Exact inference in predictive quantile regressions with an application to stock returns

Sermin Gungor* Richard Luger†
University of Western Ontario Laval University

September 12, 2017

*Department of Economics, University of Western Ontario, London, Ontario N6A 5C2, Canada. *E-mail address:* `sgungor@uwo.ca`

†Corresponding author. Department of Finance, Insurance and Real Estate, Laval University, Quebec City, Quebec G1V 0A6, Canada. E-mail address: `richard.luger@fsa.ulaval.ca`

Exact inference in predictive quantile regressions with an application to stock returns

Abstract: We develop an exact and distribution-free procedure to test for quantile predictability at several quantile levels *jointly*, while allowing for an endogenous predictive regressor with any degree of persistence. The approach proceeds by combining together the quantile regression t-statistics from each considered quantile level and uses Monte Carlo resampling techniques to control the overall significance level of the data-dependent combination in finite samples. A simulation study confirms the fact that the proposed inference procedure controls the familywise error rate and achieves good power. We use the new approach to test the ability of many commonly used variables to predict the quantiles of excess stock returns, and shed new light on tail predictability.

JEL classification: C12; C14; C22

Keywords: Predictability; Quantile regression; Multiple testing; Persistent predictor; Monte Carlo permutation test; Exact distribution-free inference

1 Introduction

In a traditional predictive regression setting, an outcome variable of interest is regressed on the lagged value of a predictor variable. The relationship under investigation is

$$E(y_t \mid x_{t-1}) = \beta_0 + \beta_1 x_{t-1},$$

where y_t is the outcome variable and x_t is the predictor variable. If $\beta_1 \neq 0$, then x_t is said to have predictive ability for the mean of the outcome distribution. The econometric method used to detect such predictability is ordinary least squares (OLS) and the null hypothesis of no predictability ($\beta_1 = 0$) is tested with a t-statistic. It is well known, since the work by Mankiw and Shapiro (1986) and Stambaugh (1999), that the usual asymptotic theory provides a poor approximation to the finite-sample distribution of the t-statistic, resulting in far too many spurious rejections of the null hypothesis. This problem of false positives worsens as the predictor variable x_t becomes more persistent and its errors are more correlated with those of the regression model. A leading example is stock return predictability using a highly persistent regressor, like the dividend-price ratio, the earnings-price ratio, the book-to-market ratio, and various interest rate and interest rate spreads.

A more recent strand of the literature uses the quantile regression method of Koenker and Bassett (1978) to examine whether the outcome distribution is predictable not just in the centre, but also in the shoulders and tails; see Cenesizoglu and Timmermann (2008), Maynard et al. (2011) and Lee (2016). Instead of the conditional expectation, these studies focus on

$$Q_\tau(y_t \mid x_{t-1}) = \beta_0(\tau) + \beta_1(\tau) x_{t-1},$$

where $Q_\tau(y_t \,|\, x_{t-1})$ is the conditional quantile of y_t at a given quantile level $\tau \in (0, 1)$. An emerging view is that many economic variables seem to have far greater predictive ability for the outer quantiles of the stock return distribution compared to the centre of the return distribution. A potential pitfall, however, is that the standard quantile regression t-statistic is also prone to the overrejection problem, just like its mean regression counterpart; see Lee (2016) and the simulation results presented here below in Section 4.

Maynard et al. (2011) address the size distortion problem by deriving the asymptotic distribution of the quantile regression t-statistic under a local-to-unity specification for the predictor variable. Lee (2016) proposes a so-called extended instrumental variable (IVX) procedure to deal with the presence of persistent regressors in predictive quantile regressions. The idea underlying his IVX-QR procedure is to generate an instrumental variable of intermediate persistence by filtering a persistent and possibly endogenous regressor. Lee (2016) suggests using the instrument in lieu of the original predictor in the quantile regression and derives the asymptotic distribution of the resulting estimator under the null of no quantile predictability, $\beta_1(\tau) = 0$.

What is common to the large-sample approaches in Maynard et al. (2011) and Lee (2016) is that they only test for quantile predictability at a single quantile level, τ. Under the null hypothesis, however, the outcome distribution should not be predictable at *any* quantile level. Yet by testing for predictability at enough quantile levels, false positive will occur. The problem then consists of combining the predictability test results from each considered quantile level, say $\tau_1, ..., \tau_q$, in such a way that controls the familywise error rate in finite samples. The familywise error rate (FWER) is defined as

$$\text{FWER} = \Pr\big(\text{Reject at least one } H_0(\tau_i) \mid \text{all } H_0(\tau_i) \text{ are true}\big),$$

where $H_0(\tau_i)$ refers to the null hypothesis of no predictability at the specific quantile level τ_i. In words, the FWER is the probability of making at least one Type I error. Given a desired significance level α, the challenge here is to devise a procedure to test the absence of quantile predictability such that FWER $\leq \alpha$, no matter the sample size.

Given a finite-sample test for single-quantile predictability, one could then of course control the FWER at level α by applying the well-known Bonferroni procedure which consists of rejecting any hypothesis $H_0(\tau_i)$ with p-value $p(\tau_i) \leq \alpha/q$, where q is the number of quantiles being tested. The corresponding Bonferroni-adjusted p-value is given by $\min(q \times p(\tau_i), 1)$. Cenesizoglu and Timmermann (2008) make the Bonferroni adjustment to (asymptotically justified) bootstrapped p-values in order to obtain a joint test across all considered quantiles which is robust to arbitrary dependencies among the individual p-values. The problem with this approach is that it will be lacking in power as q grows. Instead of the Bonferroni adjustment, Westfall and Young (1993) advocate resampling-based methods to obtain less conservative multiple testing procedures which take into account the dependence structure between test statistics. The idea that resampling can be used to estimate the joint distribution of p-values is central to resampling-based testing.

In this spirit, we propose an *exact* resampling-based procedure for controlling the joint significance of quantile predictability tests at several quantile levels. We achieve control of the FWER by exploiting a permutation principle which holds under the null hypothesis of no quantile predictability, assuming that the predictor variable evolves according to an AR(1) model – a first-order autoregressive model. The AR(1) model is a standard assumption in the mean predictability literature; see Cavanagh et al. (1995), Stambaugh (1999), Lewellen (2004), Torous et al. (2004), Amihud and Hurvich (2004), Campbell and Yogo (2006), and Polk et al. (2006) among others. The distinguishing features of our approach are that: (i) it places no restrictions on the persistence of the predictor variable, thereby allowing for unit-root and even explosive behaviour; (ii) it is invariant to the contemporaneous dependence between the errors of the predictor model and those of the predictive regression model; and (iii) it makes no parametric distributional assumptions, thereby allowing for heavy-tailed errors.

The developed approach rests on an equally likely property for the collection of quantile regression t-statistics together with the AR(1) regression t-statistic. This property paves the way for an exact distribution-free test procedure using the technique of Monte Carlo (MC) tests (Barnard, 1963; Birnbaum, 1974; Dwass, 1957); see Dufour and Khalaf (2001) for a survey of MC test techniques. Specifically, the proposed test procedure proceeds by linking: (i) each point in an exact confidence set for the AR(1) parameter, which is obtained by numerically "inverting" an MC test, with (ii) MC p-values of a test statistic which combines the quantile regression t-statistics at each considered quantile level of the outcome distribution. This simultaneous inference approach yields a test of no quantile predictability that controls the FWER in finite samples.

The paper is organized as follows. In Section 2 we describe the general framework. In Section 3 we develop the quantile predictability test procedure. After an investigation in Section 4 of its size, power, and robustness to departures from the maintained assumptions, we then present in Section 5 an empirical application to test the ability of many commonly used variables to predict the quantiles of excess stock returns. We conclude in Section 6.

2 Statistical framework

The setting involves the outcome variable of interest y_t and another variable x_{t-1}, observed at time $t-1$, which could have the ability to predict y_t. More precisely, we work within a framework involving the collection of random variables $y_1, ..., y_T$, $x_0, x_1, ..., x_T$ and we say that x_{t-1} does not predict y_t if the following condition holds:

$$y_t \text{ is independent of } x_{t-1}, \text{ for each } t = 1, ..., T. \tag{1}$$

Let $F(y_t)$ and $F(y_t \,|\, x_{t-1})$ denote, respectively, the unconditional distribution and the conditional distribution of y_t, given x_{t-1}. Condition (1) can then be stated as

$$F(y_t \,|\, x_{t-1}) = F(y_t), \text{ for each } t = 1, ..., T,$$

while, under the alternative hypothesis of predictability, x_{t-1} affects some parts of the outcome distribution such that $F(y_t \,|\, x_{t-1}) \neq F(y_t)$. In order to complete the statistical framework, a time-series model for x_t will be specified. As typically done in the mean predictability literature, x_t will be assumed to evolve according to an AR(1) model.

We formally state the null hypothesis of no predictability as

$$H_0 : \begin{cases} y_t &= \beta_0 + u_t, \\ x_t &= \mu + \phi x_{t-1} + v_t, \end{cases} \tag{2}$$

where $(u_1, v_1)', ..., (u_T, v_T)'$ are independent are identically distributed (i.i.d.) random error vectors. No parametric distributional assumption (e.g. normality) is made, thereby leaving open the possibility of heavy-tailed errors. The parameters β_0, μ, and ϕ are unknown, and $\phi \in \mathcal{D}_\phi$, where $\mathcal{D}_\phi \subseteq \mathbb{R}$ is a nonempty set of admissible values for ϕ. Depending on the context, the choice of $\mathcal{D}_\phi$ could restrict ϕ to stationary values, or it could allow for a unit root and even explosive values. Indeed, the set $\mathcal{D}_\phi$ may be $\mathbb{R}$ itself, the open interval $(-1, 1)$, the closed interval $[-1, 1]$, or any other appropriate subset of $\mathbb{R}$.

Following Cenesizoglu and Timmermann (2008), Maynard et al. (2011), and Lee (2016), we use linear quantile regression models to detect predictability at various points of y_t's distribution. For convenience, these models are written here again as

$$Q_\tau(y_t \,|\, x_{t-1}) = \beta_0(\tau) + \beta_1(\tau)x_{t-1}, \tag{3}$$

where $Q_\tau(y_t \,|\, x_{t-1})$ is the conditional quantile of y_t at a given quantile level $\tau \in (0, 1)$. Stack the regressors appearing in (3) in the vector $\boldsymbol{x}_{t-1} = (1, x_{t-1})'$ and let $\boldsymbol{\theta}(\tau) = \big(\beta_0(\tau), \beta_1(\tau)\big)'$ be the vector comprising the associated quantile-specific intercept and slope coefficients. The standard quantile regression coefficients estimates are then given by

$$\hat{\boldsymbol{\theta}}(\tau) = \underset{\boldsymbol{\theta}(\tau)}{\arg\min} \sum_{t=1}^{T} \rho_\tau\big(y_t - \boldsymbol{\theta}(\tau)'\boldsymbol{x}_{t-1}\big),$$

where $\rho_\tau(u) = u(\tau - \mathbb{I}[u < 0])$ is the quantile regression loss function (Koenker and Bassett, 1978). Here $\mathbb{I}[A]$ is the indicator function which equals 1 when event A occurs, and 0 otherwise. In matrix form, the employed regressors are $\boldsymbol{X}_{-1} = [\boldsymbol{x}_0, ..., \boldsymbol{x}_{T-1}]'$.

Like in conventional quantile regression inference, we consider the t-statistic

$$t(\tau) = \frac{\hat{\beta}_1(\tau)}{\sqrt{\widehat{\text{var}}\big(\hat{\beta}_1(\tau)\big)}}. \tag{4}$$

Here $\hat{\beta}_1(\tau)$ is the second element of $\hat{\boldsymbol{\theta}}(\tau)$ and $\widehat{\text{var}}\big(\hat{\beta}_1(\tau)\big)$ is the $(2,2)$-entry in the matrix $\omega^2(\tau)(\boldsymbol{X}'_{-1}\boldsymbol{X}_{-1})^{-1}$, where $\omega^2(\tau) = \tau(1-\tau)/f_u^2(F_u^{-1}(\tau))$ and $f_u(F_u^{-1}(\tau))$ denotes the density of u_t evaluated at the τth quantile. This resembles the usual t-statistic based on OLS except that the definition of the variance uses $\omega^2(\tau)$ instead of σ^2, the variance of u_t. We compute $\omega^2(\tau)$ following the methods described in Koenker and Bassett (1978).[1] Under certain regularity conditions Lee (2016) shows that $t(\tau) \sim N(0,1)$, asymptotically. A two-sided p-value function can then be defined as

$$p(\tau) = 2\big(1 - \Phi(|t(\tau)|)\big), \tag{5}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. A remarkable feature of the proposed MC test procedure is that it will yield an exact inference even if $t(\tau)$ is not normally distributed.

