QUANTILE REGRESSION FOR LONG MEMORY TESTING:
A CASE OF REALIZED VOLATILITY

Uwe Hassler

Paulo M. M. Rodrigues

Antonio Rubia

Please address correspondence to
Paulo M. M. Rodrigues
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal;
Tel.: 351 21 313 0831, email: Paulo M.M. Rodrigues@bportugal.pt

Banco de Portugal
EUROSYSTEM

Quantile Regression for Long Memory Testing: A case of Realized Volatility[*]

Uwe Hassler[a], Paulo M.M. Rodrigues[b] and Antonio Rubia[c]

[a] Goethe University Frankfurt

[b] Banco de Portugal, Universidade Nova de Lisboa and CEFAGE

[c] University of Alicante

Abstract

In this paper we derive a quantile regression approach to formally test for long memory in time series. We propose both individual and joint quantile tests which are useful to determine the order of integration along the different percentiles of the conditional distribution and, therefore, allow to address more robustly the overall hypothesis of fractional integration. The null distributions of these tests obey standard laws (e.g., standard normal) and are free of nuisance parameters. The finite sample validity of the approach is established through Monte Carlo simulations, showing, for instance, large power gains over several alternative procedures under non-Gaussian errors. An empirical application of the testing procedure on different measures of daily realized volatility is presented. Our analysis reveals several interesting features, but the main finding is that the suitability of a long-memory model with a constant order of integration around 0.4 cannot be rejected along the different percentiles of the distribution, which provides strong support to the existence of long memory in realized volatility from a completely new perspective.

Keywords: Fractional integration, Lagrange multiplier, integrated variance, LAD estimator.
JEL: C12, C22.

1

1 Introduction

There is a growing interest in finance and economics in modelling and forecasting dependence in the tails of the conditional distribution of a time series. An efficient way to address this issue is through the quantile regression (QR) approach introduced by Koenker and Bassett (1978), now routinely implemented in market downside risk management and other applied areas. The QR analysis distinctively deals with estimation and inference at different quantiles, allowing to address a wide range of hypotheses and offer new insights on the time-series properties of the data. For instance, using these techniques, Engle and Manganelli (2004) show that the conditional Value-at-Risk of daily returns is a strongly persistent process. The reason is that daily downside risk measures are mainly driven by volatility, which typically exhibits long-range dependence possibly generated by a fractionally-integrated process. Similarly, Koenker and Xiao (2004) report evidence of strongly persistent, yet heterogenous dynamics in U.S. short-term interest rate along the deciles of the conditional distribution. The QR analysis reveals that the largest autoregressive coefficient varies significantly from bottom to top quantiles, showing asymmetric patterns ranging from stationarity to explosiveness which can be related to different policy strategies of the Federal Reserve Board.

In this paper, we contribute to the extant literature by proposing a novel quantile regression test to detect long memory (also known as fractional integration) in the time series context. This general class of models allows for long-run dependence characterized by autocovariances that decay hyperbolically, thereby offering an intermediate case between the exponential decay of short memory and the infinite persistence of unit root processes. Consequently, long-memory models often explain convincingly the time-series dynamics exhibited by many economic and non-economic time series; see, Henry and Zaffaroni (2003) for a review. We propose a series of Lagrange Multiplier (LM) semiparametric tests for fractional integration that extend the regression procedure in Breitung and Hassler (2002) to the QR setting and which allows us to address more general hypotheses than the unit root case analyzed in this literature. By inverting these test statistics, furthermore, confidence-interval estimates of the long-memory parameter can readily be obtained.

More specifically, we discuss the asymptotic theory for both individual and joint quantile regression long memory tests (QRLM henceforth) under a fairly general class of errors in the data generating process. Individual quantile tests are intended to address the fractional integration hypothesis at a specific quantile $0 < \tau < 1$, with the conditional median $\tau = 1/2$ being a leading example, while joint tests involve sets of quantiles in closed subintervals of $(0, 1)$. We show that the asymptotic null distributions of QRLM tests do not depend on the degree of integration in the observable series nor on other nuisance parameters, and can be characterized by usual probability laws, which is in sharp contrast to existing tests for the unit root hypothesis. Individual quantile tests are asymptotically distributed as a standard normal or a Chi-squared distribution, while joint tests across quantiles are distributed according to

truncated versions of the well-known Kolmogorov and Cramer von Mises distributions. Monte Carlo experimental analysis shows that QRLM tests can yield large power improvements over suitable alternatives and tend to offer more robust inference against observations drawn from heavy-tailed distributions.

The QRLM tests allow us to gain greater insight into the dynamics of the underlying time-series process. Applying our novel procedure, we analyze the long-range properties of different measures of daily realized volatility of IBM, one of the most liquid and frequently-traded securities in the U.S. stock exchange. Realized volatility time-series characteristically display empirical autocorrelation functions that slowly decay to zero and which cannot be explained by unit root models. Consequently, a number of studies have successfully used fractionally-integrated models to capture long-range dependence and forecast these series, outperforming GARCH and stochastic volatility models; see, for instance, Andersen, Bollerslev, Diebold and Labys (2003). Nevertheless, there is currently a debate discussing whether long memory is really present in these series or the spurious consequence of model misspecification. QRLM tests allow us to determine the order of fractional integration along the different quantiles of the conditional distribution, thereby addressing more robustly this hypothesis and bringing completely new evidence to the field.

The main findings of this analysis can be summarized as follows. First, we observe meaningful differences between realized volatility measured in levels and their logarithmic transforms. In levels, the QRLM tests strongly reject the hypothesis that filtering the long-run component with a fractionally integrated model renders innovations that follow a stationary ARMA-type model. The existence of sheer differences between low and high levels of volatility, related to different regimes in these series, is the most likely reason underlying this rejection; see, among others, Diebold and Inoue (2001), Maheu and McCurdy (2002) and Baillie and Kapetanios (2007). Second, the logarithmic transform, widely used in realized volatility modelling (*e.g.*, Andersen *et al.* 2003), considerably regularizes the data and reduces the heterogeneity caused by different regimes. After applying this nonlinear transformation, the QRLM tests show that a constant long-memory parameter model may explain well the long-run of observations belonging to quantiles in the left tail and center deciles of the conditional distribution. Confidence interval-based estimators, computed by inverting the QRLM tests, infer the most likely values of the long-memory coefficient of around 0.4, the standard value reported in related literature. At top deciles, the individual analysis uncovers an upward trend in the long-memory parameter. This may be consistent with log-realized volatility being driven by different components related to 'normal' and 'high' periods of volatility, yet we also observe confidence-interval estimates that considerably widen at these quantiles, leading to greater parameter uncertainty. In this context, joint QRLM tests are particularly useful to formally determine whether a long-memory model with constant order of integration generates the data. We observe that there is sufficient regularity such that the joint tests across deciles cannot reject the suitability

of a single, constant long-memory model. The confidence-interval estimates of the common underlying long-memory parameter yield values in the region $[0.4, 0.5]$, which agrees with the estimates provided by alternative semiparametric procedures and supports the findings in the extant literature. Therefore, according to our analysis, long-range dependence is caused by a long-memory model.

This paper can be related to different strands of previous research. First, it generalizes the unit-root testing procedures put forward in the QR literature by proposing a test that can identify fractional integration in the data. Previous papers have focused on testing for the unit-root hypothesis, see, among others, Hasan and Koenker (1997), Koenker and Xiao (2004), Ling and McAleer (2004), Thompson (2004), Chan, Peng and Qi (2006), Galvao (2009) and Xiao (2009). Our analysis is more general and nests the unit root hypothesis as a particular case. Second, our paper extends reciprocally the fractional integration testing, traditionally focused on the conditional mean analysis, towards a more general setting involving other aspects of the conditional distribution. The analysis on quantiles and sets of quantiles produces more robust evidence to determine the adequacy of the long-memory filter. QRLM tests are a direct extension of the Least-Squares (LS) based tests proposed in the time-domain by Breitung and Hassler (2002) and further generalized in Demetrescu, Kuzin and Hassler (2008) and Hassler, Rodrigues and Rubia (2009); for further references see Tanaka (1999), Robinson (1991, 1994) and Hassler and Breitung (2006).

In addition, our paper can be related to the literature concerned with stochastic-trend detection when data are drawn from a heavy-tailed distribution. Indeed, an earlier motivation for the QR methodology was that it offers estimates that can exhibit better properties against non-Gaussian features of the data, particularly, excess kurtosis. The well-known Least-Absolute Deviations (LAD) procedure is simply a QR computed at the median of the distribution. Characteristically, most unit-root tests are based on LS estimation, which ensures efficiency under Gaussian conditions, but can lead to finite-sample bias under heavy-tailed distributions. Consequently, the literature on unit-root testing has suggested alternative approaches based on M estimators (Lucas 1995), LAD estimators (Knight 1989; Phillips 1991; Hercé 1996; Li and Li 2009) and the QR generalization of the latter, surveyed previously. In sharp contrast, fractional integration testing has barely received attention in this context. Li and Li (2008) discuss the asymptotic properties of LAD estimators in a fully parametric modelling context for a class of ARFIMA-GARCH models in a Laplace quasi-maximum likelihood estimation setting, while Delgado and Velasco (2005) propose a nonparametric sign test for fractional integration under zero-median errors. The QRLM test at the median $\tau = 1/2$ offers robustness against excess kurtosis and constitutes a valid alternative to these tests. Finally, our paper is related to the empirical literature concerned with realized volatility modelling. We provide both a novel procedure to detect long-memory and bring more robust evidence to the field; see, among others, Andersen, Bollerslev, Diebold and Labys (2001, 2003), Barndorff-Nielsen and Shephard

(2004) and Corsi, Mittnik, Pigorsch and Pigorsch (2008).

The remainder of the paper is organized as follows. Section 2 reviews the LS testing framework for fractional integration proposed by Breitung and Hassler (2002) and discusses the asymptotic behavior of a median-based QRLM testing procedure. Section 3 generalizes this setting to the context of QR, introducing individual and joint tests under different sets of assumptions. Section 4 presents experimental evidence on the finite sample size and power of the test and compares it to several alternative procedures. In section 5, we apply the QRLM tests to characterize the extent of long-run dependence in realized variation of stock prices. The final section summarizes and concludes. A technical appendix collects the proofs of the main theoretical statements of the paper.

In what follows, '$\rightarrow$' and '$\overset{p}{\rightarrow}$' denote weak convergence and convergence in probability, respectively, as the sample size is allowed to diverge, and $\mathbb{I}(\cdot)$ is an indicator function that takes values equal to one if the condition in parenthesis is fulfilled and zero otherwise. Finally, vectors and matrices are represented in bold letters throughout the text.

2   Least-squares and LAD testing

Consider the following data generating process

$$(1 - L)^{d+\theta} y_t = \varepsilon_t, \quad t = 1, ..., T \tag{1}$$

where L denotes the lag operator, $(d, \theta)'$ is a real-valued vector with unknown elements that are not restricted to be integer and, for the moment, we assume $\varepsilon_t \sim iid\mathcal{N}(0, \sigma^2)$. Our main aim is to test the null hypothesis that $\{y_t\}$ is fractionally integrated of order d, denoted as I(d), against the alternative I($d + \theta$), i.e., testing H$_0$: $\theta = 0$ against H$_1$: $\theta \neq 0$ for a given value of d. The standard unit-root model, $d = 1$, is a particular case in this generalized setting, since the long-memory parameter, d, may take any other real value.[1] It is common in the theoretical literature to assume initial conditions characterized by $\varepsilon_k = 0$ for all $k \leq 0$, so that the process is well-defined in mean-square sense. We maintain this condition for the moment, noting that it is not really necessary to derive the null distribution of our test statistics, as is formally proven in the technical appendix.

The theoretical setting described above has been considered, among others, in Robinson (1991, 1994) and Tanaka (1999), who derive score statistics for long memory in the frequency and time domains, respectively. Breitung and Hassler (2002) proposed a variant of the regression procedure introduced by Agiakloglou and Newbold (1994), which has been further

---

[1]The theoretical properties of $\{y_t\}$ in (1) are well-known. An I(d) process generates hyperbolically decaying autocovariances that are not summable for $d > 0$. For values $|d| < 1/2$, $\{y_t\}$ is invertible and stationary, while for $d \geq 1/2$ it does not have finite variance. When ε_t is generated from a stationary ARMA, the process is generally referred to as an ARFIMA model.

extended in Demetrescu *et al.* (2008) and Hassler *et al.* (2009). To illustrate the main features of this test, consider the time series $\{\varepsilon_{t,d}\}$ which corresponds to the values of $\{y_t\}$ differenced under the null, *i.e.*,

$$\varepsilon_{t,d} = (1-L)^d y_t \,, \tag{2}$$

and then define the regressor $x_{t-1,d}^*$ as a weighted partial sum of lags of $\varepsilon_{t,d}$ according to

$$x_{t-1,d}^* = \sum_{j=1}^{t-1} j^{-1} \varepsilon_{t-j,d} \,, \ t = 2, ..., T \tag{3}$$

following the Lagrange Multiplier principle under the Gaussian restriction.

Hence, based on (2) and (3), Breitung and Hassler (2002) show that the LM test for the null hypothesis $H_0 : \theta = 0$, say LM_{LS}, is equivalent to the squared t-statistic on the estimate $\widehat{\phi}$ to test $H_0 : \phi = 0$ in the auxiliary LS regression,

$$\varepsilon_{t,d} = \phi x_{t-1,d}^* + u_t, \tag{4}$$

i.e.,

$$LM_{LS} = \frac{\left(\sum_{t=2}^{T} \varepsilon_{t,d} x_{t-1,d}^* \right)^2}{\widehat{\sigma}_u^2 \sum_{t=2}^{T} \left(x_{t-1,d}^* \right)^2} \tag{5}$$

where $\widehat{\sigma}_u^2$ denotes the LS estimate of the residual variance of (4).

Under the conditions given above and as the sample size diverges, LM_{LS} converges to a $\chi_{(1)}^2$. Hence, the test is asymptotically equivalent to the score tests in Robinson (1991) and Tanaka (1999), both of which are derived under Gaussianity. Although this restriction is not strictly necessary (*i.e.*, $LM_{LS} \to \chi_{(1)}^2$ holds as $T \to \infty$ independently of whether errors are normal or not), it is only under such a restriction that the test achieves the efficiency established in Robinson (1994). Under large departures from the normal distribution, the test is no longer efficient and may suffer from severe power losses in finite samples.

Seeking to robustify inference against large or discordant observations, the ϕ parameter in the auxiliary regression (4) can alternatively be estimated in the LAD setting, defined as

$$\widehat{\phi}_{LAD} = \arg\min_{\phi \in \mathbb{R}} \sum_{t=2}^{T} \left| \varepsilon_{t,d} - \phi x_{t-1,d}^* \right|. \tag{6}$$

Evidently, the null hypothesis $H_0 : \theta = 0$ still implies $H_0 : \phi = 0$. Therefore, paralleling the LS analysis, this can be tested through the test statistic

$$LM_{LAD} = \left[2\widehat{\phi}_{LAD} \widehat{f}(0) \right]^2 \sum_{t=2}^{T} \left(x_{t-1,d}^* \right)^2 \tag{7}$$

where $\widehat{f}(0)$ is a consistent estimate of the (unknown) density of the distribution of the regression residuals, $\widehat{u}_{t,d} = \varepsilon_{t,d} - \widehat{\phi}_{LAD} x_{t-1,d}^*$, evaluated at the origin. The following theorem characterizes

6

the asymptotic distribution of LM_{LAD} under a set of restrictions similar to that maintained by Breitung and Hassler (2002). These will be relaxed considerably in the more general context of QR discussed in the following section.

Theorem 2.1 *Considering $\{y_t\}$ generated as in (1) with $\varepsilon_t \sim iid\,(0, \sigma^2)$ having median zero, and assuming that the innovations are distributed according to a continuous, strictly positive density f in a neighborhood of zero, it follows, under the null hypothesis $H_0 : \phi = 0$ as $T \to \infty$ that,*

$$\sqrt{T}\,\widehat{\phi}_{LAD} \to \mathcal{N}\left(0, \frac{3}{2\pi^2 f^2\,(0)\,\sigma^2}\right), \tag{8}$$

and consequently,

$$LM_{LAD} \to \chi^2_{(1)}$$

assuming that $\widehat{f}\,(0)$ is a consistent estimator of $f\,(0)$.

Proof. *See the technical appendix.*

Remark 2.1. The asymptotic variance of $\widehat{\phi}_{LAD}$ is characterized by the density of the error distribution in the LAD regression, which can be consistently estimated through a series of alternative methods; see Koenker (2005) for a review. We shall discuss this issue more carefully in the generalized QR context. Also, note that $\widehat{\phi}_{LAD}$ is the maximum likelihood estimator of ϕ when the distribution of ε_t is the double exponential distribution with coefficient $\lambda > 0$, *i.e.*, $f\,(\varepsilon_t) = \exp\,(-|\varepsilon_t|/\lambda)\,/(2\lambda)$, with $\lambda^{-1} \equiv 2f\,(0)$. Therefore, LM_{LAD} is an efficient score test based on the gradient of the likelihood function when innovations are drawn from this distribution. As in the case of the LM_{LS} test, the asymptotic convergence stated in Theorem 2.1 does not require a particular distribution to hold, but on the other hand it requires innovations having both zero mean and zero median, a technical restriction often required in robust inference; see, for instance, Delgado and Velasco (2005). We shall conveniently relax this unnecessarily restrictive assumption for practical purposes in the QR setting.

Remark 2.2. It is possible to use more general estimation procedures aiming to ensure full robustness (in the sense of obtaining high breakdown resilience) when the sample is contaminated with a substantial fraction of influential observations. These procedures can be embedded into a general class of estimators known as M estimators, of which the LAD, LS and maximum-likelihood estimators are particular cases. More specifically, let $\eta\,(\cdot)$ be a real-valued function, possibly satisfying certain regularity and smoothness conditions (*e.g.*, being twice-differentiable with a piecewise continuous second derivative) intended to suitably downweight large observations. Then, the M estimator of ϕ in (4) can generally be defined as the solution of $\min_{\phi \in \mathbb{R}} \sum_{t=2}^T \eta\,\left(\left[\varepsilon_{t,d} - \phi x^*_{t-1,d}\right]/s\right)$, where s denotes an estimate of the scale parameter σ. Denoting the resultant estimator as $\widehat{\phi}_M$, it can be shown (*e.g.*, Amemiya 1985, Section 2.3.3)

that $\sqrt{T}\left(\widehat{\phi}_M - \phi\right)$ is normally distributed with zero mean and asymptotic variance given by the limit in probability of

$$\sigma^2 \left[\frac{1}{T}\sum_{t=2}^{T}\left(x_{t-1,d}^*\right)^2 \psi_{tA}\right]^{-1} \left[\frac{1}{T}\sum_{t=2}^{T}\left(x_{t-1,d}^*\right)^2 \psi_{tB}\right] \left[\frac{1}{T}\sum_{t=2}^{T}\left(x_{t-1,d}^*\right)^2 \psi_{tA}\right]^{-1} \tag{9}$$

where $\psi_{tA} = \mathrm{E}\left\{\eta''\left(\left[\varepsilon_{t,d} - \phi x_{t-1,d}^*\right]/s\right)\right\}$ and $\psi_{tB} = \mathrm{E}\left\{\eta'\left(\left[\varepsilon_{t,d} - \phi x_{t-1,d}^*/s\right]^2\right)\right\}$, respectively. The study of this type of estimators and their relative performance warrants careful analysis and, although introduced in this paper, it is left for future research.

There are two characteristics of the LM_{LAD} test that should be highlighted in relation to other procedures in the robust unit root literature. In particular, i) LM_{LAD} is asymptotically distributed as its LS-based counterpart, and ii) it has a nuisance parameter free limiting distribution. These appealing properties allow inference to be carried out without the need for data-specific critical values and, furthermore, extend directly to the more general QR setting, as we shall show later. In sharp contrast, available tests for the unit-root hypothesis based on robust estimators do not have a pivotal distribution. For instance, the test in Hercé (1996), based on LAD estimation of the autoregressive root, has an asymptotic distribution that can be represented as $\delta\mathcal{DF} + \sqrt{1 - \delta^2}\mathcal{Z}$, where $\mathcal{DF}$ denotes the Dickey-Fuller distribution, $\mathcal{Z}$ is an independent standard normal variate, and the nuisance parameter δ measures the correlation between ε_t and $sign\left(\varepsilon_t\right)$, with $sign\left(\varepsilon_t\right) = 2\mathbb{I}\left(\varepsilon_t > 0\right) - 1$. Similar mixture processes characterize the limit distributions of M- and QR-based unit-root tests, see, for example, Lucas (1995), and Koenker and Xiao (2004), respectively. These tests require to compute specific critical values according to the sample estimates of the nuisance parameters involved.

3   Quantile regression

In this section, we discuss the asymptotic theory for an LM type test for fractional integration within the QR theoretical framework. The purpose of this analysis is twofold. First, we pursue to naturally extend the median-based analysis discussed in the previous section to a generic quantile $0 < \tau < 1$, not necessarily the median, and to sets of quantiles. In addition, we provide a theoretical discussion that follows under more general assumptions than those discussed previously. The quantile regression setting does not require strong distribution assumptions and makes no prior assumption about the conditional median of innovations. Furthermore, we can relax the restrictive i.i.d. context discussed previously.

