Purdue University Economics Department Working Paper No **1308**

Bootstrap Procedures for Detecting Multiple Persistence Shifts in a Heteroskedastic Time Series

Mohitosh Kejriwal* **Xuewen Yu†**

Purdue University *Purdue University*

November 30, 2018

Abstract

This paper proposes new bootstrap procedures for detecting multiple persistence shifts in a time series driven by nonstationary volatility. The assumed volatility process can accommodate discrete breaks, smooth transition variation as well as trending volatility. We develop wild bootstrap sup-Wald tests of the null hypothesis that the process is either stationary [$I(0)$] or has a unit root [$I(1)$] throughout the sample. We also propose a sequential procedure to estimate the number of persistence breaks based on ordering the regime-specific bootstrap p-values. The asymptotic validity of the advocated procedures is established both under the null of stability and a variety of persistence change alternatives. Monte Carlo simulations support the use of a non-recursive scheme for generating the $I(0)$ bootstrap samples and a partially recursive scheme for generating the $I(1)$ bootstrap samples, especially when the data generating process contains an $I(1)$ segment. A comparison with existing tests illustrates the finite sample improvements offered by our methods in terms of both size and power. An application to OECD inflation rates is included.

Keywords: heteroskedasticity, multiple structural changes, sequential procedure, unit root, Wald tests, wild bootstrap

JEL Classification: C22

*Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907 (mkejriwa@purdue.edu).

†Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907 (yu656@purdue.edu).

1 Introduction

Distinguishing among different forms of nonstationarity has been a topic of long-standing interest in time series econometrics. For instance, it is well known that a standard t-test for the presence of a deterministic trend diverges with the sample size if the process contains a unit root, i.e., a stochastic trend. Moreover, tests for structural change in the deterministic trend are sensitive to whether the stochastic component is assumed to be stationary or not (Harvey et al., 2009; Perron and Yabu, 2009). Perron (1989) showed that the presence of infrequent shifts in the trend function biases unit root tests toward the unit root hypothesis if these tests do not account for such shifts. Similarly, Diebold and Inoue (2001) provide analytical as well as Monte Carlo evidence to illustrate that models with stochastic regime switching can generate long memory behavior by inducing an upward bias in the estimate of the long memory or fractional differencing parameter. Kejriwal and Perron (2010) demonstrate that standard structural break detection procedures (such as information criteria or sequential testing) applied to a regression with integrated variables tend to select the maximum permissible number of breaks if the regression is spurious.

An important related problem concerns inference regarding the conditional mean in the presence of nonstationarity in variance. Structural changes in variance have been extensively documented for macroeconomic and financial time series; see, *inter alia*, McConnell and Perez Quiros (2000) and Sensier and van Dijk (2004) for empirical evidence on a variety of macroeconomic variables, and Loretan and Phillips (1994) for evidence on stock market returns and exchange rates. The non-robustness of standard unit root tests to nonstationary volatility has been well established both theoretically and numerically by Cavaliere (2005) and Cavaliere and Taylor (2007, 2008a, 2009) while Cavaliere and Taylor (2005) provide similar evidence in the context of standard stationarity tests. The basic reason is that these procedures are based on a so-called "global homoskedasticity" assumption which typically fails when the time series is generated by innovations with nonstationary volatility (Cavaliere and Taylor, 2009).

A relatively small literature has also addressed the problem of discriminating between instability in the conditional mean and nonstationary volatility. Hansen (2000) shows that standard structural change tests (e.g., Andrews, 1993) in the linear regression model do not have the correct size in the presence of nonstationarity in variance and develops a fixed-regressor bootstrap procedure to remedy the problem. Pitarakis (2004) documents the extent of size distortions suffered by the stability tests of Andrews (1993) with substantial oversizing for versions that do not correct for heteroskedasticity and substantial under-sizing for heteroskedasticity-consistent versions. Zhou and Perron (2008) provide simulation evidence which shows that ignoring a deterministic shift in variance biases the supremum likelihood ratio test for stability of the conditional mean towards rejection of the null hypothesis with the extent of size distortions increasing in the size of the variance break. Perron and Zhou (2008) develop likelihood ratio tests of the joint hypothesis that the coefficients and error variance in a linear regression model with stationary regressors are stable. Xu (2015) demonstrates that cumulative sum (CUSUM) tests for structural change in the mean have incorrect size under variance nonstationarity. These studies convincingly establish that it is important from an empirical standpoint to allow for the possibility of nonstationary volatility when testing the stability of the conditional mean.

This paper is concerned with the problem of testing for structural changes in the persistence of a univariate time series driven by nonstationary volatility. We are interested in changes which involve switches between unit root [$I(1)$] and stationary [$I(0)$] behavior as well as changes that preserve the $I(0)$ nature of the time series across regimes. While a plethora of tests for persistence change is available in the literature, most of these procedures are based on the global homoskedasticity assumption; for instance, see Kim (2000), Busetti and Taylor (2004), Harvey et al. (2006) for tests allowing a single break and Bai and Perron (1998, BP henceforth), Leybourne et al. (2007), Kejriwal et al. (2013, KPZ henceforth) for tests allowing multiple breaks and Kejriwal (2018) for procedures to determine the number of breaks. Diebold and Chen (1996) provide simulation evidence suggesting that bootstrap versions of single-break tests developed in Andrews (1993) and Andrews and Ploberger (1994) achieve better size control than their asymptotic counterparts, again assuming error homoskedasticity.[1] Cavaliere and Taylor (2008b, CT henceforth) develop bootstrap tests robust to nonstationary volatility based on the ratio of partial sums of demeaned (or detrended) residuals. Their procedure assumes (1) the process is $I(0)$ under the null hypothesis of stability; (2) a single break under the alternative that is associated with either a $I(1)$-$I(0)$ or $I(0)$-$I(1)$ shift but not an $I(0)$-$I(0)$ shift; (3) a stable trend function under both the null and alternative hypotheses.

In this paper, we provide a comprehensive treatment of issues related to testing for structural changes in persistence when the innovations are globally heteroskedastic, i.e., fail the global homoskedasticity condition pervasive in this literature. Our approach is rather

[1] Their simulation design also assumes that the time series is $I(0)$ under the null hypothesis and regime-wise $I(0)$ under the alternative hypothesis.

general in that it allows: (1) an $I(1)$ *or* $I(0)$ null hypothesis; (2) multiple structural changes where the number and timing of the changes are unknown; (3) persistence change alternatives of the $I(1)$-$I(0)$, $I(0)$-$I(1)$ and $I(0)$-$I(0)$ forms which the testing procedure does not require prior knowledge of; (4) disentangling persistence shifts from shifts in the trend function. The assumed volatility process is of a general form that can accommodate single and multiple breaks, smooth transition variation as well as trending volatility. We develop sup-Wald tests based on a wild bootstrap scheme that delivers procedures with accurate size and satisfactory power properties. These procedures include tests for a specified number of changes as well as tests that assume the number of breaks is unknown (up to a known upper bound). We also propose a sequential approach for determining the number of persistence breaks based on ordering the bootstrap p-values. The finite sample adequacy of the advocated procedures including a comparison with existing tests is evaluated through an extensive set of Monte Carlo experiments.

It is important to emphasize that unlike Xu (2008) who proposes a recursive wild bootstrap scheme for conducting inference in an autoregressive model with stationary roots, the $I(0)$ bootstrap samples in our approach are generated only from the estimated residuals under the null hypothesis and not recursively using the parameter estimates. This is because the full sample estimate of the persistence parameter (the sum of the autoregressive coefficients) is biased towards unity when the series has at least one asymptotically non-negligible $I(1)$ segment. As a result, recursively generated bootstrap samples using such estimates is not only invalid when the null hypothesis is $I(1)$ (Basawa et al., 1991) but also leads to poor power properties when the alternative involves at least one $I(1)$ segment (Gulesserian and Kejriwal, 2014). Monte Carlo simulations confirm the presence of these features in finite samples. Our approach is similar to Hansen's (2000) fixed-regressor bootstrap scheme although the latter does not consider the possibility of a unit root null hypothesis.

The rest of the paper is organized as follows. Section 2 lays out the modeling framework and the associated assumptions. Section 3 details the procedures for testing persistence change. The large sample effects of nonstationary volatility on persistence change tests constructed assuming homoskedastic innovations are studied in Section 4. The proposed bootstrap tests are presented in Section 5. Section 6 discusses extensions of the procedures to deal with deterministic trends as well as disentangling shifts in persistence from shifts in the trend function. Section 7 provides Monte Carlo evidence to assess the finite sample performance of the advocated procedures, Section 8 applies the proposed approach to test for persistence change in OECD inflation rates and Section 9 concludes. All proofs are collected in Appendix A. Additional Monte Carlo results are provided in Appendix B.

As a matter of notation, we will use $\mathcal{C} = \mathcal{C}[0,1]$ to denote the space of continuous functions on $[0,1]$ and $\mathcal{D}$ the space of right continuous with left limit processes on $[0,1]$, '$\xrightarrow{p}$' to denote convergence in probability, '$\xrightarrow{w}$' to denote weak convergence in the space $\mathcal{D}$ endowed with the Skorohod metric, and '$\xrightarrow{w}_p$' to denote weak convergence in probability under the bootstrap measure (Giné and Zinn, 1990). Further, $B_1(.)$ and $B_2(.)$ denote standard independent Brownian motions on $[0,1]$ and $B(.) = [B_1(.), B_2(.)]'$. Further, for any stochastic process $Z(.)$ defined over $[0,1]$, denote $Z^{(i)}(.)$ represent $Z(.)$ demeaned over $[\lambda_{i-1}, \lambda_i]$, i.e., $Z^{(i)}(r) = Z(r) - (\lambda_i - \lambda_{i-1})^{-1} \int_{\lambda_{i-1}}^{\lambda_i} Z,\ r \in [\lambda_{i-1}, \lambda_i]$. Finally, for brevity of presentation, all integrals of the form $\int_a^b f(r)dr$ are expressed as $\int_a^b f$.

2 The Persistence Change Model

Consider a univariate time series y_t generated by the AR(p) model

$$\left.\begin{array}{c} y_t = \mu_i + u_t \\ u_t = u_{T_{i-1}^0} + h_t \\ h_t = \alpha_i h_{t-1} + \sum_{j=1}^{p-1} \pi_{ij} \Delta h_{t-j} + e_t \\ h_{T_{i-1}^0} = \ldots = h_{T_{i-1}^0 - p+1} = 0 \end{array}\right\} \ t = T_{i-1}^0 + 1, T_{i-1}^0 + 2, \ldots, T_i^0;\ i = 1, \ldots, m+1 \qquad (1)$$

with the convention that $T_0^0 = 0$ and $T_{m+1}^0 = T$, where T is the sample size. The process is therefore subject to m breaks or $m+1$ regimes with break dates $(T_1^0, \ldots, T_m^0)$. Both the break dates and the number of breaks are assumed to be unknown. The same data generating process was considered by Leybourne et al. (2007) and Kejriwal (2018) and is designed to ensure that the successive $I(1)$ and $I(0)$ regimes join up at the breakpoints thereby avoiding the problem of spurious jumps to zero in u_t. While we assume a common lag length p across regimes, regime-specific lag lengths can be accommodated by interpreting p as the maximum lag length across the $(m+1)$ regimes.

Our analysis is based on the following assumptions on the process generating $\{y_t\}$:

Assumption A1: $T_i^0 = [T\lambda_i^0]$, where $0 < \lambda_1^0 < \ldots < \lambda_m^0 < 1$.

Assumption A2: All roots of the polynomial $\pi_i(L) = 1 - \pi_{i1}L - \pi_{i2}L^2 - \ldots - \pi_{i,p-1}L^{p-1}$ lie outside the unit circle.

Assumption A3: The error term e_t in (1) satisfies

$$e_t = \sigma_t \varepsilon_t \tag{2}$$

where $\{\varepsilon_t\}$ is an i.i.d. sequence with zero mean and unit variance and $\{\sigma_t\}$ is a strictly positive non-stochastic sequence. We assume $\sup_t E(\varepsilon_t^{4+\beta}) < \infty$ for some $\beta > 0$.

Assumption A4: For some strictly positive deterministic sequence $\{a_T\}$, we assume the sequence $\{\sigma_t\}$ satisfies

$$a_T^{-1}\sigma_{[Ts]} = g(s),\ \ s \in [0,1]$$

where $g(.) \in \mathcal{D}$ is a strictly positive, non-stochastic function with a finite number of points of discontinuity and satisfies a uniform first-order Lipschitz condition except at the points of discontinuity.

Assumption A1 facilitates the development of the asymptotic theory by requiring the breakpoints to be asymptotically distinct. Each segment is assumed to increase proportionately with the sample size. This requirement is standard in the structural change literature (see, e.g., Bai and Perron, 1998; 2003a). Assumption A2 corresponds to the requirement of at most one unit root in each regime while all remaining roots are stationary. Assumption A3 specifies that the stochastic process for $\{e_t\}$ is determined by the time-varying volatilities $\{\sigma_t\}$ and the i.i.d. innovation sequence $\{\varepsilon_t\}$ with a requirement on the existence of moments for the innovations (see, e.g., Xu, 2008). We follow Cavaliere and Taylor (2008a) in assuming that the innovations are i.i.d. although our results continue to hold under the weaker condition that for $\mathcal{F}_t = \sigma\text{-}field\{e_s, s \leq t\}$, $\{\varepsilon_t, \mathcal{F}_t\}$ is a martingale difference sequence satisfying (i) $E(\varepsilon_t^2) = 1$ for all t; (ii) $T^{-1}\sum_{t=1}^{T}\varepsilon_t^2 \xrightarrow{p} 1$; and (iii) $\sup_t E(\varepsilon_t^{4+\beta}) < \infty$ for some $\beta > 0$.

The key assumption that embodies global heteroskedasticity is Assumption A4. It allows the time series $\{y_t\}$ to be driven by a general class of nonstationary heteroskedastic errors that includes a wide variety of specifications for volatility considered in the literature. For instance, single or multiple volatility breaks, linearly trending volatility, piece-wise linear trends in variance and smooth transition shifts can all be shown to satisfy Assumption A4 with $a_T = 1$ and a particular choice of $g(.)$. Further, models of explosive deterministic volatility are allowed through appropriate choice of the scaling factor a_T (specifically, by letting a_T grow with the sample size). Cavaliere and Taylor (2008, 2009) provide a detailed discussion of the specific choices of a_T and $g(.)$ associated with each of the aforementioned volatility models.

Remark 1 *As in CT, the function $g(.)$ is assumed to be non-stochastic to enable simplification of the theoretical analysis. In particular, Assumption A4 rules out nonstationary autoregressive stochastic volatility (SV) models (Hansen,1995), SV models with jumps (Georgiev, 2008), "nonstationary nonlinear heteroskedastic" models with stochastically trending volatility (Park, 2002) and near-integrated GARCH models (Nelson,1990). This assumption can be potentially weakened to allow cases where the sequences $\{\sigma_t\}$ and $\{\varepsilon_t\}$ are stochastically independent. In such cases, our results can be interpreted as holding conditional on a given realization of $g(.)$, where $g(.)$ has sample paths satisfying Assumption A4.*

In order to accommodate $I(0)$ preserving persistence changes as in the framework of Bai and Perron (1998), we also consider the following data generating process for y_t:

$$\left.\begin{array}{c} y_t = \mu_i + u_t \\ u_t = \alpha_i u_{t-1} + \sum_{j=1}^{p-1} \pi_{ij} \Delta u_{t-j} + e_t \\ u_0 = ... = u_{-p+1} = 0 \end{array}\right\} \; t = T_{i-1}^0 + 1, T_{i-1}^0 + 2, ..., T_i^0;\ i = 1, ..., m+1 \tag{3}$$

The conditions stated in Assumptions A1-A4 are assumed to hold for (3) as well. Since the direction of persistence change is typically unknown in practice, it is important from a practical perspective for the testing framework to allow alternatives that involve switches between $I(1)$ and $I(0)$ regimes as captured by (1) as well as alternatives that do not involve a change in the order of integration across regimes as in (3).

3 Testing Procedures

The testing procedures recommended in this paper are based on bootstrap versions of the asymptotic procedures developed by BP and KPZ assuming conditional homoskedasticity which corresponds to setting $\sigma_t = \sigma$ in (2). Section 3.1 reviews the test statistics designed to detect a specified number of breaks while Section 3.2 outlines the procedures when the number of breaks is not specified. The proposed bootstrap versions of these tests will be described in Section 5.

3.1 Tests for a Specified Number of Breaks

We first consider the statistics designed to test the null hypothesis $H_0^{(1)}$: $\alpha_i = 1$ for all i in (1). The estimating regression takes the form

$$\Delta y_t = c_i + (\alpha_i - 1)y_{t-1} + \sum_{j=1}^{p-1} \pi_j \Delta y_{t-j} + e_t^* \tag{4}$$

with $c_i = (1-\alpha_i)[u_{T_{i-1}^0} + \mu_i]$ and e_t^* is the regression error. Under $H_0^{(1)}$, $c_i = 0$ for all i. The true lag order p is assumed unknown but can be estimated using standard information criteria such as the AIC or BIC. Further, the coefficients of the lagged differences in (4) are not allowed to change since, as argued in KPZ, the goal of the testing approach is to direct power against structural changes in the persistence parameter α_i and not changes in the short-run dynamics.[2]

Under the alternative, the following two models are considered depending on whether the initial regime contains a unit root or not:

Model 1a: $\alpha_i = 1$ in odd regimes and $|\alpha_i| < 1$ in even regimes.

Model 1b: $\alpha_i = 1$ in even regimes and $|\alpha_i| < 1$ in odd regimes.

Consider first the Wald test that applies when the alternative involves a fixed value $m = k$ of changes. Denote a candidate vector of break fractions by $\lambda = (\lambda_1, ..., \lambda_k)$ and the alternative hypotheses corresponding to models 1a and 1b as $H_{a,k}^{(1)}$ and $H_{b,k}^{(1)}$, respectively. The test against $H_{a,k}^{(1)}$ is defined as

$$\begin{aligned} F_{1a}(\lambda, k) &= (T-k)(SSR_0^{(1)} - SSR_{1a,k}^{(1)})/[kSSR_{1a,k}^{(1)}] \text{ if } k \text{ is even} \\ F_{1a}(\lambda, k) &= (T-k-1)(SSR_0^{(1)} - SSR_{1a,k}^{(1)})/[(k+1)SSR_{1a,k}^{(1)}] \text{ if } k \text{ is odd} \end{aligned} \tag{5}$$

while that against $H_{b,k}^{(1)}$ is defined as

$$\begin{aligned} F_{1b}(\lambda, k) &= (T-k-2)(SSR_0^{(1)} - SSR_{1b,k}^{(1)})/[(k+2)SSR_{1b,k}^{(1)}] \text{ if } k \text{ is even} \\ F_{1b}(\lambda, k) &= (T-k-1)(SSR_0^{(1)} - SSR_{1b,k}^{(1)})/[(k+1)SSR_{1b,k}^{(1)}] \text{ if } k \text{ is odd} \end{aligned} \tag{6}$$

In (5) and (6), $SSR_0^{(1)}$ denotes the sum of squared residuals under $H_0^{(1)}$ while $SSR_{1a,k}^{(1)}$ and

[2]Simulation evidence presented in KPZ illustrates that the test statistics do not have much power against pure changes in short-run dynamics but are powerful when there is a change in both persistence and these dynamics.

$SSR_{1b,k}^{(1)}$ denote, respectively, the sum of squared residuals obtained from estimating (4) under the restrictions imposed by Model 1a and Model 1b. For some arbitrary small positive number ϵ, we define the set $\Lambda_\epsilon^k = \{\lambda : |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_k \leq 1 - \epsilon\}$. The sup-Wald tests are then defined as

$$F_{1a}(k) = \sup_{\lambda \in \Lambda_\epsilon^k} F_{1a}(\lambda, k), \quad F_{1b}(k) = \sup_{\lambda \in \Lambda_\epsilon^k} F_{1b}(\lambda, k).$$

When the direction of persistence change is unknown, the relevant test statistic is given by

$$W_1(k) = \max[F_{1a}(k), F_{1b}(k)]$$

While the foregoing tests are based on the $I(1)$ null hypothesis, the stable $I(0)$ null can be tested by employing the BP procedure. This amounts to testing $H_0^{(0)}$: $c_i = c,\ \alpha_i = \alpha$, for all i with $|\alpha| < 1$ in (4). The relevant alternative hypothesis within the BP framework is $H_{1,k}^{(0)}$: $\alpha_1 \neq \alpha_2 \neq ... \neq \alpha_{k+1}$, $|\alpha_i| < 1$ for all i. The time series is thus regimewise-$I(0)$ under $H_{1,k}^{(0)}$. The BP test for a fixed number $m = k$ changes is given by

$$G_1(\lambda, k) = [T - 2(k+1)](SSR_0^{(0)} - SSR_{1,k}^{(0)})/[kSSR_{1,k}^{(0)}] \tag{7}$$

In (7), $SSR_0^{(0)}$ denotes the sum of squared residuals under $H_0^{(0)}$ while $SSR_{1,k}^{(0)}$ denotes the sum of squared residuals obtained from *unrestricted* OLS estimation of (4). The BP test is then defined as $G_1(k) = \sup_{\lambda \in \Lambda_\epsilon^k} G_1(\lambda, k)$.

Under conditional homoskedasticity, the limiting distributions of $F_{1a}(k)$, $F_{1b}(k)$ and $W_1(k)$ are pivotal under $H_0^{(1)}$ while that of $G_1(k)$ is pivotal under $H_0^{(0)}$, given the trimming choice ϵ. Asymptotic critical values are tabulated by KPZ and BP for different choices of ϵ. Further, the tests $F_{1a}(k)$ and $F_{1b}(k)$ are consistent under $H_{a,k}^{(1)}$ and $H_{b,k}^{(1)}$, respectively while $W_1(k)$ and $G_1(k)$ are consistent under each of $H_{a,k}^{(1)}$, $H_{b,k}^{(1)}$ and $H_{1,k}^{(0)}$.

As shown in Kejriwal (2018), $G_1(k)$ does not have correct asymptotic size under $H_0^{(1)}$ even in the homoskedastic case. On the other hand, the KPZ tests $[F_{1a}(k), F_{1a}(k), W_1(k)]$ diverge to positive infinity under $H_0^{(0)}$. In order to control asymptotic size when the process is either $I(1)$ or $I(0)$ under the null hypothesis, KPZ propose simultaneous application of their test and the BP test. Let $H_0 = H_0^{(1)} \cup H_0^{(0)}$. For a given significance level η, the KPZ statistic for testing H_0 is given by

$$H(k, \eta) = \min\left[W_1(k), \frac{cv_{w,k}(\eta)}{cv_{g,k}(\eta)} G_1(k)\right]$$

where, at level η, $cv_{w,k}(\eta)$ and $cv_{g,k}(\eta)$ are the critical values of the statistics $W_1(k)$ and $G_1(k)$, respectively. Under conditional homoskedasticity, the decision rule of rejecting H_0 when $H(k,\eta) > cv_{w,k}(\eta)$ has asymptotic size at most η under H_0 and unit asymptotic power against each of $H_{a,k}^{(1)}$, $H_{b,k}^{(1)}$ and $H_{1,k}^{(0)}$. The computation of $G_1(.)$ and $W_1(.)$ can be accomplished using the dynamic programming algorithms proposed in Bai and Perron (1998, 2003) and Perron and Qu (2006), respectively.

3.2 Tests when the Number of Breaks is Unknown

When the number of breaks is unknown up to an upper bound A, KPZ propose the following test statistic directed to detect processes which alternate between $I(1)$ and $I(0)$ regimes:

$$Wmax_1 = \max_{1 \leq k \leq A} W_1(k)$$

Similarly, for detecting $I(0)$-preserving changes, BP propose the statistic

$$UDmax_1 = \max_{1 \leq k \leq A} G_1(k)$$

To achieve correct size under H_0, KPZ also suggest the statistic

$$Hmax_1(\eta) = \min\left[Wmax_1, \frac{cv_{w,max}(\eta)}{cv_{g,max}(\eta)} UDmax_1\right]$$

where, at level η, $cv_{w,max}(\eta)$ and $cv_{g,max}(\eta)$ are the critical values of $Wmax_1$ and $UDmax_1$ respectively. The decision rule is to reject H_0 if $H\max_1(\eta) > cv_{w,max}(\eta)$. The tests $Wmax_1$, $UDmax_1$ and $Hmax_1(\eta)$ are all consistent under each of $H_{a,m}^{(1)}$, $H_{b,m}^{(1)}$ and $H_{1,m}^{(0)}$, where $m \leq A$ is the true number of structural changes.

4 The Large Sample Effects of Nonstationary Volatility

This section studies the large sample behavior of the KPZ and BP tests when the time series is subject to nonstationary volatility as specified in Assumptions A3 and A4. In particular, we show that the null limiting distributions of the tests are not pivotal and depend on the sample path of the volatility process $g(.)$. Therefore these tests do not have correct asymptotic size unless $g(.)$ is a constant. Since the KPZ tests are correctly sized under the $I(1)$ null hypothesis $H_0^{(1)}$ when conditional homoskedasticity holds, we study their large sample behavior under $H_0^{(1)}$ when the time series satisfies Assumptions A1-A4. Similarly, the effects of nonstationary volatility on the BP tests are investigated under $H_0^{(0)}$. Under

$H_0^{(0)}$, the KPZ tests diverge to positive infinity while under $H_0^{(1)}$, the BP tests have incorrect asymptotic size even when conditional homoskedasticity holds. It can be shown that these properties continue to hold under Assumptions A1-A4. Monte Carlo evidence indicates that the extent of size distortions in finite samples can be considerable (see Table B-1 in Appendix B).