Observe that (2) entails a joint hypothesis as it states that all points of y_t's distribution are unaffected by x_{t-1}. Let $\tau_1, ..., \tau_q$ be the quantile levels that will be used to test for predictability. Furthermore, let ϕ_0 be a specified value such that $\phi_0 \in \mathcal{D}_\phi$ and consider the following subhypothesis:

$$H_0(\phi_0) : \bigcap_{i=1}^{q} \beta_1(\tau_i) = 0,\ \phi = \phi_0, \tag{6}$$

[1] Specifically, we obtain the statistic in (4) using the `rq` command available with the R package 'quantreg'. This t-statistic can also be obtained with the `qreg` command in Stata.

which states zero slope coefficients in the quantile regressions at each considered quantile level, and an admissible value for x_t's persistence parameter. The joint null hypothesis of no predictability at quantile levels $\tau_1, ..., \tau_q$ can then be viewed as

$$H_0 : \bigcup_{\phi_0 \in \mathcal{D}_\phi} H_0(\phi_0), \tag{7}$$

where the union is taken over all admissible values for the AR(1) parameter in (2).

3 Exact inference methods

3.1 Tests of $H_0(\phi_0)$

The p-value function in (5) evaluated at each quantile level $\tau_1, ..., \tau_q$ appearing in (6) yields $p(\tau_1), ..., p(\tau_q)$. In order to test $H_0(\phi_0)$, these p-values can be combined via their minimum value:

$$p_{\min} = \min\left\{p(\tau_1), ..., p(\tau_q)\right\} \text{ and } S_{\min} = 1 - p_{\min},$$

so we reject $H_0(\phi_0)$ when $p_{\min}$ is small, or, equivalently, when $S_{\min}$ is large. The intuition here is that the null hypothesis should be rejected if at least one of the individual p-values is sufficiently small. This method of combining tests was suggested by Tippett (1931) and Wilkinson (1951).

The second method we consider – originally suggested by Fisher (1932) and Pearson (1933) – combines the p-values via their product:

$$p_\times = \prod_{i=1}^{q} p(\tau_i) \text{ and } S_\times = 1 - p_\times,$$

which may provide more information about departures from the null hypothesis compared to using only the minimum p-value. Indeed, the product of several quantile-specific p-values may well indicate a rejection of the null of no predictability even though the individual p-values may not be small enough to be significant on their own. Here we let $\mathcal{S}$ refer to either $S_{\min}$ or $S_\times$. For further discussion and other examples of the test combination technique, see Folks (1984), Westfall and Young (1993), Dufour et al. (2015), and Catani and Ahlgren (2017).

It seems natural to test the value ϕ_0 specified in $H_0(\phi_0)$ using a standard OLS-based t-statistic:

$$t^{\mathrm{OLS}}(\phi_0) = \frac{\hat{\phi}^{\mathrm{OLS}} - \phi_0}{\sqrt{\widehat{\mathrm{var}}\left(\hat{\phi}^{\mathrm{OLS}}\right)}}. \tag{8}$$

As usual, $\hat{\phi}^{\mathrm{OLS}} = (\boldsymbol{X}'_{-1}\boldsymbol{X}_{-1})^{-1}\boldsymbol{X}'_{-1}\boldsymbol{X}$, where $\boldsymbol{X}_{-1}$ was defined previously and $\boldsymbol{X} = (x_1, ..., x_T)'$. The quantity $\widehat{\mathrm{var}}\left(\hat{\phi}^{\mathrm{OLS}}\right)$ is the $(2,2)$-entry in the matrix $\hat{\sigma}^2(\boldsymbol{X}'_{-1}\boldsymbol{X}_{-1})^{-1}$ computed with $\hat{\sigma}^2 = (\boldsymbol{X}'\boldsymbol{M}\boldsymbol{X})/(T-2)$, where $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}_{-1}(\boldsymbol{X}'_{-1}\boldsymbol{X}_{-1})^{-1}\boldsymbol{X}'_{-1}$ is the matrix that projects onto the space orthogonal to the span of $\boldsymbol{X}_{-1}$. This OLS route is reasonable if one assumes that the distribution of v_t in (2) has mean zero and a finite variance.

Alternatively, if one assumes that the median of v_t is zero, then the least absolute deviations (LAD) estimator becomes the natural choice. The LAD estimates of $\boldsymbol{\vartheta} = (\mu, \phi)$ are found as

$$\hat{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}} \sum_{t=1}^{T} \rho_\tau(x_t - \boldsymbol{\vartheta}'\boldsymbol{x}_{t-1})$$

by setting $\tau = 0.5$. In this case, the loss function corresponds to $\rho_{0.5}(v) = 0.5|v|$. The LAD estimates yield a t-statistic to test $\phi = \phi_0$ of the form

$$t^{\text{LAD}}(\phi_0) = \frac{\hat{\phi}^{\text{LAD}} - \phi_0}{\sqrt{\widehat{\text{var}}\left(\hat{\phi}^{\text{LAD}}\right)}}, \tag{9}$$

whose computation is similar to (4). We use $\mathcal{T}(\phi_0)$ to denote either $|t^{\text{OLS}}(\phi_0)|$ or $|t^{\text{LAD}}(\phi_0)|$ with a view towards two-sided alternatives.

Let $\boldsymbol{z}_t = (y_t, x_t)'$ and define $\boldsymbol{\varepsilon}_t = (y_t,\ x_t - \phi_0 x_{t-1})'$ for the given value ϕ_0. Observe that $\boldsymbol{\varepsilon}_t = (\beta_0 + u_t,\ \mu + v_t)'$ and $\boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_T$ forms a collection of i.i.d. (and hence *exchangeable*) random vectors under $H_0(\phi_0)$. This means that for every permutation $d_1, ..., d_T$ of the integers $1, ..., T$, we have

$$(\boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_T) \stackrel{d}{=} (\boldsymbol{\varepsilon}_{d_1}, ..., \boldsymbol{\varepsilon}_{d_T}), \tag{10}$$

where the symbol "$\stackrel{d}{=}$" stands for the equality in distribution (Randles and Wolfe, 1979, Definition 1.3.6). The key property here is that the joint distribution of exchangeable random vectors remains invariant under permutations of their order. Note that property (10) is also invariant to the true values of β_0 and μ, and to the contemporaneous dependence structure between u_t and v_t in (2).

To simplify the notation, let $\tilde{\boldsymbol{\varepsilon}}_t = \boldsymbol{\varepsilon}_{d_t}$ and consider the sequence of random vectors $\tilde{\boldsymbol{z}}_1, ..., \tilde{\boldsymbol{z}}_T$ obtained from the recursion

$$\tilde{\boldsymbol{z}}_t = \begin{pmatrix} \tilde{y}_t \\ \tilde{x}_t \end{pmatrix} = \begin{pmatrix} 0 \\ \phi_0 \tilde{x}_{t-1} \end{pmatrix} + \tilde{\boldsymbol{\varepsilon}}_t, \text{ for } t = 2, ..., T, \tag{11}$$

with initial value $\tilde{\boldsymbol{z}}_1 = \boldsymbol{z}_1$. It is easy to see that $\boldsymbol{z}_t$ and $\tilde{\boldsymbol{z}}_t$ obey the same data-generating process when $H_0(\phi_0)$ holds. Denoting the original sample as $\boldsymbol{Z} = (\boldsymbol{z}_1, ..., \boldsymbol{z}_T)$ and an artificial sample generated according to (11) as $\tilde{\boldsymbol{Z}} = (\tilde{\boldsymbol{z}}_1, ..., \tilde{\boldsymbol{z}}_T)$, we then have under $H_0(\phi_0)$ that $\boldsymbol{Z} \overset{d}{=} \tilde{\boldsymbol{Z}}$, for each of the $(T-1)!$ possible realizations of $\tilde{\boldsymbol{Z}}$. Theorem 1.3.7 in Randles and Wolfe (1979) further implies that

$$\begin{pmatrix} \mathcal{S} \\ \\ \mathcal{T}(\phi_0) \end{pmatrix} \overset{d}{=} \begin{pmatrix} \widetilde{\mathcal{S}} \\ \\ \widetilde{\mathcal{T}}(\phi_0) \end{pmatrix}, \tag{12}$$

where $\widetilde{\mathcal{S}}$ refers to either $\widetilde{S}_{\min}$ or $\widetilde{S}_{\times}$, defined, respectively, as

$$\widetilde{S}_{\min} = 1 - \min\left\{\tilde{p}(\tau_1), ..., \tilde{p}(\tau_q)\right\} \text{ and } \widetilde{S}_{\times} = 1 - \prod_{i=1}^{q} \tilde{p}(\tau_i),$$

with $\tilde{p}(\tau_i) = 2\big(1 - \Phi(|\tilde{t}(\tau_i)|)\big)$, for $i = 1, ..., q$; and $\widetilde{\mathcal{T}}(\phi_0)$ refers to either $|\tilde{t}^{\text{OLS}}(\phi_0)|$ or $|\tilde{t}^{\text{LAD}}(\phi_0)|$. Here $\tilde{t}(\tau_i)$, $\tilde{t}^{\text{OLS}}(\phi_0)$, and $\tilde{t}^{\text{LAD}}(\phi_0)$ are the t-statistics in (4), (8), and (9), respectively, each computed with the artificial sample $\tilde{\boldsymbol{Z}}$. We see that generating such artificial samples according to (11) yields $(T-1)!$ possible values of $\big(\widetilde{\mathcal{S}}, \widetilde{\mathcal{T}}(\phi_0)\big)'$ which are all equally likely values for $\big(\mathcal{S}, \mathcal{T}(\phi_0)\big)'$, i.e.,

$$\Pr\left(\begin{pmatrix} \mathcal{S} \\ \\ \mathcal{T}(\phi_0) \end{pmatrix} = \begin{pmatrix} \widetilde{\mathcal{S}} \\ \\ \widetilde{\mathcal{T}}(\phi_0) \end{pmatrix}\right) = \frac{1}{(T-1)!}. \tag{13}$$

The equally likely property in (13) paves the way for an MC test of $H_0(\phi_0)$ as follows.[2] First, the statistics $\big(\mathcal{S}, \mathcal{T}(\phi_0)\big)'$ are computed with the original sample $\boldsymbol{Z}$. The test then proceeds by generating $B-1$ artificial samples $\tilde{\boldsymbol{Z}}_1, ..., \tilde{\boldsymbol{Z}}_{B-1}$, each one according to (11). With each such sample, the statistics are computed yielding $\big(\widetilde{\mathcal{S}}_b, \widetilde{\mathcal{T}}_b(\phi_0)\big)'$, for

[2]Note that calculating the complete randomization distribution would require enumerating $(T-1)!$ possibilities, which is well-nigh impossible. As in Dwass (1957), the MC test technique yields exact p-values without the need to enumerate the entire randomization distribution.

