Check for updates

Journal of Time Series Analysis

WILEY

SPECIAL ISSUE ARTICLE OPEN ACCESS

Moving Sum Procedure for Multiple Change Point Detection in Large Factor Models

Matteo Barigozzi[1] | Haeran Cho[2] | Lorenzo Trapani[3,4]

[1]Department of Economics, Università di Bologna, Bologna, Italy | [2]School of Mathematics, University of Bristol, Bristol, UK | [3]Department of Economics and Management, Universita' di Pavia, Pavia, Italy | [4]School of Business, University of Leicester, Leicester, UK

Correspondence: Haeran Cho (haeran.cho@bristol.ac.uk)

Received: 3 October 2024 | **Revised:** 23 July 2025 | **Accepted:** 19 September 2025

Funding: This work was supported by the Engineering and Physical Sciences Research Council (Grant No. EP/Z531327/1).

Keywords: data segmentation | factor model | MOving SUMs (MOSUM)

ABSTRACT

This paper proposes a moving sum methodology for detecting multiple change points in high-dimensional time series under a factor model, where changes are attributed to those in loadings as well as emergence or disappearance of factors. We establish the asymptotic null distribution of the proposed test for family-wise error control and show the consistency of the procedure for multiple change point estimation. Simulation studies and an application to a large dataset of volatilities demonstrate the competitive performance of the proposed method.

MSC2020 Classification: 62M10, 62H25

1 | Introduction

Factor models are, arguably, one of the most frequently employed tools to model and carry out inference when large-dimensional, vector-valued time series are available. Whilst a comprehensive review of the literature goes well above and beyond the scope of this paper, it is worth noting that factor models have a long history, dating back at least to the seminal paper by Spearman (1904). Since being popularised by the contribution by Chamberlain and Rothschild (1983), and since the development of the asymptotic theory to analyse large-dimensional factor models (Bai 2003), their usage has become *de rigueur* in social sciences and economics where they have been applied to diverse fields such as business cycle analysis, asset pricing and economic monitoring and forecasting (see, *inter alia*, the review by Stock and Watson (2011) for a comprehensive list of references).

As the time dimension increases, it is inevitable that any model (including factor models) may undergo changes in its structure, which, in turn, may affect the properties of estimation techniques and hamper predictive ability. There is a huge literature on both ex-post and online detection of change points in various contexts, and we refer to the recent book by Horváth and Rice (2024) for a review focussed in particular on ex-post detection, and to the paper by Aue and Kirch (2024) for a review of the literature on online (sequential) detection.

Where factor models are concerned, on the one hand, the analysis in Stock and Watson (2009) and Bates et al. (2013) shows that, when changes are "sufficiently small" and do not involve a change in the number of common factors, the presence of change points is inconsequential for inference on the factor spaces. On the other hand, in several applications, there is no guarantee that the factor structure only undergoes small changes and that the number of common factors is time-invariant. Hence, it is not surprising that the literature has developed several methods to test, retrospectively, whether there is a change point or not. Breitung and Eickmeier (2011), Corradi and Swanson (2014), Han and

[This article is part of the special issue "Data Segmentation"].

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Time Series Analysis* published by John Wiley & Sons Ltd.

Inoue (2015), Yamamoto and Tanaka (2015), Baltagi et al. (2017), Cheng et al. (2016) and Bai et al. (2024) propose several tests to check if there is a break, based on the idea that a change point in the loadings is observationally equivalent to a change in the covariance matrix of the common factors. Thus, assuming homoscedastic factors, a change in the covariance matrix of the estimated common factors, can only be due to a break in the loadings. See also Barigozzi and Trapani (2020) and He et al. (2024) for an alternative formulation of change point tests in an online context. Whilst the literature on change point testing is well-established, there is relatively little work on detection and estimation of (possibly) multiple change points in the factor structure with some exceptions (Barigozzi et al. 2018; Li et al. 2023; Bai et al. 2024).

In this paper, we make an advance on the change point literature by proposing a method to estimate the number and locations of (possibly) multiple change points under a factor model. Specifically, we propose a procedure based on the so-called MOving SUMs (MOSUM) process, which, since the seminal contribution by Hušková and Slabý (2001), has been shown to have desirable properties in multiple change point estimation under weak assumptions on the distribution of the data. Versatility of the MOSUM procedure has been demonstrated in the context of detecting changes in the mean (Hušková and Slabý 2001; Eichinger and Kirch 2018), trend (Kim et al. 2024) and the drift of stochastic processes (Kirch and Klein 2024); see also Kirch and Reckruehm (2024) for a general framework for multivariate time series segmentation based on an estimating function. Some recent contributions extend the use of this methodology to high-dimensional time series (Cho and Owens 2024; Cho et al. 2024), demonstrating its scalability and suggesting that it is worth exploring the performance of a MOSUM procedure in the context of large factor models.

We build on that, as discussed above, a change in the loadings is observationally equivalent to a change in the covariance matrix of the common factors, and construct a MOSUM statistic based on the (moving) partial sums of the outer products of the estimated factors. In this respect, our statistic is akin to the one based on the (maximally selected) sequence of likelihood ratio tests proposed, for example, in Duan et al. (2023), save for the fact that it is based on using *moving*, as opposed to *cumulative*, partial sums. Upon testing for changes in the factor model using the MOSUM process, if the null hypothesis of no change is rejected, we perform multiple change point detection by estimating both the total number and the locations of the breaks. We establish the asymptotic null distribution of the MOSUM process as well as the consistency of the proposed change point detection procedure in estimating the total number and locations of the change points, with the accompanying rate of estimation for individual breaks. In doing so, we derive a Strong Invariance Principle for the MOSUM process based on the outer products of the estimated factors, which may be of independent interest.

The remainder of the paper is organised as follows. We describe the MOSUM procedure in Section 2. Assumptions, the limiting distribution under the null, and the consistency under the alternative are reported in Section 3. Sections 4 and 5 provide findings from simulation studies and an application to financial data. Section 6 concludes the paper, and all the proofs of theoretical results are given in the Appendix. An implementation of the proposed method is available at https://github.com/haeran-cho/mosum.fts.

Notations. For a random variable X, we have $|X|_\nu = (\mathsf{E}(|X|^\nu))^{1/\nu}$. We denote: weak convergence on the space $D[0,1]$ with $\xrightarrow{D}$; $\overset{D}{=}$ is equality in distribution; $\to$ is the ordinary limit; $\Gamma(\cdot)$ is the Gamma function; $\log(x)$ is the natural logarithm of $x > 0$; $\mathbb{I}_{\mathcal{A}}$ is an indicator function satisfying $\mathbb{I}_{\mathcal{A}} = 1$ if the event $\mathcal{A}$ is true and $\mathbb{I}_{\mathcal{A}} = 0$ otherwise. By $\mathbf{I}$, $\mathbf{O}$ and $\mathbf{0}$, we denote an identity matrix, a matrix of zeros and a vector of zeros whose dimensions depend on the context. For a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we denote by $\mathbf{A}^\top$ its transpose and $\|\mathbf{A}\|$ its Euclidean norm, with $\Lambda_{\max}(\mathbf{A})$ and $\Lambda_{\min}(\mathbf{A})$ denoting its largest and the smallest eigenvalues in modulus. When $m = n$, we denote by $\mathsf{Vech}(\mathbf{A})$ the vector of length $m(m+1)/2$ that stacks the elements on and below the main diagonal of $\mathbf{A}$, and $\mathbf{L}_m$ the $m(m+1)/2 \times m^2$-matrix satisfying $\mathsf{Vech}(\mathbf{A}) = \mathbf{L}_m\mathsf{Vec}(\mathbf{A})$, where $\mathsf{Vec}(\cdot)$ is an operator that stacks the columns of the matrix into a vector. Conversely, we define $\mathbf{K}_m \in \mathbb{R}^{m^2\times m(m+1)/2}$ satisfying $\mathsf{Vec}(\mathbf{A}) = \mathbf{K}_m\mathsf{Vech}(\mathbf{A})$, see Appendix A.12 of Lütkepohl (2005). Finally, given two sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n = O(b_n)$ if, for some finite positive constant C there exists $N \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ such that $|a_n||b_n|^{-1} \le C$ for all $n \ge N$.

2 | MOSUM Procedure for Data Segmentation

2.1 | Model

Let $\{\mathbf{X}_t,\ 1 \le t \le T\}$ denote an N-dimensional time series that admits the following factor model representation with R change points, as

$$\mathbf{X}_t = \begin{cases} \mathbf{\Lambda}_0\mathbf{f}_t + \mathbf{e}_t & \text{for } k_0 + 1 = 1 \le t \le k_1, \\ \mathbf{\Lambda}_1\mathbf{f}_t + \mathbf{e}_t & \text{for } k_1 + 1 \le t \le k_2, \\ \vdots & \\ \mathbf{\Lambda}_R\mathbf{f}_t + \mathbf{e}_t & \text{for } k_R + 1 \le t \le k_{R+1} = T. \end{cases} \tag{1}$$

Here, the matrix of loadings $\mathbf{\Lambda}_j = [\lambda_{j,1}, \ldots, \lambda_{j,N}]^\top \in \mathbb{R}^{N\times r}$ has column rank $r_j \le r$ fixed for all N, and "loads" the vector of random factors $\mathbf{f}_t = (f_{1,t}, \ldots, f_{r,t})^\top$ onto the cross-sections of $\mathbf{X}_t$, and $\mathbf{e}_t = (e_{1,t}, \ldots, e_{N,t})^\top$ denotes the idiosyncratic component. The model in (1) allows for the changes due to emergence or disappearance of factor(s) by permitting the ranks of $\mathbf{\Lambda}_j$ to vary, as well as rotational changes in the loading matrices. At the same time, the model is not identifiable in that the changes in the loadings are (observationally) equivalent to changes in the second-order properties of the common factor $\mathbf{f}_t$ – a fact, as mentioned in Introduction, frequently explored in the relevant literature for developing change point tests; see for example, Bai et al. (2024). We further illustrate this point in the following example.

Example 1. Consider the factor model in (1) with $R = 1$ and the single change point at $k_1 = k^*$. For any $\mathbf{\Lambda}_j \in \mathbb{R}^{N\times r_j}$, $j = 0, 1$, we can find $\mathbf{\Lambda} \in \mathbb{R}^{N\times r}$ of full column rank with $r \ge \max(r_0, r_1)$ such that $\mathbf{\Lambda}_j = \mathbf{\Lambda}\mathbf{A}_j$ for $\mathbf{A}_j \in \mathbb{R}^{r\times r_j}$ of rank r_j. Then, we can re-write (1) as

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\begin{bmatrix}\mathbf{X}_{0:k^*}\\ \mathbf{X}_{k^*:T}\end{bmatrix} = \begin{bmatrix}\mathbf{F}_{0:k^*}\mathbf{\Lambda}_0^\top\\ \mathbf{F}_{k^*:T}\mathbf{\Lambda}_1^\top\end{bmatrix} + \mathbf{E}, \text{ where } \mathbf{E}^\top = \begin{bmatrix}\mathbf{e}_1, \ldots, \mathbf{e}_T\end{bmatrix},$$

$$\underset{(b-a)\times N}{\mathbf{X}_{a:b}} = \begin{bmatrix}\mathbf{X}_{a+1}^\top\\ \vdots\\ \mathbf{X}_b^\top\end{bmatrix} \text{ and } \underset{(b-a)\times r}{\mathbf{F}_{a:b}} = \begin{bmatrix}\mathbf{f}_{a+1}^\top\\ \vdots\\ \mathbf{f}_b^\top\end{bmatrix} \text{ for all } 0 \le a < b \le T.$$

It follows that

$$\begin{bmatrix}\mathbf{X}_{0:k^*}\\ \mathbf{X}_{k^*:T}\end{bmatrix} = \begin{bmatrix}\mathbf{F}_{0:k^*}\mathbf{\Lambda}_0^\top\\ \mathbf{F}_{k^*:T}\mathbf{\Lambda}_1^\top\end{bmatrix} + \mathbf{E} = \begin{bmatrix}\mathbf{F}_{0:k^*}\mathbf{A}_0^\top\\ \mathbf{F}_{k^*:T}\mathbf{A}_1^\top\end{bmatrix}\mathbf{\Lambda}^\top + \mathbf{E} =: \mathbf{G}\mathbf{\Lambda}^\top + \mathbf{E},$$

which is an observationally equivalent representation with a time-invariant loading matrix $\mathbf{\Lambda}$ and pseudo factors of dimension r contained in $\mathbf{G}$.

