5.33	0.00	4.30	3.68	73.88	18.14	46.32	7.57	45.27	0.84	0.00	0.00	2.04	97.96	0.00	0.10	24.68	75.22
7	RF	0.3	0.70	98.92	0.38	0.00	4.79	95.21	0.00	0.00	0.05	89.43	8.59	1.93	0.26	98.39	1.31	0.04	0.00	0.01	0.06	99.93	0.00	0.96	2.33	96.71
	($h = 1$)	0.5	0.00	99.67	0.33	0.00	0.03	99.97	0.00	0.00	0.00	89.65	8.51	1.84	0.00	98.72	1.25	0.03	0.00	0.01	0.07	99.92	0.00	1.12	2.71	96.17
	SF	0.3	11.16	86.01	2.71	0.12	52.54	47.44	0.02	0.00	1.65	76.63	17.06	4.66	15.49	83.05	1.43	0.03	0.00	0.82	3.50	95.68	0.08	15.95	25.85	58.12
	($h = 1, m = 1$)	0.5	6.88	90.58	2.43	0.11	42.42	57.56	0.02	0.00	0.89	78.91	15.68	4.52	9.91	88.66	1.42	0.01	0.00	0.99	3.62	95.39	0.01	17.82	24.88	57.29
8	RF	0.3	23.08	1.08	75.37	0.47	71.76	0.65	27.59	0.00	0.77	0.22	83.96	15.05	9.62	0.74	87.60	2.04	0.00	0.00	0.02	99.98	0.00	0.00	1.05	98.95
	($h = 2$)	0.5	0.44	0.01	98.97	0.58	9.74	0.12	90.14	0.00	0.00	0.00	85.62	14.38	0.07	0.00	97.89	2.04	0.00	0.00	0.02	99.98	0.00	0.00	1.29	98.71
	SF	0.3	1.01	96.44	2.44	0.11	15.51	84.47	0.02	0.00	0.06	79.20	16.66	4.08	1.50	96.93	1.54	0.03	0.00	0.74	3.56	95.70	0.00	17.05	24.55	58.40
	($h = 2, m = 1$)	0.5	0.92	96.34	2.64	0.10	14.05	85.92	0.03	0.00	0.04	79.89	16.05	4.02	1.52	96.98	1.47	0.03	0.00	0.95	3.78	95.27	0.00	18.53	24.00	57.47
9	SF	0.3	65.05	10.39	23.73	0.83	98.07	1.21	0.71	0.01	18.23	11.13	57.20	13.44	75.03	8.65	15.96	0.36	0.01	0.02	1.98	97.99	0.42	0.96	23.42	75.20
	($h = 2, m = 2$)	0.5	63.00	9.19	26.98	0.83	98.19	1.04	0.76	0.01	16.49	9.35	61.21	12.95	73.00	7.69	18.93	0.38	0.01	0.02	2.37	97.60	0.36	0.76	25.83	73.05

Note: Rows report the percentage of estimated breaks for each information criterion out of 10 000 replications and across the nine cases defined in Table I. The reduced form breaks estimation is denoted RF and is given once before every structural form case (SF) that has the same reduced form. R^2 refers to the theoretical R^2 in the reduced form DGP. The IC that yields the highest frequency for the true number of SF breaks is marked in bold.

Table III. Empirical distributions of estimated numbers of breaks, AR(1) errors, $T = 120, \beta_2 = 0.1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC									
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3							
1	RF	0.3	99.84	0.16	0.00	0.00	99.99	0.01	0.00	0.00	93.30	5.23	1.05	0.42	99.37	0.62	0.01	0.00	0.01	0.03	0.06	99.90	0.88	1.06	2.75	95.31
	($h = 0$)	0.5	99.84	0.16	0.00	0.00	99.99	0.01	0.00	0.00	93.30	5.23	1.05	0.42	99.37	0.62	0.01	0.00	0.01	0.03	0.06	99.90	0.88	1.06	2.75	95.31
	SF	0.3	58.33	19.07	14.17	8.43	93.32	5.65	0.98	0.05	16.53	14.71	22.93	45.83	67.06	17.24	10.93	4.77	0.01	0.01	0.11	99.87	0.82	0.97	4.34	93.87
	($h = 0, m = 0$)	0.5	58.26	19.19	14.12	8.43	93.31	5.68	0.95	0.06	16.43	14.56	22.91	46.10	67.00	17.30	10.93	4.77	0.00	0.01	0.13	99.86	0.54	0.85	4.16	94.45
2	SF	0.3	2.08	56.98	25.32	15.62	13.27	79.93	6.24	0.56	0.25	18.38	23.58	57.79	3.13	64.00	22.56	10.31	0.00	0.01	0.05	99.94	0.00	0.54	3.10	96.36
	($h = 0, m = 1$)	0.5	1.95	56.93	25.29	15.83	12.71	80.40	6.29	0.60	0.20	18.33	23.82	57.65	2.93	64.14	22.58	10.35	0.00	0.01	0.07	99.92	0.00	0.52	2.60	96.88
3	SF	0.3	9.33	8.10	54.02	28.55	52.33	11.71	33.55	2.41	0.94	2.02	26.63	70.41	13.59	9.52	55.21	21.68	0.00	0.00	0.04	99.96	0.02	0.02	1.84	98.12
	($h = 0, m = 2$)	0.5	8.97	7.94	54.25	28.84	51.83	11.67	34.16	2.34	0.89	1.93	26.35	70.83	13.12	9.43	55.61	21.84	0.00	0.00	0.04	99.96	0.01	0.02	1.68	98.29
4	RF	0.3	0.31	99.24	0.45	0.00	2.47	97.53	0.00	0.00	0.01	87.77	9.85	2.37	0.08	98.59	1.26	0.07	0.00	0.01	0.01	99.98	0.00	0.69	1.97	97.34
	($h = 1$)	0.5	0.00	99.57	0.43	0.00	0.02	99.98	0.00	0.00	0.00	88.45	9.30	2.25	0.00	98.63	1.30	0.07	0.00	0.01	0.01	99.98	0.00	0.74	1.99	97.27
	SF	0.3	59.91	19.21	13.26	7.62	94.26	4.95	0.74	0.05	16.83	15.43	22.95	44.79	68.29	17.38	9.87	4.46	0.00	0.00	0.10	99.90	0.56	1.03	4.37	94.04
	($h = 1, m = 0$)	0.5	59.91	18.59	13.66	7.84	94.18	4.98	0.80	0.04	16.93	14.72	22.90	45.45	68.43	16.74	10.28	4.55	0.00	0.01	0.08	99.91	0.50	0.92	3.71	94.87
5	SF	0.3	11.06	51.39	23.59	13.96	46.65	48.61	4.43	0.31	1.44	18.87	23.33	56.36	15.29	55.14	20.68	8.89	0.00	0.00	0.04	99.96	0.03	0.62	2.94	96.41
	($h = 1, m = 1$)	0.5	5.98	54.59	24.40	15.03	31.24	63.08	5.26	0.42	0.76	18.90	23.23	57.11	8.40	60.35	21.62	9.63	0.00	0.01	0.03	99.96	0.02	0.39	2.59	97.00
6	SF	0.3	11.22	11.78	49.92	27.08	56.21	15.13	26.49	2.17	1.10	3.45	26.86	68.59	16.02	13.84	50.15	19.99	0.00	0.00	0.04	99.96	0.01	0.01	1.88	98.10
	($h = 1, m = 2$)	0.5	11.84	10.23	50.60	27.33	56.98	13.41	27.47	2.14	1.07	3.04	26.83	69.06	16.69	12.02	50.99	20.30	0.00	0.00	0.04	99.96	0.00	0.04	1.86	98.10
7	RF	0.3	2.08	97.52	0.39	0.01	9.74	90.23	0.03	0.00	0.10	88.24	9.30	2.36	0.93	97.71	1.29	0.07	0.00	0.01	0.06	99.93	0.00	0.71	1.98	97.31
	($h = 1$)	0.5	0.02	99.65	0.32	0.01	0.23	99.75	0.02	0.00	0.00	88.73	9.16	2.11	0.01	98.54	1.37	0.08	0.00	0.02	0.09	99.89	0.00	0.79	2.26	96.95
	SF	0.3	8.55	50.61	26.31	14.53	35.92	58.21	5.43	0.44	1.25	17.75	24.89	56.11	11.78	55.83	23.16	9.23	0.00	0.00	0.05	99.95	0.03	0.47	3.09	96.41
	($h = 1, m = 1$)	0.5	6.66	53.35	25.60	14.39	30.54	63.57	5.44	0.45	0.91	18.42	24.06	56.61	8.99	59.24	22.44	9.33	0.00	0.00	0.07	99.93	0.01	0.57	2.44	96.98
8	RF	0.3	29.28	2.39	67.86	0.47	73.17	1.86	24.96	0.01	2.09	0.65	81.24	16.02	14.73	1.93	81.16	2.18	0.00	0.00	0.02	99.98	0.00	0.00	0.75	99.25
	($h = 2$)	0.5	1.85	0.14	97.48	0.53	16.59	0.47	82.93	0.01	0.03	0.00	84.04	15.93	0.56	0.14	96.95	2.35	0.00	0.00	0.04	99.96	0.00	0.00	1.09	98.91
	SF	0.3	2.35	57.34	26.23	14.08	15.28	78.83	5.44	0.45	0.17	18.07	24.64	57.12	3.35	63.98	23.74	8.93	0.00	0.00	0.06	99.94	0.00	0.57	2.67	96.76
	($h = 2, m = 1$)	0.5	2.28	56.66	26.26	14.80	14.47	79.07	6.05	0.41	0.24	18.62	24.29	56.85	3.39	64.31	22.84	9.46	0.00	0.00	0.07	99.93	0.00	0.52	2.90	96.58
9	SF	0.3	24.15	16.12	36.82	22.91	72.63	13.54	12.72	1.11	3.66	5.83	24.49	66.02	31.96	17.33	34.88	15.83	0.00	0.00	0.04	99.96	0.02	0.09	2.19	97.70
	($h = 2, m = 2$)	0.5	23.17	15.56	38.26	23.01	73.85	12.42	12.57	1.16	3.39	5.45	25.36	65.80	30.61	16.48	36.58	16.33	0.00	0.00	0.04	99.96	0.02	0.10	2.20	97.68