For $r \in [0,1]$, let $\widetilde{g}(r) = \left(\int_0^r g(s)^2\right)^{1/2}$, $B_{g,1}(r) = \widetilde{g}(1)^{-1}\int_0^r g(s)dB_1(s)$, $B_{g,2}(r) = \widetilde{g}(1)^{-1}\int_0^r g(s)^2 dB_2(s)$. We first state the result for the KPZ tests under $H_0^{(1)}$:

Theorem 1 *Suppose Assumptions A1-A4 hold. Then, under $H_0^{(1)}$, if k is even, we have*

$$F_{1a}(\lambda,k) \xrightarrow{w} \frac{1}{k}\sum_{i=1}^{k/2}\left[\begin{array}{c} \frac{\left[\left\{B_{g,1}^{(2i)}(\lambda_{2i})\right\}^2 - \left\{B_{g,1}^{(2i)}(\lambda_{2i-1})\right\}^2 - \widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i})^2 - \widetilde{g}(\lambda_{2i-1})^2\right\}\right]^2}{4\int_{\lambda_{2i-1}}^{\lambda_{2i}}\left[B_{g,1}^{(2i)}(r)\right]^2 dr} \\ +\frac{1}{\lambda_{2i}-\lambda_{2i-1}}\left\{B_{g,1}(\lambda_{2i}) - B_{g,1}(\lambda_{2i-1})\right\}^2 \end{array}\right]$$

$$F_{1b}(\lambda,k) \xrightarrow{w} \frac{1}{(k+2)}\sum_{i=0}^{k/2}\left[\begin{array}{c} \frac{\left[\left\{B_{g,1}^{(2i+1)}(\lambda_{2i+1})\right\}^2 - \left\{B_{g,1}^{(2i+1)}(\lambda_{2i})\right\}^2 - \widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i+1})^2 - \widetilde{g}(\lambda_{2i})^2\right\}\right]^2}{4\int_{\lambda_{2i}}^{\lambda_{2i+1}}\left[B_{g,1}^{(2i+1)}(r)\right]^2 dr} \\ +\frac{1}{\lambda_{2i+1}-\lambda_{2i}}\left\{B_{g,1}(\lambda_{2i+1}) - B_{g,1}(\lambda_{2i})\right\}^2 \end{array}\right]$$

If k is odd,

$$F_{1a}(\lambda,k) \xrightarrow{w} \frac{1}{(k+1)}\sum_{i=1}^{(k+1)/2}\left[\begin{array}{c} \frac{\left[\left\{B_{g,1}^{(2i)}(\lambda_{2i})\right\}^2 - \left\{B_{g,1}^{(2i)}(\lambda_{2i-1})\right\}^2 - \widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i})^2 - \widetilde{g}(\lambda_{2i-1})^2\right\}\right]^2}{4\int_{\lambda_{2i-1}}^{\lambda_{2i}}\left[B_{g,1}^{(2i)}(r)\right]^2 dr} \\ +\frac{1}{\lambda_{2i}-\lambda_{2i-1}}\left\{B_{g,1}(\lambda_{2i}) - B_{g,1}(\lambda_{2i-1})\right\}^2 \end{array}\right]$$

$$F_{1b}(\lambda,k) \xrightarrow{w} \frac{1}{(k+1)}\sum_{i=0}^{(k-1)/2}\left[\begin{array}{c} \frac{\left[\left\{B_{g,1}^{(2i+1)}(\lambda_{2i+1})\right\}^2 - \left\{B_{g,1}^{(2i+1)}(\lambda_{2i})\right\}^2 - \widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i+1})^2 - \widetilde{g}(\lambda_{2i})^2\right\}\right]^2}{4\int_{\lambda_{2i}}^{\lambda_{2i+1}}\left[B_{g,1}^{(2i+1)}(r)\right]^2 dr} \\ +\frac{1}{\lambda_{2i+1}-\lambda_{2i}}\left\{B_{g,1}(\lambda_{2i+1}) - B_{g,1}(\lambda_{2i})\right\}^2 \end{array}\right]$$

Corollary 1 *Denote the limits of $F_{1a}(\lambda,k)$ and $F_{1b}(\lambda,k)$ by $F_{1a}^0(\lambda,k)$ and $F_{1b}^0(\lambda,k)$, respectively. Then, by the continuous mapping theorem, we have $F_{1a}(k) \xrightarrow{w} \sup_{\lambda\in\Lambda_\epsilon^k} F_{1a}^0(\lambda,k)$, $F_{1b}(k) \xrightarrow{w} \sup_{\lambda\in\Lambda_\epsilon^k} F_{1b}^0(\lambda,k)$, $W_1(k) \xrightarrow{w} \max[F_{1a}^0(k), F_{1b}^0(k)]$, $Wmax_1 = \max_{1\leq k\leq A} W_1(k) \xrightarrow{w} \max_{1\leq k\leq A}\{\max[F_{1a}^0(k), F_{1b}^0(k)]\}$.*

Remark 2 *The process $B_{g,1}(s)$ is Gaussian with zero mean and variance $\nu(s) = \widetilde{g}(s)^2/\widetilde{g}(1)^2$ so that $B_{g,1}(.)$ is a variance-transformed Brownian motion with directing process ν. See Davidson (1994, Section 29.4) and Cavaliere (2005) for discussion of transformed Brownian motion.*

Remark 3 *The limiting distributions stated in Theorem 1 are not pivotal and depend on the sample path of the volatility process* $g(.)$*. The pivotal limit results in KPZ can be recovered only if* $g(.) = \sigma$ *in which case* $B_{g,1}(.)$ *reduces to the standard Brownian motion* $B_1(.)$*. The practical implication of this result is that the KPZ tests in general have incorrect asymptotic size in the presence of nonstationary volatility.*

Next, we state the large sample result for the BP tests under $H_0^{(0)}$:

Theorem 2 *Suppose Assumptions A1-A4 hold. Then, under* $H_0^{(0)}$*, we have*

$$
\begin{aligned}
& G_1(\lambda, k) \xrightarrow{w} \frac{1}{k}\sum_{i=1}^{k}\left[\frac{\{\lambda_i B_{g,1}(\lambda_{i+1}) - \lambda_{i+1}B_{g,1}(\lambda_i)\}^2}{\lambda_i\lambda_{i+1}(\lambda_{i+1}-\lambda_i)} + \frac{\{\widetilde{g}(\lambda_i)^2 B_{g,2}(\lambda_{i+1}) - \widetilde{g}(\lambda_{i+1})^2 B_{g,2}(\lambda_i)\}^2}{\widetilde{g}(\lambda_i)^2\widetilde{g}(\lambda_{i+1})^2\{\widetilde{g}(\lambda_{i+1})^2 - \widetilde{g}(\lambda_i)^2\}}\right] \\
\equiv \; & G_1^0(\lambda, k) \\
& G_1(k) \xrightarrow{w} \sup_{\lambda\in\Lambda_\epsilon^k} G_1^0(\lambda, k) \\
& UDmax_1 \xrightarrow{w} \max_{1\le k\le A}\left\{\sup_{\lambda\in\Lambda_\epsilon^k} G_1^0(\lambda, k)\right\}
\end{aligned}
$$

Remark 4 *The limit distributions of the BP tests under the stable* $I(0)$ *null hypothesis are not pivotal under nonstationary volatility and depend on the volatility process* $g(.)$*. Only when* $g(.) = \sigma$ *can we recover the pivotal limit distribution obtained by BP (stated in their Proposition 6). Therefore, using the critical values tabulated in BP will lead to incorrectly sized tests in general when nonstationary volatility is present. The non-robustness of the Andrews (1993) and Andrews and Ploberger (1994) tests for a single break to shifts in the marginal distribution of the regressors was earlier demonstrated by Hansen (2000).*

Remark 5 *The absence of large sample invariance of the KPZ and BP tests to unconditional heteroskedasticity continues to hold for the heteroskedasticity-robust versions of these tests even though the limits differ from those obtained in Theorems 1 and 2 [See Remark 12 in Georgiev et al. (2018)]. Monte Carlo simulations (unreported) did not reveal any particular improvements of the robust versions over their non-robust counterparts in the presence of unconditional heteroskedasticity so our analysis focuses on the latter set of tests which are simpler to implement in practice.*

5 The Wild Bootstrap

The large sample analysis in the preceding section shows that the testing procedures of KPZ and BP are not robust to the presence of nonstationary volatility. In order to solve the identified inference problem, this section proposes wild bootstrap versions of these tests and establishes their asymptotic validity under Assumption A1-A4 thereby permitting their application to a general class of heteroskedastic data generating processes. Unlike the standard residual bootstrap, the wild bootstrap procedure (Liu, 1988) can mimic the pattern of heteroskedasticity in the errors and therefore replicate the first order limit distributions of the test statistics derived in Section 4. We further show that the bootstrap KPZ and BP test statistics remain consistent under the relevant alternatives so that the procedures can be reliably employed in practice to detect instability in persistence. Since the direction of persistence change is typically unknown, our subsequent analysis will only consider the $W_1(.)$, $Wmax_1$, $G_1(.)$, $UDmax_1$ tests which are also the statistics we recommend for use in practice. Section 5.1 presents the bootstrap algorithm, Section 5.2 establishes the asymptotic validity of the procedure under the null hypothesis of no structural change in persistence, Section 5.3 demonstrates the consistency of the proposed tests, and Section 5.4 proposes and justifies a sequential procedure for estimating the number of breaks.

5.1 The Bootstrap Algorithm

Since the null hypothesis H_0 accommodates both $I(1)$ and $I(0)$ processes, the algorithm is based on generating two kinds of bootstrap samples, one of which is $I(1)$ while the other is $I(0)$, conditional on the data $\{y_t\}_{t=1}^T$. The $I(1)$ bootstrap samples will be used to approximate the finite sample distribution of the KPZ statistics while the $I(0)$ bootstrap samples will be used to approximate the distribution of the BP statistics. Note that our approach is in contrast to CT's bootstrap scheme in which the bootstrap samples are generated only under the $I(0)$ null given that they do not consider the $I(1)$ case. Further, for reasons discussed below, our proposed bootstrap scheme is not recursive as that employed in Xu (2008) for the stationary autoregressive model. Denote by $\{v_t;\ t = 1, ..., T\}$ a sequence of i.i.d. random variables with zero mean and unit variance that are independent of $\{y_t\}_{t=1}^T$. We now enumerate the steps involved in generating the two types of bootstrap samples. We start with the $I(1)$ case.

(A) I(1) Bootstrap Samples

1. Estimate the regression

$$\Delta y_t = \sum_{j=1}^{l_T} \pi_j \Delta y_{t-j} + e_t^*; \quad t = l_T + 2, ..., T \tag{8}$$

   where l_T is chosen by BIC based on (8). The estimates are denoted $(\breve{l}_T, \breve{\pi}_1, ..., \breve{\pi}_{\breve{l}_T})$. Obtain the residuals $\{\breve{e}_t\}$ as

$$\breve{e}_t = \Delta y_t - \sum_{j=1}^{\breve{l}_T} \breve{\pi}_j \Delta y_{t-j}; \quad t = \breve{l}_T + 2, ..., T$$

2. Obtain the bootstrap residuals $\{e_t^{(1)}\}$ as

$$e_t^{(1)} = \breve{e}_t v_t, \quad t = \breve{l}_T + 2, ..., T$$

3. Generate the bootstrap sample $\{y_t^{(1)}\}$ as

$$\begin{aligned} y_t^{(1)} &= y_{t-1}^{(1)} + e_t^{(1)}; \quad t = \breve{l}_T + 2, ..., T \\ y_t^{(1)} &= y_t; \; t = 1, ..., \breve{l}_T + 1 \end{aligned} \tag{9}$$

4. Construct the bootstrap versions of the $W_1(.)$ and $Wmax_1$ statistics using $\{y_t^{(1)}\}_{t=1}^T$ based on a specification that *does not* include any lagged first differences of $y_t^{(1)}$.

5. Repeat steps (2)-(4) B times to approximate the bootstrap distribution of the statistics in step (4).

(B) I(0) Bootstrap Samples

1. Estimate the regression

$$y_t = c + \alpha y_{t-1} + \sum_{j=1}^{l_T} \pi_j \Delta y_{t-j} + e_t^*; \quad t = l_T + 2, ..., T \tag{10}$$

   where l_T is chosen by BIC based on (10). The estimates are denoted $(\widetilde{c}, \widetilde{\alpha}, \widetilde{l}_T, \widetilde{\pi}_1, ..., \widetilde{\pi}_{\widetilde{l}_T})$. Obtain the residuals $\{\widetilde{e}_t\}$ as

$$\widetilde{e}_t = y_t - \widetilde{c} - \widetilde{\alpha} y_{t-1} - \sum_{j=1}^{\widetilde{l}_T} \widetilde{\pi}_j \Delta y_{t-j}; \quad t = \widetilde{l}_T + 2, ..., T$$

2. Obtain the bootstrap residuals $\{e_t^{(0)}\}$ as

$$e_t^{(0)} = \widetilde{e}_t v_t, \quad t = \widetilde{l}_T + 2, ..., T$$

3. Generate the bootstrap sample $\{y_t^{(0)}\}$ as

$$\begin{aligned} y_t^{(0)} &= e_t^{(0)}, \quad t = \widetilde{l}_T + 2, ..., T \\ y_t^{(0)} &= 0, \quad t = 1, ..., \widetilde{l}_T + 1 \end{aligned} \tag{11}$$

4. Construct the bootstrap versions of the $G_1(.)$ and $UDmax_1$ statistics using $\{y_t^{(0)}\}_{t=1}^T$ based on a specification that *does not* include any lagged first differences of $y_t^{(0)}$.

5. Repeat steps (2)-(4) B times to approximate the bootstrap distribution of the statistics in step (4).

For the $I(1)$ scheme (A), we do not introduce short-run dynamics in the bootstrap DGP (9) through the inclusion of lagged first differences, i.e., we do not "recolor" the innovations using the estimates $\{\breve{\pi}_j\}$ in the terminology of Cavaliere and Taylor (2009). The reason is that the estimated lag polynomial may fail to satisfy Assumption A2 so that recoloring the innovations leaves open the possibility that the bootstrap sample is generated from a data generating process with explosive roots or more than one unit root. Indeed, Cavaliere and Taylor (2009) note this possibility in their Monte Carlo simulations when testing the null hypothesis of a unit root.[3] Our proposed scheme (A) rules out explosive or multiple unit roots in the bootstrap DGP by generating the bootstrap sample $\{y_t^{(1)}\}$ as partial sums of the bootstrap residuals which are serially independent, conditional on the data. Such a *partially recursive* scheme was also employed by Cavaliere and Taylor (2008a) in devising bootstrap unit root tests.

The $I(0)$ bootstrap scheme (B) is *non-recursive* since we do not "add back" the conditional mean component based on the parameter estimates. Rather, our bootstrap data $\{y_t^{(0)}\}$ have constant (zero) mean and are serially independent, conditional on the data. Using a recursive scheme leads to tests with lower power relative to the non-recursive scheme when the original data contain an $I(1)$ segment. The reason is that the estimated full sample persistence parameter converges to unity at rate T so that the recursive bootstrap samples are effectively drawn from an autoregressive process with a root close to unity. This feature

[3] We observed this feature in our Monte Carlo experiments as well.

contributes to an increase in the bootstrap critical values (relative to the non-recursive bootstrap) which in turn has an adverse effect on power. The issue was illustrated by Gulesserian and Kejriwal (2014) in the context of stationarity testing based on the sieve bootstrap in the homoskedastic case. Monte Carlo simulations in the present context suggest notable power gains from using the non-recursive form of the wild bootstrap for alternatives that involve switches between $I(1)$ and $I(0)$ regimes (see Section 7).

Remark 6 *Step 4 in both (A) and (B) constructs the bootstrap statistics from an AR(1) specification instead of one that includes the lags of the first differences. The reason is that the bootstrap residuals* $\{e_t^{(1)}\}$ *in scheme (A) and the bootstrap residuals* $\{y_t^{(0)}\}$ *in scheme (B) are serially independent, conditional on the data. It is therefore not necessary to control for serial correlation through the lagged differences in the bootstrap regression. Indeed, our Monte Carlo simulations (unreported) confirmed that using an AR(1) bootstrap specification resulted in improved finite sample properties (size and power) of the testing procedures.*

Denote the bootstrap analogues of $W_1(k)$, $Wmax_1$, $G_1(k)$ and $UDmax_1$ statistics by $W_1^*(k)$, $Wmax_1^*$, $G_1^*(k)$ and $UDmax_1^*$ respectively. The associated p-values are denoted $p^*_{k,W_1}, p^*_{Wmax}, p^*_{k,G_1}$ and p^*_{UDmax}, respectively.[4] For instance, $p_{1,W_1} = 1 - D^*_{1,1}(W_1(1))$, where $D^*_{1,1}(.)$ denotes the cumulative distribution function of $W_1^*(1)$ and $W_1(1)$ denotes the value of the statistic computed using the original data $\{y_t\}_{t=1}^T$.[5] Similarly, for a given significance level η, denote the bootstrap critical values of $W_1(k)$, $Wmax_1$, $G_1(k)$ and $UDmax_1$ by $cv^*_{w,k}(\eta)$, $cv^*_{w,max}(\eta)$, $cv^*_{g,k}(\eta)$ and $cv^*_{g,max}(\eta)$, respectively. Finally, define our proposed statistics $H^*(k,\eta) = \min\left[W_1(k), \frac{cv^*_{w,k}(\eta)}{cv^*_{g,k}(\eta)}G_1(k)\right]$ and $Hmax_1^*(\eta\ ) = \min\left[Wmax_1, \frac{cv^*_{w,max}(\eta)}{cv^*_{g,max}(\eta)}UDmax_1\right]$. Henceforth, we will refer to $H^*(.,.)$ and $Hmax_1^*(.)$ as the "hybrid" tests.

5.2 Asymptotic Size

The following two theorems establish that the wild bootstrap can successfully replicate the first order asymptotic distribution of the test statistics. Let $U[0,1]$ denote a uniform distribution over $[0,1]$.

Theorem 3 *Under the conditions of Theorem 1, (i)* $W_1^*(k) \xrightarrow{w}_p \max[F^0_{1a}(k), F^0_{1b}(k)]$, $Wmax_1^* \xrightarrow{w}_p \max_{1\leq k\leq A}\ \{\max[F^0_{1a}(k), F^0_{1b}(k)]\}$; *(ii)* $p^*_{k,W_1} \xrightarrow{w} U[0,1]$, $p^*_{Wmax} \xrightarrow{w} U[0,1]$.

[4] The dependence of the p-values on the sample size T is suppressed for economy of notation.

[5] The bootstrap distribution is unknown in practice and is approximated by B Monte Carlo replications. As shown in Hansen (1996), the approximate p-value converges to the true p-value as $B \to \infty$.

Theorem 4 *Under the conditions of Theorem 2, (i)* $G_1^*(k) \xrightarrow{w}_p \sup_{\lambda \in \Lambda_\epsilon^k} G_1^0(\lambda, k)$, $UDmax_1^*$ $\xrightarrow{w}_p \max_{1 \leq k \leq A} \left\{ \sup_{\lambda \in \Lambda_\epsilon^k} G_1^0(\lambda, k) \right\}$; *(ii)* $p_{k,G_1}^* \xrightarrow{w} U[0,1]$, $p_{UDmax}^* \xrightarrow{w} U[0,1]$.

A consequence of Theorems 1 and 2 is the following corollary which states that, for a given significance level η, the statistics $H_1^*(k,\eta)$ and $Hmax_1^*(\eta)$ have asymptotic size at most η.

Corollary 2 *Suppose Assumptions A1-A4 hold. Then, under* H_0, $\lim_{T\to\infty} P(H^*(k,\eta) > cv_{w,k}^*(\eta)) \leq \eta$ *and* $P(Hmax_1^*(\eta) > cv_{w,max}^*(\eta)) \leq \eta$.

5.3 Consistency

The following result states the consistency of the test statistics when the time series is subject to $k = m$ structural changes in persistence:

Theorem 5 *Suppose Assumptions A1-A4 hold and* $\lambda^0 \in \Lambda_\epsilon^m$. *Then, under each of* $H_{a,m}^{(1)}$, $H_{b,m}^{(1)}$ *and* $H_{1,m}^{(0)}$, *we have* $p_{m,W_1}^* \xrightarrow{p} 0$, $p_{Wmax}^* \xrightarrow{p} 0$, $p_{m,G_1}^* \xrightarrow{p} 0$, $p_{UDmax}^* \xrightarrow{p} 0$.

5.4 Estimating the Number of Breaks

Based on Corollary 2 and Theorem 5, a bootstrap procedure can be devised to estimate the number of breaks based on a sequential test of the null hypothesis of l breaks against the alternative of $l+1$ breaks. Kejriwal (2018) proposes a sequential approach based on asymptotic critical values assuming conditional homoskedasticity. Dealing with heteroskedasticity entails not only the use of bootstrap critical values but also renders invalid employing the full sample critical values when testing stability in each of the $(l+1)$ segments. The latter feature is due to the fact that unlike the homoskedastic case, the time-varying nature of the volatility process leads to different limit distributions of the test statistics across the different segments. We propose a new bootstrap sequential procedure that remains valid in the heteroskedastic case and involves a non-trivial modification of the corresponding procedure that assumes homoskedasticity.

We first develop a sequential test of the null hypothesis of $l(\geq 1)$ breaks against the alternative of $(l+1)$ breaks. To this end, we partition the sample into $(l+1)$ segments using the l estimated break dates $(\hat{T}_1, ..., \hat{T}_l)$ obtained by minimizing the *unrestricted* sum of squared residuals, i.e., the intercept, slope (the persistence parameter) and the coefficients of the lagged first differences are all allowed to be regime-specific. The one break KPZ and BP statistics are then computed using data from each of the estimated $(l+1)$ regimes. These statistics are denoted $W_1^{(i)}(1)$ and $G_1^{(i)}(1)$, respectively, for $i = 1, ..., l + 1$. The parameter estimates in each of the $(l + 1)$ estimated regimes are used to generate the regime-specific $I(1)$ and $I(0)$ bootstrap samples based on schemes A and B respectively, as detailed in Section 5.1. These samples are used to compute the bootstrap p-values of the statistics $W_1^{(i)}(1)$ and $G_1^{(i)}(1)$, denoted by $p_{1,W_1}^{*,(i)}$ and $p_{1,G_1}^{*,(i)}$, respectively ($i = 1, ..., l + 1$). For a given significance level η, we reject the null of l breaks in favor of $(l + 1)$ breaks if

$$\min_{1 \leq i \leq l+1} \{p_i^*\} < \eta_{l+1} \tag{12}$$

where $p_i^* = \max\{p_{1,W_1}^{*,(i)}, p_{1,G_1}^{*,(i)}\}$ and $\eta_{l+1} = 1 - (1 - \eta)^{1/(l+1)}$. As shown in Appendix A, the decision rule (12) has asymptotic size at most η under the null hypothesis of l breaks.

The various steps associated with the implementation of the sequential procedure can now be enumerated as follows:

1. Test the null of no break (H_0) against the alternative of at least one break. For a given significance level η, we reject H_0 if $p_{max}^* = \max\{\, p_{Wmax}^*, p_{UDmax}^*\} < \eta$ and conclude in favor of at least one break; otherwise the procedure stops and the number of breaks is estimated to be zero.

2. Upon a rejection in step 1, use the decision rule (12) with $l = 1$ to determine if there is more than one break. This process is repeated by increasing l sequentially until the test fails to reject the null hypothesis of no additional structural breaks.

3. The estimate $\hat{m}$ is obtained as the total number of rejections obtained from steps 1 and 2.

The following result ensures that the probability of selecting the true number of breaks employing the above sequential approach is at least $(1 - \eta)$ in large samples:

Theorem 6 *Under the conditions of Theorem 5,* $\lim_{T\to\infty} P(\hat{m} = m) \geq 1 - \eta$.

6 Extensions

This section discusses extensions to the bootstrap procedures advocated in the preceding section to deal with the following issues: (1) the presence of deterministic trends; (2) distinguishing between a process driven by pure trend shifts from one that is accompanied by shifts in persistence. We consider each of these issues in turn.

6.1 Deterministic Trends

In order to deal with the potential presence of deterministic trends, we consider an extension of (1) that includes the possibility of m breaks in the deterministic trend:

$$\left.\begin{array}{c} y_t = \mu_0 + \beta_0 t + \sum_{j=1}^{m} \mu_j DU_{jt} + \sum_{j=1}^{m} \beta_j DT_{jt} + u_t \\ u_t = u_{T_{i-1}^0} + h_t \\ h_t = \alpha_i h_{t-1} + \sum_{j=1}^{p-1} \pi_{ij} \Delta h_{t-j} + e_t \\ h_{T_{i-1}^0} = ... = h_{T_{i-1}^0 - p+1} = 0 \end{array}\right\} \begin{array}{c} t = T_{i-1}^0 + 1, T_{i-1}^0 + 2, ..., T_i^0; \\ i = 1, ..., m+1 \end{array} \tag{13}$$

where $DU_{jt} = I(t > T_j^0),\ DT_{jt} = I(t > T_j^0)(t - T_j^0);\ j = 1, ..., m$. The data generating process (13) can be expressed as

$$y_t = c_i + b_i t + \alpha_i y_{t-1} + \sum_{j=1}^{p-1} \pi_{ij} \Delta y_{t-j} + e_t \tag{14}$$

with

$$\begin{aligned} c_i &= (1-\alpha_i)\left\{\mu_0 + \sum_{j=1}^{i-1}(\mu_j - \beta_j T_j^0) + u_{T_{i-1}^0}\right\} + \left(\alpha_i - \sum_{j=1}^{p-1} \pi_{ij}\right)\left\{\beta_0 + \sum_{j=1}^{i-1} \beta_j\right\} \\ b_i &= (1-\alpha_i)(\beta_0 + \sum_{j=1}^{i-1} \beta_j) \end{aligned} \tag{15}$$

The estimating regression therefore takes the form

$$\Delta y_t = c_i + b_i t + (\alpha_i - 1) y_{t-1} + \sum_{j=1}^{p-1} \pi_j \Delta y_{t-j} + e_t^*$$

with e_t^* denoting the regression error. KPZ propose tests of the null hypothesis $\widetilde{H}_0^{(1)}$: $c_i = c,\ \alpha_i = 1$ for all i in (14). Note that under $\widetilde{H}_0^{(1)},\ b_i = 0$ for all i so that the process follows a stable unit root process with possible drift. As in the non-trending case, KPZ consider two models under the alternative hypothesis depending on whether the initial regime is trend or difference stationary. In accordance with the notation in Section 3, the test statistics in the trending case are denoted by $F_{2a}(\lambda, k),\ F_{2b}(\lambda, k),\ W_2(k)$ and $Wmax_2$. KPZ show that the limit distributions of the statistics under $\widetilde{H}_0^{(1)}$ are pivotal under error homoskedasticity but different from those in the non-trending case.

We now turn to testing the null of a stable trend stationary process, i.e., $\widetilde{H}_0^{(0)}$: $c_i =$

$c,\ b_i = b,\ \alpha_i = \alpha$ for all i where $|\alpha| < 1$ in the model

$$y_t = c_i + b_i t + \alpha_i y_{t-1} + + \sum_{j=1}^{p-1} \pi_{ij} \Delta y_{t-j} + e_t \tag{16}$$

with $c_i = (1-\alpha_i)\left\{\mu_0 + \sum_{j=1}^{i-1}(\mu_j - \beta_j T_j^0)\right\} + \alpha_i \left\{\beta_0 + \sum_{j=1}^{i-1}\beta_j\right\}$ and b_i defined as in (15). The estimating regression takes the form

$$\Delta y_t = c_i + b_i t + (\alpha_i - 1) y_{t-1} + + \sum_{j=1}^{p-1} \pi_j \Delta y_{t-j} + e_t^* \tag{17}$$

The test statistic for a fixed number $m = k$ changes is based on

$$G_2(\lambda, k) = [T - 3(k+1)](\widetilde{SSR}_0^{(0)} - SSR_{2,k}^{(0)})/[kSSR_{2,k}^{(0)}] \tag{18}$$

In (18), $\widetilde{SSR}_0^{(0)}$ denotes the sum of squared residuals under $\widetilde{H}_0^{(0)}$, i.e., that obtained from OLS estimation of (17) subject to the restrictions $c_i = c,\ b_i = b,\ \alpha_i = \alpha$ for all i. The quantity $SSR_{2,k}^{(0)}$ denotes the sum of squared residuals obtained from *unrestricted* OLS estimation of (17). The test statistic is then defined as $G_2(k) = \sup_{\lambda \in \Lambda_\epsilon^k} G_2(\lambda, k)$. When the number of breaks is unknown, the relevant test statistic is $UDmax_2 = \max_{1 \le k \le A} G_2(k)$. The limit distributions of $G_2(.)$ and $UDmax_2$ under error homoskedasticity are derived in Kejriwal (2018).

Under Assumptions A1-A4, the above test statistics are not asymptotically pivotal and depend on the sample path of $\{\sigma_t\}$ as demonstrated in Section 3 for the non-trending case. We propose the following bootstrap algorithm that enables asymptotically valid inference in the trending case. As in Section 5, we generate both $I(1)$ and $I(0)$ bootstrap samples to ensure that the procedure has correct asymptotic size under $\widetilde{H}_0$, where $\widetilde{H}_0 = \widetilde{H}_0^{(1)} \cup \widetilde{H}_0^{(0)}$.

(A′) I(1) Bootstrap Samples

1. Estimate the regression

$$\Delta y_t = c + \sum_{j=1}^{l_T} \pi_j \Delta y_{t-j} + e_t^*; \quad t = l_T + 2, ..., T \tag{19}$$

where l_T is chosen by BIC based on (19). The estimates are denoted $(\breve{c}, \breve{l}_T, \breve{\pi}_1, ..., \breve{\pi}_{\breve{l}_T})$.

Obtain the residuals $\{\breve{e}_t\}$ as

$$\breve{e}_t = \Delta y_t - \breve{c} - \sum_{j=1}^{\breve{l}_T} \breve{\pi}_j \Delta y_{t-j}; \quad t = \breve{l}_T + 2, ..., T$$

2. Obtain the bootstrap residuals $\{e_t^{(1)}\}$ as

$$e_t^{(1)} = \breve{e}_t v_t; \quad t = \breve{l}_T + 2, ..., T$$

3. Generate the bootstrap sample $\{y_t^{(1)}\}$ as

$$\begin{aligned} y_t^{(1)} &= y_{t-1}^{(1)} + e_t^{(1)}; \quad t = \breve{l}_T + 2, ..., T \\ y_t^{(1)} &= y_t; \ t = 1, ..., \breve{l}_T + 1 \end{aligned} \tag{20}$$

4. Construct the bootstrap versions of the $W_2(.)$ and $Wmax_2$ statistics using $\{y_t^{(1)}\}_{t=1}^T$ based on a specification that *does not* include any lagged first differences of $y_t^{(1)}$.

