Article:

Iacone, Fabrizio orcid.org/0000-0002-2681-9036, Leybourne, Stephen and Taylor, A M Robert (2017) Testing for a change in mean under fractional integration. Journal of Time Series Econometrics. ISSN 1941-1928

https://doi.org/10.1515/jtse-2015-0006

Fabrizio Iacone*, Stephen J. Leybourne and A.M. Robert Taylor

Testing for a Change in Mean under Fractional Integration

DOI 10.1515/jtse-2015-0006

Abstract: We consider testing for the presence of a change in mean, at an unknown point in the sample, in data that are possibly fractionally integrated, and of unknown order. This testing problem has recently been considered in a number of papers, most notably Shao (2011, "A Simple Test of Changes in Mean in the Possible Presence of Long-Range Dependence." *Journal of Time Series Analysis* 32:598–606) and Iacone, Leybourne, and Taylor (2013b, "A Fixed-b Test for a Break in Level at an Unknown Time under Fractional Integration." *Journal of Time Series Analysis* 35:40–54) who employ Wald-type statistics based on OLS estimation and rely on a self-normalization to overcome the fact that the standard Wald statistic does not have a well-defined limiting distribution across different values of the memory parameter. Here, we consider an alternative approach that uses the standard Wald statistic but is based on quasi-GLS estimation to control for the effect of the memory parameter. We show that this approach leads to significant improvements in asymptotic local power.

Keywords: change in mean, fractional integration, Wald statistic
JEL Classification: C22

1 Introduction

In this paper we revisit the problem of testing for the presence of a change (at an unknown point) in the mean of a series, that is possibly fractionally integrated, with memory parameter δ, $I(\delta)$. It is well known that if not accounted for, a mean shift in a short memory process can induce features in the autocorrelation function and the periodogram of a time series that can be mistaken as evidence of long memory; see, for example, Diebold and Inoue (2001), Granger and Hyung (2004), Mikosch and Stărică (2004) and Iacone (2010). To avoid the possibility of spurious inference being made about the memory properties of a time series it is

***Corresponding author: Fabrizio Iacone,** Department of Economics and Related Studies, University of York, York, UK, E-mail: fabrizio.iacone@york.ac.uk
Stephen J. Leybourne, School of Economics and Granger Centre for Time Series Econometrics, University of Nottingham, Nottingham, UK
A.M. Robert Taylor, Essex Business School, University of Essex, Colchester, UK

therefore important to be able to detect a mean change and several recent papers have addressed this testing problem using sup-Wald based statistics. Berkes et al. (2006) and Qu (2011) suggest tests to discriminate between a null hypothesis of long memory without mean change and an alternative of change in the mean in an otherwise weakly autocorrelated process. A rather less restrictive approach is taken by Wang (2008), Shao (2011) and Iacone, Leybourne, and Taylor (2013b). Here the null is of no mean change, and the alternative is of mean change, but with no restriction on the memory properties of the series under either the null or alternative hypothesis (i. e. the mean change can be associated with either a short or long memory process).

Focussing on the latter group of tests because of their wider applicability, the approach they take is essentially based on OLS estimation in that the (sup-) Wald-type statistic used to test for the presence of a change in mean is that which would be computed were it known that $\delta = 0$. Self-normalisations applied to the raw test statistics ensure that they have well-defined limiting null distributions which depend only on the memory parameter, δ. It is, however, unlikely that these OLS-based statistics will provide the best available power for detecting a mean change when $\delta \neq 0$, or even when $\delta = 0$ (due to the normalisations involved). It therefore makes sense to investigate an approach that employs a quasi-GLS estimation procedure appropriate for a given δ; that is, one based on taking Δ^{δ}-differences of the series under consideration, and constructing a non-normalised (sup-) Wald statistic. We show that the corresponding quasi-GLS based Wald statistic again has a limit null distribution depending only on the memory parameter. We also demonstrate that its local asymptotic power function is significantly higher than those of the OLS-based tests of Shao (2011) and Iacone, Leybourne, and Taylor (2013b).

