DISCUSSION PAPER SERIES

DP15435

LM TESTS FOR JOINT BREAKS IN THE DYNAMICS AND LEVEL OF A LONG-MEMORY TIME SERIES

Juan J. Dolado, Heiko Rachinger and Carlos Velasco

FINANCIAL ECONOMICS

ISSN 0265-8003

LM TESTS FOR JOINT BREAKS IN THE DYNAMICS AND LEVEL OF A LONG-MEMORY TIME SERIES

Juan J. Dolado, Heiko Rachinger and Carlos Velasco

Discussion Paper DP15435
Published 08 November 2020
Submitted 07 November 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

LM TESTS FOR JOINT BREAKS IN THE DYNAMICS AND LEVEL OF A LONG-MEMORY TIME SERIES

Abstract

We consider a single-step Lagrange Multiplier (LM) test for joint breaks (at known or unknown dates) in the long memory parameter, the short-run dynamics and the level of a fractionally integrated time-series process. The regression version of this test is easily implementable and allows to identify the speci c sources of the break when the null hypothesis of parameter stability is rejected. However, its size and power properties are sensitive to the correct specification of short-run dynamics under the null. To address this problem, we propose a slight modification of the LM test (labeled LMW-type test) which also makes use of some information under the alternative (in the spirit of a Wald test). This test shares the same limiting distribution as the LM test under the null and local alternatives but achieves higher power by facilitating the correct specification of the short-run dynamics under the null and any alternative (either local or fixed). Monte Carlo simulations provide support for these theoretical results. An empirical application, concerning the origin of shifts in the long-memory properties of forward discount rates in five G7 countries, illustrates the usefulness of the proposed LMW-type test.

JEL Classification: N/A

Keywords: N/A

Juan J. Dolado - dolado@eco.uc3m.es
Universidad Carlos III de Madrid, CEPR and CEPR

Heiko Rachinger - heiko.rachinger@uib.es
Universitat de les Illes Balears

Carlos Velasco - carlos.velasco@uc3m.es
Universidad Carlos III de Madrid

Acknowledgements
Forthcoming in Journal of Business & Economic Statistics. We are indebted to a Co-editor, an Associate Editor and three anonymous referees for many useful comments that helped improve the paper substantially. We also thank Eric Zivot for making the exchange-rate dataset available to us, Juan-Carlos Escanciano, Robinson Kruse, Jesús Gonzalo, Uwe Hassler, Laura Mayoral, Peter Robinson, Abderrahim Taamouti, and participants in several seminars and workshops for helpful suggestions on earlier drafts of the paper. Financial support from the Spanish Ministerio de Economía y Competitividad (grants ECO2016-78652 and ECO2017-86009-P, ECO2017-83255-C3-P, ECO2014-57007p and MDM 2014-0431), and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444) is gratefully acknowledged. The usual disclaimer applies.

LM tests for joint breaks in the dynamics and level of a long-memory time series*

Juan J. Dolado

Universidad Carlos III de Madrid and CEPR

Heiko Rachinger

Universitat de les Illes Balears

Carlos Velasco

Universidad Carlos III de Madrid

This version: 19 October, 2020.

Abstract

We consider a single-step Lagrange Multiplier (LM) test for joint breaks (at known or unknown dates) in the long memory parameter, the short-run dynamics and the level of a fractionally integrated time-series process. The regression version of this test is easily implementable and allows to identify the specific sources of the break when the null hypothesis of parameter stability is rejected. However, its size and power properties are sensitive to the correct specification of short-run dynamics under the null. To address this problem, we propose a slight modification of the LM test (labeled LMW-type test) which also makes use of some information under the alternative (in the spirit of a Wald test). This test shares the same limiting distribution as the LM test under the null and local alternatives but achieves higher power by facilitating the correct specification of the short-run dynamics under the null and any alternative (either local or fixed). Monte Carlo simulations provide support for these theoretical results. An empirical application, concerning

*Corresponding author: Carlos Velasco, Department of Economics, Universidad Carlos III de Madrid, Getafe, 28903 Madrid, Spain; E-Mail: carlos.velasco@uc3m.es. We are indebted to a Co-editor, an Associate Editor and three anonymous referees for many useful comments that helped improve the paper substantially. We also thank Eric Zivot for making the exchange-rate dataset available to us, Juan-Carlos Escanciano, Robinson Kruse, Jesús Gonzalo, Uwe Hassler, Laura Mayoral, Peter Robinson, Abderrahim Taamouti and participants in several seminars and workshops for helpful suggestions on earlier drafts of the paper. Financial support from the Spanish Ministerio de Economía y Competitividad (grants ECO2016-78652 and ECO2017-86009-P, ECO2017-83255-C3-P, ECO2014-57007p and MDM 2014-0431), and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444) is gratefully acknowledged. The usual disclaimer applies.

the origin of shifts in the long-memory properties of forward discount rates in five G7 countries, illustrates the usefulness of the proposed LMW-type test.

JEL Classification: C13, C22

Keywords: LM Test, Structural Breaks, Long Memory, Level

1 Introduction

The confoundedness issues raised by Diebold and Inoue (2001) and Granger and Hyung (2004) have sparked controversy about the origin of long-memory features in some time series processes. Part of this debate has focused on whether long-memory is truly driven by a *fractionally integrated process* of order d, $I(d)$, or spuriously generated by level shifts in short-memory time series instead (see, e.g., Lobato and Savin (1998), and Perron and Qu (2010)). Conversely, it has been claimed that breaks in the memory parameter d could be misleadingly interpreted as breaks in the level, μ, of this type of stochastic processes (see, e.g., McCloskey (2010), and Shao (2011)).

These different views have led to two strands of research on this topic (for a general overview, see Aue and Horváth 2011). The first one deals with testing for breaks only in d. Following the rationalization of $I(d)$ processes in terms of aggregation of heterogeneous persistent processes (see Robinson 1978, and Granger 1980), it has been argued that policy regime changes can shift the long-memory component of many macro and financial variables over relevant subsamples (see, e.g. Gadea and Mayoral, 2005, for empirical evidence on these shifts in inflation). Accordingly, several tests have been proposed (both in the time and frequency domains) to test the null of a stable value of d against the alternative of a structural break at known or unknown dates; see, *inter alia*, Beran and Terrin (1996), Hassler and Scheithauer (2011) and Yamaguchi (2011).[1]

In parallel, another line of research has focused on the derivation of tests for breaks only in the level μ (or in other deterministic components) of stochastic processes with stationary long-memory disturbances where d is constant; see, e.g., Hidalgo and Robinson (1996), Lavielle and Moulines (2000), and Iacone *et al.* (2013). Lastly, there are also studies on the design of robust estimation procedures of the memory parameter in the presence of level/trend shifts (see, e.g. McCloskey and Perron, (2013)).

A common feature in most of the above-mentioned literature is that breaks are only allowed in a *single* parameter (either in d or μ). However, the more realistic case of potential *joint* breaks in both parameters, and possibly in the short-run dynamics of an I(d) process, has received much less

[1] Forerunners of this line of research are Kim et al. (2002), Busetti and Taylor (2004), and Harvey et al. (2009) who test for changes in time series from being $I(0)$ to being $I(1)$ or viceversa. Multiple changes are tackled in Leybourne et al. (2007) and Kejriwal et al. (2013).

attention. This is somewhat surprising since, following the detection of a break, it is important to find out if its origin comes from only one or several parameters at the same time.[2] Among the scant literature on this specific issue, there are a few contributions related to ours to be highlighted. To our knowledge, Gil-Alaña (2008) is the first paper to propose a single-step testing procedure based on a Chow F-test. Yet, despite conjecturing that its limiting distribution corresponds to the one derived by Bai and Perron (1998) for parameter breaks in regressions involving $I(0)$ series, no formal proof of this claim is provided. Next, Hassler and Meller (2014) have extended Robinson (1994) and Breitung and Hassler's (2002) LM test of $I(1)$ *vs.* $I(d)$ to deal with breaks in d, where level shifts are also allowed. Their proposed test is conducted in a two-step sequential fashion. Initially, the location of the mean break is detected using Hsu's (2005) semiparametric testing approach; next, the corresponding broken mean is removed from the time series to test for a break in d. However, the issue of how the two-step procedure affects the asymptotic properties of the test is not analyzed by these authors. This could be problematic in some instances: for example, μ could be very imprecisely estimated at the demeaning stage when d is close to 0.5, due to the $T^{1/2-d}$ rate of convergence of the sample mean estimator. Some of these shortcomings have been addressed by Rachinger (2017), who proposes a unified (single-step) testing procedure for modelling joint breaks. As in Gil-Alaña (2008), this author extends Bai and Perron's (1998) test from $I(0)$ to $I(d)$ processes by proposing a Likelihood Ratio (LR) version of the standard Chow test for the null of parameter stability of d and μ when $d \in [0, 0.5)$. Consistency results, T-rate convergence of the break fraction estimator and the limiting distributions of the estimated parameters under different sources of breaks are derived.

Our main goal in this paper is to propose LM alternatives to the LR test for joint parameter breaks in I(d) processes because the use of restricted estimates under the null makes LM tests computationally much simpler. In particular, as in Hassler and Meller (2014), we focus on the derivation of a regression version of the LM test which provides a linearization of the true model under local alternatives involving parameter breaks. Yet, we differ from their approach in several important respects. First, to address the shortcomings of their two-stage procedure, we derive a single-step LM testing procedure. Second, in addition to breaks in d and μ, we also allow for shifts in the short-run dynamics of an I(d) process, which are modelled using a parametric autoregressive process of order p, labeled AR(p).[3] A potential limitation of the LM test, however,

[2] Dolado *et al.* (2008) argue that it is important to distinguish between breaks in d and in μ for at least two reasons. First, because it can improve forecasting; in particular, the larger d is, the more observations are required to produce good forecasts. Second, because if d is estimated too high due to shifts in μ in bivariate systems, fractional cointegration could become a spurious outcome.

[3] Although a semiparametric approach would help us abstract from short-term dynamics when estimating d, we opt here for a parametric approach due to our interest in identifying further potential breaks in the short-term dynamics. We choose an AR(p) process to model short-run dynamics because this type of

is that its implementation requires the restrictive assumption of a known lag length of the AR(p) process under the null, which might not be the correct one under the alternative. Inspired by Wooldridge (1990), we fix this problem by means of an alternative regression-based LM test which, besides yielding consistent estimation of p when parameters are allowed to shift, exhibits higher power than the LM test for joint breaks under fixed alternatives. The insight for this power gain is that, unlike the local approximation provided by the LM test, the new test yields an exact regression representation of the true model specification under any alternative (local or fixed), where the relevant coefficients to be tested happen to be linearly related to the parameters of interest. As a result, it can also be partially interpreted as a Wald test and, for this reason, it is labeled "LMW-type" test in the sequel.

LMW-type tests have been proposed by Dolado *et al.* (2002, 2009), and Lobato and Velasco (2007) to test the nulls of $I(1)/I(0)$ against the alternative of $I(d)$ processes, with $d \in (0,1)$ under the assumption of parameter stability. We extend their testing approach by allowing for joint breaks in d, μ and the short-run dynamics when the null is an $I(d)$ process with stable parameters. Moreover, both LM and LMW-type tests can deal with shifts in $d \in (-0.5, 0.5)$ under the alternative hypothesis, covering a wider range of values than those considered by Rachinger's (2017) LR tests. This generalization can be achieved because the only requirement for implementing our LM tests is adequate performance of the constrained estimators of the parameters under the null, whereas LR tests also require good performance of the unconstrained estimators.

Summing up, by deriving single-step LM and LMW-type tests (and their asymptotic distribution under the null and alternatives), this paper contributes to the relevant literature on detecting the source of breaks in persistent time-series processes. More concretely, the proposed tests: (i) allow to test for the presence of joint or individual breaks in a wide range of parameters, involving non-stable long-memory dynamics, short-run dynamics or the level parameter; (ii) are easily implementable by means of regression methods under the joint null of parameter stability; (iii) exhibit similar asymptotic behaviour under the null and local alternatives but the LMW-type test has higher power under fixed alternatives, especially when short-run dynamics are present;[4] (iv) provide consistent estimates of the break date when considered to be unknown; and (v) can be used when either breaks in different parameters might not be coincidental in time or when there are multiple breaks. Finally, our empirical application on potential breaks in forward discount rates for several G7 countries provides new findings on their origin (in dynamics and/or in levels), an issue which has raised considerable attention in the literature on exchange rates.

The rest of the paper is structured as follows. In Section 2, we lay out the data generating

process can be easily incorporated in the regression version of the LM tests.

[4]Notice that, in spite of the nonlinear nature of our proposed tests, this result somehow echoes the well-known ranking in terms of power of Wald and LM tests in linear regression setups; see Engle (1984).

processes (DGP). In Section 3 and Section 4, we derive the LM and LMW-type tests, respectively. Sections 5 and 6 are devoted to study their asymptotic properties both under the null and under local and fixed alternatives, distinguishing between two different settings: known and unknown break dates, and briefly sketch how the tests could be generalized to account for non-coincidental and multiple breaks in time. In Section 7, we provide simulation results regarding the finite-sample performance of the tests. In Section 8 we discuss an empirical application related to the detection of structural changes in the forward discount of exchange rates. Finally, Section 9 concludes. All the proofs, some technical results, and additional simulation results are gathered in an Online Appendix.

2 Data generating process

For simplicity, we start by considering the case of a *single* breakpoint (at a *known* or *unknown* date) which changes in the asymptotics as a fraction λ_0 of the sample size; λ_0 lies in the interval $\Lambda = [\epsilon, 1-\epsilon]$, where $\epsilon > 0$ is assumed to be known. In particular, we consider an autoregressive $I(d_0)$ (i.e. ARFI (p,d)) process with long memory parameter $d_0 \in D$, where $D \subset (-0.5, 0.5)$, level μ_0 and short-run dynamics captured by a finite order AR lag polynomial $\alpha_0(L)$ during the first subsample, $t = 1, \ldots, [\lambda_0 T]$. This process may become $I(d_1)$ with $d_1 \in D$, level μ_1 and autoregression $\alpha_1(L)$ during the second subsample, $t = [\lambda_0 T]+1, \ldots, T$.[5] These assumptions lead to the following transition model, considered as the DGP in the sequel

$$\alpha_t(L)\,\Delta_t^{d_t}(y_t - \mu_t) = \varepsilon_t, \quad t = 1, 2, \ldots, \tag{1}$$

with $\varepsilon_t \sim$ i.i.d.$(0, \sigma_0^2)$. The shifting parameters are defined as:

$$\begin{aligned} \alpha_t(L)\,\Delta_t^{d_t} &= 1(t \le [T\lambda_0])\,\alpha_0(L)\,\Delta_t^{d_0} + 1(t > [T\lambda_0])\,\alpha_1(L)\,\Delta_t^{d_1}, \\ \mu_t &= 1(t \le [T\lambda_0])\,\mu_0 + 1(t > [T\lambda_0])\,\mu_1, \end{aligned}$$

where $1(\cdot)$ is an indicator function of the relevant subsample; $[x]$ denotes the integer part of x; and $\alpha_i(L) = 1 - \alpha_{1,i}L \cdots - \alpha_{p,i}L^p$ are stable AR lag polynomials of known order p with unknown coefficients $\boldsymbol{\alpha}_i = (\alpha_{1,i}, \ldots, \alpha_{p,i})'$, $i = 0, 1$, such that $\boldsymbol{\alpha}_i \in Int(\mathbf{A})$ where $\mathbf{A}$ is a compact set such

[5] Our choice of the stationary and invertible range $D \subset (-0.5, 0.5)$ is dictated in part by the result in Hualde and Nielsen (2019) showing that consistent estimation of the level in an ARFIMA (p,d,q) process with a constant term ($\gamma_0 = 1$, in their notation) and d lying in an arbitrarily large finite interval requires $d < 0.5$. However, when $d > 0.5$, the estimates of the other parameters governing the dynamics of the process are consistent and asymptotically normal, as in Hualde and Robinson (2011). Remark 4 below includes a further discussion about the implementation of our tests when $d_0, d_1 > 0.5$.

that all the roots of $\alpha_i(L)$ are outside the unit circle.[6] At $[\lambda_0 T]$, a shift in the parameters of the DGP in (1) is allowed, so that $d_1 = d_0 + \theta_0$, $\mu_1 = \mu_0 + \nu_0$, and $\alpha_1(L) = \alpha_0(L) + \beta_0(L)$, where $\beta_0(L)$ is another lag polynomial with $\beta_0(0) = 0$ and coefficients $\boldsymbol{\beta}_0 = (\beta_{1,0}, \ldots, \beta_{p,0})'$. Finally, $\Delta_t^b := \sum_{j=0}^{t-1} \pi_j(b) L^j$, where $\pi_j(b) := \frac{\Gamma(j-b)}{\Gamma(-b)\Gamma(j+1)}$, $j = 0, 1, \ldots$, denotes the (truncated or "Type II") fractional-differencing filter for $b \in D$. It should be noted that Hualde and Nielsen (2020) have previously analyzed the estimation and inference of a similar process to DGP (1) with more general short-run dynamics than the AR(p) process assumed here, but with stable parameters. To relax this last assumption and provide inference on the existence and location of a break, our proposed LM approach relies on restricted-parameter estimation under the null of no breaks.

Remark 1. Notice that the previous definition of $\Delta_t^{d_t}$ implies that the filter applied to $(y_t - \mu_t)$ is $\sum_{j=0}^{t-1} \pi_j^*(d_0, \boldsymbol{\alpha}_0)$ when $t < [\lambda_0 T]$, and $\sum_{j=0}^{t-1} \pi_j^*(d_1, \boldsymbol{\alpha}_1)$ when $t > [T\lambda_0]$, where $\alpha_i(L)\Delta_t^{d_i} := \sum_{j=0}^{t-1} \pi_j^*(d_i, \boldsymbol{\alpha}_i) L^j$. We prefer to use this truncated "Type II" filter, rather than a non-truncated "Type-I" filter, because it facilitates the treatment of non-stationary series with $d > 1/2$ after first differencing (see Remark 4 below).

Remark 2. Notice that, by rewriting the DGP as $y_t = \mu_t + \left(1 - \alpha_t(L)\Delta_t^{d_t}\right)(y_t - \mu_t) + \varepsilon_t$, and using the truncated filters $\pi_j^*(d_i, \boldsymbol{\alpha}_i)$, $i = 0, 1$ recursively, it follows that

$$\begin{aligned} y_t &= \mu_0 - \sum_{j=1}^{t-1} \pi_j^*(d_0, \boldsymbol{\alpha}_0)\{y_{t-j} - \mu_0\} + \varepsilon_t, \quad \text{for } t \leq [T\lambda_0] \\ y_t &= \mu_1 - \sum_{j=1}^{t-[T\lambda_0]-1} \pi_j^*(d_1, \boldsymbol{\alpha}_1)\{y_{t-j} - \mu_1\} - \sum_{j=t-[T\lambda_0]}^{t-1} \pi_j^*(d_1, \boldsymbol{\alpha}_1)\{y_{t-j} - \mu_0\} + \varepsilon_t, \quad \text{for } t > [T\lambda_0], \end{aligned}$$

implying that the chosen filter guarantees that the lags of y_t in the autoregression are centered around the appropriate value of the level μ_t in each of the two subsamples. Likewise, it ensures that *all* past information is discounted at the relevant value of d to generate each new observation before and after the break.

