Econometric Theory, **39**, 2023, 534–581.
doi:10.1017/S026646662200010X

ON MULTIPLE STRUCTURAL BREAKS IN DISTRIBUTION: AN EMPIRICAL CHARACTERISTIC FUNCTION APPROACH

Zhonghao Fu
Fudan University
and
Shanghai Institute of International Finance and Economics

Yongmiao Hong
Chinese Academy of Sciences
and
University of Chinese Academy of Sciences

Xia Wang
Renmin University of China

We estimate and test for multiple structural breaks in distribution via an empirical characteristic function approach. By minimizing the sum of squared generalized residuals, we can consistently estimate the break fractions. We propose a sup-F type test for structural breaks in distribution as well as an information criterion and a sequential testing procedure to determine the number of breaks. We further construct a class of derivative tests to gauge possible sources of structural breaks, which is asymptotically more powerful than the smoothed nonparametric tests for structural breaks. Simulation studies show that our method performs well in determining the appropriate number of breaks and estimating the unknown breaks. Furthermore, the proposed tests have reasonable size and excellent power in finite samples. In an application to exchange rate returns, our tests are able to detect structural breaks in distribution and locate the break dates. Our tests also indicate that the documented breaks appear to occur in variance and higher-order moments, but not so often in mean.

Fu acknowledges financial support from the National Science Foundation of China (NSFC) (Nos. 71903032 and 72121002), Hong acknowledges financial support from the NSFC's Fundamental Scientific Center Project (No. 71988101), and Wang acknowledges financial support from the NSFC (No. 71873151). We thank the Editor (Peter C.B. Phillips), the Co-Editor (Robert Taylor), three anonymous referees, and seminar participants in the 2019 City University of Hong Kong Workshop in Econometrics and Statistics, the International Workshop on Extreme Value Theory at Fudan University, and the 2019 Guangzhou Econometrics Workshop, for their insightful comments and suggestions. All remaining errors are solely ours. Address correspondence to Zhonghao Fu, School of Economics, Fudan University, Shanghai, China; e-mail: zhfu@fudan.edu.cn; Yongmiao Hong, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing, China; e-mail: yh20@cornell.edu; Xia Wang, School of Economics, Renmin University of China, Beijing, China; e-mail: wxia@ruc.edu.cn.

1. INTRODUCTION

Structural breaks are common in economic time series dynamics. Shocks induced by institutional changes, such as the transition from a fixed to a floating exchange rate mechanism, the initiation of new currencies, the introduction of new accounting standards, or the outbreak of the COVID-19 pandemic, may cause the structural parameters in an economic model to vary over time. Besides, policy switches, preference changes, and technological progress can also lead to structural changes. Many empirical studies have documented the structural instability of economic time series data (e.g., Stock and Watson, 1996; Hansen, 2001; Zhang, Osborn, and Kim, 2008; Rossi, 2013).

Structural breaks in time series have drawn considerable attention. Both econometricians and statisticians have devoted a vast amount of effort to this field. Most works investigate structural breaks in linear regression models. Examples include Bai's (1996) test for parameter constancy in linear regression model via the empirical distribution function (EDF), Bai and Perron's (1998) approach for multiple structural breaks in linear regression models, Qu and Perron's (2007) work for a system of equations, Chen and Hong (2012) nonparametric tests for smooth structural changes, Hall, Han, and Boldea's (2012) test for abrupt structural breaks, and Chen's (2015) test for smooth structural changes, both in linear regression models with endogeneity. The aforementioned studies mainly focus on estimation and testing for structural breaks in mean. They may easily miss a structural break in higher-order moments. For instance, when a structural break occurs in conditional variance, tests for structural breaks in conditional mean may fail to detect it. To examine the stability of a time series completely, one should consider structural changes in distribution. In the forecasting literature, density forecasts can provide more insight for macroeconomic risk management than point forecasts; see, e.g., Diebold, Gunther, and Tay (1998) and Diebold, Hahn, and Tay (1999). However, due to the instability of macroeconomic and financial time series data (Rossi, 2013), density forecasts may deliver suboptimal predictions under structural breaks (e.g., Rossi and Sekhposyan, 2013; González-Rivera and Sun, 2017). In financial risk management, detecting structural breaks in the tail distribution of financial returns is important for assessing the probability of extreme events (e.g., Koedijk, Schafgans, and de Vries 1990; Quintos, Fan, and Phillips, 2001; Lin and Kao, 2008). In addition, when studying nonlinear dependence of financial variables, it is necessary to test the constancy of copulas, which may change when the joint distribution has structural breaks (e.g., Busetti and Harvey, 2011; Manner, Stark, and Wied, 2019).

Aside from the remarkable progress on structural breaks in mean, attention has been paid to distributional changes. Among them, Cowell (1985) introduces a system of axioms to measure distributional changes. Dümbgen (1991) estimates a structural break in distribution and develops the corresponding asymptotic theory for independent random variables. Yakir (1997) considers the problem of raising the alarm as soon as possible for a distributional change of independent random variables. Inoue (2001) proposes two nonparametric tests for an unknown break in the distribution of a time series via the EDF. Qu (2008) proposes two tests for structural breaks in regression quantiles. Selk and Neumeyer (2013) propose a Kolmogorov–Smirnov test to check whether the innovation distribution changes in a nonparametric autoregression via the EDF. Zou, Yin, Feng, and Wang (2014) propose a nonparametric maximum likelihood approach for multiple structural breaks based on independent data. Zhou, Fu, and Zhang (2017) propose two nonparametric tests based on the empirical likelihood function and likelihood ratio to detect a distributional change for independent observations. Although the aforementioned studies provide useful tools for detecting distributional changes, the majority of them focus on testing for distributional change for independent observations with a single break. Making inference, such as determining the number of breaks, is useful in practice. If breaks are detected accurately, one can then split the whole sample into several subsamples and regard the data in each subsample as a stationary process. In this way, one can infer the regime-switch behavior of economic dynamics. For instance, it is crucial to determine which regime, the forward- or backward-looking behavior, dominates the short-run inflation dynamics in the New-Keynesian Phillips curve (e.g., Galí and Gertler, 1999). After identifying structural breaks, one can then investigate this important issue in subsamples (e.g., Zhang et al., 2008).

This paper proposes an empirical characteristic function (ECF) approach to estimating and testing for multiple structural breaks in the joint distribution of a multivariate time series. Furthermore, it develops a sequence of derivative tests to investigate structural breaks in various aspects of the joint distribution. By equivalence between a cumulative distribution function (CDF) and the corresponding CF, structural breaks in distribution are equivalent to structural breaks in the CF. Given an observed time series sample, we can estimate and test structural breaks using the ECF. The ECF is a powerful analytic tool and has been widely used in econometrics literature, such as testing for stationarity and normality (Epps, 1987, 1988), testing for independence and conditional independence (Pinkse, 1998; Hong, 1999; Hong and Lee, 2003; Su and White, 2007; Wang and Hong, 2018), estimating parameters (Singleton, 2001; Jiang and Knight, 2002; Knight and Yu, 2002; Chacko and Viceira, 2003), and testing for strict stationarity (Hong, Wang, and Wang, 2017). There is also a growing literature on detecting structural breaks in distribution using ECFs. For example, Hušková and Meintanis (2006) propose tests to detect a distributional structural change for independent observations. Hušková and Meintanis (2008) construct tests for the multivariate k-sample problem for independent data. Hlávka, Hušková, Kirch, and Meintanis (2012) develop an online procedure to monitor structural breaks in the error distribution of an autoregressive time series. Hlávka, Hušková, and Meintanis (2017) develop tests to detect a change point for multivariate independent observations and time series observations in vector autoregressive models. Hlávka, Hušková, and Meintanis (2020) propose two-sample break-detection procedures for vectorial observations based on ECFs. The existing relevant literature has made a remarkable contribution

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

to the topic of distributional structural breaks. In this paper, we also use an ECF approach to develop a new test for structural breaks in distribution. Unlike the aforementioned literature that mainly focuses on a single structural break for independent data, we consider multiple distributional structural changes of time series data. As an essential feature of our approach, we transform the CF-based structural break problem to a generalized regression representation which has an interesting interpretation as a time series functional data regression. Berkes, Gabrys, Horváth, and Kokoszka (2009) propose a functional CUSUM-type test to check a single break in the mean function under the assumption of independent and identically distributed error functions. Aue, Gabrys, Horváth, and Kokoszka (2009) derive the limiting distribution of a breakpoint estimator under the same setting. Hörmann and Kokoszka (2010) discuss detecting a shift in mean for weakly dependent functional data. Aston and Kirch (2012) focus on an epidemic change model for weakly dependent functional data. Aue, Rice, and Sönmez (2018) propose a method to detect and date structural breaks in functional data without dimension reduction. Unlike the existing literature on functional data analysis that usually assumes the error function to be identically distributed, the covariance structure of the generalized error function in our generalized regression framework is not identically distributed over time under structural breaks. As a result, our asymptotic results, especially the asymptotic local power of our test, are more complicated.

Our paper complements the existing studies on distributional changes with several important features. First, we estimate breaks by minimizing a sum of squared generalized residuals. Using the ECF, we can convert structural breaks in distribution into structural breaks in a generalized regression. Thus, by extending Bai and Perron's (1998) approach for structural breaks in a linear regression to the generalized regression, we can consistently estimate the break fractions and derive the convergence rates of the break fraction estimators for time series data. We note that our asymptotic results differ substantially from the classical literature such as Bai and Perron (1998). The key is that the structural breaks in distribution will affect the covariance structure of the generalized error function in our generalized regression framework. This complicates our asymptotic analysis and distinguishes our asymptotic results from the existing literature on structural breaks.

Second, we propose a sup-F type test for the null hypothesis of no distributional change against the alternative hypothesis of a fixed number of breaks in the joint distribution of a multivariate time series. The proposed test has nontrivial power against a class of local alternatives that converge to the null hypothesis at the rate of $\sqrt{T}$, where T is the number of time series observations. We also show that the dimension of time series data does not affect the asymptotic power of our test. It implies that our approach is particularly suitable for testing structural breaks in a multivariate time series.

Third, when the null hypothesis of no structural break in distribution is rejected, we suggest a Bayesian information criterion (BIC)-type information criterion and a sequential testing procedure to estimate the number of breaks. Our simulation studies show that both procedures perform well in consistently estimating the true number of breaks. Zou et al. (2014) also show the consistency of BIC in determining the number of breaks. However, their theory is developed for independent data.

Fourth, by the moment-generating property of a joint CF when moments exist, we develop a class of derivative tests to gauge possible sources of structural breaks, including breaks in univariate distribution, univariate moments, and conditional moments (e.g., conditional mean or conditional variance). Compared to the EDF-based approaches such as Inoue (2001), our ECF approach can provide insight into the pattern of structural breaks.

In an empirical study, we apply our approach to exchange rate returns of Euro, Japanese Yen, Chinese Yuan, and Canadian Dollar. We find strong evidence of structural breaks in distribution and locate the break dates. Our tests also suggest that the documented structural breaks occur in variance and higher-order moments.

We organize the paper as follows. In Section 2, we introduce our ECF-based estimation method for distributional breaks and investigate the asymptotic properties of the estimated break fractions. In Section 3, we propose a joint test for structural breaks and derive the asymptotic distributions under the null and local alternative hypotheses, respectively. We propose two methods to determine the number of breaks in and develop a class of derivative tests to infer patterns of structural breaks in Section 5. We study the finite sample performance of the proposed estimation procedure and tests in Section 6 and conduct an empirical application to exchange rate returns in Section 7. Section 8 concludes. The proofs of the main results and additional simulation results are relegated to the Supplementary Material.

Throughout this paper, $\mathbf{i}$ denotes the imaginary number such that $\mathbf{i} = \sqrt{-1}$. For a scalar a, $\lfloor a \rfloor$ denotes its integer part. For an $m \times n$ complex-valued matrix $A = (a_{ij})$, where a_{ij} is the (i,j)th entry for $i = 1, \ldots, m; j = 1, \ldots, n$, we denote its complex conjugate by $A^* = (a_{ji}^*)$, its transpose by $A' = (a_{ji})$, its real part by $\mathrm{Re}(A) = (\mathrm{Re}[a_{ij}])$, and its euclidean norm by $\|A\| = (\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2)^{1/2}$. The operators "$\xrightarrow{p}$," "$\xrightarrow{d}$," and "$\Rightarrow$" denote convergence in probability, convergence in distribution, and weak convergence, respectively.

2. ESTIMATION AND LIMITING DISTRIBUTION

2.1. Estimation

Let $\{Y_t\}_{t=1}^T$ be a d-dimensional time series sample, and the corresponding CDFs be $\{\mathcal{F}_{Y_t}^0(y)\}_{t=1}^T$, where $d \geq 1$ is a fixed integer. Throughout, we denote the true value of a parameter with a superscript "0." Suppose there exist M^0 breaks in the distribution of Y_t, namely $\{T_1^0, \ldots, T_{M^0}^0\}$, and the corresponding break fractions are $\{r_1^0, \ldots, r_{M^0}^0\}$, where $r_j^0 = T_j^0/T$ for $j = 1, 2, \ldots, M^0$. We let the collection of variables $\{Y_{T_{j-1}^0+1}, \ldots, Y_{T_j^0}\}$ in the jth subsample be strictly stationary for $j = 1, 2, \ldots, M^0 + 1$, where we adopt the convention that $T_0^0 = 0$ and $T_{M^0+1}^0 = T$.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

That is, the CDFs of $\{Y_t\}_{t=1}^T$ are assumed to be identical within each time segment $[T_{j-1}^0+1, T_j^0]$:

$$\mathcal{F}_{Y_t}^0(y) = \begin{cases} \mathbf{F}_1^0(y), & \text{for } t = 1,\dots,T_1^0, \\ \mathbf{F}_2^0(y), & \text{for } t = T_1^0+1,\dots,T_2^0, \\ \quad\vdots & \qquad\vdots \\ \mathbf{F}_{M^0+1}^0(y), & \text{for } t = T_{M^0}^0+1,\dots,T, \end{cases}$$

where $\{\mathbf{F}_j^0(y)\}_{j=1}^{M^0+1}$ characterize the M^0+1 regimes of the time series sample $\{Y_t\}_{t=1}^T$. We note that the aforementioned assumption can be violated if structural breaks produce a trend in the data and the drift impacts the distribution of Y_t within a subsample. We rule out that case in our setting. In general, the number of breaks M^0 can be regarded as an unknown parameter. However, we treat it as known in this section and will propose data-driven procedures to determine the number of breaks in Section 4.

Let $\phi_t^0(u) \equiv E(e^{\mathbf{i}u'Y_t})$ be the CF of Y_t at time t. Given the equivalence between a CDF and its corresponding CF, it follows

$$\phi_t^0(u) = \begin{cases} \psi_1^0(u), & \text{for } t = 1,\dots,T_1^0, \\ \psi_2^0(u), & \text{for } t = T_1^0+1,\dots,T_2^0, \\ \quad\vdots & \qquad\vdots \\ \psi_{M^0+1}^0(u), & \text{for } t = T_{M^0}^0+1,\dots,T, \end{cases}$$

where $\{\psi_j^0(u)\}_{j=1}^{M^0+1}$ are the M^0+1 regimes of the CFs $\{\phi_t^0(u)\}_{t=1}^T$ specified by the true break dates $\{T_1^0,\dots,T_{M^0}^0\}$. When there exist M^0 structural breaks, the CF $\phi_t^0(u)$ behaves like a step function of a rescaled time index $t/T \in (0,1]$ for each $u \in \mathbb{R}^d$. This motivates us to consider the following time-varying generalized regression

$$e^{\mathbf{i}u'Y_t} = \phi_t^0(u) + \varepsilon_t(u), \tag{2.1}$$

where $\varepsilon_t(u)$ is a generalized error function such that $E[\varepsilon_t(u)] = 0$, $E|\varepsilon_t(u)|^2 = 1 - |\phi_t^0(u)|^2$, and $\operatorname{cov}[\varepsilon_t(u), \varepsilon_t(v)] = E[\varepsilon_t(u)\varepsilon_t(v)^*] = \phi_t^0(u-v) - \phi_t^0(u)\phi_t^0(-v)$. Since the CF always exists for any distribution, (2.1) is a natural representation of a multivariate time series in the frequency domain. It facilitates the analysis of structural breaks in the distribution of a multivariate time series. For each fixed u, (2.1) is equivalent to a decomposition of a complex-valued random variable $e^{\mathbf{i}u'Y_t}$ into its unconditional mean $E(e^{\mathbf{i}u'Y_t})$ and a disturbance $\varepsilon_t(u)$. Hence, we can relate the distributional change point issue with the strand of literature on testing and estimating multiple structural breaks in a regression framework (e.g., Bai, 1994; Bai and Perron, 1998).

Interestingly, we note that $e^{\mathbf{i}u'Y_t}$ is a functional process of u. It implies that (2.1) is analogous to analyzing structural breaks in the mean of a functional time series (e.g., Berkes et al., 2009; Hörmann and Kokoszka, 2010; Aue et al., 2018). However, we emphasize that our generalized regression framework differs from

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

the existing literature on functional data in two aspects. First, the generalized error function in (2.1) is not identically distributed across time t since its covariance structure is inherently influenced by break dates (e.g., $\mathrm{var}(e^{\mathbf{i}u'Y_t}) = 1 - |\phi_t^0(u)|^2$). Thus, the second moment of $\varepsilon_t(u)$ shares the same break dates with the distribution of Y_t. A structure break implies not only a shift of the CF, but also a change in the covariance structure of $\varepsilon_t(u)$. As a result, our asymptotic local power analysis is different and more complicated since the covariance structure of the test statistics is much involved. In contrast, the existing literature on functional data usually assumes the error function to be exogenously given. That is, the error function's covariance structure is not affected by structural breaks. For example, Berkes et al. (2009) assume the error function to be independent and identically distributed across time t. Hörmann and Kokoszka (2010) and Aue et al. (2018) require the error function to be L^4-m-approximable, which implies that the error function is identically distributed. We note that our generalized error function $\varepsilon_t(u)$ does not satisfy the above assumptions when structural breaks occur. Second, unlike the existing literature on functional data analysis that is motivated by analyzing an observable functional data process, such as a temperature curve or an intraday stock price curve, we do not require to observe a functional data process directly. Neither do we require to observe the true CF $\phi_t^0(u)$. We note that our estimators and test statistics for structural breaks are solely based on the ECF, which can be directly computed using the observed time series sample $\{Y_t\}_{t=1}^T$. Given each observed Y_t, $e^{\mathbf{i}u'Y_t}$ is directly computable for any $u \in \mathbb{R}^d$. Then, the complex-valued time series sample $\{e^{\mathbf{i}u'Y_t}\}_{t=1}^T$ can be constructed accordingly. We note that $\{e^{\mathbf{i}u'Y_t}\}_{t=1}^T$ can be viewed as an observable continuous functional process of $u \in \mathbb{R}^d$ given the observed time series sample $\{Y_t\}_{t=1}^T$. Under a collection of break dates $\{T_j\}_{j=1}^{M^0}$, the ECF of Y_t in each subsample $\{Y_t\}_{t=T_{j-1}+1}^{T_j}$ is just a sample average of $\{e^{\mathbf{i}u'Y_t}\}_{t=T_{j-1}+1}^{T_j}$, for $j = 1, \ldots, M^0 + 1$. The ECF is a consistent estimator for the true CF when $\{T_j\}_{j=1}^{M^0}$ coincide with the true break dates. With the feasible ECF, we note that observing the time series sample $\{Y_t\}_{t=1}^T$ is sufficient to adopt our approach.