More specifically, note that, given model (1), the τ-th conditional quantile function of the filtered series $\varepsilon_{t,d}$ can be characterized as

$$Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1}) = F^{-1}(\tau) + \phi x_{t-1,d}^* = \mathbf{z}_{t-1,d}'^{*}\boldsymbol{\beta}(\tau) \tag{10}$$

where $\mathcal{F}_{t-1}$ denotes the σ-field generated by $\{\varepsilon_s, s < t\}$, $F(\cdot)$ is the cumulative distrubution function of innovations, $\mathbf{z}^*_{t-1,d} = (1, x^*_{t-1,d})'$, and $\boldsymbol{\beta}(\tau) = (\alpha(\tau), \phi)'$ with $\alpha(\tau) \equiv F^{-1}(\tau)$. Under the null hypothesis of interest, the true value of ϕ equals zero globally (*i.e.*, across the different quantiles), which provides us with a testable hypothesis to identify the order of integration in the data in this generalized setting.

Parameter estimation in this equation involves the following minimization problem,

$$\min_{\mathbf{b}(\tau) \in \mathbb{R}^2} \sum_{t=2}^T \rho_\tau(\varepsilon_{t,d} - \mathbf{z}'^*_{t-1,d}\mathbf{b}(\tau)) \equiv \min_{\mathbf{b}(\tau)} \mathcal{L}_T(\tau) \tag{11}$$

where $\rho_\tau(s) = s(\tau - \mathbb{I}(s < 0))$ is the so-called 'check' function; see Koenker and Bassett (1978). We shall denote as $\widehat{\boldsymbol{\beta}}(\tau) = \left(\widehat{\alpha}(\tau), \widehat{\phi}_{QR}(\tau)\right)'$ the estimator of $\boldsymbol{\beta}(\tau)$ obtained from minimizing $\mathcal{L}_T(\tau)$ in (11).

3.1 Quantile regression test for long memory

Given the quantile regression estimates, define $s(\tau) = [f(F^{-1}(\tau))]^{-1}$ as the reciprocal of the density function of residual evaluated at the quantile of interest, often referred to as sparsity function, and consider the following test statistic,

$$LM_{QR\tau} = \left[\frac{\widehat{\phi}_{QR}(\tau)}{\widehat{s}(\tau)\sqrt{\tau(1-\tau)}}\right]^2 \sum_{t=2}^T \left(x^*_{t-1,d}\right)^2 \tag{12}$$

with $\widehat{s}(\tau)$ representing a consistent estimate of $s(\tau)$ hereafter. The following result extends Theorem 2.1 to any quantile $\tau \in (0,1)$ under slightly more general conditions, since median-zero errors are no longer required. We will discuss further extensions later on.

Theorem 3.1 *Consider $\{y_t\}$ as given in (1), with $\varepsilon_t \sim iid\,(0, \sigma^2)$ and assume that the cumulative distribution function of ε_t, say $F(z)$, has a differentiable continuous Lebesgue density, $0 < f(z) < \infty$, and bounded derivatives on $\{z : 0 < F(z) < 1\}$. Let $\widehat{\boldsymbol{\beta}}(\tau) = \left(\widehat{\alpha}(\tau), \widehat{\phi}_{QR}(\tau)\right)'$ be the solution of $\min_{\mathbf{b}(\tau)} \mathcal{L}_T(\tau)$ for a fixed $\tau \in (0,1)$, and denote $\boldsymbol{\beta}(\tau)$ as the vector of true parameter values. Then, under the null hypothesis $H_0 : \theta = 0$ and as $T \to \infty$,*

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right) \to \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))}\mathbf{V}^{-1}\right) \tag{13}$$

with $\mathbf{V} = diag\,(1, \sigma^2\pi^2/6)$. Consequently,

$$LM_{QR\tau} \to \chi^2_{(1)}$$

if $\widehat{s}(\tau)$ is a consistent estimate of $s(\tau)$.

Proof. *See the technical appendix.*

Remark 3.1. Note that $\widehat{\boldsymbol{\beta}}(\tau)$ is the maximum likelihood estimate of the unknown parameters vector when the distribution of ε_t is the generalized Laplace distribution. Hence, the LM procedure achieves full efficiency under this restriction, but the asymptotic result stated in the theorem holds regardless of the particular distribution of the data as long as the mild standard regularity conditions apply.

At this point, it is also worth comparing this test to the well-known Dickey-Fuller test, since this has received considerable attention in the QR literature. Given (1), the unit-root hypothesis can be tested as $H_0 : \phi = 0$ in the auxiliary regression $\Delta y_t = \alpha + \phi y_{t-1} + u_t$, where $y_{t-1} = \sum_{j=1}^{t-1} \varepsilon_{t-j}$ under the null of integration. In our testing procedure, the unit-root hypothesis implies $H_0 : \phi = 0$ in the auxiliary regression $\varepsilon_{t,d} = \alpha + \phi x_{t-1,d}^* + u_t$, where under the null hypothesis $\varepsilon_{t,d} \equiv \Delta y_t$, and $x_{t-1,d}^* \equiv \sum_{j=1}^{t-1} j^{-1} \varepsilon_{t-j}$. Therefore, the only difference between our regression model and the familiar Dickey-Fuller representation lies in the introduction of the harmonic weights j^{-1}. This difference is, however, crucial to ensure power to detect fractional alternatives and, furthermore, has major implications on the rate of convergence and on the shape of the resulting limit distribution. From Koenker and Xiao (2004), the τ-th QR estimate of ϕ in $\Delta y_t = \alpha + \phi y_{t-1} + u_t$ is T-consistent under the conditions of Theorem 3.1, and converges, once adequately normalized, to $\delta_{\varepsilon\tau} \mathcal{DF} + \sqrt{1 - \delta_{\varepsilon\tau}^2} \mathcal{Z}$, where $\delta_{\varepsilon\tau}$ is a nuisance parameter depending on τ and other population characteristics of ε_t; see also Hercé (1996) and Galvao (2009). In contrast, $\widehat{\phi}_{QR\tau}$ is $\sqrt{T}$-consistent in our context and $LM_{QR\tau} \to \chi^2_{(1)}$ holds since

$$\sqrt{T}\,\widehat{\phi}_{QR}(\tau) \to \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))} \frac{6}{\sigma^2 \pi^2}\right), \tag{14}$$

which follows directly from (13). In particular, given that $T^{-1} \sum_{t=2}^{T} x_{t-1,d}^{*2} \overset{p}{\to} \sigma^2 \pi^2/6$, and the sparsity function $s(\tau)$ can be consistently estimated (see Remark 3.3 below for details), we can construct a pivotal test converging asymptotically to the standard normal distribution uniformly on τ for one-sided testing, or a squared test statistic, such as $LM_{QR\tau}$, to test the null hypothesis against a two-sided alternative.

Using this testing framework we can also derive procedures to test the hypothesis of fractional integration on a range of quantiles in any closed subinterval of $(0, 1)$. For instance, we may wish to test $H_0 : \phi = 0$ given a fixed value of d in a neighborhood of $\tau = 1/2$, since this may provide a more robust interpretation of results. It should be noted that, given model (1), the QR setting provides a useful check to formally determine the adequacy of a long-memory model with constant parameter. Under the null hypothesis, the filtered series $\varepsilon_{t,d}$ behaves as an i.i.d. process and this should be detected at any quantile of the conditional distribution. This feature suggests a diagnosis test computed over a wide range of quantiles. We therefore discuss two alternative tests for this class of hypotheses as a generalization of Theorem 3.1. The asymptotic distributions hold directly from the property of tightness of the QR process and the continuous mapping theorem.

Theorem 3.2. *Let* $\Theta = [\underline{\tau}, \overline{\tau}]$ *be a closed subset of* $(0,1)$ *of length* $\Delta = \overline{\tau} - \underline{\tau}$, *and consider an equidistant partitioning* $\tau_i = \underline{\tau} + i\Delta/T$, $i = 0, 1, ..., T$. *Define the random function* $S_T(\tau) = \widehat{\phi}_{QR}(\tau) / \sqrt{\widehat{\kappa}(\tau)/T}$, *mapping* $\tau \in (0,1)$ *into* $\mathbb{R}$, *with* $\widehat{\kappa}(\tau) = \widehat{s}^2(\tau) \left(\sum_{t=2}^{T} x_{t-1,d}^{*2}/T \right)^{-1}$. *If* $\sup_{\tau \in \Theta} |\widehat{s}(\tau) - s(\tau)| = o_p(1)$, *then under the assumptions of Theorem 3.1 and the null hypothesis it follows as* $T \to \infty$ *that*

$$\mathcal{KS} = \max_{1 \leq i \leq T} |S_T(\tau_i)| \to \sup_{\tau \in \Theta} |\mathcal{B}(\tau)| \qquad (15)$$

and

$$\mathcal{CM} = \sum_{1 \leq i \leq T} S_T^2(\tau_i)(\tau_i - \tau_{i-1}) \to \int_{\tau \in \Theta} \mathcal{B}^2(\tau) d\tau, \qquad (16)$$

where $\mathcal{B}(\tau) = W(\tau) - \tau W(1)$ *is a standard Brownian Bridge.*

Proof. *See the technical appendix.*

Remark 3.2. The limits in (15) and (16) are truncated versions of the well-known Kolmogorov-Smirnov ($\mathcal{KS}$) and Cramér von Mises ($\mathcal{CM}$) distributions which would arise when evaluating the supremum or the integral over the interval $[0,1]$. This serves as motivation to call the test statistics accordingly, noting that it is straightforward to obtain critical values for any of these distributions by simulation. Related test statistics, with limiting distributions which are familiar from the literature on parameter instability, can be designed similarly. For instance, setting $\Theta = [\tau_0, 1 - \tau_0]$ for some $\tau_0 \in (0, 0.5)$, it then holds as in Theorem 3.2 that $\max_{1 \leq i \leq T} S_T^2(\tau_i) / \tau_0(1 - \tau_0) \to \sup_{\tau \in \Theta} \mathcal{S}^*(\tau)$, where $\mathcal{S}^*(\tau) \equiv \mathcal{B}(\tau)^2 / \tau_0(1 - \tau_0)$ is usually referred to as the square of a standardized tied-down Bessel process of order one. Asymptotic critical values of this distribution can be found, for instance, in Andrews (1993, Table 1).

3.2 Short-run dependence

In applied settings, economic and financial variables often exhibit short-run dependence. It is therefore interesting to consider a more general type of data generating process to accommodate this possibility in the theoretical analysis. Hence, assume that the error term ε_t in the data generating process is driven by stationary AR(p) dynamics, *i.e.* $\varepsilon_t = \sum_{j=1}^{p} a_j \varepsilon_{t-j} + v_t$, where v_t is a white noise process. Then, the null hypothesis of fractional integration $H_0 : \theta = 0$ implies $H_0 : \phi = 0$ in this equation, which may be tested through the significance of ϕ in the p-th order augmented auxiliary regression

$$\varepsilon_{t,d} = \phi x_{t-1,d}^* + \sum_{j=1}^{p} a_j \varepsilon_{t-j,d} + u_t, \qquad (17)$$

following Demetrescu *et al.* (2008) and Hassler *et al.* (2009). The conditional quantile function of $\varepsilon_{t,d}$ can consequently be written as

$$Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1}) = F^{-1}(\tau) + \phi x_{t-1,d}^* + \sum_{j=1}^{p} a_j \varepsilon_{t-j,d} = \mathbf{z}_{pt-1,d}^{\prime*} \boldsymbol{\beta}_p(\tau) \tag{18}$$

where $\boldsymbol{\beta}_p(\tau) = (\alpha(\tau), \phi, a_1, ..., a_p)'$, $\mathbf{z}_{pt-1,d}^* = (1, x_{t-1,d}^*, \varepsilon_{t-1,d}, ..., \varepsilon_{t-p,d})'$ and $F(\cdot)$ denoting the distribution function of v_t. This approach exploits the same type of augmentation strategy that characterizes the well-known Augmented Dickey-Fuller test attempting to 'whiten' the residuals; see, for instance, Koenker and Xiao (2004). As discussed previously, parameter estimates are obtained by solving numerically $\min_{\mathbf{b}(\tau) \in \mathbb{R}^{p+2}} \sum_{t=p+1}^{T} \rho_\tau(\varepsilon_{t,d} - \mathbf{z}_{pt-1,d}^{\prime*} \mathbf{b}(\tau)) \equiv \min_{\mathbf{b}(\tau)} \mathcal{L}_{Tp}(\tau)$. The following theorem presents the basic asymptotic behavior of the estimated coefficients and allows us to extend Theorem 3.1 under the more general conditions treated here.

Theorem 3.3 *Consider* $\{y_t\}$ *generated from (1), with* $A(L)\varepsilon_t = v_t$, $A(L) = 1 - \sum_{j=1}^{p} a_j L^j$ *having all roots outside the unit root circle and* $v_t \sim iid(0, \sigma^2)$ *satisfying the distribution assumptions of Theorem 3.1. Let* $\widehat{\boldsymbol{\beta}}_p(\tau) = \arg\min_{\mathbf{b}(\tau)} \mathcal{L}_{Tp}(\tau)$ *and* $\boldsymbol{\beta}_p(\tau)$ *the vector of true parameter values. Then, under the null hypothesis* $H_0 : \theta = 0$,

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}_p(\tau) - \boldsymbol{\beta}_p(\tau)\right) \to \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))}\boldsymbol{\Omega}_p^{-1}\right) \tag{19}$$

with $\boldsymbol{\Omega}_p = \lim_{t \to \infty} E\left(\mathbf{z}_{t-1,d}^* \mathbf{z}_{t-1,d}^{*\prime}\right)$.

Proof. *See the technical appendix.*

Theorem 3.4. *Define* $\boldsymbol{\Psi}_{\tau p} = s^2(\tau)\boldsymbol{\Omega}_p^{-1}$ *and let* $\widehat{\kappa}(\tau)$ *be a consistent estimate of the* $(2,2)$ *element of* $\boldsymbol{\Psi}_{\tau p}$. *Paralleling Theorem 3.2, define the random function* $S_{Tp}(\tau) = \widehat{\phi}_{QR}(\tau)/\sqrt{\widehat{\kappa}(\tau)/T}$, *with* $\widehat{\phi}_{QR}(\tau)$ *denoting the QR estimate of* ϕ *in* $\widehat{\boldsymbol{\beta}}_p(\tau)$, *and the statistics* $\mathcal{KS} = \max_{1 \leq i \leq T} |S_{Tp}(\tau_i)|$ *and* $\mathcal{CM} = \sum_{1 \leq i \leq T} S_{Tp}^2(\tau_i)(\tau_i - \tau_{i-1})$, $\tau_i \in \Theta$. *Then, under the assumptions in Theorem 3.3, and as in Theorem 3.2,* $\mathcal{KS} \to \sup_{\tau \in \Theta} |\mathcal{B}(\tau)|$ *and* $\mathcal{CM} \to \int_{\tau \in \Theta} \mathcal{B}^2(\tau) d\tau$.

Proof. *See the technical appendix.*

Remark 3.3. Under the assumptions of Theorem 3.3, the matrix $\boldsymbol{\Omega}_p$ can be estimated consistently as $\boldsymbol{\Omega}_{Tp}^* = (T-p)^{-1} \sum_{t=p+1}^{T} \mathbf{z}_{pt-1,d}^* \mathbf{z}_{pt-1,d}^{\prime*}$. Also, following Siddiqui (1960) and Bassett and Koenker (1982) the sparsity function $s(\tau) = [f(F^{-1}(\tau))]^{-1}$ can be estimated consistently as the sample difference quotient $\bar{\mathbf{z}}_{pt-1,d}^{\prime*}\left(\widehat{\boldsymbol{\beta}}_p(\tau + h_T) - \widehat{\boldsymbol{\beta}}_p(\tau - h_T)\right)/2h_T$, with $\bar{\mathbf{z}}_{pt-1,d}^*$ denoting the sample mean of $\mathbf{z}_{pt-1,d}^*$, and h_T being a bandwidth parameter that tends to zero as the sample length increases at a suitable rate. Alternatively, we can also use kernel-type estimators proposed in the nonparametric density estimation literature; see Koenker (2005) for a review.

It is straightforward to show, as a corollary of Theorem 3.3, that the LM-type QRLM test statistic

$$LM_{QR\tau,p} = \frac{\widehat{\phi}^2_{QR}(\tau)}{\tau(1-\tau)\widehat{\omega}_{22}\widehat{s}^2(\tau)} \tag{20}$$

converges weakly to a $\chi^2_{(1)}$ distribution as $T \to \infty$ under the set of assumptions considered therein, where $\widehat{\phi}_{QR\tau,p}$ and $\widehat{\omega}_{22}$ denote, respectively, the QR estimate of ϕ from the minimization of $\mathcal{L}_{Tp}(\tau)$ at the τ-th quantile, and the second element in the diagonal of the inverse of $\sum_{t=p+1}^{T} \mathbf{z}^*_{pt-1,d}\mathbf{z}'^*_{pt-1,d}$; see also Demetrescu et $al.$ (2008) and Hassler et $al.$ (2009). Similarly, the t-ratio test, $\sqrt{LM_{QR\tau,p}} \times sign(\widehat{\phi}_{QR}(\tau))$, is distributed asymptotically as a standard normal and can be used alternatively for one-sided testing. The random function $S_{Tp}(\tau)$ that characterizes the $\mathcal{KS}$ and $\mathcal{CM}$ tests in Theorems 3.2 and 3.4 can straightforwardly be computed in the generalized context of Theorem 3.3 as $S_{Tp}(\tau) = \widehat{\phi}_{QR}(\tau)[\widehat{\omega}_{22}\widehat{s}^2(\tau)]^{-1/2}$, thereby enabling QRLM testing over arbitrary sets of quantiles in closed subintervals of $(0,1)$. Note that, whereas $\widehat{\omega}_{22}$ remains fixed across different quantiles given the available sample, $\widehat{\phi}_{QR}(\tau)$ and, particularly $\widehat{s}(\tau)$, may largely vary.

Remark 3.4. In the unit root context, Koenker and Xiao (2004) show that the limit result of the estimator of ϕ in $\Delta y_t = \alpha + \phi y_{t-1} + \sum_{k=1}^{p} a_k \Delta y_{t-k} + u_t$ under the null hypothesis is unaffected by short-run dynamics provided that the regression is suitably augmented. The QR estimates of ϕ and the augmentation-related parameters $\{a_k\}_{k=1}^{p}$ are T- and $\sqrt{T}$-consistent, respectively, which results in asymptotic negligible effects of augmentation on the null distribution of the estimator of ϕ. In our case, the estimates of $\boldsymbol{\beta}_p(\tau)$ are all $\sqrt{T}$-consistent under the null hypothesis, as shown in Theorem 3.3, and since generally $\mathrm{E}(x^*_{t-1,d}\varepsilon_{t-k,d}) \neq 0$ for $k \geq 1$, the $\boldsymbol{\Omega}_p$ matrix is not block-diagonal. Nevertheless, the null distribution of the $LM_{QR\tau,p}$ is the same as that of $LM_{QR\tau}$ in the i.i.d. context, provided that the short-run dynamics is suitably accounted for through augmentation.

Remark 3.5. In the fractional integration literature it is common to set the initial values $\varepsilon_s = 0$ for all $s \leq 0$. All the asymptotic results discussed in this paper hold under such a restriction and, more generally, if ε_s is any other finite constant or behaves as a random variable with zero mean and bounded variance; see the technical appendix for details. Additionally, a non-zero drift coefficient in the data generating process, $(1-L)^{d+\theta}(y_t - \mu) = \varepsilon_t$, can easily be accounted for by previous demeaning. Robinson (1994) suggests a procedure that delivers consistent estimates of μ uniformly on d under the null hypothesis; see also Demetrescu et $al.$ (2008, Prop.4). Denoting $\Delta^\delta = (1-L)^\delta$, it follows that $\Delta^{d+\theta} y_t = \mu \Delta^{d+\theta} + \varepsilon_t$, so μ can be identified under the null hypothesis from the linear regression of the filtered processes $\Delta^d y_t$ on the regressor $x_t = \sum_{j=0}^{t-1} \pi_{j,d}$, where $\pi_{j,d}$ denotes the corresponding weights in the (truncated) expansion of $(1-L)^d$ given the posited value of d. The residuals from this regression correspond to the filtered process $\varepsilon_{t,d}$. We shall use this method in the empirical analysis in Section 5.

Remark 3.6. As discussed in Hassler *et al.* (2009), the testing procedure can also be used to construct confidence intervals that include the true value of d with $100 \times (1 - \alpha) \%$ asymptotic nominal probability. In particular, d can be estimated through a confidence interval obtained from a grid-search on Θ, a closed subset of $\mathbb{R}$, using the general results in Theorem 3.3. Denote $LM^*_{QR\tau,p}(\delta)$ as the value of the test statistic in Theorem 3.3 when evaluated at any $\delta \in \Theta$, and consider $\mathcal{D}_{T\alpha} = \left\{ \delta : \Pr \left[\chi^2_{(1)} \leq LM^*_{QR\tau,p}(\delta) \right] \leq 1 - \alpha \right\}$, *i.e.*, the subset of Θ for which the null hypothesis cannot be rejected at the $(1 - \alpha)$ asymptotic nominal confidence level. It follows that if $\mathcal{D}_{T\alpha}$ is in the interior of Θ, then the probability of d being within $\mathcal{D}_{T\alpha}$ is at least $(1 - \alpha)$. The grid-search process is computationally feasible because the reasonable order of integration in observable data usually assumes values in a small range. We will use this technique in the empirical section.