Remark 3. An alternative DGP that could be considered is the following

$$y_t = \mu_t + \Delta_t^{-d_t} \alpha_t^{-1}(L)\varepsilon_t, \quad t = 1, 2, \ldots,$$

with $\varepsilon_t \sim$ i.i.d.$(0, \sigma_0^2)$, such that

$$\begin{aligned} \Delta_t^{-d_t} \alpha_t^{-1}(L) &= 1(t \leq [T\lambda_0])\,\Delta_t^{-d_0} \alpha_0^{-1}(L) + 1(t > [T\lambda_0])\,\Delta_t^{-d_1} \alpha_1^{-1}(L), \\ \mu_t &= 1(t \leq [T\lambda_0])\,\mu_0 + 1(t > [T\lambda_0])\,\mu_1. \end{aligned}$$

[6]Formally the previous expression for the filter $\alpha_t(L)\Delta_t^{d_t}$ should be multiplied by $1(t > 0)$ since nesting the AR(p) lag polynomial $\alpha_t(L)$ with the truncated fractional filter $\Delta_t^{d_t}$ would require using pre-sample observations (negative lags). However, for simplicity, we omit this more precise notation in the sequel.

Rather than based on an autoregressive representation as in (1), this DGP (labeled DGP-MA to distinguish it from DGP (1)) provides an alternative definition of a breaking stochastic process based on a moving average that reweights in each period the whole sequence of innovations from $t = 1$ but ignores how the observations were actually generated prior to t. Although we show that our proposed tests are also consistent under local or fixed alternatives for DGP-MA (see Corollaries 2 and 3 below), we prefer to work with the autoregressive DGP (1) on the grounds that, while the level of y_t is adjusted immediately after the break in both DGPs, DGP-MA reinitializes completely the process after the break, i.e. $y_t = \mu_1 + \Delta_t^{-d_1} \alpha_1^{-1}(L)\,\varepsilon_t$ for $t > [\lambda_0 T]$ does not depend on d_0. Instead, under DGP (1), observations y_t for $t = [\lambda_0 T]+1, [\lambda_0 T]+2, \ldots, T$ are generated by filtering *all* (centered) past observations with the new memory value d_1, so that for $t > [\lambda_0 T]$,

$$y_t - \mu_1 = \left(\Delta_t^{-d_1} - \Delta_{t-[T\lambda_0]}^{-d_1}\right) \left\{ \frac{\Delta_t^{d_1-d_0}}{\alpha_0(L)} \varepsilon_t \right\} + \Delta_{t-[T\lambda_0]}^{-d_1} \alpha_1^{-1}(L)\,\varepsilon_t,$$

where the first term on the RHS accounts for innovations prior to the break and the second term for the ones afterwards, which are treated in the same way in both DGPs. Further, DGP (1) is more amenable to analytical and numerical analysis because, in practice, the easiest way to obtain residuals is to use autoregressive fractional filters on observed data, which are easily computed by means of fractional differencing. By contrast, DGP-MA requires the use of a more complicated recursive procedure based on implicitly defined residuals. In effect, $\alpha_1(L)\,\Delta_t^{d_1}(y_t - \mu_t) \neq \varepsilon_t$ for $t > [\lambda_0 T]$ when data are generated by that DGP, as this filtering ignores that observations were integrated with d_0 up to $t = [\lambda_0 T]$. Power comparisons of the tests for data generated under either DGP are provided in Section 7 (Block (III)) showing that differences are minor.

Remark 4. Our approach can also deal with a non-stationary process with both $d_0, d_1 > 0.5$, and a potentially breaking linear trend, such that

$$\alpha_t(L)\Delta_t^{d_t}(y_t - \mu_t - \zeta_t t) = \varepsilon_t,$$

with $\zeta_t = \zeta_0 1(t \leq [\lambda_0 T]) + \zeta_1 1(t > [\lambda_0 T])$, by applying our testing procedure to the first-differenced data, Δy_t, to test for breaks in the trend slope ζ_t and in the memory $d_t - 1$. To provide initial consistent estimates of non-stationary values of d under the null, one possibility is to use Hualde and Nielsen's (2020) estimation procedure for I(d) processes (with d lying in an arbitrarily large interval). Once this initial estimate is first differenced, our proposed tests in a stationary setup would be valid.

In sum, using the previous notation for potential shifts in the memory parameter (θ_0), the stable AR component ($\boldsymbol{\beta}_0$) and in the level (ν_0), and labeling the dummy variable for the second subsample (regime) as $R_t(\lambda) = 1(t > [\lambda T])$, the following model will be considered in Sections 3 and 4 to develop our testing procedures (under coincidental breaks),[7]

$$\left(\alpha_0\left(L\right)+R_t\left(\lambda_0\right)\beta_0\left(L\right)\right)\Delta_t^{d_0+\theta_0 R_t(\lambda_0)}\left(y_t-\mu_0-\nu_0 R_t\left(\lambda_0\right)\right)=\varepsilon_t,\ \ t=1,\ldots,T.$$

Remark 5. It is important to highlight at this stage that the assumption of known lag length p under the null of no break in the short-memory parameters could be highly restrictive. In effect, if the chosen p is not the right one, this could lead to incorrect inference about the existence of breaks. In practice, following Schwert (1989), Hassler and Meller (2014) argue that many short memory processes can be approximated by letting p grow with T, e.g. according to the rule of thumb: $p=[4(T/100)^{1/4}]$. However, unlike us, these authors do not allow for breaks in the AR polynomial $\alpha(L)$. Hence, for analytical tractability, the derivation of the asymptotic distribution of the LM test will proceed momentarily under the assumption of known p, while later it will be shown that the use of the LMW-type test avoids such a restrictive assumption.[8]

3 Score-driven and regression-based LM tests

According to the LM principle, we test the null hypothesis:

$$H_0:(\theta_0,\boldsymbol{\beta}_0',\nu_0)=\mathbf{0}, \tag{H0}$$

against the alternative hypothesis where *all* parameters are allowed to shift at a fraction λ_0 of the sample size, which for the moment is assumed to be known:

$$H_1(\lambda_0):(\theta_0,\boldsymbol{\beta}_0',\nu_0)\neq\mathbf{0}. \tag{H1}$$

From (2), the following Gaussian pseudo-log-likelihood function is used

$$\mathcal{L}_T\left(\psi,\lambda\right)=-\frac{T}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t\left(\psi,\lambda\right)^2, \tag{2}$$

for every possible breakpoint λ, and $\psi=\left(\theta,\boldsymbol{\beta}',\nu,d,\boldsymbol{\alpha}',\mu,\sigma^2\right)'$, where the definition of the error term above is given by

$$\varepsilon_t\left(\psi,\lambda\right)=\left(\alpha_0\left(L\right)+R_t\left(\lambda\right)\beta_0\left(L\right)\right)\left(\Delta_t^{d_0+\theta R_t(\lambda)}\left(y_t-\mu_0\right)-\nu\Delta_t^{d_0+\theta R_t(\lambda)}R_t\left(\lambda\right)\right).$$

[7] An equivalent representation yielding identical results would be to consider the model $\left(\alpha_1\left(L\right)-S_t\left(\lambda_0\right)\beta_0\left(L\right)\right)\Delta_t^{d_1-\theta_0 S_t(\lambda_0)}\left(y_t-\mu_1+\nu_0 S_t\left(\lambda_0\right)\right)=\varepsilon_t$, with $S_t(\lambda)=1-R_t(\lambda)=1\left(t\leq[\lambda T]\right)$.

[8] Ideally the number and location of breaks, and the number of autoregressive lags of the time-series processes should all be chosen simultaneously, but this is well beyond the scope of the paper. Thus, it is preferable to shape our approach for break testing as being robust to the choice of p by taking it large enough as to provide a good fit, but without allowing it to grow with T since this would not only involve a completely different asymptotic theory but also affect the power properties of the proposed LM tests.

For each λ, the LM test is based on the derivatives of $\mathcal{L}_T(\psi,\lambda)$ in the direction of ψ, evaluated at the restricted estimates $\tilde{\psi}_T = (0,\mathbf{0}',0,\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}'_{0T},\tilde{\mu}_{0T},\tilde{\sigma}^2_{0T})'$. The last four elements of $\tilde{\psi}_T$ denote estimates of parameters $d_0,\boldsymbol{\alpha}'_0$, μ_0 and σ^2_0, respectively, under the null of no breaks, where the whole sample of observations, $t=1,\ldots,T$, is used to obtain such estimates.[9] In particular, the score-driven formulation of the LM test becomes

$$\widetilde{LM}_T(\lambda) = \left.\frac{\partial\mathcal{L}_T(\psi,\lambda)}{\partial\psi'}\right|_{\psi=\tilde{\psi}_T}\left(-\left.\frac{\partial^2\mathcal{L}_T(\psi,\lambda)}{\partial\psi\partial\psi'}\right|_{\psi=\tilde{\psi}_T}\right)^{-1}\left.\frac{\partial\mathcal{L}_T(\psi,\lambda)}{\partial\psi}\right|_{\psi=\tilde{\psi}_T}, \tag{3}$$

where the score in the directions of θ, $\boldsymbol{\beta}$ and ν can be expressed as

$$\begin{aligned}
\tilde{\mathcal{L}}_{\theta,T}(\lambda) &= \left.\frac{\partial\mathcal{L}_T(\psi,\lambda)}{\partial\theta}\right|_{\psi=\tilde{\psi}_T} = -\frac{1}{\tilde{\sigma}^2_{0T}}\sum_{t=[\lambda T]+1}^{T}(\log\Delta_t\tilde{\varepsilon}_t)\,\tilde{\varepsilon}_t\\
\tilde{\mathcal{L}}_{\boldsymbol{\beta},T}(\lambda) &= \left.\frac{\partial\mathcal{L}_T(\psi,\lambda)}{\partial\boldsymbol{\beta}}\right|_{\psi=\tilde{\psi}_T} = -\frac{1}{\tilde{\sigma}^2_{0T}}\sum_{t=[\lambda T]+1}^{T}\begin{pmatrix}\tilde{\alpha}_T^{-1}(L)\,\tilde{\varepsilon}_{t-1}\\ \cdots\\ \tilde{\alpha}_T^{-1}(L)\,\tilde{\varepsilon}_{t-p}\end{pmatrix}\tilde{\varepsilon}_t\\
\tilde{\mathcal{L}}_{\nu,T}(\lambda) &= \left.\frac{\partial\mathcal{L}_T(\psi,\lambda)}{\partial\nu}\right|_{\psi=\tilde{\psi}_T} = \frac{1}{\tilde{\sigma}^2_{0T}}\sum_{t=[\lambda T]+1}^{T}(\tilde{\alpha}_T(L)\,\Delta^{\tilde{d}_{0T}}_{t-[\lambda T]}1)\tilde{\varepsilon}_t.
\end{aligned}$$

In the previous expressions, $\log\Delta_t\tilde{\varepsilon}_t = -\sum_{j=1}^{t-1}j^{-1}\tilde{\varepsilon}_{t-j}$ depends on the restricted residuals $\tilde{\varepsilon}_t$ which are defined as follows

$$\tilde{\varepsilon}_t = \varepsilon_t\left(\tilde{\psi}_T\right) = \tilde{\alpha}_T(L)\,\Delta_t^{\tilde{d}_{0T}}(y_t-\tilde{\mu}_{0T}),\quad t=1,2,\ldots,T, \tag{4}$$

while their corresponding variance estimator is given by

$$\tilde{\sigma}^2_{0T} = \frac{1}{T}\sum_{t=1}^{T}\tilde{\varepsilon}_t^2. \tag{5}$$

As mentioned above, the restricted estimates of the parameters required to compute $\tilde{\varepsilon}_t$ result from minimizing the conditional sum of squares (CSS) over the whole sample,

$$(\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}_{0T},\tilde{\mu}_{0T}) = \arg\min_{d\in D,\boldsymbol{\alpha}\in A,\mu}\sum_{t=1}^{T}\left(\alpha(L)\,\Delta_t^d(y_t-\mu)\right)^2. \tag{6}$$

The properties of $\tilde{d}_{0T}$ have been discussed, *inter alia*, in Chung and Baillie (1993), Robinson (2006) and Hualde and Robinson (2011) in models without drift ($\mu=0$), and in Hualde and Nielsen (2020) when there is a drift.[10]

[9] See below for why the whole sample, rather than the first subsample, is chosen, and for further details on the estimation procedure.

[10] In particular, Hualde and Robinson (2019) show that the estimators $\tilde{d}_{0T}$ and $\tilde{\boldsymbol{\alpha}}_{0T}$ are $T^{1/2}$-consistent and asymptotically normal for $d_0\in Int(D)$ and $\boldsymbol{\alpha}_0\in Int(A)$, while $\tilde{\mu}_{0T}$ is $T^{1/2-d_0}$-consistent.

The relevant block of the inverse Hessian matrix concerning the subset of parameters $\left(\theta,\boldsymbol{\beta}',\nu\right)$ of ψ can be approximated (using the arguments in the proof of Theorem 1 below) as

$$\left[\left(\left.\frac{\partial^2\mathcal{L}_T\left(\psi,\lambda\right)}{\partial\psi\partial\psi'}\right|_{\psi=\tilde{\psi}_T}\right)^{-1}\right]_{[1:(2+p),1:(2+p)]} = \widetilde{P}^{-1/2}\left(\left.\frac{\partial^2\mathcal{L}_T\left(\psi,\lambda\right)}{\partial\left(\theta,\boldsymbol{\beta}',\nu\right)'\partial\left(\theta,\boldsymbol{\beta}',\nu\right)}\right|_{\psi=\tilde{\psi}_T}\right)^{-1}\widetilde{P}^{-1/2}\left(1+o_p\left(1\right)\right),$$

where $\widetilde{P}_T$ is a scaling matrix defined as

$$\widetilde{P}_T = P_T\left(\lambda;\tilde{d}_{0T}\right) = \begin{pmatrix} \lambda\cdot\mathbf{I}_{p+1} & 0 \\ 0 & \frac{L_T\left(\tilde{d}_{0T};\lambda,\lambda\right)-L_T^2\left(\tilde{d}_{0T};0,\lambda\right)}{L_T\left(\tilde{d}_{0T};\lambda,\lambda\right)} \end{pmatrix}, \tag{7}$$

which captures the effect of replacing the unknown values of d_0, $\boldsymbol{\alpha}_0$ and μ_0 by their (restricted) estimates, where $L_T(d;a,b) = T^{2d-1}\left(1-2d\right)\Gamma^2\left(1-d\right)\sum_{t=[\max(a,b)T]+1}^{T}(\Delta_{t-[aT]}^{d}1)(\Delta_{t-[bT]}^{d})$ and $\mathbf{I}_{p+1}$ is a $p+1$ dimensional identity matrix.

As in Breitung and Hassler (2002) test, the previous $\widetilde{LM}_T\left(\lambda\right)$ test statistic has a simple regression model representation, where the underlying regression provides a linearization of the true model under local alternatives involving parameter breaks. The regression-based version of the LM test is equal to T times the coefficient of determination $\tilde{R}_T^2\left(\lambda\right)$ in a linear regression model of the restricted residuals $\tilde{\varepsilon}_t$ on the scores of the general model, namely,

$$\tilde{\varepsilon}_t = \eta_0 + \eta_\lambda'\tilde{Z}_t^{(p)}\left(\lambda\right) + \eta_Z'\tilde{Z}_t^{(p)}\left(0\right) + error_t, \tag{8}$$

with the vector $\tilde{Z}_t^{(p)}\left(\lambda\right) = Z_t^{(p)}\left(\lambda,\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}_{0T},\tilde{\mu}_{0T}\right)$ being defined as

$$Z_t^{(p)}\left(\lambda,d,\boldsymbol{\alpha},\mu\right) = \begin{pmatrix} R_t\left(\lambda\right)\log\Delta_t\varepsilon_t\left(0,\mathbf{0}',0,d,\boldsymbol{\alpha},\mu\right) \\ \left\{R_t\left(\lambda\right)\Delta_{t-j}^{d}\left(y_{t-j}-\mu\right)\right\}_{j=1}^{p} \\ \alpha\left(L\right)\Delta_t^{d}R_t\left(\lambda\right) \end{pmatrix},$$

where the role of the regressor $\tilde{Z}_t^{(p)}\left(0\right) = Z_t^{(p)}\left(0,\tilde{d}_{0T},\tilde{\boldsymbol{\alpha}}_{0T},\tilde{\mu}_{0T}\right)$ is to control for the effect on the test of having estimated parameters under the null of no breaks.

Finally, when the break date is taken to be unknown, the corresponding LM test becomes

$$\widetilde{LM}_T\left(\tilde{\lambda}_T\right) = \sup_{\lambda\in\Lambda}\widetilde{LM}_T\left(\lambda\right),$$

where, as defined above, $\Lambda = [\epsilon, 1-\epsilon]$, and $\tilde{\lambda}_T = \arg\max_{\lambda\in\Lambda}\widetilde{LM}_T\left(\lambda\right).$

4 Regression-based LMW-type tests

As an alternative to the LM test based only on the restricted ML estimates, an LMW-type test (based on an auxiliary regression using information under the alternative) can be derived along the lines of Lobato and Velasco (2007). Building upon previous results by Dolado *et al.* (2002), Lobato and Velasco (2007) derive an Efficient Fractional Dickey Fuller (EFDF) test for the null hypothesis of $d = 1$ against the alternative of $d < 1$, which was later extended by Dolado *et al.* (2009) by allowing the null to be any memory $d = d_0$ against the alternative $d \neq d_0$. These authors show that, despite their asymptotic equivalence under local alternatives, the EFDF test could achieve higher power under fixed alternatives than the conventional LM test because it provides a better approximation of the DGP in such a case (see Remark 6 below). In line with these findings, our strategy here is to propose a similar test statistic designed to allow for joint breaks in $d, \boldsymbol{\alpha}$ and μ, where the null hypothesis is given by (H0) above.