Given there are M^0 breaks in $\phi_t^0(u)$, and $\psi_j^0(u)$ is the true CF in each time segment $[T_{j-1}^0 + 1, T_j^0]$ for $j = 1, \ldots, M^0 + 1$, (2.1) can be rewritten as follows:

$$e^{\mathbf{i}u'Y_t} = \begin{cases} \psi_1^0(u) + \varepsilon_t(u), & \text{for } t = 1, \ldots, T_1^0, \\ \psi_2^0(u) + \varepsilon_t(u), & \text{for } t = T_1^0 + 1, \ldots, T_2^0, \\ \quad\vdots & \quad\vdots \\ \psi_{M^0+1}^0(u) + \varepsilon_t(u), & \text{for } t = T_{M^0}^0 + 1, \ldots, T. \end{cases}$$

We now discuss how to estimate the unknown break fractions $\{r_j^0\}_{j=1}^{M^0}$. Suppose $\psi_j^0(u) \neq \psi_k^0(u), j \neq k$ for u in some nonnegligible subset of $\mathbb{R}^d$ and for $j,k = 1, \ldots, M^0 + 1$. Then for each M^0-partition $\{r_j\}_{j=1}^{M^0}$, where $r_j = T_j/T, j = 1, 2, \ldots, M^0$, we can estimate the CF for Y_t, $t \in [T_{j-1}+1, T_j]$, by the ECF based on the jth subsample:

$$\tilde{\psi}_j(u) \equiv \frac{1}{T_j - T_{j-1}} \sum_{t=T_{j-1}+1}^{T_j} e^{\mathbf{i}u'Y_t}.$$

By the law of large numbers, $\tilde{\psi}_j(u)$ converges almost surely to $\psi_j^0(u)$ as $T_j - T_{j-1} \to \infty$ and $T_j = T_j^0$ for each $j = 1, \ldots, M^0$ (see, e.g., Feuerverger and Mureika, 1977). Note that the ECF $\tilde{\psi}_j(u)$ depends on the partition $\{r_j\}_{j=1}^{M^0}$. For notational simplicity, however, we continue to write it as $\tilde{\psi}_j(u)$. Similar to Bai (1994, 1997) and Bai and Perron (1998), we impose some restrictions on the possible values of break dates to ensure that the break dates are asymptotically distinct from each other. Specifically, we define a set Π_ϵ^0 for some small positive number $\epsilon > 0$ such that $\Pi_\epsilon^0 = \{\{r_j\}_{j=1}^{M^0} : r_j - r_{j-1} \geq \epsilon, \text{ for } j = 1, \ldots, M^0 + 1\}$, where we follow the convention that $r_0 = 0$ and $r_{M^0+1} = 1$. Then the estimated break fractions $\{\hat{r}_j\}_{j=1}^{M^0}$ solve the following optimization problem:

$$\min_{\{r_1,\ldots,r_{M^0}\} \in \Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \sum_{t=T_{j-1}+1}^{T_j} \int_{\mathbb{R}^d} \left| e^{\mathbf{i}u'Y_t} - \tilde{\psi}_j(u) \right|^2 W(u) \mathrm{d}u, \tag{2.2}$$

where $T_j = \lfloor T r_j \rfloor$ for $j = 1, 2, \ldots, M^0$, and $W(\cdot) : \mathbb{R}^d \to \mathbb{R}^+$ is a nonnegative symmetric weighting function. Compared to Bai and Perron (1998) who minimize the sum of squared residuals in a linear regression, we call the objective function in (2.2) the sum of squared generalized residuals (SSGR) of the generalized regression in (2.1). Note that in order to pin down the break fractions in distribution, we need to consider all $u \in \mathbb{R}^d$. We thus require $W(u)$ to have unbounded support. With a proper choice of weighting function, the objective function can have a closed-form expression by integrating u out, which is a salient feature of our approach. Since the ECF can be directly computed for any $u \in \mathbb{R}^d$ without any measurement error, the discrete sampling issue in functional data analysis is not present in our setting. We do not need to consider discrete grids of u to obtain the objective function in (2.2). Our idea is similar to the continuous ECF method proposed by Knight and Yu (2002). We both avoid choosing discrete grid points by considering a continuous weighting function for the nuisance parameter u. We relegate the discussion of weighting functions to Section 3.

The proposed ECF approach has several appealing features. First, it is applicable for a multivariate time series with relatively large dimension. Unlike the density function-based approaches, the ECF does not require nonparametric smoothing. Second, minimizing the SSGR works naturally well for multiple structural breaks. Indeed, (2.2) is a K-means clustering issue that can determine the heterogeneity in the mean of $e^{\mathbf{i}u'Y_t}$. The key step is to determine the appropriate number of clusters. Third, with a proper choice of weighting function $W(u)$, the objective function in (2.2) can be computed without numerical integration over u. One can then

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

obtain consistent estimates $\{\hat{r}_j\}_{j=1}^{M^0}$ by grid search. Furthermore, our method does not require the existence of the probability density, thus allowing the multivariate time series Y_t or some of its components to be discrete random variables. Last but not the least, the ECF approach does not impose any moment restriction on Y_t. That is useful in detecting distributional changes in high-frequency financial time series data since their higher-order moments may not exist (Loretan and Phillips, 1994).

2.2. Assumptions

To study the consistency and limiting distributions of the estimated break fractions $\{\hat{r}_j\}_{j=1}^{M^0}$, where $\hat{r}_j = \hat{T}_j/T$ for $j = 1,2,\ldots,M^0$, we impose the following regularity conditions.

Assumption A.1. (i) Let $\{Y_t\}_{t=T_{j-1}^0+1}^{T_j^0}, j = 1,\ldots,M^0+1,$ be the partition of the time series sample $\{Y_t\}_{t=1}^T$ at the true break dates $\{T_j^0\}_{j=1}^{M^0}$. For each $j = 1,\ldots,M^0+1$, $\{Y_t\}_{t=T_{j-1}^0+1}^{T_j^0}$ is strictly stationary and (ii) $\{Y_t\}_{t=1}^T$ is strong mixing with mixing coefficient $\alpha(s) = \max_j \alpha_j(s)$ such that $\sum_{s=0}^{\infty}(s+1)^{q/2-1}\alpha(s) < \infty$ for some $q > 2$ and $\delta > 0$, where $\alpha_j(s)$ is the mixing coefficient for the jth subsample.

Assumption A.2. The weighting function $W(\cdot) : \mathbb{R}^d \to \mathbb{R}^+$ is nonnegative, symmetric, and integrable with $\int_{\mathbb{R}^d} \|u\|^4 W(u)\mathrm{d}u < \infty$.

Assumption A.3. The break dates $T_j^0 = \lfloor Tr_j^0 \rfloor$ for $0 < r_1^0 < r_2^0 < \cdots < r_{M^0}^0 < 1$, where $r_0 = 0$ and $r_{M^0+1} = 1$. There exists a small positive number $\epsilon > 0$ such that $\min_{j\in\{1,\ldots,M^0+1\}}(r_j^0 - r_{j-1}^0) \geq \epsilon$.

Assumption A.4. Let $\varepsilon_t(u) = e^{\mathbf{i}u'Y_t} - \psi_j^0(u)$ be the generalized error function for $t \in [T_{j-1}^0+1, T_j^0]$: (i) $\{\varepsilon_t(u)\}_{t=T_{j-1}^0+1}^{T_j^0}$ satisfies $E[\varepsilon_t(u)] = 0$ and the generalized long-run variance function $\Omega^{(j)}(u,v) = \sum_{l=-\infty}^{\infty} \sigma_l^{(j)}(u,v)$ with $\sigma_l^{(j)}(u,v) = E[\varepsilon_t(u)\varepsilon_{t+l}(v)^*]$; (ii) $(T_j^0 - T_{j-1}^0)^{-1/2}\sum_{t=T_{j-1}^0+1}^{T_{j-1}^0+\lfloor(T_j^0-T_{j-1}^0)r\rfloor} \varepsilon_t(u) \Rightarrow B^{(j)}(u,r)$, where $B^{(j)}(u,r)$ is a generalized Brownian motion on $\mathbb{U}\times[0,1]$ with mean zero and covariance kernel $E[B^{(j)}(u,r)B^{(j)}(v,s)^*] = \min\{r,s\}\Omega^{(j)}(u,v)$, and $\mathbb{U} \subset \mathbb{R}^d$ denotes a symmetric compact hypercube around the origin.

Assumption A.1(i) implies that the Y_t's are identically distributed within each subsample $\{Y_t\}_{t=T_{j-1}^0+1}^{T_j^0}$ for $j = 1,\ldots,M^0+1$. We note that such a condition rules out the possibility that some drift arises in Y_t after a break. We have to impose strict stationarity in each subsample to establish our asymptotic theory. Assumption A.1(ii) restricts the degree of temporal dependence in $\{Y_t\}_{t=1}^T$ and is generally adopted in time series analysis. A variety of time series processes,

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

such as autoregressive moving average, bilinear, and autoregressive conditional heteroskedastic (ARCH) processes, satisfy the strong mixing condition (Fan and Li, 1999). This implies that the weak dependence is bounded by the mixing coefficient not only within each subsample $\{Y_t\}_{t=T_{j-1}^0+1}^{T_j^0}$, but also across the whole sample. Assumption A.1 does not require that the probability density of Y_t exists. Therefore, the components of Y_t can be either continuous or discrete random variables or a mixture of both. Furthermore, it does not impose any moment restriction on Y_t. This is appealing since financial time series data often exhibit heavy tails and thus may not have finite higher-order moments.

Assumption A.2 imposes mild conditions on the weighting function $W(u)$, which ensure the existence of the integral in (2.2). Note that $W(u)$ assigns weights to various values of u such that (2.2) can detect breaks of all kinds. Intuitively, the choice of weighting function will affect the finite sample power of the proposed test since the magnitude of structural breaks may vary across u. Hence, one should adopt a weighting function that assigns weights accordingly. However, such an optimal weighting function depends on the distribution of Y_t and the very nature of structural breaks, which are usually unknown *a priori*. A possible remedy is to conduct a two-step approach, where one can get the knowledge of the structure breaks using some prespecified weighting function in the first step and then obtain the data-dependent optimal weighting function in the second step. However, such a procedure is rather involved and beyond the scope of the present paper.

Assumption A.3 requires the break dates to be asymptotically distinct, which is commonly adopted in the literature and vital for our asymptotic theory. Intuitively, if two break dates are too close to each other, one cannot distinguish them since there is not enough information in the corresponding segment. It indicates that the sample size of each segment increases proportionately as T grows, which is essential for the application of the law of large numbers and the central limit theorem. In addition, $r_0 = 0$ and $r_{M^0+1} = 1$ indicate that $r_1^0 \geq \epsilon$ and $r_{M^0}^0 \leq 1-\epsilon$, implying that the breaks are bounded away from the boundaries of the sample.

Assumption A.4 allows for serial dependence of unknown form in the generalized error function. Otherwise, it would imply that $\{Y_t\}_{t=1}^T$ is independent over time. Because for any $T_{j-1}^0 + 1 \leq t \neq s \leq T_j^0$,

$$E\left[\varepsilon_t(u)\varepsilon_s(v)^*\right] = E\left[e^{\mathbf{i}(u'Y_t - v'Y_s)}\right] - \psi_j^0(u)\psi_j^0(v)^* = 0,$$

if and only if $E[e^{\mathbf{i}(u'Y_t - v'Y_s)}] = E(e^{\mathbf{i}u'Y_t})E(e^{-\mathbf{i}v'Y_s})$, which implies independence between Y_t and Y_s. This is of course too restrictive for a time series process. We note that although the generalized error functions $\{\varepsilon_t(u)\}_{t=1}^T$ have the same mean $E[\varepsilon_t(u)] = 0$ across M^0+1 regimes, they are not identically distributed since their second moment depends on the time-varying CF of Y_t. Structural breaks in Y_t induce changes in the covariance structure of $\varepsilon_t(u)$. Therefore, we have to impose a functional central limit theorem for each time segment $[T_{j-1}^0 + 1, T_j^0]$.

2.3. Consistency and Limiting Distribution

We now show the consistency and asymptotic distributions of the estimated break fractions.

THEOREM 2.1. *Suppose Assumptions A.1–A.3 hold. Then as* $T \to \infty$,

$$\hat{r}_j \xrightarrow{p} r_j^0, \quad j = 1, 2, \ldots, M^0,$$

where $\{\hat{r}_j\}_{j=1}^{M^0}$ *are the estimated break fractions that solve (2.2).*

Theorem 2.1 implies that the solution to (2.2) consistently estimates the true break fractions $\{r_j^0\}_{j=1}^{M^0}$. This result is quite similar to Bai and Perron's (1998) Proposition 1 on estimation for break fractions in a linear regression model. Next theorem provides convergence rate of the estimated break fractions.

THEOREM 2.2. *Under Assumptions A.1–A.3, for every* $\eta > 0$, *there exists a* $\delta \in (0, \infty)$, *such that for* $T \to \infty$,

$$P\left[\left|T\left(\hat{r}_j - r_j^0\right)\right| > \delta\right] < \eta, \quad j = 1, 2, \ldots, M^0.$$

We note that the rate-T convergence in probability holds for the estimated break fraction $\hat{r}_j$, rather than the estimated break date $\hat{T}_j$. Our result is also equivalent to $P(|\hat{T}_j - T_j| > \delta) \leq \eta$ for each $j = 1, 2, \ldots, M^0$. This implies that the deviation of any single estimated break date from the true break date is bounded by some constant δ that is independent of T.

Given the consistency of the estimated break fractions, we further obtain the standard root-T asymptotic normality of the feasible ECF $\hat{\psi}_j(u)$, where

$$\hat{\psi}_j(u) \equiv \frac{1}{\hat{T}_j - \hat{T}_{j-1}} \sum_{t=\hat{T}_{j-1}+1}^{\hat{T}_j} e^{\mathbf{i}u'Y_t}.$$

THEOREM 2.3. *Suppose Assumptions A.1–A.4 hold. Then as* $T \to \infty$,

$$\sqrt{T}\left[\hat{\psi}_j(u) - \psi_j^0(u)\right] \Rightarrow (r_j^0 - r_{j-1}^0)^{-1/2} B^{(j)}(u, 1), \; j = 1, 2, \ldots, M^0 + 1,$$

for $u \in \mathbb{U}$.

Theorem 2.3 implies that treating $\hat{r}_j$ as the true parameter r_j^0 does not affect the asymptotic distribution of the feasible ECF because of the rate-T convergence of the estimated break fractions. Furthermore, based on the feasible ECF, we can also obtain the corresponding estimated density function (e.g., Shephard, 1991). Specifically, if Y_t is a continuous random vector and its probability density function

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

exists, then a consistent density estimator at the partition $[\hat{T}_{j-1}+1, \hat{T}_j]$ can be obtained by the inverse Fourier transform of $\hat{\psi}_j(u)$, i.e.,

$$\hat{f}_{Y_j}(y) = \frac{1}{2\pi}\int_{\mathbb{R}^d} e^{-\mathbf{i}u'y}\hat{\psi}_j(u)\mathrm{d}u,\ y\in\mathbb{R}^d.$$

We now show the asymptotic distribution of the estimated break fractions $\{\hat{r}_j\}_{j=1}^{M^0}$, which is useful for constructing confidence intervals of the break dates. We assume that the magnitudes of the breaks converge to zero as the sample size increases. As pointed out by Bai (1994, 1997), when the magnitudes of the shifts are nonzero constants that are independent of T, the limiting distribution of $\{\hat{r}_j\}_{j=1}^{M^0}$ is highly data-dependent and difficult to obtain. In addition, when the magnitude of structural breaks is large, estimation of the break dates is quite precise as the sample size increases. Thus, it is more meaningful to construct confidence intervals for small breaks.

Let $\delta_{T,j}(y) = \mathbf{F}_{j+1}^0(y) - \mathbf{F}_j^0(y) = v_T\delta_j(y)$ be the magnitude of the break for the CDFs between the $(j+1)$th regime and the jth regime, where $v_T \to 0$ is a scalar satisfying $v_T T^{1/2} \to \infty$, and $\delta_j(y)$ is absolutely continuous and independent of sample size T. Then, by Fourier transform, we have

$$\Delta_{T,j}(u) = \psi_{j+1}^0(u) - \psi_j^0(u) = v_T\Delta_j(u),$$

where $\Delta_j(u) \equiv \int_{\mathbb{R}^d} e^{\mathbf{i}u'y}\mathrm{d}\delta_j(y)$. To establish the limiting distribution of the estimated break fractions, we impose the following condition.

Assumption A.5. (i) $\max_j \sup_{u\in\mathbb{R}^d}|\Delta_j(u)| < \infty$ and (ii) let $\mathcal{G}^{(j)}(u,\eta)$ denote a two-sided generalized Brownian motion defined on $\mathbb{U}\times[-C,C]$ for some constant $C>0$, such that $\mathcal{G}^{(j)}(u,\eta) = \mathcal{G}_1^{(j)}(u,\eta)$ when $\eta\geq 0$ and $\mathcal{G}^{(j)}(u,\eta) = \mathcal{G}_2^{(j)}(u,-\eta)$ when $\eta<0$, where $\mathcal{G}_1^{(j)}(u,\eta)$ and $\mathcal{G}_2^{(j)}(u,\eta)$ are generalized Brownian motions with the following covariance kernels $E\left[\mathcal{G}_1^{(j)}(u,\eta_1)\mathcal{G}_1^{(j)}(v,\eta_2)^*\right] = \min\{\eta_1,\eta_2\}\Omega^{(j+1)}(u,v)$ when $\eta_1,\eta_2\geq 0$, and $E\left[\mathcal{G}_2^{(j)}(u,-\eta_1)\mathcal{G}_2^{(j)}(v,-\eta_2)^*\right] = \min\{-\eta_1,-\eta_2\}\Omega^{(j)}(u,v)$, when $\eta_1,\eta_2<0$. Here, $\Omega^{(j)}(u,v)$ is defined in Assumption A.4(ii).