Note: Same as Table II.

Table IV. Empirical distributions of estimated numbers of breaks, i.i.d. errors, $T = 240$, $\beta_2 = 0.1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC								
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3						
1	RF	0.3	100.00	0.00	0.00	0.00	100.00	0.00	0.00	97.29	2.43	0.26	0.02	99.84	0.16	0.00	0.00	0.02	0.06	0.49	99.43	2.47	2.64	7.53	87.36
	($h = 0$)	0.5	100.00	0.00	0.00	100.00	0.00	0.00	97.29	2.43	0.26	0.02	99.84	0.16	0.00	0.00	0.00	0.02	0.06	0.49	99.43	2.47	2.64	7.53	87.36
	SF	0.3	99.25	0.72	0.03	0.00	100.00	0.00	0.00	88.69	8.14	2.73	0.44	99.42	0.55	0.02	0.00	0.94	1.67	6.44	90.95	34.48	11.42	20.11	33.99
	($h = 0, m = 0$)	0.5	99.23	0.73	0.04	0.00	100.00	0.00	0.00	88.66	8.15	2.77	0.42	99.42	0.56	0.02	0.00	0.94	1.50	6.04	91.52	31.28	11.96	20.51	36.25
2	SF	0.3	0.00	98.70	1.28	0.02	0.02	99.98	0.00	0.00	85.10	12.45	2.45	0.00	98.98	1.00	0.02	0.00	1.16	4.33	94.51	0.00	21.31	24.93	53.76
	($h = 0, m = 1$)	0.5	0.00	98.76	1.21	0.03	0.02	99.98	0.00	0.00	84.98	12.53	2.49	0.00	99.03	0.95	0.02	0.00	1.05	3.99	94.96	0.00	20.24	24.89	54.87
3	SF	0.3	0.79	0.15	98.07	0.99	33.17	1.02	65.81	0.00	0.00	0.01	85.66	14.33	1.13	0.22	97.85	0.80	0.00	2.66	97.34	0.00	0.00	26.46	73.54
	($h = 0, m = 2$)	0.5	0.71	0.15	98.12	1.02	30.83	0.96	68.21	0.00	0.00	0.01	85.45	14.54	0.91	0.21	98.02	0.86	0.00	2.53	97.47	0.00	0.00	25.41	74.59
4	RF	0.3	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.76	3.94	0.30	0.00	99.68	0.32	0.00	0.00	0.05	0.36	99.59	0.00	2.65	6.73	90.62
	($h = 1$)	0.5	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.97	3.78	0.25	0.00	99.67	0.33	0.00	0.00	0.05	0.43	99.52	0.00	2.93	6.92	90.15
	SF	0.3	99.49	0.49	0.02	0.00	100.00	0.00	0.00	89.05	8.19	2.35	0.41	99.62	0.37	0.01	0.00	0.64	1.72	6.54	91.10	24.94	14.28	23.45	37.33
	($h = 1, m = 0$)	0.5	99.50	0.47	0.03	0.00	100.00	0.00	0.00	89.73	7.59	2.30	0.38	99.67	0.33	0.00	0.00	0.90	1.41	5.89	91.80	27.34	12.81	22.68	37.17
5	SF	0.3	0.67	98.07	1.23	0.03	14.70	85.30	0.00	0.00	84.40	13.21	2.39	0.84	98.18	0.95	0.03	0.00	1.11	3.86	95.03	0.01	20.51	24.32	55.16
	($h = 1, m = 1$)	0.5	0.03	98.76	1.18	0.03	2.60	97.40	0.00	0.00	84.61	13.09	2.30	0.04	99.05	0.88	0.03	0.00	0.93	4.03	95.04	0.00	20.30	24.17	55.53
6	SF	0.3	2.70	2.58	93.70	1.02	49.96	6.34	43.70	0.00	0.01	0.25	85.58	14.16	3.28	2.94	93.02	0.76	0.00	2.34	97.66	0.00	0.00	24.53	75.47
	($h = 1, m = 2$)	0.5	2.33	1.21	95.45	1.01	48.97	3.08	47.95	0.00	0.02	0.06	85.84	14.08	3.13	1.45	94.70	0.72	0.00	2.33	97.67	0.00	0.00	24.24	75.76
7	RF	0.3	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.98	3.77	0.25	0.00	99.69	0.30	0.01	0.00	0.07	0.42	99.51	0.00	2.98	6.63	90.39
	($h = 1$)	0.5	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	96.23	3.59	0.18	0.00	99.72	0.28	0.00	0.00	0.09	0.49	99.42	0.00	3.07	7.11	89.82
	SF	0.3	0.18	98.81	1.01	0.00	6.30	93.69	0.01	0.00	85.14	12.12	2.74	0.26	98.93	0.81	0.00	0.00	0.74	3.90	95.36	0.00	16.74	25.16	58.10
	($h = 1, m = 1$)	0.5	0.09	98.99	0.92	0.00	3.27	96.73	0.00	0.00	85.77	11.70	2.53	0.09	99.17	0.74	0.00	0.00	0.82	3.56	95.62	0.00	18.33	23.90	57.77
8	RF	0.3	0.16	0.02	99.67	0.15	4.23	0.05	95.72	0.00	0.00	0.00	94.84	5.16	0.00	0.00	99.48	0.52	0.00	0.19	99.81	0.00	0.00	3.94	96.06
	($h = 2$)	0.5	0.00	99.90	0.10	0.00	0.00	100.00	0.00	0.00	95.15	4.85	0.00	0.00	99.51	0.49	0.00	0.00	0.28	99.72	0.00	0.00	4.52	95.48	
	SF	0.3	0.00	98.86	1.11	0.03	0.16	99.81	0.03	0.00	84.59	12.91	2.50	0.00	99.10	0.89	0.01	0.00	0.76	3.36	95.88	0.00	16.23	24.15	59.62
	($h = 2, m = 1$)	0.5	0.00	98.97	1.01	0.02	0.15	99.83	0.02	0.00	85.37	12.27	2.36	0.00	99.13	0.86	0.01	0.00	0.79	3.70	95.51	0.00	17.65	24.02	58.33
9	SF	0.3	18.93	7.43	72.90	0.74	85.12	4.00	10.88	0.00	0.92	1.78	85.49	11.81	22.46	7.88	69.09	0.57	0.00	2.14	97.86	0.00	0.00	24.23	75.77
	($h = 2, m = 2$)	0.5	16.45	5.47	77.44	0.64	83.01	3.43	13.56	0.00	0.74	1.16	86.82	11.28	19.45	5.91	74.13	0.51	0.00	2.56	97.44	0.00	0.00	26.19	73.81

Note: Same as Table II.