4   Finite sample analysis

In this section, we evaluate the small sample properties of the QRLM test statistics. The finite-sample performance of LS-based tests given data generated as in (1) has received considerable attention in the literature under normally distributed innovations. Among others, Breitung and Hassler (2002) and Nielsen (2004) have shown the empirical performance of LS based LM tests, both in absolute terms and in relation to alternative frequency domain-based procedures. Moreover, the simulations in Delgado and Velasco (2005) show undersizing effects in the LM test proposed by Tanaka (1999) under errors with infinite variance.

Following the same approach as Koenker and Xiao (2004) and Galvao (2009), we analyze the empirical size and power of the QRLM test $LM_{QR\tau}$ for $\tau = 1/2$ under different scenarios and in relation to other test statistics. In our first experiment, we consider data generated according to $(1 - L)^{d+\theta} y_t = \varepsilon_t, \quad t = 1, ..., T$, where $\{\varepsilon_t\}$ are independent and identically distributed innovations drawn from a Student-t distribution with v degrees of freedom. In our simulations we use $v \in \{2, 3, 1000\}$ and sample lengths $T \in \{100, 250, 1000\}$. The case $v = 1000$ corresponds to the Gaussian distribution, whereas all remaining cases are characterized by heavy-tailed distributions. For $v = 2$, the tails of the Student-t distribution have such a slow decay that ε_t has infinite variance, a possibility not formally covered by the asymptotic theory discussed previously. The Student-t distribution is continuous in the degrees of freedom parameter verifying $\mathrm{E}(|\varepsilon_t|^{v+\epsilon}) < \infty$ for an arbitrarily small $\epsilon > 0$, so we can think of $v = 2$ as a 'limiting' case corresponding to the formal bound of our theory. This is common practice in experimental analyses in the robust literature. As in Breitung and Hassler (2002), we focus on the unit-root case under the null hypothesis, $d = 1$, since this is the leading case studied in the QR literature, and explore the average frequencies of rejections when testing $H_0 : \theta = 0$ against a two-sided alternative at the 5% nominal level for values of θ in the range $[-0.3, 0.3]$. The null distribution of the test does not depend on the particular value of d, and its power properties

are mainly determined by the size and sign of θ. For $\theta = 0$, the average frequency of rejection given, say 5000 replications, represents the empirical size of the test, whereas the cases $|\theta| > 0$ allow us to characterize the empirical power given T and v.

Because inference in QR requires dealing with the unknown density of the innovations, we use standard kernel techniques implementing a Gaussian kernel with a sample-length dependent bandwidth. For comparative purposes, we also use resampling methods. The interest in the latter is justified because the bootstrap approach can circumvent the problem of having to specify a bandwidth parameter for kernel estimation. In particular, we consider a fixed-matrix design that attempts to take advantage of the i.i.d. property of the estimated residuals under the null hypothesis; see Buchinsky (1995) and Hahn (1995). In particular, given the QR residuals $\widehat{u}_{t\tau} = \varepsilon_{t,d} - \mathbf{z}'^{*}_{t-1,d}\widehat{\boldsymbol{\beta}}(\tau)$, we proceed as follows: (i) we generate a bootstrap replication, say $\left\{\widehat{u}^b_{t\tau}\right\}^T_{t=1}$, with errors sampled independently and with replacement and build $\varepsilon^b_{t,d} = \mathbf{z}'^{*}_{t-1,d}\widehat{\boldsymbol{\beta}}(\tau) + \widehat{u}^b_{t\tau}$. Then, (ii) we estimate this model again via QR, obtaining the bootstrap estimate $\widehat{\boldsymbol{\beta}}^b(\tau)$. Finally, (iii) steps (i)-(ii) are repeated a large number of times, say $N = 750$, so that the covariance matrix of $\widehat{\boldsymbol{\beta}}(\tau)$ can be estimated as $N^{-1}\sum_{i=1,N}\boldsymbol{\varrho}_i\boldsymbol{\varrho}'_i$, with $\boldsymbol{\varrho}_i = \widehat{\boldsymbol{\beta}}^b_i(\tau) - \widehat{\boldsymbol{\beta}}(\tau)$. We are aware that this is not the only possible bootstrap approach, and that more sophisticated procedures may be used in more general contexts (*e.g.*, Fitzenberger 1997). Furthermore, certain aspects may provide further refinements (*e.g.*, fixing the length of the bootstrap sample to a smaller value than T), but since our main interest lies in the analysis of the empirical performance of a simple resampling method, an comprehensive investigation of replication methods is beyond the scope of this paper. We shall denote $LM^K_{QR\tau}$ and $LM^B_{QR\tau}$ as the resultant kernel- and bootstrap-based test statistics, respectively.

In addition, to evaluate the relative behavior of the test, we analyze the performance of two alternative procedures that are also based on the LM principle. One is the standard LS based test suggested by Breitung and Hassler (2002), LM_{LS}, which is efficient in the Gaussian context and formally valid for $v > 2$. Additionally, since the main interest is in a context characterized by extreme observations, the other is the sign-based LM test proposed by Delgado and Velasco (2005), which provides a robust alternative for long-memory testing. This tests is based on the same harmonic weighting structure that characterizes LM_{LS} and $LM_{QR\tau}$, but has the outstanding property of being formally valid even if $\mathrm{E}(\varepsilon^2_t) = \infty$. It requires, however, both the median and the mean of ε_t to be zero, which in practice may imply a loss of generality, but which holds true in our experimental analysis. Given $\varepsilon_{t,d} = (1-L)^d y_t$, denote $S_{t,d} = sign(\varepsilon_{t,d})$, then the test statistic proposed by Delgado and Velasco (2005) is,

$$t_{DV} = \sqrt{\frac{6}{\pi^2 T}} \sum_{j=1}^{T-1} \frac{1}{j} \left(\sum_{t=j+1}^{T} S_{t,d} S_{t-j,d} \right) \tag{21}$$

which, as $T \to \infty$, converges to a standard normal distribution under the null hypothesis and hence $t^2_{DV} \to \chi^2_{(1)}$. Delgado and Velasco (2005) provide exact critical values for t_{DV} that we shall use in our analysis (the remaining tests are based on asymptotic critical values).

[Insert Table 1 around here]

Table 1 presents the rejection frequencies of the three tests under the different configurations considered. Several features are worth commenting in detail. As expected, the Gaussian environment provides the necessary conditions for the optimality of the LS based procedure, LM_{LS}, which largely outperforms any of the alternative approaches in terms of finite sample size and power for any sample length analyzed. On the other hand, when innovations are drawn from heavy-tailed distributions, the LM_{LS} test tends to be undersized, particularly, when the degree of leptokurtosis of the underlying distribution seriously departs from that of the Gaussian. This result agrees with the evidence presented by Delgado and Velasco (2005) for the LM test of Tanaka (1999), which is asymptotically equivalent to LM_{LS}. Interestingly, our simulations reveal that it is necessary to introduce a considerable degree of leptokurtosis (as measured by v) to generate sizeable departures with respect to the Gaussian case: The differences between $v = 1000$ and $v = 3$ are not particularly dramatic in terms of size distortions nor power reductions in any of the sample lengths analyzed, so that the LM_{LS} test seems to exhibit a considerably degree of robustness against heavy-tailed distributions.

Nevertheless, in a non-Gaussian environment, LM_{LS} is no longer efficient. Even if we do not observe power reductions, other alternative procedures that exploit different estimation/inference approaches may produce better relative results. Indeed, our simulation study reveals that QRLM tests can yield fairly large gains in relative power with respect to LM_{LS} when errors are drawn from heavy-tailed distributions, while still ensuring approximately correct size even in small samples. For $v = 2$, the QRLM test shows similar undersizing as the LM_{LS}, but displays considerable power improvements (roughly doubling the power of the LS test and the power is even larger for several other configurations of the data generating process). For instance, for $T = 100$, $v = 2$ and $\theta = -0.1$, the power of LM_{LS} is approximately 17.1%, whereas $LM_{QR\tau}^K$ and $LM_{QR\tau}^B$ present rejection frequencies of 46.2% and 53.5%, respectively. The relative gains in power are asymmetric and tend to be much larger in the stationary region $(\theta < 0)$ than in the explosive direction $(\theta > 0)$. This pattern tends to disappear as $v \to 2$, for which power shows similar patterns around the origin. The performances of $LM_{QR\tau}^K$ and $LM_{QR\tau}^B$ seem to present a similar size-power balance, the former presenting mild undersizing effects as v approaches two, and the latter showing small oversizing behavior. As noted by Koenker (2005), the discrepancies obtained between reasonable alternatives to estimate the covariance matrix tend to be small in QR.

The robust sign-based LM test shows a remarkable steady empirical size, exhibiting approximately correct size in all cases, although it shows moderate undersizing when T is small. This test tends to show comparable power to the QRLM tests when $\theta > 0$, although we observe that the QRLM methodology always provides moderate gains over this test which, based on the sample lengths analyzed, are more marked when $v \to 2$. Nevertheless, when comparing results

in the direction of the stationary region, it is immediately clear that QR is better suited. Not only does our test seem to be particularly powerful when $\theta < 0$ but the ability of the sign-based test to reject the null in a two-sided testing context dramatically collapses in this region. As a result, QR show a more appealing overall performance.

In addition, we also analyze the performance of these tests when the data exhibits autoregressive short-run dependence characterized by stationary AR(1) dynamics with coefficient $a \in \{0.5, 0.75\}$, namely, $(1 - aL)(1 - L)^{d+\theta} y_t = \varepsilon_t$, and repeat the analysis under the same considerations as those previously discussed. For simplicity of exposition, and since the main qualitative results remain unaltered, we discuss the results for $LM_{QR\tau,p}$ with covariance matrix computed with a the kernel-based procedure, and the least-squares based test LM_{LS} when the auxiliary regressions in the respective analysis are augmented with one lag of the dependent variable.

[Insert Table 2 around here]

Table 2 shows the empirical rejection frequencies under short-run dynamics for the different parameter configurations that characterize this experiment. For fairly small samples, such as $T = 100$, the QRLM test shows significant oversizing effects in relation to the i.i.d. context, particularly, under Gaussian conditions, which nevertheless are quickly corrected as the sample length increases. As in the i.i.d. experiment, both QR- and LS-based tests tend to show undersizing effects when $v = 2$. In terms of power, it is evident that both tests suffer important power reductions in relation to the i.i.d. context, stemming from the augmentation required to ensure correct size. For $a = 0.5$, we observe that the power of both tests is characterized by a strong asymmetric pattern such that the alternatives $\theta < 0$ are easier detected than their counterparts $\theta > 0$, a pattern which was already noted by Demetrescu *et al.* (2008). However, this result seems to be data-dependent, and different conclusions arise when $a = 0.75$. The QR-based test shows considerably improved power over its LS alternative as the degree of leptokurtosis increases, particularly for negative values of θ. For positive values, the gains are smaller than in the i.i.d. case, and a considerable degree of leptokurtosis is necessary to beat the LS based procedure.

Consequently, the overall experimental evidence suggests that the tests proposed in this paper are well-suited for empirical studies, even in small samples, and may provide improved performance over LS based alternatives when the data is driven by heavy-tailed distributions.

5   Long-run dependence in realized stock volatility

The growing availability of intraday data on the price of individual stocks and financial indices allows us to study different aspects of the stochastic properties of returns. In this section, we analyze the long-run behavior of daily realized volatility of IBM, one of the most liquid and

frequently-traded securities in the U.S. stock exchange. Realized volatility is a theoretically consistent estimate of integrated volatility which is based on simple sums of intraday returns; see, among others, Andersen *et al.* (2001, 2003). Our interest in this variable is motivated by the findings in the extant literature suggesting that realized volatility characteristically exhibits long memory dynamics. There is an ongoing debate about the sources of long-range dependence in this literature, but at the theoretical level there is little consensus on the mechanism generating this phenomenon. Most econometric attention has been focused on the role of aggregation (*e.g.,* Lieberman and Phillips, 2008), but long-range dependence may also arise spuriously from neglected breaks or nonlinear patterns, see, for instance, Diebold and Inoue (2001) and references therein. The reader is referred to Corsi *et al.* (2008) for a recent survey of this literature.

5.1   Data and preliminary evidence of long memory

We observe continuously compounded IBM returns, sampled regularly over 5-minute intervals from 9.30 a.m. to 4.00 p.m. over the period from 04/01/1993 to 31/05/2007, totalling 156 intraday observations over $3,630$ trading days.[2] Building on the theory of semimartingales and the quadratic variation process, consistent measures of the volatility process based on simple sums of intraday returns have been suggested in the financial econometrics literature. We compute different measures of realized variation, such as daily realized volatility, here defined as the square root of the sum of squared 5-minute log-returns over the day: $\sigma_{RV}(t) = \left[\sum_{n=1}^{156} r_{(n),t}^2\right]^{1/2}$. The resultant measure is widely considered as an accurate estimate of daily integrated volatility of stocks. Additionally, we compute the unnormalized realized absolute variation (or first-order power variation) of returns, defined as the sum of absolute-valued returns over the day: $\sigma_{RPV}(t) = \sum_{m=1}^{156} |r_{(n),t}|$. Realized absolute variations is an accurate estimate of the integral of the volatility process and, under certain conditions, is robust to jumps in returns; see Barndorff-Nielsen and Shephard (2004) for details. Finally, as customary in this literature, we also consider logarithmic transformations of these variables, *i.e.*, log-realized volatility, denoted $\log \sigma_{RV}(t)$, and log-realized power variation, $\log \sigma_{RPV}(t)$; see Andersen *et al.* (2003). Figure 1 shows the dynamics followed by these measures in the sample period, while Table 3 presents descriptive statistics.

<div align="center">

[Insert Figure 1 around here]

</div>

Some comments on the main distributional features of these series are in order. Daily

---

[2]Original data comprises the record of trades and quotations available in the NYSE Transaction and Quote (TAQ) database and allows computation of returns in considerably smaller intervals. In practice, however, there exists a trade-off between the theoretical argument that support ultra-high frequency sampling and the noise that arises from bid-ask bounce and other microstructure effects embedded in such a frequency. The consensus in the literature is to sample returns over 5-minute intervals to balance these effects.

realized volatility $\sigma_{RV}(t)$ typically exhibits a considerable degree of leptokurtosis and right skewness due to the massive influence of the jump component in the data generating process of speculative returns. For IBM, the sample kurtosis over the period under analysis is 102.81, from which the assumption of normality is largely rejected. Being less sensitive to outliers, realized power variation $\sigma_{RPV}(t)$ shows a more moderate degree of kurtosis (30.03), which is still large enough to strongly reject the hypothesis of normality. As reported in previous papers (*e.g.*, Andersen *et al.* 2003), the unconditional distribution of the log transformations of $\sigma_{RV}(t)$ and $\sigma_{RPV}(t)$ appears to be approximately normal, although we note that normality is formally rejected by standard testing procedures in our analysis; see Table 3 for details.

<div align="center">

[Insert Table 3 around here]

</div>

The most important stylized feature of realized volatility measures is an autocorrelation pattern characterized by slowly decaying correlations towards zero, a distinctive feature of long-memory processes. This phenomenon is clearly visible in Figure 2, which shows the sample autocorrelation function of the series up to the 400th lag-order. Even though the first-order correlation of $\sigma_{RV}(t)$ is not particularly sizeable (0.276), the remaining correlations remain highly significant: Distant observations, which span almost two years of trading days, remain positively correlated, thereby suggesting a strong degree of temporal dependence in the series; see Table 3 for further details. A similar pattern appears in the remaining measures of daily variation, although the first-order correlation tends to be larger for these series. This is not surprising, since outliers tend to downward bias sample autocorrelation estimates.

<div align="center">

[Please Insert Figure 2 around here]

</div>

The characteristic correlation pattern displayed by these series cannot be captured by stationary ARMA-type models (for which correlations decay geometrically) nor by unit root models (for which low-order correlations should be close to one). Hence, the literature on realized volatility modelling has argued that this pattern of temporal dependence is caused by a fractionally integrated model with long-memory coefficient $0 < d < 1$. Table 3 also reports several sample-based point estimates of the long-memory parameter and their 95% asymptotic confidence intervals generated by different semiparametric procedures in the frequency domain that are usually applied in the empirical analysis of realized volatility; see, for instance, Andersen *et al.* (2001). These estimators exploit the information provided by the periodogram ordinates in the vicinity of the origin and have the outstanding advantage of not requiring a functional form specification of the short-run component. In particular, $\widehat{d}_{GPH}$ denotes the Geweke and Porter-Hudack (GPH) estimator, defined as the slope coefficient in the least-squares regression $\log(I(\lambda_j)) = \alpha - 2\log(\lambda_j) + \epsilon_j$, $j = 1,...,m$, where $\lambda_j = 2\pi j/T$ are Fourier frequencies, $I(\lambda_j) = \frac{1}{2\pi T}\sum_{t=1}^{T}|(\sigma_{RV}(t) - \bar{\sigma}_{RV})e^{-it\lambda_j}|^2$ is the periodogram, and $m = m(T)$ is a bandwidth that goes off to infinity with the sample length T. Similarly, local

Whittle estimates of the long-memory coefficient are used in practice, having the advantage over the GPH procedure of not requiring normality of the data and being more efficient for the same choice of m. These are generally defined by optimizing the (Whittle) likelihood $W(g, d) = -\frac{1}{m} \sum_{j=1}^{m} \left\{ \log \left(g \lambda_j^{-2d} \right) + \frac{\lambda_j^{2d}}{g} I(\lambda_j) \right\}$, where g is a proportionality constant that depends on the short-run dynamics of the process. Table 3 reports the estimates obtained from the exact local Whittle estimator recently suggested by Shimotsu and Phillips (2005), denoted $\widehat{d}_{ELW}$ in Table 3. In both analyses we set $m = [T^{0.6}]$, and construct confidence intervals using the asymptotic distributions $\sqrt{m} \left(\widehat{d}_{GPH} - d \right) \to \mathcal{N}(0, \pi^2/24)$ and $\sqrt{m} \left(\widehat{d}_{ELW} - d \right) \to \mathcal{N}(0, 1/4)$.

The estimates from these methods show that realized measures in levels tend to exhibit smaller values of d than their logarithmic counterparts. This is consistent with the bias originated by outliers and, more generally, large levels of kurtosis, as reported by Haldrup and Nielsen (2004). The confidence intervals for d on the log-realized measures and realized power variation include values larger than the $1/2$ cut-off limit for stationarity, showing strongly persistent dynamics. Only for the realized volatility series, $\sigma_{RV}(t)$, does the local Whittle estimator show statistical evidence suggesting strongly persistent, yet stationary, dynamics.

5.2 Quantile regression analysis

The QRLM tests discussed in this paper can shed further light on the empirical properties of realized volatility and are useful to address two different questions. The first relates to the fact that the excess kurtosis of the unconditional distribution may bias the semiparametric long-memory estimates, as previously discussed. Since median-based procedures offer robustness against influential observations, QR testing at the 50th percentile is a natural alternative to robustly address the existence of long-memory patterns. Furthermore, the QR approach allows us to analyze the general suitability of a long-run filter based on a constant value of d. If the true process is truly generated by a long-memory model with constant long-range coefficient, we should be able to identify the characteristic value of the long-memory coefficient along the different quantiles of the conditional distribution. Consequently, our testing approach allows us to offer a more robust discussion on the properties of realized volatility.

In particular, the empirical analysis is conducted in the following terms. First, at any of the percentiles $\tau \in \mathcal{Q}$, $\mathcal{Q} = \{0.1, 0.11, ..., 0.9\}$, we run the auxiliary quantile regression $\varepsilon_{t,d} = \phi x^*_{t-1,d} + \sum_{j=1}^{p} a_j \varepsilon_{t-j,d} + u_t$, and compute the t-ratio, say $t_{QR\tau,p}$, for the significance of the estimated value of ϕ, to test the sequence of two-sided hypotheses $H_0 : d = d_0$, $d_0 \in \mathcal{D}$, with $\mathcal{D} = \{0, 0.01, ..., 1\}$.[3] This analysis attempts to provide a detailed examination of the existence of long-memory patterns across the quantiles of the conditional distribution. For ease of exposition, we shall report the test statistics for the subset of hypotheses $H_0 : d = d_0$, with

---

[3]We compute t-statistics rather than squared t-statistics because the sign is informative about the overdifferencing or underdifferencing implied by the null. A positive (negative) value is indicative of potentially significant underdifferencing (overdifferencing).

