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

September 2013

Abstract

We introduce a new quantile regression approach to test for long memory in time series. The procedures proposed allow testing at individual and joint quantiles. The latter approach is useful to determine the order of integration over a set of percentiles, allowing in this way to more generally address the overall hypothesis of fractional integration. The null distributions of these tests are standard and free of nuisance parameters. The finite sample validity of the approach is established through Monte Carlo simulations, which also provides evidence of power gains over least-squares procedures under non-Gaussian errors. An empirical application of the new procedures on different measures of daily realized volatility is presented. 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, providing in this way strong support to the existence of long memory in realized volatility from a completely new perspective.

*Acknowledgements: Earlier versions of this paper were presented at an ETSERN meeting at the University of Alicante and research seminars at the University of Sydney, Monash University, and at the Robust Econometric Methods for Modelling Economic and Financial Variables conference at Banco de Portugal, 3rd Humboldt Copenhagen Conference on Financial Econometrics, the 6th Annual Society for Financial Econometrics conference in Singapore, and the 20th Finance Forum Meeting. We thank participants for useful comments and suggestions. We are also grateful to José Faias, Juan Mora, Peter C.B. Phillips, David Veredas, and Zhijie Xiao. Financial support from ECO2011-29751 and ECO2012-33619 projects is gratefully acknowledged.

†Corresponding author. Department of Financial Economics, University of Alicante. CP 03080, Alicante, Spain. Tel/Fax. (34) 965 90 34 00. E-mail: antonio.rubia@ua.es

1 Introduction

There is growing interest in the finance and economics literature on modelling dependence in the tails of the conditional distribution of time series. The Quantile Regression (QR) theory introduced by Koenker and Bassett (1978) provides a simple and convenient instrument to this end. This methodology, now routinely implemented in market downside risk management and other applied areas, distinctively deals with estimation and inference at different quantiles, allowing to address a wide range of hypotheses and offering new insights on the time-series properties of the data. For instance, Engle and Manganelli (2004) show that daily conditional Value-at-Risk is a strongly persistent process. The reason is that downside risk measures are mainly driven by volatility, which distinctively exhibits long-range dependence at the daily frequency 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 rates 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 QR-based test to detect long memory in the time domain. Long-memory (also referred to as fractionally integrated) models allow for long-run dependence characterized by autocovariances that decay hyperbolically, thereby offering an intermediate case between the characteristic exponential decay of short memory and the infinite persistence of unit root processes. This class of models often explains convincingly the time-series dynamics exhibited by many economic and non-economic variables. We propose a series of QR-based tests for fractional integration that extend the Lagrange Multiplier (LM) testing procedure in Breitung and Hassler (2002) and which allow us to address more general hypotheses than the standard unit root case analyzed in previous QR literature; see, among others, Koenker and Xiao (2004) and Galvao (2009). Furthermore, confidence intervals of the long-memory coefficient, exhibiting robust properties to non-Gaussian features of the data, can readily be obtained by inverting these test statistics.

More specifically, this paper discusses the asymptotic theory for both individual and joint quantile regression long memory tests (QRLM henceforth) under a fairly general class of errors. Individual quantile tests are useful to address the fractional integration hypothesis at a specific quantile $0 < \tau < 1$, while joint tests involve sets of quantiles in closed subintervals of $(0, 1)$. We formally show that the asymptotic null distributions of these tests do not depend on the degree of integration nor other nuisance parameters. Furthermore, and in sharp contrast to existing tests for the unit root hypothesis, the

QRLM tests can be characterized by usual probability laws, such as the standard normal distribution. Monte Carlo analysis shows that QR-based tests can yield power improvements over suitable OLS-based alternatives and tend to offer more robust inference against observations drawn from heavy-tailed and/or skewed distributions.

We use the QRLM tests proposed in this paper to formally address the existence of long-memory patterns in different measures of realized variation of IBM, one of the most liquid and actively-traded stocks in the world. These tests allow us to determine the order of fractional integration along the different quantiles of the conditional distribution of the series, thereby addressing more robustly this hypothesis and bringing completely new evidence to the field. The main conclusion from our analysis is that there is sufficient regularity in this process such that the suitability of a long-memory model with constant coefficient slightly greater than 0.40 cannot be rejected. Therefore, according to this analysis, the evidence of long-range dependence in realized volatility is caused by a long-memory model. Hence, our paper contributes to the ongoing debate on whether realized volatility is really driven by a persistent process or not providing new evidence that supports this hypothesis.

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 in this field have exclusively focused on testing the unit-root hypothesis. Our analysis is more general and nests this setting 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 of the data. The tests proposed in this paper are a direct extension of the Least-Squares (LS) based tests proposed by Breitung and Hassler (2002) and further generalized in Demetrescu *et al.* (2008) and Hassler *et al.* (2009). Finally, our paper is related to the empirical literature concerned with realized volatility modelling; see, among others, Andersen *et al.* (2001, 2003), Barndorff-Nielsen and Shephard (2004), and Corsi *et al.* (2008). We introduce a new procedure to detect long-memory and provide robust evidence supporting the existence of long-memory in realized variance.

The remainder of the paper is organized as follows. Section 2 contains the theoretical contribution and lays out the QRLM testing procedures. Section 3 presents experimental evidence on the finite sample size and power performance of the tests in relation to suitable LS-based procedures. Section 4 applies the QRLM tests to characterize the extent of long-run dependence in realized variation of IBM stock returns. Finally, Section 5 summarizes and concludes. A technical appendix collects the proofs of the theorems stated in Section 4.

2 Theoretical analysis

2.1 Assumptions and notation

Consider the fractionally integrated process,

$$(1-L)^{d+\theta}\, y_t = \varepsilon_t,\ t = 1, 2, ...,\ \varepsilon_s = 0 \text{ for all } s \leq 0 \tag{1}$$

where $\{y_t\}$ is an observable variable, L denotes the lag operator, $(d, \theta)'$ is a real-valued vector, and $\{\varepsilon_t\}$ is an invertible short-memory process with specific properties that shall be laid out later in the paper. For a pre-specified value d, our main aim is to test the null hypothesis that $\{y_t\}$ is fractionally integrated of order d, denoted as FI(d), against the alternative FI($d+\theta$), *i.e.*, testing $\mathrm{H}_0 : \theta = 0$ against $\mathrm{H}_1 : \theta \neq 0$. The standard unit-root case, $d = 1$, is encompassed as a particular case in this generalized context.

Let $\{\varepsilon_{t,d}\}$ be the series resulting from differencing $\{y_t\}$ under the null hypothesis, namely, $\varepsilon_{t,d} = (1-L)^d\, y_t$, where the fractional difference operator, $\Delta^d = (1-L)^d$, is characterized by the formal binomial expansion:

$$\Delta^d = \sum_{j=0}^{\infty} \pi_j\,(d)\, L^j, \quad \text{with } \pi_0\,(d) = 1 \text{ and } \pi_j\,(d) = \frac{j-1-d}{j}\, \pi_{j-1}\,(d)\ ,\ j > 0. \tag{2}$$

Given $\{\varepsilon_{t,d}\}$, define the filtered process $x^*_{t-1,d} = \sum_{j=1}^{t-1} j^{-1} \varepsilon_{t-j,d}\ ,\ t = 2, ..., T$. The particular form of $\left\{x^*_{t-1,d}\right\}$, characterized by an harmonic weighting of the lags of $\{\varepsilon_{t,d}\}$, results from the expansion of the polynomial $\log\,(1-L)^d$ which features the partial derivative of the (Gaussian) log-likelihood function of $\{y_t\}$ under the null hypothesis. As a result, $\left\{x^*_{t-1,d}\right\}$ is core in the construction of LM-based tests under the Gaussian condition. More generally, we can use $\left(\varepsilon_{t,d}, x^*_{t-1,d}\right)'$ to construct valid tests even if data are not normally distributed. The main theoretical results put forward in this paper formally hold under the following set of sufficient conditions.

Assumption 1. *Let $\{\varepsilon_t\}$ in (1) be an autoregressive process, $\mathcal{A}(L)\,\varepsilon_t = v_t$, with $\mathcal{A}(L) = 1-\sum_{j=1}^{p} a_j L^j$, $0 \leq p < \infty$, having all roots outside the unit root circle.*

Assumption 2. *Let $v_t \sim iid\left(0, \sigma^2\right)$, with $v_t = 0$ for all $t \leq 0$, $E\left(|v_t|^r\right) < \infty$ for some $r > 2$, and a cumulative distribution function $F(z)$ with a continuous density, $f\,(z)$, uniformly bounded away from* 0 *and infinity on $\{z : 0 < F\,(z) < 1\}$.*

These conditions are standard in the related literature; see, for instance, Koenker and Xiao (2004) and Galvao (2009). The appendix provides some technical comments on the role played by these assumptions. Given the variables $\left(\varepsilon_{t,d}, x^*_{t-1,d}\right)'$ introduced previously, consider the following auxiliary regression in order to determine whether $\mathrm{H}_0 : \theta = 0$ holds true:

$$\varepsilon_{t,d} = \phi x^*_{t-1,d} + \sum_{j=1}^{p} a_j \varepsilon_{t-j,d} + v_t,\ t = p+1, ..., T, \tag{3}$$

where $\{v_t\}$ is a disturbance term with properties as described in Assumption 2. Breitung and Hassler (2002) and Demetrescu *et al.* (2008) show that the squared t-statistic for $\mathrm{H}_0 : \phi = 0$ given the OLS estimate $\widehat{\phi}$ from (3), denoted LM_{LS} in the sequel, is asymptotically equivalent to the LM test for $\mathrm{H}_0 : \theta = 0$. The LM test was shown to be efficient under Gaussianity by Robinson (1994); see also Tanaka (1999).

Under $\mathrm{H}_0 : \theta = 0$, $x^*_{t-1,d}$ is (asymptotically) stationary and admits the causal representation $x^*_{t-1,d} = \sum_{j=0}^{t-1} \varphi_j v_{t-j-1}$, where $\left\{\varphi_j\right\}_{j\geq 0}$ is a squarely (but not absolutely) summable sequence independent of the value of d, and LM_{LS} is asymptotically distributed as a Chi-squared distribution with one degree of freedom, denoted $\chi^2_{(1)}$. Under the sequence of local alternatives $\theta = c/\sqrt{T}$, $c \neq 0$, the general characterization $\varepsilon_{t,d} = \left(c/\sqrt{T}\right) x^*_{t-1,d+\theta} + \varepsilon_t + o_p(1)$ holds true (Tanaka 1999; Demetrescu *et al.* 2008) and it follows $x^*_{t-1,d} = x^*_{t-1,d+\theta} + \left(c/\sqrt{T}\right) \sum_{j=1}^{t-1} j^{-1} x^*_{t-1-j,d+\theta} + o_p\left(1\right)$, with $x^*_{t-1,d+\theta} = \sum_{j=0}^{t-1} \varphi_j v_{t-j-1}$. As a result, $\phi = c/\sqrt{T}$ captures the extent and the direction of the departure from the null, with positive (negative) values of ϕ indicating larger (smaller) orders of integration than d. Hence, testing the fractional integration hypothesis on the basis of (3) ensures non-trivial power against local alternatives; see Demetrescu *et al.* (2008).