Table V. Empirical distributions of estimated numbers of breaks, AR(1) errors, $T = 240, \beta_2 = 0.1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC									
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3							
1	RF	0.3	100.00	0.00	0.00	0.00	100.00	0.00	0.00	97.23	2.46	0.28	0.03	99.80	0.20	0.00	0.00	0.02	0.06	0.53	99.39	2.33	2.53	7.74	87.40	
	($h = 0$)	0.5	100.00	0.00	0.00	100.00	0.00	0.00	97.23	2.46	0.28	0.03	99.80	0.20	0.00	0.00	0.02	0.06	0.53	99.39	2.33	2.53	7.74	87.40		
	SF	0.3	64.62	18.37	12.02	4.99	95.74	3.73	0.49	0.04	17.13	14.82	24.50	43.55	17.57	10.55	3.93	0.00	0.00	0.06	99.94	0.87	0.38	2.76	95.99	
	($h = 0, m = 0$)	0.5	64.69	18.36	12.05	4.90	95.71	3.74	0.51	0.04	17.26	14.78	24.19	43.77	17.60	10.61	3.91	0.00	0.01	0.05	99.94	0.55	0.40	2.56	96.49	
2	SF	0.3	0.06	63.23	24.48	12.23	1.02	93.18	5.42	0.38	0.00	18.41	23.34	58.25	0.06	66.56	23.09	10.29	0.00	0.01	0.04	99.95	0.00	0.23	1.83	97.94
	($h = 0, m = 1$)	0.5	0.05	63.09	24.42	12.44	0.97	93.20	5.46	0.37	0.00	18.24	23.49	58.27	0.06	66.26	23.29	10.39	0.00	0.01	0.02	99.97	0.00	0.30	1.57	98.13
3	SF	0.3	1.43	1.93	66.91	29.73	19.42	6.54	70.69	3.35	0.07	0.16	25.64	74.13	1.76	2.23	69.29	26.72	0.00	0.00	100.00	0.00	0.00	0.92	99.08	
	($h = 0, m = 2$)	0.5	1.35	1.76	66.95	29.94	18.58	6.38	71.61	3.43	0.07	0.14	25.57	74.22	1.63	2.09	69.40	26.88	0.00	0.00	100.00	0.00	0.00	0.85	99.15	
4	RF	0.3	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	0.00	95.63	4.10	0.27	0.00	99.61	0.39	0.00	0.00	0.01	0.34	99.65	0.00	2.31	5.85	91.84
	($h = 1$)	0.5	0.00	99.96	0.04	0.00	0.00	100.00	0.00	0.00	0.00	95.70	4.03	0.27	0.00	99.65	0.35	0.00	0.00	0.03	0.38	99.59	0.00	2.61	6.35	91.04
	SF	0.3	66.68	17.18	11.19	4.95	96.12	3.43	0.39	0.06	18.01	14.56	23.86	43.57	69.81	16.44	9.87	3.88	0.00	0.00	0.03	99.97	0.46	0.45	2.82	96.27
	($h = 1, m = 0$)	0.5	66.64	16.78	11.65	4.93	96.18	3.33	0.44	0.05	18.46	14.13	24.18	43.23	70.07	15.88	10.19	3.86	0.00	0.00	0.01	99.99	0.41	0.43	2.63	96.53
5	SF	0.3	1.94	62.40	23.59	12.07	16.32	78.70	4.72	0.26	0.11	17.95	24.30	57.64	2.23	65.30	22.45	10.02	0.00	0.00	0.02	99.98	0.01	0.25	1.93	97.81
	($h = 1, m = 1$)	0.5	0.61	63.10	23.83	12.46	6.54	88.03	5.12	0.31	0.03	18.07	23.81	58.09	0.75	66.09	22.74	10.42	0.00	0.00	0.00	100.00	0.00	0.17	1.86	97.97
6	SF	0.3	2.12	4.51	65.25	28.12	26.03	11.69	59.54	2.74	0.10	0.62	26.58	72.70	2.71	5.17	66.98	25.14	0.00	0.00	0.00	100.00	0.00	0.00	0.89	99.11
	($h = 1, m = 2$)	0.5	2.26	3.40	66.10	28.24	26.07	9.10	61.88	2.95	0.08	0.49	26.76	72.67	2.75	3.92	67.69	25.64	0.00	0.00	0.00	100.00	0.00	0.00	0.85	99.15
7	RF	0.3	0.00	99.96	0.04	0.00	0.01	99.99	0.00	0.00	0.00	95.74	3.86	0.40	0.00	99.77	0.23	0.00	0.00	0.08	0.32	99.60	0.00	2.62	5.43	91.95
	($h = 1$)	0.5	0.00	99.97	0.03	0.00	0.00	99.99	0.01	0.00	0.00	95.89	3.73	0.38	0.00	99.74	0.26	0.00	0.00	0.13	0.35	99.52	0.00	2.93	6.09	90.98
	SF	0.3	1.20	62.90	24.45	11.45	10.95	83.65	5.08	0.32	0.09	18.08	25.09	56.74	1.46	65.32	23.43	9.79	0.00	0.00	0.03	99.97	0.00	0.24	1.64	98.12
	($h = 1, m = 1$)	0.5	0.70	63.78	23.72	11.80	7.53	87.28	4.83	0.36	0.07	18.90	24.55	56.48	0.86	66.65	22.41	10.08	0.00	0.00	0.01	99.99	0.00	0.24	1.57	98.19
8	RF	0.3	0.43	0.03	99.46	0.08	7.81	0.15	92.04	0.00	0.00	94.55	5.45	0.08	0.00	99.37	0.55	0.00	0.00	0.13	0.13	99.87	0.00	3.34	96.66	
	($h = 2$)	0.5	0.00	0.00	99.96	0.04	0.02	0.00	99.98	0.00	0.00	94.73	5.27	0.00	0.00	99.49	0.51	0.00	0.00	0.24	0.24	99.76	0.00	4.03	95.97	
	SF	0.3	0.12	63.03	25.36	11.49	1.71	92.85	5.16	0.28	0.00	17.87	24.67	57.46	0.13	66.17	24.05	9.65	0.00	0.00	0.02	99.98	0.00	0.16	1.48	98.36
	($h = 2, m = 1$)	0.5	0.11	63.28	24.83	11.78	1.62	92.90	5.19	0.29	0.00	18.32	24.01	57.67	0.13	66.66	23.12	10.09	0.00	0.00	0.02	99.98	0.00	0.24	1.43	98.33
9	SF	0.3	8.74	9.96	56.92	24.38	50.04	14.65	33.67	1.64	0.50	1.73	26.21	71.56	10.36	10.62	57.34	21.68	0.00	0.00	0.02	99.98	0.00	0.01	0.90	99.09
	($h = 2, m = 2$)	0.5	8.24	8.62	58.78	24.36	48.69	13.40	36.20	1.71	0.50	1.35	26.78	71.37	9.48	9.39	59.83	21.30	0.00	0.00	0.02	99.98	0.00	0.01	0.99	99.00

Note: Same as Table II.