$d_0 \in \{0, 0.1, ..., 1\}$, at any of the deciles $\{0.1, ..., 0.9\}$ of the conditional distribution, although complete results are available upon request. Note that, owing to statistical difficulties of the QR methodology to accurately deal with inference at extreme quantiles, it is customary to avoid top and bottom percentiles. A similar decile-based analysis for the unit root hypothesis is conducted in Koenker and Xiao (2004) and Galvao (2009) on interest and exchange rates time-series, but we note that our study is more general: It not only allows us to test for the unit root hypothesis, but also for fractional integration dynamics.

The auxiliary regression in our analysis is augmented with p lags of the dependent variable which are determined according to Schwert's (1989) rule, *i.e.*, $p = \left[4 \left(T/100 \right)^{1/4} \right]$. As discussed by Demetrescu *et al.* (2008, 2011), data-driven methods of lag-length selection fail to ensure correct empirical size in long-memory testing, whereas deterministic rules, such as Schwert's rule, manage to keep empirical size close to the nominal level. The standard error of $\widehat{\phi}_{QR}(\tau)$ is estimated based on the sandwich-type estimator proposed by Powell (1991), seeking to obtain robustness against potential heteroskedasticity in the data. We use a Gaussian kernel to estimate the density of the data with deterministic bandwidth parameter, h_T, set according to the rule $0.3 \times \min \{\widehat{\sigma}_u, IQR(\widehat{u}_t)/1.34\} \times T^{-1/5}$, where $IQR(\cdot)$ denotes the interquartile range, a robust measure of scatter. To account for a non-zero constant effect, we demeaned the raw data using the approach suggested by Robinson (1994) and Demetrescu *et al.* (2008, Prop.4), as discussed in Remark 3.5. For completeness of analysis, we also compute the LS based LM test of Breitung and Hassler (2002) with a robust covariance matrix as in Demetrescu *et al.* (2008) to address long-memory under a conditional mean analysis.

Second, our testing procedure allows us to construct confidence intervals for d by identifying the non-rejection region at a desired nominal level α, as discussed in Remark 3.6. We compute the non-rejection region of $H_0 : d = d_0$ at the 5% and 1% nominal size given the quantile $\tau \in \mathcal{Q}$, denoted $\mathcal{CI}_{95\%}(d|\tau)$ and $\mathcal{CI}_{99\%}(d|\tau)$, which determine 95% and 99% confidence intervals for d, respectively. Finally, the joint tests proposed in Theorems 3.2 and 3.4 may be used to analyze whether $H_0 : d = d_0$ holds over subsets of quantiles for any of the values of $d_0 \in \mathcal{D}$. We compute the $\mathcal{KS}$ and the $\mathcal{CM}$ type tests in (15) and (16), respectively, to analyze whether $H_0 : d = d_0$, $d_0 \in \mathcal{D}$, applies uniformly over all quantiles comprised in the intervals $\mathcal{T}_1 = [0.4, 0.6]$ and $\mathcal{T}_2 = [0.1, 0.9]$. While $\mathcal{T}_1$ analyzes the suitability of any specified value of d at the center of the distribution, $\mathcal{T}_2 \equiv \mathcal{Q}$ focuses on the whole distribution after excluding top and bottom quantiles, following the same approach as Koenker and Xiao (2004) and Galvao (2009). We obtain asymptotic critical values for these tests by experimental simulation of the limit distributions $\sup_{\tau \in \mathcal{T}_i} |\mathcal{B}(\tau)|$ and $\int_{\tau \in \mathcal{T}_i} \mathcal{B}^2(\tau) d\tau$, $i = \{1, 2\}$.

Since the logarithmic transformation considerably alters the stochastic properties of realized volatility (*e.g.*, Andersen *et al.* 2003), we report and comment the different results from our analysis throughout the following subsections. The first subsection is devoted to the analysis of the log realized volatility series, $\log \sigma_{RV}(t)$ and $\log \sigma_{RPW}(t)$; the second subsection focuses

on the series in levels, $\sigma_{RV}(t)$ and $\sigma_{RPW}(t)$. Finally, the last subsection discusses the main implications of these results.

5.2.1 Long memory in logarithms of realized volatility

The results for log-realized volatility, log $\sigma_{RV}(t)$, and log-realized power variation, log $\sigma_{RPW}(t)$, are reported in Tables 4 and 5, respectively. We first discuss the main evidence from the individual test statistics computed at the measures of central positioning, namely, the QRLM test at the median and the LS based LM test for the conditional mean. Next, we discuss the results obtained from the individual and joint QRLM tests at the remaining quantiles.

Under the QR analysis at $\tau = 1/2$, the null hypothesis that daily log realized volatility is purely driven by stationary short-run dynamics $(d_0 = 0)$ is, as expected, strongly rejected, since no stationary ARMA model can produce the shape of autocorrelations depicted in Figure 2, and so is the unit root hypothesis $(d_0 = 1)$ for similar reasons. The 95% confidence interval for d at the median, namely $\mathcal{CI}_{95\%}(d\,|\tau = 1/2)$, is given by $[0.36, 0.56]$, mostly suggesting the existence of stationary long-range dependence. The LS based test for the conditional mean of the process yields a slightly larger confidence interval, namely, $[0.42, 0.62]$. These estimates are remarkably similar to those obtained by the semiparametric estimates in the frequency domain, reported in Table 3.

Turning our attention to the log $\sigma_{RPW}(t)$ series, the main picture is very similar, but we remark several minor differences. First, the discrepancies between the conditional median and conditional mean-based analysis are now smaller, which is not surprising in view that the excess kurtosis of log $\sigma_{RPW}(t)$ is smaller, as reported in Table 3. The 95% confidence interval of the QRLM test at $\tau = 1/2$ is now $[0.38, 0.52]$, while for its LS counterpart is $[0.36, 0.51]$. Second, the testing procedures on log $\sigma_{RPW}(t)$ seem to deliver more efficient estimates, since the amplitude of the confidence intervals is smaller. The overall evidence for stationary long-range dependence based on the mean analysis is also stronger, but the tests cannot reject nonstationary dynamics arising from values $d > 1/2$ for neither log $\sigma_{RV}(t)$ nor log $\sigma_{RPW}(t)$.

[Insert Tables 4 and 5 around here]

When analyzing the results from the individual QRLM tests across different quantiles, two main features emerge. First, the QRLM always find strong statistical evidence of long-range dependence. The confidence intervals for d always include values strictly greater than zero and smaller than one and reject the hypotheses of a short-memory or a unit-root model driving the long-term of the series. Second, there is an upward trend in the confidence intervals such that both the central value and the amplitude of the confidence interval tend to increase with τ. This phenomenon is clearly visible in Figure 3, which shows the 'central' values $d : \arg\inf_d |t_{QR\tau,p}|$ for which the individual QRLM test statistics are closer to zero (*i.e.*, the values of d for which we obtain maximum sample evidence for the null hypothesis given τ), as well as the

corresponding 95% and 99% confidence intervals for d. The upward trend is hardly noticeable for quantiles below the median and the confidence intervals are very similar to those discussed for the median. For instance, the $\mathcal{CI}_{95\%}(d|\tau)$ sets for $\log\sigma_{RV}(t)$ and $\log\sigma_{RPW}(t)$ at $\tau = 1/10$ are, respectively, $[0.29, 0.43]$ and $[0.33, 0.47]$, which are only slightly smaller than those at $\tau = 1/2$ discussed before, and actually share most of the values. This pattern also holds for quantiles at the central deciles above the median, but a stronger diverging effect is evident for percentiles belonging to the upper quartile. For instance, the $\mathcal{CI}_{95\%}(d|\tau)$ sets for $\log\sigma_{RV}(t)$ and $\log\sigma_{RPW}(t)$ at $\tau = 9/10$ are respectively, $[0.56, 0.83]$ and $[0.56, 0.84]$, so most of the values comprised in these intervals are above the values in $\mathcal{CI}_{95\%}(d\,|\tau = 1/2)$; see Tables 4 and 5 for details. The amplitude of the confidence intervals shows a similar shifting pattern as a function of τ: It tends to remain steady for quantiles in the lower tail and center of the distribution, but largely widens at top deciles. For instance, for the log-realized volatility (power variation) time-series, the size of the $\mathcal{CI}_{95\%}(d|\tau)$ set at $\tau = 9/10$ is almost three times larger than (twice as large as) its counterpart at $\tau = 1/10$.

<center>**[Insert Figure 3 around here]**</center>

Although most of the observations at the lower deciles and center of the distribution seem to be driven by a model with common long-memory coefficient, there is a considerable degree of parameter uncertainty for observations corresponding to the largest levels of volatility. In this context, the QRLM joint tests provide a valuable tool for disentangling formally whether there is sufficient regularity in favour of a constant long-memory parameter model. As expected from the individual analysis, the $\mathcal{KS}$ and $\mathcal{CM}$ tests cannot reject this hypothesis for quantiles in the $\mathcal{T}_1$ central interval, finding that values around $d = 0.4$ seem to fit rather well. More interestingly, the joint tests over the quantiles in the extended range $\mathcal{T}_2 = \mathcal{Q} \equiv [0.1, 0.9]$ cannot reject the null hypothesis of a constant long-memory parameter model for certain values of $d \in \mathcal{D}$. Paralleling the strategy used for individual quantile testing, we can construct confidence intervals for these values by identifying the non-rejection region of the $\mathcal{KS}$ and $\mathcal{CM}$ tests given the sets of quantiles analyzed. The resulting confidence intervals are denoted as $\mathcal{CI}_{100\times(1-\alpha)\%}(d|\mathcal{T})$ in Tables 4 and 5, with $\mathcal{T}$ representing either $\mathcal{T}_1$ or $\mathcal{T}_2$. Thus, for the $\log\sigma_{RV}(t)$ time series, the $\mathcal{CI}_{95\%}(d|\mathcal{T}_2)$ sets given by the $\mathcal{KS}$ and $\mathcal{CM}$ tests are $[0.44, 0.45]$ and $[0.41, 0.42]$, respectively. Similarly, for $\log\sigma_{RPW}(t)$ series, the resulting 95% confidence intervals for d are $[0.48, 0.51]$ and $[0.43, 0.49]$, respectively.

Consequently, the joint analysis across quantiles does not reject the suitability of a fractionally integrated model with constant long-memory parameter driving the long-run of logarithmic measures of daily integrated volatility. The range of admissible values is slightly greater than 0.4, the value around which previous literature tends to identify the long-memory coefficient in daily realized volatility time series. Andersen *et al.* (2003) have referred to this as the "typical value" in their study. Furthermore, the overall evidence agrees with the results based on the

<center>23</center>

semiparametric estimators in the frequency domain. This estimate suggests that the long-run component of realized volatility is driven by a strongly persistent, yet stationary, process.

5.2.2 Long memory in levels of realized variation

Tables 6 and 7 report the main results of our analysis for the $\sigma_{RV}(t)$ and $\sigma_{RPW}(t)$ series, respectively. As in the previous case, the overall qualitative evidence for both series is similar, so we only discuss the case of realized volatility for the sake space. The results of this analysis reveal two main features. Similarly as log-realized measures of daily variation, the QRLM shows that long-run dependence seems to be present across the different quantiles of $\sigma_{RV}(t)$. In sharp contrast, however, the extent of persistence, as measured by d, cannot be accepted to be the same. In particular, the confidence intervals for d show a much stronger upward trend in the values of the long-memory coefficient for which the null hypothesis cannot be rejected. Observations in low-volatile periods, related to lower deciles, seem to be driven by a persistent, but stationary, process. On the other hand, observations in high-volatility regimes, related to top deciles which include the spikes observable in Figure 1, are captured by fractionally integrated models with values of d around unity or greater, which suggest the presence of integrated, and even explosive dynamics. This pattern is similar to the results in Konker and Xiao (2004), showing that large observations are related to explosive patterns, whereas low observations tend to follow stationary dynamics. Not surprisingly, therefore, the $\mathcal{KS}$ and $\mathcal{CM}$ joint tests formally reject the hypothesis that a constant value of the long-memory parameter underlies simultaneously the long-range dependence of the series across quantiles in $\mathcal{T}_2$ and even at the central quantiles in the $\mathcal{T}_1$ interval. Similar conclusions emerge from the analysis on realized power variation.

[Insert Table 6 and 7 around here]

5.3 Discussion

The direct conclusion from the quantile regression analysis of $\sigma_{RV}(t)$ and $\sigma_{RPV}(t)$ is that the overall suitability of a fractionally integrated model with fixed long-memory parameter is largely rejected. Stated more precisely, filtering these measures of realized volatility with a fractionally integrated model with constant long-memory parameter does not suffice to render innovations that follow a stationary ARMA-type model uniformly over the quantiles of the distribution. This is not surprising, because neither a fractionally-integrated nor a stationary short-run model can simultaneously deal with sudden bursts of volatility (which are driven by the jump component of returns) and with periods of low or normal volatility, which characterize most of the observations in the sample (see Figure 1).

Therefore, the heterogenous evidence of long-memory in levels, suggesting stationarity at low deciles and integration or even explosive patterns at top deciles, is likely caused by the

sheer differences between low-volatility and high-volatility regimes in these series, *i.e.*, due to neglected nonlinear patterns. Values of d around unity at the top deciles are not necessarily originated by a strongly persistent process, but rather reflect the effort of the parametric structure of the long-run component to accommodate abrupt changes in volatility by adopting unit root-like dynamics. Diebold and Inoue (2001) showed analytically that stochastic-regime switching can easily be mistaken for long memory as long as only a small amount of regime switching occurs in an observed sample path, as it essentially generates similar effects as structural breaks. Maheu and McCurdy (2002) find strong evidence of regime-switching dynamics in the daily realized variance $\sigma_{RV}^2(t)$ of foreign exchange rates; see also Baillie and Kapetianos (2007). In other words, the rejection of fractional-integration models for $\sigma_{RV}(t)$ and $\sigma_{RPV}(t)$ through our diagnosis-type tests is not necessarily the consequence of a long-run misspecification implied by these models, but it must be interpreted as the overall failure of the class of ARFIMA model analyzed to convincingly accommodate the underlying (nonlinear) patterns given the sample.

In sharp contrast, the overall evidence for the log transformations of realized volatility and power variation $\log \sigma_{RV}(t)$ and $\log \sigma_{RPW}(t)$ cannot reject that a fractionally integrated model with a constant long-memory coefficient, in the region $[0.4, 0.5]$, drives the long-term component and renders short-run innovations stationary. The logarithmic function is a special case of the nonlinear Box-Cox transformation routinely applied in time-series modelling to reduce heterogeneity and smooth the sample path of the observed series. It attenuates the statistical problems related to different regimes, as it brings observations together and reduces variability, thereby accounting, at least partially, for different regimes. Furthermore, this transformation can preserve the order of fractional integration present in the original series because d is theoretically invariant respect to nonlinear transformations under several conditions, as shown by Gourieroux and Jasiak (2002). The overall evidence based on our analysis suggests, consequently, that long-memory characterizes the long-term component of realized volatility both in levels and in their logarithms, with our testing procedure being able to identify it after accounting for nonlinear patterns through the simple log transform. In the empirical modelling of these series, therefore, this transformation is important.

Finally, we note that this evidence, based on a semiparametric approach, does not necessarily imply that the log transform completely eliminates all the nonlinear features of the data. In fact, individual quantile based QRLM tests at top deciles of $\log \sigma_{RV}(t)$ and $\log \sigma_{RPW}(t)$ still suggest that these series may be generated by a time-series process with characteristics different from those that drive the remaining observations. The central point from our analysis, however, is that there is sufficient regularity in the log series for the diagnosis-type analysis based on the QRLM tests to accept that the sample is driven by a fractionally integrated model with constant long-memory parameter. This evidence completely agrees with the results based on alternative semiparametric estimation procedures applied in this paper, and provides further

support to the previous literature that argue that long-range dependence is a stylized feature of realized volatility.

6 Concluding remarks

In this paper, quantile regression based tests that allow testing for fractionally integrated patterns against integer or fractional integration at different quantiles have been introduced and discussed. An immediate application of this general setting is the LAD, or median estimator, which can outperform standard LS based procedures in settings where innovations are drawn from heavy-tailed distributions. More generally, the theory discussed in this paper allows for more general forms of hypothesis testing, by enabling inference involving the degree of persistence to be carried out at different individual quantiles, or over sets of quantiles. Our procedure, therefore, can provide further insights on the time-series properties of a time-series process.

A distinctive property of the LM-type tests proposed in this paper is that, under the null hypothesis, they will converge to a standard normal distribution or simple transformations of this, such as a Chi-squared distribution for a squared version of the test. Augmented versions of these tests are asymptotically robust against weakly-dependent errors under quite general conditions, and exhibit good statistical performance in samples of moderate size. This makes the class of QRLM test procedures introduced in this paper a valuable tool to address the order of integration of a time-series, particularly, in a non Gaussian context. LS based techniques have traditionally been preferred over alternative approaches because of their good statistical properties, simplicity and computational tractability. However, there are practical contexts, such as the realized volatility case studied in this paper, in which LS no longer provide necessarily optimal estimates, and the properties of the resulting tests can largely be improved by applying alternative procedures, such as quantile regressions. The test proposed in this paper can readily be computed together with its LS counterpart and significance evaluated on the basis of the same critical values, thereby providing, say, standard and robust inference on the extent of long-run dependence of the series.

Using individual and joint QRLM tests, we have analyzed the long-range dependence in different measures of daily integrated volatility, including realized volatility, realized power variation and their logarithmic transforms of these magnitudes. The QRLM tests proposed in this paper, implemented over the whole set of percentiles along the deciles of the conditional distribution, show that the suitability of long-memory models with constant parameter cannot be rejected on log transforms of realized volatility measures. This evidence is more robust than that based simply on the least-squares analysis and leads to conclude that long-memory is a feature of realized volatility time series.

References

[1] Agiakloglou, C. and P. Newbold (1994) Lagrange multiplier tests for fractional difference. *Journal of Time Series Analysis* 14, 253–262.

[2] Amemiya, T. (1985) *Advanced econometrics.* Harvard University Press.

[3] Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001) The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association* 96, 42-55.

[4] Andersen, T., T. Bollerslev, F.X. Diebold and P. Labys (2003) Modeling and Forecasting Realized Volatility, *Econometrica* 71, 529-626.

[5] Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-56

[6] Baillie, R.T. and G. Kapetanios (2007) Testing for neglected nonlinearity in long-memory models. *Journal of Business and Economic Statistics* 25, 447-461.

[7] Barndorff-Nielsen O. E. and N. Shephard (2004). Power and bipower variation withstochastic volatility and jumps. *Journal of Financial Econometrics* 2, 1-37.

[8] Bassett, G., and R. Koenker (1982) An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association* 77, 407-415.

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

[10] Buchinsky, M. (1995) Estimating the asymptotic covariance matrix for quantile regression models: A Monte Carlo study. *Journal of Econometrics* 68, 303-338.

[11] Chan, N.H., L. Peng and Y. Qi (2006) Quantile inference for near integrated autoregressive time series with infinite variance. *Statistica Sinica* 16, 15-28.

[12] Corsi, F., S. Mittnik, C. Pigorsch and U. Pigorsch (2008) The volatility of realized volatility. *Econometric Reviews* 27, 46-78.

[13] Davidson, J. (1994) *Stochastic Limit Theory*, Oxford University Press.

[14] Delgado, M.A. and C. Velasco (2005) Sign tests for long-memory time series. *Journal of Econometrics* 128, 215-251.

[15] Demetrescu, M., U. Hassler and V. Kuzin (2011) Pitfalls of post-model-selection testing: experimental quantification. *Empirical Economics* 40, 359-372.

[16] Demetrescu, M., V. Kuzin and U. Hassler (2008) Long memory testing in the time domain. *Econometric Theory* 24, 176-215.

[17] Diebold F. and A. Inoue (2001) Long memory and regime switching. *Journal of Econometrics* 105, 131–159.

[18] Engle, R.F. and S. Manganelli (2004) CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics* 22, 367-381.

[19] Fitzenberger B. (1997) The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics* 82, 235-287.

[20] Galvao Jr. A. F. (2009) Unit root quantile autoregression testing using covariates. *Journal of Econometrics* 152, 165-178.

[21] Gourieroux C. and J. Jasiak (2002) Nonlinear autocorrelograms: An application to inter-trade durations, *Journal of Time Series Analysis* 23(2), 127-154.

[22] Gutenbrunner, C. and J. Jurečková (1992) Regression rank scores and regression quantiles. *Annals of Statistics* 20, 305-330.

[23] Hahn, J. (1995) Bootstrapping quantile regression estimates. *Econometric Theory* 11, 105-121.

[24] Haldrup, N., M. Ø. Nielsen (2006) Estimation of fractional integration in the presence of data noise. *Computational Statistics & Data Analysis* 51, 3100-3114.

[25] Hasan, M.N. and R. Koenker (1997) Robust rank tests of the unit root hypothesis. *Econometrica* 65, 133-161.

[26] Hassler, U. and J. Breitung (2006) A residual-based LM type test against fractional cointegration. *Econometric Theory* 22, 1091-1111.

[27] Hassler, U., P.M.M. Rodrigues and A. Rubia (2009) Testing for the General Unit Root Hypothesis in the Time Domain. *Econometric Theory* 25, 1793-1828.

[28] Henry, M. and P. Zaffaroni (2003), The long range dependence paradigm for macroeconomics and finance, in P. Doukhan, G. Oppenheim & M. S. Taqqu, eds, *Theory and Applications of Long-Range Dependence*, Birkhäuser, Boston, 417–438.