For simplicity, we start with the case where the break date λ_0 and the parameters $d_0, \boldsymbol{\alpha}_0$ and μ_0 are all assumed to be known. Thus, under shifts, the data for $t = [T\lambda_0] + 1, \ldots, T$ is generated by $\alpha_1(L)\Delta_t^{d_1}(y_t - \mu_t) + \varepsilon_t$, which satisfies

$$\begin{aligned}\alpha_0(L)\Delta_t^{d_0}(y_t - \mu_0) &= \alpha_0(L)\Delta_t^{d_0}(y_t - \mu_0) - \alpha_1(L)\Delta_t^{d_1}(y_t - \mu_t) + \varepsilon_t \\ &= \alpha_0(L)\left[1 - \Delta_t^{\theta_0}\right]\Delta_t^{d_0}(y_t - \mu_0) + \beta_0(L)\Delta_t^{d_1}(y_t - \mu_0 - \nu_0 R_t(\lambda_0)) \\ &\quad + \nu_0\alpha_0(L)\Delta_t^{d_1}R_t(\lambda_0) + \varepsilon_t,\end{aligned}$$

where recall that $d_1 = d_0 + \theta_0$, $\alpha_1(L) = \alpha_0(L) + \beta_0(L)$ and $\mu_1 = \mu_0 + \nu_0$ with $\mu_t = \mu_0 + \nu_0 R_t(\lambda_0)$ and $\Delta_t^d R_t(\lambda) = \sum_{j=0}^{t-1} 1(j < t - [T\lambda])\pi_j(d) = \sum_{j=0}^{t-[T\lambda]-1}\pi_j(d) = \pi_{t-[T\lambda]-1}(d-1)$. Then, a test for the joint null of $\left(\theta_0, \boldsymbol{\beta}_0', \nu_0\right) = 0$ in (H0) can be constructed by means of testing the following null hypothesis

$$H_0 : \vartheta_1 = \vartheta_2 = \cdots = \vartheta_{2+p} = 0$$

in a regression given by

$$\begin{aligned}\alpha_0(L)\Delta_t^{d_0}(y_t - \mu_0) &= \vartheta_1\alpha_0(L)\left[\frac{1 - \Delta_t^{\theta_0 R_t(\lambda)}}{\theta_0}\right]\Delta_t^{d_0}(y_t - \mu_0) \\ &\quad + \sum_{j=1}^{p}\vartheta_{j+1}R_{t-j}(\lambda)\Delta_{t-j}^{d_1}(y_{t-j} - \mu_0 - \nu_0 R_{t-j}(\lambda)) \\ &\quad + \vartheta_{p+2}\alpha_0(L)\Delta_t^{d_1}R_t(\lambda) + \varepsilon_t,\end{aligned} \tag{9}$$

for $t = 1, \ldots, T$, and each λ. Denoting $\Theta = (\vartheta_1, \vartheta_2 \ldots, \vartheta_{2+p})'$, $\varepsilon_t^0 = \alpha_0(L)\Delta_t^{d_0}(y_t - \mu_0)$, we define $X_t^{(p)}(\lambda) = X_t^{(p)}(\lambda, \theta, \nu, d, \boldsymbol{\alpha}', \mu)$ for each $(\lambda, \theta, \nu, d, \boldsymbol{\alpha}', \mu)$,

$$X_t^{(p)}\left(\lambda, \theta, \nu, d, \boldsymbol{\alpha}', \mu\right) = \begin{pmatrix} \alpha(L)\left[\frac{1-\Delta_t^{\theta R_t(\lambda)}}{\theta}\right]\Delta_t^d(y_t - \mu) \\ \left\{R_t(\lambda)\Delta_{t-j}^{d+\theta}(y_{t-j} - \mu - \nu R_{t-j}(\lambda))\right\}_{j=1}^{p} \\ \alpha(L)\Delta_t^{d+\theta}R_t(\lambda)\end{pmatrix},$$

so that regression (9) can be rewritten in a more compact way as

$$\varepsilon_t^0 = \Theta' X_t^0(\lambda) + \varepsilon_t, \tag{10}$$

with $X_t^0(\lambda) = X_t^{(p)}(\lambda, \theta_0, \nu_0, d_0, \boldsymbol{\alpha}_0', \mu_0)$.

Remark 6. Note that the implementation of the LMW-type test based upon regression (10) is closely related to the regression-based version of the LM test since the artificial regressor $Z_t^{(p)}(\lambda)$ used in (8) corresponds to the limit of $X_t^{(p)}(\lambda)$ as $\theta, \nu \to 0$.[11] Hence, while the regression-based LMW-type test provides an exact representation of the DGP under local and fixed alternatives, the LM test only provides an accurate approximation under local alternatives. As a result, the limiting behaviour of both tests will be identical under local alternatives but it will differ under fixed alternatives (i.e. when $\theta, \nu \nrightarrow 0$).

Remark 7. As pointed out in Remark 5, the LM test fails to provide direct information on whether the choice of the lag length p ensures i.i.d. innovations under fixed alternatives. However, this is not a problem for the LMW-type test which, by making use of additional information under the alternative, is able to yield i.i.d. residuals under *both* the null and the alternative, leading to potential power gains relative to the LM test. Subsection 6.1 below provides further details on how to choose p when implementing the LMW-type test.

Under the more realistic assumption of unknown $d_0, \boldsymbol{\alpha}_0$ and μ_0, running regression (10) requires the estimation of these parameters (on top of θ and ν). With regard to $d_0, \boldsymbol{\alpha}_0$ and μ_0, our suggestion is to use the restricted estimates $\tilde{d}_{0T}$, $\tilde{\boldsymbol{\alpha}}_{0T}$ and $\tilde{\mu}_{0T}$ obtained from minimizing (6) under the null with observations for the whole sample. This facilitates comparisons with the LM test that uses the entire sample to compute these estimates under the null.[12] As regards the estimation of θ and ν, one can set $\hat{\theta}_T(\lambda) = \hat{d}_{1T}(\lambda) - \tilde{d}_{0T}$ and $\hat{\nu}_T(\lambda) = \hat{\mu}_{1T}(\lambda) - \tilde{\mu}_{0T}$ where $\hat{d}_{1T}(\lambda)$ and $\hat{\mu}_{1T}(\lambda)$ are this time the CSS estimates using observations from the second subsample. Hence, from (11), this procedure implies the following feasible regression representation of the LMW-type test

$$\tilde{\varepsilon}_t = \Theta' \tilde{X}_t^{(p)}(\lambda) + e_t \tag{11}$$

where the restricted residuals $\tilde{\varepsilon}_t = \tilde{\alpha}_{0T}(L) \Delta_t^{\tilde{d}_{0T}} (y_t - \tilde{\mu}_{0T})$ are regressed on $\tilde{X}_t^{(p)}(\lambda) = X_t^{(p)}(\lambda,$ $\hat{\theta}_T(\lambda), \hat{\nu}_T(\lambda), \tilde{d}_{0T}, \tilde{\boldsymbol{\alpha}}_{0T}, \tilde{\mu}_{0T})$.

Testing for breaks in all the parameters corresponds to testing the joint null hypothesis of $\vartheta_1 = \vartheta_2 = \cdots = \vartheta_{2+p} = 0$ in (11). Likewise, testing for a break in only a subset of the parameters can be easily accommodated. For example, a test for a break only in both memory and short-run dynamics (resp. only in μ) corresponds to testing the null hypothesis of $\vartheta_1 = \cdots = \vartheta_{1+p} = 0$ (resp. $\vartheta_{2+p} = 0$).[13] Thus, from regression (11), the LMW-type test statistic for the joint hypothesis

[11] As pointed out in LV (2007), notice that, for $\theta \to 0$, the filter $\left[\frac{1-\Delta_t^\theta}{\theta}\right]$ becomes $-\log \Delta_t$ when $\theta \to 0$, which corresponds to the well-known lag filter $\sum_{k=1}^{t-1} k^{-1} L^k$ used in the regression-based LM test.

[12] In addition, as will be illustrated in our simulation study, the size in finite samples of the LMW-type test becomes closer to the nominal size when the whole sample is used.

[13] As before, note that if only a subset of the parameters is assumed to shift, a test not allowing for a

$H_0 : \Theta = 0$ is defined as

$$\widetilde{LMW}_T(\lambda) = \tilde{\Theta}_T(\lambda)' \tilde{V}_T^{-1}(\lambda) \tilde{\Theta}_T(\lambda), \tag{12}$$

where $\tilde{\Theta}_T(\lambda) = \left(\tilde{\vartheta}_{1T}(\lambda), \tilde{\vartheta}_{2T}(\lambda), ..., \tilde{\vartheta}_{2+pT}(\lambda)\right)'$ denotes the LS estimate of Θ, while

$$\tilde{V}_T(\lambda) = \hat{\sigma}_T^2(\lambda) \widetilde{P}_T^{1/2} \left(\sum_{t=1}^{T} \tilde{X}_t^{(p)}(\lambda) \tilde{X}_t^{(p)}(\lambda)' \right)^{-1} \widetilde{P}_T^{1/2}$$

denotes its variance estimate.[14] The error variance $\hat{\sigma}_T^2(\lambda)$ is obtained as

$$\hat{\sigma}_T^2(\lambda) = \frac{1}{T} \sum_{t=1}^{[T\lambda]} \left(\hat{\alpha}_{0T}(L) \Delta_t^{\hat{d}_{0T}} (y_t - \hat{\mu}_{0T}) \right)^2 + \frac{1}{T} \sum_{t=[T\lambda]+1}^{T} \left(\hat{\alpha}_{1T}(L) \Delta_{t-[T\lambda]}^{\hat{d}_{1T}} (y_t - \hat{\mu}_{1T}) \right)^2 \tag{13}$$

which corresponds to estimation under the alternative, as the LS estimates $\tilde{\Theta}_T(\lambda)$ of regression (11) reproduce $\left(\hat{\theta}_T(\lambda), \hat{\boldsymbol{\alpha}}_{1T}(\lambda) - \hat{\boldsymbol{\alpha}}_{0T}, \hat{\nu}_T(\lambda)\right)$ to match the CSS estimators $\left(\hat{d}_{0T}, \hat{\boldsymbol{\alpha}}_{0T}, \hat{\mu}_{0T}\right)$ and $\left(\hat{d}_{1T}, \hat{\boldsymbol{\alpha}}_{1T}, \hat{\mu}_{1T}\right)$ for the first and second subsamples, respectively. As a result, the regression-based LMW-type test can be computed as T times the R^2 of regression (11) augmented by $\tilde{X}_t^{(p)}(0)$ to account for the estimation effect of the different parameters required to compute (12).

Finally, as with the LM test, the LMW-type test statistic for unknown break date becomes

$$\sup_{\lambda \in \Lambda} \widetilde{LMW}_T(\lambda) = \widetilde{LMW}_T(\tilde{\lambda}_T),$$

where $\tilde{\lambda}_T = \arg\max_{\lambda \in \Lambda} \widetilde{LMW}(\lambda)$.

5 Asymptotic properties of LM tests

5.1 Asymptotic theory of LM tests under local alternatives

We next derive the asymptotic distributions of the proposed $\widetilde{LM}_T$ tests under the following set of assumptions:

Assumption 1. The true lag length p of the stable short-run dynamics AR polynomial $\alpha(L)$ is known.

break in the non-tested parameter again should enjoy better finite sample properties (e.g. setting $\mu_0 = \mu_1$ or $\nu = 0$ in (9) when testing for a break in the dynamics, that is $H_0 : \vartheta_1 = \vartheta_2 = ... = \vartheta_{1+p} = 0$).

[14] In the case when $(d_0, \boldsymbol{\alpha}_0, \mu_0)$ are taken as known, it follows from the discussion in Wooldridge (1990) and LV (2007) that the estimation of (θ, ν) by $\left(\hat{\theta}_T(\lambda), \hat{\nu}_T(\lambda)\right)$ does not affect the limiting distribution of the LMW-type test in (10) under the null. However, this is no longer true when $(d_0, \boldsymbol{\alpha}_0, \mu_0)$ need to be estimated since these estimates affect the dependent variable in regression (11), increasing the variance of the LMW-type test statistic. This is reflected by the need to pre- and post-multiply by $\widetilde{P}_T^{1/2}$ in the definition of $\tilde{V}_T(\lambda)$ compared to the usual least-squares expression.

Assumption 2. $\varepsilon_t \sim i.i.d.\left(0, \sigma_0^2\right)$ with q moments such that $q > \max\{4, \frac{2}{1-2d_0}\}$.

Assumption 3. $d_0 \in Int\left(D\right)$, $D = \left[\underline{d}, \bar{d}\right]$, $-0.5 < \underline{d} < \bar{d} < 0.5$, $\boldsymbol{\alpha}_0 \in Int\left(\mathbf{A}\right)$, where $\mathbf{A}$ is a compact set, and $\lambda_0 \in \Lambda$.

As pointed out in Remark 5 above, Assumption 1 is adopted here for convenience: it just to facilitates a direct comparison of the asymptotic distributions of the LM and LMW-type tests, which will be shown to be identical under the null and local alternatives (see Theorem 2 and Proposition 5 further below).[15] By contrast, this assumption becomes redundant when deriving the properties of the LMW-type test, which provides a consistent data-driven choice of p under the null and alternative (see Subsection 6.1 below). For the specific case of $d_0 \in Int\left(D\right)$, Assumptions 2 and 3 are equivalent to the conditions required by Hualde and Nielsen (2020) in their treatment of the more general setup where d_0 lies in a compact set which can be arbitrarily large, but where no breaks are considered. Lastly, as in in Marinucci and Robinson (2000), Assumption 3 reflects that at least four moments are required to prove tightness for weak convergence.

To derive the asymptotic null distribution and local power of the LM test, we analyze its properties under the following sequence of local-break alternatives,

$$H_{1,T}^{d,\alpha,\mu}\left(\lambda_0\right) : \left(\theta_0, \boldsymbol{\beta}_0', \nu_0\right) = \left(\delta/T^{1/2}, \boldsymbol{\gamma}'/T^{1/2}, \eta/T^{1/2-d_0}\right), \tag{14}$$

for some $\lambda_0 \in \Lambda$, where $\boldsymbol{\gamma} = \left(\gamma_1, \ldots, \gamma_p\right)'$, and the null is recovered by setting $(\delta, \boldsymbol{\gamma}', \eta) = 0$ while leaving λ_0 unspecified.

We next derive the asymptotic distribution of the $\widetilde{LM}_T$ test in (3) in the case of an unknown break fraction λ, which is a function of both standard Brownian Motion (BM) and fractional BM (fBM).

Let $\kappa = (\kappa_1, \ldots, \kappa_p)'$ with $\kappa_k = \sum_{j=k}^{\infty} j^{-1} c_{j-k}, k = 1, \ldots, p$, where the c_j are the coefficients of L^j in the expansion of $1/\alpha_0\left(L\right)$, and $\Phi = \left\{\Phi_{k,j}\right\}, \Phi_{k,j} = \sum_{t=0}^{\infty} c_t c_{t+|k-j|}$, $k, j = 1, \ldots, p$, denotes the Fisher information matrix for $\boldsymbol{\alpha}$ under Gaussianity. Further, let

$$\varpi_p\left(\lambda, \delta, \boldsymbol{\gamma}\right) = \Xi^{1/2}\left\{B_{p+1}\left(\lambda\right) - \lambda B_{p+1}\left(1\right)\right\} + \Xi\begin{pmatrix}\delta \\ \boldsymbol{\gamma}\end{pmatrix}\left(\lambda\left(1-\lambda_0\right) - \left(\lambda - \lambda_0\right)_+\right), \ \Xi = \begin{pmatrix}\pi^2/6 & \kappa' \\ \kappa & \Phi\end{pmatrix},$$

where B_{p+1} is a $(p+1)$-dimensional standardized BM.

Next, define

$$\mathcal{A}_{d,p}^0\left(\lambda, \delta, \boldsymbol{\gamma}\right) = \frac{1}{\lambda\left(1-\lambda\right)}\varpi_p\left(\lambda, \delta, \boldsymbol{\gamma}\right)' \Xi^{-1} \varpi_p\left(\lambda, \delta, \boldsymbol{\gamma}\right)$$

and

$$\mathcal{A}_\mu^0\left(d_0, \lambda, a\right) = \frac{\left(\mathcal{W}_{d_0}\left(1-\lambda\right) - L\left(d_0; 0, \lambda\right)\mathcal{W}_{d_0}\left(1\right) + a\left(L\left(d_0; \lambda, \lambda_0\right) - L\left(d_0; 0, \lambda_0\right)L\left(d_0; 0, \lambda\right)\right)\right)^2}{L\left(d_0; \lambda, \lambda\right) - L^2\left(d_0; 0, \lambda\right)},$$

[15] In addition, this assumption helps define the robustified versions of the two tests to the true source of the break under the alternative (see Subsection 5.2.1 below).

where $L(d;a,b) = (1-2d)\int_{\max(a,b)}^{1}(s-a)^{-d}(s-b)^{-d}ds$ (so that $L(d;a,a) \equiv (1-a)^{1-2d}$ and $L(0;a,b) = 1-\max(a,b)$) and $\mathcal{W}_{d_0}(\lambda) = (1-2d_0)^{1/2}\int_0^{\lambda}(\lambda-s)^{-d_0}dB(s)$ is the standard fBM (so that $\mathcal{W}_{d_0}(1)$ has unit variance and $\mathrm{Cov}(\mathcal{W}_{d_0}(a), \mathcal{W}_{d_0}(b)) = (1-2d_0)\int_0^{\min(a,b)}(a-s)^{-d_0}(b-s)^{-d_0}ds = L(d_0;1-a,1-b)$).

Then, the following result holds.

Theorem 1 *With an unknown break fraction λ_0, under Assumptions 1, 2 and 3 and $H_{1T}(\lambda_0)$,*

$$\sup_{\lambda\in\Lambda}\widetilde{LM}_T(\lambda) \xrightarrow{d} \sup_{\lambda\in\Lambda}\left\{\mathcal{A}_{d,p}^0(\lambda,\delta,\boldsymbol{\gamma}) + \mathcal{A}_{\mu}^0\left(d_0,\lambda,\frac{\eta/\sigma_0}{\sqrt{1-2d_0}\Gamma(1-d_0)}\right)\right\},$$

where the two terms on the right hand side are independent.

Remark 8. The asymptotic distribution of the sup-$\widetilde{LM}_T(\lambda)$ test under H_0 is then given by

$$\sup_{\lambda\in\Lambda}\left\{\mathcal{A}_{d,p}^0(\lambda,0,\mathbf{0}) + \mathcal{A}_{\mu}^0(d_0,\lambda,0)\right\}.$$

Since the break fraction is not identified under the null of no structural breaks, the distribution above is non-standard. Besides, it only depends on d_0, but not on $\lambda_0, \boldsymbol{\alpha}_0$ or μ_0. Critical values of such limiting distribution are reported in Table 1 for a grid of values of d_0 and ϵ generated as in Theorem 1 above, using 10,000 grid points for the break fraction and 100,000 simulations. To compute the critical values for an unknown d_0, we interpolate between these values and replace d_0 by $\tilde{d}_{0T}$ as in (6) (see Giraitis *et al.* (2006) for a similar solution).

[Table 1 about here]

Remark 9. Under local alternatives, the two components $A_{d,p}^0$ and A_{μ}^0 in the asymptotic distribution of the sup-$\widetilde{LM}_T(\lambda)$ test capture the contributions of the local shifts of the dynamics parameters and of the level, respectively.[16] It is noteworthy that, while the term $\mathcal{A}_{d,p}^0(\lambda,\delta,\boldsymbol{\gamma})$ is symmetric around the break fraction $\lambda_0 = 0.5$, the term $\mathcal{A}_{\mu}^0\left(d_0,\lambda,\eta/(\sigma_0\sqrt{1-2d_0}\Gamma(1-d_0))\right)$ happens to be positively (resp. negatively) skewed if $d_0 > 0$ (resp. $d_0 < 0$). Hence, when there is only a break in $(d,\boldsymbol{\alpha}')$, the local power of the sup-$\widetilde{LM}_T(\lambda)$ test is maximized for $\lambda_0 = 0.5$. In

[16] Note that the limit term $A_{d,p}^0$ is similar to that obtained by Horváth and Shao (1999) in their test for a break only in d using a LR test from Whittle estimation. Likewise A_{μ}^0 is similar to the limit term derived by Iacone et al. (2013) for a break only in μ in the first-differenced version of their model, designed to test for a break in the linear trend of a I(d) process under any memory using Abadir et al.'s (2007) Extended Local Whittle estimation. Thus, our result generalizes theirs by allowing for joint breaks in both d and μ, plus in the short-run dynamics.

contrast, if there were either only breaks in μ or in both $(d, \boldsymbol{\alpha}')$ and μ, then local power would be maximized for some $\lambda_0 < 0.5$ (resp. $\lambda_0 > 0.5$) if $d_0 > 0$ (resp. $d_0 < 0$).