In what follows we use the notation: $x := y$ ($x =: y$) to indicate that x is defined by y (y is defined by x); $[\cdot]$ to denote the integer part; $\mathbb{I}(\cdot)$ to denote the indicator function whose value is one when its argument is true and zero otherwise; L to denote the lag operator; and $\xrightarrow{d}$ to denote convergence in the Skorohod J_1 topology of $D[0,1]$, the space of real-valued functions on $[0,1]$ which are continuous on the right and with finite left limit, respectively, as the sample size diverges.

2 A Quasi-GLS based Test for a Break in the Mean and Its Limit Distribution

Consider the scalar time series process, y_t, satisfying the data generating process [DGP]:

$$y_t = \alpha + \beta \mathbb{I}\,(t > T_0) + u_t, \quad t = 1, \ldots, T, \qquad [1]$$

$$y_t = 0, \quad t \le 0. \qquad [2]$$

where $T_0 := \lfloor \tau_0 T \rfloor$ for $\tau_0 \in [\tau_L, \tau_U] =: \Lambda \subset (0, 1)$; that is, $T_0 \in \Lambda_T := \{\lfloor \tau_L T \rfloor, \ldots, \lfloor \tau_U T \rfloor\}$ (τ_L and τ_U representing trimming parameters). Here u_t is a zero-mean $I(\delta)$ process. When $\beta = 0$, $E(y_t) = \alpha$ for any $t > 0$, so the mean is constant, but when $\beta \ne 0$, the mean changes at observation T_0 from α to $\alpha + \beta$. One concern lies with testing the null hypothesis $H_0 : \beta = 0$ against the two-sided alternative hypothesis $H_1 : \beta \ne 0$, without assuming knowledge of the location of the breakpoint T_0.

We obtain u_t in (1) by integrating an $I(0)$ process, η_t say, δ times; that is, let η_t be a scalar, zero-mean covariance stationary process with finite and non-zero spectral density at all frequencies, then

$$u_t := \Delta^{-\delta}\{\eta_t \mathbb{I}(t > 0)\}$$

when $t \ge 1$, $u_t := 0$ when $t \le 0$. For $\delta \in (-0.5, 0.5)$, $\Delta^{-\delta} = (1 - L)^{-\delta}$ can be expanded as $\Delta^{-\delta} = \sum_{t=0}^{\infty} \Delta_t^{(\delta)} L^t$, where $\Delta_t^{(\delta)} := \Gamma(t + \delta)/(\Gamma(\delta)\Gamma(t + 1))$, with $\Gamma(\cdot)$ denoting the Gamma function, with the conventions that $\Gamma(0) := \infty$ and $\Gamma(0)/\Gamma(0) := 1$. Therefore, when $t \ge 1$ the process u_t can be written as $u_t = \sum_{s=-\infty}^{t} \Delta_{t-s}^{(\delta)} \{\eta_s \mathbb{I}(s > 0)\}$, and, noting $\mathbb{I}(s > 0)$, as $u_t := \sum_{s=1}^{t} \Delta_{t-s}^{(\delta)} \eta_s$.

Remark 1: The process u_t is $I(\delta)$ and it belongs to the class of Type II fractionally integrated processes; see, for example, Marinucci and Robinson (1999). The assumption that $\delta \in (-0.5, 0.5)$ is common in the literature, and is also made in Shao (2011) and Iacone, Leybourne, and Taylor (2013b). Notice that α and β cannot be estimated consistently if $\delta > 0.5$. □

We follow the approach of Iacone, Leybourne, and Taylor (2013a), and take Δ^δ-differences of y_t in (1), to obtain

$$\Delta^\delta y_t = \alpha \Delta^\delta \{1\mathbb{I}(t > 0)\} + \beta \Delta^\delta \{1\mathbb{I}(t > T_0)\} + \Delta^\delta u_t, \quad t = 1, \ldots, T \qquad [3]$$

where, by definition, $\Delta^\delta u_t = \eta_t$. Let $\tau \in \Lambda$ denote a generic break fraction, with $T_a := \lfloor \tau T \rfloor \in \Lambda_T$ the associated break date. To keep the notation manageable, we also introduce