[29] Hercé, M. (1996) Asymptotic theory of LAD estimation in a unit root process with finite variance errors. *Econometric Theory* 12, 129-153.

[30] Knight, K. (1989) Limit theory for autoregressive parameter estimates in an infinite-variance random walk. *The Canadian Journal of Statistics* 17, 261-278.

[31] Knight, K. (1991) Limiting theory for M estimators in an integrated infinite variance process. *Econometric Theory* 7, 200-212.

[32] Koenker, R. (2005) *Quantile Regression*, Econometric Society Monographs.

[33] Koenker, R. and G. Bassett (1978) Regression quantiles. *Econometrica* 46, 33-49.

[34] Koenker, R. and Z. Xiao (2004) Unit root quantile autoregression inference. *Journal of the American Statistical Association* 99, 775 787.

[35] Koul, H. and A. K. Saleh (1995) Autoregression quantiles and related rank-scores processes. *The Annals of Statistics* 23, 670-689.

[36] Li, G. and W. K. Li (2008) Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Biometrika* 95, 399-414.

[37] Li, G. and W. K. Li (2009) Least absolute deviation estimation for unit root processes with GARCH errors. *Econometric Theory* 25, 1208 - 1227

[38] Lieberman, O. and P.C.B. Phillips (2008) Refined inference on long memory in realized volatility. *Econometric Reviews* 27, 254–267.

[39] Ling, S. and M. McAleer (2004). Regression quantiles for unstable autoregressive models. *Journal of Multivariate Analysis* 89, 304-328.

[40] Lucas, A. (1994) An outlier robust unit root test with an application to the extended Nelson-Plosser data. *Journal of Econometrics* 66, 153-174.

[41] Lucas, A. (1995) Unit root tests based on M estimators. *Econometric Theory* 11, 331-346.

[42] Maheu, J.M. and T.H. McCurdy (2002) Nonlinear features of realized FX volatility. *Review of Economics and Statistics* 88, 668-681.

[43] Nielsen, M.Ø. (2004) Efficient likelihood inference in nonstationary univariate models. *Econometric Theory* 20, 116-146.

[44] Phillips, P.C.B. (1991) A Shortcut to LAD estimator asymptotics. *Econometric Theory* 7, 450-463.

[45] Pollard, D. (1991) Asymptotics for least absolute deviation regression estimators. *Econometric Theory* 7, 186-199.

[46] Portnoy, S. (1984) Tightness of the sequence of empiric cdf processes defined from regression fractiles, in *Robust and nonlinear time series analysis*, eds. J. Franke, W. Hardle, and D. Martin, Springer-Verlag: New York.

[47] Powell, J.L. (1991) Estimation of monotonic regression models under quantile restrictions, in *Nonparametric and Semiparametric Methods in Econometrics* (ed. by W. Barnett, J. Powell and G. Tauchen), Cambridge: Cambridge University Press.

[48] Robinson, P.M. (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* 47, 67-84.

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

[50] Siddiqui, M. (1960) Distribution of quantiles from a bivariate population. *Journal of Research of National Bureau of Standards* 64B, 145-150.

[51] Shimotsu, K. and PC.B. Phillips (2005) Exact local Whittle estimation of fractional integration. *Annals of Statistics* 33, 1890-1933.

[52] Schwert, G.W. (1989) Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics* 7, 147-159.

[53] Tanaka, K. (1999) The nonstationary fractional unit root. *Econometric Theory* 15, 549 - 582.

[54] Taniguchi, M. and Y. Kakizawa (2000) *Asymptotic Theory of Statistical Inference for Time Series*. Springer Series in Statistics, Springer Verlag.

[55] Thomson, S.B. (2004) Robust tests of the unit root hypothesis should not be "modified". *Econometric Theory* 20, 360-381.

[56] Xiao, Z. (2009) Quantile cointegrating regression. *Journal of Econometrics* 150, 248-260.

[57] White, H. (2001). *Asymptotic Theory for Econometricians*, Academic Press.

Technical Appendix

Before proceeding, consider the following notation. For an $(n \times 1)$ vector, $||A||$ denotes the euclidean vector norm such that $||A||^2 = A'A$. For an $(n \times m)$ matrix, $||A||$ denotes the Euclidean matrix norm, $||A||^2 = tr(A'A)$. The constant K is used to refer to some generic, strictly positive and finite constant. The conventional notation $o(1), (o_p(1))$ is used to represent a series of numbers (random numbers) converging to zero (in probability), while $O(1), (O_p(1))$ denotes a series of numbers (random numbers) bounded (in probability) as the sample length is allowed to diverge. As in the main text, the notation $\rightarrow$ and $\overset{p}{\rightarrow}$ denotes weak convergence and convergence in probability of a series of random variables, while $\overset{a.s}{\rightarrow}$ denotes almost surely convergence. Finally, throughout the proofs, we shall consider the 'observable' process, $x^*_{t-1,d} = \sum_{j=1}^{t-1} j^{-1} \varepsilon_{t-j}$, and its 'theoretical' counterpart, $x^{**}_{t-1,d} = \sum_{j=1}^{\infty} j^{-1} \varepsilon_{t-j}$, using the same characteristic notation in superscripts for related variables.

The following lemma collects some preliminary results that are useful to derive the asymptotic properties given in the subsequent theorems.

Lemma A. *Let $\{u_t, \mathcal{F}_t\}_{-\infty}^{\infty}$ be a Martingale Difference Sequence (MDS) with $\sup_t E(u_t^2) < K$ for all t. For $t > 1$, define the measurable processes $m^*_{t-1} = \sum_{j=1}^{t-1} \omega_j u_{t-j}$ and $m^{**}_{t-1} = \sum_{j=1}^{\infty} \omega_j u_{t-j}$, with $\omega_j = O(1/j)$. Then:*
i) $m^*_{t-1} = O_p(1)$, and $m^{**}_{t-1} = O_p(1)$.
ii) $m^{**}_{t-1} - m^*_{t-1} = O_p(1/\sqrt{t})$.
iii) $\max_{1 \leq t \leq T} |m^{**}_t| = o_p(\sqrt{T})$ and $\max_{1 \leq t \leq T} |m^*_t| = o_p(\sqrt{T})$.

Proof. Part $i)$ follows immediately by noting from the MDS property that $E(m^{**}_t) = 0$ for all t and since $E(m^{**2}_{t-1}) = \sum_{j=1}^{\infty} \omega_j^2 E(u_{t-j}^2)$, which is bounded by $K \sum_{j=1}^{\infty} \omega_j^2$ for all t because $\sum_{j=1}^{\infty} \omega_j^2 = O(1)$, so that m^{**}_{t-1} (and, similarly, m^*_{t-1}) is bounded in probability. For part $ii)$, define the 'bias' term $b^*_{t-1} = \sum_{j=t}^{\infty} \omega_j u_{t-j}$ for $t > 1$ and note that $E(b^{*2}_{t-1}) = \sum_{j=t}^{\infty} \omega_j^2 E(u_{t-j}^2)$, so $E(b^{*2}_{t-1}) = O(1/t)$ and therefore, $m^{**}_{t-1} - m^*_{t-1} \equiv b^*_{t-1} = O_p(1/\sqrt{t})$ from Markov's inequality. For $iii)$, take $\epsilon > 0$, and note that

$$
\begin{aligned}
\Pr\left(\max_{1 \leq t \leq T} |m^{**}_t| > \epsilon\sqrt{T}\right) &\leq \sum_{t=1}^{T} \Pr\left(|m^{**}_t| > \epsilon\sqrt{T}\right) = \sum_{t=1}^{T} E\left(|m^{**}_t| \times \mathbb{I}\left(|m^{**}_t| > \epsilon\sqrt{T}\right)\right) \\
&\leq \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} E\left\{[m^{**}_t]^2 \, \mathbb{I}\left(|m^{**}_t| > \epsilon\sqrt{T}\right)\right\} \\
&= \frac{1}{\epsilon^2 T} \sum_{t=1}^{T} E\left\{\varphi^{**}_{tT}\right\}
\end{aligned}
$$

from Markov's inequality, with φ^{**}_{tT} defined implicitely. Since $E([m^{**}_t]^2) < \infty$ for all $t \geq 1$, $\Pr(|\varphi^{**}_{tT}| > 0)$ tends to zero as $T \rightarrow \infty$, from which $\varphi^{**}_{tT} \overset{p}{\rightarrow} 0$. Furthermore, $|\varphi^{**}_{tT}| \leq (m^{**}_t)^2$,

31

so $E\left(\varphi_{tT}^{**}\right)/\epsilon^2 = o_p\left(1\right)$ by virtue of the Dominated Convergence Theorem (Davidson 1994, Lemma 4.12), and we conclude that $\max_{t\geq 1}\left|m_t^{**}\right| = o_p\left(\sqrt{T}\right)$ from Markov's inequality. Finally, noting that $m_t^* \equiv m_t^{**} - b_t^*$, then $\max_{t\geq 1}\left|m_t^*\right| \leq \max_{t\geq 1}\left|m_t^{**}\right| + \max_{t\geq 1}\left|b_t\right|$ from the triangle inequality and given that $E\left\{b_t^2\mathbb{I}\left(\left|b_t\right| > \epsilon\sqrt{T}\right)\right\} \leq E\left(b_t^2\right) = O\left(1/t\right)$ as discussed in ii), then $\max_{t\geq 1}\left|b_t\right| = O_p\left(\sqrt{\log T/T}\right)$ from Markov's inequality and, therefore, $\max_{2\leq t\leq T}\left|m_t^*\right| = o_p\left(\sqrt{T}\right)$, as required. ∎

Proof of Theorem 2.1. Under the null hypothesis, the auxiliary regression $\varepsilon_{t,d} = \phi x_{t-1,d}^* + u_t$ holds true with $\phi = 0$, so that $u_t = \varepsilon_{t,d} \equiv \varepsilon_t$. Let $x_{t-1,d}^{**} = \sum_{j=1}^{\infty} j^{-1}\varepsilon_{t-j}$, and noting that $\mathrm{E}\left(x_{t-1,d}^{**2}\right) = \sigma^2\pi^2/6 \equiv v$, define $v_T = \sqrt{vT}$ and the array $s_{tT}^* = x_{t-1,d}^*/v_T$. Then Theorem 2.1 holds from Pollard (1991, Theorem 2) after noting that the set of sufficient conditions required therein are satisfied in our context: ($\mathcal{C}1$) ε_t is independent of $\mathcal{F}_{t-1}$, ($\mathcal{C}2$) s_{tT}^* is $\mathcal{F}_{t-1}$-measurable, ($\mathcal{C}3$) $\max_{2\leq t\leq T}\left|s_{tT}^*\right| = o_p\left(1\right)$ and ($\mathcal{C}4$) $\sum_{t=2}^{T} s_{tT}^{2*} = 1$. More specifically, ($\mathcal{C}1$) and ($\mathcal{C}2$) hold trivially from independence and owing to the $\mathcal{F}_{t-1}$-measurable nature of $x_{t-1,d}^*$. Condition ($\mathcal{C}3$) follows from Lemma Aiii), since $\varepsilon_t \sim iid\left(0, \sigma^2\right)$ is a restricted case of the more general conditions studied there. Finally, to check ($\mathcal{C}4$) note that $x_{t-1,d}^{**}$ is ergodic and stationary because it is defined on a measurable transformation of a stationary and ergodic process, ε_t, under the set of assumptions considered (White, 2001, Theorem 3.35). Given $s_{tT}^{**} = x_{t-1,d}^{**}/v_T$, then $\mathrm{E}(s_{tT}^{**2}) = 1$ and hence the Ergodic Theorem [ET] (White, 2001, Theorem 3.34) ensures that $\sum_{t=2}^{T} s_{tT}^{**2} \overset{a.s}{\to} 1$. Its 'observable' counterpart, $\sum_{t=2}^{T} s_{tT}^{*2}$, converges to the same limit by applying similar arguments; in particular, let $\mu_{2t-1}^* = E\left(x_{t-1,d}^{*2}/v\right) = O\left(1/t^2\right)$, and define $r_{t-1}^* = x_{t-1,d}^{*2}/v - \mu_{2t-1}^*$. Then, from the triangle inequality, we can show that,

$$\left|\sum_{t=2}^{T} s_{tT}^{2*} - 1\right| \leq \left|T^{-1}\sum_{t=2}^{T} r_{t-1}^*\right| + \left|T^{-1}\sum_{t=2}^{T}\left(\mu_{2,t-1}^* - 1\right)\right| = o_p\left(1\right)$$

so $\sum_{t=2}^{T} s_{tT}^{2*} = 1 + o_p\left(1\right)$ as required. To see this, note that $\mathrm{E}\left(\left|r_{t-1}^*\right|\right) < \infty$ with $\mathrm{E}\left(r_{t-1}^*\right) = 0$ for all $t > 1$, so $\{r_t^*\}$ retains stationarity and ergodicity in mean and hence, $T^{-1}\sum_{t=2}^{T} r_{t-1}^* \overset{a.s}{\to} 0$ from the ET. In addition, the non-stochastic sequence $\mu_{2,t-1}^*$ converges to 1 as $T \to \infty$, so for any arbitrarily small $\epsilon > 0$ there exists an $M_\epsilon > 0$ not depending on T such that $\left|\mu_{2,t-1}^* - 1\right| < \epsilon$ for all $t > M_\epsilon$ for which we further observe

$$\left|T^{-1}\sum_{t=2}^{T}\left(\mu_{2,t-1}^* - 1\right)\right| \leq \frac{M_\epsilon}{T} + T^{-1}\sum_{t=2}^{M_\epsilon}\left|\mu_{2,t-1}^*\right| + T^{-1}\sum_{t=M_\epsilon+1}^{T}\left|\mu_{2,t-1}^* - 1\right|$$
$$= O\left(T^{-1}\right) + o\left(1\right)$$

which completes the result. Therefore, from Pollard (1991, Theorem 2) and under the null hypothesis it follows that

$$2f\left(0\right)v_T\widehat{\phi}_{LAD} \to \mathcal{N}\left(0, 1\right).$$

from which the results stated in the main text follow directly. ∎

Proof of Theorem 3.1. We follow the approach of Knight (1989); see also Koenker and Xiao (2004) and Galvao (2009) for similar analyses in the unit-root testing context. For any scalar $a \neq 0$, define $\psi_\tau(a) = \tau - \mathbb{I}(a < 0)$ and let $u_{t\tau} = \varepsilon_{t,d} - \mathbf{z}_{t-1,d}'^{*}\boldsymbol{\beta}(\tau)$, noting that, under the null hypothesis, $u_{t\tau} = \varepsilon_t - F^{-1}(\tau)$. After reparameterization, the objective function in the QR estimation is equivalent to

$$\min_{\boldsymbol{\delta} \in \mathbb{R}^2} \sum_{t=2}^{T} \left[\rho_\tau(u_{t\tau} - \frac{1}{\sqrt{T}} \mathbf{z}_{t-1,d}'^{*}\boldsymbol{\delta}) - \rho_\tau(u_{t\tau}) \right]. \tag{A.1}$$

Following Knight (1989), we use $\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + \int_0^v [\mathbb{I}(u \leq s) - \mathbb{I}(u < 0)]\, ds$ to rewrite this problem as $\min_{\boldsymbol{\delta} \in \mathbb{R}^2} H_T^*(\boldsymbol{\delta})$, with

$$H_T^*(\boldsymbol{\delta}) = -\frac{1}{\sqrt{T}} \sum_{t=2}^{T} \left(\mathbf{z}_{t-1,d}'^{*}\boldsymbol{\delta} \right) \psi_\tau(u_{t\tau}) + \sum_{t=2}^{T} \int_0^{l_{tT}} [\mathbb{I}(u_{t\tau} \leq s) - \mathbb{I}(u_{t\tau} < 0)]\, ds$$

and $l_{tT} \equiv \boldsymbol{\delta}' \mathbf{z}_{t-1,d}^{*}/\sqrt{T}$. Note that the function $H_T^*(\boldsymbol{\delta})$ is a convex random variable that is minimized at $\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right)$. Then, if there exists a convex function $H(\cdot)$ with a unique minimum, and the finite-dimensional distribution of $H_T(\cdot)$ converges weakly to that of $H(\cdot)$, then convexity ensures that $\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right)$ converges in distribution to the minimizer of $H(\cdot)$.

Let us now introduce some variables to keep simplicity in notation. Let $\boldsymbol{\Delta}_T^* = \sum_{t=2}^{T} \psi_\tau(u_{t\tau}) \mathbf{z}_{t-1,d}^{*}$, define $\boldsymbol{\Delta}_T^{**} = \sum_{t=2}^{T} \psi_\tau(u_{t\tau}) \mathbf{z}_{t-1,d}^{**}$, $\left(\zeta_{t-1}^*, \zeta_{t-1}^{**}\right)' = \psi_\tau(u_{t\tau})\left(x_{t-1,d}^*, x_{t-1,d}^{**}\right)'$ and denote $z_{Tt}^*(\boldsymbol{\delta}) = \int_0^{l_{tT}} [\mathbb{I}(u_{t\tau} \leq s) - \mathbb{I}(u_{t\tau} < 0)]\, ds$. Also, use the short-hand notation $Q_{\varepsilon_t}(\tau|\mathcal{F}_{t-1}) \equiv \mathbf{z}_{t-1,d}'^{*}\boldsymbol{\beta}(\tau) = q_{\tau,t}$. Then, we can write

$$\begin{aligned}
H_T^*(\boldsymbol{\delta}) &= \left\{ -\frac{1}{\sqrt{T}} \boldsymbol{\delta}' \boldsymbol{\Delta}_T^* \right\} + \left\{ \sum_{t=2}^{T} E\left(z_{Tt}(\boldsymbol{\delta})\right) \right\} + \left\{ \sum_{t=2}^{T} \left(z_{Tt}(\boldsymbol{\delta}) - E\left(z_{Tt}(\boldsymbol{\delta})\right)\right) \right\} \\
&= \left\{ H_{1T}^*(\boldsymbol{\delta}) \right\} + \left\{ H_{2T}^*(\boldsymbol{\delta}) \right\} + \left\{ H_{3T}^*(\boldsymbol{\delta}) \right\}, \text{ say,}
\end{aligned}$$

and define $H_T^{**}(\boldsymbol{\delta}) = \sum_{h=1}^{3} H_{hT}^{**}(\boldsymbol{\delta})$ exactly in the same way as $H_T^*(\boldsymbol{\delta})$, the only difference being that $\mathbf{z}_{t-1,d}^{*}$ is replaced by $\mathbf{z}_{t-1,d}^{**}$. The proof then follows by first showing uniform convergence of the 'theoretical' process $H_{1T}^{**}(\boldsymbol{\delta})$ over the bounded sets of $\mathbb{R}^2$, and then showing that $|H_T^{**}(\boldsymbol{\delta}) - H_T^*(\boldsymbol{\delta})| = o_p(1)$, so that $H_T^{**}(\boldsymbol{\delta})$ and $H_T^*(\boldsymbol{\delta})$ share the same asymptotic representation.

First, for $H_{1T}^{**}(\boldsymbol{\delta})$, note that under the null hypothesis $\psi_\tau(u_{t\tau})$ is an i.i.d. process with discrete support $\{\tau, \tau - 1\}$ and probabilities $\{1 - \tau, \tau\}$. Also, $\psi_\tau(u_{t\tau})$ is independent of $\left(x_{t-1,d}^*, x_{t-1,d}^{**}\right)'$, and thus $\mathrm{E}(\boldsymbol{\Delta}_T^{**}) = 0$ and $\mathrm{E}(\boldsymbol{\Delta}_T^{**}\boldsymbol{\Delta}_T^{**\prime}) = \tau(1 - \tau)\mathbf{V}$, where $\mathbf{V} = \mathrm{diag}(1, v)$, recalling $v \equiv \sigma^2\pi^2/6$. Furthermore, since $\mathrm{E}(\psi_\tau(u_{t\tau})|\mathcal{F}_{t-1}) = 0$, then $\{\zeta_{t-1}^{**}, \mathcal{F}_t\}$ is a stationary, ergodic and square-integrable vector MDS. From the Lindeberg-Lévy theorem $T^{-1}\sum_{t=2}^{T} \psi_\tau(u_{t\tau}) \to \mathcal{N}(0, \tau(1 - \tau))$. Also, since $i)$ $T^{-1}\sum_{t=2}^{T} \zeta_t^{2**} \xrightarrow{p} \tau(1 - \tau)v$ from the Ergodic Theorem, and $ii)$ $\max_{2 \leq t \leq T} |\zeta_t^{**}| \leq \max_{2 \leq t \leq T} |x_{t-1,d}^{**}| = o_p\left(\sqrt{T}\right)$ because $\max_{t \geq 2} |\psi_\tau(u_{t\tau})| \leq 1$ and Lemma

33

Aiii), sufficient conditions to apply a Central Limit Theorem for MDS (Davidson, 1994, Theorem 24.3) hold, which together with the Cramér-Wold device allow to conclude that for any fixed $\boldsymbol{\delta}$, $T^{-1/2}\boldsymbol{\Delta}_T^{**} \to \mathbf{N}_\tau$ and $H_{1T}^{**}(\boldsymbol{\delta}) \to -\boldsymbol{\delta}'\mathbf{N}_\tau$, with $\mathbf{N}_\tau$ denoting a 2-dimensional normal variate with zero mean and covariance matrix $\tau(1-\tau)\mathbf{V}$.