Assumption A.5(i) restricts the magnitude of each break to be finite. The process $\mathcal{G}^{(j)}(u,\eta)$ defined in Assumption A.5(ii) is analogous to the two-sided Brownian motion in Bai (1994). Both conditions are commonly adopted to derive the asymptotic distribution of the estimated fractions. Meanwhile, given Assumptions A.4 and A.5(ii), $v_T\sum_{t=T_{j-1}^0+1}^{T_{j-1}^0+\lfloor\eta v_T^{-2}\rfloor}\varepsilon_t(u) \Rightarrow \mathcal{G}^{(j)}(u,\eta)$ in $\mathbb{U}\times[-C,C]$ when $v_T^2T\to\infty$ by invariance principle.

THEOREM 2.4. *Suppose Assumptions A.1–A.5 hold. Let* $\psi_{j+1}^0(u) - \psi_j^0(u) = v_T\Delta_j(u)$ *be the magnitude of the break for the CF from the jth regime to the*

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

$(j+1)$th regime. Then for any $v_T = T^{-a}$ with $a \in (0, \frac{1}{2})$, and $j = 1, 2, \ldots, M^0$, as $T \to \infty$,

$$Tv_T^2(\hat{r}_j - r_j^0) \xrightarrow{d} \arg\max_{\eta} \Lambda^{(j)}(\eta)$$

and

$$\text{SSGR}_{M^0}(T_j^0) - \text{SSGR}_{M^0}(T_j^0 + \lfloor \eta v_T^{-2} \rfloor) \Rightarrow \Lambda^{(j)}(\eta),$$

where $\text{SSGR}_{M^0}(T_j^0 + \lfloor \eta v_T^{-2} \rfloor)$ and $\text{SSGR}_{M^0}(T_j^0)$ denote the sum of squared generalized residuals based on partition $\{T_1^0, \ldots, T_j^0 + \lfloor \eta v_T^{-2} \rfloor, \ldots, T_{M^0}^0\}$ and the true break dates $\{T_1^0, \ldots, T_{M^0}^0\}$, respectively, and

$$\Lambda^{(j)}(\eta) = \begin{cases} 2\int_{\mathbb{R}^d} Re\left\{\mathcal{G}^{(j)}(u,\eta)\Delta_j(u)^*\right\} W(u)du - |\eta| \int_{\mathbb{R}^d} \left|\Delta_j(u)\right|^2 W(u)du, & \text{if } \eta < 0, \\ -2\int_{\mathbb{R}^d} Re\left\{\mathcal{G}^{(j)}(u,\eta)\Delta_j(u)^*\right\} W(u)du - |\eta| \int_{\mathbb{R}^d} \left|\Delta_j(u)\right|^2 W(u)du, & \text{if } \eta > 0. \end{cases}$$

The limiting distribution can be viewed as a generalization of Bai and Perron's (1998) results to dependent functional data process with multiple breaks. We note that analogous results have been obtained by Berkes et al. (2009) for an independent functional data process and Aue et al. (2018) for a dependent functional data process with a single break.

Note that the limiting distribution for each $\hat{r}_j$ only depends on its break magnitude characterized by $\Lambda^{(j)}(\eta)$. In general, it is difficult to tabulate the asymptotic critical values. To construct the confidence intervals for the estimated break fractions, one can use a suitable bootstrap procedure. Alternatively, one can standardize the estimator by estimating the long-run variance of $\mathcal{G}^{(j)}(u,\eta)$ as illustrated by Aue et al. (2018).

3. TESTS FOR STRUCTURAL BREAKS IN DISTRIBUTION

3.1. A Joint Test Against a Fixed Number of Breaks

In this subsection, we consider testing the null hypothesis of no structural breaks against the alternative hypothesis of M structural breaks in distribution. Given a prespecified number of breaks M, we test

$$\mathbb{H}_0 : \phi_t^0(u) = \phi^0(u), \text{ for all } u \in \mathbb{R}^d,$$

where $\phi^0(u)$ is a time-invariant CF, against

$$\mathbb{H}_A : \phi_t^0(u) = \begin{cases} \psi_1(u), & \text{for } t = 1, \ldots, T_1, \\ \psi_2(u), & \text{for } t = T_1 + 1, \ldots, T_2, \\ \vdots & \vdots \\ \psi_{M+1}(u), & \text{for } t = T_M + 1, \ldots, T, \end{cases}$$

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

for u in some nonzero Borel measurable subset of $\mathbb{R}^d$, some partition $\{T_j\}_{j=1}^{M}$, and corresponding collection of CFs $\{\psi_j(u)\}_{j=1}^{M+1}$. To test $\mathbb{H}_0$ against $\mathbb{H}_A$, we construct a generalized sup-F test statistic by comparing the SSGRs between the restricted model (under $\mathbb{H}_0$) and the unrestricted model (under $\mathbb{H}_A$).

We denote the SSGR under $\mathbb{H}_A$ with M structural breaks at $\{T_j\}_{j=1}^{M}$ as

$$\mathrm{SSGR}_M(r_1,\ldots,r_M) = \sum_{j=1}^{M+1} \sum_{t=T_{j-1}+1}^{T_j} \int_{\mathbb{R}^d} \left| e^{\mathbf{i}u'Y_t} - \tilde{\psi}_j(u) \right|^2 W(u)\mathrm{d}u,$$

and the SSGR under $\mathbb{H}_0$ as

$$\mathrm{SSGR}_0 = \sum_{t=1}^{T} \int_{\mathbb{R}^d} \left| e^{\mathbf{i}u'Y_t} - \tilde{\psi}(u) \right|^2 W(u)\mathrm{d}u,$$

where $\tilde{\psi}(u) = T^{-1}\sum_{t=1}^{T} e^{\mathbf{i}u'Y_t}$. Under $\mathbb{H}_0$, the ECFs $\tilde{\psi}_j(u)$ and $\tilde{\psi}(u)$ should both converge to the true time-invariant CF $\phi^0(u)$ for any prespecified number of breaks M and any partition $\{T_j\}_{j=1}^{M}$. As a result, $\mathrm{SSGR}_M(r_1,\ldots,r_M)$ should be close to SSGR_0 under $\mathbb{H}_0$. Nevertheless, under $\mathbb{H}_A$, the probability limit of $\mathrm{SSGR}_M(r_1,\ldots,r_M)$ will deviate from that of the SSGR_0 under certain collections of break dates $\{T_j\}_{j=1}^{M}$. Our test statistic is thus given by

$$\sup F = \sup_{\{r_1,\ldots,r_M\}\in\Pi_\epsilon} F_T(r_1,\ldots,r_M), \tag{3.1}$$

where $\Pi_\epsilon = \{\{r_j\}_{j=1}^{M} : r_j - r_{j-1} \geq \epsilon \text{ for } j=1,\ldots,M+1\}$ for some small $\epsilon > 0$, and

$$F_T(r_1,\ldots,r_M) = \mathrm{SSGR}_0 - \mathrm{SSGR}_M(r_1,\ldots,r_M). \tag{3.2}$$

We note that the test based on (3.1) does not require that the true number of breaks M^0 is known *a priori*. When $M = M^0$, the partition under $\mathbb{H}_A$ will coincide with the true break dates $\{T_j^0\}_{j=1}^{M^0}$. When $M < M^0$, $\{T_j\}_{j=1}^{M}$ will be a collection of M breaks in $\{T_j^0\}_{j=1}^{M^0}$ (e.g., Bai and Perron, 1998). Even when $M > M^0$, it still holds that SSGR_0 will deviate from $\mathrm{SSGR}_M(r_1,\ldots,r_M)$ for a certain collection of break dates $\{T_j\}_{j=1}^{M}$ when it contains (part of) $\{T_j^0\}_{j=1}^{M^0}$. This guarantees the consistency of our sup-F test against an unknown number of breaks.

3.2. Asymptotic Distribution

To derive the asymptotic distribution of the generalized sup-F test statistic in (3.1), we impose the following condition analogously to Assumption A.4.

Assumption A.4*. (i) $\{\varepsilon_t(u)\}_{t=1}^{T}$ satisfies $E[\varepsilon_t(u)] = 0$ and the generalized long-run variance function $\Omega(u,v) = \sum_{l=-\infty}^{\infty} \sigma_l(u,v)$, where $\sigma_l(u,v) = E[\varepsilon_t(u)\varepsilon_{t+l}(v)^*]$ and (ii) $T^{-1/2}\sum_{t=1}^{\lfloor Tr \rfloor} \varepsilon_t(u) \Rightarrow B(u,r)$, where $B(u,r)$ is a generalized Brownian

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

motion on $\mathbb{U} \times [0,1]$ with mean zero and covariance kernel $E[B(u,r)B(v,s)^*] = \min\{r,s\}\Omega(u,v)$, and $\mathbb{U}$ is a compact set defined in Assumption A.4.

Assumption A.4* can be regarded as a special case of Assumption A.4. Under $\mathbb{H}_0$, the covariance kernel of $\varepsilon_t(u)$ is the same across the whole sample. Hence, the functional central limit theorem stated in Assumption A.4(ii) degenerates to A.4*(ii). Note that Assumption A.4* allows for serial dependence of unknown form in the generalized error functions.

THEOREM 3.1. *Suppose Assumptions A.1, A.2, and A.4* hold. Let* $\mathcal{B}(u,r) \equiv B(u,r) - rB(u,1)$ *be a generalized Brownian bridge. Under* $\mathbb{H}_0$*, as* $T \to \infty$*,*

$$\sup F \xrightarrow{d} \sup_{\{r_1,\ldots,r_M\}\in\Pi_\epsilon} F(r_1,\ldots,r_M),$$

where

$$F(r_1,\ldots,r_M) = \sum_{j=1}^{M+1} \frac{1}{r_j - r_{j-1}} \int_{\mathbb{R}^d} \left|\mathcal{B}(u,r_j) - \mathcal{B}(u,r_{j-1})\right|^2 W(u)\mathrm{d}u.$$

Our generalized sup-F test is similar to the sup-F test in Bai and Perron (1998), except that the asymptotic distribution for our test statistic is not pivotal since it depends on the unknown data generating process. As revealed by Assumption A.4*(ii), the distribution of the generalized error function $\varepsilon_t(u)$ depends on u. To obtain an asymptotically pivotal test, we have to impose further restrictions on the second moment of $\varepsilon_t(u)$ so that we can integrate out u in the limiting distribution. However, such restrictions will affect the alternative hypothesis since the second moment of $\varepsilon_t(u)$ depends on structural breaks in Y_t. Therefore, we do not standardize our test statistic and its limiting distribution is data-dependent. We need to use some resampling methods to obtain the critical values in finite samples. In contrast, by assuming homogeneity in regression errors (Assumptions A8 and A9), Bai and Perron's (1998) test statistic is asymptotically pivotal. Despite being not asymptotically pivotal, our SSGR-based test statistic is easy to compute, since SSGR_M in the sup-F statistic is the value of the objective function specified by (2.2) at the optimal estimated break dates $\{\hat{T}_j\}_{j=1}^M$. Interestingly, by taking the derivatives of the ECF with respect to u at the origin, our test can detect structural breaks at various moments if they exist. We will further elaborate this point in Section 5.

We conclude this subsection by a local power analysis. Consider a class of local alternatives:

$$\mathbb{H}_A(a_T) : \mathbf{F}_j^0(y) = \mathcal{F}^0(y) + a_T \vartheta_j(y),$$

where $a_T \to 0$ as $T \to \infty$. $\mathcal{F}^0(y)$ is a time-invariant CDF of Y_t, $\vartheta_j(y)$ captures the deviation of the jth regime from $\mathcal{F}^0(y)$, and the rate a_T controls the speed at which the local alternative $\mathbb{H}_A(a_T)$ converges to the null hypothesis. Given the

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

transformation between CDF and its CF, the above local alternative is equivalent to the following representation

$\mathbb{H}_A(a_T): \psi_j^0(u) = \phi^0(u) + a_T\theta_j(u),$

where $\phi^0(u)$ is the time-invariant CF of Y_t and $\theta_j(u) = \int_{\mathbb{R}^d} e^{\mathbf{i}u'y} \mathrm{d}\vartheta_j(y)$.

THEOREM 3.2. *Suppose Assumptions A.1–A.4 hold. Then under $\mathbb{H}_A(a_T)$ with $a_T = T^{-1/2}$, as $T \to \infty$,*

$$\sup F \xrightarrow{d} \sup_{\{r_1,\ldots,r_M\}\in\Pi_\epsilon} F^A(r_1,\ldots,r_M),$$

where

$$F^A(r_1,\ldots,r_M) = \sum_{j=1}^{M+1} \frac{1}{r_j - r_{j-1}} \int_{\mathbb{R}^d} \left|G(u,r_j) - G(u,r_{j-1}) + \Gamma(u,r_j) - \Gamma(u,r_{j-1})\right|^2 W(u)\mathrm{d}u,$$

with $G(u,r_j) = \left[\sum_{k=1}^{l}(r_k^0 - r_{k-1}^0)^{1/2}B^{(k)}(u,1) + (r_{l+1}^0 - r_l^0)^{1/2}B^{(l+1)}\left(u, \frac{r_j - r_l^0}{r_{l+1}^0 - r_l^0}\right)\right] - r_j\left[\sum_{k=1}^{M^0+1}(r_k^0 - r_{k-1}^0)^{1/2}B^{(k)}(u,1)\right]$, *and* $\Gamma(u,r_j) = \left[\sum_{k=1}^{l}(r_k^0 - r_{k-1}^0)\theta_k(u) + (r_j - r_l^0)\theta_{l+1}(u)\right] - r_j\left[\sum_{k=1}^{M^0+1}(r_k^0 - r_{k-1}^0)\theta_k(u)\right]$.

The asymptotic distribution shown by Theorem 3.2 involves two processes indexed by u and r_j. The process $G(u,r_j)$ is random and it is the limiting process of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T_j}\varepsilon_t(u) - r_j\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t(u)\right]$. Under $\mathbb{H}_0$, $G(u,r_j)$ is a generalized Brownian bridge and is identical to $\mathcal{B}(u,r_j)$. However, under $\mathbb{H}_A(a_T)$, the generalized error functions are different across subsamples specified by the true break dates $\{T_j^0\}_{j=1}^{M^0}$. Suppose T_j lies in the $(l+1)$th subsample specified by the true break dates, i.e., $r_l^0 < r_j < r_{l+1}^0$. Then, by Assumption A.4, we have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T_j}\varepsilon_t(u) = \sum_{k=1}^{l}\left[\frac{1}{\sqrt{T}}\sum_{t=T_{k-1}^0+1}^{T_k^0}\varepsilon_t(u)\right] + \frac{1}{\sqrt{T}}\sum_{t=T_l^0+1}^{T_j}\varepsilon_t(u)$$
$$\Rightarrow \sum_{k=1}^{l}(r_k^0 - r_{k-1}^0)^{1/2}B^{(k)}(u,1) + (r_{l+1}^0 - r_l^0)^{1/2}B^{(l+1)}\left(u, \frac{r_j - r_l^0}{r_{l+1}^0 - r_l^0}\right),$$

where the first term is a weighted sum of the limiting processes in the first l subsamples and the second term is the limiting process in the $(l+1)$th subsample. The process $\Gamma(u,r_j)$ is nonrandom, contributed by the local alternative component $a_T\theta_k(u)$. It shares the same structure as the process $G(u,r_j)$. Compared to the existing literature that usually assumes that the limiting behavior of $\varepsilon_t(u)$ is invariant to structural breaks in mean of a functional time series, our asymptotic results under the local alternative have to take into account the underlying true break dates.

Theorem 3.2 does not require that the number of breaks specified in the test is identical to the true number of breaks. Even when one fails to pin down the correct number of breaks, Theorem 3.2 still holds. This implies that our test is robust to misspecification of the number of breaks. Theorem 3.2 also shows that our test can detect $\mathbb{H}_A(a_T)$ at the parametric rate $a_T = T^{-1/2}$, which is the same as Bai and Perron's (1998) test for breaks in conditional mean and Inoue's (2001) test for a single break in distribution. Moreover, unlike Kapetanios' (2009) test, which requires smoothed nonparametric estimation for probability density functions, the convergence rate of our test does not depend on the dimension of Y_t. As a result, our test avoids the "curse of dimensionality" problem when the dimension of Y_t is large.

3.3. Weighting Function

The weighting function $W(u)$ plays a vital role in estimating and testing for structural breaks in our framework. On the one hand, an appropriate choice of weighting function can improve the finite sample power performance of the test for structural breaks since the magnitude of structural breaks in the CF varies across u. On the other hand, when the dimension of Y_t is large, a weighting function that can deliver a closed-form expression of SSGRs is highly desirable from a computational point of view. In this subsection, we discuss some weighting functions that can meet this requirement.

We follow Hong et al. (2017) and consider the following normal weighting function:

$$W(u) = (2\pi b)^{-\frac{d}{2}} e^{-\frac{\|u\|^2}{2b}}, \tag{3.3}$$

where $b > 0$ is a scaling parameter. With this weighting function, (2.2) has the following closed-form expression:

$$\max_{\{r_1,\ldots,r_{M^0}\}\in\Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \frac{1}{T_j - T_{j-1}} \sum_{s=T_{j-1}+1}^{T_j} \sum_{r=T_{j-1}+1}^{T_j} e^{-\frac{b\|Y_s - Y_r\|^2}{2}}.$$

Such an expression can be regarded as a weighted generalized distance between observations. Intuitively, the generalization from the distance between Y_s and Y_r to their exponential squared distance can capture structural breaks in all higher-order moments of Y_t if they exist. For more discussion on this type of weighting function, see Hušková and Meintanis (2008).