5. Repeat steps (2)-(4) B times to approximate the bootstrap distribution of the statistics in step (4).

(B′) I(0) Bootstrap Samples

1. Estimate the regression

$$y_t = c + bt + \alpha y_{t-1} + \sum_{j=1}^{l_T} \pi_j \Delta y_{t-j} + e_t^*; \quad t = l_T + 2, ..., T \tag{21}$$

where l_T is chosen by BIC based on (21). The estimates are denoted $(\widetilde{c}, \widetilde{b}, \widetilde{\alpha}, \widetilde{l}_T, \widetilde{\pi}_1, ..., \widetilde{\pi}_{\widetilde{l}_T})$. Obtain the residuals $\{\widetilde{e}_t\}$ as

$$\widetilde{e}_t = y_t - \widetilde{c} - \widetilde{b}t - \widetilde{\alpha} y_{t-1} - \sum_{j=1}^{\widetilde{l}_T} \widetilde{\pi}_j \Delta y_{t-j}; \quad t = \widetilde{l}_T + 2, ..., T$$

2. Obtain the bootstrap residuals $\{e_t^{(0)}\}$ as

$$e_t^{(0)} = \widetilde{e}_t v_t; \quad t = \widetilde{l}_T + 2, ..., T$$

3. Generate the bootstrap sample $\{y_t^{(0)}\}$ as

$$\begin{aligned} y_t^{(0)} &= e_t^{(0)}; \;\; t = \widetilde{l}_T + 2, ..., T \\ y_t^{(0)} &= 0; \; t = 1, ..., \widetilde{l}_T + 1 \end{aligned} \tag{22}$$

4. Construct the bootstrap versions of the $G_2(.)$ and $UDmax_2$ statistics using $\{y_t^{(0)}\}_{t=1}^T$ based on a specification that *does not* include any lagged first differences of $y_t^{(0)}$.

5. Repeat steps (2)-(4) B times to approximate the bootstrap distribution of the statistics in step (4).

The difference between the schemes (A) and (A′) is that the residuals imposing the $I(1)$ null hypothesis in step 1 of the latter are obtained using a regression that includes a constant to account for the possible drift in the process. Similarly, the residuals in scheme (B′) are based on a regression that includes a deterministic trend. Note, however, that the bootstrap samples in both (A′) and (B′) are generated in a manner similar to those for (A) and (B), respectively. In particular, it is not necessary to include the estimated drift in step 3 of (A′) when constructing $\{y_t^{(1)}\}$ or the estimated deterministic trend in step 3 of (B′) when constructing $\{y_t^{(0)}\}$ given that the test statistics $W_2(.)$ and $G_2(.)$ are invariant to their presence in the data generating process.

The bootstrap analogues of $W_2(k)$, $Wmax_2$, $G_2(k)$ and $UDmax_2$ and the associated p-values are obtained as described in Section 5.1. The sequential procedure outlined in Section 5.4 is accordingly modified so that the bootstrap critical values are now obtained from schemes (A′) and (B′). The following result states the large sample validity of the proposed procedure in the potential presence of deterministic trends:

Theorem 7 *Suppose Assumptions A1-A4 hold. Under bootstrap schemes (A′) and (B′), (i) Corollary 2 holds with H_0 replaced by $\widetilde{H}_0$, (ii) Theorems 5 and 6 hold.*

6.2 Disentangling Trend and Persistence Shifts

An important feature of the test statistics developed in Sections 5.1 and 6.1 is that they are tests of the null hypothesis that the trend function parameters and the persistence parameter are *jointly* stable. Consequently, the tests have power against processes that are driven by pure trend shifts with no accompanying change in persistence. To distinguish between trend and persistence shifts, we can adapt the three-step approach recommended in Kejriwal (2018) for the homoskedastic case to the present context.

Consider first the non-trending case. The first step entails determining the number of breaks ($\widetilde{m}$) using the sequential procedure described in Section 5.4 and the associated breakpoint estimates ($\hat{T}_1, \hat{T}_2, ..., \hat{T}_{\hat{m}}$) obtained from the *unrestricted* model that allows all parameters including the coefficients of the lagged first differences to change across regimes. Second, using the estimated breakpoints, the Wald statistic for testing the null hypothesis of stable $I(0)$ persistence is constructed [i.e., constancy of α_i over all i in (4)] while allowing all other parameters to vary across the $(\hat{m}+1)$ regimes. To account for nonstationary volatility, the Wald statistic should be computed using a heteroskedasticity-robust estimator of the variance-covariance matrix.[6] Third, the null hypothesis of stable $I(0)$ persistence is rejected if the computed Wald statistic is significant at the specified level where the critical value is obtained from the $\chi^2(\hat{m})$ distribution. Otherwise, the null is not rejected and we conclude in favor of a model with pure level shifts.

The analysis in the trending case is complicated by the fact that the process can be either $I(1)$ (with a possibly time-varying drift) or $I(0)$ (around a broken deterministic trend). As in Section 6.1, we can develop tests separately for the $I(1)$ and $I(0)$ null hypotheses and choose as the critical region the intersection of the critical regions of the two tests. The three-step approach can be implemented as follows. In the first step, the estimate of the number of breaks (say $\breve{m}$) is obtained applying the modified sequential procedure suggested in Section 6.1. The associated breakpoints are now computed from the unrestricted specification (16). Second, we compute the Wald statistic (against using heteroskedasticity robust standard errors) for testing the null hypothesis of constant persistence allowing the trend function parameters and the coefficients of the lagged differences to change at the estimated breakpoints. In the $I(0)$ case, the statistic has a limiting $\chi^2(m)$ distribution so that standard critical values (with $\breve{m}$ degrees of freedom) can be used.[7] In the $I(1)$ case, the statistic has a non-pivotal limit even in the homoskedastic case depending in particular on the vector of break fractions λ^0 (Kejriwal, 2018). To achieve asymptotically valid inference, we can apply a second wild bootstrap scheme based on residuals estimated under the version of the model (19) that allows the level to change across regimes at the estimated breakpoints. The $I(1)$ bootstrap samples can then be obtained from a data generating process that now

[6] Phillips and Xu (2006) show that for an $AR(p)$ model, the Eicker-White standard errors are asymptotically valid in the presence of nonstationary volatility. Alternatively, a wild bootstrap test could be employed where the bootstrap samples are drawn using a non-recursive scheme similar to scheme (B) but one that accommodates the estimated level shifts in the bootstrap DGP.

[7] Alternatively, a wild bootstrap test based on a scheme similar to (B′) could be used where the bootstrap DGP is constructed allowing the trend function to change at the estimated breakpoints using the regime-specific trend function estimates.

includes the estimated regime-specific drift in (20) [in contrast to scheme (A′)]. The bootstrap distribution of the Wald statistic can then be approximated using Monte Carlo simulation from which the relevant critical values can be obtained. Finally, the null hypothesis of stable [$I(1)$ or $I(0)$] persistence is rejected if the $I(0)$ and $I(1)$ Wald statistics are both significant at the specified level using their respective critical values.

7 Monte Carlo Evidence

This section presents Monte Carlo evidence to assess the finite sample performance of the recommended procedures as well as provide a comparison with existing approaches. We consider a variety of data generating processes (DGPs) characterized by nonstationary volatility where the specifications for the volatility process are borrowed from the Monte Carlo design of CT to enable comparison with their tests. The three volatility specifications are as follows:

1. Model 1 [Single Volatility Break]: $\sigma_t = \sigma_0^* + (\sigma_1^* - \sigma_0^*) I(t \geq 0.5T)$
2. Model 2 [Trending Volatility]: $\sigma_t = \sigma_0^* + (\sigma_1^* - \sigma_0^*) \left(\frac{t-1}{T-1}\right)$
3. Model 3 [Near-Integrated Stochastic Volatility]: $\sigma_t = \sigma_0^* \exp(0.5 \upsilon b_t / \sqrt{T})$, $b_t = (1 - c/T) b_{t-1} + k_t$, $k_t \sim i.i.d.\ N(0,1)$, $b_0 = 0$.

Following CT, we set $\sigma_0^* = 1$ in all cases, $\delta := \sigma_0^*/\sigma_1^* \in \{1, 1/3, 3\}$ for Models 1 and 2, $\upsilon = 5$ and $c \in \{0, 10\}$ for Model 3. Next, a sequence $\{z_t\}_{t=1}^T$ is generated by the ARMA(1,1) process

$$
\begin{aligned}
z_t &= \rho z_{t-1} + e_t - \theta e_{t-1},\ z_0 = 0 \\
e_t &= \sigma_t \varepsilon_t \\
\varepsilon_t &\sim i.i.d.\ N(0,1)
\end{aligned}
$$

The level of trimming is set at $\epsilon = 15\%$ and 1000 Monte Carlo replications are used in all experiments. The sample size $T \in \{200, 400\}$ in all experiments except those in section 7.4 where $T \in \{400, 600\}$. We report results for the non-trending case only given that the results for the trending case are qualitatively similar.

The lag length in the KPZ and BP procedures was selected using BIC with the maximum allowable number of lags set to five.[8] We report the performance of the tests $H^*(k, \eta)$; $k =$

[8] We also experimented with larger values but found that they yielded comparable size but lower power, especially for the multiple break and sequential tests. The BIC was computed under the null model for each

$1, 2$ and $Hmax_1^*(\eta) = \max\{H^*(1,\eta), H^*(2,\eta)\}$. For brevity, we suppress the dependence on η henceforth. As a benchmark for comparison, we include CT's single break bootstrap tests of the $I(0)$ null hypothesis based on the ratio of partial sums of residuals. In their notation, they recommend the tests $\mathcal{K}_1, \mathcal{K}_1'$ and $\mathcal{K}_4 = \max(\mathcal{K}_1, \mathcal{K}_1')$, where $\mathcal{K}_1$ and $\mathcal{K}_1'$ are designed to detect the $I(0)$-$I(1)$ and $I(1)$-$I(0)$ alternatives, respectively. We do not reproduce the expressions for the tests here and refer the reader to their original article for details. The wild bootstrap for the hybrid tests are implemented using a two point distribution ($v_t \in \{-1, 1\}$ with equal probability for each value) while the standard normal distribution is used for the CT tests. While the choice of distribution was found to have little impact on the CT tests (as also noted by the authors), our simulations showed notably improved finite sample properties of the hybrid tests when using the two-point distribution relative to the normal.

7.1 Finite Sample Size

In the no persistence change case, the time series $\{y_t\}$ is generated as

- DGP-0:

$$\begin{aligned} y_t &= \alpha y_{t-1} + z_t \\ y_0 &= 0 \end{aligned} \tag{23}$$

Table 1 presents the empirical size of 5% bootstrap tests. As expected, the CT tests have rejection probabilities close to 5% under $H_0^{(0)}$, i.e., $\alpha < 1$. Under $H_0^{(1)}$, however, these tests are severely oversized for all three volatility models (including the homoskedastic case) and serial correlation configurations. The proposed tests, however, are robust to whether the process is $I(1)$ or $I(0)$. The H^* tests are more accurately sized when $\alpha \in \{.5, 1\}$ but less so when $\alpha = .7$. This is because the tests are a hybrid of the KPZ and BP tests which individually have size close to 5% when $\alpha = 1$ and $\alpha = .5$, respectively. When $\alpha = .7$, the BP tests are mildly over-sized while the KPZ tests diverge at rate T which explains the size distortions of the hybrid tests. Similar reasoning explains the slightly higher sizes observed for the hybrid tests when the sample size increases, especially when $\alpha = .7$ and the errors are MA(1).

test. No qualitative differences were observed if lag selection was implemented instead under the alternative model.

7.2 Finite Sample Power

To examine finite sample power, we consider DGPs with one and two breaks. To economize on space, the results are reported for the case $\rho = \theta = 0$ and briefly summarized for the other cases. The DGPs in the one break case are specified as follows:

	For $t \leq [T\lambda_1^0]$	For $t \geq [T\lambda_1^0] + 1$
DGP-1	$y_t = \alpha y_{t-1} + z_t$	$y_t = y_{t-1} + z_t$
DGP-2	$y_t = y_{t-1} + z_t$	$y_t - y_{[T\lambda_1^0]} = \alpha(y_{t-1} - y_{[T\lambda_1^0]}) + z_t$
DGP-3	$y_t = \alpha_1 y_{t-1} + z_t$	$y_t = \alpha_2 y_{t-1} + z_t$

For DGP-1 and DGP-2, we let $\alpha \in \{.5, .7\}$ while for DGP-3, we take $\alpha_1, \alpha_2 \in \{.2, .9\}$ and define $\alpha = \alpha_2 - \alpha_1$. The breakpoint is set at $\lambda_1^0 = .5$. The findings are presented in Table 2. We report size-adjusted power so that all tests have empirical size at most 5% under H_0.[9] Several features of the results are worth noting. First, the hybrid tests are generally more powerful than the CT tests. The exception occurs in the case of an $I(1)$ regime with (relatively) high volatility (e.g., DGP-1 with $\delta = 1/3$ and DGP-2 with $\delta = 3$). This pattern arises since such a process is dominated by the $I(1)$ component so that the sampling behavior of the tests mimics that obtained with a stable $I(1)$ process. In Appendix B (Table B-2), we show that power improves considerably if the $I(0)$ regime is longer and/or the volatility shift is less prominent. Second, the CT tests only have trivial power against $I(0)$-preserving breaks (DGP-3) which is expected given that they are designed to detect changes between $I(1)$ and $I(0)$ regimes. In contrast, the hybrid tests have substantial power in this case implying that using the KPZ tests to control size in the $I(1)$ case causes little power loss relative to using the BP tests in isolation. Third, the hybrid tests are generally more powerful when the time series is driven by deterministic volatility (Models 1 and 2) rather than stochastic volatility (Model 3). With serially correlated errors, the results (Tables B-3 and B-4 in Appendix B) are qualitatively similar except that power is lower for all the tests relative to $\rho = \theta = 0$.

[9] Size adjustment is important for power comparison given that the CT tests are considerably over-sized in the $I(1)$ case.

With two breaks, the DGPs are specified as follows:

	For $t \le [T\lambda_1^0]$	For $[T\lambda_1^0]+1 \le t \le [T\lambda_2^0]$	For $t \ge [T\lambda_2^0]+1$
DGP-4	$y_t = y_{t-1} + z_t$	$y_t - y_{[T\lambda_1^0]} = \alpha(y_{t-1} - y_{[T\lambda_1^0]}) + z_t$	$y_t = y_{t-1} + z_t$
DGP-5	$y_t = \alpha y_{t-1} + z_t$	$y_t = y_{t-1} + z_t$	$y_t - y_{[T\lambda_2^0]} = \alpha(y_{t-1} - y_{[T\lambda_2^0]}) + z_t$
DGP-6	$y_t = \alpha_1 y_{t-1} + z_t$	$y_t = \alpha_2 y_{t-1} + z_t$	$y_t = \alpha_1 y_{t-1} + z_t$

Define $\alpha = \alpha_2 - \alpha_1$ for DGP-6. The true breakpoint vector is set at $(\lambda_1^0, \lambda_2^0) = (.3, .8)$. The pattern of results, reported in Table 3, is broadly similar to the one break case with the power advantages of the hybrid tests more apparent, especially for DGP-4 and DGP-6. This is unsurprising given that the CT tests are directed against the alternative of a single break. An interesting feature of the hybrid tests is that when the first regime has lower persistence relative to the second, $H^*(1)$ has higher power than $H^*(2)$ even though the former is based on a misspecified model. However, $Hmax_1^*$ has adequate power in most cases, often close to that of the more powerful test among $H^*(1)$ and $H^*(2)$. This feature highlights the practical advantage of using $Hmax_1^*$ to detect the presence of at least one break in applications.

7.3 Comparison with Recursive Bootstrap

In order to highlight the advantages of employing the proposed bootstrap schemes A and B, we now provide a comparison with the fully recursive bootstrap schemes. The recursive counterpart of scheme A entails replacing step 3 in scheme A with the recursion

$$\begin{aligned} y_t^{(1)} &= y_{t-1}^{(1)} + \sum_{j=1}^{\breve{l}_T} \breve{\pi}_j \Delta y_{t-j}^{(1)} + u_t^{(1)}; \quad t = \breve{l}_T + 2, ..., T \\ y_t^{(1)} &= y_t; \; t = 1, ..., \breve{l}_T + 1 \end{aligned} \tag{24}$$

while the recursive counterpart of scheme B involves replacing step 3 in scheme B with the recursion

$$\begin{aligned} y_t^{(0)} &= \widetilde{c} + \widetilde{\alpha} y_{t-1}^{(0)} + \sum_{j=1}^{\widetilde{l}_T} \widetilde{\pi}_j \Delta y_{t-j}^{(0)} + u_t^{(0)}; \quad t = \breve{l}_T + 2, ..., T \\ y_t^{(0)} &= 0; \; t = 1, ..., \widetilde{l}_T + 1 \end{aligned} \tag{25}$$

Since the bootstrap data obtained from (24) and (25) are serially correlated, conditional on the original data, the bootstrap statistics will now need to be adjusted by including lagged first differences in the estimated regression as in the construction of the statistics based on the original data $\{y_t\}$. The lag length is again chosen using the BIC. Table 4 reports the empirical size and size-adjusted power (only in the single break case, for brevity) of the recursive bootstrap tests (denoted with a superscript "r") for $\rho = \theta = 0$. The procedure has accurate size in general with a tendency to under-reject in some cases. A power comparison with Table 2 reveals that the recursive bootstrap tests are generally less powerful than the hybrid tests for DGP-1 and DGP-2 which contain an $I(1)$ segment, in accordance with the discussion in Section 5.1. For DGP-3, the two approaches yield comparable power. The power gains are even more transparent if one were to *a priori* rule out the $I(1)$ null hypothesis and hence apply the BP tests in isolation (see Tables B-7 and B-8 in Appendix B). Overall, these findings favor the use of the proposed scheme over the recursive scheme in terms of its relative ability in detecting persistence change.

7.4 Number of Breaks

The sequential testing algorithm developed in Section 5.4 is assessed by its effectiveness in estimating the true number of breaks when the data are generated by DGP 0-6. We apply the algorithm with $A = 2$ and $\eta = .10$.[10] The results are presented in Table 5, where P_c denoting the probability of "correct" selection and P_o denoting the probability of "over-estimation". The procedure is generally reliable when the time series is stable (DGP-0) or is subject to a single break (DGP 1-3) consistent with the findings in Tables 1 and 2. Its performance, however, deteriorates in the two breaks case where the likelihood of underestimation can be non-negligible, especially when the first regime has mild persistence. For instance, in DGP-5 with an abrupt increase in volatility (model 1 with $\delta = 1/3$), the breakpoint estimate used to partition the sample into two segments (following evidence of at least one break in the first step) is typically close to the second true breakpoint, so that the first segment effectively includes a break from $I(0)$ to $I(1)$ while the second segment is effectively $I(0)$. Whether a second break is selected then depends on the power of the single break test in the first segment, which is shown to be relatively low (see Table 2). Similarly, with decreasing volatility, the breakpoint is estimated near the first true date so that selecting an additional break depends on the power of the single break test in the $I(1)$-$I(0)$ case. In additional simulations reported in Appendix B (Table B-9), we observed a notable improvement in

[10] We also experimented with $\eta = .05$ but found that that the underestimation probabilities were considerable in many cases and that $\eta = .10$ appeared to provide the best compromise in terms of the size-power tradeoff.

performance as the magnitude of the volatility shift decreases and/or the volatility shift is such that it occurs near the second persistence break in the increasing volatility case and near the first break otherwise. Further refinement of the algorithm to increase its reliability in detecting multiple breaks is a potentially interesting topic for future research.

8 Empirical Application: OECD Inflation Rates

The persistence of inflation plays a key role in the formulation and evaluation of quantitative macroeconomic models. Evidence in favor of high and unchanged persistence from reduced form specifications across different monetary policy regimes has generally been construed as suggesting that inflation persistence is a feature that any reasonable model for the economy should be able to replicate. The Lucas Critique, on the other hand, suggests that the parameters of macroeconometric models depend implicitly on agents' expectations of the policy process and are unlikely to remain stable as policymakers change their behavior, if agents are forward looking. An empirical finding of high and stable persistence in such a context can potentially be interpreted either in terms of the presence of a strong backward looking component in the dynamics of inflation induced through, say indexation or rule-of-thumb behavior on the part of the price setters, or in terms of historical policy shifts being of relatively modest magnitude. In contrast, Erceg and Levin (2003) suggest that inflation persistence is not an inherent characteristic of the economy but rather varies with the credibility and transparency of the monetary regime. Similarly, Orphanides and Williams (2004) show that the absence of a long-run inflation objective for the monetary authority leads to substantially higher inflation persistence relative to an environment where the inflation objective is clearly understood by price-setters.

While early empirical studies (e.g., Cogley and Sargent, 2001; Stock, 2001) examined the stability of inflation persistence within a framework that assumes constant unconditional volatility, subsequent investigations recognized the importance of accounting for time-varying volatility and its potential impact on our understanding of the nature of persistence. Cogley and Sargent (2005) estimate Bayesian VARs with drifting coefficients and stochastic volatility and conclude in favor of a decline in U.S. inflation persistence associated with a change in monetary policy while Primiceri (2005) employs similar but more general methods to stress the importance of exogenous non-policy factors in explaining the evolution of inflation. Cogley, Primiceri and Sargent (2010) decompose inflation into a trend component and an inflation gap component arguing that while inflation has remained persistent due to movements in trend inflation associated with shifts in the Federal Reserve's target, inflation gap persistence increased during the Great Inflation and declined after the Volcker disinflation. More recently, Bataa et al. (2013) examine the evidence for G-7 countries within a VAR framework utilizing the multivariate break detection procedure of Qu and Perron (2007) while Bataa et al. (2014) employ the univariate Bai and Perron approach to identify breaks in mean, dynamics and the volatility of inflation. Both studies assume that inflation is a $I(0)$ process under the null hypothesis of stability and regime-wise $I(0)$ in the presence of breaks, thereby ruling out the possibility of $I(1)$ regimes. Kejriwal (2018) investigates the persistence of OECD inflation rates allowing for unit roots but assumes conditional homoskedasticity. The present paper contributes to this literature by employing the proposed bootstrap approach to disentangle breaks in mean, persistence and volatility.

Our empirical investigation is based on monthly CPI inflation data for nineteen OECD countries used in Noriega et al. (2013) and Belaire-Franch (2017). The data span the period 1960:1-2008:6 so that $T = 582$, except for Germany and Korea where the starting point is 1960:2. The inflation rates are seasonally unadjusted and computed as $i_t = 1200(\ln P_t - \ln P_{t-1})$, where P_t denotes the CPI at time t.[11] The results are reported in Table 6. To conserve space, we present only a subset of the results here, with the full set reported in Table B-10. The analysis proceeds in six steps which we describe below.

First, we apply the sequential algorithm detailed in Section 5.4 with $A = 5$, $\epsilon = .15$ and $\eta = .10$ to estimate the number of breaks $\hat{m}$ [column (2)]. **Second**, conditional on $\hat{m}$, the breakpoint estimates are obtained by minimizing the unrestricted sum of squared residuals [column (3)]. **Third**, to distinguish persistence shifts from pure mean shifts, we conduct Wald tests (at the 10% level) of the null hypothesis that the process is subject to $\hat{m}$ mean shifts against the alternative hypothesis of $\hat{m}$ mean as well as persistence shifts (see Section 6.2). This outcome is reported in column (4) where "Yes" indicates non-rejection of the pure mean shifts hypothesis and "No" indicates otherwise. Here, a heteroskedasticity robust standard error estimate is used to construct the statistics although a wild bootstrap approach could also be used.[12]

Fourth, based on the selected model, the largest (across regimes) estimated sum of autoregressive parameters ("largest AR sum") is computed [column (5)] along with equal-tailed 90% confidence intervals [column (6)] based on the procedure advocated by Andrews and Guggenberger (2014, AG henceforth) that allows uniformly valid inference over the

[11] We prefer to use seasonally unadjusted rates since commonly used adjustment procedures such as Census X-11 or X-12 can have adverse effects on the power of structural change tests by smoothing the time series of interest (see Ghysels and Perron, 1996).

[12] Employing the wild bootstrap procedure in the present context yielded results identical to those reported.

stationary and non-stationary regions of the parameter space as well as heteroskedasticity.[13] The BIC is used to select the number of lags within each regime with a maximum allowable lag length of 12. **Fifth**, a comparison with the CT procedure is included to highlight the differences among the two approaches in terms of model selection [columns (7) and (8)]. **Sixth**, unit root tests allowing for nonstationary volatility (Cavaliere and Taylor, 2009) are conducted as a robustness check on the model selection results [column (9)].

We now turn to a discussion of the empirical results. Evidence of at least one break ($\hat{m} > 0$) is obtained for seven countries, of which two (Austria, Korea) favor an $I(0)$ process with a single mean shift. The AG interval estimates are consistent with the presence of at least one $I(1)$ segment in fourteen countries of which three are subject to at least one persistence break. Interestingly, in four out of the five countries with persistence breaks, the first break corresponds to an increase in persistence that occurs between the early and mid Seventies, a period often described as one of "the Great Inflation" and commonly believed to be associated with both a high level and high degree of persistence. In contrast, for France and Germany which experience two persistence breaks, the second break is associated with a persistence decline occurring in the Eighties and Nineties (see Table B-10).

Next, we provide a comparison between our results and those from the CT procedure. For the latter, we first apply the test $\mathcal{K}_4$ (at the 10% level with 15% trimming) designed to detect a single persistence change [$I(1)$-$I(0)$ or $I(0)$-$I(1)$]. Upon a rejection, the p-values of the tests $\mathcal{K}_1$ and $\mathcal{K}_1'$ are computed and the direction of persistence change is determined by the smaller of the two p-values. If both p-values are zero, the evidence is not conclusive, i.e., the procedure cannot distinguish between the two alternatives (see Section 7 of CT). Comparing the outcomes in columns (7) and (8) shows that the procedures agree only for Luxembourg and Netherlands while they point to different models for all other countries. The CT approach is inconclusive in seven cases. Further, in all of the eleven cases where the proposed approach decides in favor of a pure $I(1)$ process, the CT procedure suggests a break in persistence. This pattern is consistent with the size distortions incurred by the latter approach when the process is $I(1)$, as documented by the simulation evidence presented in Section 7. In the two cases where we select a pure mean shift process, the CT approach again points to a persistence break which can be potentially explained by the non-robustness of the approach to such processes.

Column (9) supplements our analysis with unit root tests applied to the regime with

[13] A caveat is that the procedure allows conditional heteroskedasticity while assuming constant unconditional heteroskedasticity.

the largest autoregressive sum estimate based on the selected model in column (7). To this end, we report the p-value of wild bootstrap ADF test proposed by Cavaliere and Taylor (2009) which is robust to nonstationary volatility.[14] The lag length in the ADF regression was selected using the Modified Akaike Information Criterion (MAIC) with the maximum lag set at $[12(T/100)^{1/4}]$. The test findings match the model selection outcomes in column (7) for fourteen out of the nineteen countries, indicating a fair degree of consistency between the two approaches.

To justify the importance of allowing for nonstationary volatility in the current context, Figure 1 plots the volatility estimates obtained by fitting a nonparametric regression to the squared residuals obtained by estimating the model selected in column (7), as suggested by Xu and Phillips (2008). A Gaussian kernel is employed with the bandwidth chosen by cross validation.[15] The estimates indicate considerable variation over time although the nature of the variation is different across countries. While a smooth trend suggests itself for some countries (e.g., France and Norway), more irregular movements are observed for others (e.g, Belgium, UK, USA). A similar overall picture is obtained if one plots the estimated variance profile as suggested by Cavaliere and Taylor (2007) indicating the nonstationary behavior of the sample volatility paths is a key feature of the inflation data that, if ignored,.might lead to potentially misleading inferential results.

Finally, to evaluate the impact of nonstationary volatility on persistence change, it is useful to compare our results with the asymptotic sequential procedure of Kejriwal (2018) which assumes homoskedasticity. Using the same dataset, Kejriwal (2018) concludes in favor of a persistence change model for six additional countries (Canada, Finland, Greece, Japan, UK, USA) all of which are found to be pure $I(1)$ processes according to our analysis that accounts for nonstationary volatility. Interestingly, Kejriwal's analysis for USA suggests a shift from a high persistence $I(0)$ regime to a low persistence $I(0)$ regime, consistent with the views expressed in Sims (2001) and Stock (2001) that the case for unstable persistence is weakened once allowance is made for shifts in the variance of the innovations.