We extend the observation made in Example 1 to the multiple change point setting and re-write the model in (1) as

$$\mathbf{X}_t = \mathbf{\Lambda}\sum_{j=0}^{R}\mathbf{A}_j\mathbf{f}_t \cdot \mathbb{I}_{\{k_j < t \le k_{j+1}\}} + \mathbf{e}_t =: \mathbf{\Lambda}\mathbf{g}_t + \mathbf{e}_t \tag{2}$$

with $\mathbf{\Lambda}_j = \mathbf{\Lambda}\mathbf{A}_j$ for all $0 \le j \le R$. Under the homoscedasticity of $\{\mathbf{f}_t\}$, we have

$$\mathsf{Cov}(\mathbf{g}_{k_j}) = \mathbf{A}_{j-1}\mathsf{Cov}(\mathbf{f}_0)\mathbf{A}_{j-1}^\top \neq \mathbf{A}_j\mathsf{Cov}(\mathbf{f}_0)\mathbf{A}_j^\top = \mathsf{Cov}(\mathbf{g}_{k_j+1})$$

for $1 \le j \le R$. Then, the problem of detecting the multiple change points in the loadings under (2), becomes that of detecting change points in the covariance of the pseudo factors $\{\mathbf{g}_t\}$.

2.2 | Methodology

Let us suppose that the number of pseudo-factors r is known. Then, we can estimate the pseudo factors $\mathbf{g}_t$ as, up to an invertible transformation, $\sqrt{T}$ times the r leading eigenvectors of the $T \times T$ matrix $(NT)^{-1}\mathbf{X}\mathbf{X}^\top$ with $\mathbf{X} = \mathbf{X}_{0:T}$. Denoting such an estimator by $\widehat{\mathbf{g}}_t$, we propose to test for a change under the model in (1) and, if any, estimate the multiple change points, by scanning the data using the following MOSUM statistic:

$$\mathcal{T}_{N,T,\gamma}(k) = \left|\mathbf{M}_{N,T,\gamma}^\top(k)\mathbf{V}_k^{-1}\mathbf{M}_{N,T,\gamma}(k)\right|^{1/2} \text{ for } \gamma \le k \le T-\gamma, \text{ with}$$

$$\mathbf{M}_{N,T,\gamma}(k) = \frac{1}{\sqrt{2\gamma}}\mathsf{Vech}\left(\sum_{t=k+1}^{k+\gamma}\widehat{\mathbf{g}}_t\widehat{\mathbf{g}}_t^\top - \sum_{t=k-\gamma+1}^{k}\widehat{\mathbf{g}}_t\widehat{\mathbf{g}}_t^\top\right), \tag{3}$$

where $\gamma \ge 1$ is a pre-selected bandwidth. The matrix $\mathbf{V}_k$ denotes the long-run covariance matrix of $\mathbf{M}_{N,T,\gamma}(k)$ which, in the absence of any change point, satisfies $\mathbf{V}_k \equiv \mathbf{V}$ for all k, with $\mathbf{V}$ defined explicitly in Theorem 1 below.

For testing the null hypothesis of no change point, $\mathcal{H}_0 : R = 0$, we compare the maximum of the MOSUM statistics against some threshold, say $D_{T,\gamma}$, and reject $\mathcal{H}_0$ if

$$\mathcal{T}_{N,T,\gamma} = \max_{\gamma \le k \le T-\gamma}\mathcal{T}_{N,T,\gamma}(k) > D_{T,\gamma}.$$

In Theorem 1 below, we derive the asymptotic null distribution of $\mathcal{T}_{N,T,\gamma}$, which enables selecting $D_{T,\gamma}$ as its upper α-quantile at a given significance level $\alpha \in (0,1)$.

Beyond testing for any change, we propose to detect and locate the multiple change points by adopting an approach proposed by Eichinger and Kirch (2018) in the univariate mean change point setting. Simply put, we select every local maximiser of $\mathcal{T}_{N,T,\gamma}(k)$ over a sufficiently large enough interval at which $\mathcal{T}_{N,T,\gamma}(k)$ exceeds the threshold. Specifically, for some fixed $\eta \in (0,1]$, we set as a change point estimator every $\widehat{k}$ that simultaneously satisfies

$$\widehat{k} = \underset{k:\ |\widehat{k}-k| \le \eta\gamma}{\arg\max}\ \mathcal{T}_{N,T,\gamma}(k) \text{ and } \mathcal{T}_{N,T,\gamma}(\widehat{k}) > D_{T,\gamma}. \tag{4}$$

Denoting such estimators by $\widehat{k}_j,\ 1 \le j \le \widehat{R}$, their total number $\widehat{R}$ is the estimator of the total number of change points R.

For the implementation of the MOSUM procedure, we require an estimator of the long-run covariance matrix $\mathbf{V}_k$. Whilst we allow it to be location-dependent to account for the heteroscedasticity in the presence of change points, estimating a long-run covariance matrix of multivariate time series is well-known to be highly challenging. In the current setting, this is augmented by that the computation of $\mathcal{T}_{N,T,\gamma}(k)$ calls for the inverse of $\mathbf{V}_k$, which may bring further numerical instabilities. Therefore, we opt to use the following HAC-type estimator in place of $\mathbf{V}_k$,

$$\widehat{\mathbf{V}} = \widehat{\mathbf{\Gamma}}(0) + \sum_{\ell=1}^{m}\left(1 - \frac{\ell}{m+1}\right)\left(\widehat{\mathbf{\Gamma}}(\ell) + \widehat{\mathbf{\Gamma}}^\top(\ell)\right), \text{ where}$$

$$\widehat{\mathbf{\Gamma}}(\ell) = \frac{1}{T}\sum_{t=\ell+1}^{T}\mathsf{Vech}\left(\widehat{\mathbf{g}}_t\widehat{\mathbf{g}}_t^\top - \mathbf{I}_r\right)\mathsf{Vech}\left(\widehat{\mathbf{g}}_{t-\ell}\widehat{\mathbf{g}}_{t-\ell}^\top - \mathbf{I}_r\right)^\top \tag{5}$$

with some bandwidth $m \ge 1$. We later show that the matrix $\widehat{\mathbf{V}}$ provides a consistent estimator of the long-run covariance matrix of $\mathbf{M}_{N,T,\gamma}(k)$ under $\mathcal{H}_0$ (see Proposition 2). The Bartlett kernel in (5) is only one of the many possible choices, and we refer to Bai et al. (2024) for a comprehensive analysis of various kernel-based estimators of $\mathbf{V}$. For consistency of multiple change point detection, we only require that a positive definite matrix is used in place of $\mathbf{V}_k$ (Theorem 3). By default, we propose to adopt $\widehat{\mathbf{V}}$ (or its diagonal entries), which is shown to work well; see Section 4 for further discussions.

3 | Theoretical Properties

3.1 | Assumptions

We begin by providing a definition of $\mathcal{L}_\nu$*-decomposable Bernoulli shifts*.

Definition 1. The d-dimensional sequence $\{\mathbf{m}_t, -\infty < t < \infty\}$ forms an L_ν-decomposable Bernoulli shift if and only if it holds that $\mathbf{m}_t = h(\boldsymbol{\eta}_t, \boldsymbol{\eta}_{t-1}, \ldots)$, where (i) $h : \mathbb{S}^\infty \to \mathbb{R}^d$ is a non-random measurable function; (ii) $\{\boldsymbol{\eta}_t\}_{t\in\mathbb{Z}}$ is an i.i.d. sequence with values in a measurable space $\mathbb{S}$; (iii) $\mathsf{E}(\mathbf{m}_t) = 0$ and $|||\mathbf{m}_t|||_\nu < \infty$; and (iv) $|||\mathbf{m}_t - \mathbf{m}_{t,l}^*|||_\nu \le c_0 l^{-a}$ for some $c_0 > 0$ and $a > 0$, where $\mathbf{m}_{t,l}^* = h(\boldsymbol{\eta}_t, \ldots, \boldsymbol{\eta}_{t-l+1}, \boldsymbol{\eta}_{t-l,t,l}^*, \boldsymbol{\eta}_{t-l-1,t,l}^*, \ldots)$,

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

with $\{\boldsymbol{\eta}^*_{s,t,l},\ -\infty < s,l,t < \infty\}$ that are i.i.d. copies of $\boldsymbol{\eta}_0$ which are independent of $\{\boldsymbol{\eta}_t\}_{t\in\mathbb{Z}}$.

Since the seminal works by Wu (2005) and Berkes et al. (2011) (see also Hörmann 2009.), decomposable Bernoulli shifts have proven a very convenient way to model and study dependent time series, mainly due to their generality, and since it is much easier to verify whether a sequence forms a decomposable Bernoulli shift than, for example, verifying mixing conditions. Virtually all of the most commonly employed models in econometrics and statistics can be shown to generate decomposable Bernoulli shifts, such as ARMA and (G)ARCH processes, non-linear time series models (e.g., random coefficient autoregressive models and threshold models), Volterra series and data generated by dynamical systems; see Berkes et al. (2011), Aue et al. (2009) and Liu and Lin (2009).

We establish the theoretical properties of the proposed MOSUM procedure under the following assumptions.

Assumption 1.

i. There exists some fixed $\epsilon \in (0,1)$ such that $\{\mathbf{f}_t\}_{t\in\mathbb{Z}}$ is an $\mathcal{L}_\nu$-decomposable Bernoulli shift with $\nu \geq 8\rho+\epsilon$ for $\rho = 1$ or $\rho = 2$, and $a > 2$.

ii. $\boldsymbol{\Sigma}_F = \mathsf{E}(\mathbf{f}_t\mathbf{f}_t^\top) \in \mathbb{R}^{r\times r}$ is positive definite.

iii. Denoting the long-run variance matrix of $\mathcal{F}_t = \mathsf{Vech}(\mathbf{f}_t\mathbf{f}_t^\top)$ by

$$\mathbf{D} = \lim_{T\to\infty}\frac{1}{T}\mathsf{E}\left\{\left[\sum_{t=1}^{T}(\mathcal{F}_t - \mathsf{E}(\mathcal{F}_t))\right]\left[\sum_{t=1}^{T}(\mathcal{F}_t - \mathsf{E}(\mathcal{F}_t))\right]^\top\right\}, \tag{6}$$

we suppose that $\mathbf{D}$ is invertible.

Assumption 2. There exists some $c_0 \in (0,\infty)$ such that:

i. λ_i is deterministic with $\|\lambda_i\| \leq c_0$ for all $1 \leq i \leq N$ and $N \in \mathbb{N}$.

ii. $\|N^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} - \boldsymbol{\Sigma}_\Lambda\| \leq c_0 N^{-1/2}$ for all $N \in \mathbb{N}$, where $\boldsymbol{\Sigma}_\Lambda \in \mathbb{R}^{r\times r}$ is positive definite.