Besides the normal weighting function, one can also use the product Laplace$(0,b)$ weighting function:

$$W(u) = (2b)^{-d} e^{-\frac{\sum_{i=1}^{d} |u_i|}{b}}. \quad \textbf{(3.4)}$$

With this weighting function, (2.2) becomes

$$\max_{\{r_1,\ldots,r_{M^0}\}\in\Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \frac{1}{T_j - T_{j-1}} \sum_{s=T_{j-1}+1}^{T_j} \sum_{r=T_{j-1}+1}^{T_j} \prod_{i=1}^{d} \frac{1}{1+b^2|Y_{is}-Y_{ir}|^2},$$

where Y_{is} is the ith entry of the $d \times 1$ vector Y_s for $i = 1,\ldots,d$.

Furthermore, we can follow Bierens and Wang (2012) and consider the uniform weighting function

$$W(u) = \frac{1}{(2c)^d} \qquad \text{for } u \in \mathbb{U} = [-c,c]^d \text{ with some } c > 0. \quad \textbf{(3.5)}$$

Then, (2.2) has the following closed-form expression:

$$\max_{\{r_1,\ldots,r_{M^0}\}\in\Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \frac{1}{T_j - T_{j-1}} \sum_{s=T_{j-1}+1}^{T_j} \sum_{r=T_{j-1}+1}^{T_j} \prod_{i=1}^{d} \frac{\sin(c(Y_{is}-Y_{ir}))}{c(Y_{is}-Y_{ir})}.$$

Since $\lim_{x\to 0} \frac{\sin x}{x} = 1$, we can replace $\frac{\sin(c(Y_{is}-Y_{ir}))}{c(Y_{is}-Y_{ir})}$ with 1 when $Y_{is} = Y_{ir}$. We note that when Y_t has unbounded support, the uniform weighting function may lead to power loss. However, by choosing a suitable value of c, one can adjust the test statistic for various data generating processes. For more discussion, see Bierens and Wang (2012).

3.4. A Resampling Procedure

Theorem 3.1 shows that the null asymptotic distribution of the proposed test is not pivotal because it depends on the unknown data generating process. To obtain the critical values of the proposed test, we propose the following moving block bootstrap procedure (e.g., Künsch, 1989).

(i) Given the data set $\{Y_t\}_{t=1}^{T}$ and a prespecified number of breaks M, compute SSGR_0 and $\text{SSGR}_M(\hat{r}_1,\ldots,\hat{r}_M)$, where $\{\hat{r}_j\}_{j=1}^{M}$ is the collection of estimated break fractions that solves (2.2). Then the test statistic $\sup F = \text{SSGR}_0 - \text{SSGR}_M(\hat{r}_1,\ldots,\hat{r}_M)$.

(ii) Pick a block length $1 < l_T < T$ such that $l_T^{-1} + l_T T^{-1/2} = o(1)$, and construct $N \equiv T - l_T + 1$ sets of block data $\{\mathcal{Y}_n\}_{n=1}^{N}$, where $\mathcal{Y}_n = \{Y_n,\ldots,Y_{n+l_T-1}\}$ is a block data set with length l_T.

(iii) Assuming $Kl_T = T$, draw i.i.d. integer-valued random variables $I_1,\ldots,I_K$ such that I_k, $k = 1,\ldots,K$, follows a discrete uniform distribution that assigns the probability $1/N$ to each value in the set $\{1,\ldots,N\}$. Construct a bootstrap data set $\Phi^\star = \{\mathcal{Y}_1^\star,\ldots,\mathcal{Y}_K^\star\}$, where $\mathcal{Y}_k^\star = \mathcal{Y}_{I_k}$ for $k = 1,\ldots,K$.

(iv) Using the bootstrap data set $\Phi^\star$, compute the bootstrap test statistic $\sup F^\star = \text{SSGR}_0^\star - \text{SSGR}_M^\star(\hat{r}_1^\star,\ldots,\hat{r}_M^\star)$, where $\text{SSGR}_0^\star$ and $\text{SSGR}_M^\star(\hat{r}_1^\star,\ldots,\hat{r}_M^\star)$ are the SSGRs under no structural breaks and under the estimated M breaks $\{\hat{r}_j^\star\}_{j=1}^M$, respectively.

(v) Repeat steps (iii) and (iv) for a total of $\mathbf{B}$ times to obtain a collection of $\mathbf{B}$ bootstrap test statistics $\{\sup F_b^\star\}_{b=1}^{\mathbf{B}}$. Then the bootstrap p-value for the proposed test is given by

$$p_J^{\mathbf{B}} = \frac{1}{\mathbf{B}}\sum_{b=1}^{\mathbf{B}} \mathbb{I}\left(\sup F \leq \sup F_b^\star\right),$$

where $\mathbb{I}(\cdot)$ is an indicator function.

To establish validity of the proposed moving block bootstrap, we impose the following condition.

Assumption A.6. (i) The block length l_T satisfies that $l_T^{-1} + l_T T^{-1/2} = o(1)$ and (ii) $\max_t E\|Y_t\|_q < \infty$ for some $q > 2$, where $\|\cdot\|_q$ denotes the L_q-norm.

Assumption A.6(i) imposes some regularity conditions on the block length. It requires that the block length increases to infinity as the sample size T grows but at a slower speed. Assumption A.6(ii) imposes a moment condition on the time series process $\{Y_t\}$. We note that such a moment condition is used to ensure the validity of the proposed resampling method.

THEOREM 3.3. *Suppose Assumptions A.1–A.3, A.4*, and A.6 hold. Then under* $\mathbb{H}_0$, $F_T^\star(r_1,\ldots,r_M) \Rightarrow^\star F(r_1,\ldots,r_M)$ *in probability as* $T \to \infty$. *Under* $\mathbb{H}_A(a_T), P^\star(\sup F > \sup F^\star) \to 1$, *provided that* $T^{1/2}a_T l_T^{-1/2} \to \infty$ *as* $T \to \infty$. *Here,* $\Rightarrow^\star$ *and* $P^\star$ *denote the weak convergence and probability under the bootstrap probability measure conditional on the observed time series sample* $\{Y_t\}_{t=1}^T$.

Theorem 3.3 shows that the proposed moving block bootstrap provides an asymptotic valid approximation to the limiting null distribution of the generalized sup-F test statistic. While under $\mathbb{H}_A(a_T)$, the sup-F test statistic will diverge to infinity at the speed of Ta_T^2. Given that the bootstrap statistic diverges at the speed of l_T under $\mathbb{H}_A$, the asymptotic validity of the proposed resampling method is guaranteed. We note that under a nonconverging global alternative, the order of magnitude of our sup-F test statistic is $O_P(T)$. We will examine the finite sample performance of this moving block bootstrap procedure in Section 6.

4. DETERMINING THE NUMBER OF BREAKS

When the generalized sup-F test in (3.1) rejects $\mathbb{H}_0$, it is crucial to determine the true number of breaks M^0. We now introduce a sequential testing procedure and an information criterion to estimate M^0.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

4.1. A Sequential Test

We first propose a sequential test of the null hypothesis of M(i.e., $M^0 = M$) breaks against the alternative hypothesis that an additional break exists. Through such a sequential search, we can pin down the true number of breaks and obtain consistent estimates for break fractions.

The idea of sequential testing is similar to the proposed generalized sup-F test. We compare the difference between the SSGRs under the null hypothesis of M breaks and under the alternative of $M+1$ breaks. They should converge to the same limit under the null hypothesis but will diverge under the alternative. We first obtain the estimated break dates $\{\hat{T}_j\}_{j=1}^M$ by solving (2.2), and denote the corresponding SSGR as $\mathrm{SSGR}_M(\hat{T}_1,\ldots,\hat{T}_M)$. Suppose there exists an additional break. We can then estimate it by minimizing SSGR conditioning on the existing break dates $\{\hat{T}_j\}_{j=1}^M$:

$$\min_{1\le j\le M+1}\ \inf_{\tau\in\Lambda_{j,\epsilon}} \mathrm{SSGR}_{M+1}(\hat{T}_1,\ldots,\hat{T}_{j-1},\tau,\hat{T}_j,\ldots,\hat{T}_M),$$

where

$$\Lambda_{j,\epsilon} = \left\{\tau : \hat{T}_{j-1} + \lfloor(\hat{T}_j-\hat{T}_{j-1})\epsilon\rfloor \le \tau \le \hat{T}_j - \lfloor(\hat{T}_j-\hat{T}_{j-1})\epsilon\rfloor\right\}$$

for some pre-specified $\epsilon > 0$. In fact, as Bai and Perron (1998) point out, minimizing SSGR by searching an additional break conditioning on the existing M breaks is equivalent to minimizing SSGR under $M+1$ unknown breaks. Hence, the test statistic is defined as

$$F_T(M+1|M) = \mathrm{SSGR}_M(\hat{T}_1,\ldots,\hat{T}_M) - \min_{1\le j\le M+1}\ \inf_{\tau\in\Lambda_{j,\epsilon}} \mathrm{SSGR}_{M+1}(\hat{T}_1,\ldots,\hat{T}_{j-1},\tau,\hat{T}_j,\ldots,\hat{T}_M).$$

Following the convention that $\hat{T}_0 = 0$ and $\hat{T}_{M+1} = T$, we note that $\mathrm{SSGR}_{M+1}(\hat{T}_1,\ldots,\hat{T}_{j-1},\tau,\hat{T}_j,\ldots,\hat{T}_M)$ is understood as $\mathrm{SSGR}_{M+1}(\tau,\hat{T}_1,\ldots,\hat{T}_M)$ for $j=1$, and as $\mathrm{SSGR}_{M+1}(\hat{T}_1,\ldots,\hat{T}_M,\tau)$ for $j = M$. They correspond to the cases when the additional break is sought in the boundary regions.

Under the null hypothesis of M breaks, the generalized error functions $\{\varepsilon_t(u)\}_{t=1}^T$ are not identically distributed, because their second moment depends on the CF of Y_t. Therefore, we shall establish the asymptotic distribution of the sequential test statistic under Assumption A.4, which imposes a functional central limit theorem for each time segment $[T_{j-1}^0+1, T_j^0]$.

THEOREM 4.1. *Suppose Assumptions A.1–A.4 hold. Let* $\mathcal{B}^{(j)}(u,r) \equiv B^{(j)}(u,r) - rB^{(j)}(u,1)$ *be a generalized Brownian bridge, where* $B^{(j)}(u,r)$ *is as defined in Assumption A.4(ii). Then under the null hypothesis of M structural breaks, as*

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

$T \to \infty$,

$$F_T(M+1|M) \xrightarrow{d} F(M+1|M),$$

where

$$F(M+1|M) = \max_{1 \le j \le M+1} \sup_{\epsilon \le r \le 1-\epsilon} \int_{\mathbb{R}^d} \frac{\left|\mathcal{B}^{(j)}(u,r)\right|^2}{r(1-r)} W(u)du.$$

Due to the heterogeneity in the Brownian bridge $\mathcal{B}^{(j)}(u,r)$ across each time segment $[T^0_{j-1}+1, T^0_j]$, the asymptotic distribution for our sequential test statistic differs from that of Bai and Perron (1998). For details, see Proposition 7 of Bai and Perron (1998).

Suppose the true number of breaks is greater than the pre-specified number of breaks M. Then there must be at least one break that is not consistently estimated. Hence, at least one segment contains an unidentified break point. According to Theorem 3.2, for this segment, the sup-F test statistic in (3.1) will diverge to infinity as the sample size increases. As a result, the test statistic $F_T(M+1|M)$ will also diverge to infinity at the parametric rate, ensuring the consistency of our sequential test.

The limiting distribution of the sequential test statistic depends on the unknown data generating process. We need to use resampling methods to obtain its critical values. We then propose the following moving block bootstrap for the sequential test. Let $1 < l^j_T < \hat{T}_j - \hat{T}_{j-1}$ be a block length for the jth subsample such that $l^j_T \to \infty$ and $l^j_T = o(T^{1/2})$, for $j = 1, \ldots, M+1$.

(i) Given the data set $\{Y_t\}_{t=1}^T$ and the hypothesized number of breaks M, estimate the break fractions $\{\hat{r}_j\}_{j=1}^M$ and compute the test statistic $F_T(M+1|M)$.

(ii) For each subsample $\{Y_t\}_{t=\hat{T}_{j-1}+1}^{\hat{T}_j}$ specified by the estimated break dates, pick a block length $1 < l^j_T < \hat{T}_j - \hat{T}_{j-1}$ such that $(l^j_T)^{-1} + l^j_T T^{-1/2} = o(1)$, and construct $N_j \equiv \hat{T}_j - \hat{T}_{j-1} - l^j_T + 1$ sets of block data $\{\mathcal{Y}^j_n\}_{n=1}^{N_j}$, where $\mathcal{Y}^j_n = \{Y_{\hat{T}_{j-1}+n}, \ldots, Y_{\hat{T}_{j-1}+n+l^j_T-1}\}$ is a block data set with length l^j_T.

(iii) Assuming $K_j l^j_T = \hat{T}_j - \hat{T}_{j-1}$, for $j = 1, \ldots, M+1$, draw i.i.d. integer-valued random variables $I^j_1, \ldots, I^j_{K_j}$ such that I_k, $k = 1, \ldots, K_j$, follows a discrete uniform distribution that assigns the probability $1/N_j$ to each value in the set $\{1, \ldots, N_j\}$. Construct a bootstrap data set $\Phi^\star_j = \{\mathcal{Y}^{j,\star}_1, \ldots, \mathcal{Y}^{j,\star}_{K_j}\}$, where $\mathcal{Y}^{j,\star}_k = \mathcal{Y}^j_{I_k}$ for $k = 1, \ldots, K_j$, and $j = 1, \ldots, M+1$. Combine the data contained in each bootstrapped subsample j, and then obtain the following bootstrap observations for the whole sample $\{\Phi^\star_1, \ldots, \Phi^\star_{M+1}\}$.

(iv) Compute the bootstrap test statistic $F^\star_T(M+1|M)$ based on the bootstrap sample $\{\Phi^\star_1, \ldots, \Phi^\star_{M+1}\}$.

(v) Repeat steps (iii) and (iv) for a total of **B** times to obtain a collection of **B** bootstrap test statistics $\{F^\star_{T,b}(M+1|M)\}_{b=1}^{\mathbf{B}}$. The bootstrap p-value for the

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

sequential test is given by

$$p_S^{\mathbf{B}} = \frac{1}{\mathbf{B}} \sum_{b=1}^{\mathbf{B}} \mathbb{I}\left[F_T(M+1|M) \leq F^{\star}_{T,b}(M+1|M)\right].$$

Supposing Assumptions A.1–A.4 and A.6 hold, the validity of this procedure can be established analogously to Theorem 3.3. We note that under the null hypothesis $\mathbb{H}_0 : M^0 = M$, the asymptotic distribution of the sequential test is equivalent to the joint test for a single break.

4.2. An Information Criterion

In this subsection, we propose an alternative data-driven choice for M^0 based on an information criterion (IC). We assume that M^0 is bounded from above by a finite integer $M_{\max}$. We denote the optimal objective function value of (2.2) for a fixed M as $\mathrm{SSGR}_M = \mathrm{SSGR}_M(\hat{r}_1, \ldots, \hat{r}_M)$. Here, we suppress the dependence of the estimators on the estimated break fractions $\{\hat{r}_j\}_{j=1}^M$. We further define

$$\hat{\sigma}^2(M) = \mathrm{SSGR}_M / T.$$

Then, we choose $\hat{M}$ as

$$\hat{M} = \arg\min_{0 \leq M \leq M_{\max}} \ln\left[\hat{\sigma}^2(M)\right] + \rho_T(M+1), \quad \textbf{(4.1)}$$

where ρ_T is a positive tuning parameter.

Assumption A.7. $\rho_T \to 0$ and $T\rho_T \to \infty$ as $T \to \infty$.

Assumption A.7 imposes standard conditions for the consistency of model selection, implying that the penalty coefficient ρ_T cannot shrink to zero too fast as $T \to \infty$. Theorem 4.2 shows that the IC procedure defined by (4.1) can consistently estimate the true number of breaks.

THEOREM 4.2. *Suppose Assumptions A.1–A.4 and A.7 hold, and let $\hat{M}$ be the number of breaks determined by (4.1). Then as $T \to \infty$,*

$$P(\hat{M} = M^0) \to 1.$$

Theorem 4.2 implies that the IC procedure can determine the true number of breaks with probability approaching one. To implement the IC procedure, one needs to choose the tuning parameter ρ_T. Motivated by the BIC, we use $\rho_T = c_\rho d \ln(T)/T$ in our simulation studies and empirical application, where d is the dimension of Y_t, and c_ρ is a deep tuning parameter at the practitioner's discretion. In our simulation studies, we examine the performance of our IC procedure with different c_ρ. Our results show that the choice of c_ρ matters only when the sample size is small. As the sample size increases, its impact becomes negligible. For more details, see Section 6.

5. INFERENCE FOR THE PATTERN OF STRUCTURAL BREAKS

When the null hypothesis of no structural break in distribution is rejected, one may need to gauge possible sources of the rejection, which can provide valuable information for modeling stationary subsamples. For instance, one may be interested in estimating and testing for structural breaks in various moments of each univariate time series or cumulants of a multivariate time series. The CF can be used to generate moments (if exist) via differentiation, which is useful for investigating structural breaks in various aspects of the joint distribution. Given the ECF-based objective function (2.2), we can develop a class of derivative estimators and tests to infer the pattern of structural breaks, such as structural breaks in (un)conditional mean, variance, and higher-order moments, respectively.

5.1. Inference for Structural Breaks in Moments

Suppose the $(p_1,\ldots,p_d)$th order moments of the d-dimensional time series variable Y_t exist. Then taking the $(p_1,\ldots,p_d)$th order partial derivative of $e^{\mathbf{i}u'Y_t}$ with respect to $u=(u_1,\ldots,u_d)$ at the origin, we have

$$\frac{\partial^{p_1+\cdots+p_d}e^{\mathbf{i}u'Y_t}}{\partial u_1^{p_1}\cdots\partial u_d^{p_d}}\bigg|_{(u_1,\ldots,u_d)=(0,\ldots,0)}=\mathbf{i}^{p_1+\cdots+p_d}Y_{1t}^{p_1}\cdots Y_{dt}^{p_d},$$

where Y_{it} is the ith entry of the $d\times 1$ vector Y_t for $i=1,\ldots,d$. Analogous to estimating and testing structural breaks in distribution, we can analyze structural breaks in the joint product moment $E(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d})$. In particular, if $d=1$, we can investigate structural breaks in certain moments of the univariate time series $\{Y_t\}_{t=1}^T$. This is important since the economic implication of breaks in various moments is different. For example, if Y_t represents the return of a financial asset, then structural breaks in the first moment implies shifts in the expected return, while breaks in the second moment indicates changes in risk.