Theorem 1 also nests the special cases where one exclusively tests for a break in dynamics or in the level, *i.e.* only in d and $\boldsymbol{\alpha}$ (so that A_μ^0 drops) or only in μ (so that $A_{d,p}^0$ drops), reflecting that these two tests are asymptotically independent under local alternatives. However, note that if *only* a subset of the parameters breaks, a testing procedure which does not allow for a break in the other parameters could lead to better power properties in finite samples. Nonetheless, estimation of the model under this null could yield misleading conclusions when the tested parameter happens to be stable while the other parameters are the ones that actually shift. We analyze this last issue in Subsection 5.2.1 below, where a robustified version of the test to the behaviour of the non-tested parameters in the DGP is provided.

Lastly, Corollary 1 below provides the asymptotic distribution of the $\widetilde{LM}_T$ test when the break fraction λ_0 is assumed to be known.

Corollary 1 *With known break fraction λ_0, under Assumptions 1, 2 and 3, and hypothesis $H_{1T}(\lambda_0)$,*

$$\widetilde{LM}_T(\lambda_0) \xrightarrow{d} \chi^2_{2+p}(c(\lambda_0)),$$

with non-centrality parameter

$$c(\lambda_0) = \omega_p^2(\delta, \boldsymbol{\gamma})\, \lambda_0 (1 - \lambda_0) + \frac{\eta^2}{\sigma_0^2} \frac{L(d_0; \lambda_0, \lambda_0) - L^2(d_0; 0, \lambda_0)}{(1 - 2d_0)\, \Gamma^2(1 - d_0)} \equiv c_{d,\boldsymbol{\alpha}}(\lambda_0) + c_\mu(\lambda_0),$$

where $\omega_p^2(\delta, \boldsymbol{\gamma}) = (\delta \quad \boldsymbol{\gamma}')\, \Xi\, (\delta \quad \boldsymbol{\gamma}')'$.

As expected, when λ_0 is known, the asymptotic distribution becomes a chi-square with $2 + p$ degrees of freedom, where the non-centrality parameter $c(\lambda_0)$ depends on the two drifts under local alternatives, namely $c_{d,\boldsymbol{\alpha}}(\lambda_0)$ and $c_\mu(\lambda_0)$. Moreover, as in the case of unknown λ_0, Corollary 1 nests the cases of testing for a break in a subset of the parameters: (i) if one tests for a break only in $(d, \boldsymbol{\alpha})$, the limiting distribution becomes $\chi^2_{1+p}(c_{d,\boldsymbol{\alpha}}(\lambda_0))$, where c_μ drops and (ii) if one tests for a break only in μ, the limiting distribution becomes $\chi^2_1(c_\mu(\lambda_0))$, where $c_{d,\boldsymbol{\alpha}}$ drops.

The next corollary shows that the results in Theorem 1 and Corollary 1 on size and local power under DGP (1) also hold under the DGP-MA discussed in Remark 3.

Corollary 2 *The conclusions of Theorem 1 and Corollary 1 also hold for data generated under the DGP-MA discussed in Remark 3.*

Remark 10. As mentioned earlier, Martins and Rodriguez (2014) and Hassler and Meller (2014) have proposed similar LM test statistics for a break in d in $I(d)$ processes, but under the assumption of a known memory parameter (d_0) in the first subsample. In such an instance, the variance of the test statistic would be smaller than when d_0 is unknown, resulting in a higher local power. Yet, since the assumption of known d_0 is quite restrictive in practice, they suggest some estimators of the memory parameter. Martins and Rodriguez (2014) plug in a parametric estimator of d to derive the asymptotic distribution of the corresponding LM test statistic. However, their approximation may not be accurate enough since it ignores the covariance between the test statistic and the estimator under the null. Hassler and Meller (2014) plug in a semiparametric estimator for d_0 but without deriving the limiting distribution of their LM test which they note could differ from the corresponding distribution under known d_0, due to the lower rate of convergence of their proposed estimator.

5.2 Consistency of LM tests

In this section we prove the consistency of the LM test for breaks in either all or a subset of the parameters. In particular, as regards the $\widetilde{LM}_T$ test for the null $H_0 : \left(\theta_0, \boldsymbol{\beta}_0', \nu_0\right) = \mathbf{0}$, we consider the following set of fixed alternative hypotheses:

$$
\begin{aligned}
H_1^{d,\alpha}(\lambda_0) &: \left(\theta_0, \boldsymbol{\beta}_0'\right)' \neq \mathbf{0} \text{ and } \nu_0 = 0, \\
H_1^{\mu}(\lambda_0) &: \left(\theta_0, \boldsymbol{\beta}_0'\right)' = \mathbf{0} \text{ and } \nu_0 \neq 0, \\
H_1^{d,\alpha,\mu}(\lambda_0) &: \left(\theta_0, \boldsymbol{\beta}_0'\right)' \neq \mathbf{0} \text{ and } \nu_0 \neq 0.
\end{aligned}
$$

where the superscripts in $H_1^{d,\alpha}(\lambda_0)$, $H_1^{\mu}(\lambda_0)$ and $H_1^{d,\alpha,\mu}(\lambda_0)$ denote, respectively, alternatives with: (i) only a break in $(d, \boldsymbol{\alpha})$, (ii) only a break in μ, and (iii) joint breaks in $(d, \boldsymbol{\alpha})$ and μ.[17] Under the corresponding alternative hypotheses, the following result holds:

Proposition 1 *Under Assumptions 1, 2 and 3, then:*
The LM test statistic for a break in all parameters, $\widetilde{LM}_T(\lambda_0)$ *and* $\sup_\lambda \widetilde{LM}_T(\lambda)$, *diverge: (i) at rate* T *under either* $H_1^{d,\alpha,\mu}(\lambda_0)$ *or* $H_1^{d,\alpha}(\lambda_0)$, *and (ii) at rate* T^{1-2d_0} *(resp.* T*) under* $H_1^{\mu}(\lambda_0)$ *with* $d_0 \geq 0$ *(resp.* $d_0 < 0$*).*

Remark 11. As anticipated above, it is important to highlight that the use of individual $\widetilde{LM}_T$ tests for breaks in a subset of the parameters – either $(d, \boldsymbol{\alpha})$ or μ– may lead to spurious rejections when the non-tested subset happens to be the only one shifting. The next subsection is devoted to analyze this issue in further detail.

As in Corollary 2, the next corollary shows that the results in Proposition 1 on consistency under DGP (1) can be extended to DGP-MA.

[17] To save space we do not consider here alternatives involving breaks only in d or α or joint breaks in $(\mu, \boldsymbol{\alpha})$, whose testing strategy would be similar to the one used for the three cases discussed in this section.

Corollary 3 *The conclusions of Proposition 1 also hold for data generated under the DGP-MA discussed in Remark 3.*

5.2.1 Robustified LM test

Whenever the joint $\widetilde{LM}_T$ test rejects the null of parameter stability, one may be interested in identifying the specific source of the break under any of the aforementioned set of fixed alternatives. To pursue this approach we propose a robustified version of the LM test against potential breaks in the non-tested parameters, first under known break fraction, and next when it is unknown.

(a) Known break fraction

For ease of exposition, besides assuming known λ_0, we consider the simple case where there are no short-run dynamics: $\alpha_0(L) = 1$, and $\boldsymbol{\beta}_0 = \mathbf{0}$. To achieve break-source identification in this case, it is convenient to derive individual LM tests under the following two simple null hypotheses,

$$H_0^d(\lambda_0) : \theta_0 = 0,$$
$$H_0^\mu(\lambda_0) : \nu_0 = 0,$$

where, unlike the individual nulls considered earlier, no assumption is explicitly made about the stability of the other (non-tested) parameter. Then, a sequential procedure can be designed to test the above simple null hypotheses. The first step consists of testing for joint breaks in d and μ by means of the $\widetilde{LM}_T$ test defined in (3). In case of rejection, the second stage entails testing the individual null $H_0^d(\lambda_0)$ (resp. $H_0^\mu(\lambda_0)$) to check if d (resp. μ) is actually breaking, irrespective of whether the other parameter shifts or not.[18] For example, to implement a robustified test of the null $H_0^d(\lambda_0)$ against the alternative $H_1^d(\lambda_0) : \theta_0 \neq 0$, rather than using the $\widetilde{LM}_T$ test based on the score in the direction of θ with H_0-restricted estimates $(\tilde{d}_{0T}, \tilde{\mu}_{0T})$ as in (6), we recommend to use the following $H_0^d(\lambda_0)$-restricted estimates

$$(\bar{d}_{0T}, \bar{\mu}_{0T}, \bar{\nu}_{0T}) = \arg \min_{d \in D, \mu, \nu} \sum_{t=1}^{T} \left(\Delta_t^d \left(y_t - \mu - \nu R_t(\lambda_0) \right) \right)^2, \tag{15}$$

where different levels are allowed in each subsample. Then, the robust individual version of the LM test for $H_0^d(\lambda_0)$, labeled $\overline{LM}_T^d(\lambda_0)$, is given by

$$\overline{LM}_T^d(\lambda_0) = \left.\frac{\partial \mathcal{L}_T(\psi, \lambda_0)}{\partial \psi}\right|_{\psi=\bar{\psi}_T} \left(- \left.\frac{\partial^2 \mathcal{L}_T(\psi, \lambda_0)}{\partial \psi \partial \psi'}\right|_{\psi=\bar{\psi}_T} \right)^{-1} \left.\frac{\partial \mathcal{L}_T(\psi, \lambda_0)}{\partial \psi}\right|_{\psi=\bar{\psi}_T},$$

where $\bar{\psi}_T = (0, \bar{\nu}_{0T}, \bar{d}_{0T}, \bar{\mu}_{0T}, \bar{\sigma}_T^2)'$ and $\bar{\sigma}_T^2 = T^{-1} \sum_{t=1}^{T} \bar{\varepsilon}_t^2$ uses the $H_0^d(\lambda_0)$-restricted residuals $\bar{\varepsilon}_t = \varepsilon_t(\bar{\psi}_T) = \Delta_t^{\bar{d}_{0T}} (y_t - \bar{\mu}_{0T} - \bar{\nu}_{0T} R_t(\lambda_0))$. Using a similar reasoning, we can define $\overline{LM}_T^\mu(\lambda_0)$

[18] Notice that, to robustify these individual tests against misleading inference, it is preferable to remain agnostic about the behaviour of the non-nested parameter.

to test H_0^{μ} based on the corresponding $H_0^{\mu}(\lambda_0)$-restricted estimation, where this time $\bar{\psi}_T = (\bar{\theta}_{0T}, 0, \bar{d}_{0T}, \bar{\mu}_{0T}, \bar{\sigma}_T^2)'$, as well as to derive robustified LM tests when allowing for short-run dynamics captured by an AR(p) process with lag polynomial $\alpha(L)$.

The following Proposition establishes the asymptotic behaviour of the robustified LM tests considered above.

Proposition 2 *Under Assumptions 1, 2 and 3:*

(a) The robustified test statistic $\overline{LM}_T^d(\lambda_0)$ for a break only in the memory, diverges at rate T under either $H_1^{d,\mu}(\lambda_0)$ or $H_1^d(\lambda_0)$. By contrast, it converges to a χ_1^2 distribution under $H_1^{\mu}(\lambda_0)$, i.e, when only μ shifts.

(b) The robustified test statistic $\overline{LM}_T^{\mu}(\lambda_0)$ for a break only in the level, diverges: (i) at rate T^{1-2d_0} (resp. T) for $0 < d < 0.5$ (resp. $-0.5 < d < 0$) under $H_1^{\mu}(\lambda_0)$; (ii) at rate T^{1-2d_1} (resp. T) for $0 \leq d_1 < 0.5$ (resp. $-0.5 < d_1 < 0$) and $H_1^{d,\mu}(\lambda_0)$. By contrast, it converges to a χ_1^2 distribution under $H_1^d(\lambda_0)$, i.e. when only d shifts.

Upon rejection of the joint null of parameter stability in the first stage of the sequential testing procedure, Proposition 2 illustrates why the robust individual tests $\overline{LM}_T^d(\lambda_0)$ and $\overline{LM}_T^{\mu}(\lambda_0)$ in the second stage help identify which specific parameter (or parameters) actually break. The insight is that the individual test of $H_0^d(\lambda_0)$ (resp. $H_0^{\mu}(\lambda_0)$) will reject asymptotically this null under $H_1^d(\lambda_0)$ (resp. $H_1^{\mu}(\lambda_0)$) but will only exhibit trivial power under $H_1^{\mu}(\lambda_0)$ (resp. $H_1^d(\lambda_0)$). It also follows from Propositions 1 and 2 that the rates of divergence of $\widetilde{LM}_T^{d,\mu}$ and $\overline{LM}_T^{\mu}(\lambda_0)$ under $H_1^{\mu}(\lambda_0)$ depend on the value of the memory parameter in the second subsample (i.e. d_0 if memory is constant, or d_1 if it breaks).

Remark 12. A brief discussion follows on how size is controlled asymptotically in the previous sequential testing approach. As is well known, Type-I errors for the joint hypothesis pile up in multiple testing when tests of individual hypotheses are implemented after *not rejecting* the previous ones, therefore requiring Bonferroni-type corrections. However, we claim that such corrections are unnecessary here because the $\overline{LM}_T^d$ test does *control size*. In effect, (under H_1^{μ}) d would be wrongly identified as the source of the break in $100\alpha\%$ cases as $T \to \infty$, whereas, since the first stage is asymptotically correct, rejection would also happen asymptotically at most in $100\alpha\%$ cases under H_0 (that is, when none of the alternatives H_1^{μ}, H_1^d or $H_1^{d,\mu}$ hold). Thus, the probability of wrongly concluding that d is breaking is controlled in both cases (i.e. under H_1^{μ} and under H_0) while, if d truly breaks (under H_1^d or $H_1^{d,\mu}$), this would be confirmed with probability tending to 1.

(b) Unknown break date

When λ_0 is unknown, its value is replaced in the test by the estimate of the break date obtained from the first step, namely, $\widetilde{\lambda}_T = \arg\max_{\lambda\in\Lambda} \widetilde{LM}_T(\lambda)$. The next two propositions (which, to our knowledge, seem to be new in the literature on break testing), justify this procedure. In particular, we use a weak convergence condition to ensure that the restricted parameter estimates converge to an interior value of the parameter space under a fixed alternative (Proposition 3 below) and under local alternatives (Proposition 4 below). This implies that the break date can be fully pinned down asymptotically by maximizing the $\widetilde{LM}_T(\lambda)$ test statistic, so that λ_0 becomes identifiable from the restricted estimates.

Proposition 3 *Under Assumptions 1, 2, 3 and* $\left(\tilde{d}, \tilde{\boldsymbol{\alpha}}', \tilde{\mu}\right) \to_p (d_A, \boldsymbol{\alpha}'_A, \mu_A)$, *where* $(d_A, \boldsymbol{\alpha}'_A) \in Int(D \times \mathbf{A})$, $H_1^{d,\alpha}(\lambda_0)$, $H_1^{\mu}(\lambda_0)$ *or* $H_1^{d,\alpha,\mu}(\lambda_0)$, $\lambda_0 \in \Lambda$, *then* $\tilde{\lambda}_T \xrightarrow{p} \lambda_0$.

Proposition 4 *Under Assumptions 1, 2, 3 and the local alternatives*

$$H_{1,T,m}^{d,\alpha,\mu}(\lambda_0) : \left(\theta_0, \boldsymbol{\beta}'_0, \nu_0\right) = m_T \left(\delta/T^{1/2}, \boldsymbol{\gamma}'/T^{1/2}, \eta/T^{1/2-d_0}\right)$$

where $\lambda_0 \in \Lambda$, $(\delta, \boldsymbol{\gamma}', \eta) \neq 0$ *and* m_T *satisfies as* $T \to \infty$

$$\frac{1}{m_T} + \frac{m_T}{T^{1/2}} \to 0,$$

then $\tilde{\lambda}_T \xrightarrow{p} \lambda_0$.

As a result of these two propositions, the conclusions of Proposition 2 also hold for the sup-version of the LM test.[19] In fact, as will be later discussed in Section 7 below, our simulation results confirm the rather satisfactory finite sample performance of the break fraction estimators $\tilde{\lambda}_T = \arg\max_{\lambda\in\Lambda} \widetilde{LM}_T(\lambda)$ and $\tilde{\lambda}_T = \arg\max_{\lambda\in\Lambda} \widetilde{LMW}_T(\lambda)$.

6 Asymptotic properties of LMW-type tests

Using estimates $(\tilde{d}_{0T}, \tilde{\boldsymbol{\alpha}}_{0T}, \tilde{\mu}_{0T})$ in place of $(d_0, \boldsymbol{\alpha}_0, \mu_0)$, we next show that the limiting distribution of LMW-type test is equivalent to the one of the LM test under a sequence of local alternatives. Recall from Remark 6 that the insight for this equivalence result is that the filter used by the LMW-type test, $(1 - \Delta_t^{\theta})/\theta$, converges to the filter used by the LM test, $-\log \Delta_t$, when $\theta \to 0$ under local alternatives. Yet, as will be shown below, the two filters could be different when θ does not converge to zero, namely, under fixed alternatives.

[19] Note further that Proposition 3 complements the results of Rachinger's (2017) on consistent estimation of the break fraction λ_0 obtained by minimization of the conditional sum of squares (CSS). Indeed, both approaches provide asymptotically the same information on λ_0 when the model is known.

Theorem 2 *Under Assumptions 1, 2 and 3 and under the local hypothesis H_{1T}, for unknown parameters $d_0, \boldsymbol{\alpha}_0$ and μ_0 and for*

(a) an unknown break fraction λ, the asymptotic behaviour of the LMW-type test $\sup_\lambda \widetilde{LMW}_T(\lambda)$ corresponds to the one derived for the $\sup_\lambda \widetilde{LM}_T(\lambda)$ test in Theorem 1.

(b) a known break fraction λ_0, the asymptotic behaviour of the LMW-type test $\widetilde{LMW}_T(\lambda_0)$ corresponds to the one derived for the $\widetilde{LM}_T(\lambda_0)$ test in Corollary 1.

In addition, we discuss the consistency of the LMW-type test for breaks in the dynamics and/or μ under fixed alternatives, where the following result holds.