$b = 1, ..., B-1$. Observe that (13) is a discrete bivariate uniform distribution, meaning that ties among the resampled values can occur. A test with *size* α can be obtained by applying the following tie-breaking rule (Dufour, 2006).[3] Draw B i.i.d. variates U_b, $b = 1, ..., B$, from a continuous uniform distribution on $[0, 1]$ and randomly assign then to create the triplets $\big(\widetilde{\mathcal{S}}_1, \widetilde{\mathcal{T}}_1(\phi_0), U_1\big)', ..., \big(\widetilde{\mathcal{S}}_{B-1}, \widetilde{\mathcal{T}}_{B-1}(\phi_0), U_{B-1}\big)', \big(\mathcal{S}, \mathcal{T}(\phi_0), U_B\big)'$. Next, compute the lexicographic ranks of $\mathcal{S}$ and $\mathcal{T}(\phi_0)$ according to

$$\begin{aligned}
\widetilde{R}_B\left[\mathcal{S}\right] &= 1 + \sum_{b=1}^{B-1} \mathbb{I}\left[\mathcal{S} > \widetilde{\mathcal{S}}_b\right] + \sum_{b=1}^{B-1} \mathbb{I}\left[\mathcal{S} = \widetilde{\mathcal{S}}_b\right] \times \mathbb{I}\left[U_B > U_b\right], \\
\widetilde{R}_B\left[\mathcal{T}(\phi_0)\right] &= 1 + \sum_{b=1}^{B-1} \mathbb{I}\left[\mathcal{T}(\phi_0) > \widetilde{\mathcal{T}}_b(\phi_0)\right] + \sum_{b=1}^{B-1} \mathbb{I}\left[\mathcal{T}(\phi_0) = \widetilde{\mathcal{T}}_b(\phi_0)\right] \times \mathbb{I}\left[U_B > U_b\right].
\end{aligned} \tag{14}$$

Upon recognizing that the triplets $\big(\widetilde{\mathcal{S}}_1, \widetilde{\mathcal{T}}_1(\phi_0), U_1\big)',...,\big(\widetilde{\mathcal{S}}_{B-1}, \widetilde{\mathcal{T}}_{B-1}(\phi_0), U_{B-1}\big)'$, $\big(\mathcal{S}, \mathcal{T}(\phi_0), U_B\big)'$ are exchangeable under $H_0(\phi_0)$, we then know from Lemma 2.3 in Dufour (2006) that the lexicographic ranks in (14) are uniformly distributed over the integers $1, ..., B$; i.e., $\Pr\big(\widetilde{R}_B\left[\mathcal{S}\right] = b\big) = 1/B$, for $b = 1, ..., B$, and $\Pr\big(\widetilde{R}_B\left[\mathcal{T}(\phi_0)\right] = b\big) = 1/B$, for $b = 1, ..., B$.

From the marginal distributions of each statistic, we can compute MC p-values as

$$\begin{aligned}
\tilde{p}_B\left[\mathcal{S}\right] &= \frac{B - \widetilde{R}_B\left[\mathcal{S}\right] + 1}{B}, \\
\tilde{p}_B\left[\mathcal{T}(\phi_0)\right] &= \frac{B - \widetilde{R}_B\left[\mathcal{T}(\phi_0)\right] + 1}{B},
\end{aligned} \tag{15}$$

where $\widetilde{R}_B\left[\mathcal{S}\right]$ and $\widetilde{R}_B\left[\mathcal{T}(\phi_0)\right]$ are the ranks of $\mathcal{S}$ and $\mathcal{T}(\phi_0)$, respectively, given by (14). The p-values in (15) may have a very complex dependence structure. Nevertheless, if we choose

[3] Here we follow the terminology in Lehmann and Romano (2005, Ch. 3) and say that a test of H_0 has *size* α if Pr(Rejecting $H_0 \,|\, H_0$ true) $= \alpha$, and that it has *level* α if Pr(Rejecting $H_0 \,|\, H_0$ true) $\leq \alpha$.

B so that $B\alpha$ is an integer (for $0 < \alpha < 1$), then these MC p-values each exactly have a size equal to α in the sense that

$$\begin{aligned} \Pr\left(\tilde{p}_B\left[\mathcal{S}\right] \leq \alpha\right) &= \alpha, \\ \Pr\left(\tilde{p}_B\left[\mathcal{T}(\phi_0)\right] \leq \alpha\right) &= \alpha, \end{aligned} \tag{16}$$

under $H_0(\phi_0)$. A decision rule could then be built from the logical equivalence that $H_0(\phi_0)$ is false if and only if $\{\bigcup_{i=1}^{q} \beta_1(\tau_i) \neq 0\}$ or $\{\phi \neq \phi_0\}$. The critical region corresponding to this decision rule is $\{\tilde{p}_B\left[\mathcal{S}\right] \leq \alpha\} \cup \{\tilde{p}_B\left[\mathcal{T}(\phi_0)\right] \leq \alpha\}$ and, by subadditivity, we obtain its level as

$$\Pr\left(\left\{\tilde{p}_B\left[\mathcal{S}\right] \leq \alpha\right\} \cup \left\{\tilde{p}_B\left[\mathcal{T}(\phi_0)\right] \leq \alpha\right\}\right) \leq 2\alpha.$$

The marginal distributions of the p-values characterized by (16) are the foundations for the tests of H_0, developed next.

3.2 Tests of H_0

The expression in (7) makes clear that ϕ is a nuisance parameter in the present context, since it is not pinned down to a specific value under H_0. In order to test such a hypothesis, which contains several distributions, we can appeal to a *minimax* argument stated as: "reject the null hypothesis whenever, for all admissible values of the nuisance parameter under the null, the corresponding point null hypothesis is rejected" (Savin, 1984).

With the statistic $\mathcal{S} = \mathcal{S}(\phi_0)$,[4] this would mean maximizing the MC p-value $\tilde{p}_B\left[\mathcal{S}(\phi_0)\right]$

[4]From this point on, we shall use the notation $\mathcal{S}(\phi_0)$ to emphasize the fact that the distribution of $\mathcal{S}$ depends on ϕ_0.

over $\phi_0 \in \mathcal{D}_\phi$. The rationale is that

$$\sup_{\phi_0 \in \mathcal{D}_\phi} \tilde{p}_B \left[\mathcal{S}(\phi_0)\right] \leq \alpha \Longrightarrow \tilde{p}_B \left[\mathcal{S}(\phi)\right] \leq \alpha,$$

where $\tilde{p}_B \left[\mathcal{S}(\phi)\right]$ is the MC p-value of $\mathcal{S}$ based on the true value ϕ. Moreover, $\Pr\big(\tilde{p}_B \left[\mathcal{S}(\phi_0)\right] \leq \alpha\big) = \alpha$ under $H_0(\phi_0)$ and for all $\phi_0 \in \mathcal{D}_\phi$. So, if αB is an integer, we then have

$$\Pr\left(\sup_{\phi_0 \in \mathcal{D}_\phi} \tilde{p}_B \left[\mathcal{S}(\phi_0)\right] \leq \alpha\right) \leq \alpha.$$

The decision rule in this case would be to reject H_0 if the maximized p-value is $\leq \alpha$. Otherwise, accept H_0 since there is not enough evidence to reject it. Note that this test has *level* α, meaning it is conservative.

Following Beaulieu et al. (2007), we can replace $\mathcal{D}_\phi$ appearing in (7) by an exact confidence set for ϕ which is valid at least under the null hypothesis. This can be interpreted as plugging in an estimator of the (perhaps unknown) set of admissible ϕ-values. Let $C_\phi(\alpha_1)$ denote a confidence set for ϕ with level $1-\alpha_1$, i.e. such that $\Pr\big(\phi \in C_\phi(\alpha_1)\big) \geq 1-\alpha_1$ under H_0. It can then be shown (see the Appendix) that

$$\Pr\left(\sup_{\phi_0 \in C_\phi(\alpha_1)} \tilde{p}_B \left[\mathcal{S}(\phi_0)\right] \leq \alpha_2\right) \leq \alpha_1 + \alpha_2. \tag{17}$$

Note as well that this is the main idea of the Bonferroni methods frequently used to deal with nuisance parameters in predictive mean regressions; see, for example, Cavanagh et al. (1995) and Campbell and Yogo (2006).[5]

[5]It is worth remarking, however, that Phillips (2014) has shown that the confidence intervals based on local-to-unity limit theory used in these predictive regression tests are invalid in the stationary case, with zero asymptotic coverage probability. In particular, this causes the popular Q-test statistic of Campbell and

Recall that a confidence set for a scalar parameter can be interpreted as the result of a collection of tests for each admissible value of the parameter. The confidence set simply reports all the values that cannot be rejected at a given nominal level. The confidence set appearing in (17) can therefore be obtained as

$$C_\phi(\alpha_1) = \big\{\phi_0 : \phi_0 \in \mathcal{D}_\phi,\ \tilde{p}_B\left[\mathcal{T}(\phi_0)\right] > \alpha_1\big\},$$

where $\tilde{p}_B\left[\mathcal{T}(\phi_0)\right]$ is the MC p-value of $\mathcal{T}(\phi_0)$ in (15). Observe that $C_\phi(\alpha_1)$ is an exact and distribution-free confidence set for ϕ, provided of course that $\mathcal{D}_\phi$ does not exclude the true value of ϕ.[6] A confidence *interval* can be defined as

$$CI_\phi(\alpha_1) = \Big[\inf\big\{\phi_0 : \phi_0 \in C_\phi(\alpha_1)\big\},\ \sup\big\{\phi_0 : \phi_0 \in C_\phi(\alpha_1)\big\}\Big], \tag{18}$$

which may be used to conduct finite-sample inference about the value of ϕ, since $\Pr\big(\phi \in CI_\phi(\alpha_1)\big) \geq 1 - \alpha_1$. This is an interesting result in itself because it is well known that the conventional t-test can yield misleading conclusions about the AR(1) parameter, particularly when it is close to 1; see Tanaka (1983), Nankervis and Savin (1985, 1988), Rayner (1990), and Nankervis and Savin (1996).

3.3 Summary of test procedure

The test of H_0 in (7) can be performed with four possible statistics depending on whether $\mathcal{S}(\phi_0)$ is set to $S_{\min}(\phi_0)$ or $S_\times(\phi_0)$, and whether $\mathcal{T}(\phi_0)$ is set to $|t^{\mathrm{OLS}}(\phi_0)|$ or $|t^{\mathrm{LAD}}(\phi_0)|$. Given

Yogo (2006) to erroneously indicate predictability with probability approaching unity even though the null of no predictability holds true.

[6] On the other hand, eliminating (truncating) inadmissible values from a confidence set does not modify its level (Abdelkhalek and Dufour, 1998).

the choice of statistics $\mathcal{S}(\phi_0)$ and $\mathcal{T}(\phi_0)$, the rest of the test procedure proceeds according to the following steps:

1. For each $\phi_0 \in \mathcal{D}_\phi$, compute the MC p-values $\left(\tilde{p}_B\left[\mathcal{S}(\phi_0)\right], \tilde{p}_B\left[\mathcal{T}(\phi_0)\right]\right)$ as in (15).

2. Compute a confidence set for ϕ with level $1 - \alpha_1$ as

$$C_\phi(\alpha_1) = \left\{\phi_0 : \phi_0 \in \mathcal{D}_\phi,\ \tilde{p}_B\left[\mathcal{T}(\phi_0)\right] > \alpha_1\right\}.$$

3. Find $p^* = \sup_{\phi_0 \in C_\phi(\alpha_1)} \tilde{p}_B\left[\mathcal{S}(\phi_0)\right]$ and reject H_0 if $p^* \leq \alpha_2$. Otherwise, accept H_0.

In practical applications, the set $\mathcal{D}_\phi$ in Step 1 is replaced by a discrete grid. Observe that only the candidate values ϕ_0 vary in Step 1. In order to control the underlying randomness in the computation of the MC p-values linked to each ϕ_0-value, the seed of the random number generator should be reset to the same value before each considered ϕ_0. In the terminology of Dufour (2006), we will refer to the test described by Steps 1–3 as a *maximized* MC (MMC) test. As a by-product, this MMC test procedure also yields an exact distribution-free confidence set for ϕ in Step 2. A confidence interval for ϕ can then be obtained from $C_\phi(\alpha_1)$ according to (18).

Given a desired significance level α, we see that there is a tradeoff between the width of the confidence set $C_\phi(\alpha_1)$ in Step 2 and the significance level $\alpha_2 = \alpha - \alpha_1$ in Step 3. While the choice of α_1, α_2 has no effect on the overall level (as long as $\alpha_1 + \alpha_2 = \alpha$), it does matter for power. Campbell and Dufour (1997) suggest that it is better to take a wider confidence set in order to have a tighter critical value when deciding whether to reject H_0. Accordingly, we carry on with the testing strategy represented by $\alpha_1 = 1\%, \alpha_2 = 4\%$ for an overall $\alpha = 5\%$.