Second, for $H_{2T}^{**}(\boldsymbol{\delta})$, following Koenker (2005, p.120) we note that

$$
\begin{aligned}
E\left(z_{Tt}^{**}(\boldsymbol{\delta})\right) &= \int_{-\infty}^{\infty}\left[\int_0^{l_{tT}}\left[\mathbb{I}\left(u_{t\tau} \le s\right) - \mathbb{I}\left(u_{t\tau} < 0\right)\right]ds\right]f(i)\,di \\
&= \int_0^{l_{tT}}\left[\int_{-\infty}^{\infty}\left[\mathbb{I}\left(u_{t\tau} \le s\right) - \mathbb{I}\left(u_{t\tau} < 0\right)\right]f(i)\,di\right]ds \\
&= \int_0^{l_{tT}}\left[F\left(q_{\tau,t}+s\right) - F\left(q_{\tau,t}\right)\right]ds
\end{aligned}
$$

because under the null hypothesis $u_{t\tau} = \varepsilon_t - q_{\tau,t}$. Using the change of variable $s = r/\sqrt{T}$,

$$
\sum_{t=2}^T E\left(z_{Tt}^{**}(\boldsymbol{\delta})\right) = \frac{1}{\sqrt{T}}\sum_{t=2}^T \int_0^{\sqrt{T}l_{tT}}\left[F\left(q_{\tau,t}+\frac{r}{\sqrt{T}}\right) - F\left(q_{\tau,t}\right)\right]dr
$$

and since for all fixed constants a and r, $\lim_{T\to\infty}\left(\frac{F\left(a+r/\sqrt{T}\right)-F(a)}{r/\sqrt{T}}\right) = f(a)$ under the assumptions in Theorem 3.1, we have

$$
\begin{aligned}
\sum_{t=2}^T E\left(z_{Tt}^{**}(\boldsymbol{\delta})\right) &= \frac{1}{T}\sum_{t=2}^T \int_0^{\sqrt{T}l_{tT}}\sqrt{T}\left[F\left(q_{\tau,t}+\frac{r}{\sqrt{T}}\right) - F\left(q_{\tau,t}\right)\right]dr \\
&= \frac{1}{T}\sum_{t=2}^T \int_0^{\sqrt{T}l_{tT}} f\left(q_{\tau,t}\right)r\,dr + o(1) \\
&= \frac{1}{2T}\sum_{t=2}^T f\left(q_{\tau,t}\right)\left[\boldsymbol{\delta}'\mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**}\boldsymbol{\delta}\right] + o(1)
\end{aligned}
$$

so, under the null hypothesis,

$$
\begin{aligned}
H_{2T}^{**}(\boldsymbol{\delta}) &= \frac{f\left(F^{-1}(\tau)\right)}{2}\boldsymbol{\delta}'\left[\frac{1}{T}\sum_{t=2}^T \mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**}\right]\boldsymbol{\delta} + o(1) \\
&\xrightarrow{p} \frac{f\left(F^{-1}(\tau)\right)}{2}\boldsymbol{\delta}'\mathbf{V}\boldsymbol{\delta}
\end{aligned}
$$

because $\mathbf{z}_{t-1,d}^{**}$ is an ergodic and stationary vector with finite variance, hence $T^{-1}\sum_{t=2}^T \mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**} \xrightarrow{a.s} \mathbf{V} = E\left(\mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**}\right)$ from the ET (see, for instance, Taniguchi and Kakizawa, 2000, Theorem 1.3.5).

Lastly, $H_{3T}^{**}(\boldsymbol{\delta})$ is the sum of a demeaned process and for a fixed $\boldsymbol{\delta}$ converges to zero in the mean square sense. To see this, note that the variance of $z_{Tt}^{**}(\boldsymbol{\delta})$ is bounded (Koenker 2005. p.122) by $\max_{2\le t\le T}|\boldsymbol{\delta}'\mathbf{z}_{t-1,d}^{**}|\sum_{t=2}^T E\left(z_{Tt}(\boldsymbol{\delta})\right)/\sqrt{T}$, and since for any finite $\boldsymbol{\delta} = (\delta_1, \delta_2)'$,

$$
\max_{2\le t\le T}|\boldsymbol{\delta}'\mathbf{z}_{t-1,d}^{**}| \le |\delta_1| + |\delta_2|\max_{2\le t\le T}|x_{t-1,d}^{**}| = O_p\left(O(1) + o_p\left(\sqrt{T}\right)\right)
$$

34

as a result $Var\left(z_{Tt}^{**}(\boldsymbol{\delta})\right) = o_p(1)$ because $\sum_{t=2}^{T} E\left(z_{Tt}^{**}(\boldsymbol{\delta})\right) = O_p(1)$. Consequently, $H_T^{**}(\boldsymbol{\delta})$ converges weakly to a random variable with known distribution, say $H(\boldsymbol{\delta})$, uniformly on $\boldsymbol{\delta}$.

We now show that $|H_T^{**}(\boldsymbol{\delta}) - H_T^*(\boldsymbol{\delta})| = o_p(1)$. To this end, note from the triangle inequality that

$$|H_T^{**}(\boldsymbol{\delta}) - H_T^*(\boldsymbol{\delta})| \leq \sum_{h=1,2} |H_{hT}^{**}(\boldsymbol{\delta}) - H_{hT}^*(\boldsymbol{\delta})| + |H_{3T}^{**}(\boldsymbol{\delta})| + |H_{3T}^*(\boldsymbol{\delta})|$$

so the required result holds by showing that the terms on the right-hand side are asymptotically negligible. Define $\varphi_{t-1} = x_{t-1,d}^{**} - x_{t-1,d}^*$, where $\varphi_{t-1}^2 = O_p(1/t)$ from Lemma Aii), and note that, for any fixed $\boldsymbol{\delta}$ with bounded norm,

$$|H_{1T}^{**}(\boldsymbol{\delta}) - H_{1T}^*(\boldsymbol{\delta})| \leq \|\boldsymbol{\delta}\| \ \left\| \left(T^{-1/2}\left(\boldsymbol{\Delta}_T^{**} - \boldsymbol{\Delta}_T^*\right)\right)\right\| = O_p\left(\sqrt{\log T/T}\right)$$

because

$$E\left\|\frac{1}{\sqrt{T}}\left(\boldsymbol{\Delta}_T^{**} - \boldsymbol{\Delta}_T^*\right)\right\|^2 \leq \frac{1}{T}\sum_{t=2}^{T} E\left[\psi_\tau^2(u_{t\tau})\left(x_{t-1,d}^{**} - x_{t-1,d}^*\right)^2\right]$$

$$\leq \frac{1}{T}\sum_{t=2}^{T} E\left[\varphi_{t-1}^2\right] = O\left(\frac{\log T}{T}\right)$$

using independence between $\psi_\tau(u_{t\tau})$ and φ_{t-1} and noting that $\mathrm{E}\left(\psi_\tau^2(u_{t\tau})\right) \leq 1$. Therefore, it follows that $\|T^{-1/2}\left(\boldsymbol{\Delta}_T^{**} - \boldsymbol{\Delta}_T^*\right)\| = o_p(1)$, and so the Asymptotic Equivalence Lemma [AEL], White (2001, Lemma 4.7), ensures $H_{1T}^*(\boldsymbol{\delta}) \to -\boldsymbol{\delta}'\mathbf{N}_\tau$. Similarly, for any fixed $\boldsymbol{\delta}$, we can show that

$$|H_{2T}^{**}(\boldsymbol{\delta}) - H_{2T}^*(\boldsymbol{\delta})| \leq \frac{f\left(F^{-1}(\tau)\right)}{2}\|\boldsymbol{\delta}\| \ \|\Omega_T^{**} - \Omega_T^*\| \ \|\boldsymbol{\delta}\| = O_p\left(T^{-1/2}\right)$$

with $\Omega_T^{**} = T^{-1}\sum_{t=2}^{T}\mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**}$ and $\Omega_T^* = T^{-1}\sum_{t=2}^{T}\mathbf{z}_{t-1,d}^*\mathbf{z}_{t-1,d}'^*$. More specifically, note that

$$x_{t-1,d}^{**2} - x_{t-1,d}^{*2} = \sum_{j=t}^{\infty}\sum_{l=t}^{\infty}\frac{\varepsilon_{t-j}\varepsilon_{t-l}}{jl} + 2\sum_{j=1}^{t-1}\sum_{l=t}^{\infty}\frac{\varepsilon_{t-j}\varepsilon_{t-l}}{jl} = \varphi_{t-1}^2 + 2x_{t-1,d}^*\varphi_{t-1}$$

so if $\mathbb{I}_{(i,j)}$ is a (2×2) indicator matrix taking value one at position (i,j) and zero elsewhere, we have from Minkowski's inequality that,

$$E\|\Omega_T^{**} - \Omega_T^*\| \leq \frac{1}{T}\sum_{t=2}^{T} E\left\|\mathbf{z}_{t-1,d}^{**}\mathbf{z}_{t-1,d}'^{**} - \mathbf{z}_{t-1,d}^*\mathbf{z}_{t-1,d}'^*\right\|$$

$$= \frac{1}{T}\sum_{t=2}^{T} E\left\|\varphi_{t-1}\left[\mathbb{I}_{(1,2)} + \mathbb{I}_{(2,1)}\right] + \left[\varphi_{t-1}^2 + 2x_{t-1,d}^*\varphi_{t-1}\right]\mathbb{I}_{(2,2)}\right\|$$

$$\leq \frac{2}{T}\sum_{t=2}^{T}\left[E\left(|\varphi_{t-1}|\right) + \frac{1}{2}E\left(\varphi_{t-1}^2\right) + E\left(|x_{t-1,d}^*\varphi_{t-1}|\right)\right]$$

$$= O\left(\frac{\log T}{T}\right) + O\left(\frac{1}{\sqrt{T}}\right)$$

35

because, from Liapunov's inequality $E\left(|\varphi_{t-1}|\right) \leq E\left(|\varphi_{t-1}|^2\right) = O\left(1/t\right)$, and the Cauchy-Schwarz inequality ensures $\mathrm{E}\left(|x^*_{t-1,d}\varphi_{t-1}|\right) \leq \sqrt{E\left(x^{*2}_{t-1,d}\right) E\left(\varphi^2_{t-1}\right)} = O\left(1/\sqrt{t}\right)$. This result and the AEL imply that $\Omega^*_T \xrightarrow{p} \mathbf{V}$ and $|H^{**}_{2T}\left(\boldsymbol{\delta}\right) - H^*_{2T}\left(\boldsymbol{\delta}\right)| = o_p\left(1\right)$ uniformly on $\boldsymbol{\delta}$. Finally, we have discussed that $H^{**}_{3T}\left(\boldsymbol{\delta}\right) = o_p\left(1\right)$, and paralleling that reasoning, we can show that $H^*_{3T}\left(\boldsymbol{\delta}\right) = o_p\left(1\right)$ by noting $\max_{t\geq 2}|x^*_{t-1,d}| = o_p\left(\sqrt{T}\right)$ from Lemma Aiii).

Consequently, $|H^{**}_T\left(\boldsymbol{\delta}\right) - H^*_T\left(\boldsymbol{\delta}\right)| = o_p\left(1\right)$ and we can claim from the AEL both $H^*_T\left(\boldsymbol{\delta}\right) = H\left(\boldsymbol{\delta}\right) + o_p\left(1\right)$ and $H^{**}_T\left(\boldsymbol{\delta}\right) = H\left(\boldsymbol{\delta}\right) + o_p\left(1\right)$, where

$$H\left(\boldsymbol{\delta}\right) = -\boldsymbol{\delta}'\mathbf{N}_\tau + \frac{f\left(F^{-1}\left(\tau\right)\right)}{2}\boldsymbol{\delta}'\mathbf{V}\boldsymbol{\delta}$$

is a convex random function which is uniquely minimized at the solution of $\partial H\left(\boldsymbol{\delta}\right)/\partial\boldsymbol{\delta}' = 0$, namely,

$$\widehat{\boldsymbol{\delta}} = \frac{1}{f\left(F^{-1}\left(\tau\right)\right)}\mathbf{V}^{-1}\mathbf{N}_\tau$$

Therefore, under the null hypothesis $\mathrm{H}_0 : \theta = 0$ and the set of assumptions considered, we conclude (Knight, 1989, 1991; Pollard, 1991) that

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}\left(\tau\right) - \boldsymbol{\beta}\left(\tau\right)\right) \to \mathcal{N}\left(0, \frac{\tau\left(1-\tau\right)}{f^2\left(F^{-1}\left(\tau\right)\right)}\mathbf{V}^{-1}\right)$$

with $\boldsymbol{\beta}\left(\tau\right) = \left(\alpha\left(\tau\right),0\right)'$, $\alpha\left(\tau\right) = F^{-1}\left(\tau\right)$, so $\widehat{\boldsymbol{\beta}}\left(\tau\right)$ is consistent and asymptotically normal distributed. Hence,

$$\sqrt{T}\widehat{\phi}_{QR}\left(\tau\right) \to \mathcal{N}\left(0, \frac{6\tau\left(1-\tau\right)}{\left[\sigma\pi f\left(F^{-1}\left(\tau\right)\right)\right]^2}\right)$$

from which the claimed results follow directly. ∎

Proof of Theorem 3.2. Portnoy (1984) and Gutenbrunner and Jurečková (1992) showed that the QR process is tight, so the limit distribution of the function $\xi_T\left(\tau\right) = \sqrt{T}\,\widehat{\phi}_{QR}\left(\tau\right)$, seen as a random function of $\tau \in \left(0,1\right)$, is a rescaled (or non-standard) Brownian bridge under the null hypothesis und the conditions in Theorem 3.1, with (14) arising for any fixed τ. Since $\widehat{\kappa}\left(\tau\right) \xrightarrow{p} 6\left[\sigma\pi f\left(F^{-1}\left(\tau\right)\right)\right]^{-2}$ uniformly on τ, following the arguments in Portnoy (1984), the scaled process $S_T\left(\tau\right) = \xi_T\left(\tau\right)/\sqrt{\widehat{\kappa}\left(\tau\right)} \to \mathcal{B}\left(\tau\right)$ in $\left(0,1\right)$. Then, the limits stated for the Kolmogorov-Smirnov and the Cramér von Mises type-tests in (15) and (16) follow directly from the continuous mapping theorem. ∎

The following lemmae comprise the main elements necessary to prove the remaining theorems stated in the paper under short-run dependence.

Lemma B1. *Consider the assumptions in Theorem 3.3 and note that $\{\varepsilon_t\}$ admits the Wold representation $\varepsilon_t = \sum_{j=0}^{\infty} b_j v_{t-j}$ with $\sum_{j=0}^{\infty} j|b_j| < \infty$. Let $\left\{\varphi_j\right\}_{j\geq 0}$ be the j-th element in the serial convolution of $\{j^{-1}\}_{j\geq 1}$ and $\{b_j\}_{j\geq 0}$. Then, $x^{**}_{t-1,d} = \sum_{j=0}^{\infty} \varphi_j v_{t-j-1}$ and $x^*_{t-1,d} = \sum_{j=0}^{t} \varphi_j v_{t-j-1}$, where $\varphi_0 = 1$ and $\varphi_j = O\left(1/j\right)$ for $j \geq 1$.*

Proof. See Hassler *et al.* (2009, Lemma B.4).

This important result shows that the essential statistical properties of $x^*_{t-1,d}$ and $x^{**}_{t-1,d}$ are preserved under the assumptions of Theorem 3.3, since $\{\varphi_j\}_{j\geq 0}$ belongs to the same space of square-summable coefficient series as $\{j^{-1}\}_{j\geq 1}$. Hence, the results discussed in Lemma A keep holding and, similarly, the asymptotic limits of the normalized sums involved in the proof of Theorem 3.1 converge at the same rates. The following lemma comprises the main additional elements to show the results of Theorem 3.3.

Lemma B2. *Let* $\mathbf{z}^{**}_{pt-1,d} = \left(1, x^{**}_{t-1,d}, \varepsilon_{t-1}, ..., \varepsilon_{t-p}\right)'$, *with* $x^{**}_{t-1,d} = \sum_{j=0}^{\infty} \varphi_j v_{t-j-1}$, $\varepsilon_t = \sum_{j=0}^{\infty} b_j v_{t-j}$, *and define* $\mathbf{\Omega}^{**}_p = T^{-1} \sum_{t=p+1}^{T} \mathbf{z}^{**}_{pt-1,d} \mathbf{z}'^{**}_{pt-1,d}$ *and* $\mathbf{\Omega}^*_p = T^{-1} \sum_{t=p+1}^{T} \mathbf{z}^*_{pt-1,d} \mathbf{z}'^*_{pt-1,d}$, *where* $\mathbf{z}^*_{pt-1,d}$ *is the finite-sample analog of* $\mathbf{z}^{**}_{pt-1,d}$. *Recall that* $\psi_\tau(a) = \tau - \mathbb{I}(a < 0)$, *and let* $u_{t\tau} = v_t - F^{-1}(\tau)$, *with* $F(\cdot)$ *denoting the cumulative distribution function of* v_t. *Then:*

i) $E\left(\mathbf{z}^{**}_{pt-1,d} \mathbf{z}'^{**}_{pt-1,d}\right) = \mathbf{\Omega}_p$, *bounded and bounded away from zero, and invertible.*

ii) $T^{-1/2} \sum_{t=p+1}^{T} \psi_\tau(u_{t\tau}) \mathbf{z}^{**}_{pt-1,d} \rightarrow \mathcal{N}\left(0, \tau(1-\tau)\mathbf{\Omega}_p\right)$.

iii) $\|\mathbf{\Omega}^{**}_p - \mathbf{\Omega}_p\| = o_p(1)$.

iv) $\left\|\mathbf{\Omega}^{**}_p - \mathbf{\Omega}^*_p\right\| = o_p(1)$.

v) $\left\|T^{-1/2} \sum_{t=p+1}^{T} \psi_\tau(u_{t\tau})\left(\mathbf{z}^{**}_{pt-1,d} - \mathbf{z}^*_{pt-1,d}\right)\right\| = o_p(1)$.