Assumption 3. There exist some $\epsilon \in (0,1)$ and $c_0 \in (0,\infty)$ such that the following holds for all $N, T \in \mathbb{N}$ and $\rho = 1$ or $\rho = 2$:

i. $\mathsf{E}(e_{it}) = 0$ and $\mathsf{E}(|e_{i,t}|^{8+\epsilon}) < \infty$ for all $1 \leq i \leq N$ and $1 \leq t \leq T$.

ii. Letting $\gamma_{s,t} = N^{-1}\sum_{i=1}^{N}\mathsf{E}(e_{i,s}e_{i,t})$, it holds that $\sum_{t=1}^{T}|\gamma_{s,t}| \leq c_0$ for all $1 \leq s \leq T$.

iii. $\mathsf{E}(|\sum_{i=1}^{N}(e_{i,t}e_{i,s} - \gamma_{s,t})|^{4+\epsilon}) \leq c_0 N^{2+\epsilon/2}$ for all $1 \leq s,t \leq T$.

iv. $\mathsf{E}(\|\sum_{i=1}^{N}\lambda_i e_{i,t}\|^{8+\epsilon}) \leq c_0 N^{4+\epsilon/2}$ for all $1 \leq t \leq T$.

v. $\sum_{j=1}^{N}\left|\mathsf{E}(e_{i,t}e_{j,t})\right| \leq c_0$ for all $1 \leq i \leq N$ and $1 \leq t \leq T$.

vi. $\mathsf{E}(|\sum_{t=1}^{T}(e_{i,t}e_{j,t} - \mathsf{E}(e_{i,t}e_{j,t}))|^{4\rho+\epsilon}) \leq c_0 T^{2\rho+\epsilon/2}$ for all $1 \leq i,j \leq N$.

Assumption 4. There exist some $\epsilon \in (0,1)$ and $c_0 \in (0,\infty)$ such that the following holds for all $N, T \in \mathbb{N}$ and $\rho = 1$ or $\rho = 2$:

i. $\mathsf{E}(\|\sum_{t=a+1}^{b}\mathbf{g}_t\sum_{i=1}^{N}(e_{i,t}e_{i,s} - \gamma_{s,t})\|^{2\rho+\epsilon}) \leq c_0(N(b-a))^{\rho+\epsilon/2}$ for all $1 \leq s \leq T$ and $0 \leq a < b \leq T$.

ii. $\mathsf{E}(|\sum_{i=1}^{N}\sum_{t=a+1}^{b}\lambda_i^\top\mathbf{g}_t e_{i,t}|^{4\rho+\epsilon}) \leq c_0(N(b-a))^{2\rho+\epsilon/2}$ for all $0 \leq a < b \leq T$.

iii. $\mathsf{E}(|\sum_{t=1}^{T}\lambda_i^\top\mathbf{g}_t e_{j,t}|^{4\rho+\epsilon}) \leq c_0 T^{2\rho+\epsilon/2}$ for all $1 \leq i,j \leq N$.

Assumptions 1–4 are closely related to those adopted in the factor model literature, see also Bai (2003). Assumption 2 is the same as Assumptions B in Bai (2003). In Assumption 1(i), we strengthen the moment condition typically employed in the literature on $\mathbf{f}_t$, switching from the existence of the 4-th moment to that of the 8th or higher moment. Also, we assume a specific form of dependence for $\{\mathbf{f}_t, -\infty < t < \infty\}$ which, as mentioned above, accommodates a wide range of time series models. This is required to derive a Strong Invariance Principle (SIP) for $\mathsf{Vech}(\mathbf{f}_t\mathbf{f}_t^\top)$, see Lemma A.9 in Appendix A.1. Assumption 3 allows for weak temporal and cross-sectional dependence in the idiosyncratic component, with similarities between (ii) and Assumption E1 in Bai (2003), (iii) and C5, and (v) and E2; part (iv) strengthens their F3 and also Assumption 6ii of Bai et al. (2024); part (vi) can be derived under more primitive conditions on $e_{i,t}$. Assumption 4(i,ii) extend Assumptions F1 and F2 of Bai (2003), to account for the scanning for multiple change points performed by the MOSUM procedure.

Generally, the strengthened conditions found in Assumptions 1(i) and 3 on the moments of $\mathbf{f}_t$ and $e_{i,t}$, are required as we go a step further from the typical factor modelling literature that focuses on establishing the consistency of the estimated factors, to control the partial sums involved in the MOSUM process. We note that $\rho = 1$ in Assumptions 1, 3 and 4, is sufficient for deriving the asymptotic null distribution of the MOSUM test statistic (Theorem 1) as well as the detection consistency of the MOSUM procedure (Theorem 3, 3), whilst $\rho = 2$ is required for establishing the rate of estimation for the change points (Theorem 3, 3).

Assumption 5.

i. There exist $\tau_j,\ 1 \leq j \leq R$, satisfying $0 < \tau_1 < \cdots < \tau_R < 1$, such that $k_j = \lfloor \tau_j T \rfloor$.

ii. $\|\mathbf{A}_j\| \leq c_0 \in (0,\infty)$ for all $0 \leq j \leq R$.

iii. Denoting by $\boldsymbol{\Sigma}_G = T^{-1}\sum_{t=1}^{T}\mathsf{E}(\mathbf{g}_t\mathbf{g}_t^\top)$, the eigenvalues of $\boldsymbol{\Sigma}_G\boldsymbol{\Sigma}_\Lambda$ are positive and distinct.

Assumption 6. There exists some $\epsilon_\circ \in (0,\infty)$ such that

$$\lim_{\min(N,T)\to\infty}\frac{T^{1/2+\epsilon_\circ}}{N} = 0.$$

When $R = 0$, Assumption 5 only requires that $\boldsymbol{\Sigma}_F\boldsymbol{\Sigma}_\Lambda$ has distinct eigenvalues, paralleling the commonly found condition such as Assumption G of Bai (2003). When $R \geq 1$, part (i) assumes that the change points are linearly spaced. The positive definiteness imposed on $\boldsymbol{\Sigma}_G$ in part (iii), together with part (i) and Assumption 1(ii), implies that any local factors are pervasive over

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

segment(s) where they are present. Finally, Assumption 6 is also found in Bai et al. (2024), and arises from that we construct the MOSUM process based on an estimate of the latent factors.

3.2 | Asymptotic Null Distribution

We present the limiting distribution of the maximally selected MOSUM statistic in (3). We write, for simplicity, $d = r(r+1)/2$ and denote by $\beta = \log(N)/\log(T)$; under Assumption 6, we have $1/2 + \epsilon_\circ < \beta$ with some $\epsilon_\circ > 0$. Define also

$$\zeta = \max\left\{\frac{2}{\nu}, 1 - \min(1,\beta)\right\}, \tag{7}$$

where ν is defined in Assumption 1. Then, we always have $\zeta \in (0, 1/2)$.

Theorem 1. *Suppose that Assumption 1–6 hold with $\rho = 1$ for Assumptions 1, 3 and 4, and let the bandwidth γ satisfy*

$$\frac{T^{2\zeta}\log(T/\gamma)}{\gamma} \to 0 \text{ and } \frac{\gamma}{T} \to 0. \tag{8}$$

Let us define $\mathbf{V} = \mathbf{L}_r(\mathbf{H}_0^\top \otimes \mathbf{H}_0^\top)\mathbf{K}_r\mathbf{D}\mathbf{K}_r^\top(\mathbf{H}_0 \otimes \mathbf{H}_0)\mathbf{L}_r^\top$, with $\mathbf{D}$ in (6), and

$$\mathbf{H}_0 = \operatorname*{plim}_{\min(N,T)\to\infty} \frac{1}{NT}(\mathbf{\Lambda}^\top\mathbf{\Lambda})(\mathbf{G}^\top\widehat{\mathbf{G}})\mathbf{\Phi}_{NT}^{-1} \tag{9}$$

where we denote with $\mathbf{\Phi}_{NT} \in \mathbb{R}^{r\times r}$ the diagonal matrix containing the r largest eigenvalues of $(NT)^{-1}\mathbf{X}\mathbf{X}^\top$ on its diagonal.

a. *Under $\mathcal{H}_0$: $R = 0$, for all $x \in \mathbb{R}$, we have*

$$\begin{aligned}\lim_{\min(N,T)\to\infty} &\mathsf{P}\left(a\left(\frac{T}{\gamma}\right)\max_{\gamma\le k\le T-\gamma} \right.\\ &\left.\times\left|\mathbf{M}_{N,T,\gamma}^\top(k)\mathbf{V}^{-1}\mathbf{M}_{N,T,\gamma}(k)\right|^{1/2} - b_d\left(\frac{T}{\gamma}\right) \le x\right)\\ &= \exp(-2\exp(-x)),\end{aligned} \tag{10}$$

where $a(x) = \sqrt{2\log(x)}$ and $b_d(x) = 2\log(x) + d\log\log(x)/2 + \log(1/2) - \log(\Gamma(d/2))$.

b. *The assertion in (a) continues to hold if $\mathbf{V}$ is replaced by a positive definite matrix $\widehat{\mathbf{V}}$ satisfying*

$$\left\|\widehat{\mathbf{V}} - \mathbf{V}\right\| = o_P\left(\log^{-1}(T/\gamma)\right). \tag{11}$$

The limiting law in Theorem 1 is analogous to those derived in Hušková and Slabý (2001) and Eichinger and Kirch (2018), modulo the fact that here, we deal with d-variate vectors and therefore the function $b_d(\cdot)$ depends on d. In contrast with a "standard" multivariate time series application, however, in our result, the cross-sectional dimension N also plays a role through the definition of ζ in (7), which enters in the condition (8) made on the bandwidth γ. As a by-product, we establish the SIP of the process $\mathsf{Vech}(\widehat{\mathbf{g}}_t\widehat{\mathbf{g}}_t^\top)$ after an appropriate centring (see Lemma A.11 in Appendix A.1), which proves crucial in deriving the asymptotic null distribution in (10).

Based on this limiting law of the maximally selected MOSUM process, we reject $\mathcal{H}_0$: $R = 0$ at the significance level $\alpha \in (0,1)$, if

$$\begin{aligned}\max_{\gamma\le k\le T-\gamma}\left|\mathbf{M}_{N,T,\gamma}^\top(k)\mathbf{V}^{-1}\mathbf{M}_{N,T,\gamma}(k)\right|^{1/2} &> \widetilde{D}_{T,\gamma}(\alpha)\\ := \frac{b_d(T/\gamma) - \log\log\left(\frac{1}{\sqrt{1-\alpha}}\right)}{a(T/\gamma)}&.\end{aligned} \tag{12}$$

The condition in (8) places both upper and lower bounds on the bandwidth γ. Specifically, γ is required to grow sufficiently faster than $T^{2\zeta}$ whilst satisfying $T^{-1}\gamma \to 0$, and the former restriction calls for larger γ when $\mathbf{f}_t$ has fewer moments or when N is small. We note that $\beta = \log(N)/\log(T)$ is known and does not need to be estimated. Conversely, ν is in general not known. We may select γ satisfying (8) by plugging in an estimate of ν, say $\widehat{\nu}$; alternatively, one can decide a value of ν, say ν^*, and test whether $|\mathbf{g}_t|_{\nu^*} < \infty$. In both cases, one difficulty is that $\mathbf{g}_t$ is not observable; however, deriving ν from the data $\mathbf{X}_t$ yields a lower bound.

Theorem 1(a) shows that when the unknown $\mathbf{V}$ is replaced by its estimator, the asymptotic null distribution continues to hold provided that (11) is met. The following Proposition 2 guarantees that this requirement is met by the estimator $\widehat{\mathbf{V}}$ proposed in (5), strengthening the observation made in Bai et al. (2024) that $\|\widehat{\mathbf{V}} - \mathbf{V}\| = o_P(1)$.