A computationally simplified procedure is obtained by replacing the confidence set by a point estimate $\hat{\phi}$, which could be either $\hat{\phi}^{\mathrm{OLS}}$ or $\hat{\phi}^{\mathrm{LAD}}$, and rejecting H_0 whenever $\tilde{p}_B\big[\mathcal{S}(\hat{\phi})\big] \leq \alpha$. Dufour (2006) refers to such tests as *local* MC (LMC) tests, and, since they are quite natural, we will study their size and power properties in the simulation study.[7]

4 Simulation experiments

In this section, we report the results of a series of simulation experiments designed to examine the size and power performance of the proposed tests. When testing for predictability at a single quantile, we use $\mathrm{LMC}^{\mathrm{OLS}}$ and $\mathrm{LMC}^{\mathrm{LAD}}$ to denote the LMC tests based on the OLS and LAD estimates of ϕ, respectively. Similarly, $\mathrm{MMC}^{\mathrm{OLS}}$ and $\mathrm{MMC}^{\mathrm{LAD}}$ will denote the MMC tests whose confidence sets for ϕ are obtained by inverting the OLS- and LAD-based t-statistics, respectively. For the joint tests involving several quantile levels, we add the subscripts "min" or "$\times$" to indicate that the p-values of the quantile regression t-statistics are combined via the minimum or the product of their values. All the tests are performed at the nominal $\alpha = 5\%$ significance level with $B - 1 = 99$ artificial samples, and the reported results are based on 1000 replications of each data-generating configuration.

For sample sizes $T = 120$ and 240, we generate data according to

$$\begin{aligned} y_t &= \beta_0 + \sigma_t u_t, \\ x_t &= \mu + \phi x_{t-1} + v_t, \end{aligned} \tag{19}$$

where $(u_t, v_t)'$ are i.i.d. random vectors following either a bivariate normal or a bivariate

[7]The term "local" reflects the fact that the underlying MC p-value is based on a specific choice for the nuisance parameter. Under additional regularity conditions, Dufour (2006) shows the asymptotic validity of LMC tests. These conditions are notably more restrictive than those under which (17) obtains.

Student-t distribution with 3 degrees of freedom. In both cases, the means are zero, scales are one, and we let ρ determine the contemporaneous covariance between u_t and v_t. The parameter ρ determines the strength of feedback from u_t to future values of the regressor variable. We set $\beta_0 = \mu = 0$, and we let $\phi = 0.95$, 0.99 and $\rho = 0$, -0.90, -0.95 to examine the effects of feedback.[8]

The null hypothesis H_0 is represented by $\sigma_t = 1$, leaving no way for x_{t-1} to affect the distribution of y_t. We let the alternative hypothesis be parameterized as $\sigma_t = |x_{t-1}|$ so the τth conditional quantile of y_t becomes

$$Q_\tau(y_t \,|\, x_{t-1}) = \beta_0(\tau) + |x_{t-1}| Q_\tau(u_t),$$

where $Q_\tau(u_t)$ is the τth quantile of u_t. This specification bears a resemblance to the absolute value ARCH model of Taylor (1986) and Schwert (1989) in which the conditional standard deviation depends on the absolute values of the volatility forcing variables. With symmetric error distributions, we then have $Q_\tau(u_t) = 0$ at $\tau = 0.5$ (the median). This means that the predictive ability of x_{t-1} is zero for the conditional median of y_t and increases (symmetrically) for the outer quantiles of y_t, because $|x_{t-1}||Q_\tau(u_t)|$ is an increasing function of $|\tau - 0.5|$. This setup is motivated by the empirical evidence in Cenesizoglu and Timmermann (2008), Maynard et al. (2011) and Lee (2016) who suggest that many economic variables have far greater predictive ability for the outer quantiles of the return distribution compared to the centre of the return distribution.

Tables 1 and 2 show the size and power results of the predictability tests applied at

[8]The negative values of ρ are in the range of error correlations typically found (e.g. when the dividend-price ratio is used as a predictor of stock returns) and are frequently used in the predictive regression literature.

various quantile levels from 0.1 to 0.9, one at a time. We do not include $\tau = 0.5$ because, by construction, there is no predictability at that quantile level. The standard t-test and the IVX-QR test of Lee (2016) serve as benchmarks for comparison purposes in these two tables. The IVX-QR method proceeds by generating an instrumental variable w_t according to the filtering rule

$$w_t = R_w w_{t-1} + \Delta x_t, \quad R_w = 1 + \frac{c_w}{T^\delta},$$

where $\delta \in (0, 1)$, $c_w < 0$, and $w_0 = 0$. In order to test the null of no quantile predictability, Lee (2016) proposes to use an ordinary quantile regression of y_t on w_{t-1}, and shows that the resulting estimator is asymptotically normal. Here we set $c_w = -5$ and $\delta = 0.5$ like in the online supplement to Lee (2016) who reports that reliable size is achieved with this choice.

Table 1 shows that the standard t-test has a tendency to over-reject the null in all cases. The IVX-QR test behaves relatively better than the t-test for the central quantiles, but still fails somewhat to control its size especially in the tails of the distribution where there are fewer observations. Indeed, we see in Table 1 that the empirical size of the t- and IVX-QR tests can be two to three times larger than the nominal level. The general pattern is that the over-rejection problem with the t-test is exacerbated by higher persistence (ϕ) and/or feedback (ρ). On the contrary, the empirical size of IVX-QR test is quite stable across (ϕ, ρ) values. Of course, the size of the IVX-QR test could be brought even closer to the nominal level by setting δ to a more conservative (smaller) value.

For the most part, the LMC tests appear to be well behaved with empirical size close to 5%. They do however display a slight tendency to over-reject, particularly when ϕ and $|\rho|$ are both high. For instance, the empirical size of the LMC tests can be around 10% when $T = 120$, and $\phi = 0.99$ and $\rho = -0.9$. Interestingly, the LMC tests behave relatively better for tail quantiles compared to the more central quantiles. In accordance with the developed theory, the MMC tests are seen to always respect the nominal 5% level constraint. This holds true regardless of the predictor persistence, feedback strength, sample size, and error distribution. Finally, the estimation method (i.e. OLS or LAD) in the MC test procedures matters not for their empirical size.

Given the size distortions observed in Table 1, the power results in Table 2 for the t-, IVX-QR, and LMC tests are based on size-adjusted critical values. The MMC tests did not require any such adjustments since they reject the null with probability not exceeding 5%. In general, the t-test displays the highest power just ahead of the LMC tests. The MMC tests are generally less powerful than the t- and LMC tests, but the power gap narrows as ϕ and/or $|\rho|$ increases. Heavier tailed t_3 errors tend to decrease test power compared to the normal case. Table 2 also shows clearly that the IVX-QR test is lacking in power. The takeaway message is that the IVX-QR approach achieves relatively better size control when compared to the standard quantile regression t-test, but this comes at a very steep price in terms of power.

To understand why the IVX-QR test is lacking in power, suppose $\phi \approx 1$ so that Δx_t behaves essentially like v_t. The IVX-QR filtering rule with $R_w < \phi$ yields a w-variable that is less persistent than the x-variable (and this is why the IVX-QR method achieves better size control than the predictive quantile regression t-test based on x_{t-1}.) And therein lies the rub because this construction makes $\{w_t\}$ less variable than $\{x_t\}$, meaning that w_{t-1} gives less information about the quantile predictability of y_t than x_{t-1}.

To the best of our knowledge, the proposed LMC and MMC tests are the only ones available thus far in the literature for joint quantile predictability testing, so we evaluate their size and power in Tables 3 and 4 relative to one another. For this purpose, we define the "Tails" of the outcome distribution by the collection of quantile levels $\{0.1, 0.2, 0.8, 0.9\}$ and its "Center" by $\{0.3, 0.4, 0.6, 0.7\}$, and we use "All" to refer to the union of all these quantile levels. As before, we see from Table 3 that LMC tests can over-reject, while the empirical size of MMC tests never exceeds the 5% level threshold. Again, the LMC over-rejections are greatest when $\phi = 0.99$ and $|\rho|$ is high.

The results in Table 4 show once again that the (size-adjusted) LMC approach yields more powerful tests. Note, however, that the MMC tests perform remarkably well in comparison, especially for the larger sample size $T = 240$. It is also important to bear in mind that size-adjusted tests are not feasible in practice, but merely serve here as a benchmark for the truly exact MMC tests. We observe from Table 4 that the MMC_{min} tests perform better than their $\text{MMC}_{\times}$ counterparts, which appear to be lacking in power notably when $T = 120$. Finally, we notice that OLS yields ever so slightly more powerful MMC_{min} tests under normal errors, while LAD appears somewhat preferable under heavier-tailed errors.

Robustness assessment

We complete the simulation experiments with an assessment of robustness to departures from the assumptions maintained under (2). First, we examine what happens to the empirical size of the joint quantile predictability tests when the predictor variable follows an AR(2) process of the form:

$$x_t = \mu + \phi_1 x_{t-1} + \phi_2 x_{t-2} + v_t,$$

where we set $\phi_1 = 1.30$ and $\phi_2 = -0.31$. These parameters values are in line with what is found empirically with the monthly stock return predictors (used in the next section.)

Second, we investigate the size performance in the presence of GARCH effects appearing as

$$\begin{aligned} y_t &= \beta_0 + \sigma_t u_t, \\ \sigma_t^2 &= \omega + a_1(y_{t-1} - \beta_0)^2 + b_1\sigma_{t-1}^2, \end{aligned}$$

where $\omega = 0.01$, $a_1 = 0.10$, and $b_1 = 0.85$. As before, normal and t_3 errors are considered, but with the GARCH(1,1) specification we use the standardized errors $u_t^* = \sqrt{1/3}u_t$ so that u_t^* has unit variance and σ_t^2 is the conditional variance of y_t.

Table 5 shows the empirical size in percentage of the proposed tests under: (i) the AR(2) process with i.i.d. errors; (ii) the GARCH(1,1) process along side a correctly assumed AR(1) model (with ϕ set to 0.99); and (iii) the AR(2) and GARCH(1,1) processes together. In each case, the feedback parameter is set as $\rho = -0.95$. Comparing the LMC and MMC tests, we see immediately that the LMC tests are far more sensitive to departures from the maintained assumptions. In particular, GARCH errors are quite damaging for the size of the LMC tests. On the contrary, the conservative MMC tests offer more robustness to these violations of the assumptions. Indeed, the empirical size of the MMC tests in Table 5 is at worst around 10%. These values compare favourably to the empirical size of the t- and IVX-QR tests seen before in Table 1.

5 Empirical application

In this section, we further illustrate the new tests with an application to the monthly excess returns on the S&P value-weighted stock index from January 1948 to December 2015. We consider six predictors commonly used in the stock return predictability literature: dividend-price ratio (d/p), earnings-price ratio (e/p), book-to-market ratio (btm), default yield spread (dfy), term spread (tms), and short rate (tbl). These data series are a subset of those analyzed by Welch and Goyal (2008) and were obtained from Amit Goyal's website. The first three predictors (d/p, e/p, btm) are valuation ratios based on stock characteristics, while the last three (dfy, tms, tbl) are related to the interest rate; see Welch and Goyal (2008) for details.

Table 6 shows the estimation results obtained by applying OLS and LAD to the assumed AR(1) model for each of the six predictors. In addition to the point estimates $\hat{\mu}$ and $\hat{\phi}$, the table shows the conventional R^2 goodness-of-fit measure and an analogous (pseudo) R_p^2 developed by Koenker and Machado (1999) for quantile regressions. The numbers in parentheses are the standard errors of the parameter estimates and the numbers in square brackets show the $1 - \alpha_1 = 99\%$ confidence intervals for ϕ obtained by inverting the MC t^{OLS}- and t^{LAD}-test according to (18). The point estimates $\hat{\phi}$ are between 0.958 and 1.002, which shows that the predictors are highly persistent. In fact, except for the term spread, the (very tight) confidence intervals reveal that the unit root hypothesis cannot be rejected.