[14] These authors also propose a wild bootstrap version of the Phillips and Perron (1988) test based on a semiparametric correction for serial correlation. Our choice is motivated by the better size control of the ADF test in finite samples (see Cavaliere and Taylor, 2009).

[15] Following Xu and Phillips (2008), we search over the bandwidths $h_i = c_i T^{-0.4}$, $i = 1, ..., 4$, where $\{c_1, c_2, c_3, c_4\} = \{.25, .4, .6, .75\}$.

9 Conclusion

This paper proposes wild bootstrap sup-Wald tests for detecting persistence change in a univariate time series driven by nonstationary volatility. We develop tests both against a specified number of structural changes as well as procedures that do not assume knowledge of the number of breaks. The set of alternative hypotheses considered include processes that are characterized by switches between $I(1)$ and $I(0)$ regimes and those that preserve the $I(0)$ nature of the time series in each regime. An alternative strategy to the bootstrap that can be employed to account for unstable volatility is an adaptive least squares approach based on a non-parametric estimate of the variance function. This approach has been taken by Xu and Phillips (2008) to obtain efficient estimators relative to ordinary least squares in an autoregressive model and by Beare (2017) to develop unit root tests robust to unstable volatility that have pivotal limiting distributions. The adaptive method can also be potentially applied in the context of persistence change testing to obtain asymptotically pivotal test statistics. A comparison of the bootstrap and adaptive approaches in finite samples would be of interest in order to evaluate their relative merits. We leave the exploration of these issues as possible avenues for future research.

References

Andrews, D.W.K. (1993), "Tests for parameter instability and structural change with unknown change point," *Econometrica* 61, 821-856.

Andrews, D.W.K., Ploberger, W. (1994), "Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica* 62, 1383-1414.

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

Bai, J., Perron, P. (2003a), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics* 18, 1-22.

Bai, J., Perron, P. (2003b), "Critical values for multiple structural change tests," *Econometrics Journal* 6, 72-78.

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H., Taylor, R.L. (1991), "Bootstrapping unstable first-order autoregressive processes," *Annals of Statistics* 19, 1098-1101.

Bataa, E., Osborn, D. R., Sensier, M., van Dijk, D. (2013), "Structural breaks in the international dynamics of inflation," *Review of Economics and Statistics* 95, 646-659.

Bataa, E., Osborn, D. R., Sensier, M., van Dijk, D. (2014), "Identifying changes in mean, seasonality, persistence and volatility for G7 and Euro Area inflation," *Oxford Bulletin of Economics and Statistics* 76, 360-388.

Beare, B. K. (2017), "Unit root testing with unstable volatility," *Journal of Time Series Analysis* DOI: 10.1111/jtsa.12279

Belaire-Franch, J. (2017), "A note on the evidence of inflation persistence around the world," *Empirical Economics* 1-11.

Busetti, F., Taylor, A.M.R. (2004), "Tests of stationarity against a change in persistence," *Journal of Econometrics* 123, 33-66.

Cavaliere, G. (2005), "Unit root tests under time-varying variances," *Econometric Reviews* 23, 259-292.

Cavaliere, G., Taylor, A.M.R. (2005), "Stationarity tests under time-varying second moments," *Econometric Theory* 21, 1112-1129.

Cavaliere, G., Taylor, A.M.R. (2007), "Testing for unit roots in time series models with non-stationary volatility," *Journal of Econometrics* 140, 919-947.

Cavaliere, G., Taylor, A.M.R. (2008a), "Bootstrap unit root tests for time series with non-stationary volatility," *Econometric Theory* 24, 43-71.

Cavaliere, G., Taylor, A.M.R. (2008b), "Testing for a change in persistence in the presence of non-stationary volatility," *Journal of Econometrics* 147, 84-98.

Cavaliere, G., Taylor, A.M.R. (2009), "Heteroskedastic time series with a unit root," *Econometric Theory* 25, 1228-1276.

Cogley, T., Primiceri, G. E., Sargent, T. J. (2010), "Inflation-gap persistence in the US," *American Economic Journal: Macroeconomics* 2, 43-69.

Cogley, T., Sargent, T. J. (2001), "Evolving post-world war II US inflation dynamics," *NBER macroeconomics annual* 16, 331-373.

Cogley, T., Sargent, T. J. (2005), "Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of Economic dynamics* 8, 262-302.

Davidson, J. (1994), "Stochastic Limit Theory," Oxford: Oxford University Press

Diebold, F.X., Chen, C. (1996), "Testing structural stability with endogenous breakpoint a size comparison of analytic and bootstrap procedures," *Journal of Econometrics* 70, 221-241.

Diebold, F.X., Inoue, A. (2001), "Long memory and regime switching," *Journal of Econometrics* 105, 131-159.

Erceg, C. J., Levin, A. T. (2003), "Imperfect credibility and inflation persistence," *Journal of monetary economics* 50, 915-944.

Georgiev, I. (2008), "Asymptotics for cointegrated processes with infrequent stochastic level shifts and outliers," *Econometric theory* 24, 587-615.

Georgiev, I., Harvey, D. I., Leybourne, S. J., Taylor, A.M.R. (2018), "Testing for parameter instability in predictive regression models," *Journal of Econometrics* 204, 101-118.

Ghysels, E., Perron, P. (1993), "The effect of seasonal adjustment filters on tests for a unit root," *Journal of Econometrics* 55, 57-98.

Giné, E., Zinn, J. (1990), "Bootstrapping general empirical measures," *Annals of Probability* 18, 851-869.

Gulesserian, S.G., Kejriwal, M. (2014), "On the power of bootstrap tests for stationarity: a Monte Carlo comparison," *Empirical Economics* 46, 973-998.

Hansen, B.E. (1995), "Regression with nonstationary volatility," *Econometrica* 63, 1113-1132.

Hansen, B.E. (1996), "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica* 64, 413-430.

Hansen, B.E. (1999), "The grid bootstrap and the autoregressive model," *Review of Economics and Statistics* 81, 594-607.

Hansen, B.E. (2000), "Testing for structural change in conditional models," *Journal of Econometrics* 97, 93-115.

Harvey, D.I., Leybourne, S.J., Taylor, A.M.R. (2006), "Modified tests for a change in persistence," *Journal of Econometrics* 134, 441-469.

Harvey, D.I., Leybourne, S.J., Taylor, A.M.R. (2009), "Simple, robust, and powerful tests of the breaking trend hypothesis," *Econometric Theory* 25, 995-1029.

Kejriwal, M. (2018), "A robust sequential procedure for estimating the number of structural changes in persistence," Working Paper, Purdue University.

Kejriwal, M., Perron, P. (2010), "Testing for multiple structural changes in cointegrated regression models," *Journal of Business & Economic Statistics* 28, 503-522.

Kim, J.Y. (2000), "Detection of change in persistence of a linear time series," *Journal of Econometrics* 54, 159-178.

Leybourne, S.J., Kim, T., Taylor, A.M.R. (2007), "Detecting multiple changes in persistence," *Studies in Nonlinear Dynamics & Econometrics* 11(3)

Liu, R.Y. (1988), "Bootstrap procedures under some non-iid models," *Annals of Statistics* 16, 1696-1708.

Loretan, M., Phillips, P.C.B. (1994), "Testing the covariance stationarity of heavy-tailed time series," *Journal of Empirical Finance* 1, 211-248.

Nelson, D.B. (1990), "Stationarity and persistence in the GARCH (1, 1) model," *Econometric theory* 6, 318-334.

Noriega, A. E., Capistrán, C., Ramos-Francia, M. (2013), "On the dynamics of inflation persistence around the world," *Empirical Economics* 44, 1243-1265.

Orphanides, A., Williams, J. (2004), "Imperfect knowledge, inflation expectations, and monetary policy" *The inflation-targeting debate* (pp. 201-246). University of Chicago Press.

Park, J.Y. (2002), "Nonstationary nonlinear heteroskedasticity," *Journal of Econometrics* 110, 383-415.

Pitarakis, J. Y. (2004), "Least squares estimation and tests of breaks in mean and variance under misspecification," *Econometrics Journal* 7, 32-54.

Perron, P., Yabu, T. (2009), "Estimating deterministic trends with an integrated or stationary noise component," *Journal of Econometrics* 151, 56-69.

Perron, P., Zhou, J. (2008), "Testing jointly for structural changes in the error variance and coefficients of a linear regression model," Unpublished Manuscript, Department of Economics, Boston University.

Phillips, P. C., Perron, P. (1988), "Testing for a unit root in time series regression," *Biometrika* 75, 335-346.

Phillips, P.C.B., Xu, K.L. (2006), "Inference in autoregression under heteroskedasticity," *Journal of Time Series Analysis* 27, 289-308.

Primiceri, G. E. (2005), "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies* 72, 821-852.

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

Sensier, M., van Dijk, D. (2004), "Testing for volatility changes in US macroeconomic time series," *Review of Economics and Statistics* 86, 833-839.

Sims, C. A. (2001), "Evolving Post-World War II US Inflation Dynamics: Comment," *NBER Macroeconomics Annual* 16, 373-379.

Stock, J. H. (2001), "Discussion of Cogley and Sargent,'Evolving Post-World War II Inflation Dynamics'," *NBER Macroeconomics Annual* 16, 379-386.

Xu, K.L. (2008), "Bootstrapping autoregression under non-stationary volatility," *Econometrics Journal* 11, 1-26.

Xu, K.L. (2015), "Testing for structural change under non-stationary variances," *Econometrics Journal* 18, 274-305.

Xu, K.L., Phillips, P.C.B. (2008), "Adaptive estimation of autoregressive models with time-varying variances," *Journal of Econometrics* 142, 265-280.

Zhou, J., Perron, P. (2008), "Testing for breaks in coefficients and error variance: Simulations and applications," Unpublished Manuscript, Department of Economics, Boston University.

Table 1: Empirical size of bootstrap tests, [$m = 0$, 5% nominal size]

T	(ρ,θ)	Test	$\alpha = 1$							$\alpha = 0.5$							$\alpha = 0.7$						
			Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
200	(0,0)	$\mathcal{K}_1$	.45	.64	.38	.58	.36	.49	.48	.05	.07	.04	.05	.04	.10	.05	.06	.09	.06	.08	.05	.13	.06
		$\mathcal{K}_1'$	.43	.37	.60	.36	.58	.48	.48	.03	.04	.05	.04	.04	.09	.04	.05	.05	.07	.04	.05	.10	.04
		$\mathcal{K}_4$	.55	.66	.62	.61	.61	.64	.59	.05	.07	.05	.05	.04	.14	.05	.07	.09	.07	.08	.06	.18	.06
		H_1^*	.06	.06	.04	.06	.03	.06	.06	.05	.03	.04	.04	.04	.03	.03	.08	.06	.04	.07	.05	.04	.03
		H_2^*	.06	.05	.06	.06	.06	.05	.06	.04	.02	.01	.05	.03	.03	.02	.07	.03	.03	.06	.04	.03	.04
		H_{max}^*	.06	.05	.05	.06	.05	.05	.05	.05	.03	.02	.04	.03	.03	.03	.08	.05	.03	.06	.04	.04	.03
	(0.5,0)	$\mathcal{K}_1$	.44	.64	.37	.60	.36	.49	.49	.04	.08	.05	.05	.04	.12	.04	.05	.08	.06	.06	.04	.14	.05
		$\mathcal{K}_1'$	.42	.34	.60	.33	.58	.46	.45	.03	.04	.05	.03	.04	.09	.03	.04	.05	.07	.04	.06	.10	.04
		$\mathcal{K}_4$	.55	.65	.62	.61	.59	.65	.58	.05	.08	.05	.06	.04	.18	.05	.05	.09	.07	.07	.06	.20	.05
		H_1^*	.06	.06	.05	.06	.05	.05	.06	.08	.04	.04	.05	.05	.04	.04	.10	.05	.05	.06	.06	.04	.04
		H_2^*	.08	.07	.07	.05	.07	.06	.06	.04	.02	.03	.04	.03	.03	.03	.07	.04	.04	.07	.05	.04	.03
		H_{max}^*	.07	.07	.07	.06	.07	.06	.06	.07	.04	.02	.04	.04	.03	.03	.09	.04	.04	.06	.04	.04	.04
	(0,0.5)	$\mathcal{K}_1$	.47	.65	.40	.61	.37	.50	.50	.03	.05	.04	.04	.04	.08	.04	.06	.08	.06	.07	.06	.10	.06
		$\mathcal{K}_1'$	.44	.37	.61	.37	.58	.48	.48	.04	.03	.03	.03	.03	.06	.04	.05	.05	.06	.05	.05	.09	.05
		$\mathcal{K}_4$	.58	.67	.62	.64	.62	.65	.62	.03	.05	.03	.04	.03	.10	.04	.07	.08	.06	.07	.05	.14	.06
		H_1^*	.08	.06	.05	.06	.04	.05	.06	.04	.02	.03	.02	.03	.02	.02	.10	.04	.06	.05	.05	.04	.05
		H_2^*	.06	.05	.04	.06	.05	.06	.04	.01	.01	.01	.02	.01	.02	.01	.06	.02	.02	.05	.03	.02	.03
		H_{max}^*	.06	.06	.04	.06	.04	.05	.05	.03	.01	.02	.02	.02	.02	.02	.09	.03	.03	.05	.04	.02	.03
400	(0,0)	$\mathcal{K}_1$	.42	.62	.36	.54	.35	.49	.46	.04	.06	.04	.06	.04	.09	.04	.04	.06	.04	.06	.04	.10	.04
		$\mathcal{K}_1'$	.42	.35	.61	.34	.59	.46	.47	.04	.05	.05	.05	.05	.06	.04	.04	.06	.05	.05	.05	.07	.04
		$\mathcal{K}_4$	.51	.62	.63	.57	.60	.67	.56	.04	.06	.05	.06	.05	.11	.04	.05	.06	.05	.06	.05	.13	.05
		H_1^*	.04	.05	.06	.04	.04	.05	.05	.07	.05	.05	.06	.05	.04	.06	.08	.06	.05	.07	.05	.05	.06
		H_2^*	.05	.05	.06	.05	.05	.05	.05	.07	.03	.03	.05	.04	.03	.04	.08	.04	.03	.07	.04	.04	.05
		H_{max}^*	.04	.04	.05	.04	.05	.05	.05	.06	.05	.05	.05	.04	.04	.04	.08	.06	.04	.08	.05	.04	.06
	(0.5,0)	$\mathcal{K}_1$	.42	.62	.36	.55	.35	.50	.46	.03	.05	.03	.06	.03	.09	.03	.04	.05	.03	.05	.03	.10	.03
		$\mathcal{K}_1'$	.41	.34	.61	.34	.57	.45	.44	.03	.05	.05	.04	.05	.06	.02	.03	.04	.05	.04	.05	.07	.03
		$\mathcal{K}_4$	.50	.63	.63	.56	.58	.68	.55	.04	.05	.05	.06	.05	.12	.03	.04	.05	.05	.05	.05	.14	.03
		H_1^*	.05	.04	.05	.04	.05	.04	.06	.07	.06	.07	.08	.07	.05	.07	.10	.08	.07	.09	.08	.05	.07
		H_2^*	.05	.06	.05	.06	.05	.05	.05	.09	.04	.04	.09	.05	.05	.05	.11	.06	.05	.10	.06	.04	.06
		H_{max}^*	.05	.05	.05	.05	.05	.04	.05	.08	.06	.06	.08	.07	.04	.06	.10	.08	.06	.09	.07	.04	.06
	(0,0.5)	$\mathcal{K}_1$	.43	.63	.37	.56	.36	.51	.47	.04	.05	.03	.06	.04	.08	.04	.05	.07	.05	.07	.04	.10	.05
		$\mathcal{K}_1'$	.43	.35	.62	.35	.58	.46	.46	.04	.05	.04	.05	.05	.06	.03	.05	.07	.06	.06	.06	.07	.05
		$\mathcal{K}_4$	.54	.64	.64	.59	.61	.68	.58	.04	.05	.04	.06	.05	.09	.03	.06	.07	.06	.07	.06	.12	.06
		H_1^*	.06	.05	.07	.04	.07	.05	.06	.05	.06	.04	.05	.04	.04	.04	.15	.10	.09	.10	.10	.07	.08
		H_2^*	.06	.05	.05	.05	.05	.05	.04	.05	.02	.02	.04	.03	.03	.02	.15	.05	.05	.12	.08	.05	.06
		H_{max}^*	.06	.05	.05	.04	.05	.04	.04	.05	.05	.04	.05	.03	.04	.03	.16	.09	.09	.11	.10	.05	.07

Table 2: Size-adjusted power of bootstrap tests, $[m = 1,\ \rho = \theta = 0,\ \lambda_1^0 = 0.5,\ 5\%]$

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	1	$\mathcal{K}_1$	.33	.65	.07	.52	.20	.43	.38	.22	.52	.03	.40	.09	.35	.26
		$\mathcal{K}_1'$	.03	.04	.03	.02	.07	.09	.05	.04	.04	.03	.02	.07	.09	.04
		$\mathcal{K}_4$	.32	.63	.05	.50	.18	.43	.36	.20	.50	.03	.39	.09	.37	.24
		H_1^*	.97	.04	.84	.28	.95	.26	.53	.61	.04	.46	.06	.65	.10	.22
		H_2^*	.57	.05	.21	.25	.45	.10	.22	.16	.05	.12	.07	.18	.06	.08
		H_{max}^*	.91	.05	.67	.23	.89	.19	.38	.43	.04	.30	.05	.48	.08	.13
	2	$\mathcal{K}_1$	.04	.03	.03	.07	.02	.09	.05	.03	.03	.04	.07	.02	.10	.04
		$\mathcal{K}_1'$	.36	.08	.62	.21	.53	.41	.35	.22	.02	.53	.10	.41	.32	.25
		$\mathcal{K}_4$	.33	.06	.61	.18	.52	.42	.34	.19	.02	.52	.09	.39	.35	.24
		H_1^*	1.0	.77	.30	.94	.76	.32	.71	.85	.41	.21	.62	.39	.18	.42
		H_2^*	.71	.35	.06	.71	.23	.16	.31	.30	.18	.05	.36	.10	.09	.16
		H_{max}^*	.98	.69	.10	.93	.53	.24	.58	.74	.35	.07	.57	.17	.13	.29
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.00	.05	.00	.03	.00	.13	.02	.00	.00	.00	.00	.00	.01	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.03	.00	.02	.11	.00
		$\mathcal{K}_4$	.00	.05	.00	.02	.00	.13	.01	.00	.00	.03	.00	.02	.11	.00
		H_1^*	1.0	.06	.76	.69	.96	.31	.71	1.0	.78	.33	.95	.89	.35	.77
		H_2^*	.88	.04	.23	.60	.71	.15	.41	.93	.32	.06	.82	.59	.18	.46
		H_{max}^*	.99	.03	.70	.67	.95	.24	.61	1.0	.72	.17	.95	.83	.28	.69
			$\alpha = 0.5$							$\alpha = 0.7$						
400	1	$\mathcal{K}_1$	.52	.77	.24	.68	.42	.54	.57	.41	.68	.11	.55	.25	.47	.44
		$\mathcal{K}_1'$	.03	.03	.02	.02	.06	.10	.06	.03	.03	.02	.02	.06	.11	.06
		$\mathcal{K}_4$	.50	.76	.20	.67	.39	.54	.55	.39	.66	.08	.54	.23	.48	.41
		H_1^*	1.0	.07	1.0	.95	1.0	.51	.84	1.0	.05	.95	.33	.99	.31	.62
		H_2^*	.99	.05	.69	.78	.96	.32	.57	.69	.04	.45	.23	.67	.15	.28
		H_{max}^*	1.0	.06	1.0	.93	1.0	.42	.75	.99	.05	.87	.26	.98	.22	.43
	2	$\mathcal{K}_1$	.04	.03	.03	.06	.02	.11	.05	.03	.03	.03	.07	.02	.11	.05
		$\mathcal{K}_1'$	.54	.24	.79	.41	.69	.53	.54	.39	.10	.69	.24	.56	.44	.41
		$\mathcal{K}_4$	.51	.20	.78	.37	.68	.55	.53	.37	.08	.68	.21	.55	.48	.40
		H_1^*	1.0	1.0	.53	1.0	.99	.56	.92	1.0	.93	.31	.99	.78	.40	.78
		H_2^*	1.0	.78	.06	.99	.67	.39	.67	.83	.59	.05	.91	.24	.23	.41
		H_{max}^*	1.0	1.0	.20	1.0	.96	.48	.84	1.0	.90	.11	.99	.54	.29	.64
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.00	.02	.00	.01	.00	.13	.01	.00	.00	.00	.00	.00	.00	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.02	.00	.01	.10	.01
		$\mathcal{K}_4$	.00	.02	.00	.01	.00	.13	.01	.00	.00	.02	.00	.01	.10	.01
		H_1^*	1.0	.45	1.0	1.0	1.0	.54	.93	1.0	1.0	.83	1.0	1.0	.56	.95
		H_2^*	1.0	.17	.91	.99	1.0	.42	.82	1.0	.92	.24	1.0	.98	.44	.84
		H_{max}^*	1.0	.27	1.0	1.0	1.0	.49	.90	1.0	1.0	.57	1.0	1.0	.52	.92

Table 3: Size-adjusted power of bootstrap tests, $[m = 2,\ \rho = \theta = 0,\ \lambda_1^0 = 0.3,\ \lambda_2^0 = 0.8,\ 5\%]$

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	4	$\mathcal{K}_1$	.03	.06	.01	.06	.01	.12	.03	.02	.06	.01	.06	.00	.11	.03
		$\mathcal{K}_1'$	.01	.00	.08	.00	.04	.11	.03	.01	.00	.07	.00	.04	.11	.02
		$\mathcal{K}_4$	.01	.06	.08	.05	.04	.21	.04	.01	.06	.07	.05	.04	.20	.04
		H_1^*	.36	.11	.42	.16	.44	.13	.20	.25	.03	.25	.07	.29	.08	.12
		H_2^*	.95	.62	.61	.89	.83	.26	.57	.68	.22	.25	.42	.44	.11	.25
		H_{max}^*	.94	.51	.58	.73	.77	.21	.51	.63	.12	.24	.22	.40	.09	.21
	5	$\mathcal{K}_1$	.44	.62	.36	.58	.33	.47	.48	.30	.54	.21	.47	.19	.39	.35
		$\mathcal{K}_1'$	.43	.36	.61	.35	.55	.44	.47	.31	.24	.51	.22	.45	.36	.35
		$\mathcal{K}_4$	.41	.59	.56	.51	.47	.58	.47	.30	.50	.48	.41	.37	.50	.36
		H_1^*	.73	.65	.82	.72	.87	.28	.41	.34	.35	.38	.40	.45	.15	.22
		H_2^*	.60	.17	.10	.54	.26	.15	.26	.16	.06	.04	.18	.08	.07	.11
		H_{max}^*	.65	.59	.52	.68	.68	.23	.31	.23	.28	.17	.35	.25	.11	.15
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.00	.00	.00	.00	.05	.00	.00	.00	.00	.00	.00	.05	.00
		$\mathcal{K}_1'$	.00	.01	.00	.00	.00	.04	.01	.00	.00	.00	.00	.00	.03	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.07	.00	.00	.00	.00	.00	.00	.08	.00
		H_1^*	.77	.74	.88	.75	.90	.29	.44	.50	.38	.69	.34	.64	.20	.27
		H_2^*	.85	.40	.26	.76	.58	.20	.38	.95	.65	.62	.85	.83	.27	.64
		H_{max}^*	.81	.70	.78	.74	.84	.25	.37	.91	.56	.70	.68	.75	.21	.51
			$\alpha = 0.5$							$\alpha = 0.7$						
400	4	$\mathcal{K}_1$	.01	.05	.00	.04	.00	.13	.02	.01	.05	.00	.55	.26	.46	.43
		$\mathcal{K}_1'$	.01	.00	.09	.00	.05	.12	.03	.01	.00	.08	.02	.06	.10	.06
		$\mathcal{K}_4$	.02	.05	.09	.03	.05	.24	.04	.01	.05	.08	.54	.22	.48	.42
		H_1^*	.46	.27	.59	.28	.54	.22	.26	.44	.19	.48	.22	.46	.14	.22
		H_2^*	1.0	.99	.96	1.0	1.0	.56	.91	.99	.79	.73	.95	.94	.33	.69
		H_{max}^*	1.0	.99	.97	1.0	.99	.45	.85	.98	.71	.71	.87	.89	.24	.57
	5	$\mathcal{K}_1'$	.53	.23	.79	.40	.69	.53	.64	.49	.41	.62	.40	.62	.50	.51
		$\mathcal{K}_4$	.51	.20	.78	.37	.68	.54	.63	.48	.64	.59	.56	.53	.64	.50
		H_1^*	.99	.98	1.0	.99	1.0	.66	.77	.91	.84	.92	.89	.95	.41	.51
		H_2^*	.98	.51	.49	.95	.82	.44	.52	.70	.13	.14	.58	.34	.22	.27
		H_{max}^*	.99	.96	.97	.98	.99	.55	.61	.77	.71	.68	.86	.84	.31	.34
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.00	.00	.00	.00	.04	.00	.00	.00	.00	.00	.00	.04	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.03	.01	.00	.00	.00	.00	.00	.03	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.06	.00	.00	.00	.00	.00	.00	.07	.00
		H_1^*	1.0	.99	1.0	1.0	1.0	.70	.84	.79	.88	.96	.73	.91	.46	.60
		H_2^*	1.0	.89	.88	1.0	.99	.53	.72	1.0	1.0	.98	1.0	1.0	.61	.93
		H_{max}^*	1.0	.99	1.0	1.0	1.0	.64	.77	1.0	.98	.98	1.0	.99	.55	.87

Table 4: Size and size-adjusted power of bootstrap recursive tests, [$\rho = \theta = 0$, $\lambda_1^0 = 0.5$, 5%]

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 1$							$\alpha = 0.5$						
200	0	H_1^r	.03	.05	.04	.05	.03	.06	.06	.04	.03	.04	.04	.03	.03	.03
		H_2^r	.02	.02	.03	.02	.02	.05	.03	.03	.02	.01	.04	.02	.03	.03
		H_{max}^r	.02	.04	.02	.05	.02	.05	.04	.04	.02	.02	.04	.02	.03	.03
			$\alpha = 0.5$							$\alpha = 0.7$						
	1	H_1^r	.95	.05	.74	.26	.92	.25	.51	.57	.04	.36	.06	.54	.08	.20
		H_2^r	.53	.04	.18	.25	.38	.11	.21	.15	.04	.07	.06	.14	.05	.07
		H_{max}^r	.89	.04	.59	.21	.84	.17	.37	.41	.04	.22	.05	.42	.07	.11
	2	H_1^r	.99	.74	.26	.92	.70	.30	.69	.79	.40	.16	.57	.28	.16	.39
		H_2^r	.66	.29	.05	.65	.22	.14	.29	.26	.14	.04	.30	.09	.08	.12
		H_{max}^r	.97	.65	.10	.90	.50	.23	.57	.69	.31	.06	.53	.16	.12	.26
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	H_1^r	.99	.08	.82	.74	.97	.34	.72	1.0	.88	.33	.99	.91	.42	.82
		H_2^r	.87	.05	.32	.64	.73	.19	.43	.93	.45	.06	.89	.59	.24	.51
		H_{max}^r	.99	.04	.77	.69	.97	.30	.64	1.0	.85	.17	.99	.85	.36	.73
			$\alpha = 1$							$\alpha = 0.5$						
400	0	H_1^r	.03	.04	.05	.04	.04	.05	.05	.05	.05	.04	.05	.05	.05	.06
		H_2^r	.02	.03	.03	.02	.02	.04	.03	.06	.03	.03	.04	.04	.03	.04
		H_{max}^r	.02	.02	.02	.04	.02	.05	.03	.05	.04	.04	.05	.04	.05	.05
			$\alpha = 0.5$							$\alpha = 0.7$						
	1	H_1^r	1.0	.07	1.0	.94	1.0	.50	.85	1.0	.06	.91	.31	.98	.31	.61
		H_2^r	.99	.03	.72	.77	.95	.30	.59	.68	.03	.42	.22	.60	.14	.28
		H_{max}^r	1.0	.04	1.0	.91	1.0	.41	.76	.98	.03	.83	.23	.96	.22	.42
	2	H_1^r	1.0	1.0	.48	1.0	.99	.58	.93	1.0	.92	.25	.99	.74	.39	.78
		H_2^r	.99	.77	.06	.99	.65	.39	.68	.82	.54	.05	.88	.22	.22	.40
		H_{max}^r	1.0	1.0	.20	1.0	.96	.50	.85	1.0	.89	.10	.99	.54	.29	.63
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	H_1^r	1.0	.47	1.0	1.0	1.0	.55	.94	1.0	1.0	.82	1.0	1.0	.60	.97
		H_2^r	1.0	.16	.96	.99	1.0	.46	.85	1.0	.96	.26	1.0	.98	.48	.87
		H_{max}^r	1.0	.30	1.0	1.0	1.0	.52	.92	1.0	1.0	.55	1.0	1.0	.55	.93