Remark 2.1. The data generating process in (1) can be generalized to allow for a non-zero constant term using the characterization $y_t = \mu + (1-L)^{-d-\theta} \varepsilon_t$ with $\mu \neq 0$. Robinson (1994) discusses a valid procedure to consistently estimate μ independently of d in this context; see also Demetrescu *et al.* (2008, Prop.4). In particular, under H_0, $\Delta^d y_t = \mu \Delta^d + \varepsilon_t$, so μ can be estimated consistently from the linear regression of $\Delta^d y_t$ on the regressor $b_{t,d} = \sum_{j=0}^{t-1} \pi_j\left(d\right)$, $t = 2, ..., T$, with $\pi_j\left(d\right)$ as defined in (2). Under the null hypothesis, the residuals from this auxiliary regression correspond to the process $\{\varepsilon_{t,d}\}$. Hence, for simplicity of notation and with no loss of generality, we can assume $\mu = 0$ in the sequel.

2.2 Quantile regression analysis

Let $\mathcal{F}_t$ be the σ-field generated by $\{v_s, s \leq t\}$ and denote $Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1})$ as the conditional quantile of $\varepsilon_{t,d}$ for some probability τ in $(0,1)$. Then, under (3), it follows that

$$Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1}) = \alpha(\tau) + \phi x^*_{t-1,d} + \sum_{j=1}^{p} a_j \varepsilon_{t-j,d} = \mathbf{z}^{*\prime}_{t-1,d}\boldsymbol{\beta}(\tau) \tag{4}$$

where $\alpha(\tau) = F^{-1}(\tau)$, $\boldsymbol{\beta}(\tau) = (\alpha(\tau), \phi, a_1, ..., a_p)'$, and $\mathbf{z}^*_{t-1,d} = \left(1, x^*_{t-1,d}, \varepsilon_{t-1,d}, ..., \varepsilon_{t-p,d}\right)'$. Although $\alpha(\tau)$ varies with τ, the remaining parameters remain fixed across quantiles, *i.e.*, they are globally identified. Therefore, under $\mathrm{H}_0 : \theta = 0$, the slope coefficient ϕ in (4) equals zero at any quantile $\tau \in (0,1)$. This property allows us to analyze whether $\mathrm{H}_0 : \phi = 0$ holds true at any arbitrary quantile and, more generally, over an arbitrary closed set of quantiles in $(0,1)$. While LS-based tests exploit the statistical information conveyed by the conditional mean, $E(\varepsilon_{t,d}|\mathcal{F}_{t-1})$, the QR analysis allows us to exploit the information in other distributional features, such as the conditional median and, more generally, at different percentiles of the conditional distribution of the data, $Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1})$. This testing strategy may ensure inference exhibiting more robust properties and enhanced power, particularly, in a non-Gaussian context.

2.2.1 Testing for fractional integration at individual quantiles

The estimation of the vector of parameters $\boldsymbol{\beta}(\tau)$ that characterizes the conditional quantile process in (4) involves the optimization problem $\min_{\mathbf{b}(\tau)\in\mathbb{R}^{p+2}} \sum_{t=p+1}^{T} \rho_\tau(\varepsilon_{t,d} - \mathbf{z}^{*\prime}_{t-1,d}\mathbf{b}(\tau))$, where $\rho_\tau(s) = s\left(\tau - \mathbb{I}_{(s<0)}\right)$ is the so-called 'check' function, with $\mathbb{I}_{(\cdot)}$ denoting the indicator function taking value 1 if the argument is true and 0 otherwise; see Koenker and Bassett (1978). Let $\widehat{\boldsymbol{\beta}}(\tau)$ be the resultant vector of estimates. The following Theorem introduces the asymptotic distribution of the scaled estimation error $\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right)$ at a fixed quantile τ under H_0 and, hence, provides the formal basis to construct QRLM tests. We use the standard notation '$\Rightarrow$' and '$\xrightarrow{p}$' to denote weak convergence and convergence in probability, respectively, throughout the following theoretical statements.

Theorem 2.1 *Let $\{y_t\}$ be generated according to (1), with Assumptions 1 and 2 holding true. Then, under the null hypothesis* $\mathrm{H}_0 : \theta = 0$*, it holds for any fixed* $\tau \in (0,1)$ *as* $T \to \infty$ *that:*

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right) \Rightarrow \mathcal{N}\left(0, \tau(1-\tau)\, s^2(\tau)\,\boldsymbol{\Omega}^{-1}\right) \tag{5}$$

where $s(\tau) = [f(\alpha(\tau))]^{-1}$*, and* $\boldsymbol{\Omega}=lim_{T\to\infty}T^{-1}\sum_{t=p+1}^{T}E\left(\mathbf{z}^*_{t-1,d}\mathbf{z}^{*\prime}_{t-1,d}\right)$*, a finite and invertible matrix. Consequently,*

$$LM_{QR}(\tau) = T\left[\frac{\widehat{\phi}(\tau)}{\widehat{\varsigma}(\tau)}\right]^2 \Rightarrow \chi^2_{(1)} \tag{6}$$

where $\widehat{\varsigma}^2(\tau)$ *denotes a consistent estimate of* $\varsigma^2(\tau) = \tau(1-\tau)\, s^2(\tau)\,\omega^2_{22}$*, with* ω^2_{ij} *denoting the* ij*-th element of* $\boldsymbol{\Omega}^{-1}$*.*

Proof. *See the technical appendix.*

Remark 2.2. The asymptotic variance of $\widehat{\phi}(\tau)$ is given by $\varsigma^2(\tau)$. Therefore, given consistent estimates of the elements that characterize the standard error of $\widehat{\phi}(\tau)$, it follows readily from Theorem 2.1 that the null distribution of the t-statistic for $\mathrm{H}_0 : \phi = 0$ in (4), namely, $t_{QR}(\tau) = \sqrt{T}\,\widehat{\phi}(\tau)/\widehat{\varsigma}(\tau)$, is asymptotically standard normal. Consequently, we can construct either a pivotal test for one-sided testing or, in analogy to least-squares LM-testing, a squared t-statistic for the two-sided alternative, $LM_{QR}(\tau) = t^2_{QR}(\tau)$, which is asymptotically distributed as a $\chi^2_{(1)}$ variate.

Remark 2.3. A consistent estimator of ω^2_{22} can readily be obtained from the diagonal of the inverse of the sample matrix $T^{-1}\sum_{t=p+1}^{T}\mathbf{z}^*_{t-1,d}\mathbf{z}^{*\prime}_{t-1,d}$. Following Siddiqui (1960) and Bassett and Koenker (1982), the sparsity-related function $s(\tau)$ can be estimated consistently under Assumptions 1 and 2 as the sample difference quotient $\bar{\mathbf{z}}^{*\prime}_d\left(\widehat{\boldsymbol{\beta}}(\tau+h_T) - \widehat{\boldsymbol{\beta}}(\tau-h_T)\right)/2h_T$, with $\bar{\mathbf{z}}^*_d$ denoting the sample mean of $\mathbf{z}^*_{t-1,d}$, and h_T being a bandwidth parameter tending to zero at a suitable rate as the sample size increases. Alternatively, the covariance matrix of $\widehat{\boldsymbol{\beta}}(\tau)$ can be estimated consistently using kernel-type estimators available in the nonparametric density estimation literature, or resampling methods; see Koenker (2005) for a review.

Remark 2.4. Although the primary purpose of the QRLM test is to conduct inference for a pre-specified value of d, the procedure can also be used to construct confidence intervals that include the true value of d with $100(1-\alpha)\%$ asymptotic coverage probability by inverting $LM_{QR}(\tau)$, as discussed in Hassler *et al.* (2009). In particular, let d_0 be the true value and let $LM_{QR,\delta}(\tau)$ denote the value of the test statistic $LM_{QR}(\tau)$ when testing $\mathrm{H}_0 : d = \delta$, with $\delta \in \Theta$, a closed interval. Define

$\mathcal{D}_{\alpha} = \left\{\delta : \Pr\left[\chi^2_{(1)} \leq LM_{QR,\delta}(\tau)\right] \leq 1-\alpha\right\}$, *i.e.*, the subset of Θ for which the null hypothesis cannot be rejected at the α significance level. It follows that if $\mathcal{D}_{\alpha}$ is in the interior of Θ, then the probability of d_0 being within $\mathcal{D}_{\alpha}$ is at least $(1-\alpha)$. Thus, the confidence interval can be constructed through a grid-search process in Θ. Similarly, confidence intervals can be constructed by identifying the non-rejection region of $t_{QR}(\tau)$ when testing H_0 against a two-sided alternative. We shall implement this technique later in the empirical section.

Remark 2.5. A generalization of the procedure in the presence of unconditional heteroskedasticity is immediate. Following Koenker (2005), it can be shown that $\sqrt{T}\left(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)\right) \to \mathcal{N}(0, \mathbf{V}(\tau))$ if $E\left(v_t^2\right) = \sigma_t^2$, with $\sigma_t^2 < \infty$ uniformly bounded. The asymptotic covariance matrix takes the White-Eicker-Huber form $\mathbf{V}(\tau) = \tau(1-\tau)\Psi_{\tau}\mathbf{\Omega}^{-1}\Psi_{\tau}$, with Ψ_{τ} denoting the limit in probability of $T^{-1}\sum_{t=p+1}^{T} z^{*}_{t-1,d} z^{*\prime}_{t-1,d}\left[f\left(Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1})\right)\right]$; see Powell (1991). This approach is similar in spirit to the (conditional) heteroskedasticity-robust LM test for fractional integration in the LS context discussed in Demetrescu *et al.* (2008) and Hassler *et al.* (2009). We shall implement both approaches in the empirical section.

It is interesting to compare $LM_{QR}(\tau)$ to alternative tests in related literature. First, note that this test is the QR analog of the LS-based test, LM_{LS}, in Breitung and Hassler (2002). While LS-based tests exhibit optimality properties when innovations are normally distributed, they are no longer efficient under departures from such assumption. Consequently, $LM_{QR}(\tau)$ may be better indicated in realistic settings in which innovations exhibit non-Gaussian features, such as excess kurtosis and skewness. Second, $LM_{QR}(\tau)$ can be used to test for fractional integration, a more general hypothesis than the standard unit-root case. Previous literature in QR has focused exclusively on the Dickey-Fuller (DF) testing approach, so QRLM tests considerably generalize this setting.