Following Cenesizoglu and Timmermann (2008), Maynard et al. (2011) and Lee (2016), we begin by testing for predictability at the individual quantile levels. For each predictor and a range of quantile levels from 0.05 to 0.95, Table 7 reports the predictive quantile regression estimates, along with associated standard errors in parentheses and the R_p^2 measure. The entries in bold are instances where the conventional t-test appears significant at the 5% level. Table 8 presents the p-values of these t-tests together with the p-values of the IVX-QR test of Lee (2016) and those of the proposed LMC and MMC tests. From the entries in bold, we can see a broad agreement among all the tests that the short rate (tbl) has predictive ability for the more central quantiles of the return distribution. There is also agreement that the default yield spread (dfy) is predictive of the quantiles in the right tail. However, there are major disagreements concerning the predictive ability of the other predictors. For example, the t- and IVX-QR tests indicate that dfy predicts the left tail, but that evidence is clearly rebuffed by the MC tests. The dividend-price (d/p) and book-to-market (btm) ratios also appear predictive of the left tail under the IVX-QR approach, but not so in light of the MC tests.

We turn next to an evaluation of joint quantile predictability. Table 9 shows the p-values of the LMC and MMC tests applied to the left-tail quantiles $\{0.05, 0.1, 0.2\}$, centre quantiles $\{0.3, 0.4, 0.5, 0.6, 0.7\}$, right-tail quantiles $\{0.8, 0.9, 0.95\}$, and all these quantiles levels. The short rate (tbl) clearly predicts the centre of the return distribution with p-values $\leq 2\%$ across the board. We see that the default yield spread (dfy) predicts the right tail also with p-values $\leq 2\%$. The evidence concerning left-tail predictability is not so clear cut. Recall from §3.3 that we reject H_0 if $p^* \leq \alpha_2$. Given that $\alpha_1 = 1\%$ and $\alpha_2 = 4\%$ here, we see that the short rate (tbl) is borderline predictive at the overall 5% level for the left tail of the return distribution, at least in light of the conservative MMC tests. Interestingly, this evidence is uncovered by the MMC tests based on LAD.

Looking back upon Table 7 we see that the short rate (tbl) enters the quantile regressions for the centre and left-tail quantiles with negative coefficients. And from the MMC test results in Table 9 we know that these coefficients are jointly significant. So increases in the short rate are associated with decreases in the central and lower quantiles of the return distribution, but there is no such effect on the upper quantiles. The MMC test results also confirm that the association between the default yield spread (dfy) and the right-tail quantiles is positive and significant. Consistent with Cenesizoglu and Timmermann (2008) and Maynard et al. (2011), the takeaway message is that the valuation ratios (dividend-price ratio, earnings-price ratio, book-to-market ratio) are insignificant for the most part, while the short rate and default yield spread contain useful information for predicting entire parts of the return distribution.

6 Conclusion

We have developed a procedure to test for quantile predictability that allows for the presence of an endogenous and persistent predictor variable along with heavy-tailed errors in both the predictive quantile regression model and the AR(1) model, assumed for the predictor variable. As far as we know, our approach is the only one that allows multiple quantile levels to be tested jointly. We achieve this by exploiting the technique of MMC tests, developed in Dufour (2006). Specifically, we combine an exact confidence set for the persistence parameter of the AR(1) model with conditional distribution-free MC tests of no quantile predictability that are linked to each point in the confidence set.

The confidence set is obtained by "inverting" a distribution-free MC t-ratio test for the AR(1) parameter, and the predictability tests are based on a data-dependent combination of quantile regression t-statistics, computed at each considered quantile level of the outcome distribution. This approach yields MMC permutation tests of the joint null hypothesis of no quantile predictability whose familywise error rate (probability of committing at least one Type I error) is kept under control no matter the sample size. To complement the results establishing the validity of our procedure, a simulation study made clear that the proposed tests deliver good discriminatory power in detecting quantile predictability. Finally, an empirical application revealed that entire parts of the distribution of stock returns are indeed predictable by variables related to the interest rate.

Appendix

The proof here closely follows that of Proposition 2 in Campbell and Dufour (1997). We wish to show that $\Pr\left(\sup_{\phi_0 \in C_\phi(\alpha_1)} \tilde{p}_B\left[\mathcal{S}(\phi_0)\right] \leq \alpha_2\right) \leq \alpha_1 + \alpha_2$. This will be true if $\Pr(A) \leq \alpha_1 + \alpha_2$, where A is the event $\tilde{p}_B\left[\mathcal{S}(\phi_0)\right] \leq \alpha_2$ for all $\phi_0 \in C_\phi(\alpha_1)$.

Define the set $I = \{\phi_0 : \phi_0 \in C_\phi(\alpha_1) \text{ and } \tilde{p}_B\left[\mathcal{S}(\phi_0)\right] > \alpha_2\}$. Then, via Bonferroni's inequality, we have that

$$\begin{aligned}
\Pr(\phi \in I) &= 1 - \Pr\left(\phi \notin C_\phi(\alpha_1) \text{ or } \tilde{p}_B\left[\mathcal{S}(\phi_0)\right] \leq \alpha_2\right]\right) \\
&\geq 1 - \Pr\left(\phi \notin C_\phi(\alpha_1)\right) - \Pr\left(\tilde{p}_B\left[\mathcal{S}(\phi_0)\right] \leq \alpha_2\right]\right) \\
&\geq 1 - \alpha_1 - \alpha_2,
\end{aligned}$$

since $\Pr\left(\phi \in C_\phi(\alpha_1)\right) \geq 1 - \alpha_1$ by definition of the confidence set for ϕ, and $\Pr\left(\tilde{p}_B\left[\mathcal{S}(\phi_0)\right] \leq \alpha_2\right]\right) = \alpha_2$ from (16). Observe that $\Pr(A) = \Pr(B^c)$, where B is the event $\tilde{p}_B\left[\mathcal{S}(\phi_0)\right] > \alpha_2$ for some $\phi_0 \in C_\phi(\alpha_1)$. Note also that $\phi \in I \Longrightarrow B$. Hence

$$\Pr(B) \geq \Pr(\phi \in I) \geq 1 - \alpha_1 - \alpha_2,$$

which implies the desired result: $\Pr(A) \leq \alpha_1 + \alpha_2$.

References

Abdelkhalek, T. and J.-M. Dufour (1998). Statistical inference for computable general equilibrium models, with application to a model of the Moroccan economy. *Review of Economics and Statistics 80*, 520–534.

Amihud, Y. and C. Hurvich (2004). Predictive regression: a reduced-bias estimation approach. *Journal of Financial and Quantitative Analysis 39*, 813–841.

Barnard, G. (1963). Comment on 'The spectral analysis of point processes' by M.S. Bartlett. *Journal of the Royal Statistical Society (Series B) 25*, 294.

Beaulieu, M.-C., J.-M. Dufour, and L. Khalaf (2007). Multivariate tests of mean-variance efficiency with possibly non-Gaussian errors. *Journal of Business and Economic Statistics 25*, 398–410.

Birnbaum, Z. (1974). Computers and unconventional test statistics. In F. Proschan and R. Serfling (Eds.), *Reliability and Biometry*, pp. 441–458. SIAM, Philadelphia.

Campbell, B. and J.-M. Dufour (1997). Exact non-parametric tests of orthogonality and random walk in the presence of a drift parameter. *International Economic Review 38*, 151–173.

Campbell, J. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*, 27–60.

Catani, P. and N. Ahlgren (2017). Combined Lagrange multiplier test for ARCH in vector autoregressive models. *Econometrics and Statistics 1*, 62–84.

Cavanagh, C., G. Elliott, and J. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory 11*, 1131–1147.

Cenesizoglu, T. and A. Timmermann (2008). Is the distribution of stock returns predictable? *SSRN Working Paper*.

Dufour, J.-M. (2006). Monte Carlo tests with nuisance parameters: a general approach to finite-sample inference and nonstandard asymptotics in econometrics. *Journal of Econometrics 133*, 443–477.

Dufour, J.-M. and L. Khalaf (2001). Monte Carlo test methods in econometrics. In B. Baltagi (Ed.), *A Companion to Theoretical Econometrics*, pp. 494–510. Basil Blackwell, Oxford, UK.

Dufour, J.-M., L. Khalaf, and M. Voia (2015). Finite-sample resampling-based combined hypothesis tests, with applications to serial correlation and predictability. *Communications in Statistics – Simulation and Computation, 44*, 2329–2347.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *Annals of Mathematical Statistics 28*, 181–187.

Fisher, R. (1932). *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.

Folks, J. (1984). Combination of independent tests. In P. Krishnaiah and P. Sen (Eds.), *Handbook of Statistics 4: Nonparametric Methods*, pp. 113–121. North-Holland, Amsterdam.

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

Koenker, R. and J. Machado (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association 94*(448), 1296–1310.

Lee, J. (2016). Predictive quantile regression with persistent covariates: IVX-QR approach. *Journal of Econometrics 192*, 105–118.

Lehmann, E. and J. Romano (2005). *Testing Statistical Hypotheses, Third Edition*. Springer, New York.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics 74*, 209–235.

Mankiw, N. and M. Shapiro (1986). Do we reject too often?: Small sample properties of tests of rational expectation models. *Economics Letters 20*, 139–145.

Maynard, A., K. Shimotsu, and Y. Wang (2011). Inference in predictive quantile regressions. *University of Guelph Working Paper*.

Nankervis, J. and N. Savin (1985). Testing the autoregressive parameter with the t statistic. *Journal of Econometrics 27*, 143–161.

Nankervis, J. and N. Savin (1988). The Student's t approximation in a stationary first order autoregressive model. *Econometrica 56*, 119–145.

Nankervis, J. and N. Savin (1996). The level and power of the bootstrap t test in the AR(1) model with trend. *Journal of Business and Economic Statistics 14*, 161–168.

Pearson, K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika 25*, 379–410.

Phillips, P. (2014). On confidence intervals for autoregressive roots and predictive regression. *Econometrica 82*, 1177–1195.

Polk, C., S. Thompson, and T. Vuolteenaho (2006). Cross-sectional forecasts of the equity premium. *Journal of Financial Economics 81*, 101–141.

Randles, R. and D. Wolfe (1979). *Introduction to the Theory of Nonparametric Statistics*. Wiley, New York.

Rayner, R. (1990). Bootstrapping p values and power in the first-order autoregression: a Monte Carlo investigation. *Journal of Business and Economic Statistics 8*, 251–263.

Savin, N. (1984). Multiple hypothesis testing. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, pp. 827–879. North-Holland, Amsterdam.

Schwert, W. (1989). Why does stock market volatility change over time? *Journal of Finance 44*, 1115–1153.

Stambaugh, R. (1999). Predictive regressions. *Journal of Financial Economics 54*, 375–421.

Tanaka, K. (1983). Asymptotic expansions associated with the AR(1) model with unknown mean. *Econometrica 51*, 1221–1231.

Taylor, S. (1986). *Modelling Financial Time Series*. John Wiley and Sons, Chichester, UK.

Tippett, L. (1931). *The Methods of Statistics*. Williams and Norgate, London.

Torous, W., R. Valkanov, and S. Yan (2004). On predicting stock returns with nearly integrated explanatory variables. *Journal of Business 77*, 937–966.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*, 1455–1508.

Westfall, P. and S. Young (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. Wiley, New York.

Wilkinson, B. (1951). A statistical consideration in psychological research. *Psychology Bulletin 48*, 156–158.