Taking the $(p_1,\ldots,p_d)$th order partial derivative of (2.1) with respect to $u=(u_1,\ldots,u_d)$ at the origin, we obtain

$$Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}=E(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d})+\nu_t, \tag{5.1}$$

where $\nu_t=Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}-E(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d})$ characterizes the deviation of $Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}$ from its expectation. Then we can consistently estimate the break fractions $\{\hat{r}_j\}_{j=1}^M$ in $E(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d})$ by solving the following optimization problem

$$\min_{\{r_1,\ldots,r_M\}\in\Pi_\epsilon^0}\sum_{j=1}^{M+1}\sum_{t=T_{j-1}+1}^{T_j}\left[Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}-\hat{m}_j(Y_{1t},\ldots,Y_{dt})\right]^2,$$

where $\hat{m}_j(Y_{1t},\ldots,Y_{dt})=(T_j-T_{j-1})^{-1}\sum_{t=T_{j-1}+1}^{T_j}Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}$. With suitable moment conditions on Y_t, it is straightforward to show that $\{\hat{r}_j\}_{j=1}^M$ converge in probability to the true breaks, analogously to Theorem 2.1.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

The test statistic for the null hypothesis of no structural breaks in $E(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d})$ against the alternative hypothesis of M structural breaks is

$$\sup F^{(p_1,\dots,p_d)} = \sup_{\{r_1,\dots,r_M\}\in\Pi_\epsilon^0} F_T^{(p_1,\dots,p_d)}(r_1,\dots,r_M),$$

where

$$F_T^{(p_1,\dots,p_d)}(r_1,\dots,r_M) = \mathrm{SSGR}_0^{(p_1,\dots,p_d)} - \mathrm{SSGR}_M^{(p_1,\dots,p_d)}(r_1,\dots,r_M),$$

with

$$\mathrm{SSGR}_0^{(p_1,\dots,p_d)} = \sum_{t=1}^{T}\left(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d} - \frac{1}{T}\sum_{t=1}^{T} Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}\right)^2,$$

and

$$\mathrm{SSGR}_M^{(p_1,\dots,p_d)}(r_1,\dots,r_M) = \sum_{j=1}^{M+1}\sum_{t=T_{j-1}+1}^{T_j}\left(Y_{1t}^{p_1}\cdots Y_{dt}^{p_d} - \frac{1}{T_j - T_{j-1}}\sum_{t=T_{j-1}+1}^{T_j} Y_{1t}^{p_1}\cdots Y_{dt}^{p_d}\right)^2,$$

which represent the sum of squared residuals under the null hypothesis of no structural breaks and the alternative hypothesis of M breaks, respectively. Under regularity conditions analogous to Assumptions A.1, A.3, and A.4*, and suitable moment restrictions, we can show that

$$F_T^{(p_1,\dots,p_d)}(r_1,\dots,r_M) \Rightarrow \sum_{j=1}^{M+1}\frac{1}{r_j - r_{j-1}}\left[\mathcal{B}^{(p_1,\dots,p_d)}(r_j) - \mathcal{B}^{(p_1,\dots,p_d)}(r_{j-1})\right]^2,$$

where $\mathcal{B}^{(p_1,\dots,p_d)}(r) = B^{(p_1,\dots,p_d)}(r) - rB^{(p_1,\dots,p_d)}(1)$ is a Brownian bridge and $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\nu_t \Rightarrow B^{(p_1,\dots,p_d)}(r)$ is a Brownian motion on $[0,1]$ with mean zero and covariance kernel $E[B^{(p_1,\dots,p_d)}(r)B^{(p_1,\dots,p_d)}(s)] = \min\{r,s\}\Omega^{(p_1,\dots,p_d)}$. Note that $\Omega^{(p_1,\dots,p_d)}$ is the long-run variance of ν_t.

The choice of derivative orders $(p_1,\dots,p_d)$ allows us to estimate and test for structural breaks in various moments. For example, the choice of $d = 1$ and $p = 1$ yields Bai's (1994, 1997) test for structural breaks in mean if the test statistic is standardized properly. The case of $p > 1$ allows us to examine structural breaks in higher-order moments.

5.2. Inference for Structural Breaks in Conditional Moments

With the development and prevalence of nonlinear and nonparametric regressions, structural breaks in nonlinear or nonparametric regression has drawn increasing attention. For example, Andrews and Fair (1988) study structural breaks in nonlinear parametric regression by extending Chow's (1960) test. Boldea and Hall (2013) provide an econometric framework to identify breaks in nonlinear regressions.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

Although parametric models provide a parsimonious way to characterize the relationship among economic variables, they may suffer from model misspecification and lead to inconsistent estimation or suboptimal prediction. Several studies have considered structural breaks in nonparametric regression models, which avoid restrictions on the functional form. For example, Müller (1992) derives a central limit theorem for the estimators of the location and size of the break point. Wu and Chu (1993) propose kernel-type estimators for the locations and sizes of jumps in a fixed-design nonparametric regression. Loader (1996) proposes an estimator for breaks based on one-sided nonparametric regression. Delgado and Hidalgo (2000) study inference for the location and size of structural breaks in a nonparametric regression model. Su and Xiao (2008) propose a CUSUM test for structural change in dynamic nonparametric regression models. Fengler, Mammen, and Vogt (2015) propose a test for structural breaks in a nonparametric additive model. Vogt (2015) proposes a nonparametric time-varying conditional mean model by nonparametric smoothing over both the regressor and the rescaled time index $t/T \in (0,1]$. Fu and Hong (2019) test for smooth structural changes in a nonparametric time series regression model via a frequency domain approach.

We now extend our approach to estimating and testing for structural breaks in nonparametric regression via appropriate differentiation of the ECF. The salient feature of our approach is that we avoid smoothed nonparametric estimation for the unknown regression function and can detect a class of local alternatives that converges to the null hypothesis at the parametric rate.

Consider the following conditional mean model:

$$Z_t^p = g_t^0(X_t) + \nu_t, \tag{5.2}$$

where Z_t is a scalar dependent variable, X_t is a $(d-1)\times 1$ explanatory vector, and $g_t^0(X_t) = E(Z_t^p|X_t)$ is the conditional moment which may vary over time.

Suppose there exists M^0 breaks in the functional form of $g_t^0(\cdot)$. Then

$$Z_t^p = \begin{cases} h_1^0(X_t) + \nu_t, & \text{for } t = 1,\dots,T_1^0, \\ h_2^0(X_t) + \nu_t, & \text{for } t = T_1^0+1,\dots,T_2^0, \\ \quad\vdots & \quad\vdots \\ h_{M^0+1}^0(X_t) + \nu_t, & \text{for } t = T_{M^0}^0+1,\dots,T, \end{cases}$$

where $g_t^0(\cdot) = h_j^0(\cdot)$ for $t \in [T_{j-1}^0+1, T_j^0]$ and $j = 1,\dots,M^0+1$ with $\{h_j^0(X_t)\}_{j=1}^{M^0+1}$ being a collection of time-invariant square-integrable functions.

One can estimate the break fractions by solving the following optimization problem

$$\min_{\{r_1,\dots,r_{M^0}\}\in\Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \sum_{t=T_{j-1}+1}^{T_j} \left[Z_t^p - \hat{h}_j(X_t)\right]^2,$$

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

where $\hat{h}_j(\cdot)$ is an estimator for $h_j^0(\cdot)$. If $h_j^0(\cdot)$ has a parametric form such as $h_j^0(\cdot) = h_j(\cdot, \beta^0)$ for some unknown parameter β^0, we can estimate β^0 and the break fractions consistently by piecewise nonlinear least square estimation. Andrews and Fair (1988) consider the inference in such nonlinear parametric models. However, if the parametric function $h_j(\cdot, \beta)$ is misspecified for some $j = 1, \ldots, M^0 + 1$, then the break fractions may not be consistently estimated. One can mitigate this misspecification issue using smoothed nonparametric estimation of $h_j^0(\cdot)$ in each subsample. However, it would become rather tedious and may suffer from the curse of dimensionality problem when the dimension of X_t is large.

Compared to the aforementioned methods, our ECF approach is free of model misspecification and curse of dimensionality. Put $Y_t = (Z_t, X_t')'$ and let $u = (u_1, \omega')'$, where $\omega = (u_2, \ldots, u_d)'$. Taking the pth order partial derivative of the CF of Y_t with respect to u_1 and let $u_1 = 0$, we have

$$\left.\frac{\partial^p E(e^{\mathbf{i}u'Y_t})}{\partial u_1^p}\right|_{u_1=0} = \mathbf{i}^p E(Z_t^p e^{\mathbf{i}\omega' X_t}).$$

Multiplying $e^{\mathbf{i}\omega' X_t}$ on both sides of (5.2) and taking expectation, we obtain

$$E(Z_t^p e^{\mathbf{i}\omega' X_t}) = E\left[g_t^0(X_t) e^{\mathbf{i}\omega' X_t}\right], \tag{5.3}$$

where $E(v_t|X_t) = 0$ implies that $E(v_t e^{\mathbf{i}\omega' X_t}) = 0$ a.s. for all $\omega \in \mathbb{R}^{d-1}$.

Suppose $\{X_t\}_{t=1}^T$ is strictly stationary, or the collection of break dates in the joint distribution of $\{X_t\}_{t=1}^T$ is a subset of $\{T_j^0\}_{j=1}^{M^0}$. Then (5.3) implies that we can pin down the break dates in $g_t(\cdot)$ by estimating structural breaks in the following generalized regression

$$Z_t^p e^{\mathbf{i}\omega' X_t} = \zeta_t^0(p, \omega) + \varepsilon_t(p, \omega),$$

where $\zeta_t^0(p, \omega) = E[g_t^0(X_t) e^{\mathbf{i}\omega' X_t}]$. Given the collection of break dates $\{T_j^0\}_{j=1}^{M^0}$, we have

$$Z_t^p e^{\mathbf{i}\omega' X_t} = \begin{cases} \xi_1^0(p, \omega) + \varepsilon_t(p, \omega), & \text{for } t = 1, \ldots, T_1^0, \\ \xi_2^0(p, \omega) + \varepsilon_t(p, \omega), & \text{for } t = T_1^0 + 1, \ldots, T_2^0, \\ \vdots & \vdots \\ \xi_{M^0+1}^0(p, \omega) + \varepsilon_t(p, \omega), & \text{for } t = T_{M^0}^0 + 1, \ldots, T, \end{cases}$$

where $\xi_j^0(p, \omega) = E\left[h_j^0(X_t) e^{\mathbf{i}\omega' X_t}\right]$ for $t \in [T_{j-1}^0, T_j^0]$ and $j = 1, \ldots, M^0 + 1$.

Following analogous reasoning in estimating and testing for structural breaks based on (2.1), we can consistently estimate break fractions $\{r_j^0\}_{j=1}^{M^0}$ in $\{Z_t^p e^{\mathbf{i}\omega' X_t}\}_{t=1}^T$ by solving the following optimization problem

$$\min_{\{r_1, \ldots, r_{M^0}\} \in \Pi_\epsilon^0} \sum_{j=1}^{M^0+1} \sum_{t=T_{j-1}+1}^{T_j} \int_{\mathbb{R}^{d-1}} \left| Z_t^p e^{\mathbf{i}\omega' X_t} - \tilde{\xi}_j(p, \omega)\right|^2 W(\omega) \mathrm{d}\omega, \tag{5.4}$$

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

where $\tilde{\xi}_j(p,\omega) = (T_j - T_{j-1})^{-1} \sum_{t=T_{j-1}+1}^{T_j} Z_t^p e^{\mathbf{i}\omega' X_t}$. Furthermore, we can also show that the estimated break fraction $\hat{r}_j = \hat{T}_j/T$ converges to the true value at the rate of T for $j = 1, \ldots, M^0$. To proceed, we replace Assumptions A.1 and A.4 with the following conditions:

Assumption A.1c. (i) $\{X_t\}_{t=1}^T$ is a strictly stationary and α-mixing sequence with mixing coefficients $\alpha(s)$ such that $\sum_{s=0}^{\infty}(s+1)^{q/2-1}\alpha(s) < \infty$ for some $q > 2$ and $\delta > 0$; (ii) ν_t is weakly stationary, with $E(\nu_t|X_t) = 0$ and $\mathrm{var}(\nu_t|X_t) = \sigma^2(X_t) < \infty$; and (iii) $\{h_j^0(\cdot) : \mathbb{R}^{d-1} \to \mathbb{R}\}_{j=1}^{M^0+1}$ is a collection of square-integrable functions such that $g_t(\cdot) = h_j^0(\cdot)$ for each $t = T_{j-1}^0 + 1, \ldots, T_j^0$.

Assumption A.4c. Let $\varepsilon_t(p,\omega) \equiv Z_t^p e^{\mathbf{i}\omega' X_t} - \xi_j^0(p,\omega)$ be the generalized error function for $t \in [T_{j-1}^0 + 1, T_j^0]$, $j = 1, 2, \ldots, M^0 + 1$, and $\mathcal{U} \subset \mathbb{R}^{d-1}$ denote a symmetric compact hypercube around the origin: (i) $\{\varepsilon_t(p,\omega)\}_{t=T_{j-1}^0+1}^{T_j^0}$ satisfies $E[\varepsilon_t(p,\omega)] = 0$ and the generalized long-run variance function $\Omega_p^{(j)}(\omega_1,\omega_2) = \sum_{l=-\infty}^{\infty} \sigma_{p,l}^{(j)}(\omega_1,\omega_2)$ with $\sigma_{p,l}^{(j)}(\omega_1,\omega_2) = E\left[\varepsilon_t(p,\omega_1)\varepsilon_{t+l}(p,\omega_2)^*\right]$ and (ii) $\left(T_j^0 - T_{j-1}^0\right)^{-1/2} \sum_{t=T_{j-1}^0+1}^{T_{j-1}^0+\left\lfloor \left(T_j^0 - T_{j-1}^0\right) r\right\rfloor} \varepsilon_t(p,\omega) \Rightarrow B_p^{(j)}(\omega, r)$, where $B_p^{(j)}(\omega,r)$ is a generalized Brownian motion on $\mathcal{U} \times [0,1]$ with mean zero and covariance kernel $E\left[B_p^{(j)}(\omega_1,r)\, B_p^{(j)}(\omega_2,s)^*\right] = \min\{r,s\}\, \Omega_p^{(j)}(\omega_1,\omega_2)$.

Similar to Assumption A.1(ii), Assumption A.1c(i) restricts the temporal dependence of X_t. Assumption A.1c(ii) imposes some regularity conditions on the true regression error ν_t. It implies that X_t is exogenous and allows for conditional heteroskedasticity. Assumption A.1c(iii) restricts that the unknown conditional moment function $h_j^0(\cdot)$ is square-integrable for each regime j. It is equivalent to a moment restriction on Z_t^p to ensure that (5.2) always exists. Assumption A.4c is analogous to Assumption A.4 except that the generalized error function for testing and estimating structural breaks in distribution is replaced by $\varepsilon_t(p,\omega)$.

Under Assumptions A.1c, A.2, and A.3, we can obtain similar conclusions to those of Theorems 2.1 and 2.2 for the estimated break fractions. Adding Assumptions A.4c and A.5, we can obtain the limiting distribution for break fractions in conditional moments, which is the same as Theorem 2.4 except that $\Omega^{(j)}(u,v)$ is replaced with $\Omega_p^{(j)}(u,v)$.

We note that Assumption A.1c requires that X_t is strictly stationary. It is possible that the distribution of X_t has structural breaks. As a result, the optimization problem given in (5.4) will deliver the union of the estimated dates for both X_t and the conditional moment $E(Z_t^p|X_t)$. We regard this as a price to pay for our proposed model-free approach. To mitigate this issue, one can first test and estimate breaks in the joint distribution of X_t via the procedures developed in Sections 2 and 4, and then detect structural breaks via solving the optimization problem given by (5.4) in Section 5.2. Ideally, a subset of break dates in $E(Z_t^p e^{\mathbf{i}\omega' X_t})$ will coincide with the estimated break dates of the distribution of X_t. Hence, one may rule them out and treat the remaining break fractions as breaks for the conditional moments.

We can detect structural breaks in a nonparametric time series regression via an analogous joint test developed in Section 3 such that it can detect a class of local alternatives at the parametric rate $\sqrt{T}$. Hence, it is asymptotically more efficient than the existing tests for structural changes based on smoothed nonparametric regression (e.g., Fengler et al., 2015; Vogt, 2015; Fu and Hong, 2019). This is an advantage of our ECF approach, which avoids smoothed nonparametric estimation and the curse of dimensionality. Moreover, since we do not impose any functional form on the conditional mean model, our estimators and tests for structural breaks are model-free. It avoids model misspecification induced by a nonlinear parametric model as in Andrews and Fair (1988) and Boldea and Hall (2013).

6. SIMULATION STUDIES

We now study the finite sample performance of our proposed estimators and tests via Monte Carlo.

6.1. Data Generating Processes

We generate data under the following data generating processes (DGPs):

DGP.S1: $Y_t = \kappa_t$;
DGP.S2: $Y_t = 0.2Y_{t-1} + \kappa_t$;
DGP.S3: $Y_t = \sqrt{h_t}\kappa_t$, $h_t = 0.2 + 0.3Y_{t-1}^2$;
DGP.S4: $Y_t \sim N(\mu, V)$, with $\mu = (0,0,0)'$, and $V = \begin{pmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 1 & 0.5 \\ 0.2 & 0.5 & 1 \end{pmatrix}$;
DGP.S5: $Y_t = (X_t, Z_t)'$ with $X_t = 0.5X_{t-1} + \kappa_t$ and $Z_t = 1 + 0.5X_t + \eta_t$, $\eta_t \sim i.i.d.N(0, 0.1^2)$;
DGP.P1 [*A single structural break in mean*]: $Y_t = \kappa_t \mathbb{I}(t \le 0.5T) + (1+\kappa_t)\mathbb{I}(t > 0.5T)$;
DGP.P2 [*A single structural break in variance*]: $Y_t = \kappa_t \mathbb{I}(t \le 0.5T) + 2\kappa_t \mathbb{I}(t > 0.5T)$;
DGP.P3 [*A single structural break in higher-order moments*]: $Y_t = (1 + \sqrt{2}\kappa_t)\mathbb{I}(t \le 0.5T) + \kappa_t^2 \mathbb{I}(t > 0.5T)$;
DGP.P4 [*A single structural break in dependence*]: $Y_t = (X_t, Z_t)'$, $X_t = 0.5X_{t-1} + \kappa_t$, $\eta_t \sim i.i.d.N(0, 0.1^2)$, and

$$Z_t = \begin{cases} 1 + 0.5X_t + \eta_t, & \text{if } t \le 0.3T, \\ 2 + 0.8X_t + \eta_t, & \text{otherwise;} \end{cases}$$

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

DGP.P5 [*Multiple structural breaks in dependence*]: $Y_t = (X_t, Z_t)'$, $X_t = 0.5X_{t-1} + \kappa_t$, $\eta_t \sim i.i.d.N(0, 0.1^2)$, and

$$Z_t = \begin{cases} 1 + X_t^2 + \eta_t, & \text{if } 1 \le t \le 0.3T, \\ 0.5 + 0.1X_t + \eta_t, & \text{if } 0.3T + 1 \le t \le 0.6T, \\ 3 + 0.5X_t + \eta_t, & \text{if } 0.6T + 1 \le t \le T, \end{cases}$$

where $\{\kappa_t\}$ is an i.i.d.$N(0,1)$ sequence.