$$\mu_t := \Delta^\delta \{1\mathbb{I}(t > 0)\},\ \mu_t(\tau) := \Delta^\delta \{1\mathbb{I}(t > T_a)\}$$

so that (3) can be written more succinctly as

$$\Delta^\delta y_t = \alpha \mu_t + \beta \mu_t(\tau_0) + \eta_t.$$

Because the location of the putative mean shift is not assumed known, we consider the Wald statistic to test $H_0: \beta = 0$ when evaluated at the generic break point T_a: for $x_t(\tau;\delta) := [\mu_t, \mu_t(\tau)]'$, let $M_{xx}(\tau;\delta) := \sum_{t=1}^{T} x_t(\tau;\delta) x_t(\tau;\delta)'$, $M_{xy}(\tau;\delta) := \sum_{t=1}^{T} x_t(\tau;\delta)\Delta^{\delta} y_t$, and, for $c := [0,1]'$, $\widehat{\beta}(\tau;\delta) := cM_{xx}(\tau;\delta)^{-1}M_{xy}\ (\tau;\delta)$. The Wald statistic is given by

$$\mathcal{W}(\tau;\delta) := \frac{\widehat{\beta}(\tau;\delta)^2}{\sigma^2 cM_{xx}(\tau;\delta)^{-1}c'} \qquad [4]$$

where σ^2 is the long run variance of η_t. We then consider the supremum of $\mathcal{W}(\tau;\delta)$ taken over $\tau \in \Lambda$; that is,

$$\mathcal{SW}_\delta := \sup_{\tau \in \Lambda} \mathcal{W}(\tau;\delta)$$

As a practical matter, the statistic $\mathcal{W}(\tau;\delta)$ in (4) is computed for each of the candidate dates $T_a \in \Lambda_T$ and the maximum value of these is then taken.

In Assumption 1 we now state the necessary regularity conditions such that we can evaluate the limiting distribution of $\mathcal{SW}_\delta$ under a sequence of local alternatives of the form $H_{1L}: \beta = \kappa\sigma T^{\delta-1/2}$. This will be subsequently given in Theorem 1.

Assumption 1: *Let η_t be the linear process satisfying $\eta_t := A(L)\ \varepsilon_t := \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}$. The weights $\{A_j\}$ are such that $\sum_{j=0}^{\infty} j|A_j| < \infty$ and ε_t is an independent, identically distributed sequence with $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = 1$, and $E(|\varepsilon_t|^q) < \infty$, for $q > \max\{4, \frac{2}{1-2\delta}\}$.*

Remark 2: Under Assumption 1, the long run variance of η_t is given as $\sigma^2 = A(1)^2$. □

Because η_t is $I(0)$ and $x_t(\tau;\delta)$ comprises deterministic regressors, it is straightforward to establish that for a fixed value of τ the Wald statistic $\mathcal{W}(\tau;\delta)$ has a χ_1^2 limiting distribution under $H_0: \beta = 0$ and under the regularity conditions of Assumption 1. However, because the location of the potential break is unknown and consequently we take the supremum of $\mathcal{W}(\tau;\delta)$ over all possible values of τ, we need to treat $\mathcal{W}(\tau;\delta)$ as a function of τ. As a consequence, we need to establish the limit of this function rather than simply establish the marginal limit for a given value of τ. The key to doing so is to use the result that under Assumption 1,

$$T^{\delta-1/2} \sum_{t=1+\lfloor \tau T \rfloor}^{T} (t - \lfloor \tau T \rfloor)^{-\delta} \eta_t \xrightarrow{d} \sigma \int_{\tau}^{1} (r-\tau)^{-\delta} dB(r) \qquad [5]$$

where $B(r)$ denotes a standard Brownian motion process on $r \in [0, 1]$; see Lemma A.1 of Iacone, Leybourne, and Taylor (2013a).