Table 1: Predictability tests at individual quantiles: empirical size

		Normal errors								t_3 errors							
	τ	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9
$\phi = 0.95$, $\rho = 0.00$																	
$T = 120$	LMC$^{\text{OLS}}$	5.5	4.8	4.3	5.4	4.6	4.9	5.3	3.8	4.9	5.2	4.6	4.6	4.8	5.4	4.6	5.1
	LMC$^{\text{LAD}}$	5.1	4.4	4.4	5.5	5.1	4.9	4.5	4.1	4.4	5.4	4.5	4.8	4.7	5.7	4.8	4.6
	MMC$^{\text{OLS}}$	1.1	1.1	0.3	1.5	0.9	1.0	1.2	1.1	1.1	1.1	1.2	1.1	1.6	1.0	1.7	0.9
	MMC$^{\text{LAD}}$	1.1	1.0	0.3	1.5	0.8	0.9	1.2	1.0	1.2	1.1	1.3	1.1	1.8	1.1	1.8	1.0
	t-test	12.0	8.4	7.8	7.8	6.9	7.3	8.8	10.1	15.2	9.5	6.6	6.4	6.8	8.1	9.9	13.1
	IVX-QR	12.0	8.5	8.9	7.9	6.0	7.5	8.3	12.5	15.7	11.0	8.1	8.6	7.0	8.1	9.6	14.5
$T = 240$	LMC$^{\text{OLS}}$	5.1	6.2	6.2	4.7	4.4	4.7	4.9	4.8	4.1	4.5	5.3	4.6	4.8	5.8	4.8	4.6
	LMC$^{\text{LAD}}$	5.1	6.2	5.3	5.4	4.9	4.8	5.2	4.3	4.0	4.4	5.5	5.8	4.9	5.5	5.2	4.4
	MMC$^{\text{OLS}}$	1.6	1.9	1.9	1.9	1.7	1.7	1.3	1.4	1.5	1.2	1.8	1.9	1.6	1.5	1.1	1.3
	MMC$^{\text{LAD}}$	1.5	1.9	1.8	1.6	1.6	1.7	1.2	1.4	1.5	1.3	1.8	2.0	1.6	1.6	1.1	1.2
	t-test	9.2	9.4	7.9	6.3	6.3	6.6	8.1	8.7	9.1	7.2	7.5	6.5	5.6	7.2	8.5	10.7
	IVX-QR	9.4	7.9	7.1	6.1	7.0	7.7	8.5	9.6	10.2	9.0	7.9	6.4	5.8	7.0	8.0	9.6
$\phi = 0.95$, $\rho =$ -0.90																	
$T = 120$	LMC$^{\text{OLS}}$	6.8	4.8	5.3	6.8	5.0	4.7	6.3	6.2	5.4	5.7	6.9	6.1	6.2	6.1	5.8	6.7
	LMC$^{\text{LAD}}$	6.4	4.7	5.9	6.4	5.8	4.0	5.7	6.0	4.8	5.8	6.1	5.7	6.9	5.7	5.2	5.8
	MMC$^{\text{OLS}}$	1.2	1.3	1.7	1.7	0.8	0.8	1.2	1.9	1.1	1.3	1.3	1.5	1.9	1.6	1.2	2.0
	MMC$^{\text{LAD}}$	1.2	1.3	1.5	1.7	0.6	0.7	1.1	1.8	1.1	1.3	1.3	1.5	2.1	1.6	1.4	2.0
	t-test	13.7	9.3	10.3	10.6	9.2	9.3	11.8	15.5	16.8	12.2	11.1	9.4	10.5	11.3	12.9	17.0
	IVX-QR	12.2	9.9	6.9	5.7	5.8	6.3	8.7	11.8	14.9	10.1	8.4	7.7	9.0	10.0	9.4	15.6
$T = 240$	LMC$^{\text{OLS}}$	4.9	5.5	5.4	5.3	5.9	6.0	5.3	5.6	4.5	5.8	4.9	4.8	5.7	4.9	4.8	5.4
	LMC$^{\text{LAD}}$	4.6	4.4	5.4	5.2	6.3	5.7	5.7	5.6	4.3	5.2	5.2	4.1	5.8	5.1	4.7	4.4
	MMC$^{\text{OLS}}$	1.2	1.0	1.5	1.5	1.9	1.3	1.5	1.4	0.8	1.0	1.1	0.9	1.9	1.4	1.3	1.3
	MMC$^{\text{LAD}}$	1.3	0.9	1.4	1.3	1.7	1.2	1.5	1.3	1.0	1.3	1.0	1.0	2.0	1.3	1.5	1.4
	t-test	9.8	8.2	7.7	8.4	8.8	10.1	9.0	12.4	10.7	9.9	8.3	6.3	7.3	7.6	8.5	11.7
	IVX-QR	9.8	8.3	5.9	4.5	5.7	7.1	7.7	9.8	10.1	7.2	7.3	7.8	6.3	7.8	6.9	9.7
$\phi = 0.99$, $\rho =$ -0.90																	
$T = 120$	LMC$^{\text{OLS}}$	6.4	7.6	10.7	10.2	9.4	9.4	8.1	6.9	6.5	6.7	7.6	7.3	7.7	7.9	6.6	7.0
	LMC$^{\text{LAD}}$	6.5	7.6	10.3	10.3	9.6	10.1	7.4	5.6	4.9	6.6	7.3	7.6	6.7	7.0	5.9	5.4
	MMC$^{\text{OLS}}$	1.9	1.7	2.5	3.5	2.3	2.2	1.9	1.9	1.4	1.4	2.0	1.2	2.0	2.0	1.2	1.6
	MMC$^{\text{LAD}}$	1.6	1.7	2.3	3.3	2.5	2.0	1.8	1.9	1.4	1.3	1.9	1.2	2.0	2.1	1.2	1.6
	t-test	17.6	15.8	17.7	15.6	16.9	16.9	16.0	16.6	19.9	15.3	13.1	13.5	13.1	13.9	16.7	20.2
	IVX-QR	12.3	8.3	7.9	7.5	6.4	7.8	10.2	13.0	12.2	9.2	6.6	5.1	6.1	6.4	9.5	14.8
$T = 240$	LMC$^{\text{OLS}}$	5.0	7.5	7.8	8.5	9.3	8.3	7.5	6.1	6.2	7.2	6.8	6.8	8.1	6.5	6.9	6.7
	LMC$^{\text{LAD}}$	5.7	6.9	7.5	7.7	8.7	8.7	7.2	6.8	5.6	7.0	6.4	6.9	8.3	6.3	6.1	6.4
	MMC$^{\text{OLS}}$	1.2	2.1	2.3	2.9	2.3	2.8	2.6	2.1	2.3	2.9	1.9	1.9	2.6	1.9	1.9	2.7
	MMC$^{\text{LAD}}$	1.2	2.3	2.2	2.8	2.1	2.8	2.6	2.1	2.4	3.1	2.0	1.9	2.7	1.9	2.0	2.8
	t-test	13.7	13.1	12.8	13.8	14.7	13.2	12.4	14.1	14.9	13.2	12.0	11.1	11.4	9.5	12.8	15.1
	IVX-QR	9.7	7.9	5.9	5.1	5.6	7.6	7.9	9.1	10.0	7.8	8.1	7.6	6.1	7.4	6.7	9.5
$\phi = 0.99$, $\rho =$ -0.95																	
$T = 120$	LMC$^{\text{OLS}}$	7.4	7.6	9.0	10.0	10.8	10.0	7.5	4.9	6.5	7.0	9.3	8.2	7.9	8.4	9.0	6.4
	LMC$^{\text{LAD}}$	6.8	7.9	9.0	9.3	10.3	10.7	7.6	5.0	6.5	6.4	8.4	8.3	7.4	7.5	6.8	6.1
	MMC$^{\text{OLS}}$	2.0	2.6	1.4	1.8	2.2	2.1	2.2	1.5	1.7	1.9	2.5	2.6	2.2	1.5	1.7	1.4
	MMC$^{\text{LAD}}$	1.8	2.6	1.4	1.9	2.2	2.1	2.1	1.4	1.6	2.0	2.7	2.5	2.2	1.5	1.8	1.5
	t-test	17.6	17.0	17.0	17.1	16.7	18.7	16.0	15.8	20.9	17.0	16.4	16.0	13.2	14.8	17.0	21.2
	IVX-QR	12.0	10.2	7.8	7.1	7.6	8.3	9.6	14.2	14.5	9.2	7.1	7.7	7.6	7.8	10.1	13.2
$T = 240$	LMC$^{\text{OLS}}$	5.5	7.1	7.9	8.3	9.4	9.5	7.3	6.1	7.2	8.2	8.0	7.3	7.4	6.3	7.4	7.3
	LMC$^{\text{LAD}}$	6.2	7.0	8.2	8.7	9.4	8.9	7.6	6.6	6.2	8.4	7.1	7.0	6.5	7.0	6.9	7.0
	MMC$^{\text{OLS}}$	2.1	1.2	2.4	2.2	2.3	2.6	2.5	2.3	2.1	2.9	2.4	2.2	2.2	2.2	2.8	2.5
	MMC$^{\text{LAD}}$	2.0	1.3	2.3	2.1	2.3	2.3	2.6	2.1	2.2	3.0	2.5	2.2	2.2	2.4	2.7	2.5
	t-test	13.7	13.4	13.6	14.6	14.4	14.5	12.4	14.1	16.7	15.4	13.8	11.7	11.4	12.9	12.4	14.9
	IVX-QR	8.8	7.1	5.9	5.4	5.6	7.3	6.1	10.7	10.5	7.5	7.4	7.2	7.3	7.1	6.6	10.1

Note: This table reports the empirical size (in percentage) of the proposed LMC and MMC tests, along with the standard t-test and the IVX-QR test of Lee (2016). The individual quantile predictability tests are performed with a nominal level $\alpha = 5\%$.