Proposition 5 *The LMW-type tests for a break in all parameters, $\widetilde{LMW}_T^{d,\alpha,\mu}(\lambda_0)$ and $\sup_\lambda \widetilde{LMW}_T^{d,\alpha,\mu}(\lambda)$, behave like the $\widetilde{LM}_T$ tests for joint breaks in Proposition 1.*

A robustified version of the LMW test can be obtained in a similar fashion as for the LM test by regressing the restricted residuals $\bar{\varepsilon}_t = \varepsilon_t\left(\bar{\psi}_T\right) = \Delta_t^{\bar{d}_{0T}}\left(y_t - \bar{\mu}_{0T} - \bar{\nu}_{0T} R_t(\lambda)\right)$ on $\tilde{X}_t^{(p)}(\lambda) = X_t^{(p)}\left(\lambda, \hat{\theta}_T(\lambda), \hat{\nu}_T(\lambda), \bar{d}_{0T}, \bar{\mu}_{0T}, \bar{\nu}_{0T}\right)$, where $(\bar{d}_{0T}, \bar{\mu}_{0T}, \bar{\nu}_{0T})$ are the restricted estimates in (15) with λ evaluated at $\tilde{\lambda}_T$.

Thus, under fixed alternatives, Proposition 5 implies that,the regression-based versions of the two tests exhibit the same rates of divergence. However, as anticipated above, their drift terms will be different, a feature which affects their relative asymptotic power. Our main finding in this respect is that the drift term of the LMW-type test is larger than the corresponding drift term of the LM test. To illustrate this result, let us consider the alternative of only a break in d at a known fraction λ_0, with $\alpha_0(L) = 1$ and $\boldsymbol{\beta}_0 = \mathbf{0}$. Then, it follows that

$$\begin{aligned} p\lim_{T\to\infty} \frac{\widetilde{LM}(\lambda)}{T} &= \frac{1-\lambda_0}{\lambda_0} C_{LM}(d_1, d_A); \\ p\lim_{T\to\infty} \frac{\widetilde{LMW}(\lambda)}{T} &= \frac{1-\lambda_0}{\lambda_0} C_{LMW}(d_1, d_A), \end{aligned}$$

where the drift terms (C_{LM} and C_{LMW}) in the previous expressions are given by

$$\begin{aligned} C_{LM}(d_1, d_A) &= \frac{\left(\sum_{j=1}^{\infty}\left(\sum_{k=1}^{j} \frac{\Gamma(j-k+d_1-d_A)}{k\Gamma(j-k+1)}\right) \frac{\Gamma(j+d_1-d_A)}{\Gamma(j+1)}\right)^2}{\frac{\bar{\sigma}^2_{d,LM}}{\sigma^2} \sum_{j=1}^{\infty}\left(\sum_{k=1}^{j} \frac{\Gamma(j-k+d_1-d_A)}{k\Gamma(j-k+1)}\right)^2}, \\ C_{LMW}(d_1, d_A) &= \frac{\Gamma(1+2(d_A-d_1))}{\Gamma^2(1+(d_A-d_1))} - 1, \end{aligned}$$

such that d_A and $\bar{\sigma}^2_{d,LM}$ are the probability limits of the restricted estimate $\widetilde{d}_{0T}$ (obtained from (6)) and the estimated variance in the $\widetilde{LM}$ test, respectively, under the alternative $H_1^d(\lambda_0)$. For the specific simplified case considered here, they can be shown to be equal to,

$$d_A = \arg\min_{d \in D_i} \left\{ \lambda_0 \frac{\Gamma\left(1-2\left(d_0-d\right)\right)}{\Gamma^2\left(d-d_0+1\right)} + \left(1-\lambda_0\right) \frac{\Gamma\left(1-2\left(d_1-d\right)\right)}{\Gamma^2\left(d-d_1+1\right)} \right\}, \tag{16}$$

$$\bar{\sigma}^2_{d,LM} = \sigma_0^2 \left(\lambda_0 \frac{\Gamma\left(1+2\left(d_A-d_0\right)\right)}{\Gamma^2\left(1+\left(d_A-d_0\right)\right)} + \left(1-\lambda_0\right) \frac{\Gamma\left(1+2\left(d_A-d_1\right)\right)}{\Gamma^2\left(1+\left(d_A-d_1\right)\right)} \right),$$

As illustrated in Figure 1, where $\lambda_0 = 0.5$, $d_0 = 0$, and $\sigma_0 = 1$, it easy to show that the drift terms above satisfy the following inequality: $C_{LMW}\left(d_1, d_A\right) > C_{LM}\left(d_1, d_A\right)$, with this difference becoming steeper when $d_1 < d_0 < 0$ (see Dolado et al, 2017). Therefore, under $H_1^d(\lambda_0)$, the LMW-type tests tend to dominate the LM tests in terms of asymptotic power due to their greater non-centrality parameters.

[**Figure 1 about here**]

6.1 Model specification for LMW-type test

To determine the value of p in practice, we need to ensure that the proposed model and the regression underlying the LMW-type test (11) are correctly specified. The key assumption to check is that the residuals are approximately i.i.d., as specified in Assumption 2. However, our method for testing residual serial correlation should account for the special features of this artificial regression, namely, that both the dependent variable and the regressors are generated variables, and that they depend on $\tilde{\lambda}_T$.

To address the parameter-estimation effect both in the definition of $\tilde{\varepsilon}_t$ and $\tilde{X}_t^{(p)}\left(\tilde{\lambda}_T\right)$ and the computation of the residuals $\hat{e}_t = \hat{e}_t\left(\tilde{\lambda}_T\right) = \tilde{\varepsilon}_t - \hat{\Theta}'\tilde{X}_t\left(\tilde{\lambda}_T\right)$, we propose to apply a Breusch-Godfrey (BG) test for the null of no serial autocorrelation in the residuals against the alternative of autocorrelation of order P (see Breusch (1979) and Godfrey (1978)). In our setup, the BG test consists of an OLS regression of the residuals $\hat{e}_t$ on its first P lags, the regressors $\tilde{X}_t^{(p)}\left(\tilde{\lambda}_T\right)$ and the residual derivatives with respect to the set of estimated parameters $\left(\tilde{d}_{0T}, \tilde{\boldsymbol{\alpha}}_{0T}, \tilde{\mu}_{0T}\right)$ in the dependent variable $\tilde{\varepsilon}_t$, i.e., $\tilde{Z}_t^{(p)}(0)$, which are similar to $\tilde{X}_t^{(p)}\left(\tilde{\lambda}_T\right)$ but without restricting the estimation to the second part of the sample. Hence, the testing dynamic residual regression

$$\hat{e}_t = \eta_0 + \eta_1 \hat{e}_{t-1} + \cdots + \eta_P \hat{e}_{t-P} + \eta_\lambda' \tilde{X}_t^{(p)}\left(\tilde{\lambda}_T\right) + \eta_Z' \tilde{Z}_t^{(p)}(0) + error_t,$$

is fitted and the LM test statistic TR_T^2 for the significance of the coefficients of $\hat{e}_{t-1}, \ldots, \hat{e}_{t-P}$, $H_0 : \eta_1 = \cdots = \eta_P = 0$, is compared to a χ_P^2 critical value. Here P should be chosen to be larger than p to be able to identify dynamics not properly described by the specified model. With and without breaks in the model, $\tilde{X}_t^{(p)}\left(\tilde{\lambda}_T\right)$ should not be significant in the residual regression, so dependence on $\tilde{\lambda}_T$ would not affect asymptotic inference. Despite the fact that lag length order selection could impact the properties of the test under some data configuration, as will be shown in our Monte Carlo simulations below, this simple procedure seems to provide good size accuracy under the null of no breaks and good power when breaks occur in some of the parameters.

6.2 Multiple breaks

An additional advantage of the LMW-type test is that it can be easily extended to allow for the presence of multiple regimes, therefore allowing for breaks in $d, \boldsymbol{\alpha}$ and μ at different periods of time. In this fashion, our maintained assumption that breaks are coincidental in time can be relaxed. We briefly sketch in the sequel how to implement the tests in this more general setup where, for notational simplicity, we consider again the case of no short-run dynamics with $\alpha(L) = 1$.

Denoting the number of regimes by m, let us consider the following DGP for $i = 0, \ldots, m-1$,

$$\Delta_t^{d_t}(y_t - \mu_t) = \varepsilon_t,\ t = [\lambda_i T] + 1, \ldots, [\lambda_{i+1} T],$$

with

$$\mu_t = \sum_{i=0}^{m-1} \mu_i R_t^{(i+1)}(\boldsymbol{\lambda})$$

where $R_t^{(i+1)}(\boldsymbol{\lambda}) = R_t^{(i+1)}(\lambda_i, \lambda_{i+1}) = 1([\lambda_i T] < t \leq [\lambda_{i+1} T]),\ \lambda_0 = 0, \lambda_m = 1,\ \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{m-1})'$ and d_t is defined similarly. For example, when testing for 0 *versus* 2 breaks (so that $m = 3$), implementation of the LMW-type test relies on the following regression model,

$$\begin{aligned}\Delta_t^{d_0}(y_t - \mu_0) &= \left(\vartheta_1 \left[\frac{1-\Delta_t^{\theta_1}}{\theta_1}\right] \Delta_t^{d_0}(y_t - \mu_0) + \vartheta_2 \Delta_t^{d_0} 1\right) R_t^{(2)}(\boldsymbol{\lambda}) \\ &\quad + \left(\vartheta_3 \left[\frac{1-\Delta_t^{\theta_2}}{\theta_2}\right] \Delta_t^{d_0}(y_t - \mu_0) + \vartheta_4 \Delta_t^{d_0} 1\right) R_t^{(3)}(\boldsymbol{\lambda}) + \varepsilon_t,\end{aligned}$$

where a test of $H_0 : \vartheta_1 = \vartheta_2 = \vartheta_3 = \vartheta_4 = 0$ corresponds to testing for two breaks in both parameters, while testing $H_0' : \vartheta_1 = \vartheta_3 = 0$ (resp. $\vartheta_2 = \vartheta_4 = 0$) is equivalent to testing for two breaks only in d (resp. μ). As for the non-coincidental breaks, e.g. testing $H_0' : \vartheta_1 = \vartheta_4 = 0$ would correspond to testing for a break in d, followed by a break in μ.

Finally, the LMW-type test defined in (9) can also be extended to a sequential testing of multiple breaks, e.g. by testing the null of k breaks (denoted $H_0^{(k)}$, with $k = 0, 1,..$) against the alternative of $k+1$ breaks (denoted $H_1^{(k+1)}$), as we carry out in the empirical application reported in Section 7. In this case, unlike before, we do recommend the use of Bonferroni conservative critical values because multiple tests are performed in many subsamples to test for a further break. In particular, if the null $H_0^{(0)}$ is rejected in favour of $H_1^{(1)}$, the sequential testing procedure implies that the same test would be applied again in each of the two different subsamples to identify further breaks, but this time using $\alpha/2$ nominal critical values to keep the size of the overall test of $H_0^{(1)}$ vs. $H_1^{(2)}$ at the right level α.

Remark 13. To test the origin of breaks in the potentially non-coincidental case, one could apply the robustified tests discussed in Section 5 at the second stage once the aforementioned sequential procedure has determined the number of breaks.

7 Finite sample simulations

In this section we report some Monte-Carlo simulation results regarding size and power of the regression-based LM and LMW-type tests in finite samples. In some simulations the break fraction will be assumed to be known while in others it will be taken to be unknown. We consider a wide range of setups, which are organized in different simulation blocks as follows.

(I) Size and power of the LM and LMW-type tests (no short-run dynamics)

In the first set of simulations we abstract from short-run dynamics (i.e. $\alpha_0(L) = 1$ and $\boldsymbol{\beta}_0 = \mathbf{0}$) and consider only shifts in d and/or μ at an unknown break fraction $\lambda \in \Lambda = [\epsilon, 1-\epsilon]$ of the sample, with $\epsilon = 0.15$. The chosen significance level is 0.05 and the sample sizes are $T \in \{200, 500, 1000\}$ regarding size, and $T = 200$ regarding power. We set an error variance $\sigma_0^2 = 1$ and take draws from a $N(0,1)$ distribution. To compute size, we consider $d_0 \in \{-0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4\}$ and $\mu_0 = 0$ while, to compute power, we consider $\lambda_0 = 0.5$, $d_0 \in \{-0.2, 0, 0.2\}$, $d_1 \in \{-0.4, -0.2, 0, 0.2, 0.4\}$, $\mu_0 = 0$ and $\mu_1 \in \{0, 0.25, 0.5\}$. The number of simulations is $10,000$.

Table 2 (panels a and b) displays the size of the two tests for joint breaks in d and μ at an unknown break fraction. The main finding is that both tests exhibit satisfactory size properties for $T = 500$ and 1000, though they can be slightly oversized for $T = 200$. Table 2 (panels c and d) displays the power results of the two tests for a break in d and/or μ at $\lambda_0 = 0.5$. The simulation results confirm that there are some power gains from using the $\widetilde{LMW}_T$ tests in finite samples.[20] As can be inspected, power for both tests is increasing in the magnitude of the shifts in d and μ. For example, looking at the second block in panel (d), for $\mu_0 = \mu_1 = 0$, a shift in d from 0 to 0.2 increases the power of the $\widetilde{LMW}_T$ test by 11.6 pp. ($= 18.2 - 6.6$) whereas, for $d_0 = d_1 = 0$, a shift in μ from 0 to 0.25 raises power by 22.7 pp. ($= 29.3 - 6.6$). The corresponding gains in power when d shifts from 0 to 0.4 (for $\mu_0 = \mu_1 = 0$) and when μ shifts from 0 to 0.5 (for $d_0 = d_1 = 0$) are

[20] We have checked whether the higher power of the $\widetilde{LMW}$-type tests relative to the $\widetilde{LM}$ test could be due to the differences in their effective sizes but size-corrected power of the former test remains higher, though to a slightly lesser extent than when the nominal size is used.

49.6 pp. and 63.0 pp., respectively. Lastly, as expected, the power arising from breaks in μ is the lower the larger d. For instance, using the shift in μ from 0 to 0.25 (this time with $d_0 = d_1 = 0.2$, instead of $d_0 = d_1 = 0$), only raises the power of the $\widetilde{LMW}_T$ test by 9.9 pp. (= 16.0 − 6.1).

[Table 2 about here]

(II) Estimates of the break fraction (no short-run dynamics)

With regard to the estimation of the break fraction when it is considered to be unknown, Figure 2 shows a second set of simulations about the finite sample performance of the break-fraction estimators $\tilde{\lambda}_T = \arg\max_{\lambda\in\Lambda} \widetilde{LM}_T(\lambda)$ and $\tilde{\lambda}_T = \arg\max_{\lambda\in\Lambda} \widetilde{LMW}_T(\lambda)$ discussed in Proposition 3 and 4. Their values have been simulated for breaks in d (from 0 to 0.4) and in μ (from 0 to 0.5) at $\lambda_0 = 0.5$, with $T \in \{200, 500\}$. As can be inspected, the distributions of these break-fraction estimates are well centered around their true value in all these simulations, and their variance decreases as the sample size increases.[21]

[Figure 2 about here]

(III) Size and power of the LM and LMW-type tests (known AR(1) short-run dynamics)

Next, we report a third set of size and power simulation results in Table 3, now allowing for known short-run dynamics captured by an AR(1) process, with all parameters potentially shifting at a known fraction $\lambda_0 = 0.5$. As regards size, we consider $T \in \{200, 500, 1000\}$ and $\mu_0 = 0$, $d_0 \in \{-0.4, -0.2, 0, 0.2, 0.4\}$, and $\alpha_0 \in \{-0.5, 0.5\}$ while for power we choose, $T = 200$, $d_0 \in \{-0.2, 0, 0.2\}$ and $d_1 \in \{-0.4, -0.2, 0, 0.2, 0.4\}$, $\mu_1 \in \{0, 0.5\}$ and $\alpha_1 \in \{\alpha_0 - 0.3, \alpha_0, \alpha_0 + 0.3\}$. Finite sample size (panels a and b) is relatively well controlled for both tests, converging to 5 percent as the sample size increases. As regards power (panels c and d), the advantage of using the $\widetilde{LMW}_T$ test, instead of the $\widetilde{LM}_T$ test, becomes much more substantial than in the previous sets of simulations, reaching a 40.4 pp. (= 94.6 − 54.2) power gain under joint shifts in d (from 0.2 to 0.4), μ (from 0 to 0.5), and α (from 0.5 to 0.8). As pointed out in Remark 6, the insight is that, unlike the $\widetilde{LM}_T$ test, the $\widetilde{LMW}_T$ test is able to yield i.i.d. residuals in its underlying regression when substantial breaks in the short-run dynamics are present (i.e. under fixed alternatives), and thus its use should be strongly recommended in those cases. In addition, panel (e), illustrates that the $\widetilde{LMW}_T$ test also has power against breaks under the alternative DGP-MA.[22] Thus, the choice of DGP (1) does not seem to play a crucial role in the obtained results.

[21] Unreported simulations show a comparable performance of the break fraction estimator for the alternative DGP-MA.

[22] Unreported simulations confirm these findings for the $\widetilde{LM}_T$ test as well.

[Table 3 about here]

(IV) Size and power of the LMW-type test (unknown AR(p) short-run dynamics)

In Table 4 we provide a fourth set of simulations similar to those presented in Block (III), but this time considering an unknown lag order p of the autoregressive short-run dynamics. Due to its higher power, results are only reported for the $\widetilde{LMW}_T$ test here and in the next two sets of simulations. As discussed in Section 6.1, in this case the BG testing procedure is implemented to check for lack of autocorrelation in the residuals estimated under the alternative. Thus, the lag order p of the ARFI(p, d) process should be augmented until the null of i.i.d. residuals is not rejected. The parameter configuration for this simulation exercise corresponds to those in Table 3. The true lag order is set to $p = 1$, and we allow for an ARFI(p, d) structure with $p \leq 3$, such that $P = 5$ lags of the residuals are considered in the BG testing procedure. Besides the size (panel a) and power (panel c) properties of the chosen test, we also report the proportion of the simulations in which the BG testing procedure selects the true value of p (panel b). As can be observed, while size and power are comparable to those reported for the $\widetilde{LMW}_T$ test in the simulations above, the proposed BG testing procedure correctly selects the true value of p in most instances, especially when $T = 500$ and 1000.[23]

[Table 4 about here]

Related to the previous set of simulations, we next consider breaks in the dynamics and levels for an ARFI$(3, d)$ process with known p. The autoregressive coefficients in the AR(3) lag polynomial $\left(1 - \alpha_{01}L - \alpha_{02}L^2 - \alpha_{03}L^3\right)$ are taken to be $\alpha_{01} \in \{-0.5, 0.5\}$, $\alpha_{02} = 0.5\alpha_{01}$ and $\alpha_{03} = 0.25\alpha_{01}$. As for power, we consider two alternative breaks in the AR parameter α_{01}, which takes values $\alpha_{11} \in \{-0.5, -0.2\}$ in the first case, and $\alpha_{11} \in \{0.2, 0.5\}$ in the second case. In both instances, the remaining parameters of the AR(3) process are set such that $\alpha_{12} = 0.5\alpha_{11}$ and $\alpha_{13} = 0.25\alpha_{11}$. The results of this simulation, reported in Table 5, support the previous findings about the fairly good size control and the satisfactory power of the $\widetilde{LMW}_T$ test.