It is worth discussing the similitudes and differences with the DF test. For ease of exposition, assume that Assumption 1 holds true with $p = 0$. Then, the DF test addresses $\mathrm{H}_0 : d = 1$ by testing $\mathrm{H}_0 : \phi = 0$ in $\Delta y_t = \alpha + \phi y_{t-1} + v_t$, with $y_{t-1} = \sum_{j=1}^{t-1}\varepsilon_{t-j}$. Similarly, $\mathrm{H}_0 : d = 1$ implies $\mathrm{H}_0 : \phi = 0$ in $\varepsilon_{t,d} = \alpha + \phi x^{*}_{t-1,d} + v_t$ in our setting, with $x^{*}_{t-1,d} = \sum_{j=1}^{t-1} j^{-1}\varepsilon_{t-j}$. Therefore, the main methodological difference between these two procedures lies exclusively in the different weights used to construct the right-hand side variable in the auxiliary regression. The harmonic weighting that characterizes our approach ensures power to detect fractional alternatives in the generalized context studied here and, furthermore, has major implications on the asymptotic null distributions of the tests. In particular, note that the QR estimate of ϕ in the DF equation is T-consistent and converges, once adequately scaled, to a mixture of the Dickey-Fuller and a standard normal distribution under Assumptions 1 and 2; see Koenker and Xiao (2004) and Galvao (2009). In contrast, as formally shown in Theorem 2.1, $\widehat{\phi}(\tau)$ is $\sqrt{T}$-consistent and converges to the standard normal distribution upon appropriate scaling. Consequently, the DF approach provides a more powerful tool to test the unit-root hypothesis $\mathrm{H}_0 : d = 1$ as it builds on super-consistent estimates, but it cannot be used to test more general hypotheses.

2.2.2 Testing fractional integration over a set of quantiles

The QR setting provides a natural diagnosis tool to formally analyze whether a model with a constant long memory coefficient fits the data across the quantiles of the conditional distribution. The central idea is to address whether there is sufficient regularity in the data not to jointly reject $\mathrm{H}_0 : \theta = 0$ for a constant value of d over the different quantiles of an arbitrary closed subinterval of $(0, 1)$. The following Theorem states the asymptotic distribution of two different, but related, testing procedures intended for this purpose.

Theorem 2.2. *Let $\mathcal{T} = [\underline{\tau}, \overline{\tau}]$ be a closed subset of $(0,1)$ of length $\Delta = \overline{\tau} - \underline{\tau}$, and consider an equidistant partitioning $\tau_i = \underline{\tau} + i\Delta/T$, $i = 0, 1, ..., T$. Let the random function $S_T(\tau) = \sqrt{\tau(1-\tau)}\sqrt{T}\,\widehat{\phi}(\tau)/\widehat{\varsigma}(\tau)$, with $\sup_{\tau\in\mathcal{T}} |\widehat{\varsigma}(\tau) - \varsigma(\tau)| = o_p(1)$, and $\varsigma^2(\tau) = \tau(1-\tau)\, s^2(\tau)\,\omega_{22}^2$ as in Theorem 2.1. Then, under the assumptions of Theorem 2.1, it follows as $T \to \infty$ that*

$$\mathcal{KS}_{\mathcal{T}} = \max_{1\leq i\leq T} |S_T(\tau_i)| \to \sup_{\tau\in\mathcal{T}} |\mathcal{B}(\tau)| \tag{7}$$

$$\mathcal{CM}_{\mathcal{T}} = \sum_{1\leq i\leq T} S_T^2(\tau_i)(\tau_i - \tau_{i-1}) \to \int_{\tau\in\mathcal{T}} \mathcal{B}^2(\tau)\, d\tau, \tag{8}$$

where $\mathcal{B}(\tau)$ is a standard Brownian Bridge.

Proof. *See the technical appendix.*

Remark 2.6 The limit distributions in (7) and (8) are truncated versions of the well-known Kolmogorov-Smirnov and Cramér-von Mises distributions, respectively, which would arise when evaluating the supremum or the integral over the $[0, 1]$ interval. This serves as motivation to call the test statistics accordingly. The limit distributions of $\mathcal{KS}_{\mathcal{T}}$ and $\mathcal{CM}_{\mathcal{T}}$ are free of nuisance parameters and, therefore, it is straightforward to obtain critical values by direct simulation of the functionals involved given the arbitrary choice of $\mathcal{T}$. For instance, for $\mathcal{T} = [0.1, 0.9]$, the 95% critical values of $\mathcal{KS}_{\mathcal{T}}$ and $\mathcal{CM}_{\mathcal{T}}$ are 1.35 and 0.44, respectively.

3 Finite sample analysis

In this section, we evaluate experimentally the small sample properties of the individual and joint-quantile QRLM test statistics previously introduced. Following Koenker and Xiao (2004) and Galvao (2009), we first analyze the empirical size and power of the individual QRLM test $LM_{QR}(\tau)$ at $\tau = 1/2$ (conditional median). To this end, we consider a data generating process (DGP) characterized by $(1-aL)(1-L)^{d+\theta}y_t = \varepsilon_t,\ \ t = 1, ..., T$, with $a \in \{0, 0.5, 0.75\}$, $T \in \{100, 250, 1000\}$, and $\theta \in [-0.3, 0.3]$. Since the unit-root is the leading case studied in the QR literature, we set the (unknown) value of d to one, and test $\mathrm{H}_0 : d = 1$ given the different values of θ. Consequently, the frequencies of rejection for $\theta = 0$ are informative of the empirical size, whereas the cases $\theta \neq 0$ allow us to characterize the power function. As shown in the theoretical section, the null distribution of $LM_{QR}(\tau)$ does not depend on the particular value of d, while power is mainly dictated by the magnitude of θ, as discussed previously. We stress, therefore, that this particular choice $d = 1$ does not imply any loss of generality of our analysis.

According to Assumption 2, we assume that $\{\varepsilon_t\}$ are i.i.d. innovations drawn from a Student-t distribution with $v \in \{2, 3, 1000\}$ degrees of freedom. The case $v = 1000$ corresponds to the Gaussian distribution, whereas the 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. Since the Student-t distribution is continuous in the degrees of freedom parameter and verifies $E\left(|\varepsilon_t|^{v+\epsilon}\right) < \infty$ for any arbitrarily small $\epsilon > 0$, we can think of $v = 2$ as a 'limiting' case corresponding to the formal bound of our theory. This is a usual approach in the robust literature.

The QRLM test is implemented using a Gaussian kernel-based estimator of the covariance matrix. Simulations are conducted considering 5000 replications and the usual 5% significance level. To evaluate the relative behavior of the QRLM test, we also address the performance of the LS-based test, LM_{LS}, introduced by Breitung and Hassler (2002). This test is efficient in the Gaussian context and, like the QRLM test, is formally valid for $v > 2$. Table 1 presents the rejection frequencies of the these tests under the different DGPs considered, focusing on the simple i.i.d. case ($a = 0$) and on a more general context in which errors exhibit stationary first-order autoregressive dependence ($a = 0.75$). For conciseness, we omit the results for $a = 0.5$, but note that these results are available upon request.

[Insert Table 1 around here]

We first discuss the main results for the i.i.d. case ($a = 0$). Both the LM_{LS} and $LM_{QR}(\tau)$ test tend to show empirical sizes close to the 5% nominal level even in the small sample $T = 100$, and increasing power on $|\theta|$ and/or T, thereby showing the consistency of both procedures. As expected, the Gaussian environment provides the necessary conditions for optimality of LM_{LS}, which outperforms in terms of finite-sample size and power. Since LM_{LS} is based on LS estimates, which involves matrix algebra, it tends to present more stable size in small samples than $LM_{QR}(\tau)$, which involves numerical optimization. On the other hand, when innovations are drawn from heavy-tailed distributions, LM_{LS} tends to be undersized as kurtosis (directly related to the degrees of freedom parameter) increases. Interestingly, a considerable degree of leptokurtosis is necessary to generate sizeable departures in this test. The differences between $v = 1000$ and $v = 3$ are not particularly dramatic in terms of size-power distortions in any of the sample lengths analyzed, so the LM_{LS} test seems to exhibit considerable resilience against heavy-tailed distributions.

Nevertheless, even if heavy-tailed distributions do not cause massive distortions, LS-based tests are no longer optimal, and alternative testing approaches may lead to more efficient results. Indeed, the QRLM tests can yield important gains in power with respect to the LS-based test while still ensuring approximately correct size in finite samples. In particular, for $v = 2$, $LM_{QR}(\tau)$ suffers a similar undersizing effect as LM_{LS}, but on the other hand exhibits power which is roughly twice as large as that of the LS test, and even larger for certain DGP configurations. For instance, for $v = 2$ and $\theta = -0.1$, the power of LM_{LS} in the sample with $T = 100$ ($T = 250$) observations is approximately 17.10% (46.16%), whereas $LM_{QR}(\tau)$ presents rejection frequencies of 46.2% (87.12%), respectively. The empirical size of LM_{LS} and $LM_{QR}(\tau)$ is similar in both scenarios (3.62% and 3.68% for $T = 250$, respectively), so the sheer difference in power in this highly non-Gaussian environment cannot merely be attributed to differences in empirical size. 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 approaches 2, for which power shows similar patterns around the origin.

We now turn our attention to the context in which the DGP exhibits autoregressive short-run dependence with $a = 0.75$. In this case, both the QRLM and LS tests are computed using an auxiliary regression augmented with one lag of the dependent variable. As in the i.i.d. case, the tests show empirical sizes that approach the nominal level as the sample size increases, and consistency to reject the null under the sequence of alternatives analyzed. Table 1 shows that, 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 be conservative 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. As expected, 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. These asymmetric patterns in power as a function of θ are data-dependent and, hence, different conclusions may arise when considering different parameter configurations. For instance, for $a = 0.5$ (results not reported), the power of both tests is characterized by a strong asymmetric pattern, but in this case the alternatives $\theta < 0$ are easier detected than their counterparts $\theta > 0$, a pattern which was already noted by Demetrescu *et al.* (2008) in the LS context.

In addition to the individual QRLM test, we analyze the consistency of the joint QRLM test to reject a false long-memory model. To this end, we consider a DGP characterized by $(1-L)^{\overline{d}+\eta_t} y_t = \varepsilon_t$, $t = 1, ..., T$, $T \in \{100, 1000\}$, and ε_t being driven by the Student-t process with i.i.d. observations and $v \in \{2, 3, 1000\}$ degrees of freedom, as discussed previously. In this model, $\eta_t \sim iid\mathcal{N}\left(0, \gamma^2\right)$ is a noise term independent of ε_t, with γ controlling the variability of the process. For $\gamma = 0$, $\{y_t\}$ is FI($\overline{d}$), however, for $\gamma > 0$, $\{y_t\}$ is generated from a model with time-varying, random long-memory coefficients $d_t = \overline{d} + \eta_t$ centered around $\overline{d}$, with γ controlling the extent of dispersion around this coefficient. In this case, the observable process is not a fractionally integrated model as defined in Section 2 and, consequently, the null hypothesis $\mathrm{H}_0 : d = \overline{d}$ should be rejected. We set $\overline{d} = 1/2$ in this experiment and analyze the average rejection frequencies of the $\mathcal{KS}$ test computed over the percentiles in the intervals $\mathcal{T}_1 = [0.4, 0.6]$ and $\mathcal{T}_2 = [0.1, 0.9]$, with γ taking values in $[0, 0.3]$. For $\gamma = 0$ the experiment provides information about the empirical size of the joint tests, while values $\gamma > 0$ allow us to characterize their power in finite samples.