Table 2: Predictability tests at individual quantiles: power

		Normal errors								t_3 errors							
	τ	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9
$\phi = 0.95$, $\rho = 0.00$																	
$T = 120$	LMC^{OLS}	30.2	44.1	50.2	54.2	54.4	52.0	43.4	36.1	31.4	40.0	50.6	53.6	54.7	49.0	44.5	27.6
	LMC^{LAD}	30.5	44.9	50.9	54.0	51.8	52.3	44.5	35.2	31.2	40.0	50.3	53.8	55.6	48.5	44.5	29.6
	MMC^{OLS}	14.7	27.7	35.4	41.7	40.3	37.8	27.1	17.1	13.9	26.6	36.6	38.8	41.9	35.7	27.4	14.2
	MMC^{LAD}	14.7	27.5	35.2	41.6	40.2	37.2	27.2	16.9	14.3	26.9	36.7	38.7	42.4	35.9	27.4	14.4
	t-test	36.0	49.3	53.4	53.6	53.7	55.5	48.9	40.9	32.5	44.5	53.9	52.7	56.0	52.8	46.3	32.7
	IVX-QR	14.0	13.9	14.6	7.3	8.4	11.8	11.4	10.9	7.1	10.4	13.9	13.9	15.9	17.9	10.0	9.6
$T = 240$	LMC^{OLS}	33.1	43.7	46.0	55.7	57.7	51.9	47.3	35.0	33.0	48.2	50.9	57.5	58.6	51.8	46.8	34.4
	LMC^{LAD}	31.5	41.3	48.3	54.0	58.2	52.2	45.1	34.3	35.5	48.0	50.0	56.0	58.2	50.7	44.9	32.7
	MMC^{OLS}	19.6	32.0	38.8	43.5	46.9	37.9	32.3	19.3	18.1	32.2	40.5	46.9	46.9	39.5	31.1	16.7
	MMC^{LAD}	19.2	31.6	38.3	43.2	46.5	37.8	32.0	18.9	18.6	32.5	40.9	47.3	47.4	39.9	31.8	17.1
	t-test	36.2	44.4	50.1	55.6	59.7	54.7	47.6	37.6	36.8	50.7	53.3	57.4	59.7	54.3	49.8	36.2
	IVX-QR	12.0	15.2	12.4	11.5	10.2	12.1	12.5	14.3	13.1	15.4	18.8	17.8	18.9	18.7	18.3	14.2
$\phi = 0.95$, $\rho =$ -0.90																	
$T = 120$	LMC^{OLS}	24.8	40.7	45.9	48.2	46.2	44.3	35.0	27.7	25.2	39.1	45.1	47.5	49.5	43.1	35.3	22.8
	LMC^{LAD}	29.0	40.2	46.0	48.1	46.2	43.8	36.6	27.1	26.8	38.5	45.3	47.7	45.6	41.9	37.4	24.8
	MMC^{OLS}	14.1	23.7	32.1	37.3	34.3	27.9	23.1	14.8	12.9	25.8	34.3	38.5	34.7	30.4	23.6	12.1
	MMC^{LAD}	13.9	23.7	32.2	37.2	33.9	27.8	23.0	14.5	13.2	26.1	34.4	38.4	35.0	30.7	23.9	12.3
	t-test	33.4	44.5	50.4	48.4	48.0	48.7	39.8	31.6	29.5	41.7	50.4	50.5	49.1	48.7	41.7	28.7
	IVX-QR	11.2	11.3	10.8	9.8	10.6	13.3	11.1	11.7	9.9	12.4	17.1	19.6	16.8	17.0	14.7	10.3
$T = 240$	LMC^{OLS}	33.8	42.9	49.6	52.8	50.8	49.4	43.8	30.7	33.1	43.3	53.5	53.7	50.7	49.3	46.3	29.4
	LMC^{LAD}	35.8	45.5	48.8	52.5	49.6	48.2	44.7	30.2	33.4	44.3	49.7	55.4	53.9	46.9	44.9	31.3
	MMC^{OLS}	18.6	29.6	35.8	40.9	39.6	35.2	29.9	17.7	17.5	29.4	37.7	43.4	42.4	36.2	30.0	16.7
	MMC^{LAD}	18.5	29.6	35.5	40.5	39.4	35.1	29.5	17.3	18.0	29.3	37.9	43.7	42.6	36.9	30.6	16.8
	t-test	36.3	48.1	53.4	54.4	53.2	50.6	47.4	32.3	34.8	45.2	53.7	57.5	56.0	50.4	47.5	33.4
	IVX-QR	14.1	13.8	13.6	12.3	12.3	12.3	15.7	14.2	15.5	17.6	18.0	17.4	18.5	16.2	20.1	14.2
$\phi = 0.99$, $\rho =$ -0.90																	
$T = 120$	LMC^{OLS}	51.3	53.6	50.6	39.3	44.3	52.8	53.4	50.5	41.0	51.6	51.5	42.9	46.5	53.6	50.6	41.9
	LMC^{LAD}	48.2	52.4	50.4	37.5	43.3	52.8	56.3	48.1	43.7	52.3	49.6	40.4	45.3	54.6	52.1	44.8
	MMC^{OLS}	35.9	41.5	44.0	34.9	38.0	46.2	43.6	36.4	28.6	39.3	41.7	34.1	36.2	45.2	38.2	29.9
	MMC^{LAD}	35.4	41.3	44.1	35.0	38.1	46.0	43.8	36.2	28.7	39.4	41.8	34.2	36.4	45.5	38.5	30.3
	t-test	55.3	57.0	54.5	40.5	46.0	55.2	60.5	55.3	46.1	55.0	52.3	44.4	47.5	57.9	54.9	48.5
	IVX-QR	8.9	11.5	10.7	7.6	7.9	9.4	10.4	10.6	7.1	8.9	13.2	14.1	14.4	13.2	9.4	8.1
$T = 240$	LMC^{OLS}	61.4	63.4	62.5	54.1	54.1	59.2	63.5	57.9	55.2	63.0	66.0	58.5	55.8	64.4	64.6	54.5
	LMC^{LAD}	61.0	64.6	65.6	56.4	52.9	61.7	65.7	58.5	56.5	64.9	66.0	58.7	53.8	65.6	65.4	53.6
	MMC^{OLS}	49.8	59.5	58.9	49.5	50.3	58.9	60.1	50.2	46.4	60.4	59.5	52.2	50.1	59.6	57.4	45.7
	MMC^{LAD}	49.8	59.5	58.5	49.3	50.2	58.8	60.1	50.0	46.5	60.5	59.8	52.5	50.3	59.9	57.8	46.3
	t-test	63.4	66.9	67.1	56.6	55.5	64.3	66.7	62.1	56.6	67.4	70.2	60.7	57.3	67.9	67.0	56.9
	IVX-QR	11.5	11.8	14.1	10.7	11.5	10.3	14.1	12.4	12.0	13.2	12.6	11.9	15.0	13.6	13.8	9.4
$\phi = 0.99$, $\rho =$ -0.95																	
$T = 120$	LMC^{OLS}	48.7	57.7	53.9	43.2	38.6	46.5	56.4	54.0	38.9	50.0	48.1	42.3	46.7	52.9	46.8	41.8
	LMC^{LAD}	48.8	53.1	52.1	38.9	36.7	50.7	55.5	51.4	38.7	49.6	45.2	38.7	41.9	52.1	48.5	42.6
	MMC^{OLS}	36.9	44.3	45.5	36.2	35.1	44.5	44.0	34.4	27.6	37.8	39.7	33.1	36.0	42.9	38.7	26.8
	MMC^{LAD}	36.6	44.2	45.3	36.2	34.9	44.4	43.9	34.4	28.0	38.2	39.8	33.3	36.4	43.0	39.1	26.5
	t-test	53.8	60.9	58.2	43.3	42.3	52.8	59.3	58.2	40.5	52.0	50.0	41.8	46.8	56.1	52.0	42.5
	IVX-QR	11.3	7.9	9.8	9.5	8.2	11.3	11.1	9.6	7.3	9.7	12.4	9.8	12.1	15.0	9.5	8.5
$T = 240$	LMC^{OLS}	60.0	65.5	62.0	58.1	56.6	62.5	62.6	61.9	56.8	63.4	65.1	55.3	56.0	65.1	63.5	51.2
	LMC^{LAD}	57.3	65.4	62.4	54.8	54.3	62.4	62.0	59.4	59.4	63.5	63.9	59.0	57.8	64.8	63.5	53.4
	MMC^{OLS}	50.4	59.5	58.2	49.4	51.5	58.2	58.5	52.2	46.8	58.9	58.4	52.3	50.5	58.4	57.6	46.5
	MMC^{LAD}	49.9	59.4	58.1	49.2	51.5	58.1	58.6	51.8	47.3	59.2	58.5	52.4	50.4	58.4	57.9	46.8
	t-test	59.8	68.0	66.8	58.5	57.0	65.0	65.5	62.0	59.6	66.1	67.3	59.2	59.5	67.9	65.6	55.0
	IVX-QR	11.9	14.3	13.4	10.9	11.0	12.3	15.3	11.2	11.0	15.2	11.0	12.2	13.8	14.0	14.6	10.8

Note: The LMC, t-, and IVX-QR tests are based on size-adjusted critical values to ensure that their Type I error probabilities are equal to $\alpha = 5\%$ (the nominal level). Such adjustments were not applied to the MMC tests, since the probability of committing a Type I error with these test is $\leq \alpha$.

Table 3: Predictability tests at multiple quantiles: empirical size

		Normal errors			t_3 errors			Normal errors			t_3 errors		
		Centre	Tails	All	Centre	Tails	All	Centre	Tails	All	Centre	Tails	All
		$\phi = 0.95$, $\rho = 0$						$\phi = 0.95$, $\rho =$ -0.90					
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	4.4	3.9	3.7	5.3	4.9	4.7	5.5	6.5	6.3	7.0	5.8	6.4
	$\text{LMC}^{\text{OLS}}_{\times}$	4.5	4.2	5.1	5.4	5.0	4.7	5.8	5.8	6.5	5.9	5.4	6.8
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	5.0	4.9	4.3	5.1	4.4	4.7	6.1	5.9	6.2	6.8	5.1	5.7
	$\text{LMC}^{\text{LAD}}_{\times}$	4.7	4.7	4.7	4.8	4.6	4.7	5.5	5.9	6.6	6.6	5.8	6.4
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	0.7	1.2	1.2	1.6	1.6	1.6	1.5	1.8	2.2	1.6	1.1	1.3
	$\text{MMC}^{\text{OLS}}_{\times}$	0.9	0.8	0.6	1.2	1.5	1.1	1.4	1.6	0.9	1.8	1.3	1.1
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	0.7	1.1	1.1	1.6	1.5	1.5	1.5	1.8	2.2	1.6	1.1	1.3
	$\text{MMC}^{\text{LAD}}_{\times}$	0.8	0.8	0.6	1.2	1.5	1.1	1.3	1.6	1.0	1.9	1.3	1.2
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	5.8	5.6	5.8	5.5	4.5	4.8	6.2	5.1	6.2	4.7	5.5	4.8
	$\text{LMC}^{\text{OLS}}_{\times}$	4.6	5.3	5.5	4.9	4.0	4.2	5.9	5.8	6.3	5.2	4.3	5.2
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	5.8	5.8	5.6	6.3	3.8	5.1	6.3	5.1	5.2	4.9	4.9	4.4
	$\text{LMC}^{\text{LAD}}_{\times}$	5.4	5.4	4.6	5.2	4.4	4.0	6.0	5.3	6.3	5.5	4.5	6.0
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	1.5	1.9	1.7	1.9	1.5	1.4	1.5	1.2	1.1	1.1	1.2	0.9
	$\text{MMC}^{\text{OLS}}_{\times}$	1.6	1.5	1.5	1.9	1.0	1.0	2.0	1.6	1.8	1.4	1.2	0.7
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	1.5	1.9	1.6	2.0	1.5	1.5	1.5	1.2	1.1	1.2	1.2	0.9
	$\text{MMC}^{\text{LAD}}_{\times}$	1.4	1.5	1.5	1.9	1.1	1.0	1.8	1.6	1.8	1.5	1.1	0.9
		$\phi = 0.99$, $\rho =$ -0.90						$\phi = 0.99$, $\rho =$ -0.95					
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	11.6	7.2	7.3	7.4	5.3	5.4	11.2	7.9	8.8	9.0	7.2	7.2
	$\text{LMC}^{\text{OLS}}_{\times}$	12.7	8.7	11.4	8.0	6.0	8.3	12.4	8.9	10.0	10.5	8.4	9.6
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	10.6	6.6	7.1	7.0	5.2	5.6	10.8	8.1	9.1	7.7	5.8	6.5
	$\text{LMC}^{\text{LAD}}_{\times}$	11.5	8.6	10.9	8.1	5.2	6.6	13.3	8.2	10.3	9.9	7.3	9.7
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	2.3	2.0	2.1	2.1	0.9	1.1	2.2	1.4	1.5	1.5	1.2	1.3
	$\text{MMC}^{\text{OLS}}_{\times}$	2.4	2.1	1.8	1.4	0.9	0.7	1.8	1.6	1.8	2.3	1.2	1.2
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	2.3	2.0	2.1	2.1	1.0	1.2	2.1	1.4	1.5	1.7	1.0	1.4
	$\text{MMC}^{\text{LAD}}_{\times}$	2.3	2.1	1.9	1.5	0.9	0.8	1.8	1.7	1.4	2.4	1.1	1.2
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	9.1	6.8	7.3	7.8	7.9	9.0	10.4	6.7	7.8	7.7	8.1	8.7
	$\text{LMC}^{\text{OLS}}_{\times}$	10.0	6.8	9.6	8.4	7.4	8.4	10.2	8.0	10.0	8.3	9.1	9.9
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	9.0	6.0	7.2	7.5	7.0	7.9	10.2	7.2	8.7	8.0	7.2	7.5
	$\text{LMC}^{\text{LAD}}_{\times}$	10.3	7.3	9.4	7.6	8.1	8.6	10.0	7.9	9.2	8.2	8.6	9.4
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	2.1	1.7	1.7	2.5	2.7	2.7	2.5	1.8	2.0	2.7	2.7	2.4
	$\text{MMC}^{\text{OLS}}_{\times}$	2.8	1.8	2.4	2.5	2.4	2.3	2.1	2.2	2.5	3.0	2.3	2.5
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	2.0	1.6	1.4	2.8	2.7	2.7	2.5	1.7	1.9	2.8	2.7	2.4
	$\text{MMC}^{\text{LAD}}_{\times}$	2.6	1.8	2.4	2.5	2.4	2.4	2.2	2.2	2.4	2.9	2.5	2.6

Note: This table reports the empirical size (in percetnage) of the proposed LMC and MMC test for quantile predictability at multiple quantile levels. The joint quantile predictability tests are performed with a nominal level $\alpha = 5\%$.