Proof. For *i*), note that $\mathbf{\Omega}_p$ can be partitioned as

$$\left(\begin{array}{cc} [\mathbf{\Sigma}_{11}]_{2\times 2} & [\mathbf{\Sigma}'_{12}]_{2\times p} \\ [\mathbf{\Sigma}_{12}]_{p\times 2} & [\mathbf{\Sigma}_{22}]_{p\times p} \end{array}\right)$$

with $\mathbf{\Sigma}_{11} = diag\left(1, \sigma^2 \sum_{j=0}^{\infty} \varphi_j^2\right)$, $\mathbf{\Sigma}_{22} = \sigma^2 \sum_{j=0}^{\infty} \mathbf{b}_j \mathbf{b}'_j$, $\mathbf{b}_j = (b_{j-1}, ..., b_{j-p})'$ and $b_l = 0$ for all $l < 0$, and $\mathbf{\Sigma}_{12} = (\mathbf{0}_p, \mathbf{\Upsilon}_p)$, with $\mathbf{\Upsilon}_p = E\left(x^{**}_{t-1,d}\varepsilon_{t-1}, .., x^{**}_{t-1,d}\varepsilon_{t-p}\right)$ and $\mathbf{0}_p$ a conformable column of zeros. This matrix is trivially bounded away from zero, and since $\mathbf{\Sigma}_{11}$ and $\mathbf{\Sigma}_{22}$ are bounded, so is $\mathbf{\Sigma}_{12}$ noting that $\|\mathbf{\Sigma}_{12}\| \leq \|\mathbf{\Sigma}_{11}\|^{1/2}\|\mathbf{\Sigma}_{22}\|^{1/2}$ from the Cauchy-Schwarz inequality. Finally, $\mathbf{\Omega}_p$ is non-singular, as the elements of $\mathbf{z}^{**}_{pt-1,d}$ are not linearly dependent. Statement *ii*) holds since, as under the assumptions of Theorem 3.1, $\left\{\mathbf{\Lambda}^{**}_{\tau p}, \mathcal{G}_t\right\}$, with $\mathbf{\Lambda}^{**}_{\tau p} = \psi_\tau(u_{t\tau})\mathbf{z}^{**}_{pt-1,d}$, is a stationary and ergodic vector MDS bounded under the L_2 norms, with $\mathcal{G}_{t-1}$ being the σ-field generated by $\{v_s, s < t\}$, and $\psi_\tau(u_{t\tau})$ independent of $\mathbf{z}^{**}_{pt-1,d}$. From the ergodic theorem, $T^{-1} \sum_{t=p+1}^{T} \mathbf{\Lambda}^{**}_{\tau p} \mathbf{\Lambda}'^{**}_{\tau p} \overset{a.s}{\rightarrow} \tau(1-\tau)\mathbf{\Omega}_p$, and since $\max_{t\geq p+1}\left|x^*_{t-1,d}\right| = o_p\left(\sqrt{T}\right)$ and $\max_{t\geq p+1}\left|\varepsilon_{t-k}\right| = o_p\left(\sqrt{T}\right)$, $k = 1, ..., p$, we can show the required result from the CLT for MDS (Davidson, 1994, Theorem 24.3) and the Cramér-Wold device. Result *iii*) follows directly from the ergodic theorem. Part *iv*) holds true if (a) $\|\mathbf{\Sigma}^{**}_{11,T} - \mathbf{\Sigma}^*_{11}\| = o_p(1)$, (b) $\|\mathbf{\Sigma}^{**}_{22} - \mathbf{\Sigma}^*_{22}\| = o_p(1)$ and (c) $\|\mathbf{\Sigma}^{**}_{12} - \mathbf{\Sigma}^*_{12}\| = o_p(1)$, where $\mathbf{\Sigma}^{**}_{ij}$ and $\mathbf{\Sigma}^*_{ij}$ are the corresponding submatrices of $\mathbf{\Omega}^{**}_p$ and $\mathbf{\Omega}^*_p$, respectively. Paralleling the proof of Theorem 3.1 above we can show that (a) holds true with $\|\mathbf{\Sigma}^{**}_{11} - \mathbf{\Sigma}^*_{11}\| = O\left(1/\sqrt{T}\right)$; and (b) holds trivially because $\mathbf{\Sigma}^{**}_{22} = \mathbf{\Sigma}^*_{22}$. To prove

(c), note that $\varepsilon_{t-k}\left(x^{**}_{t-1,d} - x^{*}_{t-1,d}\right) \equiv \varepsilon_{t-k}\varphi_{t-1}$, and since $E\left(\varepsilon_{t-k}\varphi_{t-1}\right) = O\left(1/\sqrt{t}\right)$ from the Cauchy-Schwarz inequality and Lemma Aii), it follows that

$$E\|\boldsymbol{\Sigma}^{**}_{12} - \boldsymbol{\Sigma}^{*}_{12}\| \leq \frac{1}{T-p}\sum_{t=p+1}^{T}\sum_{k=1}^{p} E|\varepsilon_{t-k}\varphi_{t-1}| = O\left(1/\sqrt{T}\right)$$

and therefore $\left\|\boldsymbol{\Omega}^{**}_{p} - \boldsymbol{\Omega}^{*}_{p}\right\| = O_p\left(1/\sqrt{T}\right) = o_p\left(1\right)$. Together with iii), this implies $\boldsymbol{\Omega}^{*}_{p} \xrightarrow{p} \boldsymbol{\Omega}_p$ from the AEL. Finally, part v) holds if $\left\|T^{-1/2}\left(\boldsymbol{\Lambda}^{**}_{\tau p} - \boldsymbol{\Lambda}^{*}_{\tau p}\right)\right\| = o_p\left(1\right)$, but, since

$$E\left\|T^{-1/2}\left(\boldsymbol{\Lambda}^{**}_{\tau p} - \boldsymbol{\Lambda}^{*}_{\tau p}\right)\right\|^2 \leq T^{-1}\sum_{t=2}^{T} E\left[\psi^2_\tau\left(u_{t\tau}\right)\varphi^2_{t-1}\right] = O\left(\frac{\log T}{T}\right)$$

as in the proof of Theorem 3.1, we have $\left\|T^{-1/2}\left(\boldsymbol{\Lambda}^{**}_{\tau p} - \boldsymbol{\Lambda}^{*}_{\tau p}\right)\right\| = O_p\left(\sqrt{\log T/T}\right) = o_p\left(1\right)$ from Markov's inequality. Together with ii) above, this implies that $T^{-1/2}\sum_{t=p+1}^{T}\psi_\tau\left(u_{t\tau}\right)\mathbf{z}^{*}_{pt-1,d} \to \mathcal{N}\left(0, \tau\left(1-\tau\right)\boldsymbol{\Omega}_p\right)$. $\blacksquare$

Proof of Theorem 3.3. The proof is now obvious in view of Lemma Aiii) and Lemma B2 and follows parallel to that of Theorem 3.1. Hence, for the sake of space we do not present the details, but these are available upon request. $\blacksquare$

Proof of Theorem 3.4. Follows directly from tightness of the QR process and the continuous mapping theorem as in Theorem 3.2. $\blacksquare$

Figures and Tables

Figure 1: Daily measures of realized variation of IBM from 04/01/1993 to 31/05/2007 estimated from 5-minute log-returns. These are realized volatility $\sigma_{RV}(t) = \left[\sum_{n=1}^{156} r_{(n),t}^2\right]^{1/2}$, (unnormalized) realized power variation $\sigma_{RPV}(t) = \sum_{m=1}^{156} |r_{(n),t}|$, and logarithmic transforms of these variables.
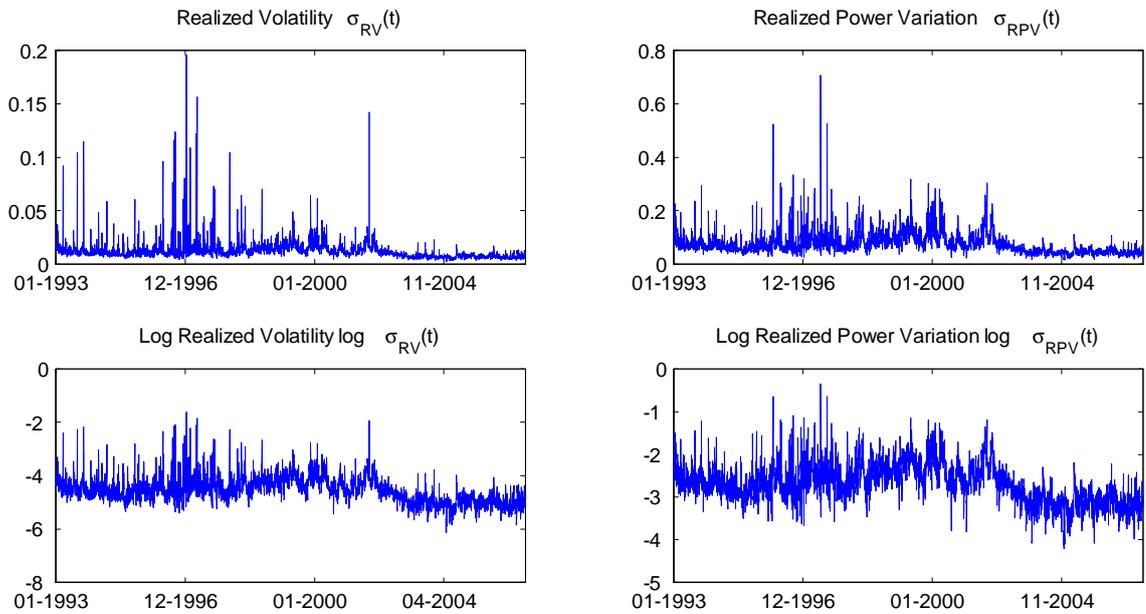
Figure 2: Sample Autocorrelation Function (ACF) of the measures of daily realized variation in Figure 1 together with upper 95% confidence band (dashed red line).
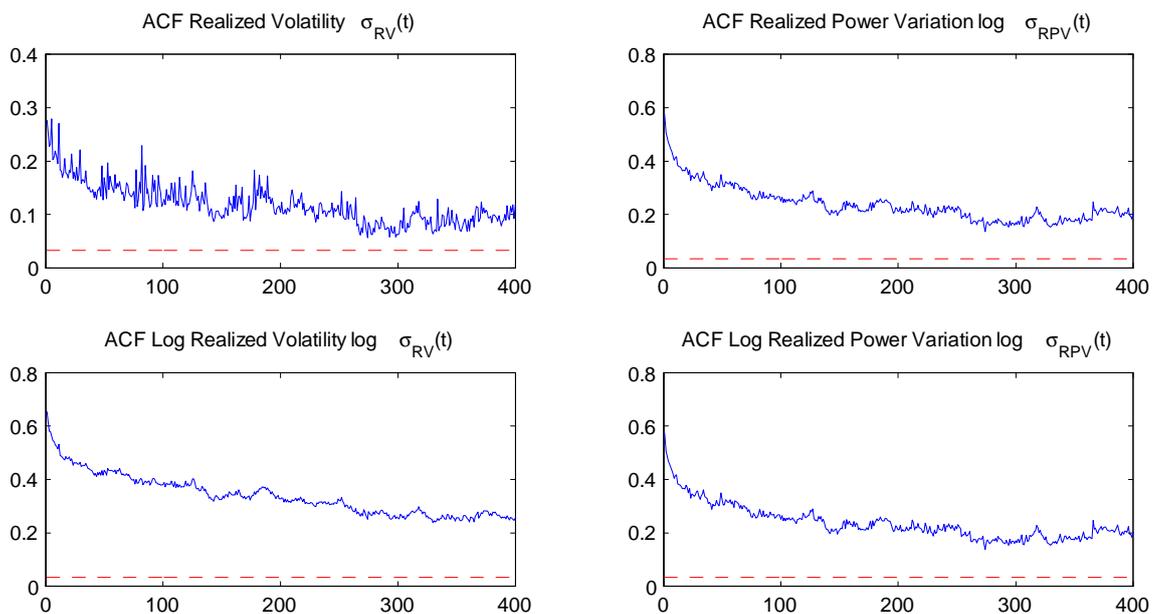
Figure 3: Estimates of the long-memory parameter of $\log \sigma_{RV}(t)$ and $\log \sigma_{RPV}(t)$ from the QRLM testing procedure and respective 95% and 99% confidence intervals. For any quantile $\tau \in \mathcal{Q}$, 'Central' denotes the value of $d \in \mathcal{D}$ for which the test statistic $|t_{QR\tau,p}|$ is closer to zero, *i.e.*, the value which provides maximum sample evidence for the null hypothesis. The remaining entries correspond to the upper and lower bands of the confidence intervals $\mathcal{CI}_{95\%}(d|\tau)$ and $\mathcal{CI}_{99\%}(d|\tau)$ constructed by inverting $t_{QR\tau,p}$.
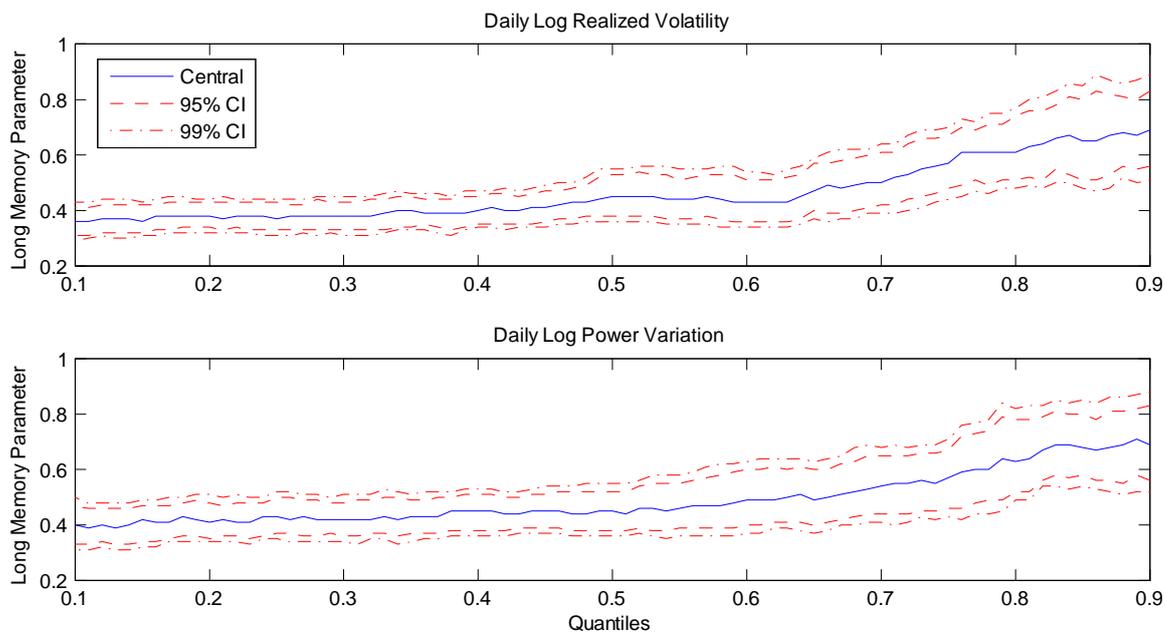
Table 1: Empirical rejection frequencies at the 5% nominal size of long-memory tests under i.i.d. errors. Data are generated according to $(1-L)^{1+\theta}\,y_t=\varepsilon_t$, with ε_t being an i.i.d. sample drawn from a Student-t with v degrees of freedom, $t=1,...,T$. Tests statistics are computed under $\mathrm{H}_0:\theta=0$. The entries $LM_{\mathrm{QR}\tau}^{\mathrm{K}}$ and $LM_{\mathrm{QR}\tau}^{\mathrm{B}}$ denote the rejection frequencies in percentages of the QRLM test $\mathrm{LR}_{QR\tau}$ at $\tau=1/2$ with covariance matrix computed with a kernel density estimate and bootstrap scheme, respectively. The entries LM_{LS} and LM_{DV} denote the rejection frequencies of the LM tests in Breitung and Hassler (2002) and the nonparametric in Delgado and Velasco (2005), respectively

	$v=2$				$v=3$				$v=1000\ (Gaussian)$			
θ	$LM_{\mathrm{QR}\tau}^{K}$	$LM_{\mathrm{QR}\tau}^{B}$	LM_{LS}	LM_{DV}	$LM_{\mathrm{QR}\tau}^{K}$	$LM_{\mathrm{QR}\tau}^{B}$	LM_{LS}	LM_{DV}	$LM_{\mathrm{QR}\tau}^{K}$	$LM_{\mathrm{QR}\tau}^{B}$	LM_{LS}	LM_{DV}
						T=100						
-0.30	98.80	99.00	94.02	19.12	94.24	95.93	93.70	23.38	76.22	77.13	92.44	24.38
-0.20	89.80	91.70	66.04	11.76	74.88	77.90	64.36	12.40	49.04	49.55	62.70	11.84
-0.10	46.18	53.48	17.08	5.86	30.78	37.33	19.54	5.10	22.16	24.18	22.62	4.18
0.00	4.72	6.37	3.78	3.56	6.04	8.20	4.34	3.60	7.56	8.13	4.64	3.12
0.10	36.64	44.42	22.64	35.34	19.82	23.90	24.32	23.94	10.98	12.38	24.68	16.86
0.20	83.18	85.70	71.34	76.34	63.84	66.43	69.64	64.50	34.42	35.73	68.68	48.52
0.30	96.82	97.53	93.70	94.16	90.08	90.73	93.54	88.38	68.98	67.00	93.38	78.04
						T=250						
-0.30	99.98	100.00	99.94	62.00	100.00	99.93	99.98	77.2	98.64	100.00	100.00	82.62
-0.20	99.88	100.00	97.88	44.98	98.72	98.73	97.22	54.64	85.14	85.23	97.04	54.40
-0.10	87.12	89.93	46.16	23.28	64.32	67.70	47.46	23.64	37.54	38.68	48.04	19.40
0.00	3.78	6.23	3.62	5.06	5.36	6.38	4.14	5.4	6.42	7.05	4.88	5.18
0.10	81.50	84.13	55.02	65.40	55.68	59.70	54.12	47.1	27.24	27.38	53.40	31.24
0.20	99.72	100.00	97.46	98.32	97.86	98.25	97.10	94.06	79.10	81.60	96.76	82.30
0.30	99.98	100.00	99.86	99.96	99.94	99.93	99.96	99.8	98.46	100.00	99.94	98.60
						T=1000						
-0.30	99.98	100.00	100.00	96.50	100.00	100.00	100.00	99.96	100.00	100.00	100.00	100.00
-0.20	100.00	100.00	100.00	89.32	99.98	100.00	100.00	98.98	99.96	100.00	100.00	99.00
-0.10	99.98	100.00	98.46	64.42	99.68	99.70	98.12	73.14	88.32	89.70	98.28	62.14
0.00	3.76	6.27	3.94	5.36	4.14	5.58	4.48	5.40	5.34	5.93	4.66	4.66
0.10	100.00	100.00	98.08	99.46	99.44	99.43	97.94	94.74	86.58	86.23	97.82	81.06
0.20	100.00	100.00	99.98	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.96
0.30	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 2: Empirical rejection frequencies at the 5% nominal size of long-memory tests under short-run dependence. Data are generated according to $(1 - L)^{1+\theta} y_t = \varepsilon_t$, with $(1 - aL)\varepsilon_t = v_t$, and v_t being an i.i.d. sample drawn from a Student-t with v degrees of freedom, $t = 1, ..., T$. Tests statistics are computed under $H_0 : \theta = 0$. The entries $LM_{QR\tau,p}$ and LM_{LS} denotes the rejection frequencies in percentages of the augmented QRLM test at $\tau = 1/2$ with covariance matrix computed with a kernel density, and the least-squares test from an augmented regression, respectively.

| | $v = 2$ | | | | $v = 3$ | | | | $v = 1000$ (*Gaussian*) | | | |
| | $LM_{QR\tau,p}$ | | LM_{LS} | | $LM_{QR\tau,p}$ | | LM_{LS} | | $LM_{QR\tau,p}$ | | LM_{LS} | |
θ	$a = 0.5$	$a = 0.75$	$a = 0.5$	$a = 0.75$	$a = 0.5$	$a = 0.75$	$a = 0.5$	$a = 0.75$	$a = 0.5$	$a = 0.75$	$a = 0.5$	$a = 0.75$
						T=100						
-0.30	43.14	5.68	18.42	3.50	31.42	6.60	18.98	3.96	23.10	9.90	21.36	4.60
-0.20	22.06	4.44	9.90	2.62	17.90	6.02	10.50	3.38	17.16	9.48	11.66	3.94
-0.10	10.20	3.78	5.56	2.68	10.64	6.10	5.88	3.72	13.26	9.58	6.70	4.36
0.00	5.22	5.16	3.74	3.52	7.22	8.04	4.98	4.86	11.06	11.76	5.20	4.92
0.10	4.36	8.30	4.34	4.94	6.44	11.44	5.48	6.20	9.46	15.06	6.06	6.18
0.20	5.70	14.40	6.28	7.10	6.18	18.48	7.26	8.56	10.18	19.92	8.02	8.60
0.30	7.84	25.56	7.72	11.46	7.44	29.06	8.50	12.58	10.44	29.20	9.08	13.10
						T=250						
-0.30	84.94	12.76	47.38	4.52	63.76	8.74	48.70	5.58	36.18	7.78	47.90	5.80
-0.20	54.88	5.48	18.72	3.10	35.28	6.10	22.50	3.80	20.86	6.44	22.20	3.68
-0.10	17.78	3.12	7.00	2.88	14.14	4.78	8.06	3.68	11.86	6.16	8.70	3.14
0.00	3.88	4.08	3.52	3.54	6.06	5.46	4.78	4.92	7.24	7.30	4.72	4.20
0.10	8.14	6.32	7.18	5.48	6.04	8.62	8.08	6.40	6.06	9.88	8.82	6.16
0.20	20.48	12.86	15.52	7.22	9.22	15.24	16.54	9.08	6.50	15.42	16.44	9.32
0.30	28.16	23.20	22.52	11.12	12.32	25.80	22.38	12.92	6.82	24.18	23.42	13.38
						T=1000						
-0.30	100.00	56.34	98.98	13.86	99.76	26.34	98.82	17.28	89.10	13.78	98.62	16.42
-0.20	99.30	24.14	78.56	5.74	89.02	10.50	76.20	7.26	56.16	7.38	76.58	7.20
-0.10	67.06	7.42	19.98	3.40	35.76	5.18	23.54	4.50	19.38	5.52	22.86	4.00
0.00	3.06	3.20	3.96	3.92	4.58	5.14	5.28	5.16	5.42	5.94	4.82	4.84
0.10	48.24	8.62	21.22	6.52	19.62	7.56	21.86	7.94	9.54	7.44	22.20	8.10
0.20	86.88	16.84	56.44	11.32	53.20	12.62	54.82	12.84	23.12	12.24	54.32	12.28
0.30	94.86	24.98	74.60	15.18	72.10	20.90	73.02	16.94	34.10	21.04	72.24	17.72

Table 3: Descriptive statistics and semiparametric estimators of the long-memory parameter for different measures of daily realized variation: realized volatility $\sigma_{RV}(t) = \left[\sum_{n=1}^{156} r_{(n),t}^2\right]^{1/2}$, realized power variation $\sigma_{RPV}(t) = \sum_{n=1}^{156} |r_{(n),t}|$ and logarithmic transformations of these series. The statistic JB is the nonparametric Jarque-Bera test for unconditional normality of the series (p-values in brackets), distributed as $\chi_{(2)}^2$. The statistic $\widehat{\rho}_k$ denotes the k-th order sample autocorrelation. The Geweke-Porter-Hudak (1983) estimator of d is denotes as $\widehat{d}_{GPH}$, whereas the exact local Whittle estimator of Shimotsu and Phillips (2005) is denoted as $\widehat{d}_{ELW}$. $\mathcal{CI}_{95\%}(d)$ denotes the 95% asymptotic confidence interval for d for any of these estimates.