[Insert Table 2 around here]

Table 2 reports the main results of this experiment. For $\gamma = 0$, the $\mathcal{KS}$ test shows some size distortions in the small sample, which nevertheless tends to disappear as the sample length increases. Consequently, using the joint test to construct a confidence interval for the fixed coefficient $\overline{d} = 1/2$, following the strategy discussed in Remark 2.4 will deliver intervals with correct coverage ratio in large samples under the set of assumptions considered. For values $\gamma > 0$, the test rejects the false null with an increasing probability on γ and/or the sample length T. Consequently, the $\mathcal{KS}$ test can consistently reject the suitability of a fractionally integrated model. As in the case of the individual QRLM test, the power function of the joint test is affected by sample-dependent aspects related to the distribution of the data, and tends to be greater under heavy-tailed distributions.

In addition to the previous experiments, we further analyzed several generalizations. For the QRLM tests at $\tau = 1/2$, we analyzed the finite-sample statistical properties using random innovations drawn from Hansen's Skewed-t distribution. This distribution is characterized by two parameters that control for excess kurtosis and skewness. The set of sufficient conditions in Assumption 2 imposes only mild restrictions on the distribution of the data that do not require symmetry, a restriction often required in the robust literature. Consequently, the main conclusions when permitting skewness in addition to excess kurtosis are completely similar to those reported previously and show that the QR-test leads to more efficient testing than LS-based tests when innovations are drawn from heavy-tailed and/or skewed distributions. These results are omitted for saving space, but available upon request. Similarly, we analyzed the performance of the individual tests building on covariance-matrix estimates based on resampling methods, as suggested in Koenker (2005), which showed a similar performance. These results are available in the working-paper version of the paper. Finally, we analyzed the power of the joint test to detect alternatives characterized by other mechanisms that generate time-varying values of the long-memory coefficient. In particular, we focused on values of d_t generated from (i) a random sample of uniform observations in $[0, 1]$; (ii) regime-switching models; (iii) stationary AR(1)-type dynamics, and (iv) pure random-walk based dynamics. The main evidence from this analysis is qualitatively similar to that discussed previously and shows the general consistency of the joint test. Those results are not presented, but available upon request.

In summary, the main picture that emerges from the Monte Carlo analysis is that both the individual and joint QRLM tests proposed in this paper are well-suited for empirical purposes. They tend to exhibit approximately correct size even in small samples, with finite-sample departures from the nominal level that vanish as the sample is allowed to grow. The QR-based tests can consistently reject the null hypothesis under the alternative and, as discussed in the related literature, can improve power performance over LS-based alternatives when the data is driven by non-Gaussian distributions with large excess kurtosis and/or skewness. We shall use these procedures to address the existence of long-memory patterns in realized volatility in the following section.

4 Long-run dependence in realized stock volatility

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 based on sums of intraday returns; see, among others, Andersen *et al.* (2001, 2003). Our interest in this process is motivated by the well-established fact that realized volatility exhibits persistent autocorrelations that are consistent with long memory dynamics. There is an ongoing debate about the sources of long-range dependence in this literature, since at the theoretical level there is little consensus on the mechanism generating this phenomenon.

4.1 Data and preliminary evidence of long memory

We record continuously-compounded returns of the IBM stock sampled regularly over 5-minute intervals from 9.30 a.m. to 4.00 p.m., totalling 78 intraday observations. The initial sample we analyze covers the period from 04/01/1993 to 31/05/2007, totalling $3,630$ trading days. We shall focus on an extended sample, including the financial crisis period, as a robustness check. Using these series, we compute two basic measures of realized variation. First, we consider daily realized volatility, defined as the square root of the sum of squared 5-minute log-returns over the day, $\sigma_{RV}(t) = \left[\sum_{n=1}^{78} r_{(n),t}^2\right]^{1/2}$. 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, $\sigma_{RPV}(t) = \sum_{n=1}^{78} |r_{(n),t}|$. Under certain conditions, this measure is robust to jumps; see Barndorff-Nielsen and Shephard (2004) for details. As customary in this literature, we analyze the long-run properties of realized variation focusing on the logarithmic transformation of these variables, *i.e.*, log-realized volatility, $\log \sigma_{RV}(t)$, and log-realized power variation, $\log \sigma_{RPV}(t)$; see Andersen *et al.* (2003). Figure 1 presents plots of these measures over the sample period considered.

[Insert Figure 1 around here]

Table 3 provides a summary of the usual descriptive statistics of these series. Daily realized volatility $\sigma_{RV}(t)$ typically exhibits a considerable degree of leptokurtosis and right skewness due to the influence of the jump component in the DGP 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. On the other hand, the logarithmic transformation removes partially the influence of outliers and considerably reduces the leptokurtosis of the resulting variables. The series $\log \sigma_{RV}(t)$ and $\log \sigma_{RPV}(t)$ exhibit kurtosis coefficients of 5.27 and 3.44, respectively. Since skewness is considerably attenuated as well, the log measures of realized variation exhibit an unconditional distribution closer to the Gaussian distribution. However, normality is still rejected on the basis of standard testing procedures, such as the Jarque-Bera test, owing mainly to excess kurtosis.

[Insert Table 3 around here]

The most important stylized feature of realized variation measures is a slowly decaying autocorrelation pattern, 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. Even though the first-order correlation of $\sigma_{RV}(t)$ is not particularly sizeable (0.276), the subsequent correlations remain highly significant. Distant observations, which span almost two years of trading days, remain positively correlated; see Table 3 for further details. A similar pattern appears in the other measures of daily variation, although the first-order correlation tends to be larger for these series.

Previous literature argues that this pattern of temporal dependence is caused by a fractionally integrated process with long-memory coefficient $0 < d < 1$; see, among others, Andersen *et al.* (2001, 2003). Together with the usual descriptive statistics, Table 3 reports point estimates of the long-memory parameter obtained with semi-parametric estimators in the frequency domain. In particular, the table reports the Geweke and Porter-Hudak (1984) estimator, $\widehat{d}_{GPH}$, and the exact local Whittle estimator suggested by Shimotsu and Phillips (2005), $\widehat{d}_{ELW}$. According to standard practice, we compute these estimates using a bandwidth parameter $m_T = [T^{0.6}]$, with $[\cdot]$ denoting the integer argument. Table 3 also reports the 95% confidence intervals for d based on the asymptotic distribution of these estimators, noting that $\sqrt{m_T}\left(\widehat{d}_{GPH} - d\right) \Rightarrow \mathcal{N}\left(0, \pi^2/24\right)$ and $\sqrt{m_T}\left(\widehat{d}_{ELW} - d\right) \Rightarrow \mathcal{N}\left(0, 1/4\right)$. According to these estimates, realized measures in levels tend to exhibit smaller values of d than their logarithmic counterparts, which is consistent with biases originated by excess kurtosis; see Haldrup and Nielsen (2007). Although all estimates are significantly smaller than one, all confidence intervals, except for one, include values of d in the nonstationarity region ($d \geq 0.5$).

[Insert Figure 2 around here]

4.2 Quantile regression analysis

Consistent with previous literature, the preliminary analysis reported in Table 3 suggests the existence of long memory patterns in the different measures of realized variation. The QRLM tests introduced in this paper can shed further light on the empirical properties of these series. The QR approach can naturally be motivated in this context by two main considerations. First, the non-Gaussian features of daily volatility discussed previously may cause biases in the semi-parametric estimates of the long-memory coefficient, as discussed by Haldrup and Nielsen (2007). Hence, alternative procedures which exhibit more robust properties against influential observations may produce more accurate inference in this context. Consequently, the QR testing at the 50th percentile seems a natural approach. Second, the QR methodology allows us to analyze the general suitability of a fractionally integrated filter with constant parameter d. If the realized volatility is truly generated by such a process, we should be able to identify its characteristic coefficient along the different quantiles of the conditional distribution of the series. Hence, the QR analysis allows us to address a more robust discussion on the properties of realized volatility and bring completely new empirical evidence to this field.

The empirical analysis is carried out in the following terms. At any of the quantiles $\tau \in \mathcal{Q}$, $\mathcal{Q} = \{0.1, 0.11, ..., 0.9\}$, we run the quantile regression (4) and compute the QRLM t-statistic $t_{QR}(\tau)$ for $\mathrm{H}_0 : \phi = 0$ in (4) given $\mathrm{H}_0 : d = \delta$, with $\delta \in \mathcal{D}$, $\mathcal{D} = \{0, 0.01, ..., 1\}$. This analysis attempts to provide a detailed examination of the existence of long-memory patterns across the quantiles of the conditional distribution for different long-memory coefficients. Following standard practice, we exclude top and bottom quantiles from $\mathcal{Q}$ owing to statistical difficulties of the QR methodology in accurately dealing with inference at extreme quantiles. We compute t-statistics rather than squared t-statistics because the sign of $\widehat{\phi}(\tau)$ is informative about the sign of θ. More specifically, we compute $t_{QR}(\tau)$ after accounting for the likely possibility of a non-zero constant mean in the series, using the procedure described in Remark 2.1. Given the resultant series, the auxiliary regression is augmented with p lags of the dependent variable according to the rule $p = \left[4(T/100)^{1/4}\right]$. 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 $\widehat{\phi}(\tau)$ is computed based on the estimator proposed by Powell (1991) seeking to obtain robustness against potential heteroskedasticity in the data, as discussed in Remark 2.5. We use a Gaussian kernel to estimate the density of the data with deterministic bandwidth parameter, h_T, set according to the rule $\left[0.3 \min\{\widehat{\sigma}_u, IQR(\widehat{v}_t)/1.34\} T^{-1/5}\right]$, where $IQR(\cdot)$ denotes the interquartile range. For completeness of analysis, we also compute the LS-based LM test with a covariance matrix estimate robust to unknown (conditional) heteroskedasticity as in Demetrescu *et al.* (2008), using the same augmentation criterion described previously.

As underlined in Remark 2.4, the QRLM testing procedure allows us to construct confidence intervals for d which are characterized by the non-rejection region of the tests involved. Thus, at any $\tau \in \mathcal{Q}$, we construct 95% and 99% confidence intervals for d, denoted as $\mathcal{CI}_{95\%}(d|\tau)$ and $\mathcal{CI}_{99\%}(d|\tau)$, respectively. In addition, we compute the $\mathcal{KS}$ and the $\mathcal{CM}$ tests in Theorem 2.2 to analyze whether $\mathrm{H}_0 : d = \delta$, $\delta \in \mathcal{D}$, 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, $\mathcal{T}_2$ focuses on the whole distribution after excluding top and bottom quantiles.