[Table 5 about here]

(V) Size and power of the LMW-type test (ARMA short-run dynamics)

Whereas the simulation results in Block (IV) have shown that the proposed BG testing procedure can deal with ARFI(p, d) of any finite order p, it is interesting to check whether this procedure

[23] Notice, however, that the lag length selection could impact the properties of the test under some data configurations, e.g. when the errors have an MA root close to 1 (Ng and Perron, (2001)).

behaves properly when the short run dynamics follow an ARMA(p, q) process instead of an AR(p) process. Our conjecture is that the BG procedure should choose a sufficiently long autoregressive lag order in this case to render approximately uncorrelated residuals. To evaluate the performance of the $\widetilde{LMW}_T$ test in such a case, the following ARFIMA$(1, d, 1)$ process is considered as the DGP of this simulation exercise

$$(1 - \alpha_t L)\, \Delta_t^{d_t} (y_t - \mu_t) = (1 - \xi_0 L)\, \varepsilon_t, \quad t = 1, 2, \ldots,$$

with $\xi_0 = 0.2$ and $\alpha_0 \in \{0.3, 0.5, 0.8\}$ for size, and $\alpha_0 = 0.5$ and $\alpha_1 \in \{0.2, 0.5, 0.8\}$ for power.[24] The results of this fifth set of simulations of the $\widetilde{LMW}_T$ test of joint breaks in d, μ, and α are displayed in Table 6. As can be inspected, the size and power properties of the $\widetilde{LMW}_T$ test are very close to those presented in Table 3 above, implying that the BG testing procedure fares well in correcting for this type of autocorrelation by selecting a sufficiently long order p in the implementation of the test.

[Table 6 about here]

(VI) Size and power of the LM and LMW-type tests with heavy-tailed innovations

In all the previous simulations, innovations have been assumed to be Gaussian. Yet, given that the distribution of financial data is often heavy tailed, we report next a sixth set of size and power simulations with the same DGP used in Table 3 above for both tests, but this time with innovations being drawn from a $t(6)$ distribution, rather than from a $N(0, 1)$. Comparing the results in Table 7 with the previous results, we conclude that the size and power properties of both tests remain similar under this much heavier-tailed distribution of innovations.

[Table 7 about here]

(VII) Size and power of the LMW-type tests with multiple breaks

We finally report here some simulation results of the proposed sequential testing procedure for multiple breaks based on $\widetilde{LMW}_T$. In particular, we assume breaks at $\lambda_0 = 1/3$ and $\lambda_1 = 2/3$, splitting the series into three regimes. The sample sizes are $T = 200, 500$. For simplicity, we abstract from short-run dynamics and only consider breaks in d and μ. The setup is thus the same as in Table 2, but with two breaks instead of a single break. Since we first test for 0 vs 1 breaks in the sequential procedure, the size of the test is the one reported in Table 2. Thus, Table 8 reports power. In particular, we compute the proportion of times that the test locates 0, 1, 2 (the correct number of breaks) and 3 breaks. In addition, whenever 2 breaks are detected, we report the mean

[24] Results for $\xi_0 = -0.2$ are provided in the online appendix. They are fairly similar.

and variance of the estimated break fractions. Three different break scenarios are considered: (a) breaks in d from 0 to 0.4 and then back to 0, plus in μ from 0 to 0.5 and then back to 0; (b) breaks in d from -0.4 to 0 and then up to 0.4, and in μ from 0 to 0.5 and then up to 1; and (c) breaks in d from 0 to 0.2 and then down to -0.2, plus the same breaks in μ as in (b). The main takeaways from Table 8 are that: (i) as expected, acceptable power performance requires sizeable breaks as well as large sample sizes ($T = 500$), in line with the consistency of the sequential procedure; and (ii) the consistency of the break fraction estimates holds in all cases, as reflected by the means becoming closer to 1/3 and 2/3 with decreasing standard deviations.

[**Table 8 about here**]

8 Empirical application

In this section we apply the proposed testing methodology to the analysis of the forward discount in exchange rate markets. As is well known, rational expectations and risk neutrality, combined with covered and uncovered interest rate parity, lead to the so-called forward exchange rate unbiasedness hypothesis (FRUH) whereby the (logged) forward rate, f_t, is an unbiased predictor of the future (logged) spot exchange rate, s_{t+1}, i.e. $E_t(s_{t+1}) = f_t$. In particular, testing FRUH corresponds to a test of the null $H_0 : \delta_0 = 0$, $\delta_1 = 1$ in the following regression model

$$\Delta s_{t+1} = \delta_0 + \delta_1 (f_t - s_t) + \varepsilon_{t+1},$$

where $(f_t - s_t)$ is the forward discount. This null has often been rejected in empirical applications where typically the OLS point estimate of δ_1 is small or even negative (see, e.g., Engel, 1996, for an overview of this literature), leading to what has been coined the *forward discount anomaly*. It has been argued that this finding may result from the unbalanced nature of the previous regression. In effect, while the dependent variable Δs_{t+1} is conventionally found to be $I(0)$, there is a large body of literature documenting that $(f_t - s_t)$ follows a $I(d)$ process with d generally lying in the non-stationary range, $0.5 < d < 1$ (see, e.g., Baillie and Bollerslev (1994), and Maynard and Phillips, (2001)). However, Choi and Zivot (2007) have shown that estimates of d are likely to be upward biased when structural instabilities in the level of $(f_t - s_t)$ are ignored. Using the residuals of the forward discount monthly series for five G7 countries, Choi and Zivot (2007) first adjust the level of these series for several structural breaks detected by means of Bai and Perron's (1998) methodology. Next, they estimate d non-parametrically (using Kim and Phillips's (2000) log-periodogram regression approach) from these break-adjusted series. Their evidence points out that $(f_t - s_t)$ is subject to several breaks and that accounting for these breaks leads to much lower estimates of d than those previously found in the literature, with $0 < d < 0.5$.

Following this controversy, we provide here a brief empirical application of our proposed tests using the forward discount data for five G7 countries examined by Choi and Zivot (2007). Their dataset includes monthly forward discount rates for the period 1976:1-1996:1 corresponding to the exchange rates in terms of US dollars for Canada, Germany, France, Italy, and U.K., where f_t is defined as the (logged) 30-day forward rate.

Figure 3 displays the five time series at hand. Choi and Zivot (2007) find five breaks in the level of $(f_t - s_t)$ for Germany and U.K., four breaks for France and Italy, and three breaks for Canada. The dates of all these breaks are displayed using dashed vertical lines in Figure 3. As can be seen, about half of them take place before 1981, while most of the remaining ones occur between the late 1980s and early 1990s.

[**Figure 3 about here**]

However, Choi and Zivot (2007) also report that reversing the testing procedure (i.e., first d is estimated from the time series of $(f_t - s_t)$ without allowing for level shifts, and then Bai and Perron's (1998) procedure is used to detect multiple breaks in the filtered series $\Delta_t^d(f_t - s_t)$), leads to a much smaller number of breaks (none for Germany and France, one for Italy, and three for Canada and U.K.). These contrasting findings possibly reflect the shortcomings of using a two-step testing procedure instead of a single-stage approach, as the one proposed here. Moreover, given that the level and the dynamics could shift simultaneously (an event which is not considered by these authors), both sources of breaks could easily get confused. Our single-step testing approach is therefore better suited to address this problem since it yields more reliable estimates both of the number of breaks and their origin. Note that, despite the fact that financial series at daily frequencies exhibit pronounced volatility clustering, there are little ARCH effects in the monthly forward discount rates and persistence in such series is well captured by an I(d) process (see e.g. Baillie and Bollerslev (1989)), making our testing procedures suitable for their analysis. At any rate, since our tests require i.i.d. innovations, we correct for any potential heteroskedasticity left by standardizing the variance of the residuals in the regressions performed in each of the different regimes.

To check this possibility of breaks in levels and dynamics, we consider an ARFI(p, d) model with drift and an AR process of unknown order p for each of the five series, allowing for simultaneous breaks in all three parameters $(d, \boldsymbol{\alpha}, \mu)$. Given its better power performance, we apply our sup-$\widetilde{LMW}_T$ test statistic, where the lag order p is selected according to the proposed BG testing procedure. To allow for multiple breaks (see the discussion in Subsection 6.2), we test sequentially (0 *vs.* 1 break and, upon rejection, 1 *vs.* 2, and so forth) and use the critical values reported in Table 1 to determine the number of breaks together with the break fractions. Table 9 (panel a) and

Figure 3 display the break dates as the vertical solid lines. In addition, horizontal solid lines depict the estimated value of d in each of the relevant subsamples where breaks have been identified.

[Table 9 about here]

At the 5 % level, the sup-$\widetilde{LMW}_T$ test detects only one statistically significant break for Canada and Germany, two for France and Italy and three for the U.K., although there is some evidence in favour of a second break for both Canada and Germany when using a 10% significance level. Thus, the number of breaks found with our single-step testing procedure (9 to 11) is in between the two contrasting numbers (21 and 7) reported by Choi and Zivot (2007) in their original and reversed two-stage testing approach. As for the lag orders in the AR(p) process, the BG testing procedure selects $p = 1$ for Germany and Italy, $p = 2$ for U.K., and $p = 3$ for Canada and France. When using the Bonferroni correction in the sequential procedure, as suggested in Subsection 6.2, all the previous breaks remain statistically significant but this time only at the 10% level.

In general, the breaking dates estimates gather around the second half of the 1980s and early 1990s, the latter possibly as a result of the collapse of the ERM in 1992. Our results agree with Choi and Zivot's (2007) in that we find a break in the early and late 1980s for U.K. and in the early 1990s for Canada, Germany, and the U.K. However, with the exception of the U.K. in 1981, their earlier breaks turn out to be statistically insignificant according to our single-step testing procedure.

In order to provide further comparisons with Choi and Zivot (2007)' findings, we also apply the robustified $\overline{LMW}_T$ test discussed in Subsection 5.2 to identify which specific parameters shift at each of the previously identified break points (see Remark 12 above for a justification of why this sequential approach does not involve a multiple testing size problem). Table 9 (panel b) presents the results obtained from applying this test to detect breaks in the dynamics as a whole (d, $\boldsymbol{\alpha}$) and in the level μ. As can be inspected, in line with these authors, we find that all detected breaks involve shifts in the levels for the specific break dates identified here. However, a majority of them (the exception are the only break in Germany and the first break in the U.K.) also involve parameter shifts in either the short-run dynamics or in the long memory parameter. For example, a comparison of the estimates of d reported in Figure 3 with the test outcomes in Table 9 (panel b) shows that, on top of μ, both d and $\boldsymbol{\alpha}$ shift after the 1987(12) break in Italy, whereas d remains stable but $\boldsymbol{\alpha}$ shifts after the second break in 1994(2). Moreover, unlike these authors, we find instances (e.g. France and the second break in the U.K.) where d increases, even after allowing for breaks in $\boldsymbol{\alpha}$ and μ. Therefore, it look like many of the breaks interpreted by Choi and Zivot (2007) as being exclusively due to shifts in the level of forward discount rates also involve shifts in the memory and short-run dynamics, and that these novel findings can only be uncovered by our proposed testing approach involving joint parameter breaks.

9 Conclusions

Our motivation for this paper is that the joint modeling of breaks in the (memory and short-run) dynamics and the level of fractionally integrated stochastic processes is a relevant issue to analyze on which research has been limited so far. By considering breaks in all parameters simultaneously, potential confounding problems about the sources of shifts in the persistence of a time-series process can be avoided. Our contribution here is to extend the well-known LM test for breaks only in the memory parameter of an I(d) process to further account for breaks in the level as well as in the short-run dynamics. As a by-product of our analysis, we derive: (i) a novel regression-based LM test *cum* Wald interpretation, labeled LMW-type test, for ARFI(d, p) processes with drift that also accounts for all these shifts; (ii) individual tests for the stability of a given parameter which are robust to the behaviour of the non-tested parameters, and (iii) consistent estimates of the break dates.

The proposed tests share several nice features. While LM tests are computationally attractive by only requiring estimation under the null, LMW-type tests can exploit further information about the alternative, potentially leading to higher power without increasing computational complexity. In addition, in contrast to LM tests, LMW-type tests allow for a consistent specification of the short-run dynamics, as long as these are restricted to AR(p) processes, although in simulations we show that they can also accommodate some ARMA processes and innovations with heavy-tailed distributions. Our Monte-Carlo simulations, based on analytical results, show in particular that LMW-type tests for joint breaks can yield substantial power gains relative to LM tests in several instances and are robust to different specifications of the DGP. Finally, our empirical application on potential breaks in forward discount rates for several G7 countries provides new findings on the origin of these breaks (in both components of the dynamics, as well as in levels), which have been subject to considerable attention in the literature but without considering shifts in all those parameters at the same time.

References

[1] Abadir, K.M., Di Staso, W. and Giraitis, L. (2007). "Nonstationarity-extended local Whittle estimation", *Journal of Econometrics* 141, 1353-1384.

[2] Aue, A. and Horváth, L. (2013), "Structural breaks in time series". *Journal of Time Series Analysis*, 34, 1–16.

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

[4] Baillie, R. T. and Bollerslev, T. (1989) "The message in daily exchange rates: A conditional variance tale". *Journal of Business and Economic Statistics*, 7, 297-305.

[5] Baillie, R.T. and Bollerslev, T. (1994) "The long memory of the forward premium". *Journal of International Money and Finance*, 13, 555-571.

[6] Beran, J. and Terrin, N. (1996)" Testing for a change of the long-memory parameter". *Biometrika* 83, 627-638.

[7] Breitung, J. and Hassler, U. (2002) "Inference on the cointegration rank in fractionally integrated processes," *Journal of Econometrics*, 110, 167-185.

[8] Breusch, T. S. (1978) "Testing for autocorrelation in dynamic linear models". *Australian Economic Papers*. 17: 334–355.

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

[10] Choi, K. and Zivot, E. (2007) "Long memory and structural changes in the forward discount: An empirical investigation," *Journal of International Money and Finance*, 26, 342-363.

[11] Chung, C.F. and Baillie, R.T. (1993) "Small sample bias in conditional sum-of-squares estimators of fractionally integrated arma models". *Empirical Economics* 18, 791-806.

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

[13] Dolado, J. & Gonzalo, J. and L. Mayoral (2002) "A Fractional Dickey-Fuller test for unit roots," *Econometrica* 70, 1963-2006.

[14] Dolado, J.J., Gonzalo, J., and Mayoral, L. (2009) "Simple Wald tests of the fractional integration parameter: an overview of new results". In J. Castle & N. Shephard (eds.) *The Methodology and Practice of Econometrics*. A Festschift in Honour of David F. Hendry, 300-321, Oxford University Press.

[15] Dolado, J.J., Gonzalo, J., and Mayoral L. (2008) "Wald tests of I(1) against I(d) alternatives: some new properties and an extension to processes with trending components". *Studies in Nonlinear Dynamics & Econometrics* 12, 1-33.

[16] Dolado, J., Rachinger, H. and Velasco, C. (2017) "LM tests for joint breaks in the memory and level of a time series" shttps://sites.google.com/site/heikorachinger/DRV2017.

[17] Engel, E. (1996) "The forward discount anomaly and the risk premium: A survey of recent evidence". *Journal of Empirical Finance*, 3, 123-192.

[18] Engle, R. (1984). "Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics". *Handbook of Econometrics*, vol II, ed. Griliches and Intrilligator (Amsterdam: North Holland, 1984), 775-826

[19] Gadea, M. D. and Mayoral, L. (2006). "The persistence of inflation in OECD countries: A fractionally integrated approach". *International Journal of Central Banking*, 2, 51-104.

[20] Gil-Alaña, L. A. (2008) "Fractional integration and structural breaks at unknown periods of time". *Journal of Time Series Analysis* 29, 163-185.

[21] Giraitis, L, Leipus, R. and Philippe, A. (2006) "A test for stationarity versus trends and unit roots for a wide class of dependent errors". *Econometric Theory*, 22.6, 989-1029.

[22] Godfrey, L. G. (1978) "Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables". *Econometrica*. 46: 1293–1301.

[23] Granger, C. W. J. and Hyung, N. (2004) "Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns". *Journal of Empirical Finance* 11, 399-421.

[24] Granger, C. W. J. (1980) "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*. 14(2), 227-238.

[25] Hassler, U. and Meller, B. (2014) "Detecting multiple breaks in long memory: the case of U.S. inflation". *Empirical Economics.* 46, 653-680.

[26] Hassler, U. and Scheithauer, J. (2011) "Detecting changes from short to long memory". *Statistical Papers* 52, 847-870.

[27] Hidalgo, J. and Robinson, P.M. (1996) "Testing for structural change in a long memory environment", *Journal of Econometrics* 73, 261-284.

[28] Horváth, L. and Shao, Q.-M. (1999) "Limit theorems for quadratic forms with applications to Whittle's estimate". *The Annals of Applied Probability*. 9, 146–187.

[29] Hsu, C.C. (2005) "Long memory or structural changes: An empirical examination on inflation rates". *Economics Letters* 88, 289-294

[30] Hualde, J. and Nielsen, M.Ø., (2020) "Truncated sum of squares estimation of fractional time series models with deterministic trends". *Econometric Theory*, 36, 751-772.

[31] Hualde, J. and Robinson, P.M. (2011) "Gaussian pseudo-maximum likelihood estimation of fractional time series models". *Annals of Statistics* 39, 3152-3181.

[32] Iacone, F., Leybourne, S. J., and Taylor, A.M.R. (2013) " Testing for a break in trend when the order of integration is unknown". *Journal of Econometrics*, 176, 30-45.

[33] Kim, J.-Y., Belaire-Franch, J., and Amador, R.B. (2002) "Corrigendum to detection of change in persistence of a linear time series". *Journal of Econometrics* 109, 389-392

[34] Lavielle, M. and Moulines E. (2000) "Least squares estimation of an unknown number of shifts in a time series". *Journal of Time Series Analysis* 21, 33-59.

[35] Lobato, I. N. and Savin, N. E. (1998) "Real and spurious long-memory properties of stock-market data". *Journal of Business & Economic Statistics* 16, 261-268.

[36] Lobato, I. N. and Velasco, C. (2007) "Efficient Wald tests for fractional unit roots". *Econometrica* 75, 575-589.

[37] Lobato, I. N. and Velasco, C. (2008) "Power comparison among tests for fractional unit roots". *Economics Letters* 99, 152-154.

[38] Marinucci, D., Robinson, P.M., 2000. "Weak convergence of multivariate fractional processes". *Stochastic Processes and their Applications* 83, 103-120.

[39] Martins, L.F. and Rodrigues, P.M.M. (2014) "Testing for persistence change in fractionally integrated models: An application to world inflation rates". *Computational Statistics and Data Analysis*. 76(C), 502-522.

[40] Maynard, A. and Phillips, P.C.B. (2001) "Rethinking an old puzzle: Econometric evidence on the forward discount anomaly" *Journal of Applied Econometrics*, 16, 671-708.

[41] McCloskey, A. (2010). "Semiparametric testing for changes in memory of otherwise stationary time series". Mimeo. January 2010.