Table 5: Break selection probabilities, $[\rho = \theta = 0,\ \eta = .10]$

T	m	DGP	$\alpha/(\alpha_1, \alpha_2)$		Model 1: δ			Model 2: δ		Model 3: c	
				δ/c	1	1/3	3	1/3	3	0	10
400	0	0	1	P_c	.91	.91	.91	.92	.90	.90	.91
				P_o	.09	.09	.09	.08	.10	.10	.10
			0.5	P_c	.86	.91	.90	.88	.89	.91	.90
				P_o	.14	.09	.10	.12	.11	.10	.10
							$\lambda_1^0 = 0.5$				
	1	1	0.5	P_c	.90	.07	.91	.82	.89	.39	.68
				P_o	.10	.04	.09	.14	.11	.15	.14
		2	0.5	P_c	.88	.87	.21	.88	.79	.47	.73
				P_o	.12	.12	.12	.12	.20	.14	.15
		3	(0.2, 0.9)	P_c	.89	.42	.87	.88	.89	.46	.80
				P_o	.11	.08	.13	.12	.11	.13	.13
			(0.9, 0.2)	P_c	.88	.82	.62	.88	.85	.48	.81
				P_o	.12	.18	.14	.12	.15	.13	.12
							$\lambda_1^0 = 0.3, \lambda_2^0 = 0.8$				
	2	4	0.5	P_c	.86	.70	.46	.85	.70	.27	.59
				P_o	.10	.17	.16	.11	.13	.11	.14
		5	0.5	P_c	.83	.07	.08	.39	.28	.18	.36
				P_o	.15	.08	.06	.12	.18	.09	.10
		6	(0.2, 0.9)	P_c	.89	.08	.11	.84	.73	.24	.57
				P_o	.12	.07	.09	.12	.16	.09	.11
			(0.9, 0.2)	P_c	.90	.72	.62	.89	.84	.32	.75
				P_o	.09	.23	.20	.10	.12	.12	.11
600	0	0	1	P_c	.89	.88	.89	.89	.89	.88	.89
				P_o	.11	.12	.11	.11	.11	.12	.11
			0.5	P_c	.87	.89	.88	.88	.88	.91	.89
				P_o	.13	.11	.12	.12	.12	.09	.11
							$\lambda_1^0 = 0.5$				
	1	1	0.5	P_c	.88	.11	.89	.87	.89	.49	.80
				P_o	.12	.05	.11	.13	.11	.17	.11
		2	0.5	P_c	.89	.88	.37	.89	.86	.51	.79
				P_o	.11	.12	.21	.11	.14	.17	.15
		3	(0.2, 0.9)	P_c	.85	.71	.84	.85	.85	.56	.86
				P_o	.15	.17	.17	.15	.15	.12	.11
			(0.9, 0.2)	P_c	.83	.82	.75	.86	.82	.55	.85
				P_o	.17	.18	.22	.14	.18	.14	.12
							$\lambda_1^0 = 0.3, \lambda_2^0 = 0.8$				
	2	4	0.5	P_c	.86	.76	.64	.86	.78	.34	.69
				P_o	.11	.20	.15	.12	.12	.14	.16
		5	0.5	P_c	.87	.05	.08	.76	.53	.27	.51
				P_o	.13	.07	.08	.13	.22	.10	.13
		6	(0.2, 0.9)	P_c	.86	.33	.30	.85	.84	.33	.69
				P_o	.14	.12	.13	.15	.16	.11	.13
			(0.9, 0.2)	P_c	.90	.83	.78	.91	.89	.38	.81
				P_o	.10	.17	.18	.09	.11	.14	.13

Table 6: Break selection in OECD Inflation rates

Country	$\hat{m}$	Break Dates	Pure Mean Shifts	Largest AR sum	90% Band	Selected Model (H^*)	Selected Model ($\mathcal{K}$)	ADF p-value
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Austria	1	1969:7	Yes	.65	[.47, .87]	I(0) with one mean shift	I(0)-I(1)	.16
Belgium	1	1974:9	No	.82	[.71, 1.05]	I(0)-I(1)	I(1)-I(0)	.11
Canada	0	-	No	.88	[.81, 1.04]	I(1)	I(1)-I(0)	.14
Finland	0	-	No	.87	[.78, 1.06]	I(1)	I(0)-I(1)	.27
France	2	1973:1; 1983:4	No	.65	[.51, .89]	I(0)-I(0)-I(0)	I(0)-I(1)/I(1)-I(0)	.10
Germany	2	1970:9; 1993:7	No	.52	[.42, .64]	I(0)-I(0)-I(0)	I(0)-I(1)	.00
Greece	0	-	No	.81	[.64, 1.12]	I(1)	I(0)-I(1)/I(1)-I(0)	.10
Italy	2	1972:6; 1979:10	No	.88	[.82, 1.02]	I(0)-I(0)-I(1)	I(0)-I(1)/I(1)-I(0)	.07
Japan	0	-	No	.84	[.66, 1.18]	I(1)	I(0)-I(1)	.31
Korea	1	1981:9	Yes	.14	[.00, .29]	I(0) with one mean shift	I(0)-I(1)/I(1)-I(0)	.00
Luxembourg	1	1999:1	No	.87	[.76, 1.10]	I(1)-I(0)	I(1)-I(0)	.51
Netherlands	0	-	No	.04	[-.03, .12]	I(0)	I(0)	.00
Norway	0	-	No	.78	[.66, 1.02]	I(1)	I(0)-I(1)/I(1)-I(0)	.08
Portugal	0	-	No	.83	[.71, 1.08]	I(1)	I(0)-I(1)/I(1)-I(0)	.16
Spain	0	-	No	.88	[.77, 1.09]	I(1)	I(0)-I(1)/I(1)-I(0)	.35
Sweden	0	-	No	.82	[.71, 1.05]	I(1)	I(0)-I(1)	.10
Switzerland	0	-	No	.82	[.67, 1.11]	I(1)	I(0)-I(1)	.07
UK	0	-	No	.88	[.76, 1.11]	I(1)	I(0)-I(1)	.20
USA	0	-	No	.90	[.83, 1.05]	I(1)	I(1)-I(0)	.16

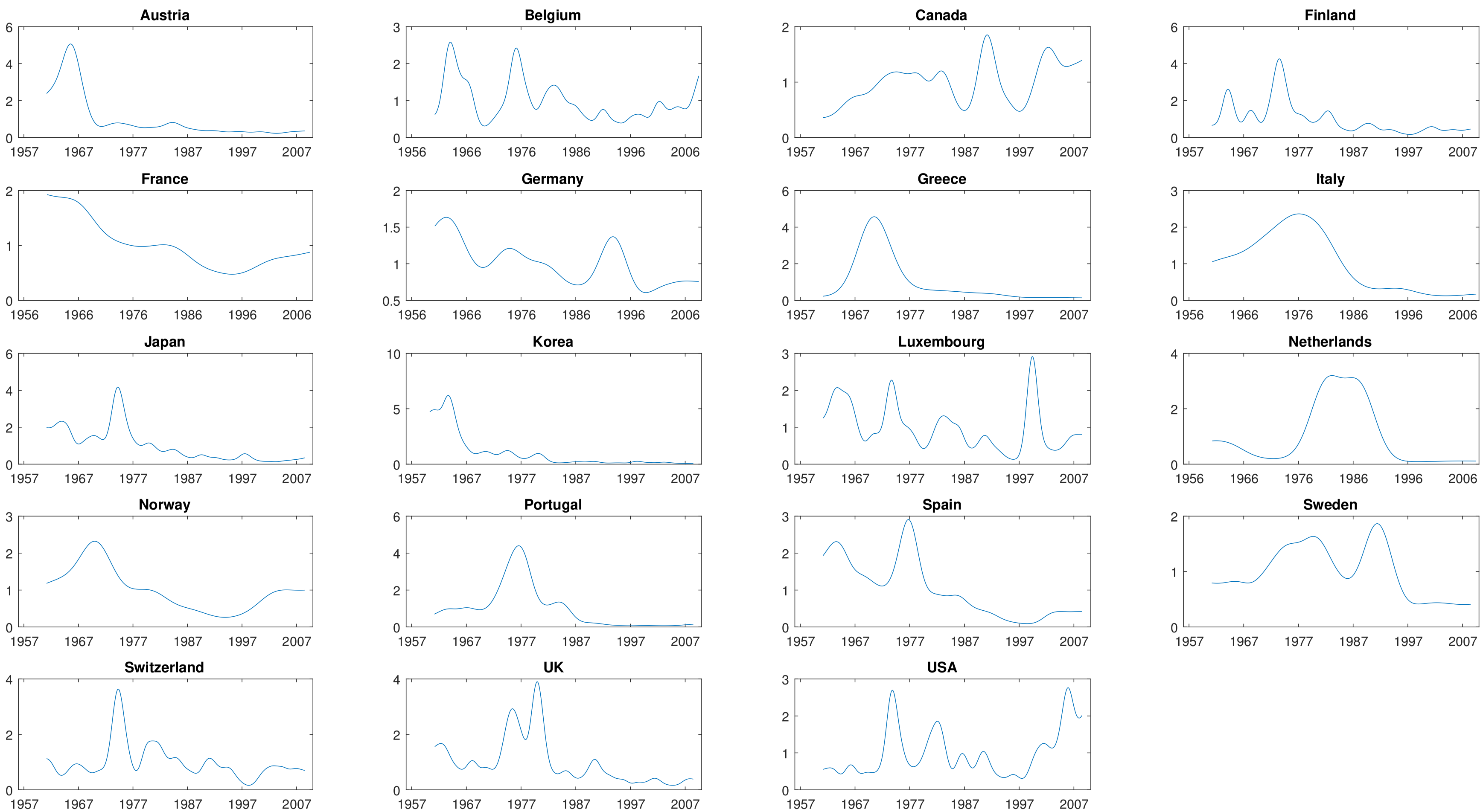

Figure 1: Nonparamatric volatility estimates (percentage of the total) for 19 OECD countries

Notes to Tables

1. Table 1 reports the empirical size of bootstrap tests with nominal size 5%. The tests $\mathcal{K}_1, \mathcal{K}_1', \mathcal{K}_4$ are the tests recommended by Cavaliere and Taylor (2008, CT) and H_1^*, H_2^*, H_{max}^* are our proposed tests.

2. Table 2 reports the size-adjusted power of 5% bootstrap tests in the single break case with breakpoint $\lambda_1^0 = .5$ and serially uncorrelated errors ($\rho = \theta = 0$).

3. Table 3 reports the size-adjusted power of 5% bootstrap tests in the two breaks case with breakpoint vector $(\lambda_1^0, \lambda_2^0) = (.3, .8)$ and serially uncorrelated errors ($\rho = \theta = 0$).

4. Table 4 reports size and size-adjusted power of 5% bootstrap recursive tests in the single break case with breakpoint $\lambda_1^0 = .5$ and serially uncorrelated errors ($\rho = \theta = 0$).

5. Table 5 reports the probabilities of selecting the true number of breaks from the sequential procedure with serially uncorrelated errors ($\rho = \theta = 0$) and level $\eta = .10$.

6. Table 6 reports the empirical results based on monthly OECD inflation rates data over 1960:1-2008:6. Column (1) reports the country name; column (2) reports the estimate $\hat{m}$ obtained from applying the sequential algorithm in Section 5.4 with $\eta = .10$ and $A = 5$; column (3) reports the estimated break dates obtained by minimizing the unrestricted sum of squared residuals with $\hat{m}$ breaks; column (4) reports the outcome of the test for the null hypothesis of pure mean shifts; column (5) reports the OLS estimate of the largest sum of AR coefficients across the estimated regimes; column (6) reports Andrews and Guggenberger's (2014) 90% confidence band for the largest sum of AR coefficients; column (7) reports the model selected by the sequential algorithm; column (8) reports the model selected by the CT procedure; column (9) reports the p-value of the wild bootstrap ADF test of Cavaliere and Taylor (2009).

Appendix A: Proofs

For a $(d \times 1)$ vector v, $\|v\| = \left(\sum_{i=1}^{d} v_i^2\right)^{1/2}$ denotes the standard Euclidean norm while for a random variable v, $\|v\|_q = (E(|v|^q))^{1/q}$ denotes the $L_q (q \geq 1)$ norm. For a matrix B, $\|B\|$ denotes the Frobenius norm, i.e., $\|B\| = \sqrt{tr(B'B)}$ and $M_B = I - P_B$, $P_B = B(B'B)^{-1}B'$. Let P^* denote the bootstrap probability measure and E^* the expectation with respect to P^*. Define the following quantities: (i) $V(r) = diag(g^2(r)I_p, g(r))$; (ii) $D_T = diag(a_T^{-2}T^{-1}, a_T^{-1}T^{-1/2})$; (iii) For $i = 1, ..., k+1$, $Z_i = (z_{T_{i-1}+1}, ..., z_{T_i})'$ where $z_t = (y_{t-1}, 1)'$ for $t = T_{i-1}+1, ..., T_i$, $Z = (z_1, ..., z_T)'$, $Y_{-1} = (y_0, ..., y_{T-1})$, $\iota_{(T\times 1)} = (1, ..., 1)'$; (iv) $\bar{z}_i = (T_i - T_{i-1})^{-1}\sum_{t=T_{i-1}+1}^{T_i} z_t$ and $\bar{z}_{i,-1} = (T_i - T_{i-1})^{-1}\sum_{t=T_{i-1}+1}^{T_i} z_{t-1}$, $\bar{z} = T^{-1}\sum_{t=1}^{T} z_t$, $\bar{z}_{-1} = T^{-1}\sum_{t=1}^{T} z_{t-1}$.

We first state two lemmas that will be useful in developing the proofs of the results.

Lemma A.1 *[Xu, 2008] Suppose $\{y_t\}$ is generated by the AR(p) model*

$$y_t = \mu + \sum_{j=1}^{p} \theta_j (y_{t-j} - \mu) + e_t$$

where all roots of $\theta(L) = 1 - \sum_{j=1}^{p} \theta_j L^j$ are outside the unit circle and $\{e_t\}$ satisfies Assumptions A3-A5. Let $\widetilde{y}_{t-j} = y_{t-j} - \mu$, $y_{-p,t} = (\widetilde{y}_{t-1}, ..., \widetilde{y}_{t-p})'$ and $x_t = (y'_{-p,t}, 1)'$. Also, define the $[(p+1)\times(p+1)]$ matrix $\Upsilon_T = diag(T^{1/2}, ..., T^{1/2}, T^{1/2}a_T^{-1})$. Then

(a) $y_{-p,t} = \sum_{j=1}^{\infty} b_j e_{t-j}$ with $b_j = (\psi_{j-1}, ..., \psi_{j-p})$ if $j \geq 1$, $\psi_j = 0$ if $j < 0$, where $\theta(L)^{-1} = \sum_{j=0}^{\infty} \psi_j L^j$, $\psi_0 = 1$, $\sum_{j=0}^{\infty} j\left|\psi_j\right| < \infty$.

(b) $a_T^{-2}\Upsilon_T^{-1}\left(\sum_{t=1}^{T} x_t x_t'\right)\Upsilon_T^{-1} \xrightarrow{p} Q$ where $Q = \begin{pmatrix} \Omega \int g^2 & 0_{(p\times 1)} \\ 0_{(1\times p)} & 1 \end{pmatrix}$ and $\Omega = \sum_{j=1}^{\infty} b_j b_j'$.

(c) $a_T^{-2}\Upsilon_T^{-1}\sum_{t=1}^{T} x_t e_t \xrightarrow{w} \int V dB_{p+1}$, where $B_{p+1} = (B_p', B_1)'$ with B_p is a p-vector Brownian motion with covariance matrix Ω and B_1 is a standard Brownian motion independent of B_p.

Lemma A.2 *Suppose $\{y_t\}$ is generated by the AR(p) model with $\alpha = 1$:*

$$y_t = \alpha y_{t-1} + + \sum_{j=1}^{p-1} \pi_j \Delta y_{t-j} + e_t$$

where $\{\pi_j\}$ satisfies Assumption A2 and $\{e_t\}$ satisfies Assumptions A3-A5. Let $e = (e_1, ..., e_T)'$, $v_t = \Delta y_t$, $w_t = (\Delta y_{t-1}, ..., \Delta y_{t-p+1})'$, $W = (w_1, ..., w_T)'$, $W_j = (w_{T_{j-1}+1}, ..., w_{T_j})'[j = 1, ..., k+1]$ and $\Pi = (\pi_1, ..., \pi_{p-1})'$. Then

(a) $a_T^{-1}T^{-1/2}\sum_{t=1}^{[Tr]} e_t \xrightarrow{w} \int_0^r g dB_1 \equiv \widetilde{g}(1)B_{g,1}(r)$.

(b) $a_T^{-1}T^{-1/2}\sum_{t=1}^{[Tr]} v_t \xrightarrow{w} d(1)\int_0^r g dB_1 \equiv d(1)\widetilde{g}(1)B_{g,1}(r)$, if $d(1) \neq 0$, where $v_t = \sum_{j=0}^{\infty} d_j e_{t-j}$ with

$\sum_{j=0}^{\infty} j\,|d_j| < \infty$.

(c) $a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]} y_{t-1}e_t \xrightarrow{w} (1/2)d(1)[\widetilde{g}(1)^2 B_{g,1}^2(r) - \widetilde{g}(r)^2]$.

(d) $\left\|(a_T^{-2}T^{-1}W'W)^{-1}\right\| = O_p(1)$.

(e) $\|D_T Z_{2i}' W_{2i}\| = O_p(1)$.

(f) $\left\|a_T^{-2}T^{-1/2}W'e\right\| = O_p(1)$.

(g) $\left\|\left[a_T^{-2}T^{-1}W'W - a_T^{-2}T^{-1}\sum_{i=1}^{k/2} W_{2i}^{*\prime}Z_{2i}(Z_{2i}'Z_{2i})^{-1}Z_{2i}'W_{2i}^{*}\right]^{-1}\right\| = O_p(1)$.

Proof of Lemma A.2: (a) The result follows from Lemma 1 in Cavaliere and Taylor (2009).

(b) By Assumption A2, $\Delta y_t = v_t = \sum_{j=0}^{\infty} d_j e_{t-j}$ with $\sum_{j=0}^{\infty} j\,|d_j| < \infty$. Then, with the additional restriction $d(1) \neq 0$, the sequence $\{v_t\}$ satisfies Assumption 1' in Cavaliere and Taylor (2009) and hence by their Theorem 3, $a_T^{-1}T^{-1/2}\sum_{t=1}^{[Tr]} v_t \Rightarrow d(1)\int_0^r g dB_1$.

(c) Note that from the Beveridge-Nelson decomposition, we have $T^{-1}\sum_{t=1}^{[Tr]} y_{t-1}e_t = d(1)T^{-1}\sum_{t=1}^{[Tr]}\{\sum_{j=1}^{t-2} e_j\}e_t + o_p(1)$. Next, using the fact that

$$T^{-1}\sum_{t=1}^{[Tr]}\{\sum_{j=1}^{t-2} e_j\}e_t = (1/2)\left[\left(T^{-1/2}\sum_{t=1}^{[Tr]} e_t\right)^2 - T^{-1}\sum_{t=1}^{[Tr]} e_t^2\right] + o_p(1)$$

the result follows from (a) since $a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]} e_t^2 \xrightarrow{w} \int_0^r g(s)^2 \equiv \widetilde{g}(r)^2$.

(d) The entries in the matrix $a_T^{-2}T^{-1}W'W$ are of the form

$$T^{-1}\sum_{t=1}^{T} \Delta y_{t-j}\Delta y_{t-j'}, \quad j, j' \in \{1, ..., p-1\}$$

When $\alpha = 1$, $\{\Delta y_t\}$ is an $AR(p-1)$ process with all roots outside the unit circle. Then by Lemma A.1(b), $T^{-1}\sum_{t=1}^{T} \Delta y_{t-j}\Delta y_{t-j'} = O_p(1)$ and the result follows.

(e) We have $a_T^{-1}T^{-1/2}y_{[Tr]} = O_p(1)$ uniformly in $r \in [0,1]$. Hence, for a fixed $j \in \{1, ..., p-1\}$, $a_T^{-1}T^{-1/2}\sum_{t=T_{2i-1}+1}^{T_{2i}} \Delta y_{t-j} = a_T^{-1}T^{-1/2}y_{T_{2i}-j} - a_T^{-1}T^{-1/2}y_{T_{2i-1}+1-j} = O_p(1) - O_p(1) = O_p(1)$. Further, $a_T^{-2}T^{-1}\sum_{t=T_{2i-1}+1}^{T_{2i}} y_{t-1}\Delta y_{t-j} = \sum_{t=T_{2i-1}+1}^{T_{2i}}(a_T^{-1}T^{-1/2}y_{t-1})\,(a_T^{-1}T^{-1/2}\Delta y_{t-j}) = O_p(1)$. Hence, all entries in the matrix $D_T Z_{2i}' W_{2i}^*$ are $O_p(1)$ and the result follows.

(f) The result follows by applying Lemma A.1(c) to the sequence $\{\Delta y_t\}$.

(g) First, observe that $a_T^{-2}T^{-1}W'W = O_p(1)$ by Lemma A.1(b). Next, $a_T^{-2}T^{-1}\sum_{i=1}^{k/2} [W_{2i}'Z_{2i}D_T]\,[(a_T^2 D_T Z_{2i}'Z_{2i}D_T)^{-1}]\,[D_T Z_{2i}'W_{2i}] = T^{-1}\sum_{i=1}^{k/2} O_p(1).O_p(1).O_p(1) = O_p(T^{-1}) =$

$o_p(1)$ by (e) and Lemma A.1(b). ▲

Proof of Theorem 1: We prove the result for Model 1a and k even. The proofs for the other tests are very similar and omitted for brevity. Let $\widetilde{E}_i^*$ and $\hat{E}_i^*$ be the vector of residuals in the i-th regime under $H_0^{(1)}$ and $H_{a,k}^{(1)}$, respectively, for $i = 1, ..., k+1$. Denote $\hat{\gamma}_{2i} = (\hat{\alpha}_{2i} - 1, \hat{c}_{2i})'$, $i = 1, ..., k/2$, where $\hat{\alpha}_{2i}$ and $\hat{c}_{2i}$ are the OLS estimates obtained from regime $2i$. Then we have

$$\begin{aligned} \widetilde{E}_i^* &= \Delta Y_i - W_i \breve{\Pi}, && \text{for } i = 1, ..., k+1 \\ \hat{E}_{2i}^* &= \Delta Y_{2i} - W_{2i}\hat{\Pi} - Z_{2i}\hat{\gamma}_{2i}, && \text{for } i = 1, ..., k/2 \\ \hat{E}_{2i+1}^* &= \Delta Y_{2i+1} - W_{2i+1}\hat{\Pi}, && \text{for } i = 0, ..., k/2 \end{aligned} \tag{A.1}$$

where $\breve{\Pi} - \Pi = (W'W)^{-1} W'e$ under $H_0^{(1)}$. Further, $\hat{\Pi}$ and $\hat{\gamma}_{2i}$ satisfy the first order conditions

$$Z_{2i}'\hat{E}_{2i}^* = 0, \text{ for } i = 1, ..., k/2 \tag{A.2}$$

$$\sum_{i=1}^{k/2} W_{2i}\hat{E}_{2i}^* + \sum_{i=0}^{k/2} W_{2i+1}\hat{E}_{2i+1}^* = 0 \tag{A.3}$$

Under $H_0^{(1)}$, from $(A.3)$, we have $\hat{\Pi} - \Pi = (W'W)^{-1}(W'e - \sum_{i=1}^{k/2} W_{2i}'Z_{2i}\hat{\gamma}_{2i})$. Next, from $(A.2)$,

$$a_T^{-2} D_T^{-1}\hat{\gamma}_{2i} = (a_T^2 D_T Z_{2i}' Z_{2i} D_T)^{-1} \left[D_T Z_{2i}' W_{2i}(\Pi - \hat{\Pi}) + D_T Z_{2i}' E_{2i} \right] \tag{A.4}$$

for $i = 1, ..., k/2$, where $E_{2i} = (e_{T_{2i-1}+1}, ..., e_{T_{2i}})'$. Solving for $(\hat{\Pi} - \Pi)$ we get

$$\hat{\Pi} - \Pi = [W'W - \sum_{i=1}^{k/2} \{W_{2i}'Z_{2i}(Z_{2i}'Z_{2i})^{-1}Z_{2i}'W_{2i}\}]^{-1}[W'e - \sum_{i=1}^{k/2} \{W_{2i}'Z_{2i}(Z_{2i}'Z_{2i})^{-1}Z_{2i}'E_{2i}\}] \tag{A.5}$$

so that using Lemma A.2,

$$\begin{aligned} |\hat{\Pi} - \Pi|| &\leq \left\| [W'W - \sum_{i=1}^{k/2} \{W_{2i}'Z_{2i}(Z_{2i}'Z_{2i})^{-1}Z_{2i}'W_{2i}\}]^{-1} \right\| \times \\ &\quad \left[\|W'e\| + \sum_{i=1}^{k/2} \{\|W_{2i}'Z_{2i}D_T\| \left\|(D_T Z_{2i}'Z_{2i}D_T)^{-1}\right\| \|D_T Z_{2i}'E_{2i}\|\} \right] \\ &= [O_p(a_T^{-2}T^{-1})][O_p(a_T^2 T^{1/2}) + \sum_{i=1}^{k/2} O_p(1).O_p(a_T^2).O_p(1)] = O_p(T^{-1/2}) \end{aligned}$$

Also,

$$\begin{aligned} ||(D_T Z_{2i}'Z_{2i}D_T)^{-1}D_T Z_{2i}'W_{2i}(\Pi - \hat{\Pi})|| &\leq ||(D_T Z_{2i}'Z_{2i}D_T)^{-1}||||D_T Z_{2i}'W_{2i}||||(\Pi - \hat{\Pi})|| \\ &= O_p(a_T^2).O_p(1)O_p(T^{-1/2}) = O_p(a_T^2 T^{-1/2}) \end{aligned} \tag{A.6}$$

Using $(A.6)$ in $(A.4)$, we have

$$a_T^{-2} D_T^{-1} \hat{\gamma}_{2i} = (a_T^{-2} D_T Z_{2i}' Z_{2i} D_T)^{-1} D_T Z_{2i}' E_{2i} + o_p(1) \tag{A.7}$$

Next, $\hat{\Pi} - \breve{\Pi} = -(W'W)^{-1} \sum_{i=1}^{k/2} \{W_{2i}' Z_{2i} \hat{\gamma}_{2i}\}$ so that

$$\begin{aligned} ||\hat{\Pi} - \breve{\Pi}|| &\leq \left\|(W'W)^{-1}\right\| \sum_{i=1}^{k/2} \left\|W_{2i}' Z_{2i} D_T\right\| \left\|D_T^{-1} \hat{\gamma}_{2i}\right\| \\ &= O_p(a_T^{-2} T^{-1}). \sum_{i=1}^{k/2} O_p(1).O_p(a_T^2) = O_p(T^{-1}) \end{aligned} \tag{A.8}$$