Proposition 2. *Suppose that Assumption 1–6 hold with $\rho = 1$ for Assumptions 1, 3 and 4. Also, let the bandwidth m satisfy*

$$\frac{\log(T/\gamma)}{m} \to 0 \text{ and } \frac{m\log(T/\gamma)}{\sqrt{\min(N,T)}} \to 0. \tag{13}$$

Then, as $\min(N,T) \to \infty$, the estimator $\widehat{\mathbf{V}}$ in (5) satisfies the condition in (11).

3.3 | Consistency in Multiple Change Point Estimation

To establish the consistency of the MOSUM procedure in multiple change points detection, we make the following assumption on the size of changes.

Assumption 7.

i. $\min_{0\le j\le R}\Delta_j \ge 2\gamma$, where $\Delta_j = k_{j+1} - k_j$.

ii. At each $1 \le j \le R$, let $\boldsymbol{\delta}_j = \mathbf{A}_j\mathbf{\Sigma}_F\mathbf{A}_j^\top - \mathbf{A}_{j-1}\mathbf{\Sigma}_F\mathbf{A}_{j-1}^\top$ and $d_j = \|\boldsymbol{\delta}_j\|$. Then for $\omega_T^{(1)} \to \infty$ arbitrarily slowly, it holds that

$$\frac{\omega_T^{(1)}\sqrt{\log(T/\gamma)}}{\min_{1\le j\le R} d_j\sqrt{\gamma}} = o(1).$$

Assumption 7(i) ensures that there exists at most a single change point over each moving window, and is implied jointly by Assumption 5(ii) and the condition (8) on γ. Condition 7 permits local changes with $d_j \to 0$, at a sufficiently slow rate, which is the case, for example, when the neighbouring loading matrices are rotations of one another with $\|\mathbf{A}_j - \mathbf{A}_{j-1}\| \to 0$.

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Theorem 3. *Suppose that Assumption 1–6 hold with $\rho = 1$ for Assumptions 1, 3 and 4. Additionally, let Assumption 7 hold and the bandwidth γ satisfy (8), and suppose that some positive definite matrix $\widetilde{\mathbf{V}}$ is used in place of $\mathbf{V}_k$ in $\mathcal{T}_{N,T,\gamma}(k)$, see (3).*

a. *For any $\alpha, \eta \in (0,1)$, there exists some sequence $\omega_T^{(1)} \to \infty$ arbitrarily slowly, such that the MOSUM procedure with $D_{T,\gamma} = \widetilde{D}_{T,\gamma}(\alpha) \cdot \omega_T^{(1)}$ as the threshold, returns $\{\widehat{k}_j,\ 1 \le j \le \widehat{R} : \widehat{k}_1 < \cdots < \widehat{k}_{\widehat{R}}\}$ which satisfies*

$$\mathrm{P}\left(\widehat{R} = R;\ \max_{1\le j\le R}|\widehat{k}_j - k_j| \le \eta\gamma\right) \to 1 \text{ as } \min(N,T) \to \infty.$$

b. *Further, if Assumptions 1, 3 and 4 hold with $\rho = 2$, there exists some sequence $\omega_T^{(2)} \to \infty$ arbitrarily slowly, such that*

$$\mathrm{P}\left(\widehat{R} = R;\ \max_{1\le j\le R} d_j^2|\widehat{k}_j - k_j| \le \omega_T^{(2)}\right) \to 1 \text{ as } \min(N,T) \to \infty.$$

Theorem 3 shows that the MOSUM procedure consistently estimates the total number and the locations of the change points. Here, we adopt a fixed, positive definite matrix $\widetilde{\mathbf{V}}$ in place of $\mathbf{V}_k$ which, without being a consistent estimator of the latter at some k, still leads to consistency in multiple change point detection. This flexibility in the choice of $\widetilde{\mathbf{V}}$ is particularly favourable since, as noted earlier in Section 2.2, the estimation of (time-varying) long-run covariance matrix for multivariate time series is challenging. Whilst the asymptotic null distribution in Theorem 1 allows for testing the null hypothesis of no change point with the family-wise error controlled, strengthening of the threshold is necessary for consistently detecting the number of change points, see for example, Eichinger and Kirch (2018) and Bai et al. (2024) where they set $\alpha = \alpha_T \to 0$ at a suitable rate. Instead, we introduce an additional multiplicative factor of $\omega_T^{(1)} \to \infty$ in the threshold $D_{T,\gamma}$; see Section 4.1 where we discuss the choice of the threshold. Under a stronger moment assumption, we obtain the rate of estimation which is inversely proportional to the squared size of change as $|\widehat{k}_j - k_j| = O_P(d_j^{-2})$, This indicates that dominant changes are located with better accuracy, such as those accompanied by a change in the dimension of the factor space due to the introduction or disappearance of factor(s).

4 | Numerical Experiments

4.1 | Tuning Parameter Selection

Empirical performance of the MOSUM procedure depends on the choice of tuning parameters.

Inspecting the proof of Theorem 1, we observe that the requirement on ν in Assumption 1(i) may be weakened if the r largest eigenvalues of $(NT)^{-1}\mathbf{X}\mathbf{X}^\top$ are bounded away from zero deterministically. Inspired by this and the condition in (8), we propose to select the bandwidth as $\gamma = \lfloor T^{2\zeta} \cdot \log^{\varrho}(T) \rfloor$ with $\zeta = \max(2/5, 1 - \log(N)/\log(T))$. Thus-selected bandwidth with $\varrho = 1.1$ works reasonably well in our simulation studies where datasets of dimensions $N \le 500$ and $T \le 1000$ are considered. For the real data application in Section 5 with $T \ge 4000$, we set $\varrho = 0.5$.

We set the threshold $D_{T,\gamma}$ as described in Theorem 3, namely $D_{T,\gamma} = \widetilde{D}_{T,\gamma}(\alpha) \cdot \omega_T^{(1)}$ where $\widetilde{D}_{T,\gamma}(\alpha)$ is given by (12) according to the asymptotic null distribution in Theorem 1. As for $\omega_T^{(1)}$, we have considered $\log^{\kappa}(T/\gamma)$ with $\kappa \in \{0, 0.1, 0.2, 0.3\}$, and observed that the choice of $\kappa = 0.2$ returned stably good performance in all experiments, see Appendix B.2 for full detail. Compared to the choice of $\omega_T^{(1)}$, the selection of the fixed significance level $\alpha \in (0,1)$ has relatively little influence and in all our studies, we set $\alpha = 0.05$. Finally, for the detection rule in (4), we set $\eta = 0.6$.

For estimating the number of pseudo factors r, we apply the approach proposed by Alessi et al. (2010) in combination with the three information criteria of Bai and Ng (2002): Addressing the arbitrariness in the choice of the penalty, it looks for a stable estimate of the factor number as the minimiser of the information criterion over sub-samples of varying dimensions and sample sizes. We take the median of the estimates from the three information criteria if they do not agree. On simulated datasets, we find that this approach consistently identifies the correct number of factors over 90% of the realisations.

Finally, in place of $\mathbf{V}_k$ in (3), we plug in the estimator $\widehat{\mathbf{V}}$ in (5) with bandwidth $m = \lfloor T^{1/4} \rfloor$. Due to the presence of change points, the number of pseudo factors r can be large in which case inverting the $d \times d$-matrix $\widehat{\mathbf{V}}$ with $d = r(r+1)/2$, may bring numerical instability, see Kirch et al. (2015) for the alternative approaches to handling similar difficulties in multivariate time series segmentation. Therefore, we explore two approaches, one performing the standardisation using the full matrix $\widehat{\mathbf{V}}$, and the other using its diagonal entries only, respectively referred to as MOSUM-full and MOSUM-diagonal. We remark that MOSUM-diagonal meets the requirement in Theorem 3 provided that all diagonal entries of $\widehat{\mathbf{V}}$ are positive. Our numerical experiments indicate that MOSUM-diagonal is to be preferred between the two, see Section 4.3 for further discussions.

4.2 | Settings

We consider the following data generating processes considered in Li et al. (2023) and Duan et al. (2023).

M1. Adopted from Li et al. (2023), we fix $T = 400$, $N = 200$ and $r_0 = 5$, and introduce $R = 2$ change points at $(k_1, k_2) = (133, 267)$ as follows: $X_{it} = \chi_{it} + \sqrt{0.5}e_{it}$, where

$$\chi_{it} = \mathbf{\Lambda}_j \mathbf{f}_t \text{ with } \mathbf{f}_t \sim \mathcal{N}_{r_0}(\mathbf{0}, \mathbf{\Sigma}_j) \text{ for} k_{j-1} + 1 \le t \le k_j, \text{ and}$$

$$\mathbf{\Sigma}_0 = [\sigma_{0,ij}] = \mathbf{D}\mathbf{\Sigma}_F\mathbf{D} \text{ with}$$

$$\mathbf{D} = \mathrm{diag}(d_{ii},\ 1 \le i \le r_0),\ d_{ii} \sim_{\mathrm{iid}} \mathcal{U}[0.5, 1.5],$$

$$\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = [\sigma_{1,ij}] \text{ with } \sigma_{1,ij} = \sigma_{1,ji}$$

$$= \begin{cases} 0.9\sqrt{\sigma_{0,11}\sigma_{0,22}} & \text{for } (i,j) = (1,2), \\ 1.3^2\sigma_{0,55} & \text{for } (i,j) = (5,5), \\ 0.5^{|i-5|}\sqrt{\sigma_{0,ii}\sigma_{0,55}} & \text{for } 1 \le i \le 4, \\ \sigma_{0,ij} & \text{otherwise,} \end{cases}$$

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

with $\mathbf{\Sigma}_F = [0.5^{|i-j|},\ 1 \le i, j \le r_0]$. The loadings are generated as $\mathbf{\Lambda}_0 = \mathbf{\Lambda}_1 = [\lambda_{0,ij},\ 1 \le i \le p,\ 1 \le j \le r_0]$ with $\lambda_{0,ij} \sim_{\text{iid}} \mathcal{U}[-1, 1]$, and $\mathbf{\Lambda}_2 = [\lambda_{2,ij},\ 1 \le i \le p,\ 1 \le j \le r_0]$ with $\lambda_{2,ij} \sim_{\text{iid}} \mathcal{U}[-1, 1]$ for $j \le 2$, whilst $\lambda_{2,ij} = \lambda_{0,ij}$ for $j \ge 3$. Within each segment, the number of factors remains constant at $r_0 = 5$ whilst the overall factor number is $r = r_0 + 2$ due to the increase of factor space after k_2. The idiosyncratic component is generated as an independent Gaussian random vector whose covariance undergoes changes at $t = 100, 200$ and 300 (with the proportion of changes set at 0.1), which are not to be detected by the proposed MOSUM method.

M2. We join together three single change point scenarios from Duan et al. (2023) to form a multiple change point one: Setting $R = 3$ and $r_0 = 3$, we generate

$$\mathbf{f}_t = \rho_f \mathbf{f}_{t-1} + \boldsymbol{\varepsilon}_{f,t},\ \boldsymbol{\varepsilon}_{f,t} \sim_{\text{iid}} \mathcal{N}_{r_0}(\mathbf{0}, \mathbf{I}_{r_0}),$$

$$\mathbf{e}_t = \rho_e \mathbf{e}_{t-1} + \boldsymbol{\varepsilon}_{e,t},\ \boldsymbol{\varepsilon}_{e,t} \sim \mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma}_e) \text{ with}$$

$$\mathbf{\Sigma}_e = [(0.3)^{|i-j|},\ 1 \le i, j \le p],$$

and $\mathbf{\Lambda}_0 = [\lambda_{0,ij},\ 1 \le i \le p,\ 1 \le j \le r_0]$ with $\lambda_{0,ij} \sim_{\text{iid}} \mathcal{N}(0, 1/r_0)$. The change points are introduced at $k_j = Tj/4$, $1 \le j \le 3$, at each of which the loading matrix undergoes a shift to $\mathbf{\Lambda}_j = \mathbf{\Lambda}_0 \mathbf{C}_j$, where

$$\mathbf{C}_1 = \begin{bmatrix} 0.5 & 0 & 0 \\ c_{1,21} & 1 & 0 \\ c_{1,31} & c_{1,32} & 1.5 \end{bmatrix},$$

$$\mathbf{C}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{C}_3 = [c_{3,ij},\ 1 \le i, j \le r_0],$$

with $c_{1,ij} \sim_{\text{iid}} \mathcal{N}(0, 1)$ and $c_{3,ij} \sim_{\text{iid}} \mathcal{N}(0, 1/r_0)$. The factor number varies as $(r_0, \ldots, r_3) = (3, 3, 2, 3)$, whilst the number of pseudo factors increases from 3 to 6 due to the change at k_3. We vary $T \in \{400, 600, 800, 1000\}$ and $N \in \{100, 200, 500\}$ as well as $(\rho_f, \rho_e) = \{(0, 0), (0.7, 0.3)\}$.