The null hypothesis of no structural breaks holds under DGPs.S1–S5. We use these DGPs to study the size performance of the proposed joint test and derivative tests. Specifically, DGPs.S2 and S3 allow us to examine the performance of our tests under serial dependence and conditional heteroskedasticity, and DGPs.S4 and S5 are designed to examine the performance of our tests under multivariate cases. DGPs.P1–P5 provide various types of structural breaks. We use them to examine the consistency of our estimator for break fractions as well as the power of our tests. DGPs.P1 and P2 contain a single break in mean and variance, respectively, whereas DGP.P3 has both time-invariant mean and variance but a single break in higher-order moments. DGPs.P4 and P5 have a single break and multiple breaks, respectively, in a multivariate context.

6.2. Determining the Number of Breaks

In this subsection, we examine the finite sample performance of the proposed IC procedure and sequential tests in determining the number of breaks. We also compare our approach with closely related methods for conditional mean models in the literature, including Bai and Perron's (1998) sequential tests, the conventional BIC procedure, and Liu, Wu, and Zidek's (1997, LWZ) information criterion procedure.

As noted earlier, for our IC, we set the turning parameter $\rho_T = c_\rho d \ln(T)/T$ and examine the sensitivity of its performance to the choice of the deep turning parameter c_ρ. We consider $c_\rho = 1, 1.5$, and 2, but for space, we only report the results for $c_\rho = 1$. The results with different choices of c_ρ are quite similar when the sample size is large. For details, see the Supplementary Material. We also show the performance of the proposed sequential testing procedure in determining the number of breaks. We test the null hypothesis of M breaks against the alternative of $M+1$ breaks sequentially for $M = 0, 1, \ldots, M_{\max}$. If we cannot reject the null hypothesis for some specific value of $\bar{M}$, then the sequential procedure stops and $\bar{M}$ is the selected number of breaks. Since the results of sequential tests depend on the significance level, we adopt the most commonly used 5% and 10% significance levels. We set the prespecified trimming parameter $\epsilon = 0.15$, which is a common choice in the literature. We also examine the performance of the methods under study with $\epsilon = 0.1$, and the results (not reported) are similar. Furthermore, we use the normal weighting function in (3.3) with $b = 1$ to compute the SSGR. We also examine the performance of our approach using the Laplace weighting function in (3.4) and the uniform weighting function in (3.5).

Since our tests are not asymptotically pivotal, we use the proposed moving block bootstrap to obtain the critical values. Under each DGP, we simulate 1,000 data sets with sample size $T = 100, 200$, and 500, respectively. To estimate the true number of breaks, we set $M_{\max} = 5$ for sequential tests and report the estimated number of breaks when we fail to reject the alternative hypothesis at given significance levels. For the choice of the block length, we adopt Politis and White's (2004) automatic block-length selection procedure. We also report the estimated number of breaks using IC. To evaluate the performance of these methods, we consider the average number of breaks and the percentage of correct selection over 1,000 replications. Bai and Ng (2002) use the former measure when they determine the number of common factors in a large dimensional factor model, while LWZ consider the latter when they determine the number of breaks in a multivariate linear regression.

Table 1 reports the average number of breaks and the percentage of correct selection over 1,000 replications with various methods of determining the number of breaks. As shown in Table 1, both the IC and sequential tests perform fairly well under DGPs.P1–P5. For linear regressions with structural breaks in mean as specified by DGPs.P1 and P4, the percentage of correct selection by our IC is slightly lower than those by BIC and LWZ when the sample size is small, but it improves as the sample size grows. For structural breaks in variance and higher-order moments under DGPs.P2 and P3, respectively, the number of breaks detected by BIC and LWZ tend to be zero, and the percentages of correct selection shrink to zero as the sample size increases. In contrast, our IC performs well under both DGPs, and the percentage of correct selection is close to one. It is not difficult to understand the failure of BIC and LWZ, which are constructed for detecting structural breaks in linear regression models. As a result, they are unable to detect structural breaks in higher-order moments. We also find that BIC and LWZ fail to detect structural breaks in nonlinear conditional mean models. Both of them suffer from a severe over-selection problem under DGP.P5. In contrast, our IC performs well and is free of model misspecification. Similar results are documented for the proposed sequential tests. As shown in Table 1, the percentage of correct selection of our sequential tests is higher than Bai and Perron's (1998) sequential tests under DGPs.P1–P5, respectively.

6.3. Estimation for Break Dates

We use DGPs.P1–P5 to evaluate the performance of the proposed estimators for break dates. We compare our estimation results with those proposed by Bai and Perron (1998) and Inoue (2001). To measure the accuracy of the estimators, we define the bias and root-mean-squared errors (RMSE) of the estimated break fractions as

$$\text{Bias} = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}(\hat{r}_{ji} - r_j^0),$$

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

TABLE 1. Performance of various methods in determining the number of breaks.

DGP	T	Average number of breaks							Percentage of correct selection						
		BIC	LWZ	ST5	ST10	DIC	DST5	DST10	BIC	LWZ	ST5	ST10	DIC	DST5	DST10
P1	100	1.064	0.840	1.194	1.291	1.254	0.890	0.041	91.8	84.0	79.1	74.3	79.4	84.2	89.0
	200	1.030	0.994	1.120	1.178	1.140	1.028	1.081	97.1	99.4	88.6	84.1	88.9	96.4	92.1
	500	1.014	1.000	1.064	1.123	1.072	1.044	1.104	98.6	100	93.8	88.5	94.0	95.7	90.6
P2	100	0.306	0.007	0.165	0.238	0.842	0.699	0.869	8.1	0.7	10.5	14.3	66.4	66.0	74.3
	200	0.185	0.001	0.107	0.182	1.014	1.014	1.071	5.1	0.1	8.5	12.8	92.3	94.6	92.1
	500	0.098	0.000	0.059	0.119	1.025	1.049	1.104	3.2	0.0	4.9	9.5	97.7	95.3	90.0
P3	100	0.183	0.008	0.166	0.252	0.980	0.659	0.833	4.7	0.4	11.2	16.2	61.0	61.6	71.1
	200	0.088	0.000	0.125	0.192	1.070	0.986	1.054	3.7	0.0	10.8	15.2	87.1	92.4	90.3
	500	0.031	0.000	0.109	0.169	1.039	1.036	1.098	2.1	0.0	9.9	14.7	96.5	96.4	90.6
P4	100	1.060	1.000	2.020	2.225	1.093	0.970	1.164	94.6	100	40.1	33.7	91.1	84.9	82.2
	200	1.017	1.000	1.412	1.544	1.081	1.077	1.169	98.4	100	68.5	59.9	92.2	92.3	84.6
	500	1.005	1.000	1.175	1.268	1.048	1.067	1.132	99.5	100	83.6	75.8	95.7	93.5	87.5
P5	100	4.867	4.469	3.441	3.692	1.774	1.334	1.973	0.0	2.3	13.7	11.1	68.1	42.2	70.8
	200	4.856	4.297	2.758	2.921	2.062	2.031	2.158	0.1	2.5	38.7	34.1	94.1	91.7	85.6
	500	4.872	4.101	2.284	2.386	2.034	2.062	2.154	0.1	7.2	74.3	67.0	96.6	94.2	86.9

Notes: (i) BIC and LWZ denote the Schwarz criterion and the modified Schwarz criterion proposed in Liu et al. (1997); (ii) ST5 and ST10 denote Bai and Perron's (1998) sequential tests under 5% and 10% significance levels, respectively; (iii) DIC denotes the information criterion proposed in this paper; (iv) DST5 and DST10 denote the sequential tests proposed in this paper under 5% and 10% significance levels, respectively; and (v) the main entries report the results based on 1,000 replications.

TABLE 2. Performance on estimating break dates.

		This paper		Inoue (2001)		Bai and Perron (1998)	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
P1	$T = 100$	4.510	75.753	0.1900	**57.412**	4.690	57.678
	$T = 200$	0.170	37.535	0.305	31.693	−0.705	**28.796**
	$T = 500$	−0.760	12.470	−0.370	14.406	−0.612	**10.654**
P2	$T = 100$	21.750	107.953	23.380	**104.866**	156.490	232.856
	$T = 200$	13.085	**50.821**	21.460	76.567	156.690	237.673
	$T = 500$	4.886	**20.257**	13.124	41.878	165.258	234.759
P3	$T = 100$	−15.050	122.732	−23.740	**91.318**	−7.450	232.574
	$T = 200$	−4.960	66.474	−16.455	**60.490**	−10.715	236.683
	$T = 500$	−1.204	27.806	−9.058	**27.138**	−2.882	231.878
P4	$T = 100$	5.870	41.904	8.680	59.200	−0.030	**1.049**
	$T = 200$	2.835	14.708	7.960	38.379	−0.010	**0.806**
	$T = 500$	0.982	3.492	3.268	16.162	0.008	**0.219**
P5	$T = 100$	4.855	**35.836**	–	–	−22.790	51.931
	$T = 200$	0.108	**3.747**	–	–	−13.680	35.491
	$T = 500$	0.003	**0.656**	–	–	−4.109	16.482

Notes: The main entries report bias and RMSEs ×1,000. The bold entries highlight the smallest RMSE in each case.

$$\text{RMSE} = \sqrt{\frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}(\hat{r}_{ji} - r_j^0)^2},$$

where M is the number of breaks, N is the number of replications, r_j^0 is the true value of the jth break fraction, and $\hat{r}_{ji}$ is the estimator for the jth break fraction in the ith replication.

Table 2 reports the bias and RMSE for the estimators of Bai and Perron (1998), Inoue (2001), and ours, based on 1,000 replications. As shown in Table 2, both the bias and RMSE of our estimator decline as the sample size T increases. Bai and Perron's (1998) estimator outperforms ours under DGPs.P1 and P4, which depict linear regressions with structural breaks in mean. However, it does not perform as well as our estimator under other DGPs, especially under DGPs.P2 and P3, which exhibit structural breaks in variance and higher-order moments with time-invariant mean. Note that our method is applicable to nonlinear regression models with structural breaks under DGP.P5, where Bai and Perron's (1998) estimator is less accurate than ours. Inoue's (2001) estimator is also reasonable under DGPs.P1–P4. The RMSEs of his estimator are quite similar to ours under DGPs.P1–P3 when the sample size is large. It is a little bit worse than ours under DGP.P4. However, Inoue's (2001) estimator is only designed for a single break model, so we do

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

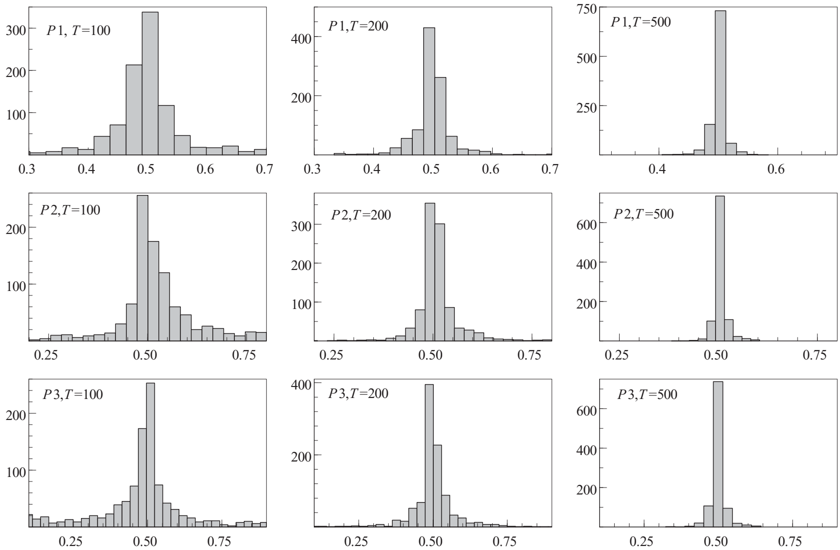

FIGURE 1. Histograms of estimated break fractions under DGPs.P1–P3.

not report the results for DGP.P5. We also plot the histograms of our estimated break fractions under DGPs.P1–P3 in Figure 1. Under each DGP, the range of $\hat{r}_j$ becomes smaller as the sample size increases, which, combined with the RMSE results in Table 2, indicates the convergence of the estimated break fractions to the true values. We note that the trimodality that is commonly present in the finite sample theory does not show up in the reported histograms. This is because the considered DGPs have a large signal-to-noise ratio. If we decrease the break size from 1 to 0.5 in mean under DGP.P1 and the break size from 2 to 1.5 in variance under DGP.P2, then the trimodality appears in the corresponding histograms. As is shown in Figure C.1 of the Supplementary Material, the trimodality is apparent when the sample size $T = 100$. It indicates that our results are consistent with the existing findings that the asymptotic distribution does not conform to the exact distribution when the sample size is small. To better approximate the finite sample distribution, it is possible to follow Jiang, Wang, and Yu (2018) to develop an in-fill asymptotic theory. However, the analytical expressions for our estimators are quite involved, and hence we have to pursue this important issue in subsequent research.

6.4. Size and Power Performance of Tests

We now examine the finite sample performance of our tests in comparison with Inoue's (2001) test for a single break in distribution. In addition, we also consider our derivative tests $F_T^{(1)}$ and $F_T^{(2)}$, which can detect structural breaks in the first two

TABLE 3. Size of tests under DGPs.S1–S5.

		$F_T, M=1$		$F_T, M=2$		$F_T^{(1)}, M=1$		$F_T^{(2)}, M=1$		In01		BP_{het}	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	$T=100$	6.1	11.7	5.6	11.5	5.1	10.0	4.6	9.3	4.1	7.7	7.8	14.5
	$T=200$	4.2	9.7	4.6	10.3	4.5	11.1	4.7	9.2	4.3	8.7	4.8	11.2
	$T=500$	5.2	10.1	5.7	10.8	5.8	12.3	6.8	12.4	3.3	6.6	5.2	10.6
S2	$T=100$	7.5	13.5	6.4	16.6	7.2	14.9	5.8	11.5	6.6	12.0	10.4	16.7
	$T=200$	6.2	13.8	6.5	16.3	5.7	11.0	4.9	10.5	5.6	11.7	6.3	11.3
	$T=500$	5.4	11.7	7.5	14.2	7.3	13.1	6.6	12.9	6.2	10.6	3.6	9.6
S3	$T=100$	6.3	12.1	7.2	13.6	5.4	11.3	12.3	19.8	4.7	9.1	10.2	17.4
	$T=200$	5.6	11.1	8.7	13.5	5.0	10.6	13.4	23.4	3.8	7.9	6.2	13.1
	$T=500$	4.1	10.5	7.5	13.9	3.8	8.2	14.0	22.0	3.5	7.9	4.7	10.2
S4	$T=100$	3.4	7.2	2.4	5.6	3.6	7.8	2.6	8.0	4.3	9.2	18.8	27.0
	$T=200$	4.0	8.1	3.2	7.2	4.7	8.9	3.3	6.8	4.5	8.3	9.9	16.8
	$T=500$	4.5	9.7	3.8	9.0	4.5	8.4	3.8	8.3	3.9	9.1	7.9	13.5
S5	$T=100$	7.4	13.5	5.6	14.7	9.2	18.1	6.8	14.2	7.4	15.9	16.2	25.1
	$T=200$	5.7	13.2	6.8	15.1	8.3	15.7	6.2	12.6	5.5	14.5	9.7	16.7
	$T=500$	5.0	12.5	7.3	14.1	8.0	15.2	6.2	13.7	5.6	12.8	6.0	14.0

Notes: (i) $F_T, M = 1$ and $F_T, M = 2$ denote the results of our joint test against the alternative hypothesis of a single break and two breaks, respectively; (ii) $F_T^{(1)}, M = 1$ and $F_T^{(2)}, M = 1$ denote the results of our derivative tests for structural changes in the first and second moments against a single break, respectively; (iii) In01 denotes the results of Inoue's (2001) test; (2) BP_{het} denote Bai and Perron's (1998) serial correlation and heteroskedasticity robust sup-F test. The main entries report the percentage of rejections.

moments, respectively. We compare the derivative test $F_T^{(1)}$ with Bai and Perron's (1998) test for structural breaks in conditional mean.

For each DGP, we simulate 1,000 data sets with the sample size $T = 100, 200$, and 500, respectively. For our tests and Inoue's (2001) test, we set the number of bootstrapping $\mathbf{B} = 200$. We set the trimming parameter $\epsilon = 0.15$ for all the tests under study. We also examine the performance of all the tests with $\epsilon = 0.10$, and we find the results (not reported here) are similar. In addition, to examine the impact of the pre-specified number of breaks under the alternative on the performance of our joint test for distributional breaks, we consider the number of breaks $M = 1$ and 2, respectively.

Table 3 reports the size of the tests under DGPs.S1–S5 at the 5% and 10% significance levels. Our joint test with $M = 1$ and 2 has reasonable size performance based on bootstrapped critical values, given that the empirical rejection rates are close to the corresponding nominal significance levels. Our derivative tests $F_T^{(1)}$ and $F_T^{(2)}$ also perform reasonably well, although $F_T^{(2)}$ tends to over-reject under DGP.S3, which is an ARCH process. However, the over-rejection is alleviated as the sample

TABLE 4. Power of tests under DGPs.P1–P5.