Table VI. Empirical distributions of estimated numbers of breaks, i.i.d. errors, $T = 120, \beta_2 = 1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC								
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3						
1	0.3	99.81	0.19	0.00	100.00	0.00	0.00	93.54	5.24	1.00	0.22	99.39	0.58	0.03	0.00	0.01	0.04	0.13	99.82	0.94	1.02	2.99	95.05		
($h = 0$)	0.5	99.81	0.19	0.00	100.00	0.00	0.00	93.54	5.24	1.00	0.22	99.39	0.58	0.03	0.00	0.01	0.04	0.13	99.82	0.94	1.02	2.99	95.05		
	0.3	98.49	1.40	0.11	0.00	99.97	0.03	0.00	86.17	9.16	3.66	1.01	99.16	0.82	0.02	0.00	8.24	4.13	14.48	73.15	70.28	7.56	11.57	10.59	
($h = 0, m = 0$)	0.5	98.50	1.38	0.12	0.00	99.97	0.03	0.00	86.12	9.28	3.59	1.01	99.17	0.81	0.02	0.00	8.50	4.15	14.20	73.15	70.33	7.64	11.40	10.63	
2	0.3	0.00	90.67	8.22	1.11	0.02	99.59	0.37	0.02	0.00	65.22	22.37	12.41	0.00	93.75	5.74	0.51	0.00	3.77	8.27	87.96	0.00	35.14	22.65	42.21
($h = 0, m = 1$)	0.5	0.00	90.67	8.16	1.17	0.00	99.61	0.38	0.01	0.00	65.17	22.54	12.29	0.00	93.69	5.70	0.61	0.00	3.69	7.93	88.38	0.00	35.58	22.62	41.80
3	0.3	0.30	0.19	94.09	5.42	23.10	1.74	75.00	0.16	0.00	0.01	73.42	26.57	0.69	0.38	95.52	3.41	0.00	0.00	7.69	92.31	0.00	0.00	43.21	56.79
($h = 0, m = 2$)	0.5	0.00	0.01	94.28	5.71	3.75	0.41	95.63	0.21	0.00	0.00	72.99	27.01	0.03	0.04	96.36	3.57	0.00	0.00	7.26	92.74	0.00	0.00	43.80	56.20
4	0.3	0.07	99.56	0.37	0.00	0.96	99.04	0.00	0.00	0.00	88.35	9.59	2.06	0.00	98.73	1.22	0.05	0.00	0.12	99.86	0.00	0.82	2.21	96.97	
($h = 1$)	0.5	0.00	99.65	0.35	0.00	0.00	99.99	0.01	0.00	0.00	88.79	9.09	2.12	0.00	98.74	1.20	0.06	0.00	0.03	0.10	99.87	0.00	0.99	2.36	96.65
	0.3	99.49	0.46	0.05	0.00	100.00	0.00	0.00	92.97	4.27	2.17	0.59	99.77	0.21	0.02	0.00	5.93	4.88	14.55	74.64	61.13	11.41	14.45	13.01	
($h = 1, m = 0$)	0.5	99.58	0.39	0.03	0.00	99.99	0.01	0.00	93.43	3.95	2.03	0.59	99.81	0.17	0.02	0.00	7.07	4.48	13.64	74.81	65.91	9.35	12.92	11.82	
5	0.3	0.75	90.59	7.36	1.30	12.19	87.45	0.34	0.02	0.04	67.01	21.07	11.88	1.14	93.12	5.15	0.59	0.00	3.46	7.08	89.46	0.00	34.61	21.40	43.99
($h = 1, m = 1$)	0.5	0.00	91.13	7.62	1.25	0.03	99.53	0.42	0.02	0.00	67.16	20.72	12.12	0.00	94.18	5.22	0.60	0.00	3.74	7.35	88.91	0.00	35.45	21.02	43.53
6	0.3	3.89	10.10	81.37	4.64	50.23	18.48	31.24	0.05	0.16	2.19	72.15	25.50	7.02	12.49	77.71	2.78	0.00	0.00	6.34	93.66	0.01	0.01	40.30	59.68
($h = 1, m = 2$)	0.5	0.25	0.36	93.72	5.67	14.53	3.33	81.98	0.16	0.00	0.03	73.85	26.12	0.48	0.54	95.60	3.38	0.00	0.00	6.74	93.26	0.00	0.00	41.09	58.91
7	0.3	0.70	98.92	0.38	0.00	4.79	95.21	0.00	0.00	0.05	89.43	8.59	1.93	0.26	98.39	1.31	0.04	0.00	0.01	0.06	99.93	0.00	0.96	2.33	96.71
($h = 1$)	0.5	0.00	99.67	0.33	0.00	0.03	99.97	0.00	0.00	0.00	89.65	8.51	1.84	0.00	98.72	1.25	0.03	0.00	0.01	0.07	99.92	0.00	1.12	2.71	96.17
	0.3	1.90	94.18	3.59	0.33	22.13	77.79	0.08	0.00	0.12	76.01	16.11	7.76	2.81	94.92	2.13	0.14	0.00	2.71	6.37	90.92	0.04	31.00	23.66	45.30
($h = 1, m = 1$)	0.5	0.00	96.61	3.01	0.38	0.19	99.77	0.04	0.00	0.00	77.57	13.73	8.70	0.00	98.01	1.79	0.20	0.00	3.18	6.24	90.58	0.00	34.08	21.32	44.60
8	0.3	23.08	1.08	75.37	0.47	71.76	0.65	27.59	0.00	0.77	0.22	83.96	15.05	9.62	0.74	87.60	2.04	0.00	0.00	0.02	99.98	0.00	0.00	1.05	98.95
($h = 2$)	0.5	0.44	0.01	98.97	0.58	9.74	0.12	90.14	0.00	0.00	0.00	85.62	14.38	0.07	0.00	97.89	2.04	0.00	0.00	0.02	99.98	0.00	0.00	1.29	98.71
	0.3	0.17	92.69	6.14	1.00	3.55	95.72	0.70	0.03	0.01	76.67	15.97	7.35	0.12	96.49	3.05	0.34	0.00	2.35	7.23	90.42	0.00	27.42	27.06	45.52
($h = 2, m = 1$)	0.5	0.00	96.05	3.34	0.61	0.27	99.46	0.25	0.02	0.00	77.98	14.02	8.00	0.00	97.63	2.12	0.25	0.00	2.86	7.06	90.08	0.00	30.81	24.47	44.72
9	0.3	29.50	6.14	62.77	1.59	89.04	3.57	7.39	0.00	3.06	1.38	81.41	14.15	36.09	4.88	58.32	0.71	0.00	0.01	6.61	93.38	0.12	0.03	42.74	57.11
($h = 2, m = 2$)	0.5	4.88	1.67	91.81	1.64	51.41	3.58	45.01	0.00	0.32	0.27	85.81	13.60	7.99	1.94	89.28	0.79	0.00	0.00	7.67	92.33	0.01	0.00	46.31	53.68

Note: Same as Table II.