	$\sigma_{RV}(t)$	$\log \sigma_{RV}(t)$	$\sigma_{RPV}(t)$	$\log \sigma_{RPV}(t)$
Mean	0.012	-4.547	0.076	-2.690
Median	0.010	-4.580	0.066	-2.714
Std.Dev.	0.009	0.481	0.041	0.456
Skewness	7.775	0.752	3.255	0.3148
Kurtosis	102.81	5.271	30.028	3.448
JB	1.54e+006 (0.00)	1.12e+003 (0.00)	1.16e+005 (0.00)	90.32 (0.00)
$\widehat{\rho}_1$	0.276	0.653	0.575	0.724
$\widehat{\rho}_{400}$	0.112	0.258	0.202	0.277
$\widehat{d}_{GPH}$	0.438	0.512	0.475	0.545
$\mathcal{CI}_{95\%}^{GPH}(d)$	[0.33,0.55]	[0.40,0.62]	[0.36,0.58]	[0.44,0.65]
$\widehat{d}_{ELW}$	0.397	0.488	0.464	0.508
$\mathcal{CI}_{95\%}^{ELW}(d)$	[0.31,0.48]	[0.40,0.57]	[0.38,0.54]	[0.42,0.59]

Table 4: Quantile regression test statistics for long memory in logs of daily realized volatility. The top of the table presents the individual t-statistics for H_0: $d = d_0$, with $d_0 = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{(1-\alpha)\times 100\%}(d|\tau)$ columns show the $(1 - \alpha) \times 100\%$ confidence intervals for d determined as the non-rejection region of the test at a $\alpha \times 100\%$ nominal level given the value of τ. The entry LS shows the corresponding test and confidence interval based on the least-squares statistic for the conditional mean. The bottom part of the table reports joint test statistics of H_0: $d = d_0$ in the set of quantiles $\mathcal{T}_1 =$[0.4,0.6] and $\mathcal{T}_2 =$[0.1,0.9]. The test statistics are the Kolmogorov-Smirnov ($\mathcal{KS}$) and Cramer von Mises ($\mathcal{CM}$) type tests described in Corollary 3.1 computed over these intervals. The $\mathcal{CI}_{(1-\alpha)\times 100\%}(d|\mathcal{T})$ columns show confidence interval for d determined as the non-rejection region of the joint test at a $\alpha \times 100\%$ nominal level given the $\mathcal{T}$ quantile intervals (see Table 6 for critical values). All statistics have been computed from an auxiliary regression augmented with p lags of the dependent variable according to Schwert's rule with $p = [4(T/100)^{1/4}]$.

	Panel A: Individual Test: $H_0 : d = d_0$ at τ											$\mathcal{CI}_{95\%}(d|\tau)$	$\mathcal{CI}_{99\%}(d|\tau)$
τ	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**		
0.9	3.09	4.00	4.39	4.39	3.65	2.65	1.28	-0.18	-1.64	-2.62	-3.50	[0.56,0.83]	[0.51,0.89]
0.8	5.27	5.89	5.86	5.13	4.06	2.06	0.27	-1.35	-2.90	-4.51	-6.11	[0.51,0.74]	[0.38.0.77]
0.7	6.31	6.67	5.70	4.17	2.30	0.11	-1.81	-3.49	-5.02	-6.82	-8.51	[0.42,0.62]	[0.38,0.64]
0.6	7.86	7.66	6.21	3.62	0.70	-1.63	-3.52	-5.16	-6.89	-8.48	-9.86	[0.36,0.51]	[0.34,0.54]
0.5	8.42	8.22	6.59	4.10	1.32	-1.24	-3.54	-5.79	-7.67	-9.41	-10.96	[0.38,0.53]	[0.36,0.56]
0.4	9.20	8.69	6.79	3.79	0.12	-3.41	-6.23	-8.32	-9.90	-10.65	-11.84	[0.35,0.45]	[0.34,0.46]
0.3	9.10	8.53	5.96	2.84	-0.83	-3.93	-6.65	-8.86	-10.25	-11.24	-11.72	[0.33,0.43]	[0.31,0.45]
0.2	10.44	9.36	7.05	3.26	-0.73	-4.54	-7.34	-9.16	-10.64	-11.68	-11.91	[0.34,0.43]	[0.32,0.44]
0.1	9.20	8.02	5.82	2.07	-1.44	-4.50	-7.18	-8.59	-9.12	-9.36	-10.23	[0.31,0.41]	[0.29,0,43]
LS	7.68	8.22	7.40	5.50	3.09	0.59	-1.93	-4.42	-6.67	-8.66	-10.46	[0.45,0.60]	[0.42,0.62]
	Panel B: Quantile Regression based Joint Tests $H_0 : d = d_0$ over $\mathcal{T}$											$\mathcal{CI}_{95\%}(d|\mathcal{T})$	$\mathcal{CI}_{99\%}(d|\mathcal{T})$
$\mathcal{KS}$ **[0.4,0.6]**	4.82	4.48	3.54	2.15	0.69	1.67	3.05	4.08	4.85	5.41	6.04	[0.36,0.48]	[0.35,0.49]
$\mathcal{CM}$ **[0.4,0.6]**	4.01	3.62	2.28	0.78	0.05	0.24	1.04	2.12	3.43	4.79	6.34	[0.38,0.48]	[0.35,0.50]
$\mathcal{KS}$ **[0.1,0.9]**	4.82	4.48	3.54	2.22	1.63	1.99	3.18	4.23	4.85	5.46	6.12	[0.44,0.45]	[0.40,0.47]
$\mathcal{CM}$ **[0.1,0.9]**	10.74	9.81	6.19	2.31	0.50	1.28	3.71	6.78	10.07	13.42	16.88	[0.41,0.42]	[0.38,0.46]

Table 5: Quantile regression test statistics for long memory in logs of daily power variation. The top of the table presents the individual t-statistics for H$_0$: $d = d_0$, with $d_0 = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{(1-\alpha)\times100\%}(d|\tau)$ columns show the $(1-\alpha)\times100\%$ confidence intervals for d determined as the non-rejection region of the test at a $\alpha\times100\%$ nominal level given the value of τ. The entry LS shows the corresponding test and confidence interval based on the least-squares statistic for the conditional mean. The bottom part of the table reports joint test statistics of H$_0$: $d = d_0$ in the set of quantiles $\mathcal{T}_1 = [0.4,0.6]$ and $\mathcal{T}_2 = [0.1,0.9]$. The test statistics are the Kolmogorov-Smirnov ($\mathcal{KS}$) and Cramer von Mises ($\mathcal{CM}$) type tests described in Corollary 3.1 computed over these intervals. The $\mathcal{CI}_{(1-\alpha)\times100\%}(d|\mathcal{T})$ columns show confidence interval for d determined as the non-rejection region of the joint test at a $\alpha\times100\%$ nominal level given the $\mathcal{T}$ quantile intervals (see Table 6 for critical values). All statistics have been computed from an auxiliary regression augmented with p lags of the dependent variable according to Schwert's rule with $p = [4(T/100)^{1/4}]$.

45

	Panel A: Individual Test: H$_0$: $d = d_0$ at τ											$\mathcal{CI}_{95\%}(d|\tau)$	$\mathcal{CI}_{99\%}(d|\tau)$
τ	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**		
0.9	3.18	4.65	5.14	4.84	3.88	2.87	1.47	-0.24	-1.63	-2.85	-4.10	[0.56,0.84]	[0.52,0.88]
0.8	4.74	5.22	5.56	5.01	3.82	2.31	0.54	-0.82	-2.26	-3.50	-4.64	[0.52,0.78]	[0.49,0.82]
0.7	5.64	6.12	5.54	4.34	2.65	0.73	-1.03	-2.79	-4.48	-6.11	-7.63	[0.44,0.65]	[0.41,0.69]
0.6	6.79	6.60	5.51	3.92	1.80	-0.18	-1.93	-3.53	-5.01	-6.42	-7.65	[0.39,0.60]	[0.37,0.63]
0.5	8.66	8.12	6.56	3.83	1.41	-1.33	-3.67	-5.48	-7.52	-8.37	-9.12	[0.38,0.52]	[0.36,0.55]
0.4	7.99	7.87	6.77	4.30	1.26	-1.59	-4.08	-6.25	-8.08	-10.08	-11.06	[0.38,0.51]	[0.36,0.53]
0.3	8.31	7.66	6.13	3.67	0.69	-2.23	-4.98	-6.80	-7.76	-9.00	-9.96	[0.36,0.48]	[0.34,0.51]
0.2	7.50	7.36	6.28	3.70	0.47	-2.24	-4.68	-6.66	-7.88	-9.30	-10.50	[0.35,0.48]	[0.34,0.51]
0.1	7.46	7.14	5.48	2.78	0.06	-2.57	-4.36	-5.91	-7.09	-8.05	-8.45	[0.33,0.47]	[0.31,0.50]
LS	7.48	8.11	7.44	5.68	3.35	0.90	-1.60	-4.06	-6.30	-8.25	-10.00	[0.36,0.51]	[0.34,0.54]
	Panel B: Quantile Regression based Joint Tests H$_0$: $d = d_0$ over $\mathcal{T}$											$\mathcal{CI}_{95\%}(d|\mathcal{T})$	$\mathcal{CI}_{99\%}(d|\mathcal{T})$
$\mathcal{KS}$ **[0.4,0.6]**	4.33	4.09	3.47	2.29	0.90	0.87	2.19	3.27	4.18	4.96	5.59	[0.37,0.52]	[0.35,0.54]
$\mathcal{CM}$ **[0.4,0.6]**	3.27	3.11	2.12	0.88	0.11	0.08	0.62	1.48	2.54	3.61	4.60	[0.39,0.52]	[0.37,0.54]
$\mathcal{KS}$ **[0.1,0.9]**	4.33	4.09	3.47	2.29	1.84	1.22	2.28	3.27	4.18	4.96	5.59	[0.48,0.51]	[0.43,0.54]
$\mathcal{CM}$ **[0.1,0.9]**	8.13	8.10	5.83	2.70	0.66	0.50	1.89	4.12	6.76	9.54	12.41	[0.43,0.49]	[0.40,0.52]

Table 6: Quantile regression test statistics for long memory in daily realized volatility. The top of the table presents the individual t-statistics for H$_0$: $d = d_0$, with $d_0 = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{(1-\alpha)\times100\%}(d|\tau)$ columns show the $(1-\alpha)\times 100\%$ confidence intervals for d determined as the non-rejection region of the test at a $\alpha \times 100\%$ nominal level given the value of τ. The entry LS shows the corresponding test and confidence interval based on the least-squares statistic for the conditional mean. The bottom part of the table reports joint test statistics of H$_0$: $d = d_0$ in the set of quantiles $\mathcal{T}_1 = [0.4, 0.6]$ and $\mathcal{T}_2 = [0.1, 0.9]$. The test statistics are the Kolmogorov-Smirnov ($\mathcal{KS}$) and Cramer von Mises ($\mathcal{CM}$) type tests described in Theorem 3.4 computed over these intervals. The C.V. 95% and C.V. 99% columns show the critical values of the corresponding distributions, approached trhough experimental simulation. All statistics have been computed from an auxiliary regression augmented with p lags of the dependent variable according to Schwert's rule with $p = \left[4(T/100)^{1/4}\right]$.

	Panel A: Individual Test: H$_0$: $d = d_0$ at τ											$\mathcal{CI}_{95\%}(d|\tau)$	$\mathcal{CI}_{99\%}(d|\tau)$
τ	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**		
0.9	6.99	10.83	13.75	16.36	15.33	12.65	9.50	6.00	2.83	-0.23	-1.51	[0.84,1.04]	[0.81,1.07]
0.8	10.32	12.97	14.18	13.82	11.84	8.08	3.35	-0.33	-3.32	-5.91	-8.49	[0.63,0.74]	[0.62,0.77]
0.7	13.05	12.27	10.50	7.79	3.83	-1.17	-5.54	-9.66	-13.47	-16.67	-20.35	[0.44,0.51]	[0.43,0.53]
0.6	16.70	13.72	8.33	2.14	-3.69	-9.47	-14.36	-17.79	-20.30	-23.18	-26.31	[0.32,0.37]	[0.30,0.37]
0.5	15.36	10.52	4.28	-2.68	-9.56	-14.98	-17.74	-20.15	-22.80	-24.90	-26.34	[0.24,0.28]	[0.23,0.29]
0.4	17.31	10.47	1.11	-8.78	-16.93	-22.29	-26.88	-27.57	-27.00	-27.32	-29.81	[0.19,0.23]	[0.18,0.24]
0.3	17.98	8.60	-2.07	-12.56	-21.24	-28.60	-29.68	-27.72	-26.20	-23.86	-22.57	[0.17,0.22]	[0.16,0.22]
0.2	19.25	8.54	-4.39	-17.18	-26.13	-28.20	-27.12	-22.47	-20.27	-18.91	-17.85	[0.15,0.19]	[0.15,0.20]
0.1	15.37	5.15	-7.61	-18.90	-28.64	-32.12	-26.74	-18.74	-16.93	-15.66	-14.00	[0.13,0.15]	[0.12,0.16]
LS	8.51	7.68	5.74	3.22	0.54	-2.10	-4.68	-7.13	-9.40	-11.52	-13.53	[0.35,0.48]	[0.33,0.52]
	Panel B: Quantile Regression based Joint Tests H$_0$: $d = d_0$ over $\mathcal{T}$											C.V. 95%	C.V. 99%
$\mathcal{KS}$ **[0.4,0.6]**	9.01	6.72	4.08	4.30	8.29	10.92	13.17	13.50	13.23	14.35	15.39	1.25	1.51
$\mathcal{CM}$ **[0.4,0.6]**	16.29	7.71	1.35	1.28	7.13	15.89	23.04	27.76	32.29	38.75	45.31	0.16	0.27
$\mathcal{KS}$ **[0.1,0.9]**	9.01	6.74	5.67	6.96	11.03	13.40	14.05	13.80	13.35	14.35	15.39	1.35	1.60
$\mathcal{CM}$ **[0.1,0.9]**	40.91	19.82	9.29	17.86	38.83	59.85	67.20	66.81	68.33	74.90	84.55	0.44	0.72

Table 7: Quantile regression test statistics for long memory in daily power variation. The top of the table presents the individual t-statistics for H$_0$: $d = d_0$, with $d_0 = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{(1-\alpha)\times 100\%}(d|\tau)$ columns show the $(1-\alpha)\times 100\%$ confidence intervals for d determined as the non-rejection region of the test at a $\alpha \times 100\%$ nominal level given the value of τ. The entry LS shows the corresponding test and confidence interval based on the least-squares statistic for the conditional mean. The bottom part of the table reports joint test statistics of H$_0$: $d = d_0$ in the set of quantiles $\mathcal{T}_1 = [0.4, 0.6]$ and $\mathcal{T}_2 = [0.1, 0.9]$. The test statistics are the Kolmogorov-Smirnov ($\mathcal{KS}$) and Cramer von Mises ($\mathcal{CM}$) type tests described in Theorem 3.4 computed over these intervals. The C.V. 95% and C.V. 99% columns show the critical values of the corresponding distributions, approached trhough experimental simulation. All statistics have been computed from an auxiliary regression augmented with p lags of the dependent variable according to Schwert's rule with $p = \left[4(T/100)^{1/4}\right]$.

	Panel A: Individual Test: H$_0 : d = d_0$ at τ											$\mathcal{CI}_{95\%}(d|\tau)$	$\mathcal{CI}_{99\%}(d|\tau)$
τ	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**		
0.9	5.01	9.34	11.99	14.73	15.17	13.31	10.48	6.82	3.71	1.92	0.81	[0.87,1.36]	[0.86,1.42]
0.8	6.63	9.41	11.70	12.73	12.63	10.85	7.97	5.54	2.33	-0.05	-1.79	[0.81,1.01]	[0.79,1.09]
0.7	7.34	9.33	9.23	8.71	7.35	5.14	1.98	-1.90	-4.15	-6.45	-8.22	[0.59,0.70]	[0.59,0.73]
0.6	9.77	9.43	7.76	5.49	2.49	-0.24	-2.78	-5.43	-7.95	-10.14	-12.06	[0.41,0.56]	[0.39,0.59]
0.5	10.56	8.89	5.39	0.95	-3.09	-6.98	-9.21	-11.19	-12.61	-14.13	-14.69	[0.28,0.37]	[0.28,0.39]
0.4	10.86	8.23	3.67	-1.85	-7.15	-11.57	-14.76	-16.80	-18.27	-17.21	-18.17	[0.23,0.30]	[0.22,0.31]
0.3	11.68	7.42	1.31	-5.58	-11.94	-16.19	-18.40	-18.32	-18.24	-17.47	-18.47	[0.19,0.25]	[0.18,0.25]
0.2	11.16	5.75	-1.77	-10.01	-17.16	-20.05	-18.37	-17.84	-16.03	-15.95	-15.18	[0.16,0.21]	[0.16,0.22]
0.1	10.16	3.38	-5.14	-12.72	-19.24	-20.96	-18.11	-15.97	-15.40	-14.79	-12.59	[0.12,0.16]	[0.11,0.17]
LS	7.60	7.57	6.40	4.44	2.14	-0.21	-2.54	-4.81	-6.90	-8.82	-10.61	[0.42,0.56]	[0.38,0.61]
	Panel B: Quantile Regression based Joint Tests H$_0 : d = d_0$ over $\mathcal{T}$											**C.V. 95%**	**C.V. 99%**
$\mathcal{KS}$ **[0.4,0.6]**	5.87	4.70	3.80	2.69	3.50	5.67	7.23	8.23	8.95	8.87	8.90	1.25	1.51
$\mathcal{CM}$ **[0.4,0.6]**	6.09	4.18	1.70	0.34	0.85	2.87	5.23	7.47	9.35	10.71	12.46	0.16	0.27
$\mathcal{KS}$ **[0.1,0.9]**	5.87	4.70	5.06	5.63	7.05	8.11	8.57	8.77	8.96	8.87	9.02	1.35	1.60
$\mathcal{CM}$ **[0.1,0.9]**	15.67	11.14	7.74	9.40	16.03	21.90	23.83	24.94	26.13	27.59	30.24	0.44	0.72

WORKING PAPERS

2010

1/10 MEASURING COMOVEMENT IN THE TIME-FREQUENCY SPACE
— *António Rua*

2/10 EXPORTS, IMPORTS AND WAGES: EVIDENCE FROM MATCHED FIRM-WORKER-PRODUCT PANELS
— *Pedro S. Martins, Luca David Opromolla*

3/10 NONSTATIONARY EXTREMES AND THE US BUSINESS CYCLE
— *Miguel de Carvalho, K. Feridun Turkman, António Rua*

4/10 EXPECTATIONS-DRIVEN CYCLES IN THE HOUSING MARKET
— *Luisa Lambertini, Caterina Mendicino, Maria Teresa Punzi*

5/10 COUNTERFACTUAL ANALYSIS OF BANK MERGERS
— *Pedro P. Barros, Diana Bonfim, Moshe Kim, Nuno C. Martins*

6/10 THE EAGLE. A MODEL FOR POLICY ANALYSIS OF MACROECONOMIC INTERDEPENDENCE IN THE EURO AREA
— *S. Gomes, P. Jacquinot, M. Pisani*

7/10 A WAVELET APPROACH FOR FACTOR-AUGMENTED FORECASTING
— *António Rua*

8/10 EXTREMAL DEPENDENCE IN INTERNATIONAL OUTPUT GROWTH: TALES FROM THE TAILS
— *Miguel de Carvalho, António Rua*

9/10 TRACKING THE US BUSINESS CYCLE WITH A SINGULAR SPECTRUM ANALYSIS
— *Miguel de Carvalho, Paulo C. Rodrigues, António Rua*

10/10 A MULTIPLE CRITERIA FRAMEWORK TO EVALUATE BANK BRANCH POTENTIAL ATTRACTIVENESS
— *Fernando A. F. Ferreira, Ronald W. Spahr, Sérgio P. Santos, Paulo M. M. Rodrigues*

11/10 THE EFFECTS OF ADDITIVE OUTLIERS AND MEASUREMENT ERRORS WHEN TESTING FOR STRUCTURAL BREAKS IN VARIANCE
— *Paulo M. M. Rodrigues, Antonio Rubia*

12/10 CALENDAR EFFECTS IN DAILY ATM WITHDRAWALS
— *Paulo Soares Esteves, Paulo M. M. Rodrigues*

13/10 MARGINAL DISTRIBUTIONS OF RANDOM VECTORS GENERATED BY AFFINE TRANSFORMATIONS OF INDEPENDENT TWO-PIECE NORMAL VARIABLES
— *Maximiano Pinheiro*

14/10 MONETARY POLICY EFFECTS: EVIDENCE FROM THE PORTUGUESE FLOW OF FUNDS
— *Isabel Marques Gameiro, João Sousa*

15/10 SHORT AND LONG INTEREST RATE TARGETS
— *Bernardino Adão, Isabel Correia, Pedro Teles*

16/10 FISCAL STIMULUS IN A SMALL EURO AREA ECONOMY
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix, José Francisco Maria*

17/10 FISCAL INSTITUTIONS AND PUBLIC SPENDING VOLATILITY IN EUROPE
— *Bruno Albuquerque*