		$F_T, M=1$		$F_T, M=2$		$F_T^{(1)}, M=1$		$F_T^{(2)}, M=1$		In01		BP_{het}	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	$T=100$	88.6	96.6	65.2	85.1	93.6	98.3	4.5	11.3	88.3	92.7	98.7	99.3
	$T=200$	99.7	100	84.6	97.9	99.6	100	4.6	9.1	98.9	99.1	100	100
	$T=500$	100	100	100	100	100	100	4.9	9.6	100	100	100	100
P2	$T=100$	67.9	81.1	55.7	71.3	9.0	14.7	84.6	94.5	29.3	46.9	18.1	28.5
	$T=200$	98.0	99.3	91.7	96.3	8.6	15.6	99.4	99.8	56.9	66.6	16.5	27.7
	$T=500$	100	100	100	100	10.9	17.2	100	100	72.3	72.4	13.3	22.3
P3	$T=100$	63.7	77.1	49.3	64.3	5.8	9.7	10.1	17.2	31.3	44.3	10.1	17.2
	$T=200$	95.5	97.4	89.8	94.8	5.3	10.9	10.0	16.6	59.3	67.1	6.7	13.9
	$T=500$	100	100	100	100	5.7	10.2	9.2	17.9	69.4	69.4	5.2	10.5
P4	$T=100$	90.7	98.3	76.3	94.4	20.6	39.8	8.7	19.0	56.0	87.5	100	100
	$T=200$	100	100	98.6	100	50.4	76.8	11.7	25.6	89.3	98.8	100	100
	$T=500$	100	100	100	100	99.1	100	44.4	65.5	100	100	100	100
P5	$T=100$	86.0	99.0	86.2	99.1	12.0	23.3	3.7	8.2	9.7	49.6	81.7	83.7
	$T=200$	99.5	100	99.9	100	27.3	51.3	2.2	6.6	31.7	88.7	72.3	77.2
	$T=500$	100	100	100	100	92.7	98.6	5.2	25.4	98.5	100	70.9	80.6

Note: See the notes in Table 3.

size increases. Inoue's (2001) test also has reasonable size performance. Based on asymptotic critical values, Bai and Perron's (1998) test tends to over-reject under DGPs.S3–S5, especially when the sample size is small.

Table 4 reports the power of the tests under DGPs.P1–P5 at the 5% and 10% significance levels. Our joint test F_T for distributional breaks with $M = 1$ and 2 is powerful in detecting all the five DGPs and the empirical rejection rates increase to one quickly as T grows. Note that DGPs.P1–P4 have a single structural break, while DGP.P5 has two structural breaks. Thus, the number of breaks is misspecified for DGPs.P1–P4 when we set $M = 2$ and for DGP.P5 when we set $M = 1$. As shown in Table 4, our test is powerful and robust to the misspecification of the number of breaks, which is consistent with Theorem 3.2. Inoue's (2001) test is also powerful, but the rejection rates are relatively lower than ours. It is interesting to observe that $F_T^{(1)}$ is powerful in capturing various forms of structural breaks in mean under DGPs.P1, P4, and P5, and is robust to structural breaks in higher-order moments, such as under DGPs.P2 and P3. Similarly, $F_T^{(2)}$ performs well in capturing structural breaks in variance under DGP.P2 and is robust to structural breaks in mean and higher-order moments under DGPs.P1 and P3. On the other hand, Bai and Perron's (1998) test is most powerful under DGP.P1, which has a structural break in mean. However, it has little power under DGPs.P2 and P3, where structural breaks occur in variance and higher-order moments. This is

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

consistent with the fact that Bai and Perron's (1998) test is designed to capture structural changes in mean.

6.5. The Choice of Weighting Function

The weighting function $W(u)$ plays an important role in our framework. In this subsection, we examine the impact of various choices of $W(u)$. In particular, we consider the three types of weighting function in Section 3.3, i.e., (a) the normal weighting function $N(0,1)$, (b) the Laplace weighting function $L(0,1/\sqrt{2})$, and (c) the uniform weighting function $U(-\sqrt{3},\sqrt{3})$, respectively. The mean and variance of these weighting functions have been standardized to 0 and 1 for each dimension.

As shown in Table 5, the size of our test is not sensitive to the choice of weighting function $W(u)$ as long as they are finite and integrable. Furthermore, Table 6 shows that the choice of weighting function $W(u)$ can affect the finite sample power of our test under various DGPs. However, as the sample size increases, the proposed test with different weighting functions achieves unity power quickly. This implies that the impact of weighting function is little in large samples.

7. APPLICATION TO FOREIGN EXCHANGE RATES

The foreign exchange market is among the most important financial markets in the world. It plays an important role in international trade, capital market stability, and international portfolio management. The fluctuation of exchange rates affects the domestic country's economy. Hence, it is important to infer and forecast the pattern of exchange rate variation. Moreover, as documented by the vast literature (e.g., Boothe and Glassman, 1987; Hsieh, 1988), the distribution of the exchange rate returns usually exhibits a sharper peak and fatter tails than the normal distribution. Therefore, the conditional mean model cannot fully describe the important features of an exchange rate return dynamics. In addition, a large stream of literature also shows that the exchange rates may suffer from structural breaks due to the reforms of exchange rate system and other changing factors such as policy shifts.

In this section, we apply our approach to infer possible structural breaks in exchange rate returns. Following Liu and He (1991) and Rime, Sarno, and Sojli (2010), we check the stability of four exchange rates: EUR (Euro), JPY (Japanese Yen), CNY (Chinese Yuan), and CAD (Canadian Dollar). The data are measured by JPY, CNY, and CAD to one U.S. Dollar, and one U.S. dollar to EUR, respectively. Similar to Inoue (2001), we use the weekly return series measured by the log-difference of the average values of exchange rates within one week. We also considered Wednesday returns, and the results are quite similar. All data are collected from the website of the Federal Reserve Bank of St. Louis, spanning from January 5, 2000 to July 10, 2019, with 1,018 observations for EUR, JPY, and CAD. Since CNY switched from a fixed to a managed floating exchange rate mechanism in July 21, 2005, we use the data that started from July 25, 2005 for CNY. The time series plots of these exchange rate returns are given in Figure 2.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

TABLE 5. Size of tests under DGPs.S1–S5 with various weighting functions.

		$N(0,1),1$		$N(0,1),2$		$L(0,\frac{1}{\sqrt{2}}),1$		$L(0,\frac{1}{\sqrt{2}}),2$		$U(-\sqrt{3},\sqrt{3}),1$		$U(-\sqrt{3},\sqrt{3}),2$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	$T=100$	6.1	11.7	5.6	11.5	5.3	11.3	5.4	9.5	7.9	12.4	7.4	13.2
	$T=200$	4.2	9.7	4.6	10.3	3.6	9.2	3.8	9.0	5.2	10.7	5.1	10.6
	$T=500$	5.2	10.1	5.7	10.8	5.0	11.0	7.0	12.6	6.0	10.8	6.6	12.0
S2	$T=100$	7.5	13.5	6.4	16.6	7.0	14.2	5.8	12.6	8.8	14.9	8.3	15.7
	$T=200$	6.2	13.8	6.5	16.3	7.1	12.8	7.0	12.8	7.3	14.7	7.4	14.8
	$T=500$	5.4	11.7	7.5	14.2	4.4	11.2	4.8	11.2	5.0	11.2	5.4	11.6
S3	$T=100$	6.3	12.1	7.2	13.6	6.0	12.5	8.1	14.0	6.7	12.1	9.0	16.1
	$T=200$	5.6	11.1	8.7	13.5	6.1	12.2	6.4	11.4	6.6	10.8	6.8	14.4
	$T=500$	4.1	10.5	7.5	13.9	4.6	9.8	5.6	11.2	5.0	10.0	6.4	11.2
S4	$T=100$	3.4	7.2	2.4	5.6	3.6	7.8	3.0	6.6	3.9	9.5	3.9	8.5
	$T=200$	4.0	8.1	3.2	7.2	5.1	11.3	3.2	9.1	6.1	11.2	4.3	11.5
	$T=500$	4.5	9.7	3.8	9.0	5.0	12.4	4.2	9.2	7.2	12.2	5.6	11.0
S5	$T=100$	7.4	13.5	5.6	14.7	8.7	16.2	7.0	16.5	9.0	17.2	8.7	16.8
	$T=200$	5.7	13.2	6.8	15.1	8.0	15.4	7.0	15.7	7.1	14.4	8.2	15.9
	$T=500$	5.0	12.5	7.3	14.1	7.2	13.4	5.6	12.2	6.4	13.4	6.2	12.4

Notes: (i) $N(0,b),M$ denotes our joint test with normal $N(0,b)$ weighting function for the alternative hypothesis of M breaks; (ii) $L(0,b),M$ denotes our joint test with Laplace $L(0,b)$ weighting function for the alternative hypothesis of M breaks; and (iii) $U(-c,c),M$ denotes our joint test with uniform$(-c,c)$ weighting function for the alternative hypothesis of M breaks. The main entries report the percentage of rejections.

TABLE 6. Power of tests under DGPs.P1–P5 with various weighting functions.

		$N(0,1),1$		$N(0,1),2$		$L(0,\frac{1}{\sqrt{2}}),1$		$L(0,\frac{1}{\sqrt{2}}),2$		$U(-\sqrt{3},\sqrt{3}),1$		$U(-\sqrt{3},\sqrt{3}),2$	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	$T = 100$	88.6	96.6	65.2	85.1	88.8	96.8	65.7	85.6	90.4	96.8	68.6	86.9
	$T = 200$	99.7	100	84.6	97.9	99.1	100	88.3	98.5	99.3	99.9	91.0	98.6
	$T = 500$	100	100	100	100	100	100	100	100	100	100	100	100
P2	$T = 100$	67.9	81.1	55.7	71.3	68.9	82.1	54.3	69.5	71.8	83.8	60.4	74.2
	$T = 200$	98.0	99.3	91.7	96.3	97.6	99.3	92.4	96.8	97.1	98.9	92.7	96.0
	$T = 500$	100	100	100	100	100	100	100	100	100	100	100	100
P3	$T = 100$	63.7	77.1	49.3	64.3	66.4	81.6	50.5	66.6	76.5	86.3	63.3	75.2
	$T = 200$	95.5	97.4	89.8	94.8	96.2	98.7	89.5	95.3	97.8	99.0	93.2	97.4
	$T = 500$	100	100	100	100	100	100	100	100	100	100	100	100
P4	$T = 100$	90.7	98.3	76.3	94.4	84.0	97.7	68.7	90.6	96.6	99.6	89.8	98.7
	$T = 200$	100	100	98.6	100	99.3	100	95.6	99.7	100	100	99.2	100
	$T = 500$	100	100	100	100	100	100	100	100	100	100	100	100
P5	$T = 100$	86.0	99.0	86.2	99.1	74.8	98.1	88.6	99.7	95.8	100	95.5	100
	$T = 200$	99.5	100	99.9	100	98.2	99.8	99.5	100	99.6	100	99.9	100
	$T = 500$	100	100	100	100	100	100	100	100	100	100	100	100

Note: See the notes in Table 5.

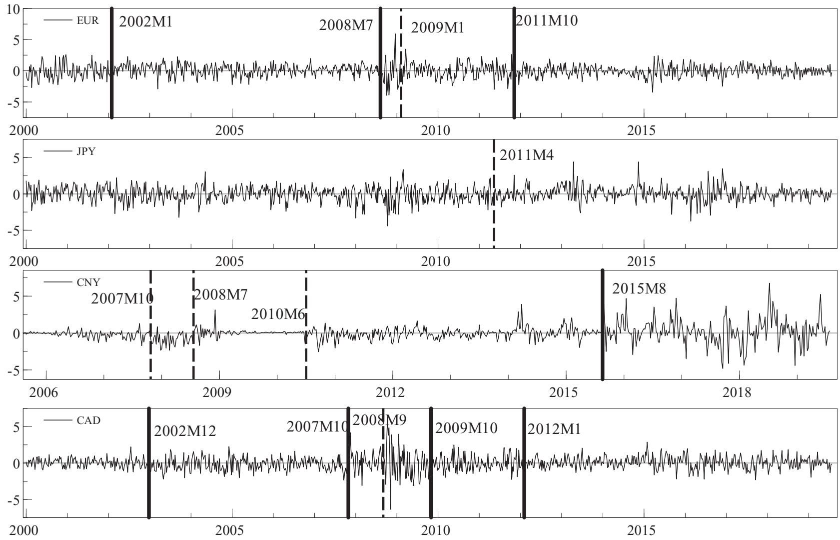

FIGURE 2. Time series plots of exchange rate returns and break dates.

We first check structural breaks using our joint test and derivative tests. We also use Bai and Perron's (1998) tests for structural breaks in mean and Inoue's (2001) test for a structural break in distribution. For Bai and Perron's (1998) tests, we consider the sup-F test against M breaks, for $M = 1, 2, \ldots, M_{\max}$, and the Dmax-$F$ and WDmax-F tests against an unknown number of breaks. The results of these tests are similar and the conclusions are almost the same. For space, we only report the Dmax-F test. For our joint test and derivative tests, we consider the cases that $M = 1$ and M is determined by IC, respectively. We report the results for the alternative hypothesis of a single break. We set the trimming parameter $\epsilon = 0.05$. We also consider the cases of $\epsilon = 0.1, 0.15$ and the results are similar. The maximum number of breaks is set to $M_{\max} = 5$. The other settings are identical to those in our simulation studies.

Table 7 reports the results of various tests at the 5% and 10% significance levels. Our joint test significantly rejects the null hypothesis of no structural breaks in distribution for all four exchange rate returns at the 10% significance level, and all except JPY at the 5% significance level. Inoue's (2001) test can only reject the null hypothesis for EUR, CNY, and CAD at the 5% significance level, but does not detect any structural changes in distribution for JPY. Bai and Perron's (1998) test can only detect structural changes for CNY at the 10% significance level. This finding is consistent with the results of our derivative test $F_T^{(1)}$. Since both the $F_T^{(1)}$ test and Bai and Perron's (1998) test can only capture structural changes in mean, the results in Table 7 show that EUR, JPY, and CAD have no structural change in mean. Furthermore, $F_T^{(2)}$ rejects the null hypothesis of no structural changes in variance for EUR, JPY, and CAD at the 10% significance level, and CNY at the

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

TABLE 7. Tests for structural breaks in exchange rate returns.

	F_T			$F_T^{(1)}$			$F_T^{(2)}$			In01			BP_{het}		
	F_T	5%	10%	$F_T^{(1)}$	5%	10%	$F_T^{(2)}$	5%	10%	In01	5%	10%	BP_{het}	5%	10%
EUR	**11.35**	8.41	7.34	5.12	13.83	10.53	**22.65**	24.22	17.13	**0.10**	0.09	0.06	4.52	10.17	8.78
JPY	**7.37**	8.03	7.30	5.67	13.55	11.97	**13.15**	14.54	11.59	0.06	0.07	0.06	4.83	10.17	8.78
CNY	**54.35**	18.89	17.15	**24.29**	30.80	24.13	**84.03**	37.01	26.69	**0.54**	0.20	0.16	**9.32**	10.17	8.78
CAD	**12.08**	7.59	6.38	6.39	10.70	9.41	**19.07**	20.74	16.54	**0.08**	0.07	0.06	4.82	10.17	8.78

Notes: (i) F_T, $F_T^{(1)}$, and $F_T^{(2)}$ denote the results of our joint test, and derivative tests for structural breaks in the first and second moments, respectively; (ii) In01 and BP_{het} denote the results of Inoue's (2001) test and Bai and Perron's (1998) serial correlation and heteroskedasticity robust Dmax-F test, respectively; (iii) Columns under F_T, $F_T^{(1)}$, $F_T^{(2)}$, In01, and BP_{het} report the values of the corresponding test statistics; and (iv) Columns under "5%" and "10%" report the corresponding bootstrapped critical values or asymptotic critical values. Bold entries indicate significance at the 10% significance level.

TABLE 8. Number of breaks for exchange rate returns.

	Distribution		Y_t			Y_t^2		
	DIC	DST5	LWZ	BIC	ST5	LWZ	BIC	ST5
EUR	4	3	0	0	0	2	3	1
JPY	1	0	0	0	0	0	2	1
CNY	4	1	0	4	0	1	3	2
CAD	5	4	0	0	0	2	3	4

Notes: (i) DIC and DST5 denote the information criterion and the sequential tests proposed in this paper under 5% significance level; (ii) BIC and LWZ denote the Schwarz criterion and the modified Schwarz criterion proposed in Liu et al. (1997); (iii) ST5 denotes Bai and Perron's (1998) sequential tests under 5% significance level; and (iv) the main entries report the number of breaks determined by the corresponding methods.

5% significance level. Hence, our derivative test $F_T^{(2)}$ documents the existence of structural breaks in variance for all four exchange rate returns. Moreover, we note that the signal of rejecting the null hypothesis of no structural break in distribution is stronger than that of rejecting no structural break in variance. For example, for both EUR and CAD, we can reject the null hypothesis of no structural break in distribution at the 5% significance level, but can only reject the null hypothesis of no structural break in variance at the 10% significance level. This implies that there may exist structural breaks in higher-order moments.

Once we have detected structural breaks in distribution for exchange rate returns, we need to move on to estimating the number and location of breaks. Table 8 reports the number of breaks determined by the proposed IC and sequential test at the 5% significance level. For comparison, we also report the number of breaks determined by LWZ and Bai and Perron's (1998) sequential tests, and the number of breaks in the first two moments of Y_t. As shown in Table 8, our method detects more breaks. In particular, the number of breaks detected in second moment is less than that in distribution. Intuitively, this implies that structural breaks may have occurred in higher-order moments of the exchange rate returns. Moreover, the number of breaks determined by IC differs from that determined by our sequential test.

The detected break dates are shown in Figures 2 and 3. We mark the break dates determined by both our IC and sequential test at the 5% significance level with solid lines, and those determined by IC only with dashed lines. These break dates are related to some important economic events. For EUR, three break dates are detected by the sequential test at 5% significance level, which are January 2002, July 2008, and October 2011, respectively. An additional break in January 2009 is detected by IC. Although Euro was introduced in January 1999, it was only used as a "digital currency" before 2002. It was officially circulated as a new currency in January 2002, which coincides with the first identified break date. The break dates of July 2008 and January 2009 might have been caused by the 2008 financial crisis. And the break date of October 2011 was related to the deterioration of the

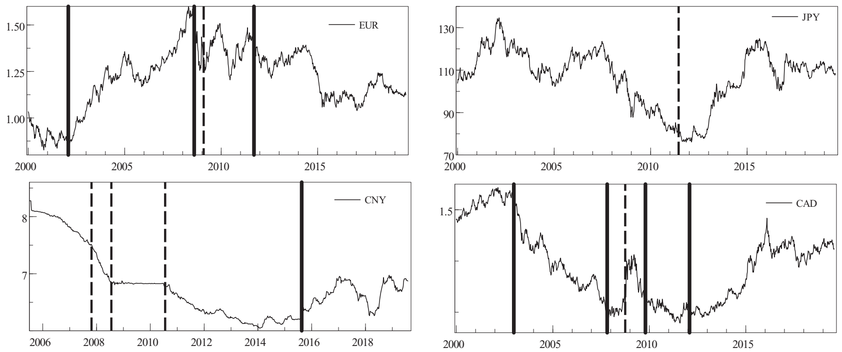

FIGURE 3. Time series plots of exchange rates and break dates.