M3. Additionally, we consider the "null" model with $R = 0$ by generating the data from the model corresponding to the first segment of 4.2 for each setting.

4.3 | Results

For each setting, we generate 200 realisations and report the distribution of $\widehat{R} - R$, and the accuracy of change point estimators measured by

$$\frac{1}{200} \sum_{i=1}^{200} \mathbb{I}\left[\min_{1 \le \ell \le \widehat{R}^{(i)}} |\widehat{k}_\ell^{(i)} - k_j| \le \log(T) \right]$$

for each $1 \le j \le R$, as proposed by Li et al. (2023), where $\widehat{k}_\ell^{(i)}$, $1 \le \ell \le \widehat{R}^{(i)}$, refer to the change point estimators from the i-th realisation. We apply the MOSUM procedure with the tuning parameters chosen as described in Section 4.1, and consider the two choices of the standardisation matrix (MOSUM-full and MOSUM-diagonal). Additionally, we include the two competitors:

i. Proposed by Li et al. (2023), BSCOV scans for changes in the covariance of $\{\mathbf{g}_t\}$ under (2) via an extension of the binary segmentation, produces a path of solutions and selects the final change point model by minimising an information criterion.

ii. Proposed by Bai et al. (2024), BDH recursively applies the likelihood ratio test via binary segmentation to detect the multiple change points under (2).

Both methods are applied with default tuning parameters and in-built factor number estimators that are based on the information criterion proposed by Bai and Ng (2002); for BDH, we set the proportion of the data trimmed off at each recursion to be 0.1. Tables 1–4 report the summary of the results over 200 realisations, and Figures 1–2 plot the histograms of change point estimators returned by the proposed MOSUM procedure, BSCOV and BDH under 4.2 when $N = 100$; see also Appendix B for the additional results.

Overall, we observe that the MOSUM procedure ('MOSUM') demonstrates competitive performance across all scenarios, both in detection and estimation. BSCOV tends to return spurious estimators under 4.2 and 4.2 when serial dependence is present with $(\rho_f, \rho_e) = (0.7, 0.3)$. On the other hand, BDH suffers from a lack of detection power against those changes that transform the loading matrix whilst not altering the number of local factors, such as k_1 under 4.2 and 4.2, particularly for smaller T. For such a change point, MOSUM is able to detect its presence, although with less accuracy. Generally, k_j's which are associated with changes in the

TABLE 1 | 4.2 with $R = 2$: Summary of change point estimators returned by MOSUM, BSCOV and BDH. The results for BSCOV have been taken from Li et al. (2023).

		$\widehat{R} - R$					Accuracy	
Method	**LRV**	**$-2 \le$**	**-1**	**0**	**1**	**≥ 2**	**$j = 1$**	**$j = 2$**
MOSUM	Diagonal	0	0	0.985	0.015	0	0.7	0.925
	Full	0	0.005	0.94	0.055	0	0.595	0.8
BSCOV	—	0	0.03	0.97	0	0	0.64	0.95
BDH	—	0	0.965	0.015	0.005	0.015	0	1

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | 4.2 with $(\rho_f, \rho_e) = (0, 0)$ and $R = 3$: Summary of change point estimators returned by MOSUM and BSCOV over 200 realisations.

				$\widehat{R} - R$					Accuracy		
n	p	Method	LRV	$-2 \leq$	-1	0	1	≥ 2	$j = 1$	$j = 2$	$j = 3$
400	100	MOSUM	Diagonal	0	0.01	0.99	0	0	0.82	0.88	0.985
			Full	0	0.025	0.975	0	0	0.75	0.895	0.93
		BSCOV	—	0.02	0	0.98	0	0	0.725	0.85	0.98
		BDH	—	0.02	0.935	0.045	0	0	0.045	0.98	1
	200	MOSUM	Diagonal	0	0	1	0	0	0.795	0.875	0.97
			Full	0	0.035	0.965	0	0	0.73	0.86	0.93
		BSCOV	—	0.01	0.005	0.985	0	0	0.79	0.855	0.99
		BDH	—	0	0.93	0.07	0	0	0.07	1	1
	500	MOSUM	Diagonal	0	0.02	0.98	0	0	0.8	0.88	0.975
			Full	0	0.03	0.97	0	0	0.74	0.895	0.95
		BSCOV	—	0.01	0	0.98	0.01	0	0.8	0.89	0.98
		BDH	—	0	0.885	0.115	0	0	0.115	1	1
600	100	MOSUM	Diagonal	0	0	1	0	0	0.83	0.88	0.99
			Full	0	0	0.995	0.005	0	0.75	0.87	0.945
		BSCOV	—	0	0	0.99	0.01	0	0.81	0.895	0.995
		BDH	—	0	0.7	0.3	0	0	0.3	1	1
	200	MOSUM	Diagonal	0	0	1	0	0	0.81	0.89	0.98
			Full	0	0.005	0.99	0.005	0	0.715	0.905	0.95
		BSCOV	—	0	0	0.995	0.005	0	0.875	0.945	0.995
		BDH	—	0	0.65	0.35	0	0	0.35	1	1
	500	MOSUM	Diagonal	0	0	1	0	0	0.83	0.92	0.99
			Full	0	0	1	0	0	0.755	0.885	0.945
		BSCOV	—	0.005	0	0.995	0	0	0.865	0.95	0.995
		BDH	—	0	0.615	0.385	0	0	0.385	1	1
800	100	MOSUM	Diagonal	0	0	0.995	0.005	0	0.785	0.905	0.975
			Full	0	0	0.985	0.01	0.005	0.715	0.88	0.925
		BSCOV	—	0	0.015	0.795	0.185	0.005	0.885	0.91	0.995
		BDH	—	0	0.46	0.54	0	0	0.535	1	1
	200	MOSUM	Diagonal	0	0	1	0	0	0.805	0.92	0.965
			Full	0	0	0.985	0.015	0	0.705	0.875	0.9
		BSCOV	—	0	0	0.78	0.215	0.005	0.89	0.965	0.995
		BDH	—	0	0.345	0.655	0	0	0.655	1	1
	500	MOSUM	Diagonal	0	0	1	0	0	0.775	0.89	0.975
			Full	0	0	0.98	0.02	0	0.765	0.92	0.92
		BSCOV	—	0	0	0.8	0.2	0	0.86	0.95	0.995
		BDH	—	0	0.375	0.625	0	0	0.62	1	1
1000	100	MOSUM	Diagonal	0	0	1	0	0	0.865	0.9	0.97
			Full	0	0	0.94	0.06	0	0.715	0.88	0.925
		BSCOV	—	0	0.005	0.835	0.16	0	0.87	0.925	0.99
		BDH	—	0	0.245	0.755	0	0	0.75	1	1
	200	MOSUM	Diagonal	0	0	0.985	0.015	0	0.83	0.935	0.98
			Full	0	0	0.925	0.075	0	0.735	0.88	0.92
		BSCOV	—	0	0	0.86	0.13	0.01	0.91	0.955	0.99
		BDH	—	0	0.185	0.815	0	0	0.81	1	1
	500	MOSUM	Diagonal	0	0	0.995	0.005	0	0.805	0.895	0.98
			Full	0	0	0.925	0.075	0	0.73	0.905	0.935
		BSCOV	—	0	0.005	0.83	0.16	0.005	0.875	0.965	0.99
		BDH	—	0	0.205	0.795	0	0	0.795	1	1

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | 4.2 with $(\rho_f, \rho_e) = (0.7, 0.3)$ and $R = 3$: Summary of change point estimators returned by MOSUM and BSCOV over 200 realisations.

				$\widehat{R} - R$					Accuracy		
n	p	Method	LRV	$-2 \leq$	-1	0	1	≥ 2	$j = 1$	$j = 2$	$j = 3$
400	100	MOSUM	Diagonal	0.005	0.08	0.915	0	0	0.595	0.765	0.905
			Full	0	0.06	0.935	0.005	0	0.49	0.74	0.85
		BSCOV	—	0.005	0.12	0.84	0.035	0	0.6	0.375	0.92
		BDH	—	0.005	0.845	0.14	0.01	0	0.12	0.995	1
	200	MOSUM	Diagonal	0.005	0.11	0.885	0	0	0.605	0.73	0.93
			Full	0	0.125	0.875	0	0	0.44	0.67	0.88
		BSCOV	—	0	0.145	0.825	0.03	0	0.625	0.32	0.9
		BDH	—	0	0.875	0.12	0.005	0	0.09	1	1
	500	MOSUM	Diagonal	0	0.115	0.885	0	0	0.535	0.71	0.95
			Full	0.005	0.115	0.88	0	0	0.485	0.68	0.855
		BSCOV	—	0	0.11	0.85	0.04	0	0.605	0.32	0.915
		BDH	—	0	0.815	0.185	0	0	0.145	1	1
600	100	MOSUM	Diagonal	0	0.025	0.97	0.005	0	0.565	0.69	0.905
			Full	0	0.045	0.925	0.03	0	0.48	0.705	0.85
		BSCOV	—	0.005	0.095	0.75	0.14	0.01	0.695	0.355	0.95
		BDH	—	0	0.575	0.39	0.035	0	0.375	1	1
	200	MOSUM	Diagonal	0	0.03	0.96	0.01	0	0.535	0.75	0.935
			Full	0	0.025	0.945	0.03	0	0.535	0.68	0.875
		BSCOV	—	0	0.1	0.785	0.11	0.005	0.665	0.38	0.945
		BDH	—	0	0.54	0.405	0.045	0.01	0.38	1	1
	500	MOSUM	Diagonal	0	0.005	0.995	0	0	0.57	0.77	0.935
			Full	0	0.025	0.945	0.03	0	0.495	0.685	0.865
		BSCOV	—	0.02	0	0.825	0.145	0.01	0.6	0.735	0.96
		BDH	—	0	0.53	0.4	0.065	0.005	0.415	1	1
800	100	MOSUM	Diagonal	0	0.01	0.96	0.03	0	0.54	0.69	0.92
			Full	0	0.005	0.885	0.11	0	0.49	0.685	0.85
		BSCOV	—	0.01	0	0.115	0.625	0.25	0.61	0.705	0.97
		BDH	—	0	0.345	0.605	0.05	0	0.58	1	1
	200	MOSUM	Diagonal	0	0	0.955	0.045	0	0.575	0.785	0.925
			Full	0	0.03	0.85	0.12	0	0.455	0.69	0.83
		BSCOV	—	0.01	0	0.08	0.595	0.315	0.625	0.805	0.975
		BDH	—	0	0.275	0.655	0.065	0.005	0.685	1	1
	500	MOSUM	Diagonal	0	0.01	0.95	0.04	0	0.6	0.725	0.915
			Full	0	0.01	0.835	0.155	0	0.495	0.63	0.8
		BSCOV	—	0.015	0	0.085	0.58	0.32	0.655	0.775	0.97
		BDH	—	0	0.295	0.63	0.065	0.01	0.64	1	1
1000	100	MOSUM	Diagonal	0	0	0.945	0.055	0	0.585	0.765	0.91
			Full	0	0.005	0.805	0.18	0.01	0.49	0.725	0.81
		BSCOV	—	0.005	0.005	0.195	0.5	0.295	0.64	0.825	0.95
		BDH	—	0	0.19	0.715	0.09	0.005	0.76	1	1
	200	MOSUM	Diagonal	0	0	0.945	0.05	0.005	0.545	0.77	0.92
			Full	0	0.01	0.81	0.17	0.01	0.47	0.74	0.84
		BSCOV	—	0	0.02	0.21	0.405	0.365	0.64	0.825	0.945
		BDH	—	0	0.135	0.765	0.08	0.02	0.825	1	1
	500	MOSUM	Diagonal	0	0	0.93	0.07	0	0.605	0.73	0.915
			Full	0	0	0.78	0.21	0.01	0.515	0.68	0.84
		BSCOV	—	0	0.01	0.205	0.45	0.335	0.67	0.82	0.96
		BDH	—	0	0.16	0.755	0.07	0.015	0.795	1	1

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | 4.2 with $R = 0$: Distribution of $\widehat{R} - R$ returned by MOSUM and BSCOV over 200 realisations for $(\rho_f, \rho_e) \in \{(0,0),(0.7,0.3)\}$.