Before we characterise the limiting distribution of $\mathcal{SW}_\delta$ in Theorem 1, we first need to introduce some additional notation. To that end, we define

$$C(\tau;\ \delta) := \frac{1}{(\Gamma(1-\delta))^2}\begin{bmatrix} \int_0^1 r^{-2\delta}dr & \int_\tau^1 r^{-\delta}(r-\tau)^{-\delta}dr \\ \int_\tau^1 r^{-\delta}(r-\tau)^{-\delta}dr & \int_\tau^1 (r-\tau)^{-2\delta}dr \end{bmatrix}$$

$$D(\tau;\ \delta) := \frac{1}{\Gamma(1-\delta)}\begin{bmatrix} \int_0^1 r^{-\delta}dB(r) \\ \int_\tau^1 (r-\tau)^{-\delta}dB(r) \end{bmatrix}$$

$$\Psi(\tau, \tau_0;\ \delta) := \frac{1}{(\Gamma(1-\delta))^2}\begin{bmatrix} \int_{\tau_0}^1 r^{-\delta}(r-\tau_0)^{-\delta}dr - \int_\tau^1 r^{-\delta}(r-\tau)^{-\delta}dr \\ \int_{\tau_M}^1 (r-\tau)^{-\delta}(r-\tau_0)^{-\delta}dr - \int_\tau^1 (r-\tau)^{-2\delta}dr \end{bmatrix}$$

where $\tau_M := \max(\tau, \tau_0)$.

Theorem 1: *Let y_t be generated according to* (1)-(2) *and let Assumption 1 hold. Under $H_{1L} : \beta = \kappa\sigma T^{\delta-1/2}$,*

$$\mathcal{SW}_\delta \xrightarrow{d} \sup_{\tau\in\Lambda} \mathcal{L}(\tau, \tau_0, \kappa;\ \delta) \qquad [6]$$

where

$$\mathcal{L}(\tau, \tau_0, \kappa;\ \delta) := \frac{\left(c[C(\tau;\ \delta)]^{-1}D(\tau;\ \delta) + \kappa\left(1 + c[C(\tau;\ \delta)]^{-1}\,\Psi(\tau, \tau_0;\ \delta)\right)\right)^2}{c[C(\tau;\ \delta)]^{-1}c'}. \qquad [7]$$

Remark 3: Setting $\kappa = 0$ in (6) and (7) in Theorem 1 it immediately follows that under the null hypothesis, $H_0 : \beta = 0$, $\mathcal{SW}_\delta \xrightarrow{d} \sup_{\tau\in\Lambda} \mathcal{L}(\tau, \tau_0, 0;\ \delta)$ where $\mathcal{L}(\tau, \tau_0, 0;\ \delta) := (c[C(\tau;\ \delta)]^{-1}D(\tau;\ \delta))^2/c[C(\tau;\ \delta)]^{-1}c'$. □

Remark 4: The proof of Theorem 1 follows along similar lines to the proof of Theorem 1 of Iacone, Leybourne, and Taylor (2013a) and is therefore omitted. Observe that the limiting distribution given for $\mathcal{SW}_\delta$ in Theorem 1 does not depend on σ^2 nor does it depend on the form of the weights A_j, as long as Assumption 1 is met. Under H_0, both the numerator and the denominator, when scaled by appropriate functions of T, converge to limits that depend on σ^2; see for example (5) and notice the presence of σ^2 in (4). However, and as is common with Wald-type statistics, under the null hypotheses these terms cancel out from the limiting distribution. Under H_{1L} the additional non-centrality term present in the limit distribution of $\mathcal{SW}_\delta$ does not depend on σ^2 by virtue of the fact that the local break magnitude in H_{1L} is scaled by σ. □

As the limiting null distribution of $\mathcal{SW}_\delta$ depends on δ, the critical values for the test also depend on δ. Moreover, the critical values also depend on Λ, as does the test statistic $\mathcal{SW}_\delta$. Both these observations are typical of tests of this kind; cf. Shao (2011) and Iacone, Leybourne, and Taylor (2013b).