Table VII. Empirical distributions of estimated numbers of breaks, AR(1) errors, $T = 120, \beta_2 = 1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC									
		0	1	≥3	0	1	≥3	0	1	≥3	0	1	≥3	0	1	≥3	0	1	≥3							
1	RF	0.3	99.84	0.16	0.00	99.99	0.01	0.00	0.00	93.30	5.23	1.05	0.42	99.37	0.62	0.01	0.00	0.01	0.03	0.06	99.90	0.88	1.06	2.75	95.31	
	($h = 0$)	0.5	99.84	0.16	0.00	99.99	0.01	0.00	0.00	93.30	5.23	1.05	0.42	99.37	0.62	0.01	0.00	0.01	0.03	0.06	99.90	0.88	1.06	2.75	95.31	
	SF	0.3	85.10	10.57	3.67	0.66	98.91	1.00	0.09	0.00	46.10	18.69	19.75	15.46	89.31	8.09	2.31	0.29	0.26	0.47	2.55	96.72	22.70	6.94	15.70	54.66
	($h = 0, m = 0$)	0.5	85.05	10.60	3.65	0.70	98.91	1.00	0.08	0.00	45.99	18.81	19.31	15.89	89.37	8.06	2.27	0.30	0.15	0.49	2.71	96.65	21.46	6.62	15.21	56.71
2	SF	0.3	0.01	69.58	21.44	8.97	0.14	94.78	4.72	0.36	0.00	32.76	29.52	37.72	0.01	75.55	18.51	5.93	0.00	0.16	0.96	98.88	0.00	5.21	11.64	83.15
	($h = 0, m = 1$)	0.5	0.00	68.62	22.03	9.35	0.01	94.69	4.93	0.37	0.00	32.41	29.29	38.30	0.00	74.91	18.85	6.24	0.00	0.10	0.99	98.91	0.00	4.51	11.33	84.16
3	SF	0.3	0.81	1.25	75.66	22.28	20.95	5.07	71.79	2.19	0.05	0.09	41.95	57.91	1.70	1.75	80.14	16.41	0.00	0.00	0.61	99.39	0.00	0.00	10.25	89.75
	($h = 0, m = 2$)	0.5	0.15	0.20	76.06	23.59	6.51	1.77	88.94	2.78	0.00	0.01	40.61	59.38	0.28	0.30	81.63	17.79	0.00	0.00	0.67	99.33	0.00	0.00	10.13	89.87
4	RF	0.3	0.31	99.24	0.45	0.00	2.47	97.53	0.00	0.00	0.01	87.77	9.85	2.37	0.08	98.59	1.26	0.07	0.00	0.01	0.01	99.98	0.00	0.69	1.97	97.34
	($h = 1$)	0.5	0.00	99.57	0.43	0.00	0.02	99.98	0.00	0.00	0.00	88.45	9.30	2.25	0.00	98.63	1.30	0.07	0.00	0.01	0.01	99.98	0.00	0.74	1.99	97.27
	SF	0.3	91.23	6.07	2.19	0.51	99.58	0.38	0.04	0.00	57.40	15.72	14.64	12.24	94.28	4.31	1.18	0.23	0.14	0.54	2.70	96.62	17.89	8.04	17.05	57.02
	($h = 1, m = 0$)	0.5	91.58	5.73	2.16	0.53	99.65	0.32	0.03	0.00	58.05	14.23	14.95	12.77	94.69	3.93	1.16	0.22	0.12	0.51	2.26	97.11	18.04	7.22	16.19	58.55
5	SF	0.3	0.90	69.02	21.34	8.74	13.82	81.98	3.90	0.30	0.08	32.71	30.05	37.16	1.66	75.27	17.45	5.62	0.00	0.11	0.66	99.23	0.00	4.34	10.81	84.85
	($h = 1, m = 1$)	0.5	0.01	69.29	21.30	9.40	0.25	94.64	4.71	0.40	0.00	32.78	29.53	37.69	0.01	76.05	17.76	6.18	0.00	0.11	0.73	99.16	0.00	4.31	10.67	85.02
6	SF	0.3	4.62	9.34	66.55	19.49	41.67	20.04	37.15	1.14	0.22	1.81	41.43	56.54	7.41	11.80	66.91	13.88	0.00	0.00	0.51	99.49	0.01	0.00	8.99	91.00
	($h = 1, m = 2$)	0.5	0.68	0.87	75.66	22.79	15.74	5.40	76.73	2.13	0.03	0.04	41.18	58.75	1.30	1.37	80.33	17.00	0.00	0.00	0.56	99.44	0.00	0.00	8.75	91.25
7	RF	0.3	2.08	97.52	0.39	0.01	9.74	90.23	0.03	0.00	0.10	88.24	9.30	2.36	0.93	97.71	1.29	0.07	0.00	0.01	0.06	99.93	0.00	0.71	1.98	97.31
	($h = 1$)	0.5	0.02	99.65	0.32	0.01	0.23	99.75	0.02	0.00	0.00	88.73	9.16	2.11	0.01	98.54	1.37	0.08	0.00	0.02	0.09	99.89	0.00	0.79	2.26	96.95
	SF	0.3	2.28	74.81	17.30	5.61	22.81	74.87	2.19	0.13	0.13	39.00	28.71	32.16	3.17	80.12	13.49	3.22	0.00	0.11	0.98	98.91	0.03	4.94	11.33	83.70
	($h = 1, m = 1$)	0.5	0.00	77.55	15.81	6.64	0.61	97.13	2.15	0.11	0.00	39.94	24.85	35.21	0.04	83.99	12.12	3.85	0.00	0.08	0.78	99.14	0.00	4.73	10.61	84.66
8	RF	0.3	29.28	2.39	67.86	0.47	73.17	1.86	24.96	0.01	2.09	0.65	81.24	16.02	14.73	1.93	81.16	2.18	0.00	0.00	0.02	99.98	0.00	0.00	0.75	99.25
	($h = 2$)	0.5	1.85	0.14	97.48	0.53	16.59	0.47	82.93	0.01	0.03	0.00	84.04	15.93	0.56	0.14	96.95	2.35	0.00	0.00	0.04	99.96	0.00	0.00	1.09	98.91
	SF	0.3	0.42	74.56	18.42	6.60	5.52	90.98	3.25	0.25	0.02	40.47	29.95	29.56	0.36	82.14	13.93	3.57	0.00	0.09	0.90	99.01	0.00	3.77	11.56	84.67
	($h = 2, m = 1$)	0.5	0.00	77.38	16.47	6.15	0.90	96.11	2.84	0.15	0.00	41.54	27.57	30.89	0.01	83.57	12.86	3.56	0.00	0.11	0.95	98.94	0.00	4.17	11.13	84.70
9	SF	0.3	23.02	10.04	56.23	10.71	77.41	9.67	12.82	0.10	2.90	2.44	48.75	45.91	29.19	8.94	54.83	7.04	0.00	0.00	0.59	99.41	0.02	0.03	10.97	88.98
	($h = 2, m = 2$)	0.5	6.84	3.30	76.83	13.03	46.62	7.67	45.25	0.46	0.51	0.75	51.97	46.77	9.77	3.89	77.55	8.79	0.00	0.00	0.58	99.42	0.00	0.01	11.16	88.83

Note: Same as Table II.

Table VIII. Empirical distributions of estimated numbers of breaks, i.i.d. errors, $T = 240$, $\beta_2 = 1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC								
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3						
1 ($h = 0$)	RF	0.3	100.00	0.00	0.00	0.00	0.00	0.00	97.29	2.43	0.26	0.02	99.84	0.16	0.00	0.00	0.02	0.06	0.49	99.43	2.47	2.64	7.53	87.36	
	SF	0.5	100.00	0.00	0.00	100.00	0.00	0.00	97.29	2.43	0.26	0.02	99.84	0.16	0.00	0.00	0.02	0.06	0.49	99.43	2.47	2.64	7.53	87.36	
		0.3	99.37	0.59	0.04	0.00	100.00	0.00	0.00	89.12	7.78	2.65	0.47	99.54	0.45	0.01	0.00	10.20	4.11	15.54	70.15	81.58	4.41	7.94	6.07
2 ($h = 0, m = 0$)	RF	0.5	99.33	0.62	0.05	0.00	100.00	0.00	0.00	89.06	7.85	2.62	0.47	99.52	0.46	0.02	0.00	10.49	4.64	15.70	69.17	81.88	4.70	7.66	5.76
	SF	0.3	0.00	94.62	4.99	0.39	0.00	99.72	0.28	0.00	71.11	20.04	8.85	0.00	95.36	4.29	0.35	0.00	3.83	8.20	87.97	0.00	35.68	21.78	42.54
		0.5	0.00	94.46	5.12	0.42	0.00	99.68	0.32	0.00	71.17	20.00	8.83	0.00	95.35	4.33	0.32	0.00	3.99	8.63	87.38	0.00	35.77	22.24	41.99
3 ($h = 0, m = 1$)	RF	0.3	0.00	97.20	2.80	0.02	0.00	99.93	0.05	0.00	79.22	20.78	0.00	0.00	97.62	2.38	0.00	0.00	0.00	9.22	90.78	0.00	0.00	46.12	53.88
	SF	0.5	0.00	97.22	2.78	0.00	0.00	99.95	0.05	0.00	79.15	20.85	0.00	0.00	97.67	2.33	0.00	0.00	0.00	9.42	90.58	0.00	0.00	46.90	53.10
		0.5	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.76	3.94	0.30	0.00	99.68	0.32	0.00	0.00	0.05	0.36	99.59	0.00	2.65	6.73	90.62
4 ($h = 1$)	RF	0.5	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.97	3.78	0.25	0.00	99.67	0.33	0.00	0.00	0.05	0.43	99.52	0.00	2.93	6.92	90.15
	SF	0.3	99.84	0.16	0.00	0.00	100.00	0.00	0.00	94.79	3.66	1.21	0.34	99.89	0.11	0.00	0.00	7.27	4.71	15.41	72.61	74.18	7.24	10.52	8.06
		0.5	99.87	0.13	0.00	0.00	100.00	0.00	0.00	95.14	3.24	1.28	0.34	99.91	0.09	0.00	0.00	8.46	4.40	15.38	71.76	77.88	5.69	9.14	7.29
5 ($h = 1, m = 0$)	RF	0.3	0.00	94.62	4.95	0.43	0.08	99.72	0.20	0.00	71.23	19.58	9.19	0.00	95.42	4.22	0.36	0.00	3.78	7.63	88.59	0.00	35.97	19.94	44.09
	SF	0.5	0.00	94.63	4.97	0.40	0.00	99.73	0.27	0.00	71.83	19.37	8.80	0.00	95.39	4.26	0.35	0.00	4.29	8.21	87.50	0.00	36.79	20.62	42.59
		0.3	0.01	96.67	2.76	2.10	7.69	90.17	0.04	0.00	0.07	79.51	20.42	0.01	0.71	97.00	2.28	0.00	0.00	7.57	92.43	0.00	0.00	42.93	57.07
6 ($h = 1, m = 2$)	RF	0.5	0.00	97.25	2.75	0.00	0.04	99.92	0.04	0.00	79.58	20.42	0.00	0.00	97.73	2.27	0.00	0.00	0.00	8.78	91.22	0.00	0.00	44.94	55.06
	SF	0.3	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	95.98	3.77	0.25	0.00	99.69	0.30	0.01	0.00	0.07	0.42	99.51	0.00	2.98	6.63	90.39
		0.5	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	96.23	3.59	0.18	0.00	99.72	0.28	0.00	0.00	0.09	0.49	99.42	0.00	3.07	7.11	89.82
7 ($h = 1$)	RF	0.3	0.00	98.20	1.67	0.13	0.52	99.45	0.03	0.00	82.11	11.65	6.24	0.00	98.58	1.34	0.08	0.00	2.84	6.12	91.04	0.00	32.16	21.52	46.32
	SF	0.5	0.00	98.49	1.41	0.10	0.00	99.98	0.02	0.00	82.93	10.41	6.66	0.00	98.81	1.12	0.07	0.00	3.33	5.95	90.72	0.00	35.83	18.28	45.89
		0.5	0.00	99.67	0.15	4.23	0.05	95.72	0.00	0.00	0.00	94.84	5.16	0.00	0.00	99.48	0.52	0.00	0.00	0.19	99.81	0.00	0.00	3.94	96.06
8 ($h = 2$)	RF	0.3	0.00	99.90	0.10	0.00	0.00	100.00	0.00	0.00	95.15	4.85	0.00	0.00	99.51	0.49	0.00	0.00	0.28	99.72	0.00	0.00	4.52	95.48	
	SF	0.5	0.00	97.82	1.92	0.26	0.02	99.78	0.20	0.00	80.82	13.04	6.14	0.00	98.36	1.49	0.15	0.00	2.38	7.39	90.23	0.00	27.36	26.12	46.52
		0.5	0.00	98.18	1.61	0.21	0.00	99.97	0.03	0.00	82.23	11.39	6.38	0.00	98.62	1.23	0.15	0.00	2.94	7.09	89.97	0.00	31.24	22.52	46.24
9 ($h = 2, m = 1$)	RF	0.3	1.00	0.15	98.15	0.70	33.62	1.44	64.93	0.01	0.02	0.00	89.90	10.08	1.32	0.15	98.04	0.49	0.00	6.71	93.29	0.00	0.00	46.07	53.93
	SF	0.5	0.01	0.01	99.41	0.57	2.26	0.14	97.60	0.00	0.00	90.84	9.16	0.01	0.01	99.53	0.45	0.00	0.00	8.26	91.74	0.00	0.00	49.60	50.40
		0.5	0.01	0.01	99.41	0.57	2.26	0.14	97.60	0.00	0.00	90.84	9.16	0.01	0.01	99.53	0.45	0.00	0.00	8.26	91.74	0.00	0.00	49.60	50.40