				$(\rho_f, \rho_e) = (0,0)$			$(\rho_f, \rho_e) = (0.7,0.3)$		
n	***p***	**Method**	**LRV**	**0**	**1**	**≥ 2**	**0**	**1**	**≥ 2**
400	100	MOSUM	Diagonal	0.985	0.015	0	0.895	0.08	0.025
			Full	0.99	0.01	0	0.965	0.025	0.01
		BSCOV	—	1	0	0	0.75	0.215	0.035
		BDH	—	1	0	0	1	0	0
	200	MOSUM	Diagonal	0.995	0.005	0	0.88	0.11	0.01
			Full	0.995	0.005	0	0.955	0.04	0.005
		BSCOV	—	0.995	0.005	0	0.705	0.25	0.045
		BDH	—	1	0	0	1	0	0
	500	MOSUM	Diagonal	0.99	0.01	0	0.88	0.11	0.01
			Full	0.99	0.01	0	0.95	0.045	0.005
		BSCOV	—	0.995	0.005	0	0.75	0.22	0.03
		BDH	—	1	0	0	1	0	0
600	100	MOSUM	Diagonal	0.99	0.01	0	0.855	0.125	0.02
			Full	0.99	0.01	0	0.905	0.075	0.02
		BSCOV	—	1	0	0	0.82	0.155	0.025
		BDH	—	1	0	0	1	0	0
	200	MOSUM	Diagonal	0.99	0.01	0	0.845	0.14	0.015
			Full	0.995	0.005	0	0.88	0.105	0.015
		BSCOV	—	1	0	0	0.77	0.18	0.05
		BDH	—	1	0	0	1	0	0
	500	MOSUM	Diagonal	0.985	0.015	0	0.855	0.11	0.035
			Full	0.995	0.005	0	0.875	0.1	0.025
		BSCOV	—	1	0	0	0.85	0.135	0.015
		BDH	—	1	0	0	1	0	0
800	100	MOSUM	Diagonal	1	0	0	0.905	0.085	0.01
			Full	1	0	0	0.95	0.045	0.005
		BSCOV	—	1	0	0	0.835	0.15	0.015
		BDH	—	1	0	0	1	0	0
	200	MOSUM	Diagonal	1	0	0	0.94	0.05	0.01
			Full	0.995	0.005	0	0.97	0.02	0.01
		BSCOV	—	1	0	0	0.875	0.115	0.01
		BDH	—	1	0	0	1	0	0
	500	MOSUM	Diagonal	1	0	0	0.92	0.075	0.005
			Full	1	0	0	0.94	0.055	0.005
		BSCOV	—	1	0	0	0.86	0.12	0.02
		BDH	—	1	0	0	1	0	0
1000	100	MOSUM	Diagonal	1	0	0	0.895	0.1	0.005
			Full	1	0	0	0.945	0.05	0.005
		BSCOV	—	1	0	0	0.915	0.085	0
		BDH	—	1	0	0	1	0	0
	200	MOSUM	Diagonal	1	0	0	0.9	0.09	0.01
			Full	1	0	0	0.955	0.045	0
		BSCOV	—	1	0	0	0.895	0.09	0.015
		BDH	—	1	0	0	1	0	0
	500	MOSUM	Diagonal	1	0	0	0.91	0.085	0.005
			Full	1	0	0	0.95	0.05	0
		BSCOV	—	1	0	0	0.91	0.085	0.005
		BDH	—	1	0	0	1	0	0

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

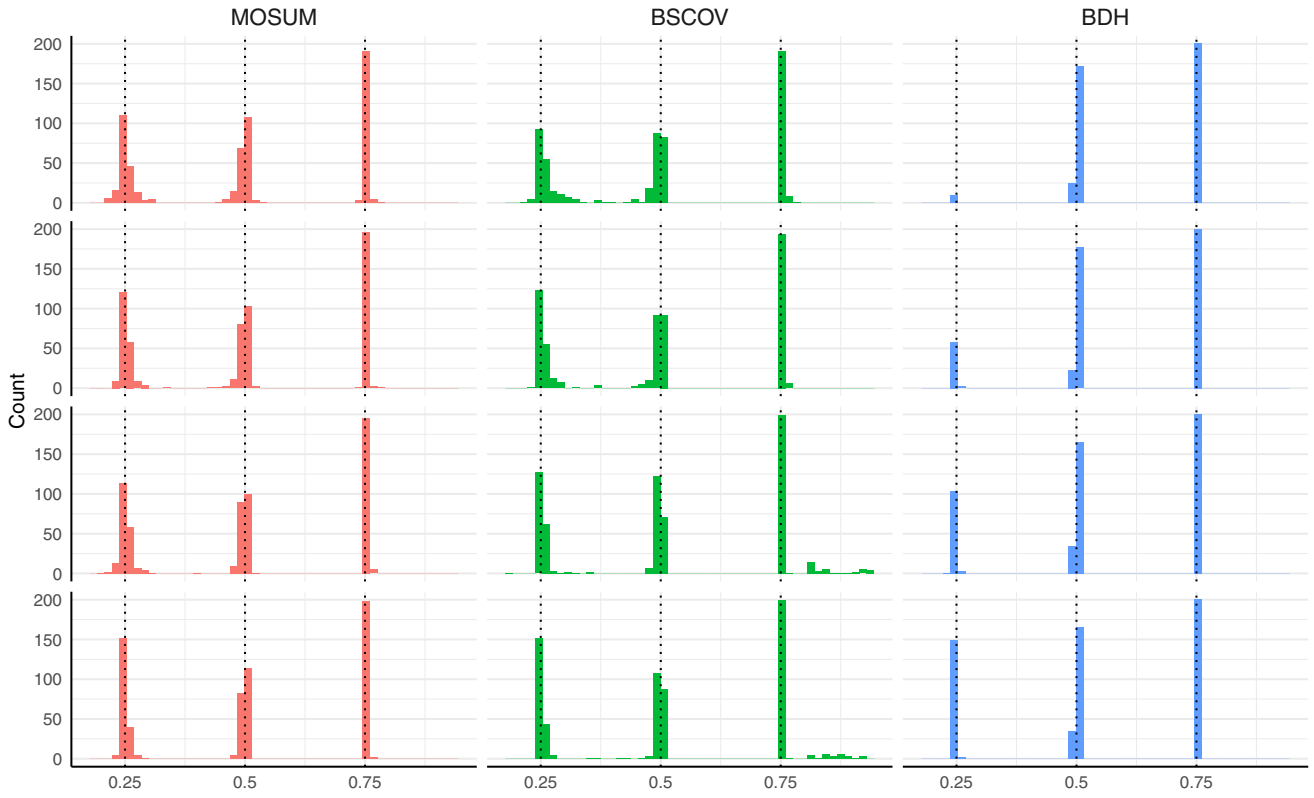

FIGURE 1 | 4.2 Histogram of the change point estimators returned by MOSUM-diagonal, BSCOV and BDH when $N = 100$, $(\rho_f, \rho_e) = (0, 0)$ and varying $T \in \{400, 600, 800, 1000\}$ (top to bottom). The scaled locations of the true change points, k_j/T, at $(1/4, 1/2, 3/4)$ are marked by vertical dotted lines.

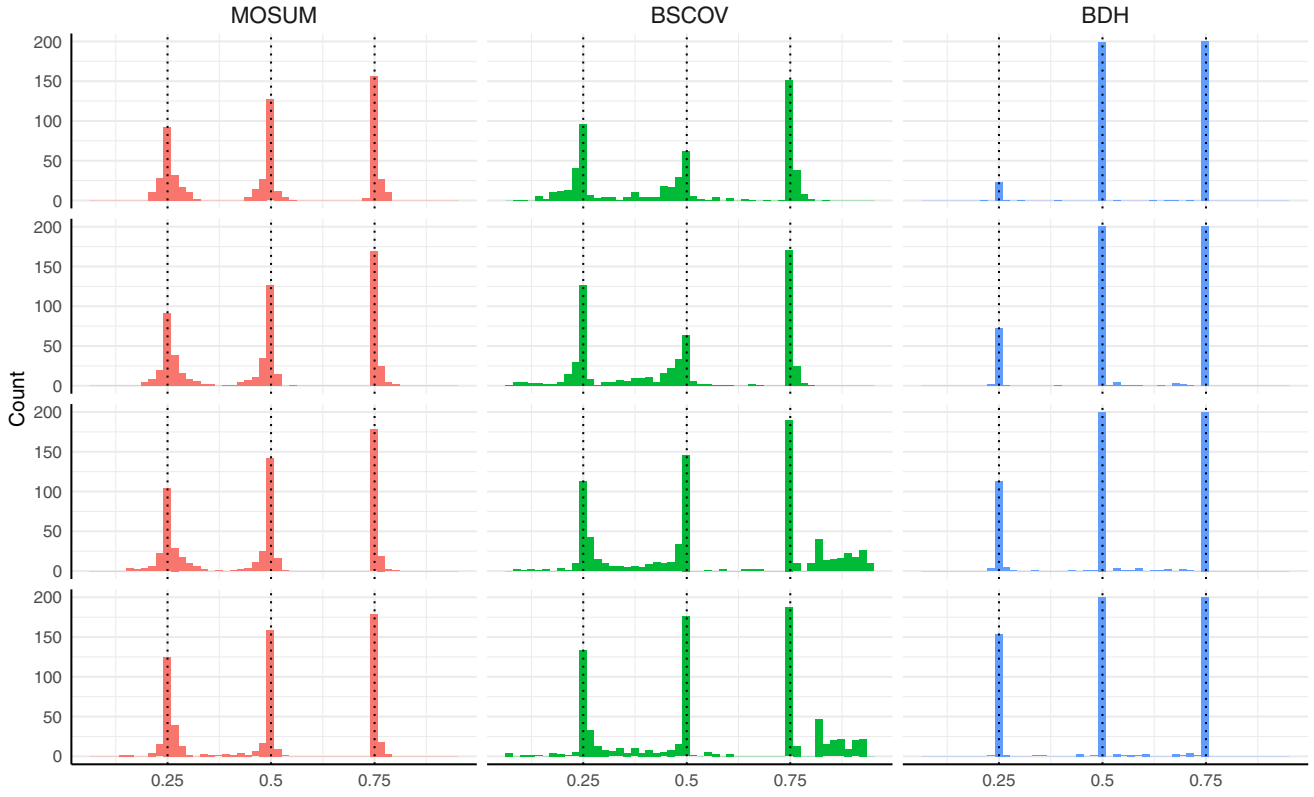

FIGURE 2 | 4.2 Histogram of the change point estimators returned by MOSUM-diagonal, BSCOV and BDH with $p = 100$, $(\rho_f, \rho_e) = (0.7, 0.3)$ and varying $T \in \{400, 600, 800, 1000\}$ (top to bottom). The scaled locations of the true change points, k_j/T, at $(1/4, 1/2, 3/4)$ are marked by vertical dotted lines.