Table 1 below reports asymptotic null (i. e. $\kappa = 0$) critical values for $\mathcal{SW}_\delta$, for $\Lambda = [0.15, 0.85]$ and various values of δ, $\delta = \{-0.4, -0.3, \ldots, 0.0, \ldots 0.3, 0.4\}$, for each of the 0.10, 0.05 and 0.01 (upper tail) significance levels. The results were obtained using Gauss 9.0, via direct simulation of the limiting distribution $\mathcal{L}(\tau, \tau_0, 0; \delta)$ from Theorem 1, using a discretisation of r that uses 2000 steps, $B(r)$ simulated using *IID* $N(0,1)$ variates, and 10000 Monte Carlo replications. With regard to the trimming parameters, we set $\tau_L = 0.15$ and $\tau_U = 0.85$. Our choice of Λ follows Shao (2011) and Iacone, Leybourne, and Taylor (2013b), to facilitate comparison, and is common in the literature.

Table 1: Asymptotic critical values of the $\mathcal{SW}_\delta$ test with $\Lambda = [0.15, 0.85]$.

Sig. level δ	**−0.4**	**−0.3**	**−0.2**	**−0.1**	**0**	**0.1**	**0.2**	**0.3**	**0.4**
0.10	5.38	5.69	6.06	6.52	7.15	7.99	9.18	10.62	12.62
0.05	6.87	7.14	7.47	8.02	8.70	9.46	10.74	12.37	14.16
0.01	9.94	10.20	10.73	11.30	11.96	13.01	14.39	15.85	17.72

3 Asymptotic Local Power

We next analyze the asymptotic local power properties of the $\mathcal{SW}_\delta$ test as functions of κ, with the local limit distributions $\mathcal{L}(\tau, \tau_0, \kappa; \delta)$ being simulated in the same way as in Section 2 above. We evaluate powers for nominal 0.05-level tests, using the asymptotic critical values from Table 1. For δ we consider $\delta \in \{-0.40, -0.20, 0.00, 0.20, 0.40\}$. Although we considered $\tau_0 \in \{0.25, 0.50, 0.75\}$ as values for the break fraction, we only report results here for $\tau_0 = 0.5$ because the results for $\tau_0 = 0.25$ and $\tau_0 = 0.75$ were qualitatively similar to those for $\tau_0 = 0.5$ and are therefore omitted in the interests of brevity. The results for $\tau_0 = 0.5$ are shown in Figure 1, where we also give asymptotic local powers (under the assumption of a Type II fractionally integrated process) of the self-normalized OLS based tests of Shao (2011) and Iacone, Leybourne, and Taylor (2013b), denoted $\mathcal{SG}$ and $\mathcal{SW}$ respectively. Here the $\mathcal{SW}$ test is constructed using a fixed-b estimate of the long run variance with $b = 0.1$ and a Bartlett kernel; see Kiefer and Vogelsang (2005). Notice that the $\mathcal{SG}$ statistic implicitly employs $b = 1$.