4.2.1 Long memory in logarithms of realized volatility

For simplicity of exposition and owing to space constraints, we focus on the log realized volatility series (*e.g.*, Andersen *et al.* 2003), noting that a complete discussion is available in the accompanying working paper; see Hassler *et al.* (2012). Tables 4 and 5 present the main results from the QRLM analysis for $\log\sigma_{RV}(t)$ and $\log\sigma_{RPV}(t)$, respectively. In these tables, Panel A reports the values of $t_{QR}(\tau)$ on the subset $\mathrm{H}_0 : d = \delta_i,\ \delta_i \in \{0, 0.1, ..., 1\}$ and deciles $\tau \in \{0, 1, 0.2, ..., 0.9\}$ for ease of exposition, as well as the confidence intervals $\mathcal{CI}_{95\%}(d|\tau)$ and $\mathcal{CI}_{99\%}(d|\tau)$, the (conditional) heteroskedasticity-robust LM_{LS} statistic of Demetrescu *et al.* (2008), and the confidence-interval estimates for d based on this test. Similarly, Panel B reports the $\mathcal{KS}_{\mathcal{T}}$ and $\mathcal{CM}_{\mathcal{T}}$ test statistics for joint hypotheses as well as 95% and 99% confidence intervals for d based on the non-rejection region of these tests.

[Insert Tables 4 and 5 around here]

To save space, and since results are entirely similar, we discuss the main evidence for the analysis on $\log\sigma_{RV}(t)$, referring the interested reader to the companion paper for an extended discussion; see Hassler *et al.* (2012). Some features are worth discussing in detail. First, the QR analysis at the conditional median ($\tau = 1/2$) strongly rejects the hypothesis that daily log realized volatility is purely driven by either stationary short-run dynamics ($d = 0$) or a unit root process ($d = 1$). In particular, the 95% confidence interval for d at this quantile is $[0.38, 0.53]$, mostly supporting the existence of (stationary) long-range dependence. Similarly, the LS-based test LM_{LS} rejects both FI(0) and FI(1) dynamics, although it yields slightly larger confidence intervals for the long-memory coefficient which are shifted to the right. For instance, the 95% confidence interval is $[0.45, 0.60]$. Note that these estimates are remarkably similar to the frequency domain-based semiparametric estimates reported in Table 3.

Second, the analysis of $t_{QR}(\tau)$ computed along different quantiles allows to reach two main conclusions. On the one hand, the individual QRLM tests always find strong statistical evidence of long-range dependence, since the confidence intervals for d always include values strictly greater than zero and smaller than one at any $\tau \in \mathcal{Q}$. On the other hand, there is an upward trend in the confidence intervals such that both their central value and their amplitude tend to increase with τ. This phenomenon is clearly visible in Figure 3, which shows the central values $d_{central} : \arg\inf_d |t_{QR}(\tau)|$ for which $t_{QR}(\tau)$

is 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 95% and 99% confidence bands. The upward trend is hardly noticeable for quantiles below the median since the confidence intervals tend to be very stable. For instance, $\mathcal{CI}_{95\%}(d|\tau) = [0.31, 0.41]$ at $\tau = 0.1$, just slightly smaller than that at $\tau = 1/2$ indicated above. While this pattern still holds for quantiles at the central deciles, a stronger diverging effect is evident for percentiles belonging to the upper quartile. For instance, $\mathcal{CI}_{95\%}(d|\tau) = [0.56, 0.83]$ at $\tau = 0.9$, so most of these values are above the upper limit of $\mathcal{CI}_{95\%}(d\,|\tau = 1/2)$; see Table 4. 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. Note, for instance, that the size of $\mathcal{CI}_{95\%}(d|\tau)$ at $\tau = 9/10$ is almost three times larger than that at $\tau = 1/10$.

[Insert Figure 3 around here]

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 realized 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. For quantiles in the $\mathcal{T}_1$ central area, the $\mathcal{KS}$ and $\mathcal{CM}$ tests formally show that values around $d = 0.4$ cannot be rejected. Similarly, for quantiles in the extended range $\mathcal{T}_2 = [0.1, 0.9]$, none of the joint tests can reject the null hypothesis of a constant long-memory parameter model for values of d around that level; see Tables 4 and 5, Panel B, for details.

Paralleling the strategy used in the individual quantile analysis, we can construct confidence intervals for the globally admissible value of d by identifying the non-rejection region of the $\mathcal{KS}$ and $\mathcal{CM}$ tests. The resulting confidence intervals, denoted generally as $\mathcal{CI}_{100(1-\alpha)\%}(d|\mathcal{T})$ with $\mathcal{T}$ representing either $\mathcal{T}_1$ or $\mathcal{T}_2$, respectively, are reported in Panel B. According to these estimates, the $\mathcal{CI}_{95\%}(d|\mathcal{T}_2)$ for d given by the $\mathcal{KS}$ and $\mathcal{CM}$ tests are $[0.44, 0.45]$ and $[0.41, 0.42]$, 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) refer to this as the "typical value" in their study. This estimate suggests that the long-run component of realized volatility is driven by a strongly persistent, yet stationary, process. Similar results hold for the measure of log-realized power variation.

In contrast to log-realized measures, the analysis of the variables in levels (not reported here), strongly rejects the existence of a fractionally integrated model with constant parameter. Stated more precisely, filtering either $\sigma_{RV}(t)$ or $\sigma_{RPV}(t)$ with a fractionally integrated model does not suffice to render stationary ARMA-type innovations uniformly over the quantiles of the distribution. This feature is not surprising, because neither a fractionally-integrated nor a stationary short-run model can simultaneously deal with the regimes of sudden volatility bursts (driven by the jump component of returns) and low or normal volatility which characterize the dynamics of these variables and which are perfectly visible in Figure 1. The logarithmic transformation reduces the heterogeneity of the series and enables the identification of the long-memory coefficient in the QR setting.

The main results from the QR analysis suggest that a fractionally integrated model with a constant long-memory coefficient in $(0.4, 0.5)$ drives the long-term component of log realized volatility and power variation, rendering short-run innovations stationary. This evidence 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 $\log \sigma_{RPW}(t)$ suggest that these series may be generated by a time-series process with characteristics different from those that drive the remaining observations. The critical point from our analysis, however, is that there is sufficient regularity for the diagnosis-type analysis based on the QRLM tests to conclude that the sample is driven by a fractionally integrated model with constant long-memory parameter. Consequently, a remarkably robust picture emerges from the overall analysis carried out in this paper, since these results essentially agree with the results based on least-squares LM testing and frequency-domain semi-parametric estimation. The overall evidence from this paper, therefore, provides support to the stylized feature of long-range dependence of realized volatility discussed in previous literature.

4.3 Robustness checks

A) Conditional heteroskedasticity

Daily observations of financial variables characteristically exhibit time-varying conditional heteroskedasticity. The conditional variance of log-realized volatility is time-varying and can be theoretically related to the so-called quarticity, a process that can be estimated consistently by integrating fourth-order powers of intraday returns over the day; see Corsi *et al.* (2008). However, dealing with conditional heteroskedasticity in quantile-regressions is far from trivial owing to theoretical difficulties. In particular, the volatility process must be specified parametrically (so inference becomes model-dependent), and the parameters that characterize the conditional mean and volatility processes jointly estimated in a non-linear quantile regression, since they both feature the conditional quantile. In this context, the asymptotic covariance matrix of the parameter estimates cannot be shown to be well-defined uniformly on the parameter space, so the problem of how to conduct inference in this generalized context is far from trivial and, to the best of our knowledge, has not been solved yet.

Owing to these difficulties, the papers that deal with QR analysis on financial series either do not approach time-varying second-order moments (*e.g.*, Gaglianone *et al.* 2011) or use a suitable proxy of this latent process in a location-scale model; see, for instance, Adrian and Brunnermeier (2011) and López-Espinosa *et al.* (2012) for recent examples. The testing strategy adopted in section 4.2 corresponds to the former approach. Nevertheless, since the evidence from the QR analysis in this paper is complemented with other test statistics, we can argue that the main conclusions from this analysis are not affected or driven by (neglected) conditional heteroskedasticity. In particular, note that the LM test in Demetrescu *et al.* (2008) builds on a White-Eicker-Huber estimator of the asymptotic covariance matrix which ensures robustness against unknown conditional heteroskedasticity. The analysis in section 4.2 shows that the individual QRLM confidence intervals for d at $\tau = 1/2$, namely, $\mathcal{CI}_{100(1-\alpha)\%}\left(d|\tau = 1/2\right)$, are not markedly different from the confidence intervals based on this procedure. As discussed in the Monte Carlo section, the QR- and LS-based tests for long memory at the conditional median and mean can exhibit similar properties in large samples provided mild departures from the Gaussian assumption. The important point to note here, however, is that the similitude between the QR-based estimates and the estimates of a procedure robust to conditional heteroskedasticity implies that $\mathcal{CI}_{100(1-\alpha)\%}\left(d|\tau = 1/2\right)$ cannot be largely affected by neglected conditional heteroskedasticity. Furthermore, since the joint confidence intervals $\mathcal{CI}_{100(1-\alpha)\%}\left(d|\mathcal{T}\right)$, $\mathcal{T} \in \{\mathcal{T}_1, \mathcal{T}_2\}$, are necessarily subsets of $\mathcal{CI}_{100(1-\alpha)\%}\left(d|\tau = 1/2\right)$ if d holds constant across quantiles, we can realistically argue that neglected conditional heteroskedasticity is unlikely to have a major qualitative influence on the conclusions discussed previously.

As an additional robustness check, we adopt a location-scale modelling approach, augmenting the conditional quantile equation (4) with an observable proxy to which the volatility of the dependent variable is likely related, namely, $Q_{\varepsilon_{t,d}}(\tau|\mathcal{F}_{t-1}) = \mathbf{z}^{*\prime}_{t-1,d}\boldsymbol{\beta}\left(\tau\right) + \gamma q_t$, with $q_t = \log\left[\sum_{n=1}^{78} r^4_{(n),t}\right]$. The logarithmic transformation is used to smooth the outliers that arise from raising returns to the fourth-order power. Whereas this approach may not be undisputable, it allows us to shed further light on the robustness of the main conclusions obtained previously. A more precise analysis constitutes perhaps an interesting topic for future research.