We can write, from $(A.1)$, for $i = 1, ..., k/2$, $\widetilde{E}_{2i}^* = \hat{E}_{2i}^* + Z_{2i} \hat{\gamma}_{2i} + W_{2i}(\hat{\Pi} - \breve{\Pi})$ and for $i = 0, ..., k/2$, $\widetilde{E}_{2i+1}^* = \hat{E}_{2i+1}^* + W_{2i+1}(\hat{\Pi} - \breve{\Pi})$. Thus the numerator of the F statistic can be written as

$$\begin{aligned} SSR_0^{(1)} - SSR_{1a,k}^{(1)} &= \sum_{i=1}^{k/2} \{\widetilde{E}_{2i}^{*\prime} \widetilde{E}_{2i}^* - \hat{E}_{2i}^{*\prime} \hat{E}_{2i}^*\} + \sum_{i=0}^{k/2} \{\widetilde{E}_{2i+1}^{*\prime} \widetilde{E}_{2i+1}^* - \hat{E}_{2i+1}^{*\prime} \hat{E}_{2i+1}^*\} \\ &= \sum_{i=1}^{k/2} (D_T^{-1} \hat{\gamma}_{2i})' (D_T Z_{2i}' Z_{2i} D_T) D_T^{-1} \hat{\gamma}_{2i} + (\hat{\Pi} - \breve{\Pi})' \sum_{i=1}^{k/2} (W_{2i}' Z_{2i} D_T)(D_T^{-1} \hat{\gamma}_{2i}) \end{aligned} \tag{A.9}$$

where

$$\begin{aligned} ||(\hat{\Pi} - \breve{\Pi})' \sum_{i=1}^{k/2} (W_{2i}' Z_{2i} D_T)(D_T^{-1} \hat{\gamma}_{2i})|| &\leq ||\hat{\Pi} - \breve{\Pi}|| \sum_{i=1}^{k/2} \left\|(W_{2i}' Z_{2i} D_T)\right\| \left\|(D_T^{-1} \hat{\gamma}_{2i})\right\| \\ &= O_p(T^{-1}). \sum_{i=1}^{k/2} O_p(1).O_p(a_T^2) = O_p(a_T^2 T^{-1}) \end{aligned}$$

Then, using $(A.7)$ in $(A.9)$, we have

$$\begin{aligned} a_T^{-2}(SSR_0^{(1)} - SSR_{1a,k}^{(1)}) &= \sum_{i=1}^{k/2} \left\{E_{2i}' Z_{2i} D_T (a_T^2 D_T Z_{2i}' Z_{2i} D_T)^{-1} D_T Z_{2i}' E_{2i}\right\} + o_p(1) \\ &= \sum_{i=1}^{k/2} [\frac{\{a_T^{-2} T^{-1} \sum_{t=T_{2i-1}+1}^{T_{2i}} (y_{t-1} - \bar{y}_{2i,-1}) e_t\}^2}{a_T^{-2} T^{-2} \sum_{t=T_{2i-1}+1}^{T_{2i}} (y_{t-1} - \bar{y}_{2i,-1})^2} \\ &\quad + \frac{T}{T_{2i} - T_{2i-1}} \{a_T^{-1} T^{-1/2} \sum_{t=T_{2i-1}+1}^{T_{2i}} e_t\}^2] \end{aligned} \tag{A.10}$$

Using Lemma A.2(a),(c) in $(A.10)$, we have

$$a_T^{-2}(SSR_0^{(1)} - SSR_{1a,k}^{(1)}) \xrightarrow{w} \widetilde{g}(1)^2 \frac{1}{4k} \sum_{i=1}^{k/2} \left[\begin{array}{c} \frac{\left[\left\{B_{g,1}^{(2i)}(\lambda_{2i})\right\}^2 - \left\{B_{g,1}^{(2i)}(\lambda_{2i-1})\right\}^2 - \widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i})^2 - \widetilde{g}(\lambda_{2i-1})^2\right\}\right]^2}{\int_{\lambda_{2i-1}}^{\lambda_{2i}} \left[B_{g,1}^{(2i)}(r)\right]^2 dr} \\ + \frac{1}{\lambda_{2i} - \lambda_{2i-1}} \left\{B_{g,1}(\lambda_{2i}) - B_{g,1}(\lambda_{2i-1})\right\}^2 \end{array} \right]$$

Finally, noting that $T^{-1}SSR^{(1)}_{1a,k} \xrightarrow{p} \int_0^1 g^2 \equiv \widetilde{g}(1)^2$, the result follows.▲

Proof of Theorem 2: We can write

$$SSR_0^{(0)} - SSR_{1,k}^{(0)} = D^R(1, k+1) - \sum_{i=1}^{k+1} D^U(i,i)$$

where $D^U(i,j)$ [$D^U(i,j)$, resp.] is the sum of squared residuals from the unrestricted [restricted, resp.) using data from segments i to j (inclusively). Let $Y_{(-1)1,i}, Z_{1,i}, W_{1,i}$, and $E_{1,i}$ denote the vectors or matrices containing elements of Y_{-1}, Z, W, and e, respectively, belonging to the partition from segment 1 to segment i (inclusively), for $i = 1, ..., k+1$. Further, define $S_i = Z'_{1,i}E_{1,i}$, $H_i = Z'_{1,i}Z_{1,i}$, $K_i = Z'_{1,i}W_{1,i}$, $L_i = W'_{1,i}W_{1,i}$, and $M_j = W'_{1,i}E_{1,i}$ for $i = 1, ..., k+1$. Finally, let $A_T = (W'M_ZW)^{-1}W'M_Ze$ and $\bar{A}_T = (W'M_{\bar{Z}}W)^{-1}W'M_{\bar{Z}}e$ where $\bar{Z} = diag(Z_1, ..., Z_{k+1})$. Then, from equations (39) and (41) [pg. 73-74] in Bai and Perron (1998),

$$SSR_0^{(0)} - SSR_{1,k}^{(0)} = \sum_{i=1}^{k} F_{T,i} + D^R(1,1) - D^U(1,1)$$

where

$$\begin{aligned}
F_{T,i} &= \left[-S'_{i+1}H^{-1}_{i+1}S_{i+1} + S'_iH_i^{-1}S_i + (S_{i+1} - S_i)[H_{i+1} - H_i]^{-1}(S_{i+1} - S_i)\right] \\
&\quad + \left[2S'_{i+1}H^{-1}_{i+1}K_{i+1}A_T - 2S'_iH_i^{-1}K_iA_T - 2(S_{i+1} - S_i)'[H_{i+1} - H_i]^{-1}(K_{i+1} - K_i)\bar{A}_T\right] \\
&\quad + \left[2(M_{i+1} - M_i)'(\bar{A}_T - A_T) + (\bar{A}_T - A_T)'(L_{i+1} - L_i)(\bar{A}_T - A_T)\right] \\
&= T1 + T2 + T3 \qquad\qquad (A.11)
\end{aligned}$$

We now analyze each of the terms $T1$-$T3$ in $(A.11)$.

$T1$:

$$\begin{aligned}
T1 &= -S'_{i+1}H^{-1}_{i+1}S_{i+1} + S'_iH_i^{-1}S_i + (S_{i+1} - S_i)'[H_{i+1} - H_i]^{-1}(S_{i+1} - S_i) \\
&= -\lambda_{i+1}^{-1}\left[\left\{T^{-1/2}\sum_{t=1}^{T_{i+1}} e_t\right\}^2 + \left\{T^{-1}\sum_{t=1}^{T_{i+1}} \widetilde{y}_{t-1}^2\right\}^{-1}\left\{T^{-1/2}\sum_{t=1}^{T_{i+1}} \widetilde{y}_{t-1}e_t\right\}^2\right] \\
&\quad +\lambda_i^{-1}\left[\left\{T^{-1/2}\sum_{t=1}^{T_i} e_t\right\}^2 + \left\{T^{-1}\sum_{t=1}^{T_i} \widetilde{y}_{t-1}^2\right\}^{-1}\left\{T^{-1/2}\sum_{t=1}^{T_i} \widetilde{y}_{t-1}e_t\right\}^2\right] \\
&\quad +(\lambda_{i+1} - \lambda_i)^{-1}\left[\left\{T^{-1/2}\sum_{t=T_i+1}^{T_{i+1}} e_t\right\}^2 + \left\{T^{-1}\sum_{t=T_i+1}^{T_{i+1}} \widetilde{y}_{t-1}^2\right\}^{-1}\left\{T^{-1/2}\sum_{t=T_i+1}^{T_{i+1}} \widetilde{y}_{t-1}e_t\right\}^2\right] \\
&\quad +o_p(a_T^2)
\end{aligned}$$

using Lemma A.1(b) where $\widetilde{y}_{t-j} = y_{t-j} - \mu$. Then, from Lemma A.1, we have

$$\begin{aligned}
& a_T^{-2} T1 \\
& \xrightarrow{w} \widetilde{g}(1)^2 \left[-\lambda_{i+1}^{-1} B_{g,1}^2(\lambda_{i+1}) + \lambda_i^{-1} B_{g,1}^2(\lambda_i) + (\lambda_{i+1} - \lambda_i)^{-1} [B_{g,1}(\lambda_{i+1}) - B_{g,1}(\lambda_i)]^2 \right] \\
& + \widetilde{g}(1)^2 \left[\begin{array}{c} -\left\{ \widetilde{g}^2(\lambda_{i+1}) \right\}^{-1} B_{g,2}^2(\lambda_{i+1}) + \left\{ \widetilde{g}^2(\lambda_i) \right\}^{-1} B_{g,2}^2(\lambda_i) + \\ \left\{ \widetilde{g}^2(\lambda_{i+1}) - \widetilde{g}^2(\lambda_i) \right\}^{-1} [B_{g,2}(\lambda_{i+1}) - B_{g,2}(\lambda_i)]^2 \end{array} \right] \\
\equiv \;& \widetilde{g}(1)^2 \left[\frac{\left\{ \lambda_i B_{g,1}(\lambda_{i+1}) - \lambda_{i+1} B_{g,1}(\lambda_i) \right\}^2}{\lambda_i \lambda_{i+1} (\lambda_{i+1} - \lambda_i)} + \frac{\left\{ \widetilde{g}(\lambda_i)^2 B_{g,2}(\lambda_{i+1}) - \widetilde{g}(\lambda_{i+1})^2 B_{g,2}(\lambda_i) \right\}^2}{\widetilde{g}(\lambda_i)^2 \widetilde{g}(\lambda_{i+1})^2 \left\{ \widetilde{g}(\lambda_{i+1})^2 - \widetilde{g}(\lambda_i)^2 \right\}} \right]
\end{aligned}$$

$T2$:

$$\begin{aligned}
T2 \;=\; & 2(T^{-1/2} S_{i+1})'(T^{-1} H_{i+1})^{-1} T^{-1} K_{i+1} T^{1/2} A_T - 2(T^{-1/2} S_i)'(T^{-1} H_i)^{-1} T^{-1} K_i T^{1/2} A_T \\
& -2[T^{-1/2}(S_{i+1} - S_i)]'[T^{-1}(H_{i+1} - H_i)]^{-1} T^{-1}(K_{i+1} - K_i) T^{1/2} \bar{A}_T
\end{aligned}$$

Define $\widetilde{\Omega}_{p-1} = (\Omega_{11} - \Omega_{12}, \Omega_{12} - \Omega_{13}, ..., \Omega_{1(p-1)} - \Omega_{1p})'$, where Ω_{ij} is the (i,j) element of Ω defined in Lemma A.1. Then, using Lemma A.1(a)-(c), we have

$$\begin{aligned}
& a_T^{-2}(T^{-1/2} S_{i+1})'(T^{-1} H_{i+1})^{-1} T^{-1} K_{i+1} \xrightarrow{w} (1/\Omega_{11})^{1/2} B_{g,2}(\lambda_{i+1}) \widetilde{\Omega}'_{p-1} \\
& a_T^{-2}(T^{-1/2} S_i)'(T^{-1} H_i)^{-1} T^{-1} K_i \xrightarrow{w} (1/\Omega_{11})^{1/2} B_{g,2}(\lambda_i) \widetilde{\Omega}'_{p-1} \\
& a_T^{-2}[T^{-1/2}(S_{i+1} - S_i)]'[T^{-1}(H_{i+1} - H_i)]^{-1} T^{-1}(K_{i+1} - K_i) \xrightarrow{w} (1/\Omega_{11})^{1/2} [B_{g,2}(\lambda_{i+1}) - B_{g,2}(\lambda_i)] \widetilde{\Omega}'_{p-1}
\end{aligned}$$

Using Lemma A.1, it can further be shown that

$$\begin{aligned}
& T^{-1} W' P_Z W \xrightarrow{w} \Omega_{11}^{-1} \widetilde{\Omega}_{p-1} \widetilde{\Omega}'_{p-1} \widetilde{g}^2(1), \quad T^{-1} W' P_{\bar{Z}} W \xrightarrow{w} \Omega_{11}^{-1} \widetilde{\Omega}_{p-1} \widetilde{\Omega}'_{p-1} \widetilde{g}^2(1) \\
& T^{-1/2} W' P_Z e \xrightarrow{w} (1/\Omega_{11})^{1/2} \widetilde{\Omega}_{p-1} B_{g,2}(1), \quad T^{-1/2} W' P_{\bar{Z}} e \xrightarrow{w} (1/\Omega_{11})^{1/2} \widetilde{\Omega}_{p-1} B_{g,2}(1)
\end{aligned}$$

so that $\bar{A}_T - A_T \xrightarrow{p} 0$. Hence, $a_T^{-2} T2 = o_p(1)$.

$T3$:

$$a_T^{-2} T3 = \; a_T^{-2}[2(M_{i+1} - M_i)'(\bar{A}_T - A_T) + (\bar{A}_T - A_T)'(L_{i+1} - L_i)(\bar{A}_T - A_T)] \xrightarrow{p} 0$$

since $\bar{A}_T - A_T \xrightarrow{p} 0$. From $(A.11)$, we then get

$$a_T^{-2} F_{T,i} \xrightarrow{w} \widetilde{g}(1)^2 \left[\frac{\left\{ \lambda_i B_{g,1}(\lambda_{i+1}) - \lambda_{i+1} B_{g,1}(\lambda_i) \right\}^2}{\lambda_i \lambda_{i+1} (\lambda_{i+1} - \lambda_i)} + \frac{\left\{ \widetilde{g}(\lambda_i)^2 B_{g,2}(\lambda_{i+1}) - \widetilde{g}(\lambda_{i+1})^2 B_{g,2}(\lambda_i) \right\}^2}{\widetilde{g}(\lambda_i)^2 \widetilde{g}(\lambda_{i+1})^2 \left\{ \widetilde{g}(\lambda_{i+1})^2 - \widetilde{g}(\lambda_i)^2 \right\}} \right]$$

The result follows by noting that $[T - 2(k+1)]^{-1} a_T^{-2} SSR_{1,k}^{(0)} \xrightarrow{p} \widetilde{g}(1)^2$.▲

Proof of Theorem 3: We will prove the theorem for the bootstrap test based on $F_{1a}(\lambda, k)$ for k even. The bootstrap statistic is given by

$$F_{1a}^*(\lambda, k) = (T - k)(SSR_0^{*,(1)} - SSR_{1a,k}^{*,(1)})/[kSSR_{1a,k}^{*,(1)}]$$

where

$$SSR_0^{*,(1)} = \sum_{t=1}^{T} \left(y_t^{(1)} - y_{t-1}^{(1)} \right)^2 \tag{A.12}$$

$$\begin{aligned} SSR_{1a,k}^{*,(1)} &= \sum_{i=1}^{k/2} \sum_{t=T_{2i-1}+1}^{T_{2i}} \left(y_t^{(1)} - \bar{y}_{2i}^{(1)} - \hat{\alpha}_{2i}^{(1)} (y_{t-1}^{(1)} - \bar{y}_{2i,-1}^{(1)}) \right)^2 \\ &\quad + \sum_{i=0}^{k/2} \sum_{t=T_{2i}+1}^{T_{2i+1}} \left(y_t^{(1)} - y_{t-1}^{(1)} \right)^2 \end{aligned} \tag{A.13}$$

In $(A.13)$, $\hat{\alpha}_{2i}^{(1)}$ denotes the slope estimate from an OLS regression of $y_t^{(1)}$ on a constant and $y_{t-1}^{(1)}$ $[t = T_{2i-1}+1, ..., T_{2i}; i = 1, ..., k/2]$. Since $y_t^{(1)} = y_{t-1}^{(1)} + e_t^{(1)}$ for $t \le T$, we have

$$\begin{aligned} & a_T^{-2}(SSR_0^{*,(1)} - SSR_{1a,k}^{*,(1)}) \\ = & \sum_{i=1}^{k/2} \left[(T_{2i} - T_{2i-1}) \left[a_T^{-1} \bar{e}_{2i}^{(1)} \right]^2 + \frac{\left[a_T^{-2} T^{-1} \sum_{t=T_{2i-1}+1}^{T_{2i}} \left\{ (y_{t-1}^{(1)} - \bar{y}_{2i,-1}^{(1)}) e_t^{(1)} \right\} \right]^2}{a_T^{-2} T^{-2} \sum_{t=T_{2i-1}+1}^{T_{2i}} \left(y_{t-1}^{(1)} - \bar{y}_{2i,-1}^{(1)} \right)^2} \right] \end{aligned} \tag{A.14}$$

Next, we establish an invariance principle for the sequence $\left\{ a_T^{-1} e_t^{(1)}; t = 1, ..., T \right\}$. To this end, let $\mathcal{F}_t^*$ be the σ-field generated by $\{v_s; s \le t\}$. Since $e_t^{(1)} = \breve{e}_t v_t$, $\{ a_T^{-1} e_t^{(1)}, \mathcal{F}_t^* \}$ is a martingale difference array. Further, uniformly over $r \in [0, 1]$,

$$a_T^{-2} T^{-1} \sum_{t=1}^{[Tr]} \left[e_t^{(1)} \right]^2 - a_T^{-2} T^{-1} \sum_{t=1}^{[Tr]} \breve{e}_t^2 \xrightarrow{p^*} 0$$

since

$$E^* \left\{ a_T^{-2} T^{-1} \sum_{t=1}^{[Tr]} \left(\left[e_t^{(1)} \right]^2 - \breve{e}_t^2 \right) \right\}^2 = E^* \left\{ a_T^{-2} T^{-1} \sum_{t=1}^{[Tr]} \left(\breve{e}_{tt}^2 (v_t^2 - 1 \right) \right\}^2 \le C T^{-2} \sum_{t=1}^{[Tr]} (a_T^{-1} \widetilde{e}_t)^4 = o_p(1)$$

(where C is a positive constant), using the fact that under $H_0^{(1)}$, $a_T^{-1} \breve{e}_t = a_T^{-1} e_t + a_T^{-1} w_t' (\Pi - \breve{\Pi}) = a_T^{-1} e_t + O_p(T^{-1/2})$. Also, $a_T^{-2} T^{-1} \sum_{t=1t}^{[Tr]} \breve{e}_t^2 \xrightarrow{w} \int_0^r g^2$ uniformly over $r \in [0, 1]$ (by Lemma 2 in Cavaliere and Taylor, 2008). Then, applying Theorem 2.1 in Hansen (1992) with $S_T(.) = T^{-1/2} \sum_{t=1}^{[T.]} v_t$, we get

$$T^{-1/2} \sum_{t=1}^{[Tr]} a_T^{-1} e_t^{(1)} = a_T^{-1} \int_0^r \breve{e}_{[Ts]} dS_T(s) \xrightarrow{w}_p \int_0^r g(s) dB_1(s) = \widetilde{g}(1) B_{g,1}(r) \tag{A.15}$$

Utilizing $(A.15)$, we have

$$
\begin{aligned}
&a_T^{-2}T^{-2}\sum_{t=T_{2i-1}+1}^{T_{2i}}\left(y_{t-1}^{(1)}-\bar{y}_{2i,-1}^{(1)}\right)^2 \xrightarrow{w}_p \widetilde{g}(1)^2\int_{\lambda_{2i-1}}^{\lambda_{2i}}\left[B_{g,1}^{(2i)}(s)\right]^2 \\
&a_T^{-2}T^{-1}\sum_{t=T_{2i-1}+1}^{T_{2i}}\left\{(y_{t-1}^{(1)}-\bar{y}_{2i,-1}^{(1)})e_t^{(1)}\right\} \xrightarrow{w}_p (1/2)\widetilde{g}(1)^2\left[\begin{array}{c}\left\{B_{g,1}^{(2i)}(\lambda_{2i})\right\}^2-\left\{B_{g,1}^{(2i)}(\lambda_{2i-1})\right\}^2 \\ -\widetilde{g}(1)^{-2}\left\{\widetilde{g}(\lambda_{2i})^2-\widetilde{g}(\lambda_{2i-1})^2\right\}\end{array}\right] \\
&a_T^{-1}T^{1/2}\bar{e}_{2i}^{(1)} \xrightarrow{w}_p (\lambda_{2i}-\lambda_{2i-1})^{-1}\widetilde{g}(1)[B_{g,1}(\lambda_{2i})-B_{g,1}(\lambda_{2i-1})]
\end{aligned}
\tag{A.16}
$$

Substituting $(A.16)$ in $(A.14)$ and noting that $(T-k)^{-1}SSR_{1a,k}^{*,(1)} \xrightarrow{p} \widetilde{g}(1)^2$, we get $F_{1a}^*(\lambda,k) \xrightarrow{w}_p F_{1a}^0(\lambda,k)$, where $F_{1a}^0(\lambda,k)$ is the weak limit of $F_{1a}(\lambda,k)$ as stated in Theorem 1. The rest of the proof follows from the proof of Theorem 5 in Hansen (2000).▲

Proof of Theorem 4: The bootstrap BP test for k breaks is given by

$$G_1^*(k)=[T-2(k+1)](SSR_0^{*,(0)}-SSR_{1,k}^{*,(0)})/[kSSR_{1,k}^{*,(0)}]$$

where

$$SSR_0^{*,(0)} = \sum_{t=1}^{T}\left(e_t^{(0)}-\bar{e}^{(0)}-\widetilde{\alpha}^{(0)}(e_{t-1}^{(0)}-\bar{e}_{-1}^{(0)})\right)^2 \tag{A.17}$$

$$SSR_{1,k}^{*,(0)} = \sum_{i=1}^{k+1}\sum_{t=T_{i-1}+1}^{T_i}\left(e_t^{(0)}-\bar{e}_i^{(0)}-\hat{\alpha}_i^{(0)}(e_{t-1}^{(0)}-\bar{e}_{i,-1}^{(0)})\right)^2 \tag{A.18}$$

In $(A.17)$, $\widetilde{\alpha}^{(0)}$ denotes the slope estimate from an OLS regression of $e_t^{(0)}$ on a constant and $e_{t-1}^{(0)}$ $[t=1,...,T]$. In $(A.18)$, $\hat{\alpha}_i^{(0)}$ denotes the slope estimate from an OLS regression of $e_t^{(0)}$ on a constant and $e_{t-1}^{(0)}$ $[t=T_{i-1}+1,...,T_i]$. We can write, after some algebra,

$$
\begin{aligned}
a_T^{-2}(SSR_0^{*,(0)}-SSR_{1,k}^{*,(0)}) &= -T\left[a_T^{-1}\bar{e}^{(0)}\right]^2-\frac{\left[a_T^{-2}T^{-1/2}\sum_{t=1}^{T}\left\{(e_{t-1}^{(0)}-\bar{e}_{i,-1}^{(0)})e_t^{(0)}\right\}\right]^2}{a_T^{-2}T^{-1}\sum_{t=1}^{T}\left(e_{t-1}^{(0)}-\bar{e}_{i,-1}^{(0)}\right)^2} \\
&+\sum_{i=1}^{k+1}\left[(T_i-T_{i-1})\left[a_T^{-1}\bar{e}_i^{(0)}\right]^2+\frac{\left[a_T^{-2}T^{-1/2}\sum_{t=T_{i-1}+1}^{T_i}\left\{(e_{t-1}^{(0)}-\bar{e}_{i,-1}^{(0)})e_t^{(0)}\right\}\right]^2}{a_T^{-2}T^{-1}\sum_{t=T_{i-1}+1}^{T_i}\left(e_{t-1}^{(0)}-\bar{e}_{i,-1}^{(0)}\right)^2}\right]
\end{aligned}
\tag{A.19}
$$

Next, we establish an invariance principle for the sequence $\left\{a_T^{-1}e_t^{(0)}; t=1,...,T\right\}$. In particular, we will show that for $r \in [0,1]$,

$$T^{-1/2}\sum_{t=1}^{[Tr]} a_T^{-1}e_t^{(0)} \xrightarrow{w}_p \int_0^r g(s)dB_1(s) \tag{A.20}$$

To this end, let $\mathcal{F}_t^*$ be the σ-field generated by $\{v_s; s \le t\}$. Since $e_t^{(0)} = \widetilde{e}_t v_t$, $\{\ a_T^{-1}e_t^{(0)}, \mathcal{F}_t^*\}$ is a martingale difference array. Further, uniformly over $r \in [0,1]$,

$$a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\left[e_t^{(0)}\right]^2 - a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\widetilde{e}_t^2 \xrightarrow{p^*} 0$$

since

$$E^*\left\{a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\left(\left[e_t^{(0)}\right]^2 - \widetilde{e}_t^2\right)\right\}^2 = E^*\left\{a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\left(\widetilde{e}_t^2(v_t^2-1\right)\right\}^2 \le CT^{-2}\sum_{t=1}^{[Tr]}(a_T^{-1}\widetilde{e}_t)^4 = o_p(1)$$

using the fact that $a_T^{-1}\widetilde{e}_t = a_T^{-1}e_t + O_p(T^{-1/2})$ [eq. (A.7) in Xu, 2008]. Also, $a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\widetilde{e}_t^2 \xrightarrow{p} \int_0^r g^2$. Then, again applying Theorem 2.1 in Hansen (1992) with $S_T(.) = T^{-1/2}\sum_{t=1}^{[T.]} v_t$, we get

$$T^{-1/2}\sum_{t=1}^{[Tr]} a_T^{-1}e_t^{(0)} = a_T^{-1}\int_0^r \widetilde{e}_{[Ts]}dS_T(s) \xrightarrow{w}_p \int_0^r g(s)dB_1(s)$$

Noting that $\{\ a_T^{-2}e_t^{(0)}e_{t-1}^{(0)}, \mathcal{F}_t^*\}$ is a martingale difference array, we can show, using similar arguments as above, that

$$T^{-1/2}\sum_{t=1}^{[Tr]} a_T^{-2}e_t^{(0)}e_{t-1}^{(0)} \xrightarrow{w}_p \int_0^r g^2(s)dB_2(s) \tag{A.21}$$

for $r \in [0,1]$, where $B_2(.)$ is independent of $B_1(.)$. Finally, since $a_T^{-2}T^{-1}SSR_{1,k}^{*,(0)} \xrightarrow{p} \widetilde{g}(1)^2$, $G_1^*(\lambda,k) \xrightarrow{w}_p G_1^0(\lambda,k)$ using $(A.20)$ and $(A.21)$ in $(A.19)$, where $G_1^0(\lambda,k)$ is the weak limit of $G_1(\lambda,k)$ as defined in Theorem 2. Hence, following the proof of Theorem 5 in Hansen (2000), $p_{k,G_1}^b \xrightarrow{w} U[0,1]$, $p_{UDmax}^b \xrightarrow{w} U[0,1]$.▲

Proof of Theorem 5: We will prove $p_{m,W_1}^b \xrightarrow{p} 0$ and $p_{m,G_1}^b \xrightarrow{p} 0$ under $H_{a,m}^{(1)}$ with m even. Consequently, $p_{Wmax}^b \xrightarrow{p} 0$ and $p_{UDmax}^b \xrightarrow{p} 0$. The proofs for the alternatives $H_{b,m}^{(1)}$ and $H_{1,m}^{(0)}$ can be established using similar arguments. The proof proceeds in two steps: (i) We first show that the bootstrap counterparts $F_{1a}^*(m)$, $F_{1b}^*(m)$ and $G_1^*(m)$ of $F_{1a}(m)$, $F_{1b}(m)$ and $G_1(m)$, respectively, are each $O_p(1)$ under $H_{a,m}^{(1)}$; (ii) $F_{1a}(m)$ [hence $W_1(m)$] and $G_1(m)$ both diverge with T.