Note: Same as Table II.

Table IX. Empirical distributions of estimated numbers of breaks, AR(1) errors, $T = 240, \beta_2 = 1$

Case	R^2	BIC			SBBIC			HQIC			SBHQIC			AIC			SBAIC									
		0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3	0	1	≥ 3							
1	RF	0.3	100.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	97.23	2.46	0.28	0.03	99.80	0.20	0.00	0.00	0.02	0.06	0.53	99.39	2.33	2.53	7.74	87.40
	($h = 0$)	0.5	100.00	0.00	0.00	100.00	0.00	0.00	0.00	97.23	2.46	0.28	0.03	99.80	0.20	0.00	0.00	0.02	0.06	0.53	99.39	2.33	2.53	7.74	87.40	
	SF	0.3	88.60	8.53	2.46	0.41	99.35	0.64	0.01	0.00	47.51	19.53	13.70	90.42	7.39	1.90	0.29	0.19	0.27	1.74	97.80	32.53	4.30	12.28	49.89	
	($h = 0, m = 0$)	0.5	88.64	8.47	2.49	0.40	99.36	0.63	0.01	0.00	47.39	19.59	13.71	90.33	7.46	1.90	0.31	0.17	0.28	1.59	97.96	32.55	4.00	12.49	50.96	
2	SF	0.3	0.00	73.93	18.98	7.09	0.00	95.96	3.72	0.32	0.00	33.58	30.26	36.16	0.00	76.21	17.79	6.00	0.00	0.12	0.74	99.14	0.00	4.15	9.93	85.92
	($h = 0, m = 1$)	0.5	0.00	73.70	19.16	7.14	0.00	95.74	3.93	0.33	0.00	33.61	30.34	36.05	0.00	76.07	17.78	6.15	0.00	0.07	0.80	99.13	0.00	4.06	9.71	86.23
3	SF	0.3	0.00	0.01	81.48	18.51	0.27	0.11	97.66	1.96	0.00	0.00	43.34	56.66	0.00	0.01	83.42	16.57	0.00	0.00	0.63	99.37	0.00	0.00	9.07	90.93
	($h = 0, m = 2$)	0.5	0.00	0.00	80.97	19.03	0.01	0.00	97.92	2.07	0.00	0.00	43.30	56.70	0.00	0.01	82.90	17.10	0.00	0.00	0.54	99.46	0.00	0.00	9.14	90.86
4	RF	0.3	0.00	99.95	0.05	0.00	0.00	100.00	0.00	0.00	0.00	95.63	4.10	0.27	0.00	99.61	0.39	0.00	0.00	0.01	0.34	99.65	0.00	2.31	5.85	91.84
	($h = 1$)	0.5	0.00	99.96	0.04	0.00	0.00	100.00	0.00	0.00	0.00	95.70	4.03	0.27	0.00	99.65	0.35	0.00	0.00	0.03	0.38	99.59	0.00	2.61	6.35	91.04
	SF	0.3	93.31	4.81	1.55	0.33	99.83	0.17	0.00	0.00	57.97	15.02	15.65	11.36	94.46	4.14	1.14	0.26	0.09	0.27	1.87	97.77	27.37	5.30	13.37	53.96
	($h = 1, m = 0$)	0.5	93.56	4.55	1.53	0.36	99.83	0.17	0.00	0.00	58.81	14.20	15.62	11.37	94.76	3.91	1.06	0.27	0.10	0.34	1.81	97.75	28.69	4.59	12.55	54.17
5	SF	0.3	0.00	73.40	19.31	7.29	0.20	96.02	3.52	0.26	0.00	33.82	29.36	36.82	0.00	76.00	17.87	6.13	0.00	0.07	0.49	99.44	0.00	3.84	9.36	86.80
	($h = 1, m = 1$)	0.5	0.00	73.72	19.07	7.21	0.00	96.02	3.71	0.27	0.00	34.33	29.25	36.42	0.00	76.01	17.78	6.21	0.00	0.09	0.56	99.35	0.00	4.28	9.48	86.24
6	SF	0.3	0.11	0.86	80.73	18.30	4.41	9.14	84.82	1.63	0.00	0.03	43.71	56.26	0.12	1.15	82.63	16.10	0.00	0.00	0.44	99.56	0.00	0.00	7.64	92.36
	($h = 1, m = 2$)	0.5	0.00	0.00	81.07	18.93	0.11	0.07	97.85	1.97	0.00	0.00	43.66	56.34	0.00	0.00	83.24	16.76	0.00	0.00	0.54	99.46	0.00	0.00	8.18	91.82
7	RF	0.3	0.00	99.96	0.04	0.00	0.01	99.99	0.00	0.00	0.00	95.74	3.86	0.40	0.00	99.77	0.23	0.00	0.00	0.08	0.32	99.60	0.00	2.62	5.43	91.95
	($h = 1$)	0.5	0.00	99.97	0.03	0.00	0.00	99.99	0.01	0.00	0.00	95.89	3.73	0.38	0.00	99.74	0.26	0.00	0.00	0.13	0.35	99.52	0.00	2.93	6.09	90.98
	SF	0.3	0.00	81.77	13.26	4.97	1.05	97.38	1.53	0.04	0.00	40.39	25.50	34.11	0.02	84.02	11.97	3.99	0.00	0.03	0.50	99.47	0.00	3.72	8.97	87.31
	($h = 1, m = 1$)	0.5	0.00	82.81	11.50	5.69	0.00	98.58	1.35	0.07	0.00	41.70	23.10	35.20	0.00	84.89	10.54	4.57	0.00	0.12	0.52	99.36	0.00	4.01	8.34	87.65
8	RF	0.3	0.43	0.03	99.46	0.08	7.81	0.15	92.04	0.00	0.00	0.00	94.55	5.45	0.08	0.00	99.37	0.55	0.00	0.00	0.13	99.87	0.00	0.00	3.34	96.66
	($h = 2$)	0.5	0.00	0.00	99.96	0.04	0.02	0.00	99.98	0.00	0.00	0.00	94.73	5.27	0.00	0.00	99.49	0.51	0.00	0.00	0.24	99.76	0.00	0.00	4.03	95.97
	SF	0.3	0.00	81.16	14.11	4.73	0.05	97.52	2.29	0.14	0.00	41.23	28.46	30.31	0.00	83.25	12.85	3.90	0.00	0.02	0.60	99.38	0.00	3.06	9.30	87.64
	($h = 2, m = 1$)	0.5	0.00	82.36	12.32	5.32	0.00	98.40	1.46	0.14	0.00	43.22	25.36	31.42	0.00	84.53	11.03	4.44	0.00	0.05	0.57	99.38	0.00	3.36	9.11	87.53
9	SF	0.3	2.10	1.07	86.38	10.45	31.90	4.40	63.23	0.47	0.03	0.06	53.64	46.27	2.53	1.22	87.34	8.91	0.00	0.00	0.34	99.66	0.01	0.00	8.04	91.95
	($h = 2, m = 2$)	0.5	0.07	0.08	89.88	9.97	4.86	0.93	93.54	0.67	0.00	0.00	54.92	45.08	0.08	0.11	91.31	8.50	0.00	0.00	0.47	99.53	0.00	0.00	9.39	90.61