[Insert Figure 4 around here]

For simplicity, we summarize the main outcomes of this analysis graphically focusing on the $\log \sigma_{RV}(t)$ time series. To this end, Figure 4 shows the $\mathcal{KS}_{\mathcal{T}_2}$ values for given values $d \in \mathcal{D}$ reported in the previous section, which are confronted with the corresponding values of the $\mathcal{KS}_{\mathcal{T}_2}$ test statistic computed from the location-scale model. Figure 4 also shows the asymptotic critical values of the test at the 95% and 99% confidence level in continuous and dotted horizontal lines, respectively. Note, for instance, that the values of d for which $\mathcal{KS}_{\mathcal{T}_2}$ is below those critical values define the non-rejection region of $\mathcal{KS}_{\mathcal{T}_2}$ and, hence, feature the confidence intervals for d based on this testing approach, *i.e.*, $\mathcal{CI}_{95\%}(d|\mathcal{T}_2)$ and $\mathcal{CI}_{99\%}(d|\mathcal{T}_2)$. The analysis shows that the $\mathcal{KS}_{\mathcal{T}_2}$ test statistics in the location-scale quantile regression take values below those in the non-augmented setting. The most noticeable effect is that the confidence intervals for d widen, particularly in the upper tail, and now include values in the nonstationary region; for instance, $\mathcal{CI}_{95\%}(d|\mathcal{T}_2) = [0.39, 0.71]$. Larger confidence intervals are the consequence of greater parameter uncertainty, which is not particularly surprising in view that we are using a crude volatility proxy in the approach. Nevertheless, $\mathcal{CI}_{95\%}(d|\mathcal{T}_2)$ is now even closer to the 95% confidence interval obtained by heteroskedasticity-robust LS-based LM test, namely, $[0.45, 0.60]$ (see Table 1), suggesting that the proxy variable used in this analysis is not unreasonable.

In spite of these differences, the main conclusions drawn in section 4.2 remain unaltered, as expected from the previous considerations. Whereas both the FI(0) and FI(1) hypotheses are strongly rejected in the location-scale model, the QRLM tests support the existence of fractional integration dynamics with a constant value of d fitting the data along the different quantiles of the conditional distribution. Although the interval now includes a wider set of values in the nonstationary region, it is a remarkable fact that the minimum value of $\mathcal{KS}_{\mathcal{T}_2}$ (*i.e.*, the value which provides maximum evidence for the null) is the the same in both approaches, namely, $d = 0.45$. Consequently, the conclusions drawn initially do not seem to be affected in a major way by the existence of time-varying conditional heteroskedasticity.

B) Long memory and the financial crisis

We extend the sample analyzed previously to span the period from 04/01/1993 to 31/05/2012, a total of 4,891 daily observations. This period includes the financial crisis, in which market volatility increased considerably over the average level of previous years as a consequence of the global uncertainty. This issue suggests the potential existence of structural breaks or other forms of nonlinear effects associated to this event in the dynamics of realized volatility measures. The main conclusions of the analysis in section 4.2, being based on a sample that ends before the crisis, are not spuriously driven by neglected breaks or nonlinearities related to this episode. By analyzing an extended sample including this period, we can address the robustness of these results and, furthermore, characterize the effects of the financial crisis on the estimates of the long-memory coefficient.

[Insert Figure 5 around here]

Figure 5 confronts the values of the $\mathcal{KS}_{\mathcal{T}_2}$ test statistic computed over the sub-sample 04/01/1993 to 31/05/2007, analyzed in Section 4.2, to the corresponding values of this test computed over the extended sample. The heterogeneity associated to the episode of the financial crisis increases parameter uncertainty in the total sample, widening the amplitude of the confidence intervals $\mathcal{CI}_{100(1-\lambda)\%}\left(d|\mathcal{T}\right)$ for d. Furthermore, it generates stronger evidence of nonstationary dynamics, shifting $\mathcal{CI}_{100(1-\lambda)\%}\left(d|\mathcal{T}\right)$ to the right. The global confidence intervals for d are now $\mathcal{CI}_{95\%}\left(d|\mathcal{T}_2\right) = [0.45, 0, 63]$ and $\mathcal{CI}_{99\%}\left(d|\mathcal{T}_2\right) = [0.42, 0.66]$, and the smallest value of the $\mathcal{KS}_{\mathcal{T}_2}$ test statistic is achieved at $d = 0.52$. This evidence largely agrees with the results of the LS-based LM test computed over the extended sample, which generates a 95% confidence interval for d given by $[0.51, 0.68]$, and the results of the semi-parametric estimates in the frequency domain, which now generate larger point-estimates for d, namely, $\widehat{d}_{GPH} = 0.66$ and $\widehat{d}_{ELW} = 0.60$, with right-shifted 95% asymptotic confidence intervals given by $[0.56, 0.76]$ and $[0.52, 0.68]$, respectively.

This evidence suggests that including the financial crisis period in the sample leads to greater evidence of nonstationary dynamics according to any of the different testing procedures considered. This feature is consistent with the inclusion of heterogenous data in the sample being generated by a different generating process. Nevertheless, as in the pre-crisis sample, the QRLM analysis and the alternative testing procedures for long memory largely reject both the FI(0) and FI(1) hypotheses, but cannot reject that $\log \sigma_{RV}(t)$ is driven by a long memory model. Furthermore, and in spite of greater global evidence of nonstationary long-range dependence, the QRLM analysis still cannot reject that a stationary long-memory model with constant coefficient slightly greater than $d = 0.4$ underlies the total sample analyzed.

5 Concluding remarks

In this paper, quantile regression based tests against integer or fractional integration at different quantiles have been introduced and discussed. The theory provided 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.

A distinctive property of the LM-type statistics proposed in this paper is that their null distributions converge to a standard normal distribution or simple transformations of this, such as a Chi-squared distribution. 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 necessarily provide 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 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 and their logarithmic transforms. 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 a constant fractional integration order 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 us to conclude that long-memory is a feature of realized volatility time series.

References

[1] Adrian, T., and Brunnermeier, M. (2011), "CoVaR," Princeton University Working Paper.

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

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

[4] Barndorff-Nielsen O. E., and Shephard, N. (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics* 2, 1-37.

[5] Bassett, G., and Koenker, R. (1982), "An Empirical Quantile Function for Linear Models with IID Errors," *Journal of the American Statistical Association* 77, 407-415.

[6] Breitung, J., and Hassler, U. (2002), "Inference on the Cointegration Rank in Fractionally Integrated Processes," *Journal of Econometrics* 110, 167-185.

[7] Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008), "The Volatility of Realized Volatility," *Econometric Reviews* 27, 46-78.

[8] Demetrescu, M., Hassler, U., and Kuzin, V. (2011), "Pitfalls of Post-Model-Selection Testing: Experimental Quantification," *Empirical Economics* 40, 359-372.

[9] Demetrescu, M., Kuzin, V., and Hassler, U. (2008), "Long Memory Testing in the Time Domain," *Econometric Theory* 24, 176-215.

[10] Engle, R.F., and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business and Economic Statistics* 22, 367-381.

[11] Gaglianone, W.P., Lima, L.R., Linton, O., and Smith, D.R. (2011), "Evaluating Value-at-Risk Models via Quantile Regression," *Journal of Business and Economic Statistics* 29, 150-160.

[12] Galvao, Jr. A. F. (2009), "Unit Root Quantile Autoregression Testing Using Covariates," *Journal of Econometrics* 152, 165-178.

[13] Geweke, J., and Porter-Hudak, S. (1983), "The Estimation and Application of Long Memory Time Series Model," *Journal of Time Series Analysis* 4, 221 238

[14] Gutenbrunner, C., and Jurečková, J. (1992), "Regression Rank Scores and Regression Quantiles," *Annals of Statistics* 20, 305-330.

[15] Haldrup, N., and Nielsen, M. Ø. (2007), "Estimation of Fractional Integration in the Presence of Data Noise," *Computational Statistics & Data Analysis* 51, 3100-3114.

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

[17] Hassler, U., Rodrigues, P. M. M. and Rubia, A. (2012), Quantile Regression for Long Memory Testing: A case of Realized Volatility, Working Paper 07/12, Banco de Portugal.

[18] Knight, K. (1998), "Limiting Distributions for L_1 Regression Estimators Under General Conditions," *The Annals of Statistics* 26, 755-770 .

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

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

[21] Koenker, R., and Xiao, Z. (2004), "Unit Root Quantile Autoregression Inference," *Journal of the American Statistical Association* 99, 775 787.

[22] Koenker, R., and Zhao, Q. (1996), "Conditional Quantile Estimation and Inference for ARCH Models," *Econometric Theory* 12(5), 793-813.

[23] Koul, H., and Saleh, A. K. (1995), "Autoregression Quantiles and Related Rank-Scores Processes," *The Annals of Statistics* 23, 670-689.

[24] López-Espinosa, G., Moreno, A., Rubia, A., and Valderrama, L. (2012), "Short-Term Wholesale Funding and Systemic Risk: A Global CoVaR Approach," *Journal of Banking and Finance* 36(12), 3150-3162.

[25] Portnoy, S. (1984), "Tightness of the Sequence of Empiric cdf Processes Defined from Regression Fractiles," in *Robust and nonlinear time series analysis*, eds. Franke, J., Hardle, W., and Martin, D., Springer-Verlag: New York.

[26] 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.

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

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

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

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

Technical Appendix

This appendix shows the formal proofs of Theorems 2.1 and 2.2 under Assumptions 1 and 2. It should be noted that the condition that $\{v_t\}$ is uniformly bounded under the L_r norms for some $r > 2$ in Assumption 2 is slightly stronger than necessary, but it allows us to offer a considerably simplified, yet rigorous, proof in this appendix. Furthermore, under the i.i.d. condition in Assumption 2, this technical restriction is not particularly severe and we can set $r = 2 + \epsilon$ with $\epsilon > 0$ arbitrarily small. A more detailed and involving proof, considering weaker conditions, is available in the accompanying working paper; see Hassler *et al.* (2012). We first provide a technical lemma for Theorem 2.1.

Lemma A *Under Assumption 2 and the null hypothesis,* $\max_{1\le t\le T}|x^*_{t-1,d}| = o_p(\sqrt{T})$.