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

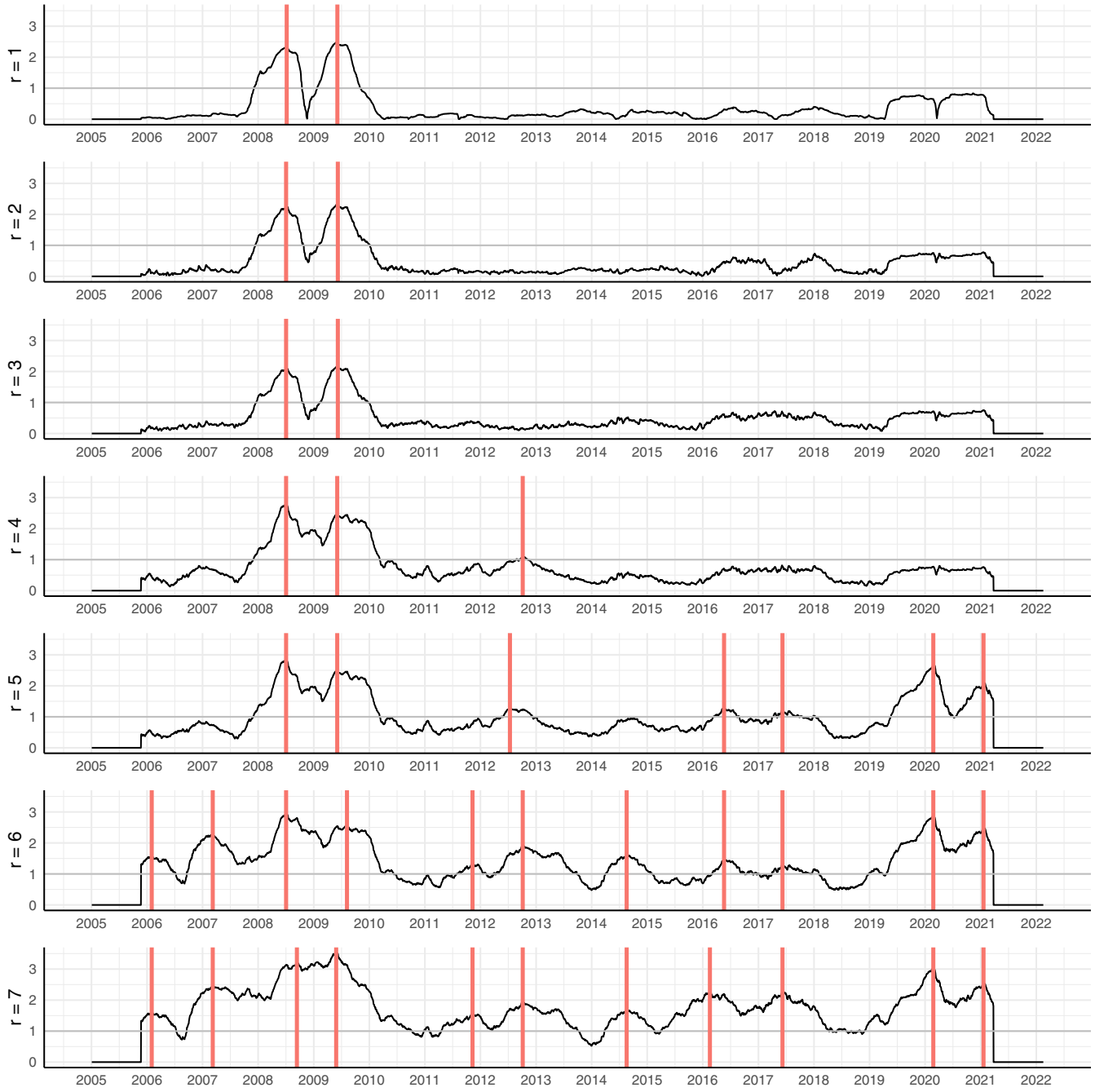

FIGURE 3 | US blue chip data: Standardised MOSUM statistics obtained with $r \in \{1, \ldots, 7\}$ (top to bottom). Vertical lines denote the change point estimators returned by MOSUM, and the horizontal line is at $y = 1$.

number of pseudo factors (such as k_2 under 4.2 and k_3 under 4.2) are estimated with higher accuracy, which agrees with the observations made in Duan et al. (2023) and also below Theorem 3.

Between MOSUM-full and MOSUM-diagonal, the latter demonstrates better accuracy in estimating both the total number and locations of change points when $R \geq 1$. This is explained by that $d = r(r+1)/2$ is as large as $d = 28$ under 4.2 and $d = 21$ under 4.2. Retaining the diagonal elements of $\widehat{\mathbf{V}}$ only effectively performs standardisation without suffering from the numerical instability inherent in inverting a large matrix. When $R = 0$, MOSUM-full performs marginally better as here, the factor number is kept at $r = 3$ under 4.2, leading to $d = 6$. Based on these observations, in practice, we recommend the use of MOSUM-diagonal when the number of (pseudo) factors is moderately large.

5 | Real Data Application

We consider daily stock prices from 72 US blue-chip companies across industry sectors between January 3, 2005 and February 16, 2022. We measure the volatility as the daily high-low range as $\sigma_{it}^2 = 0.361(p_{it}^{\text{high}} - p_{it}^{\text{low}})^2$ where p_{it}^{high} (resp. p_{it}^{low}) denotes the maximum (resp. minimum) price of stock i on day t, and set $X_{it} = \log(\sigma_{it}^2)$, see, for example, Diebold and Yilmaz (2014).

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

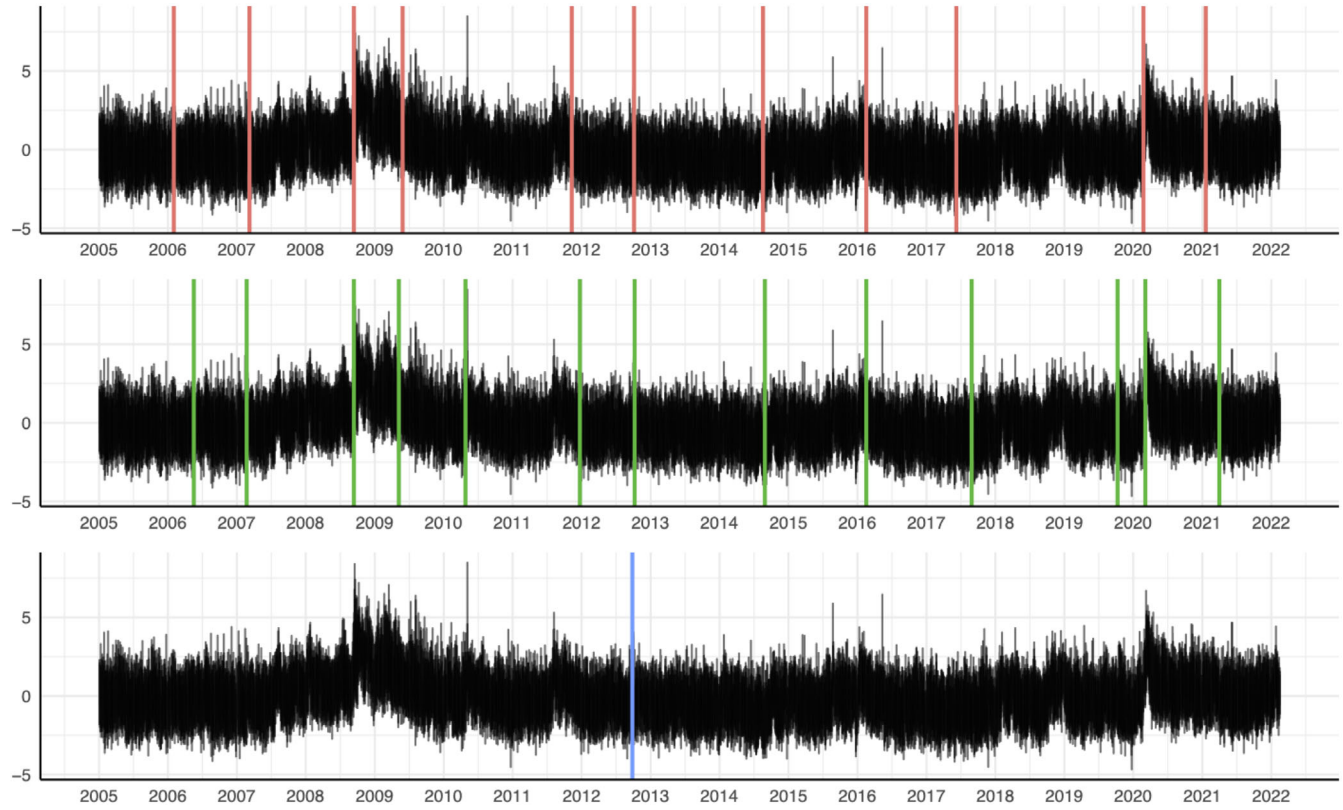

FIGURE 4 | US blue chip data: Volatilities from the 72 companies. Vertical lines denote the change point estimators detected by MOSUM, BSCOV and BDH (top to bottom), with $r = 7$.

The sub-sampling-based factor number estimator discussed in Section 4.1 returns $r = 7$ as the factor number. However, the panel data is unbalanced with the dimension $N = 72$ being considerably smaller than the sample size $T = 4312$, a situation that does not favour the sub-sampling approach as pointed out by Onatski (2024). The information criteria of Bai and Ng (2002) return $r = 5$, whilst the approach based on inspecting the ratio of successive eigenvalues (Ahn and Horenstein 2013) returns $r = 1$, which implies that any change point we detect would solely be attributed to heteroscedasticity of the single factor. Whilst the former is known to detect weakly pervasive factors (Bai and Ng 2023), the latter tends to recover only strongly pervasive ones.

In the presence of some uncertainty in the number of factors, a situation commonly faced in real data analysis, we choose to apply the proposed MOSUM procedure with varying $r \in \{1, \ldots, 7\}$ and inspect its outputs. We set other tuning parameters as described in Section 4.1 and adopt the standardisation based on the diagonal entries of $\widehat{\mathbf{V}}$ only (referred to as 'MOSUM-diagonal' in Section 4); this gives a bandwidth $\gamma = 227$ corresponding almost to one trading year. Figure 3 illustrates the series of standardised MOSUM statistics $D_{T,\gamma}^{-1} \cdot \mathcal{T}_{N,T,\gamma}(k)$, $\gamma \le k \le T - \gamma$, obtained for different values of r, where the standardisation is applied to ensure that the MOSUM series derived from $\mathsf{Vech}(\widehat{\mathbf{g}}_t \widehat{\mathbf{g}}_t^\top)$ of different dimensions are comparable. See also Figure 4 and Table 5 for the visualisation and the list of change point estimators returned by MOSUM and those competitors considered in the simulation studies, namely BSCOV (Li et al. 2023) and BDH (Bai et al. 2024).