[42] McCloskey, A. and Perron, P. (2013) "Memory parameter estimation in the presence of level shifts and deterministic trends" *Econometric Theory*, 29, 1196-1237.

[43] Ng, S. and Perron, P. (2001) " Lag length selection and the construction of unit root tests with good size and power" *Econometrica*, 69, 1519-1574.

[44] Perron, P. and Qu, Z. (2010) "Long-memory and level shifts in the volatility of stock market return indices". *Journal of Business & Economic Statistics* 28, 275-290.

[45] Rachinger, H. (2017) "Multiple breaks in long memory time series". University of Vienna, mimeo. https://sites.google.com/site/heikorachinger/Rachinger2017.pdf

[46] Robinson, P.M. (1978) "Statistical inference for a random coefficient autoregressive model", *Scandinavian Journal of Statistics* 5, 163-168.

[47] Robinson, P.M. (1994) "Efficient tests of nonstationary hypotheses". *Journal of the American Statistical Association*, 89, 1420-1437

[48] Robinson, P.M. (2006) "Conditional-sum-of-squares estimation of models for stationary time series with long memory", *Time Series and Related Topics: In Memory of Ching-Zong Wei.* (H.-C. Ho, C.-K. Ing and T.L. Lai, eds.). IMS Lecture Notes - Monograph Series, 52, 130-137.

[49] Shao, X. (2011). "A simple test of changes in mean in the possible presence of long-range dependence". *Journal of Time Series Analysis* 32, 598-606.

[50] Wooldridge, J. M. (1990) "A unified approach to robust, regression-based specification tests", *Econometric Theory* 6, 17-43.

[51] Yamaguchi, K. (2011) "Estimating a change point in the long memory parameter", *Journal of Time Series Analysis* 32, 304-314.

10 TABLES AND FIGURES

Table 1: **Critical Values of LM tests for breaks in (d, μ) or in $(d, \boldsymbol{\alpha}, \mu)$ for an unknown break fraction λ.**

a) Critical Values of LM tests for breaks in (d, μ) , $p = 0$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	11.3	11.5	11.5	11.6	11.9	12.3	12.9	14.0	15.9	18.4	20.6
0.15	10.9	10.9	11.0	11.2	11.4	11.8	12.4	13.5	15.4	18.0	20.2
0.2	10.5	10.5	10.6	10.7	11.0	11.3	12.0	13.1	14.9	17.5	19.8

b) Critical Values of LM tests for breaks in $(d, \boldsymbol{\alpha}, \mu)$, $p = 1$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	14.2	14.2	14.2	14.3	14.5	14.7	15.2	16.2	17.8	20.2	22.4
0.15	13.6	13.6	13.7	13.8	14.0	14.2	14.7	15.6	17.3	19.7	21.9
0.2	13.1	13.2	13.2	13.3	13.5	13.7	14.2	15.2	16.8	19.2	21.5

c) Critical Values of LM tests for breaks in $(d, \boldsymbol{\alpha}, \mu)$, $p = 2$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	16.5	16.5	16.6	16.7	16.7	16.9	17.4	18.1	19.7	21.9	24.1
0.15	15.9	16.0	16.0	16.1	16.2	16.4	16.8	17.6	19.1	21.4	23.7
0.2	15.4	15.4	15.4	15.6	15.7	15.9	16.3	17.1	18.7	20.9	23.2

d) Critical Values of LM tests for breaks in $(d, \boldsymbol{\alpha}, \mu)$, $p = 3$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	18.6	18.7	17.7	18.8	18.8	18.9	19.4	20.0	21.5	23.6	25.7
0.15	18.0	18.1	18.1	18.1	18.3	18.4	18.8	19.4	20.8	23.1	25.3
0.2	17.5	17.5	17.6	17.7	17.7	17.9	18.2	18.9	20.3	22.6	24.9

e) Critical Values of LM-tests for breaks in $(d, \boldsymbol{\alpha}, \mu)$, $p = 4$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	20.6	20.6	20.7	20.7	20.8	20.9	21.3	21.9	23.2	25.2	27.4
0.15	19.8	20.0	20.0	20.1	20.2	20.3	20.6	21.3	22.6	24.8	26.9
0.2	19.5	19.5	19.5	19.6	19.6	19.8	20.1	20.7	22.1	24.2	26.5

f) Critical Values of LM-tests for breaks in $(d, \boldsymbol{\alpha}, \mu)$, $p = 5$.

$\epsilon \setminus d_0$	-0.49	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.49
0.1	22.5	22.5	22.5	22.6	22.7	22.7	23.1	23.7	24.9	26.9	29.0
0.15	21.8	21.9	21.9	22.0	22.0	22.1	22.4	23.0	24.2	26.4	28.5
0.2	21.3	21.3	21.3	21.4	21.4	21.6	21.9	22.4	23.7	25.8	28.1

Note: Unknown break fraction $\lambda \in [\epsilon, 1-\epsilon]$. 5% Significance level.

Based on 10,000 grid points for the break fraction and 100,000 simulations.

Table 2: **Simulated size and power of the LM and LMW-type tests for a joint break in long memory and level (unknown break fraction)**

a) $\widetilde{LM}$ test: Size

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	6.8	6.6	6.5	6.6	6.4	6.1	6.0	5.2	4.8
500	6.0	6.7	6.4	6.4	6.0	5.9	5.7	6.0	5.7
1000	5.6	6.5	6.2	5.8	5.4	5.2	5.1	5.6	4.9

b) $\widetilde{LMW}$-type test: Size

$T \setminus d_0$	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
200	7.4	7.6	6.7	7.4	6.6	6.2	6.1	4.6	5.3
500	6.6	6.9	6.3	6.3	6.9	6.6	6.0	4.2	4.4
1000	5.8	6.6	6.8	6.2	6.0	5.6	5.0	4.7	4.2

c) $\widetilde{LM}$ test: Power ($T = 200$)

d_0	-0.2				0					0.2			
$\mu_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
0	26.6	**6.5**	17.8	58.5	74.7	27.0	**6.4**	14.7	53.4	71.7	22.0	**6.0**	13.6
0.25	94.7	71.3	50.8	67.5	89.4	57.5	24.2	25.2	55.9	76.3	30.5	10.2	14.8
0.5	100	99.6	92.8	86.0	99.4	93.2	65.8	45.0	62.6	86.5	54.0	21.3	20.1

d) $\widetilde{LMW}$-type test: Power ($T = 200$)

d_0	-0.2				0					0.2			
$\mu_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
0	29.9	**6.7**	21.1	63.9	75.3	28.0	**6.6**	18.2	56.2	73.4	23.1	**6.1**	16.0
0.25	95.9	75.7	58.5	73.9	94.3	63.9	29.3	28.7	61.0	80.1	35.1	16.2	19.7
0.5	100	99.8	93.9	87.7	99.4	94.8	69.6	49.0	65.7	93.9	58.1	26.2	27.3

Note: Rejection probabilities of 5% test for joint break in d and μ, $\epsilon = 0.15, \lambda_0 = 0.5, \mu_0 = 0, \sigma_0^2 = 1$. Figures in bold characters correspond to simulated sizes.

Table 3: **Simulated size and power of the LM and LMW-type tests for a joint break in long memory, level and AR short-run dynamics of known order**

a) $\widetilde{LM}$ test: Size

α_0	-0.5				0.5			
$T \setminus d_0$	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
200	6.5	6.6	6.3	6.5	4.1	4.4	4.9	4.7
500	6.1	5.8	5.9	5.4	4.9	5.4	5.4	4.9
1000	5.7	5.5	5.7	5.0	5.0	5.2	5.3	5.2

b) $\widetilde{LMW}$- type test: Size

α_0	-0.5				0.5			
$T \setminus d_0$	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
200	7.1	6.6	6.1	6.0	6.2	5.8	5.7	5.8
500	5.8	5.5	6.2	5.7	6.5	6.5	6.4	6.1
1000	5.8	5.8	5.5	5.1	5.7	5.3	5.3	5.5

c) $\widetilde{LM}$ test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	43.8	7.8	13.4	58.5	86.9	41.1	7.7	15.1	60.8	85.7	37.5	7.7	15.1
0	-0.5	-0.5	33.7	**6.5**	27.8	77.2	81.2	31.5	**6.6**	27.9	78.8	80.3	27.1	**6.3**	28.6
		-0.2	12.6	15.1	58.8	95.3	50.3	12.3	15.0	58.9	95.7	50.7	12.2	14.4	61.4
		0.2	85.8	40.2	8.6	11.3	97.7	83.1	38.5	9.4	13.1	97.3	82.8	37.7	7.4
0	0.5	0.5	20.5	**4.1**	15.2	57.9	61.5	18.9	**4.4**	17.1	60.5	64.0	18.2	**4.9**	18.2
		0.8	5.3	29.1	69.6	91.1	9.4	6.4	30.9	69.1	81.1	10.4	7.4	31.0	53.8
		-0.8	100	100	100	99.2	99.8	100	99.3	90.9	84.2	96.7	91.8	66.7	47.5
0.5	-0.5	-0.5	100	100	100	99.1	100	100	99	88.2	90.4	97.3	88.0	58.6	51.6
		-0.2	100	100	100	99.5	99.9	99.4	97.1	90.9	97.2	91.7	73.2	58.5	75.2
		0.2	99.4	93	55.8	31.2	98.4	93.7	66.1	21.2	17.3	98.2	86.6	46.4	12.4
0.5	0.5	0.5	80.8	51.7	43.1	67.2	80.5	46.4	14.9	24.8	61.9	69.3	25.3	7.5	19.6
		0.8	37.6	51.5	73.2	92.4	24.9	15.9	34.6	71.2	81.4	13.2	8.2	33.5	54.2

d) $\widetilde{LMW}$ -type test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	91.9	62.8	50.1	71.5	99.8	91.0	64.0	51.0	73.4	99.8	90.8	62.8	51.6
0	-0.5	-0.5	34.9	**7.1**	27.4	76.7	83.6	31.5	**6.6**	27.0	78.3	82.3	29.1	**6.1**	28.8
		-0.2	32.8	45.9	85.5	99.2	57.3	29.5	48.2	86.6	99.2	58.4	28.7	47.7	85.9
		0.2	93.6	51.0	13.8	22.0	99.8	92.5	44.8	13.4	23.0	99.6	90.1	42.9	12.5
0	0.5	0.5	34.3	**6.2**	22.3	74.0	80.4	29.0	**5.8**	23.1	75.0	81.8	25.7	**5.7**	24.6
		0.8	11.3	54.3	92.8	99.8	12.0	12.3	53.9	93.6	99.7	12.7	12.4	58.1	94.3
		-0.8	100	100	100	99.2	100	100	99.9	96.7	88.8	99.9	99.6	92.3	72.6
0.5	-0.5	-0.5	100	100	100	99	100	100	98.5	86.9	90.8	97.9	87.3	55.9	50.6
		-0.2	100	100	100	100	99.9	99.6	98.9	97.8	99.6	92.7	83.2	78.4	92.1
		0.2	99.5	96.5	70.4	49.6	99.6	96.4	74.4	32.4	30.7	99.8	92.6	56.6	18.9
0.5	0.5	0.5	90	67.9	58.2	83.6	87.3	58.8	23.5	34	78.5	83.6	34.8	10.7	27.9
		0.8	58.7	74.1	95.2	99.9	32.9	27.9	59.4	95.1	99.7	18.0	17.2	61.3	94.6

e) $\widetilde{LMW}$ -type test: Power under DGP-MA ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	89.5	63.3	49.9	71.3	99.3	90.6	62.3	51.5	73.7	99.5	91.0	63.2	52.7
0	-0.5	-0.5	26.6	**7.1**	25.2	78.3	77.8	27.5	**6.6**	27.0	80.4	79.8	28.8	**6.1**	29.9
		-0.2	28.1	46.3	86.7	99.5	48.2	27.7	47.0	86.7	99.5	51.5	28.6	46.7	88.2
		0.2	91.6	49.7	14.1	23.3	99.6	90.3	45.4	12.2	25.3	99.6	90.6	46.2	12.3
0	0.5	0.5	27.0	**6.2**	23.3	75.5	75.2	26.3	**5.8**	23.4	81.3	78.2	28.0	**5.7**	29.4
		0.8	12.1	52.2	94.4	99.9	8.9	12.0	54.9	96.1	99.9	10.6	12.7	62.1	97.8
		-0.8	100	100	100	99.4	100	100	99.9	96.6	87.9	100	99.4	92.1	75.7
0.5	-0.5	-0.5	100	100	100	99.1	100	99.9	98.2	87.2	89.4	99.1	88.1	57.0	51.8
		-0.2	100	100	100	99.9	100	99.9	98.7	97.6	99.4	95.2	82.8	78.5	91.3
		0.2	100	96.6	69.8	49.5	100	97.8	75.3	32.4	34.5	99.8	93.9	56.2	18.2
0.5	0.5	0.5	92.6	67.5	59.5	82.7	92.2	57.9	24.6	35.0	82.1	83.3	37.1	10.8	33.3
		0.8	59.5	73.2	96.1	99.9	33.1	26.3	62.3	96.0	100	18.8	18.7	65.3	98.1

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5$, $\mu_0 = 0$, $\sigma_0^2 = 1$. Figures in bold characters correspond to simulated sizes.

Table 4: **Simulated size and power of the LMW-type test for a joint break in memory, level and AR short-run dynamics of unknown order.**

a) $\widetilde{LMW}$- type test: Size

α_0	-0.5					0.5				
$T \setminus d_0$	-0.4	-0.2	0	0.2	0.4	-0.4	-0.2	0	0.2	0.4
200	7.5	7.3	7.0	6.8	7.0	7.2	6.9	7.2	7.0	6.1
500	6.7	6.7	6.7	6.4	6.2	6.6	6.4	6.6	5.9	5.6
1000	6.1	5.5	5.7	5.6	5.7	6.2	5.7	5.5	4.8	5.1

b) Correct lag order selected by BG procedure under H_0

α_0	-0.5					0.5				
$T \setminus d_0$	-0.4	-0.2	0	0.2	0.4	-0.4	-0.2	0	0.2	0.4
200	75.0	89.4	91.3	91.3	91.7	64.6	65.3	65.9	75.7	87.6
500	92.7	93.1	92.9	93.0	93.6	88.6	91.0	90.5	90.7	92.1
1000	93.5	93.2	93.1	93.3	93.6	92.0	92.5	92.9	93.0	93.4

Note: Proportion (in %) of cases where the BG correctly selects the correct AR(p) structure

c) $\widetilde{LMW}$- type test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	92.0	65.0	50.6	72.0	99.6	91.6	64.6	52.4	72.0	99.7	91.4	62.5	52.9
0	-0.5	-0.5	35.1	**7.3**	27.8	78.3	84.2	33.5	**7.0**	27.4	79.4	84.7	31.5	**6.8**	29.9
		-0.2	33.1	50.2	87.9	99.5	62.4	33.0	49.7	88.2	99.4	60.3	30.7	49.4	88.2
		0.2	87.9	41.6	10.5	14.2	99.5	88.4	37.7	10.0	16.1	99.6	87.5	41.0	11.3
0	0.5	0.5	27.7	**6.9**	19.5	68.3	75.6	24.8	**7.2**	22.4	71.9	78.4	25.5	**7.0**	25.2
		0.8	10.1	42.6	90.8	99.7	11.3	10.8	51.9	93.4	99.8	10.9	12.1	55.0	93.6
		-0.8	100	100	100	99.4	100	100	99.9	96.8	89.3	99.9	99.5	92.4	75.4
0.5	-0.5	-0.5	100	100	99.8	99.2	100	99.9	97.5	87.0	89.8	97.2	88.2	59.6	53.0
		-0.2	100	100	100	99.8	99.8	99.6	98.7	98.1	99.8	93.2	84.4	80.7	93.2
		0.2	99.0	91.7	58.0	35.0	98.8	95.0	64.2	25.3	22.3	99.5	91.0	47.8	15.9
0.5	0.5	0.5	80.3	55.9	47.7	73.4	82.6	51.6	22.3	32.6	76.1	80.4	33.4	10.7	27.4
		0.8	40.1	60.4	92.6	99.7	28.4	21.9	57.0	93.5	99.8	16.6	16.5	56.8	94.9

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5,\ \mu_0 = 0, \sigma_0^2 = 1$. Figures in bold characters correspond to simulated sizes.

Table 5: **Simulated size and power of the LMW-type test for a joint break in memory, level and AR(3) short-run dynamics**

a) $\widetilde{LMW}$- type test: Size

α_{00}	-0.5					0.5				
$T \ \backslash \ d_0$	-0.4	-0.2	0	0.2	0.4	-0.4	-0.2	0	0.2	0.4
200	7.1	6.9	7.1	7.4	7.3	4.7	5.3	6.1	5.5	5.6
500	6.7	6.7	6.9	5.9	6.1	4.4	4.0	4.9	5.0	5.1
1000	5.9	6.4	5.7	5.7	5.8	4.1	4.1	4.3	5.0	5.1

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5,\ \mu_0 = 0,\ \sigma_0^2 = 1$.

b) $\widetilde{LMW}$- type test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_{00}	$\alpha_1 \backslash \ d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
0	-0.5	-0.5	27.2	**6.9**	21.6	70.5	77.0	27.8	**7.1**	24.2	72.0	79.0	26.5	**7.4**	23.5
		-0.2	14.7	33.7	80.4	99.0.	35.8	16.4	36.4	81.4	98.8	35.4	15.6	34.5	81.5
0		0.2	97.6	73.2	25.9	12.0	100	97.5	72.8	25.2	11.5	100	97.5	73.6	26.7
	0.5	0.5	20.9	**5.3**	17.2	61.9	72.4	22.1	**6.1**	18.1	65.8	72.9	21.6	**5.5**	18.6
0.5	-0.5	-0.5	100	100	99.9	99.0	99.9	99.7	96.5	86.5	87.2	96.9	86.5	58.7	51.5
		-0.2	100	100	99.9	99.9	99.2	98.3	96.8	96.3	99.5	85.2	74.1	70.7	88.1
0.5		0.2	95.8	76.5	33.6	16.7	99.7	97.6	74.2	29.5	13.7	99.9	97.4	74.0	27.4
	0.5	0.5	27.2	7.8	19.5	64.3	74.5	24.5	6.3	19.5	67.4	74.0	25.3	7.8	19.8

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5,\ \mu_0 = 0,\ \sigma_0^2 = 1$. $\alpha_{00} = -0.5$ and $\alpha_{10} \in \{-0.5, -0.2\}$, and $\alpha_{00} = 0.5$ and $\alpha_{10} \in \{0.2, 0.5\}$ respectively. $\alpha_{i1} = 0.5\alpha_{i0}$ and $\alpha_{i2} = 0.25\alpha_{i0},\ i = 0, 1$. Figures in bold characters correspond to simulated sizes.