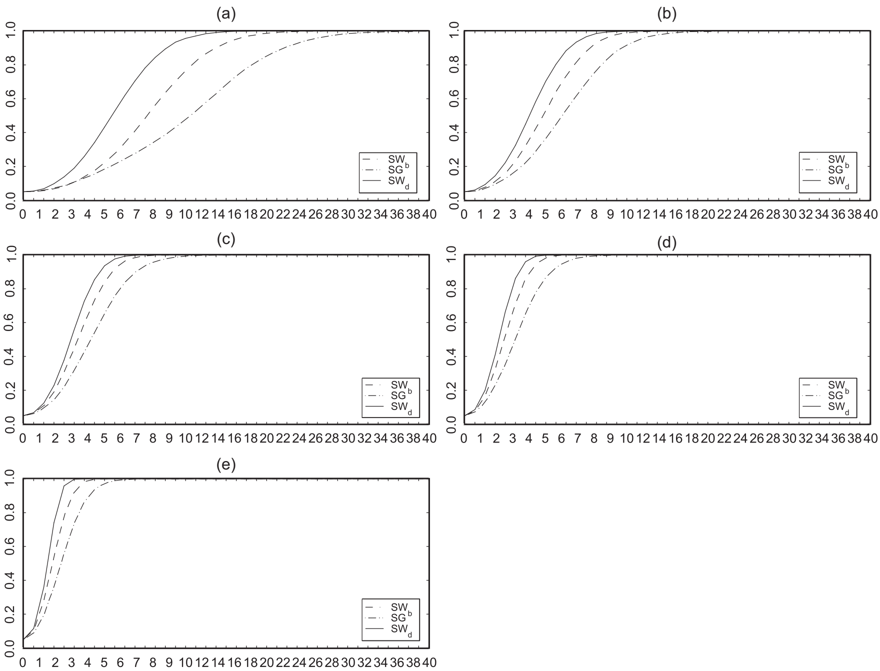

Figure 1: (a) Power of the texts for $\delta = -0.4$, $\tau_0 = 0.5$; (b) Power of the texts for $\delta = -0.2$, $\tau_0 = 0.5$; (c) Power of the texts for $\delta = 0$, $\tau_0 = 0.5$; (d) Power of the texts for $\delta = 0.2$, $\tau_0 = 0.5$; (e) Power of the texts for $\delta = 0.4$, $\tau_0 = 0.5$.

It is immediately evident from Figure 1 that, for all the values of δ considered the $\mathcal{SW}_\delta$ test offers higher asymptotic power than either of the OLS based tests, and we also observe that $\mathcal{SG}$ is outperformed by $\mathcal{SW}$. Broadly speaking, the power advantage of $\mathcal{SW}_\delta$ is at its strongest the farther δ is away from 0, and is substantial in such cases; however, $\mathcal{SW}_\delta$ still has notably higher asymptotic local power than the $\mathcal{SW}$ test even for $\delta = 0$. This last result is a consequence of the self-normalization being employed by the $\mathcal{SW}$ statistic.

Finally, for practical application one would require a feasible version of the $\mathcal{SW}_\delta$ statistic, with δ and σ^2, replaced by consistent estimators thereof, $\hat{\delta}$ and $\hat{\sigma}^2$ say (the OLS based tests require estimation of δ, but not σ^2 as such). It can be shown that provided $\hat{\delta} - \delta = O_p(T^{-\lambda})$ for some $\lambda > \max(0, \delta)$ and $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, then the feasible variant of $\mathcal{SW}_\delta$ obtains the limiting properties stated for $\mathcal{SW}_\delta$ in Theorem 1. Iacone, Leybourne, and Taylor (2013a) discuss suitable local Whittle-type choices for $\hat{\delta}$ (and $\hat{\sigma}^2$) which are easily adapted to the current modelling framework.

References

Berkes, I., L. Horváth, P. Kokoszka, and Q. Shao. 2006. "On Discriminating Between Long-Range Dependence and Changes in Mean." *Annals of Statistics* 34:1140–65.

Diebold, F.X., and A. Inoue. 2001. "Long Memory and Structural Change." *Journal of Econometrics* 105:131–59.

Granger, C.W.J., and N. Hyung. 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.

Iacone, F. 2010. "Local Whittle Estimation of the Memory Parameter in Presence of Deterministic Components." *Journal of Time Series Analysis* 31:37–49.

Iacone, F., S.J. Leybourne, and A.M.R. Taylor. 2013a. "Testing for a Break in Trend When the Order of Integration Is Unknown." *Journal of Econometrics* 176:30–45.

Iacone, F., S.J. Leybourne, and A.M.R. Taylor. 2013b. "A Fixed-*b* Test for a Break in Level at an Unknown Time Under Fractional Integration." *Journal of Time Series Analysis* 35:40–54.

Kiefer, N.M., and T.J. Vogelsang. 2005. "A New Asymptotic Theory for Heteroskedasticity – Autocorrelation Robust Tests." *Econometric Theory* 21:1130–64.

Marinucci, D., and P.M. Robinson. 1999. "Alternative Forms of Fractional Brownian Motion." *Journal of Statistical Planning and Inference* 80:111–22.

Mikosch, T., and C. Stărică. 2004. "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects." *Review of Economics and Statistics* 86:378–90.

Qu, Z. 2011. "A Test Against Spurious Long Memory." *Journal of Business and Economic Statistics* 29:423–38.

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.

Wang, L. 2008. "Change-in-Mean Problem for Long Memory Time Series Models with Applications." *Journal of Statistical Computation and Simulation* 78:653–68.

Note: We thank the Editor and an anonymous referee for their helpful and constructive comments on an earlier version of this paper.