Note: Same as Table II.

In case 2, where a single break applies in the SF but the RF is stable ($h = 0, m = 1$), SBBIC is the most accurate method in terms of the correct detection of the SF break with i.i.d. disturbances in Tables IV, VI and VIII. This is followed by SBHQIC, which is marginally more accurate than SBBIC in Table II; these results correspond to the good performances of these two criteria in our OLS study (Hall *et al.*, 2013b). While BIC does well in Table II, when the magnitudes of the SF coefficients and changes are relatively small, it has a tendency to over-estimate the number of breaks for the larger magnitudes in Table VI. However, and not surprisingly, HQIC has a greater tendency than BIC to over-estimate the number of breaks.

The presence of any break is more difficult to detect when two reverting breaks apply with a stable RF (case 3, with $h = 0, m = 2$) in these tables, so that BIC, SBBIC and SBHQIC often erroneously imply no breaks are present, with this being particularly a feature of Table II. However, this is largely eliminated for the larger sample size and/or larger coefficient magnitudes (Tables IV, VI and VIII), except sometimes for SBBIC. Note in Table VI that the RF R^2 plays an important role for the performance of SBBIC, with the poor fitted values resulting from the DGP with relatively low explanatory power causing this criterion to often detect no breaks in the SF, whereas the higher R^2 value leads to much improved detection of the two breaks.

As in case 1 where no breaks occur, the presence of unmodelled AR(1) disturbances leads to an increased tendency for all criteria to detect (additional) spurious breaks in the SF; compare cases 2 and 3 of Table II with those of Table III and similarly of Tables IV and V, Tables VI and VII, and Tables VIII and IX. Although with $T = 240$ and $\beta_2 = 1$ in Table IX SBBIC has very good accuracy for detection of the true number of breaks in the SF with positively autocorrelated errors, its performance is less impressive at the other extreme of $T = 120$ and $\beta_2 = 0.1$ in Table III where it often under-specifies the numbers of true breaks, especially in case 3 when $m = 2$. On the other hand, BIC, SBHQIC and (especially) HQIC often over-estimate the number of breaks for cases 2 and 3 in this latter table.

Since the consistent criteria BIC, SBBIC, HQIC and SBHQIC correctly detect the presence of a single RF break in the vast majority of replications across cases 4 to 7, the characteristics just discussed largely continue to apply when $h = 1$. This can be seen particularly in Tables VIII and IX, where $T = 240$ and $\beta_2 = 1$, and the results for cases 4 to 7 overall reflect the corresponding cases 1 to 3 where $h = 0$. Other settings, however, show a greater influence from estimation of RF breaks.

For the same SF coefficients as in Tables VIII and IX, but with the smaller sample size of $T = 120$, Tables VI and VII illustrate the additional difficulties that apply when the DGP exhibits RF breaks. Compared with results for $h = 0$, SBBIC more often under-estimates the number of SF breaks for cases 4 to 7 when the RF R^2 is low at 0.3, but the performance more closely matches that for $h = 0$ when $R^2 = 0.5$. Further, the relative timing of breaks in the RF and SF plays a role with this criterion. In Tables VI and VII, for example, $m = 1$ is more often correctly specified using SBBIC for case 5 (when $\pi_1 = \lambda_1 = 0.5$) with $R^2 = 0.3$ than for case 7 (where $\pi_1 = 0.3, \lambda_1 = 0.6$). On the other hand, for the smaller breaks in Tables II and III, where $\beta_2 = 0.1$, SBBIC (and in general BIC) has better performance for case 7 than case 5. Overall, the performance of SBHQIC is more robust to the timing of these breaks.

When $h = 2$ in cases 8 and 9 with $T = 120$, BIC and (to a greater extent) SBBIC can miss the presence of any RF break, particularly when the coefficients are of smaller magnitude and $R^2 = 0.3$. Table VI, for example, shows how this leads to a deterioration in the performance of these criteria for the detection of SF breaks compared with the situation when $R^2 = 0.5$, with this being particularly clear for case 9 with two SF breaks. This feature is also seen, but to a lesser extent, in the number of SF breaks detected by SBHQIC. The different performances of SBBIC for the two RF scenarios in case 9 extends also to $T = 240$ in Tables VIII and IX. This applies despite the criterion correctly detecting two RF breaks in at least 92% of replications, suggesting a role for the estimation of the RF break dates themselves, and not simply the number of these.

Overall, these simulation results indicate that the best performing criteria are SBBIC and SBHQIC. The former works well for the detection of breaks in both the RF and SF equations across many of the cases considered but can fail to detect any breaks when two breaks of the reverting form are present in the SF. With breaks of such reverting form, the use of SBHQIC more satisfactorily detects the presence of breaks, but at the cost of over-specifying the number of breaks in other cases. The heavier weighting of break date estimation implied by the use of $k = 3$ in

(14) and (15) generally works better than $k = 1$, while the inconsistent AIC-based criteria do not appear to be useful if the correct detection of the number of breaks is an important consideration.

5. CONCLUSIONS

This article is, to the best of our knowledge, the first to investigate the use of IC for inference on structural breaks when the coefficients of a linear model with endogenous regressors may experience multiple changes. Hall *et al.* (2012) provide a methodology for such inference using a hypothesis testing approach, but we are able to avoid the possible inconsistency inherent in using sequential hypothesis tests with a fixed type 1 error and also the sample splitting their method involves. We believe that the use of an (appropriate) information criterion offers a different route to inference on the number of breaks in a structural equation, which we hope will be useful to practitioners.

The paper makes two specific contributions. First, we show that suitably defined IC yield estimators of the number of breaks, when employed in the second stage of a 2SLS procedure with breaks in the RF taken into account in the first stage. Second, a Monte Carlo analysis investigates the finite sample performance of a range of criteria based on BIC, HQIC and AIC for equations estimated by 2SLS. Versions of the consistent criteria BIC and HQIC perform well overall when the penalty term weights estimation of each break point more heavily than estimation of each coefficient. However, AIC is inconsistent and badly over-estimates the number of true breaks; we recommend that this criterion should not be used for empirical estimation of the number of structural breaks.