Proof. Define the non-decreasing deterministic sequence $\xi_{t-1} = \sum_{j=1}^{t-1} j^{-1}$. Then, for some $\eta > 0$,

$$\Pr\left(\max_{1\le t\le T}|x^*_{t-1,d}| > \eta\sqrt{T}\right) \le \Pr\left(\max_{1\le t\le T}\sum_{j=1}^{t-1} j^{-1}|\varepsilon_{t-j,d}| > \eta\sqrt{T}\right) \le \Pr\left(\max_{1\le t\le T}|\varepsilon_{t,d}| > \frac{\eta\sqrt{T}}{\xi_{T-1}}\right).$$

Following Koenker and Zhao (1996, Lemma A.1) and noting that, under Assumption 2 $E\left(|\varepsilon_{t,d}|^{2+\epsilon}\right) \le K < \infty$, it then follows from Markov's inequality that

$$\Pr\left(\max_{1\le t\le T}|\varepsilon_{t,d}| > \frac{\eta\sqrt{T}}{\xi_{T-1}}\right) \le \frac{K}{\eta^{2+\epsilon}}\frac{\xi^{2+\epsilon}_{T-1}}{T^{\epsilon/2}} = o\left(1\right).$$

This upper bound vanishes asymptotically because ξ_{t-1} grows at logarithmic rate, $\log(t-1)$, such that $\xi_{T-1} = O\left(\log T\right)$ and, hence, $\xi^{2+\epsilon}_{T-1}T^{-\epsilon/2} \to 0$ as $T \to \infty$ for all $\epsilon > 0$. ■

Proof of Theorem 2.1

Following Knight (1998), Koenker (2005, Theo. 4.1) established the asymptotic normality of the estimates of a linear quantile regression under a set of sufficient conditions, namely, Conditions A1 and A2. The proof of Theorem 2.1 follows by direct verification of these conditions under the set of assumptions considered. Condition A1 is covered by Assumption 2. Condition A2 holds if *(i)* $T^{-1}\sum_{t=2}^{T}\mathbf{z}^*_{t-1,d}\mathbf{z}^{*\prime}_{t-1,d} \xrightarrow{p} \mathbf{\Omega}$, a positive definite limit; and *(ii)* $\max_{1\le t\le T}||\mathbf{z}^*_{t-1,d}|| = o_p(\sqrt{T})$. To show *(i)*, define the stationary sequence with an infinite past $\mathbf{z}^{**}_{t-1,d} = \left(1, x^{**}_{t-1,d}, \varepsilon_{t-1,d}, ..., \varepsilon_{t-p,d}\right)'$, noting that $E\left(\mathbf{z}^{**}_{t-1,d}\mathbf{z}^{**\prime}_{t-1,d}\right) = \mathbf{\Omega} < \infty$. Since $||T^{-1}\sum_{t=2}^{T}\mathbf{z}^*_{t-1,d}\mathbf{z}^{*\prime}_{t-1,d} - \mathbf{\Omega}|| = o_p(1)$ (*c.f.* Hassler *et al.* 2009, Lemma B.6), and $\mathbf{\Omega}$ is non-singular (c.f. Demetrescu *et al.* 2008, Lemma 5), *(i)* is fulfilled. Regarding *(ii)*, note that (4) corresponds to a quantile autoregression augmented by $x^*_{t-1,d}$. Quantile autoregressions have been dealt with by Koul and Saleh (1995), establishing *(ii)* under Assumptions 1 and 2 and $x^*_{t-1,d} = 0$. Consequently, we can ignore the lagged endogenous regressors and focus on the i.i.d. case aiming to show that *(ii)* holds when $x^*_{t-1,d} \neq 0$, since the proof then readily extends to the augmented context as in Koul and Saleh (1995). The required result follows directly from Lemma A and, consequently, the proof of Theorem 2.1 is complete by Koenker (2005, Theo. 4.1). ■

Proof of Theorem 2.2

Portnoy (1984) and Gutenbrunner and Jurečková (1992) showed that the QR process is tight, so the limit distribution of the function $\sqrt{T}\ \widehat{\phi}\left(\tau\right)$, seen as a random function of $\tau \in (0,1)$, is a rescaled (or non-standard) Brownian bridge under the null hypothesis and the conditions in Theorem 2.1, with a normal distribution arising for any fixed τ. Since $\widehat{s}\left(\tau\right)\widehat{\omega}_{22}$ converges to its theoretical counterpart uniformly on τ, following the arguments in Portnoy (1984), the scaled process $S_T\left(\tau\right) \rightarrow \mathcal{B}\left(\tau\right)$ in $(0,1)$, where $\mathcal{B}\left(\tau\right)$ is a standard Brownian bridge. Then, the limits stated for the Kolmogorov-Smirnov and the Cramér-von Mises type statistics in (7) and (8) follow directly from the continuous mapping theorem. ■

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 $\sigma_{RV}(t) = \left[\sum_{n=1}^{78} r_{(n),t}^2\right]^{1/2}$, (unnormalized) realized power variation $\sigma_{RPV}(t) = \sum_{m=1}^{78} |r_{(n),t}|$, and logarithmic transforms of these variables.

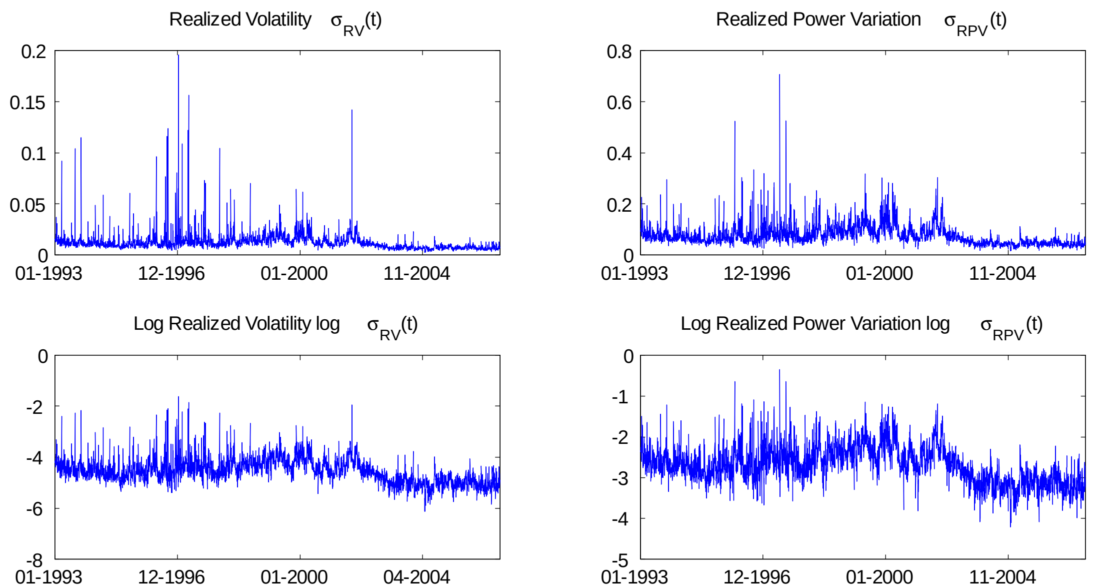

Figure 2: Sample Autocorrelation Function (ACF) of the measures of daily realized variation in Figure 1 together with upper 95% confidence band (dashed line).

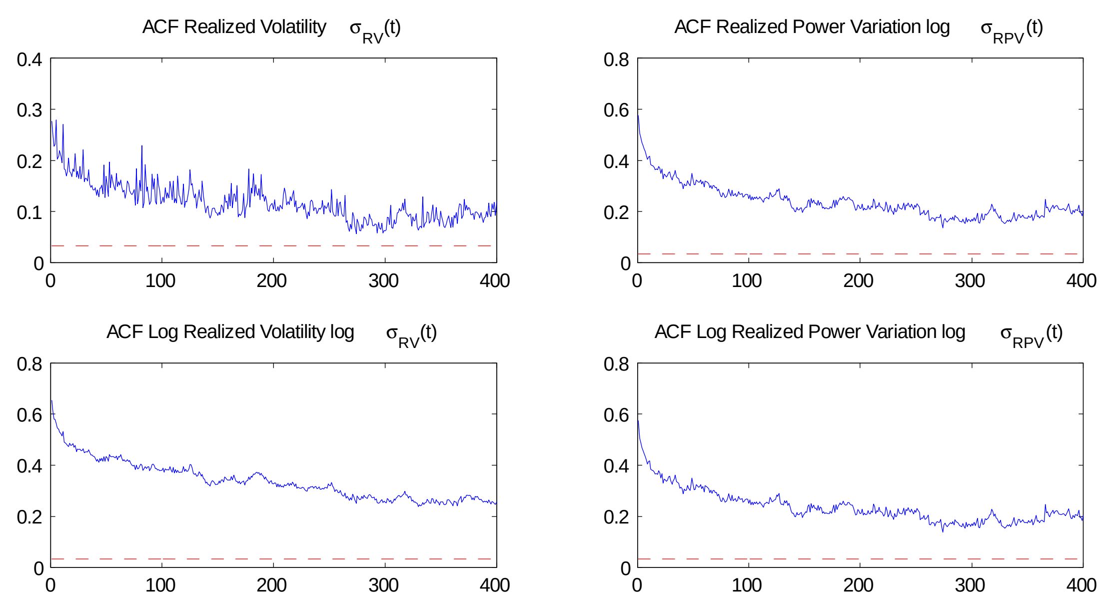

Figure 3: Estimates of the long-memory parameter of $\log \sigma_{RV}(t)$ and $\log \sigma_{RPV}(t)$ from the QRLM testing procedure and respective 95% and 99% confidence intervals. For any quantile $\tau \in \mathcal{Q}$, 'Central' denotes the value of $d \in \mathcal{D}$ for which the test statistic $|t_{QR}(\tau)|$ 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 $t_{QR}(\tau)$.

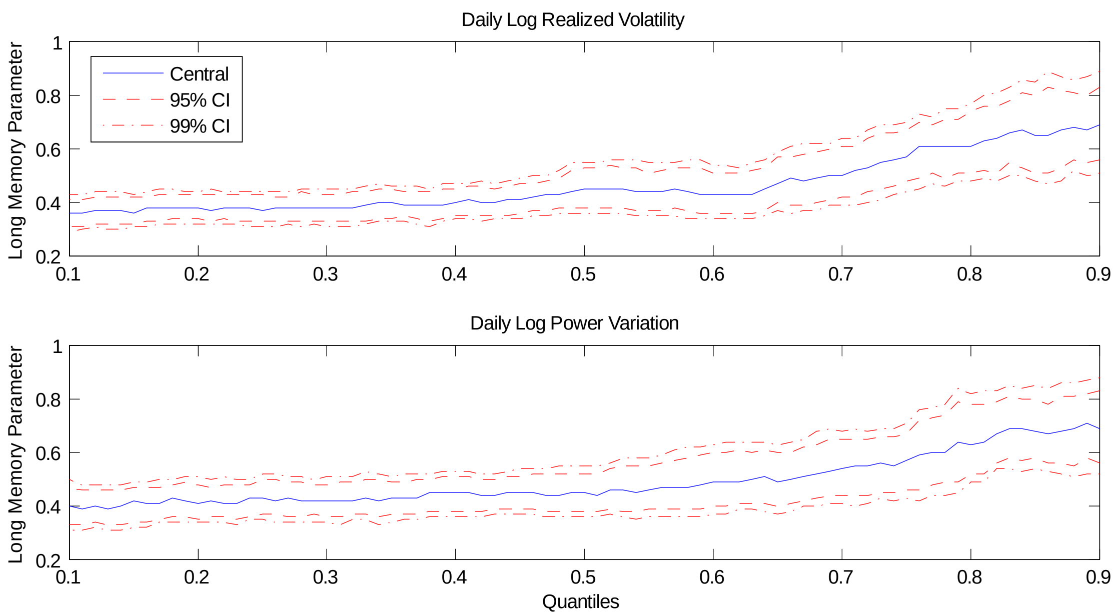

Figure 4: Values of the $\mathcal{KS}_{\mathcal{T}_2}$ test statistic on $\log \sigma_{RV}(t)$ for different values of d computed from the auxililary regression (4) and in a location-scale model.