Table 6: **Simulated size and power of the LMW-type test for a joint break in long memory, level and ARFIMA(1,d,1) short-run dynamics**

a) $\widetilde{LMW}$ -type test: Size for ARMA(1,1) innovations

α_0	0.3					0.5					0.8				
$T \setminus d_0$	-0.4	-0.2	0	0.2	0.4	-0.4	-0.2	0	0.2	0.4	-0.4	-0.2	0	0.2	0.4
200	8.5	7.3	7.0	7.1	5.9	6.8	7.4	6.8	6.7	4.9	5.9	5.6	4.9	4.3	5.1
500	7.3	7.0	6.7	6.8	6.3	6.7	7.6	7.3	6.7	5.3	7.9	7.5	6.1	5.8	4.9
1000	6.4	6.2	6.2	5.7	5.5	5.7	6.5	6.3	6.4	5.6	6.0	5.7	5.3	5.3	5.0

b) $\widetilde{LMW}$- type test: Power for ARMA(1,1) innovations ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		0.2	86.3	37.5	10.6	16.4	99.5	85.7	37.4	11.0	17.5	99.7	85.5	36.7	10.6
0	0.5	0.5	26.0	**7.4**	22.7	68.5	76.2	25.3	**6.8**	21.2	68.7	75.6	23.8	**6.7**	21.6
		0.8	12.9	46.1	90.2	99.6	10.8	13.9	48.3	90.8	99.7	10.1	11.5	48.9	90.1
		0.2	96.9	80.6	45.3	33.7	98.6	91.3	55.1	20.4	24.6	99.4	88.3	44.4	15.2
0.5	0.5	0.5	66.4	36.6	42.6	73.3	78.2	39.6	16.7	28.3	69.9	76.6	26.3	9.9	22.8
		0.8	35.2	61.1	92.7	99.7	22.3	20.1	54.6	91.3	99.6	13.3	15.2	53.4	91.8

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5$, $\mu_0 = 0$, $\xi_0 = 0.2$, $\sigma_0^2 = 1$. Figures in bold characters correspond to simulated sizes.

Figure 1: **Drift of the tests for a break in long memory parameter**

Drift terms of the LM and LMW-type tests as a function of the break magnitude $d_1 - d_0$. $\widetilde{LM}$ test (dashed line) and $\widetilde{LMW}$ -type test (solid line) ($\lambda_0 = 0.5$, $d_0 = 0$).

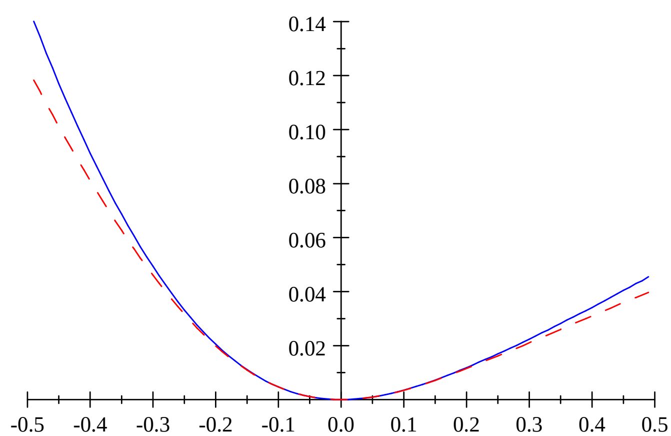


Table 7: **Simulated size and power of the LM and LMW-type tests for joint breaks in memory, level and AR component with t(6) innovations.**

a) $\widetilde{LM}$ test: Size

α_0	-0.5				0.5			
$T \setminus d_0$	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
200	6.9	6.6	6.8	5.9	4.1	4.4	4.2	4.3
500	6.0	6.1	5.8	5.7	5.1	5.1	4.9	5.0
1000	5.5	5.8	5.5	5.5	4.9	5.3	5.0	4.7

b) $\widetilde{LMW}$ -type test: Size

α_0	-0.5				0.5			
	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
	6.8	6.5	6.2	6.2	6.6	6.2	5.9	5.5
	6.0	5.9	5.6	5.4	6.4	6.3	5.6	5.4
	5.3	5.7	5.4	5.5	5.8	5.9	5.4	5.3

c) $\widetilde{LM}$ test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	42.2	8.0	13.8	57.8	86.9	40.1	7.2	15.1	59.1	87.0	38.1	6.4	15.4
0	-0.5	-0.5	31.5	**6.9**	24.7	77.5	80.9	30.5	**6.6**	27.2	78.3	81.6	27.9	**6.8**	28.8
		-0.2	12.0	13.6	58.7	95.1	51.0	11.7	14.2	61.2	95.7	50.3	11.1	14.8	61.4
		0.2	86.2	41.3	8.5	10.4	97.7	83.0	38.4	8.8	11.7	97.2	81.6	37.1	8.9
0	0.5	0.5	19.3	**4.1**	15.7	57.4	61.1	19.2	**4.4**	16.9	59.6	60.9	18.1	**4.2**	17.6
		0.8	5.8	27.8	69.7	92.4	9.2	6.1	28.8	70.2	78.4	9.7	6.2	29.9	53.5
		-0.8	100	100	99.9	97.3	99.4	99.4	96.2	80.2	78.0	94.6	84.6	51.2	37.5
0.5	-0.5	-0.5	100	100	99.7	97.4	99.4	99.2	93.7	78.6	86.4	94.4	77.7	43.7	45.7
		-0.2	100	100	99.6	99.0	99.0	97.1	89.8	85.9	96.9	84.7	59.6	45.1	69.7
		0.2	98.1	85.8	42.6	25.4	98.5	92.1	59.0	16.9	15.2	97.8	84.9	41.4	10.3
0.5	0.5	0.5	70.3	37.8	36.1	63.6	76.3	37.1	11.9	21.9	61.0	65.5	22.8	5.4	19.5
		0.8	29.6	44.9	73.0	91.8	21.0	11.8	33.6	69.9	79.7	12.5	7.0	30.9	52.3

d) $\widetilde{LMW}$ -type test: Power ($T = 200$)

		d_0	-0.2				0					0.2			
μ_1	α_0	$\alpha_1 \setminus d_1$	-0.4	-0.2	0	0.2	-0.4	-0.2	0	0.2	0.4	-0.2	0	0.2	0.4
		-0.8	92.0	65.2	49.9	71.3	99.7	92.0	63.1	51.0	71.5	99.6	91.8	62.7	51.8
0	-0.5	-0.5	32.0	**6.8**	24.3	77.0	82.7	31.0	**6.5**	26.4	77.8	83.6	28.6	**6.2**	27.5
		-0.2	30.9	45.4	85.4	99.3	57.7	30.0	45.2	86.0	99.40	56.9	29.0	46.0	86.8
		0.2	93.4	50.6	13.4	21.4	99.7	90.9	45.1	13.2	21.2	99.6	90.3	42.5	13.1
0	0.5	0.5	32.4	**6.6**	23.4	74.8	81.1	28.0	**6.2**	23.1	75.0	80.6	26.2	**5.9**	23.2
		0.8	12.6	51.6	93.2	99.8	11.80	11.9	52.0	93.4	99.7	11.0	12.3	54.2	93.6
		-0.8	100	100	100	97.8	100	100	99.5	91.3	85.0	99.9	98.7	87.4	68.5
0.5	-0.5	-0.5	100	100	99.5	96.9	99.7	99.3	92.3	75.7	86.1	95.2	77.4	42.2	45.6
		-0.2	100	100	99.9	99.9	99.1	98.3	95.8	95.7	99.6	86.5	72.0	70.5	89.9
		0.2	98.9	91.6	57.5	41.4	99.4	95.4	66.6	26.0	27.4	99.7	92.0	49.6	16.1
0.5	0.5	0.5	81.3	52.8	50.4	79.8	84.6	48.6	18.4	31.6	75.9	82.0	33.1	9.0	26.0
		0.8	46.7	68.8	94.8	99.8	27.5	20.8	57.6	94.0	99.7	15.3	15.4	56.1	93.9

Note: Rejection probabilities of 5% test for joint break in d, α and μ at $\lambda_0 = 0.5$, $\mu_0 = 0, \sigma_0^2 = 1$. Figures in bold characters correspond to simulated sizes.

Table 8: **Simulated performance of sequential testing procedure for the number of breaks with LMW-type tests for a joint break in long memory and level (unknown break fractions)**

	Detected number of breaks				Break fraction estimation	
	0	1	2	3	$\hat{\lambda}_0$	$\hat{\lambda}_1$
a) $(d_0, d_1, d_2) = (0, 0.4, 0)$ and $(\mu_0, \mu_1, \mu_2) = (0, 0.5, 0)$						
T=200	60.6	17.4	**20.2**	1.8	0.37 (0.077)	0.65 (0.057)
T=500	15.5	7.0	**71.9**	5.6	0.35 (0.052)	0.65 (0.050)
b) $(d_0, d_1, d_2) = (-0.4, 0, 0.4)$ and $(\mu_0, \mu_1, \mu_2) = (0, 0.5, 1)$						
T=200	1.4	47.1	**46.2**	5.3	0.34 (0.046)	0.67 (0.075)
T=500	0	16.3	**78.0**	5.7	0.34 (0.023)	0.67 (0.057)
c) $(d_0, d_1, d_2) = (0, 0.2, -0.2)$ and $(\mu_0, \mu_1, \mu_2) = (0, 0.5, 1)$						
T=200	1.3	79.8	**16.8**	2.1	0.35 (0.077)	0.66 (0.076)
T=500	0	39.3	**54.6**	6.1	0.35 (0.060)	0.66 (0.046)

Note: Proportion (in %) of detected number of breaks (left) and means and standard errors (in brackets) of break fraction estimates for the cases in which correctly two breaks are detected (right).
5% test for joint break in d and μ, $\epsilon = 0.15$, $\lambda_0 = 1/3$, $\lambda_1 = 2/3$, $\sigma_0^2 = 1$.

Table 9: **Breaks in the forward discount series**

a) Detection of breaks by the LMW-type test with AR lag order selection by the BG procedure

Country	number of breaks	break dates
Canada	1	1992(10)
France	2	1988(3), 1992(9)
Germany	1	1992(9)
Italy	2	1987(12), 1994(2)
U.K.	3	1981(8), 1990(2), 1992(7)

Note: Break dates from applying the sup-$\widetilde{LMW}_T$ -type test for joint breaks in d, μ and α.

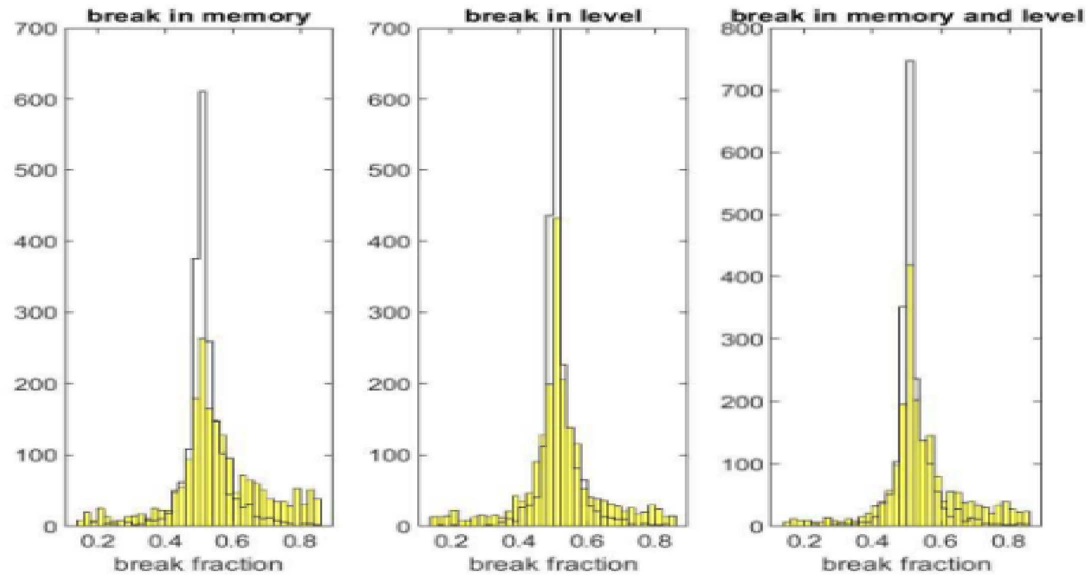
b) Detection of the source of breaks by the robustified LMW-type test

Country	break dates		
Canada	1992(10): $(d, \alpha)^{*}, \mu^{**}$		
France	1988(3): $(d, \alpha)^{**}, \mu^{**}$	1992(9): $(d, \alpha)^{**}, \mu^{**}$	
Germany	1992(9): μ^{**}		
Italy	1987(12): $(d, \alpha)^{**}, \mu^{**}$	1994(2): $(d, \alpha)^{**}, \mu^{**}$	
U.K.	1981(8): μ^{**}	1990(2): $(d, \alpha)^{**}, \mu^{**}$	1992(7): $(d, \alpha)^{**}, \mu^{**}$

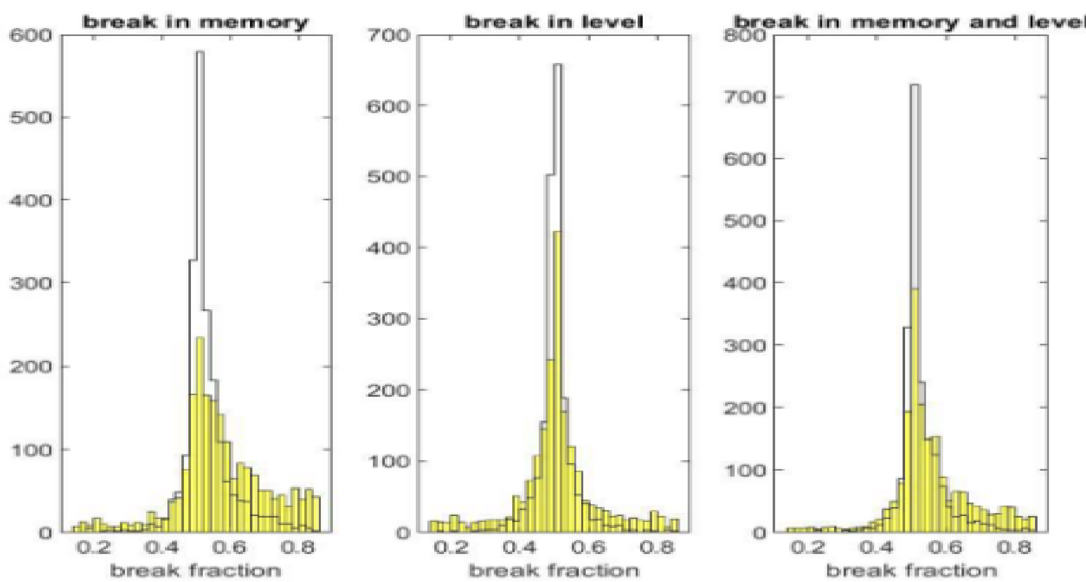
Note: * and ** denote statistical significance of the $\overline{LMW}_T$ -type test at the 10% and 5% levels, respectively.

Figure 2: **Distribution of break fraction estimates for LM and LMW-type tests**

a) LM test

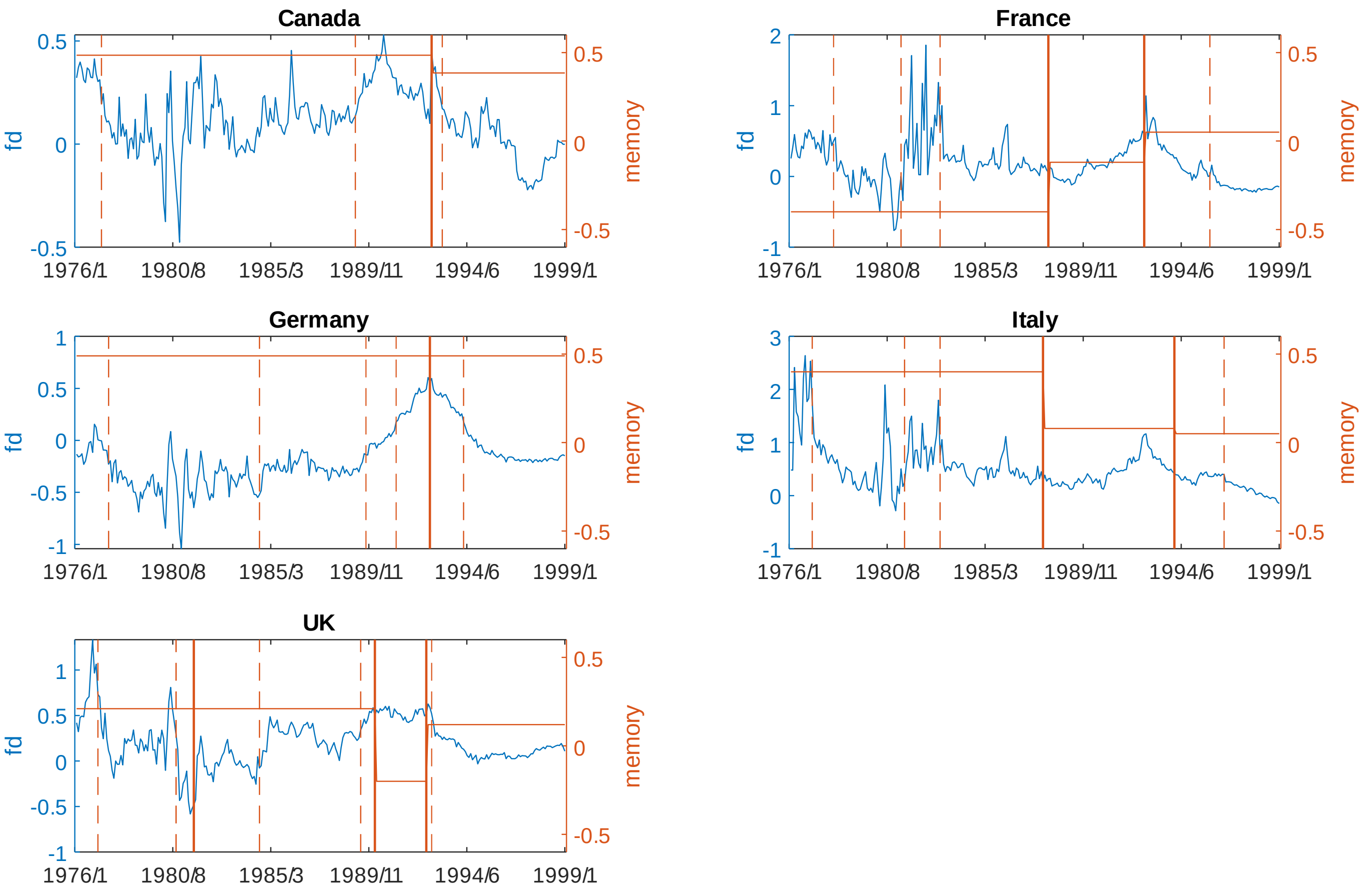


b) LMW-type test



Note: Break at $\lambda_0 = 0.5$ in memory from 0 to 0.4 and/or in level from 0 to 0.5. T=200 (yellow) and 500 (blank).

Figure 3: **Forward discount series: break dates and long memory estimates**



Note: Vertical dashed and solid lines display break dates reported by Choi and Zivot (2007) and by the LMW-type test procedure respectively. Horizontal lines indicate the estimated long memory parameters within each regime.