For (i), first note that for $s \in [0,1]$,

$$a_T^{-1}\breve{e}_{[Ts]} = a_T^{-1}e_{Ts]} + a_T^{-1}\sum_{i=1}^{m/2}(\alpha_{2i}-1)h_{[Ts]-1}I([Ts] \in [T_{2i-1}^0+1, T_{2i}^0]) + O_p(T^{-1/2})$$

so that

$$a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\breve{e}_t^2 = a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}e_t^2 + a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\sum_{i=1}^{m/2}(\alpha_{2i}-1)^2h_{t-1}^2I(t \in [T_{2i-1}^0+1, T_{2i}^0]) + o_p(1)$$

$$\xrightarrow{p} \int_0^r g^2 + \sum_{i=1}^{m/2}(\alpha_{2i}-1)^2(\Omega_{11}^{(2i)})\int_0^r g(s)^2 I(s \in [\lambda_{2i-1}^0, \lambda_{2i}^0])ds = \breve{V}(r) \text{ (say)} \tag{A.22}$$

where $a_T^{-2}T^{-1}\sum_{t=T_{2i-1}^0+1}^{[Ts]}h_{t-1}^2 \xrightarrow{p} \Omega_{11}^{(2i)}\int_{\lambda_{2i-1}^0}^s g^2$ if $s \in [\lambda_{2i-1}^0, \lambda_{2i}^0]$ and $\Omega_{11}^{(2i)}$ is the $(1,1)$ element of $\Omega^{(2i)}$, where $\Omega^{(2i)}$ is defined analogously to Ω in Lemma A.1 but is now specific to regime $2i$. Therefore, we have for $r \in [0,1]$,

$$a_T^{-1}T^{-1/2}\sum_{t=1}^{[Tr]}e_t^{(1)} = a_T^{-1}\int_0^r \breve{e}_{[Ts]}dS_T(s) \xrightarrow{w}_p \int_0^r g_1(s)dB_1(s) = \breve{B}_{g,1}(r) \text{ (say)} \tag{A.23}$$

where $g_1(s) = g(s)[1+\sum_{i=1}^{m/2}(\alpha_{2i}-1)^2(\Omega_{11}^{(2i)})I(s \in [\lambda_{2i-1}^0, \lambda_{2i}^0])ds]^{1/2}$. Note that $\int_0^r g_1(s)^2 = \breve{V}(r)$. Then, the results stated in $(A.16)$ all hold with $B_{g,1}(.)$ replaced by $\breve{B}_{g,1}(.)$. Further, $(T-k)^{-1}SSR_{1a,k}^{*,(1)} \xrightarrow{w}_p \int_0^1 g_1(s)^2 = \breve{V}(1)$. Thus, $F_{1a}^*(m) = O_p(1)$. Entirely analogous arguments can be used to establish $F_{1b}^*(m) = O_p(1)$.

Next, we show that $G_1^*(m)$ is stochastically bounded under $H_{a,m}^{(1)}$. First, note that we can write

$$\widetilde{e}_t = y_t - \bar{y} - \widetilde{\alpha}(y_{t-1}-\bar{y}_{-1}) - (w_t - \bar{w})'\widetilde{\Pi}$$

where $T(\widetilde{\alpha}-1) = O_p(1)$ since $H_{a,m}^{(1)}$ involves a mix of $I(1)$ and $I(0)$ regimes. Further,

$$\begin{aligned}\bar{y} - \widetilde{\alpha}\bar{y}_{-1} &= \bar{y} - \bar{y}_{-1} - (\widetilde{\alpha}-1)\bar{y}_{-1} = T^{-1}(y_T - y_0) - (\widetilde{\alpha}-1)\bar{y}_{-1} \\ &= O_p(a_T T^{-1/2}) - O_p(a_T T^{-1/2}) = O_p(a_T T^{-1/2})\end{aligned}$$

Thus, in an $I(1)$ regime, i.e., $t \in [T_{2i}+1, ..., T_{2i+1}],\ i = 0, ..., m/2$, we have

$$\begin{aligned}a_T^{-1}\widetilde{e}_t &= a_T^{-1}e_t + a_T^{-1}(1-\widetilde{\alpha})y_{t-1} + O_p(T^{-1/2}) = a_T^{-1}e_t + O_p(T^{-1})O_p(T^{1/2}) + O_p(T^{-1/2}) \\ &= a_T^{-1}e_t + O_p(T^{-1/2})\end{aligned} \tag{A.24}$$

In an $I(0)$ regime, i.e., $t \in [T_{2i-1}+1, ..., T_{2i}],\ i = 1, ..., m/2$, we have

$$a_T^{-1}\widetilde{e}_t = a_T^{-1}e_t + (\alpha_{2i}-1)a_T^{-1}h_{t-1} + O_p(T^{-1/2}) \tag{A.25}$$

Combining $(A.24)$ and $(A.25)$, we can write for $t = 1, ..., T$,

$$a_T^{-1}\widetilde{e}_t = a_T^{-1}e_t + \sum_{i=1}^{m/2}(\alpha_{2i} - 1)a_T^{-1}h_{t-1}I(t \in [T_{2i-1}^0 + 1, ..., T_{2i}^0])$$

so that for $r \in [0, 1]$,so that

$$\begin{aligned} a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\widetilde{e}_t^2 &= a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}e_t^2 + a_T^{-2}T^{-1}\sum_{t=1}^{[Tr]}\sum_{i=1}^{m/2}(\alpha_{2i} - 1)^2 h_{t-1}^2 I(t \in [T_{2i-1}^0 + 1, T_{2i}^0]) + o_p(1) \\ &\xrightarrow{p} \breve{V}(r) \end{aligned}$$

where $\breve{V}(r)$ is defined in $(A.22)$. Hence, $a_T^{-1}T^{-1/2}\sum_{t=1}^{[Tr]}e_t^{(0)} = a_T^{-1}\int_0^r \widetilde{e}_{[Ts]}dS_T(s) \xrightarrow{w}_p \breve{B}_{g,1}(r)$ and the limits in $(A.20)$ and $(A.21)$ now hold with $g(.)$ replaced by $g_1(.)$. Also, $a_T^{-2}T^{-1}SSR_{1,k}^{*,(0)} \xrightarrow{p} \breve{V}(1)$. Thus, $G_1^*(m) = O_p(1)$.

To show (ii), note that since $\lambda^0 \in \Lambda_\epsilon^m$ and $F_{1a}(m) = \sup_{\lambda \in \Lambda_\epsilon^m} F_{1a}(\lambda, m)$, it is sufficient to show that $F_{1a}(\lambda^0, m) = O_p(T)$. Define

$$\begin{aligned} \breve{\Pi} &= (\sum_{t=1}^{T} w_t w_t')^{-1}\sum_{t=1}^{T} w_t \Delta y_t \\ \widetilde{\mu}_{2i} &= \mu_{2i} + y_{T_{2i-1}^0} - \mu_{2i-1}, \; i = 1, ..., m/2 \end{aligned}$$

Then $h_{t-1} = y_{t-1} - \widetilde{\mu}_{2i}, \; t \in [T_{2i-1}^0 + 1, T_{2i}^0]$. We can write

$$\begin{aligned} SSR_0^{(1)} &= \sum_{t=1}^{T}(\Delta y_t - w_t'\breve{\Pi})^2 = \sum_{i=0}^{m/2}\sum_{t=T_{2i}^0+1}^{T_{2i+1}^0}\left\{w_t'(\Pi - \breve{\Pi}) + e_t\right\}^2 \\ &\quad + \sum_{i=1}^{m/2}\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}\left\{(\alpha_{2i} - 1)h_{t-1} + w_t'(\Pi - \breve{\Pi}) + e_t\right\}^2 \\ &= \sum_{t=1}^{T}e_t^2 + \sum_{i=1}^{m/2}(\alpha_{2i} - 1)^2\left(\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}h_{t-1}^2\right) + 2(\Pi - \breve{\Pi})'\sum_{i=1}^{m/2}(\alpha_{2i} - 1)\left(\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}h_{t-1}w_t\right) \\ &\quad + (\Pi - \breve{\Pi})'(W'W)(\Pi - \breve{\Pi}) + 2(\Pi - \breve{\Pi})'W'e + 2\sum_{i=1}^{m/2}(\alpha_{2i} - 1)\left(\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}h_{t-1}e_t\right) \end{aligned}$$

Let $\bar{Z}^{(1)} = diag(\widetilde{Z}_1^{(1)}, ..., \widetilde{Z}_{m+1}^{(1)})$, where $\widetilde{Z}_i^{(1)}$ is the first column of $\widetilde{Z}_i = (h_{T_{2i-1}^0}, ..., h_{T_{2i}^0-1})$ and the $[(m+1) \times 1]$ vector $\gamma_1 = (0, \alpha_2 - 1, 0, \alpha_4 - 1, .., 0)'$. Noting that $\breve{\Pi} - \Pi = (W'W)^{-1}W'e + (W'W)^{-1}\sum_{i=1}^{m/2}(\alpha_{2i} - 1)\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}h_{t-1}w_t$, we can write

$$SSR_0^{(1)} = \sum_{t=1}^{T}e_t^2 + \gamma_1'\bar{Z}^{(1)\prime}M_W\bar{Z}^{(1)}\gamma_1 - e'W(W'W)^{-1}W'e + 2\gamma_1'\bar{Z}^{(1)\prime}e \tag{A.26}$$

Now we use the following facts: (i) $e'W(W'W)^{-1}W'e = a_T^2[a_T^{-2}T^{-1/2}e'W(a_T^{-2}T^{-1}W'W)^{-1}\cdot a_T^{-2}T^{-1/2}W'e] = a_T^2 O_p(1) = O_p(a_T^2)$; (ii) $\gamma_1'\bar{Z}^{(1)\prime}e = O_p(a_T^2 T^{1/2})$. Thus, from $(A.26)$, we get

$$a_T^{-2}SSR_0^{(1)} = a_T^{-2}\sum_{t=1}^{T} e_t^2 + a_T^{-2}\gamma_1'\bar{Z}^{(1)\prime}M_W\bar{Z}^{(1)}\gamma_1 + O_p(T^{1/2})$$

Next, we have (with $\hat{\Pi}$ denoting the estimate of Π under the alternative model),

$$SSR_{1a,m}^{(1)} = \sum_{i=0}^{m/2}\sum_{t=T_{2i}^0+1}^{T_{2i+1}^0}\left\{w_t'(\Pi-\hat{\Pi})+e_t\right\}^2 + \sum_{i=1}^{m/2}\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0}\left\{\begin{array}{c}(\alpha_{2i}-\hat{\alpha}_{2i})(y_{t-1}-\bar{y}_{2i,-1})\\ +(w_t-\bar{w}_{2i})'(\Pi-\hat{\Pi})+e_t\end{array}\right\}^2$$

Then, noting that

$$\begin{aligned}\hat{\Pi}-\Pi &= (W'W)^{-1}[\sum_{i=1}^{m/2}(\alpha_{2i}-\hat{\alpha}_{2i})\sum_{t=T_{2i-1}^0+1}^{T_{2i}^0} w_t(y_{t-1}-\bar{y}_{2i,-1}) + W'e]\\ &= [O_p(a_T^{-2}T^{-1})][O_p(a_T^2T^{1/2}) + O_p(a_T^2T^{1/2})] = O_p(T^{-1/2})\end{aligned}$$

we can show, after some simplification,

$$a_T^{-2}SSR_{1a,m}^{(1)} = a_T^{-2}\sum_{t=1}^{T}e_t^2 + O_p(1) \tag{A.27}$$

Combining $(A.26)$ and $(A.27)$,

$$a_T^{-2}(SSR_0^{(1)} - SSR_{1a,m}^{(1)}) = a_T^{-2}\gamma_1'\bar{Z}^{(1)\prime}M_W\bar{Z}^{(1)}\gamma_1 + O_p(T^{1/2}) \tag{A.28}$$

Now, since regime $2i$ $(i=1,...,m/2)$ is $I(0)$, $a_T^{-2}\gamma_1'\bar{Z}^{(1)\prime}M_W\bar{Z}^{(1)}\gamma_1 = a_T^{-2}O_p(a_T^2T) = O_p(T)$. Since this term is positive and dominant in $(A.28)$, $F_{1a}(\lambda^0, m)$ diverges to positive infinity at rate T. Entirely analogous arguments can be used to show the divergence of $G_1(m)$ at rate T. The details are omitted.▲

Proof of Theorem 6: To prove this result, it is sufficient to show that

$$\lim_{T\to\infty} P(\min_{1\le i\le l+1}\{p_i^*\} < \eta_{l+1}) \le \eta \tag{A.29}$$

as the rest of the proof follows the same arguments as in the proof of Theorem 2 in Kejriwal (2018). First, note that

$$\begin{aligned}P(\min_{1\le i\le l+1}\{p_i^*\} &< \eta_{l+1}) = 1 - P(\min_{1\le i\le l+1}\{p_i^*\} \ge \eta_{l+1})\\ &= 1 - \Pi_{i=1}^{l+1}\left[P(p_i^* \ge \eta_{l+1})\right]\\ &= 1 - \Pi_{i=1}^{l+1}\left[1 - P(\{p_{1,W_1}^{*,(i)} < \eta_{l+1}\}\cap\{p_{1,G_1}^{*,(i)} < \eta_{l+1}\})\right]\end{aligned} \tag{A.30}$$

where the second equality follows from the independence of the test statistics across segments and the third from the fact that $p_i^* = \max\{p_{1,W_1}^{*,(i)}, p_{1,G_1}^{*,(i)}\}$. Next, from Theorems 3 and 4, it follows that under the null hypothesis of l breaks, we have for any segment $i \in \{1, ..., l+1\}$,

$$\begin{aligned} P(\{p_{1,W_1}^{*,(i)} &< \eta_{l+1}\} \cap \{p_{1,G_1}^{*,(i)} < \eta_{l+1}\}) \leq P(\{p_{1,W_1}^{*,(i)} < \eta_{l+1}\}) \to \eta_{l+1} \text{ if } i \text{ is } I(1) \\ P(\{p_{1,W_1}^{*,(i)} &< \eta_{l+1}\} \cap \{p_{1,G_1}^{*,(i)} < \eta_{l+1}\}) \leq P(\{p_{1,G_1}^{*,(i)} < \eta_{l+1}\}) \to \eta_{l+1} \text{ if } i \text{ is } I(0) \quad \text{(A.31)} \end{aligned}$$

Thus, using $(A.31)$ in $(A.30)$, we have

$$\lim_{T\to\infty} P(\min_{1\leq i\leq l+1}\{p_i^*\} < \eta_{l+1}) \leq 1 - \left[1 - \eta_{l+1}\right]^{l+1} = \eta$$

which proves $(A.29)$.▲

Proof of Theorem 7: The proof uses arguments very similar to those used in proving Theorems 3-6 and is hence omitted.▲

Appendix B: Additional Monte Carlo Results

Notes to Tables

1. Table B-1 reports the empirical size of asymptotic tests with nominal size 5%. The tests $\mathcal{K}_1, \mathcal{K}_1', \mathcal{K}_4$ are the ratio-based tests of Kim (2000) and Busetti and Taylor (2004) and H_1, H_2, H_{max} are the tests of Kejriwal et al. (2013).

2. Table B-2 reports the size-adjusted power of 5% bootstrap tests under different change points and volatility intensity in the single break case with breakpoint $\lambda_1^0 = .5$ and serially uncorrelated errors ($\rho = \theta = 0$) with abrupt volatility change (Model 1).

3. Table B-3 reports the size-adjusted power of 5% bootstrap tests in the single break case with breakpoint $\lambda_1^0 = .5$ and AR(1) errors ($\rho = .5, \theta = 0$).

4. Table B-4 reports the size-adjusted power of 5% bootstrap tests in the single break case with breakpoint $\lambda_1^0 = .5$ and MA(1) errors ($\rho = 0, \theta = .5$).

5. Table B-5 reports the size-adjusted power of 5% bootstrap tests in the two breaks case with breakpoint vector $(\lambda_1^0, \lambda_2^0) = (.3, .8)$ and AR(1) errors ($\rho = .5, \theta = 0$).

6. Table B-6 reports the size-adjusted power of 5% bootstrap tests in the two breaks case with breakpoint vector $(\lambda_1^0, \lambda_2^0) = (.3, .8)$ and MA(1) errors ($\rho = 0, \theta = .5$).

7. Table B-7 reports the empirical power of 5% bootstrap recursive and the proposed non-recursive BP tests in the single break case with breakpoint $\lambda_1^0 = .5$ and AR(1) errors ($\rho = .5, \theta = 0$).

8. Table B-8 reports the empirical power of 5% bootstrap recursive and the proposed non-recursive BP tests in the single break case with breakpoint $\lambda_1^0 = .5$ and MA(1) errors ($\rho = 0, \theta = .5$).

9. Table B-9 reports the probabilities of selecting the true number of breaks from the sequential procedure under different abrupt volatility break points and intensities in the two breaks case with breakpoint vector $(\lambda_1^0, \lambda_2^0) = (.3, .8)$, serially uncorrelated errors ($\rho = \theta = 0$) and level $\eta = .10$.

10. Table B-10 reports the regime-wise parameter estimates and wild bootstrap p-values of the ADF test for the OECD countries with persistence breaks.

Table B-1: Empirical size of asymptotic tests, [$m = 0$, 5% nominal size]

T	(ρ,θ)	Test	$\alpha = 1$							$\alpha = 0.5$							$\alpha = 0.7$						
			Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
200	(0,0)	$\mathcal{K}_1$	.55	.87	.31	.77	.37	.57	.60	.08	.66	.01	.40	.01	.39	.21	.10	.68	.02	.45	.02	.41	.25
		$\mathcal{K}_1'$	.55	.32	.84	.37	.76	.56	.57	.06	.01	.60	.01	.36	.35	.16	.08	.03	.62	.02	.40	.37	.17
		$\mathcal{K}_4$	.76	.87	.85	.82	.81	.86	.79	.09	.56	.52	.32	.27	.65	.25	.13	.61	.53	.37	.31	.66	.30
		H_1^*	.06	.39	.21	.26	.11	.49	.32	.05	.15	.14	.10	.08	.28	.15	.07	.18	.15	.14	.10	.31	.19
		H_2^*	.04	.33	.35	.07	.14	.41	.38	.03	.15	.14	.07	.05	.23	.15	.05	.21	.19	.09	.07	.25	.20
		H_{max}^*	.05	.44	.37	.17	.16	.56	.46	.04	.19	.16	.10	.08	.31	.17	.06	.24	.19	.14	.09	.34	.22
	(0.5,0)	$\mathcal{K}_1$	.56	.86	.32	.79	.37	.58	.60	.10	.68	.02	.44	.01	.41	.24	.13	.70	.04	.49	.03	.42	.27
		$\mathcal{K}_1'$	.54	.30	.84	.36	.75	.56	.56	.07	.02	.60	.02	.37	.35	.16	.09	.03	.61	.02	.39	.36	.18
		$\mathcal{K}_4$	.75	.87	.85	.82	.80	.85	.79	.12	.60	.51	.36	.28	.65	.28	.15	.64	.53	.41	.31	.66	.33
		H_1^*	.06	.42	.24	.27	.12	.51	.33	.07	.18	.16	.13	.11	.31	.17	.09	.22	.18	.16	.13	.31	.19
		H_2^*	.05	.33	.36	.07	.16	.44	.35	.03	.19	.19	.05	.07	.23	.18	.05	.24	.22	.09	.09	.26	.21
		H_{max}^*	.06	.46	.39	.17	.18	.58	.45	.05	.23	.20	.11	.10	.33	.20	.08	.29	.24	.14	.12	.36	.24
	(0,0.5)	$\mathcal{K}_1$	.55	.87	.31	.78	.36	.57	.59	.06	.65	.00	.37	.00	.38	.18	.09	.68	.01	.43	.01	.40	.22
		$\mathcal{K}_1'$	.53	.31	.84	.35	.76	.56	.55	.04	.00	.58	.00	.32	.33	.13	.07	.01	.60	.01	.37	.35	.16
		$\mathcal{K}_4$	.75	.86	.85	.82	.80	.85	.79	.05	.53	.47	.28	.22	.63	.22	.10	.59	.51	.35	.27	.65	.27
		H_1^*	.05	.35	.23	.21	.12	.48	.32	.03	.12	.14	.08	.07	.27	.13	.09	.20	.18	.16	.12	.31	.19
		H_2^*	.05	.26	.29	.06	.10	.37	.31	.01	.07	.08	.02	.02	.17	.12	.04	.16	.15	.07	.07	.23	.15
		H_{max}^*	.06	.39	.32	.14	.12	.53	.39	.03	.13	.14	.07	.07	.27	.15	.08	.24	.23	.14	.11	.33	.21
400	(0,0)	$\mathcal{K}_1$	.53	.86	.31	.77	.36	.61	.58	.06	.64	.00	.36	.00	.42	.17	.08	.66	.01	.38	.01	.43	.19
		$\mathcal{K}_1'$	.54	.30	.87	.35	.78	.54	.55	.06	.01	.63	.00	.38	.36	.14	.08	.01	.64	.01	.38	.36	.16
		$\mathcal{K}_4$	.72	.86	.87	.79	.80	.88	.77	.06	.54	.53	.28	.29	.68	.22	.08	.55	.54	.30	.31	.68	.25
		H_1^*	.05	.40	.24	.25	.13	.53	.35	.06	.18	.17	.14	.10	.32	.22	.07	.20	.18	.14	.12	.35	.22
		H_2^*	.05	.38	.44	.09	.18	.45	.42	.06	.21	.20	.09	.08	.30	.25	.06	.24	.25	.11	.10	.32	.26
		H_{max}^*	.05	.50	.45	.19	.21	.61	.51	.06	.22	.22	.14	.11	.36	.24	.07	.25	.24	.15	.13	.40	.27
	(0.5,0)	$\mathcal{K}_1$	.54	.85	.31	.76	.36	.61	.57	.07	.65	.01	.38	.01	.42	.19	.08	.67	.01	.41	.01	.43	.20
		$\mathcal{K}_1'$	.54	.30	.87	.35	.78	.54	.55	.06	.01	.63	.00	.37	.35	.15	.07	.01	.64	.01	.38	.36	.16
		$\mathcal{K}_4$	.72	.86	.87	.79	.80	.88	.76	.07	.55	.53	.29	.29	.68	.24	.09	.57	.54	.32	.31	.68	.25
		H_1^*	.05	.41	.25	.26	.13	.53	.36	.07	.22	.20	.16	.13	.34	.22	.09	.22	.22	.18	.15	.37	.23
		H_2^*	.06	.38	.42	.09	.17	.45	.41	.07	.25	.27	.13	.12	.33	.27	.09	.30	.30	.14	.13	.35	.30
		H_{max}^*	.06	.49	.45	.18	.19	.61	.50	.08	.26	.26	.17	.15	.39	.27	.10	.28	.29	.20	.18	.42	.30
	(0,0.5)	$\mathcal{K}_1$	.53	.86	.31	.77	.37	.61	.58	.05	.64	.00	.36	.00	.41	.17	.07	.65	.01	.38	.00	.42	.19
		$\mathcal{K}_1'$	.53	.29	.87	.35	.78	.54	.55	.05	.00	.62	.00	.36	.35	.14	.06	.01	.64	.00	.38	.36	.15
		$\mathcal{K}_4$	.72	.86	.87	.79	.80	.88	.77	.05	.54	.51	.26	.27	.69	.21	.07	.55	.53	.30	.29	.69	.23
		H_1^*	.06	.37	.25	.23	.14	.53	.34	.04	.17	.15	.11	.09	.32	.20	.14	.28	.26	.21	.17	.38	.29
		H_2^*	.06	.37	.37	.09	.15	.44	.38	.04	.18	.16	.08	.07	.29	.22	.14	.33	.30	.18	.17	.35	.32
		H_{max}^*	.06	.48	.40	.18	.17	.60	.47	.05	.19	.18	.12	.11	.35	.23	.16	.36	.34	.23	.21	.42	.35

Table B-2: Size-adjusted power of bootstrap tests under different change points and volatility intensity, [Model 1, DGP 1, $m = 1$, $\rho = \theta = 0$, $\alpha = 0.5$, 5%]

T	λ_1^0	Test	Model 1: δ				
			1/5	1/3	1/2.5	1/1.5	1/1.1
200	0.2	H_1^*	.05	.05	.04	.07	.28
		H_2^*	.04	.04	.04	.05	.09
		H_{max}^*	.05	.04	.05	.06	.20
	0.3	H_1^*	.05	.04	.05	.15	.61
		H_2^*	.04	.04	.04	.08	.20
		H_{max}^*	.04	.05	.04	.11	.46
	0.5	H_1^*	.04	.05	.06	.56	.94
		H_2^*	.04	.04	.04	.24	.55
		H_{max}^*	.05	.04	.05	.44	.89
	0.6	H_1^*	.08	.17	.27	.87	.98
		H_2^*	.03	.04	.07	.41	.70
		H_{max}^*	.04	.08	.12	.77	.95
	0.7	H_1^*	.33	.51	.61	.92	.96
		H_2^*	.19	.24	.27	.58	.77
		H_{max}^*	.23	.32	.44	.86	.94
	0.9	H_1^*	.31	.39	.44	.58	.62
		H_2^*	.30	.31	.32	.43	.47
		H_{max}^*	.29	.33	.37	.55	.60
400	0.2	H_1^*	.06	.06	.06	.30	.88
		H_2^*	.05	.05	.05	.08	.38
		H_{max}^*	.05	.05	.05	.16	.74
	0.3	H_1^*	.06	.06	.06	.73	1.0
		H_2^*	.05	.05	.05	.26	.73
		H_{max}^*	.05	.05	.05	.55	.98
	0.5	H_1^*	.06	.08	.20	.99	1.0
		H_2^*	.05	.05	.06	.78	.99
		H_{max}^*	.05	.05	.10	.97	1.0
	0.6	H_1^*	.58	.84	.94	1.0	1.0
		H_2^*	.07	.18	.31	.93	1.0
		H_{max}^*	.30	.65	.83	1.0	1.0
	0.7	H_1^*	.96	.99	1.0	1.0	1.0
		H_2^*	.67	.75	.82	.98	1.0
		H_{max}^*	.88	.96	.98	1.0	1.0
	0.9	H_1^*	.84	.88	.90	.93	.93
		H_2^*	.78	.79	.80	.84	.86
		H_{max}^*	.80	.84	.87	.93	.93

Table B-3: Size-adjusted power of bootstrap tests, [$m = 1$, $\rho = 0.5$, $\theta = 0$, $\lambda_1^0 = 0.5$, 5%]

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	1	$\mathcal{K}_1$	.39	.68	.11	.57	.24	.43	.43	.27	.57	.04	.46	.13	.34	.31
		$\mathcal{K}_1'$	.03	.03	.02	.02	.06	.10	.05	.03	.03	.03	.02	.06	.09	.05
		$\mathcal{K}_4$	.37	.67	.07	.55	.21	.44	.41	.25	.55	.04	.44	.11	.37	.28
		H_1^*	.82	.06	.53	.15	.75	.16	.32	.39	.05	.25	.06	.43	.06	.13
		H_2^*	.30	.05	.10	.17	.26	.10	.16	.10	.05	.07	.08	.11	.06	.08
		H_{max}^*	.63	.06	.29	.13	.59	.12	.22	.22	.05	.12	.07	.26	.05	.09
	2	$\mathcal{K}_1$	.03	.03	.03	.08	.02	.11	.06	.03	.03	.03	.08	.02	.11	.06
		$\mathcal{K}_1'$	.36	.08	.63	.20	.53	.40	.38	.23	.02	.52	.11	.41	.30	.24
		$\mathcal{K}_4$	.34	.07	.62	.18	.52	.42	.36	.21	.03	.51	.10	.40	.33	.23
		H_1^*	.96	.41	.24	.75	.49	.24	.53	.70	.21	.18	.43	.28	.15	.33
		H_2^*	.46	.14	.05	.53	.16	.14	.27	.22	.09	.04	.28	.09	.09	.14
		H_{max}^*	.89	.32	.08	.75	.29	.15	.40	.53	.16	.06	.42	.13	.10	.21
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.01	.08	.00	.04	.00	.14	.01	.00	.00	.00	.00	.00	.01	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.04	.00	.02	.10	.00
		$\mathcal{K}_4$	.01	.07	.00	.04	.00	.14	.01	.00	.00	.04	.00	.02	.10	.00
		H_1^*	.81	.03	.47	.29	.71	.16	.36	.95	.44	.40	.76	.76	.23	.57
		H_2^*	.34	.02	.14	.27	.27	.08	.16	.45	.17	.03	.48	.25	.13	.26
		H_{max}^*	.69	.03	.38	.27	.64	.11	.23	.89	.42	.16	.75	.61	.17	.42
			$\alpha = 0.5$							$\alpha = 0.7$						
400	1	$\mathcal{K}_1$	.53	.76	.22	.68	.40	.54	.53	.40	.68	.10	.57	.26	.45	.42
		$\mathcal{K}_1'$	.03	.03	.03	.01	.06	.09	.05	.03	.03	.03	.02	.06	.09	.05
		$\mathcal{K}_4$	.51	.75	.19	.67	.38	.55	.51	.39	.66	.08	.56	.23	.47	.40
		H_1^*	1.0	.08	.99	.71	1.0	.43	.74	.96	.07	.90	.17	.96	.26	.48
		H_2^*	.91	.05	.63	.44	.89	.27	.45	.46	.05	.38	.15	.48	.14	.20
		H_{max}^*	1.0	.06	.97	.63	1.0	.35	.61	.89	.06	.76	.13	.92	.18	.30
	2	$\mathcal{K}_1$	.03	.02	.03	.08	.02	.12	.06	.02	.02	.03	.07	.02	.11	.06
		$\mathcal{K}_1'$	.54	.20	.78	.39	.69	.49	.54	.39	.09	.68	.23	.57	.40	.41
		$\mathcal{K}_4$	.52	.18	.77	.35	.68	.52	.51	.37	.08	.67	.21	.56	.43	.39
		H_1^*	1.0	.99	.37	1.0	.92	.50	.87	.99	.87	.25	.97	.62	.34	.68
		H_2^*	.96	.68	.06	.97	.41	.33	.55	.63	.39	.05	.76	.16	.19	.34
		H_{max}^*	1.0	.97	.14	1.0	.83	.41	.76	.98	.77	.09	.96	.37	.25	.52
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.00	.03	.00	.01	.00	.12	.01	.00	.00	.00	.00	.00	.00	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.03	.00	.01	.09	.00
		$\mathcal{K}_4$	.00	.02	.00	.01	.00	.12	.01	.00	.00	.03	.00	.01	.09	.00
		H_1^*	1.0	.15	.97	.97	.99	.44	.84	1.0	.95	.72	1.0	1.0	.52	.92
		H_2^*	.99	.05	.81	.89	.98	.31	.63	.99	.81	.19	.99	.90	.36	.73
		H_{max}^*	1.0	.07	.95	.96	.99	.39	.78	1.0	.94	.47	1.0	.99	.45	.87