It is noteworthy that the sets of change point estimator output by the MOSUM procedure are *nested* as the number of pseudo factors increases. That is, denoting by $\widehat{\mathcal{K}}(r)$ the set of change point estimators with r as the factor number, we have $\widehat{\mathcal{K}}(r) \subset \widehat{\mathcal{K}}(r')$ for any $r < r'$, when accommodating the possible bias in the change point estimators (up to 3 months). Specifically, with $r \le 3$, we detect two prominent changes in mid-2008 and 2009 which, being associated with the Great Financial Crisis in 2007–2009, are detected invariably with all r. With $r = 4$, we additionally detect 2012-10-05 as a change point, which is subsequently detected for all $r \ge 5$. With $r \ge 5$, 2020-02-24 and 2021-01-19 emerge as change points, which are accounted for by the stock market crash in February 2020 and the ensuing recession due to the COVID-19 pandemic. With $r \in \{6, 7\}$, MOSUM outputs almost identical sets of change point estimators. These results offer an interpretation as to how different factors 'encode' different structural changes, and demonstrate that the MOSUM procedure is insensitive to the specified number of factors within certain ranges (i.e., $\{1, 2, 3\}$, $\{6, 7\}$).

Similarly to MOSUM, BSCOV also returns nested sets of change point estimators as r increases, and many of its estimators overlap with those returned by MOSUM. When $r = 7$, the estimators returned by MOSUM form a subset of those returned by BSCOV, and some changes detected solely by BSCOV do not appear as estimators detected with $r < 7$ by any method. BDH estimates the number of factors by the information criterion of Bai and Ng (2002) at each iteration of the binary segmentation algorithm, and as such, when applied with a fixed number of factors, its output lacks the nested property.

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 | US blue chip data: List of change point estimators obtained with $r \in \{1, \ldots, 7\}$ by MOSUM, BSCOV (Li et al. 2023) and BDH (Bai et al. 2024).

r	MOSUM	BSCOV	BDH
1	2008-07-07, 2009-06-05	2008-09-12, 2009-03-25	—
2	2008-07-07, 2009-06-05	2008-09-12, 2009-03-25	—
3	2008-07-07, 2009-06-05	2008-09-12, 2009-03-25	2008-06-27, 2029-08-07
4	2008-07-03, 2009-06-04, 2012-10-05	2008-09-12, 2009-05-08, 2012-10-05	2008-07-03, 2009-07-23, 2012-06-21, 2014-09-29, 2020-02-24
5	2008-07-03, 2009-06-04, 2012-07-13, 2016-05-20, 2017-06-08, 2020-02-24, 2021-01-19	2008-09-12, 2009-05-08, 2012-10-08, 2014-10-24, 2020-03-05	2007-10-31, 2012-10-05
6	2006-02-01, 2007-03-09, 2008-07-03, 2009-08-07, 2011-11-10, 2012-10-05, 2014-08-19, 2016-05-20, 2017-06-08, 2020-02-24, 2021-01-19	2007-01-11, 2008-09-12, 2009-05-08, 2011-12-23, 2012-10-08, 2014-08-29, 2020-03-05	2012-10-05
7	2006-02-01, 2007-03-09, 2008-09-12, 2009-05-28, 2011-11-10, 2012-10-05, 2014-08-19, 2016-02-17, 2017-06-08, 2020-02-24, 2021-01-19	2006-05-18, 2007-02-22, 2008-09-12, 2009-05-08, 2010-04-26, 2011-12-23, 2012-10-08, 2014-08-29, 2016-02-17, 2017-08-28, 2019-10-10, 2020-03-05, 2021-04-01	2012-09-26

6 | Conclusions

This paper proposes a MOSUM procedure for change point analysis under a static factor model that is popularly adopted in econometrics and statistics. In addition to deriving the asymptotic null distribution of the maximally selected MOSUM statistic, we establish the consistency of the procedure in multiple change point estimation with the accompanying rate of estimation, contributing to the relatively scarce literature on multiple change point detection in factor models. On a range of simulated datasets and in a real data application, we demonstrate the competitiveness of the proposal empirically. At the same time, the success of the proposed single-scale MOSUM procedure hinges on the availability of the bandwidth γ that fulfils a set of assumptions, which may not exist in the presence of multiscale change points (Cho and Kirch 2024). One natural avenue for an extension is to apply the MOSUM procedure with a range of bandwidths, which we leave for future research.

Acknowledgments

The authors are grateful to the editors and the anonymous reviewer for their constructive comments and suggestions to improve the contents of this paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The dataset analysed in Section 5 is retrieved from the Wharton Research Data Services at https://wrds-www.wharton.upenn.edu.

References

Ahn, S. C., and A. R. Horenstein. 2013. "Eigenvalue Ratio Test for the Number of Factors." *Econometrica* 81, no. 3: 1203–1227.

Alessi, L., M. Barigozzi, and M. Capasso. 2010. "Improved Penalization for Determining the Number of Factors in Approximate Static Factor Models." *Statistics & Probability Letters* 80: 1806–1813.

Aue, A., S. Hörmann, L. Horváth, and M. Reimherr. 2009. "Break Detection in the Covariance Structure of Multivariate Time Series Models." *Annals of Statistics* 37, no. 6B: 4046–4087.

Aue, A., and C. Kirch. 2024. "The State of Cumulative Sum Sequential Changepoint Testing 70 Years After Page." *Biometrika* 111, no. 2: 367–391.

Bai, J. 2003. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica* 71, no. 1: 135–171.

Bai, J., J. Duan, and X. Han. 2024. "The Likelihood Ratio Test for Structural Changes in Factor Models." *Journal of Econometrics* 238, no. 2: 105631.

Bai, J., and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70, no. 1: 191–221.

Bai, J., and S. Ng. 2023. "Approximate Factor Models With Weaker Loadings." *Journal of Econometrics* 235, no. 2: 1893–1916.

Baltagi, B. H., C. Kao, and F. Wang. 2017. "Identification and Estimation of a Large Factor Model With Structural Instability." *Journal of Econometrics* 197, no. 1: 87–100.

Barigozzi, M., H. Cho, and P. Fryzlewicz. 2018. "Simultaneous Multiple Change-Point and Factor Analysis for High-Dimensional Time Series." *Journal of Econometrics* 206, no. 1: 187–225.

Barigozzi, M., and L. Trapani. 2020. "Sequential Testing for Structural Stability in Approximate Factor Models." *Stochastic Processes and their Applications* 130, no. 8: 5149–5187.

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bates, B. J., M. Plagborg-Møller, J. H. Stock, and M. W. Watson. 2013. "Consistent Factor Estimation in Dynamic Factor Models With Structural Instability." *Journal of Econometrics* 177, no. 2: 289–304.

Berkes, I., S. Hörmann, and J. Schauer. 2011. "Split Invariance Principles for Stationary Processes." *Annals of Probability* 39, no. 6: 2441–2473.

Breitung, J., and S. Eickmeier. 2011. "Testing for Structural Breaks in Dynamic Factor Models." *Journal of Econometrics* 163, no. 1: 71–84.

Chamberlain, G., and M. Rothschild. 1983. "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets." *Econometrica* 51, no. 5: 1281–1304.

Cheng, X., Z. Liao, and F. Schorfheide. 2016. "Shrinkage Estimation of High-Dimensional Factor Models With Structural Instabilities." *Review of Economic Studies* 83, no. 4: 1511–1543.

Cho, H., and C. Kirch. 2024. "Data Segmentation Algorithms: Univariate Mean Change and Beyond." *Econometrics and Statistics* 30: 76–95.

Cho, H., H. Maeng, I. A. Eckley, and P. Fearnhead. 2024. "High-Dimensional Time Series Segmentation via Factor-Adjusted Vector Autoregressive Modeling." *Journal of the American Statistical Association* (forthcoming) 119: 2038–2050.

Cho, H., and D. Owens. 2024. "High-Dimensional Data Segmentation in Regression Settings Permitting Heavy Tails and Temporal Dependence." *Electronic Journal of Statistics* 18, no. 1: 2620–2664.

Corradi, V., and N. R. Swanson. 2014. "Testing for Structural Stability of Factor Augmented Forecasting Models." *Journal of Econometrics* 182, no. 1: 100–118.

Diebold, F. X., and K. Yilmaz. 2014. "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms." *Journal of Econometrics* 182, no. 1: 119–134.

Duan, J., J. Bai, and X. Han. 2023. "Quasi-Maximum Likelihood Estimation of Break Point in High-Dimensional Factor Models." *Journal of Econometrics* 233, no. 1: 209–236.

Eichinger, B., and C. Kirch. 2018. "A MOSUM Procedure for the Estimation of Multiple Random Change Points." *Bernoulli* 24, no. 1: 526–564.

Han, X., and A. Inoue. 2015. "Tests for Parameter Instability in Dynamic Factor Models." *Econometric Theory* 31: 1117–1152.

He, Y., X. Kong, and L. Yu. 2024. "Online Change-Point Detection for Matrix-Valued Time Series With Latent Two-Way Factor Structure." *Annals of Statistics* 52, no. 4: 1646–1670.

Hörmann, S. 2009. "Berry-Esseen Bounds for Econometric Time Series." *Latin American Journal of Probability and Mathematical Statistics* 6: 377–397.

Horváth, L., and G. Rice. 2024. *Change Point Analysis for Time Series*. Springer.

Hušková, M., and A. Slabỳ. 2001. "Permutation Tests for Multiple Changes." *Kybernetika* 37, no. 5: 605–622.

Kim, J., H.-S. Oh, and H. Cho. 2024. "Moving Sum Procedure for Change Point Detection Under Piecewise Linearity." *Technometrics* 66, no. 3: 358–367.

Kirch, C., and P. Klein. 2024. "Moving Sum Data Segmentation for Stochastics Processes Based on Invariance." *Statistica Sinica* 30: 76–95.

Kirch, C., B. Muhsal, and H. Ombao. 2015. "Detection of Changes in Multivariate Time Series With Application to Eeg Data." *Journal of the American Statistical Association* 110, no. 511: 1197–1216.

Kirch, C., and K. Reckruehm. 2024. "Data Segmentation for Time Series Based on a General Moving Sum Approach." *Annals of the Institute of Statistical Mathematics* 76, no. 3: 393–421.

Li, Y.-N., D. Li, and P. Fryzlewicz. 2023. "Detection of Multiple Structural Breaks in Large Covariance Matrices." *Journal of Business & Economic Statistics* 41, no. 3: 846–861.

Liu, W., and Z. Lin. 2009. "Strong Approximation for a Class of Stationary Processes." *Stochastic Processes and their Applications* 119, no. 1: 249–280.

Lütkepohl, H. 2005. *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.

Onatski, A. 2024. "The Hallin-Liška Criterion Through the Lens of the Random Matrix Theory." In *Recent Advances in Econometrics and Statistics: Festschrift in Honour of Marc Hallin*, edited by M. Barigozzi, S. Hörmann, and D. Paindaveine. Springer.

Spearman, C. 1904. "General Intelligence Objectively Determined and Measured." *American Journal of Psychology* 15, no. 2: 201–293.

Stock, J. H., and M. W. Watson. 2009. "Forecasting in Dynamic Factor Models Subject to Structural Instability." In *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, edited by J. Castle and N. Shephard, 173–205. Oxford University Press.

Stock, J. H., and M. W. Watson. 2011. "Dynamic Factor Models." In *Oxford Handbook of Economic Forecasting*, edited by M. Clements and D. Hendry. Oxford University Press.

Wu, W. B. 2005. "Nonlinear System Theory: Another Look at Dependence." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 40: 14150–14154.

Yamamoto, Y., and S. Tanaka. 2015. "Testing for Factor Loading Structural Change Under Common Breaks." *Journal of Econometrics* 189, no. 1: 187–206.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1**: Supporting Information.

14679892, 2026, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jtsa.70028 by Area Sistemi Dipart & Doc, Wiley Online Library on [27/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License