Greek sovereign debt crisis. For JPY, we only find one break in April 2011, which may be due to the 2011 Earthquake off the Pacific coast of Tohoku which occurred in March 2011. We identify four break dates for CNY, which are October 2007, July 2008, June 2010, and August 2015, respectively. In May 2007, the People's Bank of China (PBC) announced the intention to expand the daily fluctuation range of CNY to USD from 0.3% to 0.5%. Since then, the CNY has increased its fluctuations gradually, and this corresponds to the first break date of October 2007. To cope with the impact of the 2008 financial crisis, PBC narrowed the fluctuation range of CNY, and the CNY to USD rate was stable at around 6.85 during the period from July 2008 to June 2010. On June 19, 2010, PBC decided to promote the exchange rate reform further. These events coincide with the identified break dates of July 2008 and June 2010. The last break date of CNY coincides with the exchange rate reform occurring on August 11, 2015. For CAD, we detect structural breaks in December 2002, October 2007, September 2008, October 2009, and January 2012, respectively. The Canadian dollar is widely viewed as a commodity currency. The first break date of December 2002 coincides with the period when oil prices began to rise. And at the last break date of January 2012, the oil prices fluctuated and started to fall from its peak. As a result, the Canadian dollar began to appreciate from December 2002 and depreciate from January 2012. The other three break dates (October 2007, September 2008, and October 2009) coincide with the start, outbreak, and recovery of the U.S. subprime debt crisis. The CAD was very volatile during the periods of September 2008 to October 2009 and tended to be stable after October 2009.

As suggested by one referee, we also checked for possible structural breaks in higher-order moments after accounting for changes in mean and variance. We divided the whole sample into some subsamples according to the estimated break dates using IC. Then we standardized the time series within each subsample. Specifically, for each exchange rate return, we define $Z_t = \left(Y_t - \bar{Y}_k\right)/\sqrt{V_k}$ if the tth observation belongs to the kth regime, where Y_t denotes the weekly exchange

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

Table 9. Tests for structural breaks in the transformed time series.

	Number of Breaks			F_T			$F_T^{(3)}$			$F_T^{(4)}$		
	DIC	DST5	DST10	F_T	5%	10%	$F_T^{(3)}$	5%	10%	$F_T^{(4)}$	5%	10%
EUR	1	1	1	**10.32**	7.12	6.49	3.94	11.13	10.02	**19.68**	15.27	11.18
JPY	1	0	1	**8.39**	8.41	7.15	5.50	12.30	8.83	4.49	13.47	10.59
CNY	3	1	1	**19.17**	9.21	8.76	22.63	881.09	540.11	22.19	2301.30	767.99
CAD	1	1	1	**14.00**	6.66	6.15	**9.03**	10.50	8.86	**16.15**	12.34	10.29

Notes: (i) DIC, DST5, and DST10 denote the information criterion and the sequential tests proposed in this paper under 5% and 10% significance levels; (ii) F_T, $F_T^{(3)}$, and $F_T^{(4)}$ denote the results of our joint test, and derivative tests for structural breaks in the third and fourth moments, respectively; (iii) columns under F_T, $F_T^{(3)}$, and $F_T^{(4)}$ report the values of the corresponding test statistics; and (iv) columns under "5%" and "10%" report the corresponding bootstrapped critical values. Bold entries indicate significance at the 10% significance level.

rate return, and $\bar{Y}_k$ and V_k are the sample mean and variance of the kth subsample, respectively. Hence, we have purged the possible structural breaks in the first two moments. If the detected breaks solely result from a changing mean or variance, the transformed sample should become strictly stationary. We then tested for structural breaks of the transformed series Z_t using our joint test and derivative tests $F_T^{(3)}$ and $F_T^{(4)}$. Table 9 reports the number of breaks determined by our IC and sequential tests at the 5% and 10% significance levels, and the results of various tests for the transformed time series. As shown in Table 9, both our IC and sequential tests identify additional structural breaks in the transformed series. Our joint test also rejects the null of no distributional structural change for the transformed series. Moreover, the derivative tests detect structural changes in the fourth moment for EUR and structural changes in the third and fourth moments for CAD. We cannot identify structural breaks in the third and fourth moments for JPY and CNY, which may be due to the finite sample problem. Table 10 summarizes the sample mean, variance, skewness, and kurtosis for all four exchange rate return series during different periods. It shows that in addition to structural breaks in mean and variance, the higher-order moments such as the third and fourth moments captured by skewness and kurtosis vary across different time periods.

8. CONCLUSION

In this paper, we propose an ECF approach to estimating and testing multiple structural breaks in distribution with unknown break dates for a multivariate time series. Based on the equivalence between CDF and CF, we can characterize structural breaks in distribution by a pseudo generalized regression representation in the frequency domain, which has an interesting interpretation of a complex-valued functional time series regression. By minimizing the SSGR in the generalized regression, we can consistently estimate the breaks and derive the convergence

Table 10. Summary statistics for the exchange rate returns during different periods.

	Period	Mean	Variance	Skewness	Kurtosis
EUR	2000/1/1–2002/1/25	−0.1633	1.7437	0.3837	2.6565
	2002/1/23–2008/7/11	0.1754	0.9290	−0.3774	3.1100
	2008/7/14–2009/1/16	−0.7098	6.0386	0.9091	4.0264
	2009/1/19–2011/10/28	0.0292	1.7151	−0.0007	2.7635
	2011/10/31–2019/7/10	−0.0547	0.8112	−0.0970	3.5439
JPY	2000/1/1–2011/4/29	−0.0425	1.2754	−0.3350	3.5267
	2011/5/2–2019/7/ 10	0.0657	1.0676	0.3716	5.8679
CNY	2000/1/1–2007/10/12	−0.0659	0.0125	−0.6634	5.2058
	2007/10/15–2008/7/11	−0.2396	0.0441	−0.1129	2.1544
	2008/7/14–2010/6/11	−0.0028	0.0166	1.6393	19.0296
	2010/6/14–2015/7/31	−0.0355	0.0383	0.5142	6.4642
	2015/8/3–2019/7/10	0.0496	0.2350	0.4054	5.2287
CAD	2000/1/1–2002/12/6	0.0641	0.3533	0.1866	3.4344
	2002/12/9–2007/10/19	−0.1870	0.7543	0.1284	2.5776
	2007/10/22–2008/9/12	0.1929	1.8651	0.1206	3.7408
	2008/9/15–2009/10/23	−0.0007	5.6819	0.1835	3.2248
	2009/10/26–2012/1/27	−0.0363	1.1365	0.5159	2.7087
	2012/1/30–2019/7/10	0.0676	0.6813	0.0547	3.2788

Notes: The main entries report the sample mean, variance, skewness, and kurtosis.

rate for the estimated break fractions, which does not depend on the dimension of the multivariate time series. We propose a sup-F type test for multiple structural breaks. We also propose a BIC-type information criterion and a sequential testing procedure to determine the number of breaks. As an advantage of using the ECF, we can take derivatives of the ECF to develop a class of derivative tests to gauge possible sources of structural breaks, which can deliver a similar version of Bai's (1994) test for breaks in mean as a special case of our approach. When testing for breaks in nonparametric regression, a class of derivative tests avoids smoothed nonparametric regression and is asymptotically more efficient than existing tests such as Vogt (2015) and Fu and Hong (2019). We propose and justify a moving block bootstrap procedure to obtain the critical values of the proposed tests. Simulations studies show that our method successively estimates each break fraction, and the proposed tests have reasonable size and excellent power. In an application to the exchange rate markets, we find significant evidence of structural breaks in distribution, which may be ignored by Bai and Perron's (1998) test. It appears that most structure breaks in exchange rate returns occur in variance and higher-order moments rather than in mean.

https://doi.org/10.1017/S026646662200010X Published online by Cambridge University Press

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit: https://doi.org/10.1017/S026646662200010X

REFERENCES

Andrews, D.W. & R.C. Fair (1988) Inference in nonlinear econometric models with structural change. *Review of Economic Studies* 55, 615–640.

Aston, J.A. & C. Kirch (2012) Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis* 109, 204–220.

Aue, A., R. Gabrys, L. Horváth, & P. Kokoszka (2009) Estimation of a change-point in the mean function of functional data. *Journal of Multivariate Analysis* 100, 2254–2269.

Aue, A., G. Rice, & O. Sönmez (2018) Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 80, 509–529.

Bai, J. (1994) Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis* 15, 453–472.

Bai, J. (1996) Testing for parameter constancy in linear regressions: An empirical distribution function approach. *Econometrica* 64, 597–622.

Bai, J. (1997) Estimation of a change point in multiple regression models. *Review of Economics and Statistics* 79, 551–563.

Bai, J. & S. Ng (2002) Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

Bai, J. & P. Perron (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

Berkes, I., R. Gabrys, L. Horváth, & P. Kokoszka (2009) Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71, 927–946.

Bierens, H.J. & L. Wang (2012) Integrated conditional moment tests for parametric conditional distributions. *Econometric Theory* 28, 328–362.

Boldea, O. & A.R. Hall (2013) Estimation and inference in unstable nonlinear least squares models. *Journal of Econometrics* 172, 158–167.

Boothe, P. & D. Glassman (1987) The statistical distribution of exchange rates: Empirical evidence and economic implications. *Journal of International Economics* 22, 297–319.

Busetti, F. & A. Harvey (2011) When is a copula constant? A test for changing relationships. *Journal of Financial Econometrics* 9, 106–131.

Chacko, G. & L.M. Viceira (2003) Spectral GMM estimation of continuous-time processes. *Journal of Econometrics* 116, 259–292.

Chen, B. (2015) Modeling and testing smooth structural changes with endogenous regressors. *Journal of Econometrics* 185, 196–215.

Chen, B. & Y. Hong (2012) Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica* 80, 1157–1183.

Chow, G.C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28, 591–605.

Cowell, F.A. (1985) Measures of distributional change: An axiomatic approach. *The Review of Economic Studies* 52, 135–151.

Delgado, M.A. & J. Hidalgo (2000) Nonparametric inference on structural breaks. *Journal of Econometrics* 96, 113–144.

Diebold, F.X., T.A. Gunther, & A.S. Tay (1998) Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39, 863–883.

Diebold, F.X., J. Hahn, & A.S. Tay (1999) Multivariate density forecast evaluation and calibration in financial risk management: High-frequency returns on foreign exchange. *Review of Economics and Statistics* 81, 661–673.

Dümbgen, L. (1991) The asymptotic behavior of some nonparametric change-point estimators. *Annals of Statistics* 19, 1471–1495.

Epps, T. (1987) Testing that a stationary time series is Gaussian. *Annals of Statistics* 15, 1683–1698.

Epps, T. (1988) Testing that a Gaussian process is stationary. *Annals of Statistics* 16, 1667–1683.

Fan, Y. & Q. Li (1999) Root-n-consistent estimation of partially linear time series models. *Journal of Nonparametric Statistics* 11, 251–269.

Fengler, M.R., E. Mammen, & M. Vogt (2015) Specification and structural break tests for additive models with applications to realized variance data. *Journal of Econometrics* 188, 196–218.

Feuerverger, A. & R.A. Mureika (1977) The empirical characteristic function and its applications. *Annals of Statistics* 5, 88–97.

Fu, Z. & Y. Hong (2019) A model-free consistent test for structural change in regression possibly with endogeneity. *Journal of Econometrics* 211, 206–242.

Galí, J. & M. Gertler (1999) Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* 44, 195–222.

González-Rivera, G. & Y. Sun (2017) Density forecast evaluation in unstable environments. *International Journal of Forecasting* 33, 416–432.

Hall, A.R., S. Han, & O. Boldea (2012) Inference regarding multiple structural changes in linear models with endogenous regressors. *Journal of Econometrics* 170, 281–302.

Hansen, B.E. (2001) The new econometrics of structural change: Dating breaks in US labor productivity. *Journal of Economic Perspectives* 15, 117–128.

Hlávka, Z., M. Hušková, C. Kirch, & S.G. Meintanis (2012) Monitoring changes in the error distribution of autoregressive models based on Fourier methods. *TEST* 21, 605–634.

Hlávka, Z., M. Hušková and S.G. Meintanis (2017) Change point detection with multivariate observations based on characteristic functions. In D. Ferger, W. González Manteiga, T. Schmidt, JL Wang (eds.), *From Statistics to Mathematical Finance*, pp. 273–290. Springer.

Hlávka, Z., M. Hušková & S.G. Meintanis (2020) Change-point methods for multivariate time-series: Paired vectorial observations. *Statistical Papers* 61, 1351–1383.

Hong, Y. (1999) Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. *Journal of the American Statistical Association* 94, 1201–1220.

Hong, Y. & T.-H. Lee (2003) Diagnostic checking for the adequacy of nonlinear time series models. *Econometric Theory* 19, 1065–1121.

Hong, Y., X. Wang, & S. Wang (2017) Testing strict stationarity with applications to macroeconomic time series. *International Economic Review* 58, 1227–1277.

Hörmann, S. & P. Kokoszka (2010) Weakly dependent functional data. *Annals of Statistics* 38, 1845–1884.

Hsieh, D.A. (1988) The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of International Economics* 24, 129–145.

Hušková, M. & S.G. Meintanis (2006) Change point analysis based on empirical characteristic functions. *Metrika International Journal for Theoretical and Applied Statistics* 63, 145–168.

Hušková, M. & S.G. Meintanis (2008) Tests for the multivariate k-sample problem based on the empirical characteristic function. *Journal of Nonparametric Statistics* 20, 263–277.

Inoue, A. (2001) Testing for distributional change in time series. *Econometric Theory* 17, 156–187.

Jiang, G.J. & J.L. Knight (2002) Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business and Economic Statistics* 20, 198–212.

Jiang, L., X. Wang, & J. Yu (2018) New distribution theory for the estimation of structural break point in mean. *Journal of Econometrics* 205, 156–176.

Kapetanios, G. (2009) Testing for strict stationarity in financial variables. *Journal of Banking and Finance* 33, 2346–2362.

Knight, J.L. & J. Yu (2002) Empirical characteristic function in time series estimation. *Econometric Theory* 18, 691–721.

Koedijk, C.G., M. Schafgans, & C.G. de Vries (1990) The tail index of exchange rate returns. *Journal of International Economics* 29, 93–108.

Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17, 1217–1241.

Lin, C.-H. & T.-C. Kao (2008) Multiple structural changes in the tail behavior: Evidence from stock index futures returns. *Nonlinear Analysis: Real World Applications* 9, 1702–1713.

Liu, C.Y. & J. He (1991) A variance-ratio test of random walks in foreign exchange rates. *The Journal of Finance* 46, 773–785.

Liu, J., S. Wu, & J.V. Zidek (1997) On segmented multivariate regression. *Statistica Sinica* 7, 497–525.

Loader, C.R. (1996) Change point estimation using nonparametric regression. *Annals of Statistics* 24, 1667–1678.

Loretan, M. & P.C.B. Phillips (1994) Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of Empirical Finance* 1, 211–248.

Manner, H., F. Stark, & D. Wied (2019) Testing for structural breaks in factor copula models. *Journal of Econometrics* 208, 324–345.

Müller, H.-G. (1992) Change-points in nonparametric regression analysis. *Annals of Statistics* 20, 737–761.

Pinkse, J. (1998) A consistent nonparametric test for serial independence. *Journal of Econometrics* 84, 205–231.

Politis, D.N. & H. White (2004) Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23, 53–70.

Qu, Z. (2008) Testing for structural change in regression quantiles. *Journal of Econometrics* 146, 170–184.

Qu, Z. & P. Perron (2007) Estimating and testing structural changes in multivariate regressions. *Econometrica* 75, 459–502.

Quintos, C., Z. Fan, & P.C.B. Phillips (2001) Structural change tests in tail behaviour and the Asian crisis. *Review of Economic Studies* 68, 633–663.

Rime, D., L. Sarno, & E. Sojli (2010) Exchange rate forecasting, order flow and macroeconomic information. *Journal of International Economics* 80, 72–88.

Rossi, B. (2013) Advances in forecasting under instability. In G. Elliott, A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 2, no. 2, pp. 1203–1324. Elsevier.

Rossi, B. & T. Sekhposyan (2013) Conditional predictive density evaluation in the presence of instabilities. *Journal of Econometrics* 177, 199–212.

Selk, L. & N. Neumeyer (2013) Testing for a change of the innovation distribution in nonparametric autoregression: the sequential empirical process approach. *Scandinavian Journal of Statistics* 40, 770–788.

Shephard, N.G. (1991) From characteristic function to distribution function: a simple framework for the theory. *Econometric Theory* 7, 519–529.

Singleton, K.J. (2001) Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics* 102, 111–141.

Stock, J.H. & M.W. Watson (1996) Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* 14, 11–30.

Su, L. & H. White (2007) A consistent characteristic function-based test for conditional independence. *Journal of Econometrics* 141, 807–834.

Su, L. & Z. Xiao (2008) Testing structural change in time-series nonparametric regression models. *Statistics and Its Interface* 1, 347.

Vogt, M. (2015) Testing for structural change in time-varying nonparametric regression models. *Econometric Theory* 31, 811–859.

Wang, X. & Y. Hong (2018) Characteristic function based testing for conditional independence: A nonparametric regression approach. *Econometric Theory* 34, 815–849.

Wu, J. & C. Chu (1993) Kernel-type estimators of jump points and values of a regression function. *Annals of Statistics* 21, 1545–1566.

Yakir, B. (1997) A note on optimal detection of a change in distribution. *Annals of Statistics* 25, 2117–2126.

Zhang, C., D.R. Osborn, & D.H. Kim (2008) The new Keynesian Phillips curve: From sticky inflation to sticky prices. *Journal of Money, Credit and Banking* 40, 667–699.

Zhou, Y., L. Fu, & B. Zhang (2017) Two non parametric methods for change-point detection in distribution. *Communications in Statistics-Theory and Methods* 46, 2801–2815.

Zou, C., G. Yin, L. Feng, & Z. Wang (2014) Nonparametric maximum likelihood approach to multiple change-point problems. *Annals of Statistics* 42, 970–1002.