APPENDIX A:

Proof of Theorem 1

First, we state Lemma 1, which presents the limiting behaviour of $RSS_j(\tau(n); n, \hat{\pi})$; the proof can be found in the Supporting Information. The notation $\xrightarrow{p,u}$ is used to denote convergence in probability uniformly in τ over the specified intervals. □

Lemma 1. Let y_t be generated by (1), x_t by (2) and $\hat{x}_t(\hat{\pi})$ by (4), and Assumptions 1–9 hold. Then, for segment j of the data, $t = [\tau_{j-1}T] + 1, \dots, [\tau_j T]$:

(i) If $\lambda_{i-1}^0 \leq \tau_{j-1}, \tau_j \leq \lambda_i^0$, then

$$T^{-1}RSS_j(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} (\tau_j - \tau_{j-1})\Gamma_i.$$

(ii) If there exists i and $\kappa > 0$ such that $\lambda_i^0, \lambda_{i+1}^0, \dots, \lambda_{i+\kappa}^0 \in [\tau_{j-1}, \tau_j]$, then

$$T^{-1}RSS_j(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} (\lambda_i^0 - \tau_{j-1})\Gamma_i + (\lambda_{i+1}^0 - \lambda_i^0)\Gamma_{i+1} + \dots + (\lambda_{i+\kappa}^0 - \lambda_{i+\kappa-1}^0)\Gamma_{i+\kappa} + (\tau_j - \lambda_{i+\kappa}^0)\Gamma_{i+\kappa+1} + F_j$$

where, in both cases, Γ_i is defined by (13), the limit in probability exists uniformly in a segment defined by $\tau_{j-1} + \epsilon < \tau_j$, for $\epsilon > 0$ and $\tau_{j-1}, \tau_j \in [0, 1]$, while F_j is a positive constant that depends on τ_{j-1}, τ_j , certain limit matrices and the parameters of the model.

Denote

$$\Gamma(\lambda^0, m, \beta^0) = \sum_{j=1}^{m+1} (\lambda_j^0 - \lambda_{j-1}^0) \left(\sigma_u^2 + 2\Sigma_{uv}\beta_j^0 + \beta_j^{0'}\Sigma_v\beta_j^0 \right)$$

where $\lambda^0 = (\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)'$ and $\beta^0 = (\beta_1^0, \beta_2^0, \dots, \beta_{m+1}^0)'$. $\Gamma(\lambda^0, m, \beta^0)$ is then the sum of the Γ_i of (13) across all true segments of the data.

Now, consider the behaviour of the information criterion (8) in cases where the researcher selects n that may over-fit, under-fit or correctly identify the true number of breaks (m) in the structural form model:

(1) $n = m$. The correct number of breaks is employed, and one of the following two scenarios applies.

(1.1) If $\tau(n) = \lambda^0$, the m break dates are correctly identified and, using part (i) of Lemma 1 together with Assumptions 10 and 11,

$$\text{IC}(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} \Gamma(\lambda^0, m, \beta^0)$$

(1.2) If $\tau(n) \neq \lambda^0$, there must exist j s.t. $[\tau_{j-1}T] + 1, \dots, [\tau_jT]$ contains at least one neglected break, and therefore Lemma 1 part (ii), with Assumptions 10 and 11, implies that

$$\text{IC}(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} \Gamma(\lambda^0, m, \beta^0) + F(\tau(n), \lambda^0)$$

where $F(\tau(n), \lambda^0) > 0$.

(2) $n < m$. The model is under-fitted, and there must exist a segment j s.t. $[\tau_{j-1}T] + 1, \dots, [\tau_jT]$ that contains at least one neglected break, and, as in case (1.2),

$$\text{IC}(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} \Gamma(\lambda^0, m, \beta^0) + F(\tau(n), \lambda^0)$$

where $F(\tau(n), \lambda^0) > 0$.

(3) $n > m$. Then the following two scenarios are possible:

(3.1) If $\tau(n)$ does not contain λ^0 , then there must exist j s.t. $[\tau_{j-1}T] + 1, \dots, [\tau_jT]$ includes at least one λ_i^0 and, again as in case (1.2),

$$\text{IC}(\tau(n); n, \hat{\pi}) \xrightarrow{p,u} \Gamma(\lambda^0, m, \beta^0) + F(\tau(n), \lambda^0)$$

where $F(\tau(n), \lambda^0) > 0$.

(3.2) If $\tau(n)$ contains λ^0 , consider

$$D_T = \{\text{IC}(\tau(n); n, \hat{\pi}) - \text{IC}(\lambda^0; m, \hat{\pi})\}$$

and

$$\begin{aligned} D_T &= T \ln \left\{ \hat{\Gamma}(\tau(n); n, \hat{\beta}) / \Gamma(\lambda^0; m, \beta^0) \right\} + T \{K(n, T) - K(m, T)\} \\ &= -\text{QLR}_T + T \{K(n, T) - K(m, T)\} \end{aligned}$$

where QLR_T is the quasi-likelihood ratio test for $H_0 : \tau(m) = \lambda^0, \beta^* = \beta^0$ vs. the alternative in which $\tau(n)$ contains λ^0 and $n > m$. Under its H_0 by standard arguments $QLR_T = Op(1)$, and since $T \{K(n, T) - K(m, T)\} \rightarrow +\infty$ from Assumption 10, it follows that

$$D_T \rightarrow \infty.$$

Taken together, cases (1), (2) and (3) imply the stated result.

ACKNOWLEDGEMENTS

The authors acknowledge the support from the ESRC (UK), under grant RES-062-23-1351. We are grateful to two anonymous referees of this journal for their constructive reports on an earlier version of the paper. The second author was a colleague of John Nankervis at the University of Manchester early in their respective academic careers. He was a friend of many and is sadly missed.

REFERENCES

- Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**: 716–723.
- Andrews DWK. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**: 821–856.
- Andrews DWK. 2003. End-of-sample stability tests. *Econometrica* **71**: 1661–1694.
- Bai J. 2000. Vector autoregressive models with structural changes in regression coefficients and in variance–covariance matrices. *Annals of Economics and Finance* **1**: 303–339.
- Bai J, Perron P. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* **66**: 47–78.
- Bai J, Perron P. 2003. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* **18**: 1–22.
- Bai J, Perron P. 2006. *Multiple structural change models: a simulation analysis*. In *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, Corbae D, Durlauf SN, Hansen BE (eds.) Cambridge University Press: New York, pp. 212–237.
- Boldea O, Hall AR, Han S. 2012. Asymptotic distribution theory for break point estimators in models estimated via 2SLS. *Econometric Reviews* **31**: 1–33.
- Dufour J-M, Ghysels E, Hall AR. 1994. Generalized predictive tests and structural change in econometrics. *International Economic Review* **35**: 199–229.
- Hahn J, Inoue A. 2002. A Monte Carlo comparison of various asymptotic approximations to the distribution of instrumental variables estimators. *Econometric Reviews* **21**: 309–336.
- Hall AR. 2005. *Generalized Method of Moments*. Oxford, UK: Oxford University Press.
- Hall AR, Han S, Boldea O. 2012. Inference regarding multiple structural changes in linear models with endogenous regressors. *Journal of Econometrics* **170**: 281–302.
- Hall AR, Osborn DR, Sakkas N. 2013a. Inference on structural breaks using information criteria. *Manchester School Paper* **81**: Supplement S3, 54–81.
- Hall AR, Osborn DR, Sakkas N. 2013b. ‘The asymptotic expectation of the residual sum of squares in linear models with multiple break points,’ Discussion paper, Department of Economics, University of Manchester, Manchester, UK.
- Hannan EJ, Quinn BG. 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society* **41**: 190–195.
- Liu J, Wu S, Zidek JV. 1997. On segmented multivariate regression. *Statistica Sinica* **7**: 497–525.
- Ng S, Perron P. 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* **69**: 1519–1554.
- Ninomiya Y. 2005. Information criterion for Gaussian change-point model. *Statistics & Probability Letters* **72**: 237–247.
- Pesaran MH, Timmermann A. 1995. Predictability of stock returns: robustness and economic significance. *Journal of Finance* **50**: 1201–1228.
- Schwarz GE. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**: 461–464.

- Shibata R. 1976. Selection of the order of an autoregressive model by Akaike's criterion. *Journal of the Royal Statistical Society B* **63**: 117–126.
- Wooldridge J, White H. 1988. Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* **4**: 210–230.
- Yao YC. 1988. Estimating the number of change-points via Schwarz' criterion. *Statistics & Probability Letters* **6**: 181–9.
- Zhang NR, Siegmund DO. 2007. A modified Bayes information criterion with applications to the analysis of the comparative genomic hybridization data. *Biometrics* **63**: 22–33.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.