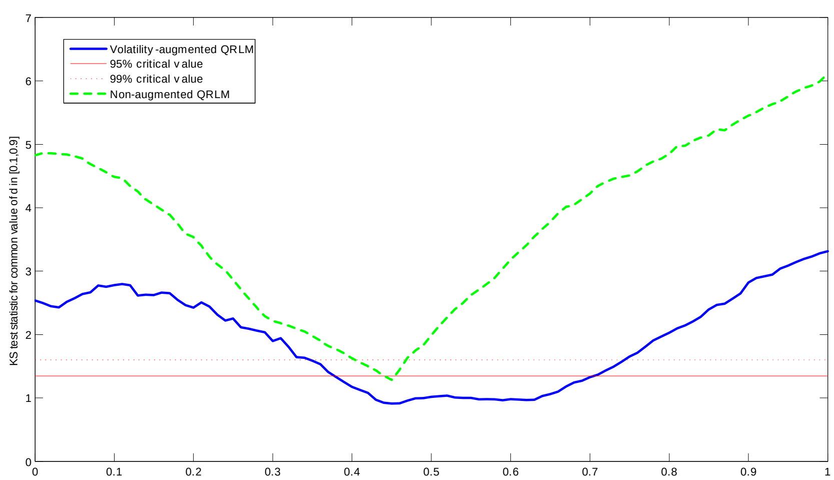

Figure 5: Values of the $\mathcal{KS}_{\mathcal{T}_2}$ test statistic on $\log \sigma_{RV}(t)$ for different values of d computed from the pre-crisis sample and the extended sample.

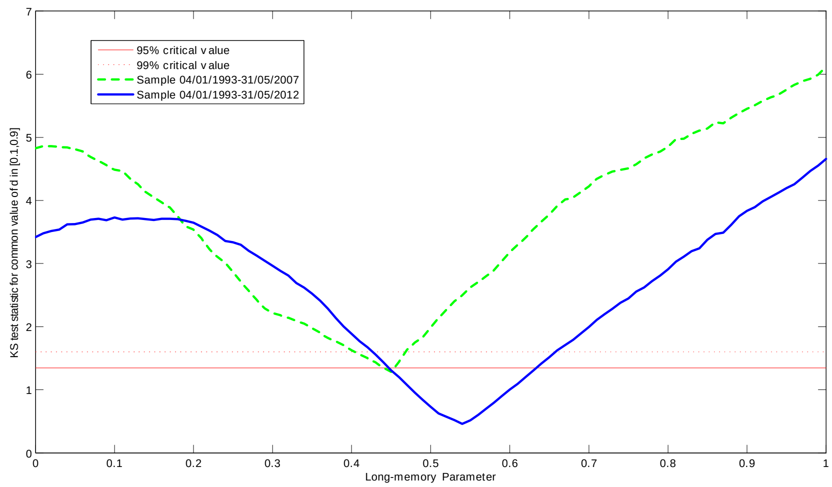

Table 1: Empirical rejection frequencies at the 5% nominal size of long-memory tests under short-run dependence. Data are generated according to $(1-L)^{1+\theta} y_t = \varepsilon_t$, $t=1,...,T$, with $(1-aL)\,\varepsilon_t = v_t$, and v_t being an i.i.d. sample drawn from a Student-t with v degrees of freedom. Tests statistics are computed under $\mathrm{H}_0: \theta = 0$. The entries LM_{QR} and LM_{LS} denote 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.

θ	$v=2$				$v=3$				$v=1000$ $(Gaussian)$			
	LM_{QR}		LM_{LS}		LM_{QR}		LM_{LS}		LM_{QR}		LM_{LS}	
	$a=0$	$a=0.75$	$a=0$	$a=0.75$	$a=0$	$a=0.75$	$a=0$	$a=0.75$	$a=0$	$a=0.75$	$a=0$	$a=0.75$
						T=100						
-0.30	98.80	5.68	94.02	3.50	94.24	6.60	93.70	3.96	76.22	9.90	92.44	4.60
-0.20	89.80	4.44	66.04	2.62	74.88	6.02	64.36	3.38	49.04	9.48	62.70	3.94
-0.10	46.18	3.78	17.08	2.68	30.78	6.10	19.54	3.72	22.16	9.58	22.62	4.36
0.00	4.72	5.16	3.78	3.52	6.04	8.04	4.34	4.86	7.56	11.76	4.64	4.92
0.10	36.64	8.30	22.64	4.94	19.82	11.44	24.32	6.20	10.98	15.06	24.68	6.18
0.20	83.18	14.40	71.34	7.10	63.84	18.48	69.64	8.56	34.42	19.92	68.68	8.60
0.30	96.82	25.56	93.70	11.46	90.08	29.06	93.54	12.58	68.98	29.20	93.38	13.10
						T=250						
-0.30	99.98	12.76	99.94	4.52	100.00	8.74	99.98	5.58	98.64	7.78	100.00	5.80
-0.20	99.88	5.48	97.88	3.10	98.72	6.10	97.22	3.80	85.14	6.44	97.04	3.68
-0.10	87.12	3.12	46.16	2.88	64.32	4.78	47.46	3.68	37.54	6.16	48.04	3.14
0.00	3.78	4.08	3.62	3.54	5.36	5.46	4.14	4.92	6.42	7.30	4.88	4.20
0.10	81.50	6.32	55.02	5.48	55.68	8.62	54.12	6.40	27.24	9.88	53.40	6.16
0.20	99.72	12.86	97.46	7.22	97.86	15.24	97.10	9.08	79.10	15.42	96.76	9.32
0.30	99.98	23.20	99.86	11.12	99.94	25.80	99.96	12.92	98.46	24.18	99.94	13.38
						T=1000						
-0.30	99.98	56.34	100.00	13.86	100.00	26.34	100.00	17.28	100.00	13.78	100.00	16.42
-0.20	100.00	24.14	100.00	5.74	99.98	10.50	100.00	7.26	99.96	7.38	100.00	7.20
-0.10	99.98	7.42	98.46	3.40	99.68	5.18	98.12	4.50	88.32	5.52	98.28	4.00
0.00	3.76	3.20	3.94	3.92	4.14	5.14	4.48	5.16	5.34	5.94	4.66	4.84
0.10	100.00	8.62	98.08	6.52	99.44	7.56	97.94	7.94	86.58	7.44	97.82	8.10
0.20	100.00	16.84	99.98	11.32	100.00	12.62	100.00	12.84	100.00	12.24	100.00	12.28
0.30	100.00	24.98	100.00	15.18	100.00	20.90	100.00	16.94	100.00	21.04	100.00	17.72

Table 2: Empirical rejection frequencies of the $\mathcal{KS}_{\mathcal{T}}$ test at the 5% nominal size, with $\mathcal{T} = [0.1, 0.9]$ and $\mathcal{T} = [0.4, 0.6]$. Data are generated from $(1-L)^{\bar{d}+\eta_t} y_t = \varepsilon_t$, $t = 1, ..., T$, with $\bar{d} = 1/2$, $\eta_t \sim iid\mathcal{N}\left(0, \gamma^2\right)$, and ε_t being driven by the Student-t process with i.i.d observations and v degress of fredom.

T=100	v=2		v=3		v=1000	
γ	**[0.1,0.9]**	**[0.4,0.6]**	**[0.1,0.9]**	**[0.4,0.6]**	**[0.1,0.9]**	**[0.4,0.6]**
0.00	0.09	0.04	0.13	0.07	0.12	0.08
0.05	0.61	0.61	0.14	0.08	0.19	0.13
0.10	0.98	0.96	0.25	0.19	0.32	0.23
0.15	1.00	1.00	0.41	0.33	0.45	0.35
0.20	1.00	1.00	0.58	0.50	0.61	0.50
0.25	1.00	1.00	0.71	0.60	0.77	0.66
0.30	1.00	1.00	0.81	0.76	0.88	0.80

T=1000	v=2		v=3		v=1000	
γ	**[0.1,0.9]**	**[0.4,0.6]**	**[0.1,0.9]**	**[0.4,0.6]**	**[0.1,0.9]**	**[0.4,0.6]**
0.00	0.07	0.04	0.05	0.04	0.04	0.03
0.05	0.87	0.87	0.33	0.33	0.27	0.25
0.10	1.00	1.00	0.82	0.84	0.80	0.77
0.15	1.00	1.00	0.99	0.99	1.00	0.98
0.20	1.00	1.00	1.00	1.00	1.00	1.00
0.25	1.00	1.00	1.00	1.00	1.00	1.00
0.30	1.00	1.00	1.00	1.00	1.00	1.00

Table 3: Descriptive statistics and semiparametric estimators of the long-memory parameter for different measures of daily realized variation: realized volatility $\sigma_{RV}(t) = \left[\sum_{n=1}^{78} r_{(n),t}^2\right]^{1/2}$, realized power variation $\sigma_{RPV}(t) = \sum_{n=1}^{78} |r_{(n),t}|$ and logarithmic transformations of these series. The statistic JB is the 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 denoted 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: $d = \delta$, with $\delta = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{100(1-\alpha)\%}(d|\tau)$ columns show the $100\alpha\%$ confidence intervals for d determined as the non-rejection region of the test at a $100\alpha\%$ 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: $d = \delta$ 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 2.2 computed over these intervals. The $\mathcal{CI}_{100(1-\alpha)\%}(d|\mathcal{T})$ columns show confidence interval for d determined as the non-rejection region of the joint test at a $100\alpha\%$ nominal level given the $\mathcal{T}$ quantile intervals. 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: $\mathbf{H}_0 : d = \delta$ 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 $\mathbf{H}_0 : d = \delta$ 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: $d = \delta$, with $\delta = 0, 0.1, ..., 1$ at the deciles $\tau = 0.1, \ldots, 0.9$ in rows. The $\mathcal{CI}_{100(1-\alpha)\%}(d|\tau)$ columns show the $100(1-\alpha)\%$ confidence intervals for d determined as the non-rejection region of the test at a $100\alpha\%$ 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: $d = \delta$ 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 2.2 computed over these intervals. The $\mathcal{CI}_{100(1-\alpha)\%}(d|\mathcal{T})$ columns show confidence interval for d determined as the non-rejection region of the joint test at a $100\alpha\%$ nominal level given the $\mathcal{T}$ quantile intervals. 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: $H_0 : d = \delta$ 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 $H_0 : d = \delta$ 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]