Table B-4: Size-adjusted power of bootstrap tests, $[m = 1,\ \rho = 0,\ \theta = 0.5,\ \lambda_1^0 = 0.5,\ 5\%]$

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	1	$\mathcal{K}_1$	.37	.65	.09	.50	.24	.43	.40	.24	.53	.03	.41	.11	.33	.26
		$\mathcal{K}_1'$	.02	.02	.01	.01	.05	.08	.04	.02	.02	.01	.01	.04	.08	.05
		$\mathcal{K}_4$	.34	.63	.04	.49	.18	.41	.37	.21	.51	.02	.39	.09	.34	.24
		H_1^*	.83	.04	.83	.21	.89	.24	.45	.49	.04	.45	.08	.60	.12	.22
		H_2^*	.34	.04	.12	.20	.31	.09	.20	.11	.04	.07	.08	.10	.05	.09
		H_{max}^*	.69	.03	.65	.18	.80	.16	.31	.33	.04	.30	.06	.46	.07	.14
	2	$\mathcal{K}_1$	.02	.03	.03	.06	.02	.10	.05	.03	.03	.03	.06	.01	.11	.05
		$\mathcal{K}_1'$	.33	.10	.60	.21	.50	.39	.36	.20	.03	.49	.10	.37	.29	.24
		$\mathcal{K}_4$	.30	.05	.59	.16	.48	.39	.33	.18	.02	.48	.08	.36	.32	.22
		H_1^*	.91	.77	.19	.86	.56	.28	.58	.66	.41	.12	.53	.29	.16	.34
		H_2^*	.49	.18	.05	.43	.23	.13	.26	.23	.09	.04	.20	.10	.07	.15
		H_{max}^*	.84	.66	.09	.83	.41	.20	.47	.56	.33	.05	.49	.19	.10	.24
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.01	.04	.00	.02	.01	.14	.02	.00	.00	.00	.00	.00	.00	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.02	.00	.01	.09	.01
		$\mathcal{K}_4$	.00	.04	.00	.02	.00	.11	.01	.00	.00	.01	.00	.01	.08	.00
		H_1^*	.93	.04	.48	.41	.87	.25	.58	.95	.44	.11	.82	.66	.26	.60
		H_2^*	.61	.01	.11	.37	.45	.09	.29	.70	.08	.03	.51	.39	.12	.34
		H_{max}^*	.86	.02	.40	.40	.83	.18	.48	.91	.41	.05	.81	.58	.20	.52
			$\alpha = 0.5$							$\alpha = 0.7$						
400	1	$\mathcal{K}_1$	.57	.78	.31	.67	.47	.60	.57	.41	.66	.12	.55	.27	.48	.42
		$\mathcal{K}_1'$	.03	.03	.02	.01	.06	.07	.05	.02	.03	.01	.02	.05	.07	.04
		$\mathcal{K}_4$	.54	.77	.20	.66	.40	.55	.54	.39	.65	.07	.54	.23	.46	.40
		H_1^*	1.0	.06	1.0	.69	1.0	.48	.74	.94	.05	.93	.27	.98	.31	.50
		H_2^*	.84	.04	.61	.54	.84	.27	.41	.44	.05	.25	.20	.47	.14	.20
		H_{max}^*	.99	.03	1.0	.65	1.0	.41	.62	.88	.03	.86	.23	.94	.22	.35
	2	$\mathcal{K}_1$	.02	.02	.02	.06	.01	.10	.06	.02	.03	.03	.06	.01	.10	.06
		$\mathcal{K}_1'$	.58	.33	.78	.47	.68	.53	.58	.40	.14	.67	.27	.55	.41	.42
		$\mathcal{K}_4$	.54	.22	.78	.40	.68	.50	.54	.37	.09	.66	.23	.54	.42	.40
		H_1^*	1.0	1.0	.37	1.0	.90	.54	.85	.98	.92	.21	.98	.59	.36	.66
		H_2^*	.93	.66	.04	.94	.51	.34	.55	.63	.36	.04	.74	.19	.20	.33
		H_{max}^*	1.0	1.0	.13	1.0	.80	.46	.75	.95	.87	.07	.97	.40	.28	.52
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	3	$\mathcal{K}_1$	.01	.03	.00	.02	.01	.14	.03	.00	.00	.00	.00	.00	.00	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.00	.00	.01	.00	.03	.01	.01	.10	.01
		$\mathcal{K}_4$	.00	.03	.00	.02	.00	.12	.02	.00	.00	.03	.00	.01	.08	.01
		H_1^*	1.0	.23	.97	.95	1.0	.44	.86	1.0	.95	.47	1.0	.98	.49	.93
		H_2^*	.99	.10	.82	.92	.98	.34	.67	1.0	.78	.11	.99	.91	.36	.73
		H_{max}^*	1.0	.11	.97	.94	1.0	.41	.80	1.0	.95	.23	1.0	.96	.44	.87

Table B-5: Size-adjusted power of bootstrap tests, [$m = 2$, $\rho = 0.5$, $\theta = 0$, $\lambda_1^0 = 0.3$, $\lambda_2^0 = 0.8$, 5%]

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	4	$\mathcal{K}_1$	.03	.07	.01	.07	.01	.12	.04	.03	.06	.01	.07	.01	.11	.03
		$\mathcal{K}_1'$	.01	.00	.09	.00	.06	.11	.02	.01	.00	.08	.00	.05	.10	.02
		$\mathcal{K}_4$	.02	.07	.09	.05	.05	.21	.04	.02	.06	.08	.05	.05	.20	.03
		H_1^*	.21	.05	.29	.08	.28	.09	.11	.16	.03	.17	.06	.20	.06	.08
		H_2^*	.86	.37	.42	.74	.69	.19	.41	.48	.15	.18	.31	.32	.10	.20
		H_{max}^*	.80	.30	.38	.49	.56	.13	.29	.40	.10	.15	.14	.26	.07	.14
	5	$\mathcal{K}_1$	.48	.65	.40	.61	.37	.48	.50	.35	.55	.26	.50	.25	.40	.39
		$\mathcal{K}_1'$	.46	.40	.62	.37	.58	.45	.48	.33	.26	.52	.24	.47	.36	.37
		$\mathcal{K}_4$	.44	.62	.57	.53	.48	.59	.49	.33	.52	.49	.44	.40	.50	.38
		H_1^*	.54	.48	.56	.54	.62	.24	.30	.29	.29	.27	.30	.30	.15	.17
		H_2^*	.26	.12	.08	.34	.14	.15	.20	.10	.07	.05	.14	.06	.07	.10
		H_{max}^*	.35	.42	.27	.54	.41	.19	.23	.15	.24	.11	.30	.16	.11	.12
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.01	.00	.00	.00	.07	.01	.00	.00	.00	.00	.00	.05	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.03	.00	.00	.00	.00	.00	.00	.03	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.08	.00	.00	.00	.00	.00	.00	.09	.00
		H_1^*	.50	.52	.53	.53	.60	.21	.26	.21	.03	.45	.05	.40	.11	.11
		H_2^*	.41	.18	.07	.41	.23	.14	.18	.67	.26	.42	.54	.63	.17	.38
		H_{max}^*	.41	.48	.34	.53	.46	.16	.18	.63	.17	.44	.33	.53	.12	.22
			$\alpha = 0.5$							$\alpha = 0.7$						
400	4	$\mathcal{K}_1$	.02	.06	.00	.05	.00	.12	.03	.02	.05	.00	.05	.00	.12	.03
		$\mathcal{K}_1'$	.02	.00	.09	.00	.05	.10	.02	.01	.00	.08	.00	.05	.11	.02
		$\mathcal{K}_4$	.02	.05	.09	.03	.05	.22	.03	.02	.04	.08	.03	.05	.22	.04
		H_1^*	.28	.12	.40	.15	.38	.19	.18	.29	.09	.35	.12	.34	.12	.16
		H_2^*	1.0	.94	.94	1.0	.99	.50	.83	.97	.53	.62	.88	.84	.27	.53
		H_{max}^*	1.0	.91	.90	.98	.97	.38	.72	.94	.45	.59	.66	.74	.21	.42
	5	$\mathcal{K}_1$	.62	.75	.56	.73	.54	.63	.62	.49	.66	.42	.63	.39	.51	.51
		$\mathcal{K}_1'$	.63	.60	.74	.56	.73	.57	.62	.48	.43	.63	.40	.60	.46	.48
		$\mathcal{K}_4$	.58	.71	.69	.66	.62	.71	.59	.47	.62	.58	.56	.51	.62	.49
		H_1^*	.97	.94	.99	.97	1.0	.58	.68	.75	.75	.76	.79	.82	.36	.43
		H_2^*	.89	.26	.34	.80	.64	.38	.43	.43	.10	.10	.34	.21	.20	.23
		H_{max}^*	.92	.85	.90	.95	.97	.48	.53	.54	.54	.50	.74	.65	.28	.29
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.00	.00	.00	.00	.04	.00	.00	.00	.00	.00	.00	.04	.00
		$\mathcal{K}_1'$	.00	.00	.00	.00	.00	.03	.00	.00	.00	.00	.00	.00	.02	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.06	.00	.00	.00	.00	.00	.00	.06	.00
		H_1^*	.99	.98	.99	.98	1.0	.63	.75	.38	.43	.83	.27	.68	.31	.31
		H_2^*	.97	.63	.58	.92	.88	.43	.55	1.0	.92	.95	.99	.99	.52	.86
		H_{max}^*	.98	.96	.96	.98	.99	.55	.62	1.0	.77	.92	.91	.95	.44	.72

Table B-6: Size-adjusted power of bootstrap tests, [$m = 2$, $\rho = 0$, $\theta = 0.5$, $\lambda_1^0 = 0.3$, $\lambda_2^0 = 0.8$, 5%]

T	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
		δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
			$\alpha = 0.5$							$\alpha = 0.7$						
200	4	$\mathcal{K}_1$	.01	.04	.00	.04	.00	.09	.03	.01	.04	.00	.05	.00	.09	.02
		$\mathcal{K}_1'$	.01	.00	.08	.00	.04	.09	.02	.01	.00	.08	.00	.04	.10	.02
		$\mathcal{K}_4$	.01	.04	.08	.03	.04	.17	.03	.01	.04	.08	.04	.04	.18	.02
		H_1^*	.40	.14	.39	.18	.46	.13	.20	.27	.06	.23	.09	.27	.07	.13
		H_2^*	.87	.47	.46	.77	.73	.20	.48	.55	.17	.18	.36	.35	.09	.21
		H_{max}^*	.86	.36	.42	.65	.69	.15	.38	.51	.11	.18	.23	.33	.06	.16
	5	$\mathcal{K}_1$	.43	.60	.35	.55	.35	.44	.45	.31	.51	.21	.45	.21	.37	.34
		$\mathcal{K}_1'$	.38	.34	.51	.31	.48	.38	.39	.26	.21	.42	.19	.38	.29	.29
		$\mathcal{K}_4$	.40	.58	.49	.50	.44	.53	.42	.28	.49	.40	.41	.32	.46	.32
		H_1^*	.26	.35	.55	.31	.56	.17	.19	.15	.22	.31	.18	.30	.10	.13
		H_2^*	.34	.09	.07	.33	.16	.13	.21	.13	.05	.03	.14	.06	.07	.10
		H_{max}^*	.30	.30	.34	.29	.40	.13	.16	.14	.16	.17	.16	.18	.08	.09
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.00	.00	.00	.00	.03	.00	.00	.00	.00	.00	.00	.04	.00
		$\mathcal{K}_1'$	.00	.01	.00	.00	.00	.02	.00	.00	.00	.00	.00	.00	.03	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.04	.00	.00	.00	.00	.00	.00	.07	.00
		H_1^*	.25	.28	.53	.22	.56	.16	.17	.51	.37	.54	.36	.58	.17	.30
		H_2^*	.54	.16	.16	.40	.29	.11	.26	.80	.38	.39	.61	.53	.17	.45
		H_{max}^*	.40	.25	.43	.24	.46	.13	.18	.71	.40	.46	.50	.59	.17	.39
			$\alpha = 0.5$							$\alpha = 0.7$						
400	4	$\mathcal{K}_1$	.01	.03	.00	.04	.00	.11	.02	.01	.03	.00	.03	.00	.11	.02
		$\mathcal{K}_1'$	.01	.00	.07	.00	.06	.09	.03	.01	.00	.08	.00	.05	.10	.03
		$\mathcal{K}_4$	.01	.03	.07	.02	.06	.19	.03	.01	.03	.08	.02	.05	.20	.03
		H_1^*	.47	.34	.60	.31	.57	.26	.30	.38	.21	.40	.23	.42	.17	.23
		H_2^*	1.0	.92	.88	.99	.98	.51	.82	.94	.59	.53	.87	.78	.31	.58
		H_{max}^*	1.0	.89	.86	.98	.96	.42	.74	.90	.49	.46	.73	.69	.25	.47
	5	$\mathcal{K}_1$	.62	.75	.58	.73	.57	.64	.63	.48	.65	.42	.60	.40	.51	.48
		$\mathcal{K}_1'$	.60	.61	.69	.57	.69	.55	.60	.45	.43	.60	.39	.57	.43	.46
		$\mathcal{K}_4$	.60	.73	.68	.69	.63	.70	.61	.46	.63	.57	.55	.50	.61	.48
		H_1^*	.71	.82	.93	.82	.94	.52	.46	.46	.59	.76	.58	.76	.33	.30
		H_2^*	.81	.25	.28	.77	.59	.38	.41	.40	.08	.07	.42	.21	.21	.25
		H_{max}^*	.81	.68	.81	.80	.87	.47	.41	.42	.42	.47	.55	.59	.27	.22
			$\alpha_1 = 0.2, \alpha_2 = 0.9, \alpha = 0.7$							$\alpha_1 = 0.9, \alpha_2 = 0.2, \alpha = -0.7$						
	6	$\mathcal{K}_1$	.00	.00	.01	.00	.00	.04	.01	.00	.00	.00	.00	.00	.03	.00
		$\mathcal{K}_1'$	.00	.02	.00	.01	.00	.03	.01	.00	.00	.00	.00	.00	.02	.00
		$\mathcal{K}_4$	.00	.00	.00	.00	.00	.04	.00	.00	.00	.00	.00	.00	.05	.00
		H_1^*	.80	.88	.97	.82	.98	.49	.53	.82	.89	.94	.83	.92	.46	.64
		H_2^*	.98	.63	.68	.95	.93	.42	.57	1.0	.94	.90	.99	.97	.52	.87
		H_{max}^*	.96	.80	.94	.87	.97	.45	.54	.99	.92	.93	.96	.96	.49	.79

Table B-7: Empirical power of bootstrap recursive and non-recursive G_1 tests, [$m = 1$, $\theta = 0$, $\lambda_1^0 = 0.5$, 5%]

T	ρ	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
			δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
				$\alpha = 0.5$ /$\alpha_1 = 0.2, \alpha_2 = 0.9$							$\alpha = 0.7/\alpha_1 = 0.9, \alpha_2 = 0.2$						
200	0.5	1	G_1^r	.89	.20	.60	.42	.82	.37	.62	.57	.16	.31	.23	.48	.27	.37
			G_1^*	.98	.72	.86	.88	.97	.63	.88	.90	.66	.64	.71	.83	.52	.73
		2	G_1^r	.94	.50	.43	.76	.66	.44	.72	.68	.26	.35	.45	.42	.32	.51
			G_1^*	.99	.82	.83	.96	.92	.65	.91	.93	.61	.74	.83	.79	.55	.79
		3	G_1^r	.90	.14	.63	.56	.83	.37	.65	.98	.60	.59	.87	.88	.49	.82
			G_1^*	.97	.42	.74	.85	.92	.48	.78	.99	.77	.78	.95	.96	.56	.90
	0.8	1	G_1^r	.62	.20	.36	.34	.55	.31	.45	.40	.17	.17	.24	.29	.23	.29
			G_1^*	.94	.74	.80	.81	.90	.61	.79	.83	.70	.62	.71	.77	.53	.67
		2	G_1^r	.74	.28	.40	.54	.49	.37	.55	.53	.15	.36	.29	.38	.27	.40
			G_1^*	.96	.75	.82	.91	.86	.62	.85	.91	.55	.78	.77	.77	.54	.76
		3	G_1^r	.64	.10	.37	.21	.58	.26	.44	.90	.38	.72	.74	.74	.43	.72
			G_1^*	.87	.38	.57	.60	.82	.43	.63	.96	.65	.84	.92	.88	.56	.83
				$\alpha = 0.5$ /$\alpha_1 = 0.2, \alpha_2 = 0.9$							$\alpha = 0.7/\alpha_1 = 0.9, \alpha_2 = 0.2$						
400	0.5	1	G_1^r	1.0	.24	.99	.86	1.0	.57	.86	.97	.16	.89	.40	.96	.42	.70
			G_1^*	1.0	.81	1.0	.99	1.0	.78	.97	.99	.70	.96	.91	.99	.67	.91
		2	G_1^r	1.0	.99	.61	1.0	.97	.68	.95	.99	.84	.47	.95	.72	.53	.82
			G_1^*	1.0	1.0	.95	1.0	1.0	.79	.98	1.0	.95	.87	.99	.95	.71	.94
		3	G_1^r	1.0	.36	.98	.99	1.0	.58	.95	1.0	.98	.85	1.0	1.0	.64	.97
			G_1^*	1.0	.75	.99	1.0	1.0	.63	.97	1.0	.99	.94	1.0	1.0	.67	.98
	0.8	1	G_1^r	.94	.20	.89	.47	.95	.45	.72	.74	.16	.62	.28	.74	.32	.50
			G_1^*	1.0	.75	.99	.93	1.0	.72	.92	.98	.70	.90	.80	.96	.64	.82
		2	G_1^r	.99	.83	.48	.96	.74	.53	.81	.90	.51	.41	.76	.53	.42	.65
			G_1^*	1.0	.98	.88	1.0	.97	.74	.95	.99	.87	.83	.97	.90	.65	.89
		3	G_1^r	.99	.11	.85	.60	.97	.44	.78	1.0	.84	.76	.98	.93	.54	.89
			G_1^*	1.0	.36	.93	.91	.99	.53	.85	1.0	.94	.87	.99	.98	.63	.92

Table B-8: Empirical power of bootstrap recursive and non-recursive G_1 tests, [$m = 1$, $\rho = 0$, $\lambda_1^0 = 0.5$, 5%]

T	θ	DGP	Test	Model 1: δ			Model 2: δ		Model 3: c		Model 1: δ			Model 2: δ		Model 3: c	
			δ/c	1	1/3	3	1/3	3	0	10	1	1/3	3	1/3	3	0	10
				$\alpha = 0.5$ $/\alpha_1 = 0.2, \alpha_2 = 0.9$							$\alpha = 0.7/\alpha_1 = 0.9, \alpha_2 = 0.2$						
200	0.5	1	G_1^r	.92	.32	.84	.63	.91	.45	.75	.76	.28	.53	.44	.68	.33	.56
			G_1^*	.99	.74	.95	.90	.99	.64	.90	.93	.69	.74	.79	.88	.55	.82
		2	G_1^r	.95	.89	.46	.93	.70	.54	.82	.77	.62	.38	.72	.50	.41	.64
			G_1^*	.99	.96	.81	.98	.90	.66	.92	.93	.82	.72	.90	.77	.56	.80
		3	G_1^r	.97	.32	.79	.80	.95	.45	.80	.99	.78	.37	.96	.81	.45	.87
			G_1^*	.99	.51	.81	.91	.97	.49	.87	1.0	.78	.55	.97	.92	.50	.90
	0.8	1	G_1^r	.88	.58	.75	.75	.85	.54	.76	.82	.55	.56	.69	.77	.45	.71
			G_1^*	.96	.81	.75	.89	.91	.61	.86	.93	.77	.61	.83	.86	.54	.80
		2	G_1^r	.89	.87	.60	.90	.72	.54	.81	.83	.78	.54	.82	.61	.48	.71
			G_1^*	.96	.87	.78	.93	.87	.61	.87	.93	.83	.74	.88	.79	.54	.80
		3	G_1^r	.83	.37	.39	.69	.62	.31	.60	.80	.19	.35	.47	.66	.29	.56
			G_1^*	.82	.41	.37	.70	.59	.31	.60	.80	.17	.38	.46	.68	.31	.56
				$\alpha = 0.5$ $/\alpha_1 = 0.2, \alpha_2 = 0.9$							$\alpha = 0.7/\alpha_1 = 0.9, \alpha_2 = 0.2$						
400	0.5	1	G_1^r	1.0	.34	1.0	.90	1.0	.63	.89	.97	.22	.93	.59	.97	.53	.78
			G_1^*	1.0	.82	1.0	.99	1.0	.78	.97	1.0	.73	.97	.94	.99	.70	.91
		2	G_1^r	1.0	1.0	.64	1.0	.96	.71	.97	.99	.95	.49	.99	.77	.60	.86
			G_1^*	1.0	1.0	.93	1.0	.99	.79	.98	1.0	.97	.85	1.0	.95	.70	.94
		3	G_1^r	1.0	.62	.99	.99	1.0	.61	.96	1.0	1.0	.75	1.0	.99	.62	.97
			G_1^*	1.0	.77	.99	1.0	1.0	.64	.97	1.0	1.0	.83	1.0	1.0	.64	.97
	0.8	1	G_1^r	.98	.53	.98	.83	.99	.67	.89	.94	.50	.91	.71	.96	.61	.83
			G_1^*	.99	.81	.97	.95	1.0	.74	.95	.98	.77	.91	.87	.98	.69	.89
		2	G_1^r	.99	.98	.66	1.0	.87	.70	.93	.94	.96	.60	.98	.73	.62	.86
			G_1^*	1.0	.98	.86	.99	.94	.75	.95	.98	.96	.80	.98	.87	.68	.90
		3	G_1^r	.99	.60	.82	.96	.96	.48	.85	.99	.68	.58	.95	.95	.47	.85
			G_1^*	.99	.65	.80	.96	.95	.48	.85	.99	.66	.65	.95	.95	.47	.84

Table B-9: Break selection probabilities under different abrupt volatility break points and intensities, [Model 1, DGP 5 and 6, $m = 2$, $\rho = \theta = 0$, 5%]

T	DGP	$\alpha/(\alpha_1, \alpha_2)$	τ		Model 1: δ								
					1/3	1/2.5	1/1.5	1/1.1	1	1.1	1.5	2.5	3
400	5	0.5	0.2	P_c	.47	.54	.76	.83	.84	.84	.84	.63	.44
				P_o	.13	.15	.17	.16	.15	.14	.14	.30	.45
			0.3	P_c	.11	.11	.62	.83	.84	.85	.80	.50	.35
				P_o	.06	.07	.14	.15	.14	.14	.16	.37	.46
			0.5	P_c	.07	.06	.66	.84	.84	.82	.66	.08	.08
				P_o	.08	.10	.14	.14	.15	.16	.17	.08	.06
			0.8	P_c	.68	.77	.87	.86	.84	.80	.61	.18	.15
				P_o	.15	.13	.12	.13	.15	.17	.19	.08	.06
			0.9	P_c	.89	.88	.87	.85	.84	.83	.78	.69	.66
				P_o	.09	.11	.12	.14	.15	.15	.17	.17	.16
	6	(0.2, 0.9)	0.2	P_c	.86	.85	.89	.89	.90	.90	.89	.72	.55
				P_o	.12	.14	.11	.11	.11	.10	.11	.27	.42
			0.3	P_c	.18	.41	.88	.89	.89	.90	.88	.67	.50
				P_o	.07	.09	.11	.11	.11	.10	.12	.29	.40
			0.5	P_c	.08	.28	.88	.88	.90	.89	.88	.31	.11
				P_o	.07	.09	.12	.12	.10	.11	.11	.12	.09
			0.8	P_c	.80	.86	.88	.89	.88	.89	.86	.46	.32
				P_o	.14	.12	.12	.11	.12	.11	.12	.12	.08
			0.9	P_c	.90	.90	.90	.89	.89	.88	.88	.86	.86
				P_o	.09	.10	.10	.11	.11	.12	.12	.13	.13
600	5	0.5	0.2	P_c	.78	.80	.85	.86	.87	.87	.88	.74	.58
				P_o	.15	.16	.15	.14	.13	.13	.12	.25	.41
			0.3	P_c	.09	.12	.83	.86	.87	.87	.88	.71	.54
				P_o	.05	.06	.15	.14	.13	.13	.12	.27	.40
			0.5	P_c	.05	.07	.86	.88	.86	.85	.83	.13	.07
				P_o	.07	.08	.13	.12	.14	.15	.16	.10	.08
			0.8	P_c	.82	.86	.87	.87	.86	.86	.76	.25	.18
				P_o	.15	.13	.13	.14	.14	.14	.21	.13	.09
			0.9	P_c	.88	.89	.88	.86	.87	.87	.85	.81	.81
				P_o	.12	.11	.12	.14	.14	.14	.15	.18	.18
	6	(0.2, 0.9)	0.2	P_c	.84	.85	.86	.85	.86	.86	.86	.80	.67
				P_o	.16	.15	.15	.15	.14	.14	.14	.20	.33
			0.3	P_c	.57	.76	.86	.86	.86	.86	.87	.79	.67
				P_o	.12	.15	.14	.14	.14	.14	.13	.21	.32
			0.5	P_c	.32	.68	.85	.86	.85	.86	.88	.64	.30
				P_o	.11	.15	.15	.14	.15	.14	.12	.14	.13
			0.8	P_c	.84	.85	.86	.86	.86	.87	.86	.74	.56
				P_o	.16	.16	.14	.14	.14	.14	.14	.15	.12
			0.9	P_c	.86	.86	.86	.86	.85	.86	.86	.87	.85
				P_o	.14	.14	.14	.14	.15	.14	.14	.14	.15

Table B-10: Regime-wise estimates of OECD countries with breaks in persistence

Country	$\hat{m}$	Regime	AR sum	90% Band	ADF p-value
(1)	(2)	(3)	(4)	(5)	(6)
Belgium	1	1st	.57	[.44, .75]	.00
		2nd	.82	[.71, 1.05]	.11
France	2	1st	.10	[-.12, .38]	.00
		2nd	.65	[.51, .89]	.10
		3rd	.63	[.51, .79]	.00
Germany	2	1st	.21	[.10, .36]	.00
		2nd	.52	[.42, .64]	.00
		3rd	.06	[-.51, 1.15]	.08
Italy	2	1st	.01	[-.13, .19]	.00
		2nd	.51	[.36, .73]	.00
		3rd	.88	[.82, 1.02]	.07
Luxembourg	1	1st	.87	[.76, .1.10]	.51
		2nd	-.13	[-.80, 1.17]	.01