Table 4: Predictability tests at multiple quantiles: power

		Normal errors			t_3 errors			Normal errors			t_3 errors		
		Centre	Tails	All	Centre	Tails	All	Centre	Tails	All	Centre	Tails	All
		$\phi = 0.95$, $\rho = 0$						$\phi = 0.95$, $\rho =$ -0.90					
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	75.4	45.9	67.5	71.2	43.5	60.7	67.7	38.2	58.5	68.7	39.3	53.7
	$\text{LMC}^{\text{OLS}}_{\times}$	75.8	53.4	70.3	70.5	48.3	67.9	68.0	44.8	57.6	67.8	45.7	55.0
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	73.3	45.1	64.9	72.0	42.3	60.9	68.5	41.7	58.0	68.5	39.1	56.8
	$\text{LMC}^{\text{LAD}}_{\times}$	76.4	53.0	71.8	73.1	50.7	68.7	67.0	47.9	57.3	67.5	45.5	54.2
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	61.0	26.8	42.0	58.0	22.4	36.5	52.2	24.5	37.0	55.4	19.8	32.4
	$\text{MMC}^{\text{OLS}}_{\times}$	60.6	35.3	33.8	59.2	28.4	19.4	51.6	27.9	7.3	52.8	20.0	3.6
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	60.6	26.7	41.3	58.2	22.6	37.3	52.3	24.5	36.6	55.5	20.3	33.1
	$\text{MMC}^{\text{LAD}}_{\times}$	60.4	35.1	32.4	59.9	28.8	20.5	51.4	27.8	7.0	53.3	20.4	4.3
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	75.3	48.1	66.4	75.3	49.1	70.8	71.7	47.1	62.8	76.7	46.1	69.7
	$\text{LMC}^{\text{OLS}}_{\times}$	75.1	52.8	70.4	76.0	56.1	75.1	69.6	53.3	65.8	73.5	54.2	71.2
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	75.6	47.9	65.6	73.5	48.8	68.4	70.6	48.1	65.9	76.0	48.2	70.5
	$\text{LMC}^{\text{LAD}}_{\times}$	73.5	53.0	72.3	73.7	57.5	72.8	70.7	53.3	65.7	73.2	54.9	71.4
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	63.5	32.1	52.0	65.4	31.8	53.4	59.0	31.1	48.6	63.7	28.2	51.0
	$\text{MMC}^{\text{OLS}}_{\times}$	61.9	41.0	60.4	65.1	40.6	60.4	55.8	37.6	46.2	60.2	36.6	46.8
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	63.3	32.0	51.4	66.0	32.6	53.7	58.7	30.5	48.4	64.0	28.8	51.8
	$\text{MMC}^{\text{LAD}}_{\times}$	61.6	41.0	60.2	65.3	41.1	60.8	55.7	37.5	45.6	60.7	37.4	48.3
		$\phi = 0.99$, $\rho =$ -0.90						$\phi = 0.99$, $\rho =$ -0.95					
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	71.7	62.5	73.7	67.4	60.0	68.9	71.0	59.4	73.2	66.6	57.2	67.6
	$\text{LMC}^{\text{OLS}}_{\times}$	65.6	64.4	45.7	67.8	64.8	61.3	64.5	65.1	58.0	66.2	60.6	52.2
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	66.2	61.2	71.7	69.4	57.0	65.8	69.5	57.9	66.8	64.9	55.1	64.8
	$\text{LMC}^{\text{LAD}}_{\times}$	65.2	64.0	45.4	66.0	63.1	55.2	63.5	69.0	46.0	65.3	58.2	46.4
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	62.7	48.0	57.3	59.5	39.5	47.3	63.1	50.5	58.0	58.2	35.7	44.2
	$\text{MMC}^{\text{OLS}}_{\times}$	63.7	54.3	18.6	60.7	39.9	12.0	61.1	52.8	17.6	56.6	36.6	9.6
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	62.6	47.8	56.9	59.6	40.0	47.8	63.0	50.1	57.7	58.2	35.9	44.5
	$\text{MMC}^{\text{LAD}}_{\times}$	63.4	54.1	17.6	60.9	40.0	13.1	61.1	52.5	16.6	56.7	37.1	10.3
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	80.0	70.7	80.6	80.5	67.2	76.8	80.8	71.9	78.5	80.2	70.8	78.6
	$\text{LMC}^{\text{OLS}}_{\times}$	79.4	74.7	77.3	79.9	71.5	74.5	79.8	75.3	77.0	80.1	73.2	76.0
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	79.1	70.5	81.0	81.8	68.9	77.1	80.8	72.2	79.6	79.6	70.7	80.4
	$\text{LMC}^{\text{LAD}}_{\times}$	79.3	75.0	76.6	78.2	71.9	71.6	79.0	74.9	76.2	79.4	72.3	72.2
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	76.7	63.4	75.0	76.5	62.8	72.9	76.9	65.0	75.9	77.5	63.9	74.5
	$\text{MMC}^{\text{OLS}}_{\times}$	76.4	69.4	66.1	76.3	69.0	63.5	76.4	69.5	64.7	76.9	69.6	60.8
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	76.7	63.2	74.9	76.5	63.0	73.0	76.8	64.6	75.7	77.4	64.1	74.7
	$\text{MMC}^{\text{LAD}}_{\times}$	76.4	69.1	65.7	76.5	69.1	64.1	76.3	69.5	64.5	76.9	69.6	61.4

Note: The LMC tests are based on size-adjusted critical values to ensure that their Type I error probabilities are equal to $\alpha = 5\%$ (the nominal level). Such adjustments were not applied to the MMC tests, since the probability of committing a Type I error with these test is $\leq \alpha$.

Table 5: Predictability tests at multiple quantiles: empirical size under departures from the maintained assumptions

		AR(2)			GARCH(1,1)			AR(2)+GARCH(1,1)		
		Centre	Tails	All	Centre	Tails	All	Centre	Tails	All
Normal errors										
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	8.0	6.6	6.9	14.3	14.3	14.7	10.9	10.6	11.3
	$\text{LMC}^{\text{OLS}}_{\times}$	8.3	7.3	8.5	12.4	16.1	17.4	10.6	13.0	13.1
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	8.1	6.3	6.7	14.0	13.7	15.5	11.9	9.7	9.8
	$\text{LMC}^{\text{LAD}}_{\times}$	8.8	7.1	8.1	13.4	15.1	16.3	11.4	12.1	12.5
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	2.0	1.2	1.6	3.1	4.6	4.8	2.7	3.4	3.6
	$\text{MMC}^{\text{OLS}}_{\times}$	2.2	1.5	1.2	3.0	4.6	4.3	2.1	4.2	3.3
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	2.1	1.2	1.6	3.1	4.5	4.7	2.8	3.5	3.7
	$\text{MMC}^{\text{LAD}}_{\times}$	2.3	1.6	1.3	2.9	4.6	4.2	2.2	4.2	3.2
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	6.2	7.0	7.0	11.3	14.5	14.4	9.6	12.3	13.5
	$\text{LMC}^{\text{OLS}}_{\times}$	6.4	6.9	7.4	11.6	17.5	16.8	9.9	14.7	13.7
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	6.1	6.5	6.8	12.3	15.4	15.4	9.4	12.0	12.2
	$\text{LMC}^{\text{LAD}}_{\times}$	7.3	6.4	7.4	12.1	16.4	16.6	9.7	14.7	13.3
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	1.9	2.6	2.3	4.3	5.9	5.7	4.4	6.0	6.1
	$\text{MMC}^{\text{OLS}}_{\times}$	2.3	2.0	2.3	3.4	6.5	6.0	3.5	7.5	5.6
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	1.9	2.7	2.3	4.4	5.9	5.6	4.4	6.0	6.2
	$\text{MMC}^{\text{LAD}}_{\times}$	2.3	1.9	2.3	3.4	6.4	6.2	3.5	7.5	5.6
t_3 errors										
$T = 120$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	8.0	6.8	6.9	14.9	14.6	15.3	12.4	13.0	13.5
	$\text{LMC}^{\text{OLS}}_{\times}$	7.7	7.9	8.6	13.7	16.8	17.2	11.7	14.8	15.9
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	6.9	6.2	5.8	13.7	13.7	14.7	12.6	11.5	12.6
	$\text{LMC}^{\text{LAD}}_{\times}$	7.6	6.3	8.0	13.3	15.5	17.0	11.6	13.3	15.0
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	1.6	2.7	2.9	5.6	4.9	5.4	4.6	6.4	6.4
	$\text{MMC}^{\text{OLS}}_{\times}$	1.6	2.8	2.0	4.2	6.1	3.4	3.8	6.4	4.4
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	1.8	3.0	3.2	5.8	5.0	5.5	4.7	6.5	6.5
	$\text{MMC}^{\text{LAD}}_{\times}$	1.8	3.1	2.4	4.3	6.1	3.4	4.0	6.6	5.1
$T = 240$	$\text{LMC}^{\text{OLS}}_{\text{min}}$	6.2	4.8	5.1	13.3	18.3	17.8	12.2	17.4	17.3
	$\text{LMC}^{\text{OLS}}_{\times}$	7.4	6.1	6.7	11.3	21.0	17.7	10.8	18.1	17.2
	$\text{LMC}^{\text{LAD}}_{\text{min}}$	6.0	4.6	4.4	11.9	16.8	16.2	13.0	16.5	16.6
	$\text{LMC}^{\text{LAD}}_{\times}$	7.1	5.6	6.9	11.3	18.1	16.5	10.1	18.3	15.9
	$\text{MMC}^{\text{OLS}}_{\text{min}}$	1.8	1.1	1.4	5.2	8.4	8.4	6.7	9.7	9.0
	$\text{MMC}^{\text{OLS}}_{\times}$	2.7	2.1	2.1	4.9	9.9	8.2	5.8	10.6	9.1
	$\text{MMC}^{\text{LAD}}_{\text{min}}$	1.8	1.2	1.4	5.1	8.5	8.7	7.0	10.3	9.3
	$\text{MMC}^{\text{LAD}}_{\times}$	2.7	2.1	2.1	5.0	10.1	8.5	6.0	10.7	9.4

Note: This table reports the empirical size (in percentage) of the proposed LMC and MMC tests under violations of the maintained assumptions, i.e. the AR(1) model assumption and the i.i.d. errors assumption. The joint quantile predictability tests are performed with a nominal level $\alpha = 5\%$.

Table 6: Estimation results for the monthly predictors

	OLS			LAD		
	$\hat{\mu}$	$\hat{\phi}$	R^2	$\hat{\mu}$	$\hat{\phi}$	R_p^2
d/p	-0.021	0.994	99.04%	-0.021	0.995	90.64%
	(0.012)	(0.003)		(0.011)	(0.003)	
		[0.982, 1.004]			[0.985, 1.007]	
e/p	-0.027	0.991	98.21%	-0.022	0.993	88.65%
	(0.013)	(0.005)		(0.012)	(0.004)	
		[0.975, 1.006]			[0.980, 1.005]	
btm	0.003	0.994	98.75%	-0.002	0.999	90.83%
	(0.002)	(0.004)		(0.002)	(0.003)	
		[0.981,0 1.009]			[0.988, 1.008]	
dfy	0.000	0.972	94.32%	0.000	0.983	79.76%
	(0.000)	(0.008)		(0.000)	(0.006)	
		[0.948, 1.004]			[0.967, 1.005]	
tms	0.001	0.958	91.73%	0.000	0.969	78.34%
	(0.000)	(0.010)		(0.000)	(0.006)	
		[0.933, 0.991]			[0.949, 0.983]	
tbl	0.000	0.992	98.27%	0.000	1.002	91.29%
	(0.000)	(0.005)		(0.000)	(0.001)	
		[0.977, 1.007]			[1.000, 1.005]	

Note: This table presents the OLS and LAD estimation results for the AR(1) model, assumed in (2). The sample period is January 1948–December 2015. For each predictor, the table shows point estimates and goodness-of-fit measures. The numbers in parentheses are standard errors of the estimates and the numbers in brackets show the 99% confidence intervals for ϕ obtained according to (18). Values < 0.001 are reported as zero.

Table 7: Predictive quantile regression estimates

τ	d/p	e/p	btm	dfy	tms	tbl
0.05	0.012	**0.014**	0.005	**-1.869**	0.165	-0.152
	(0.009)	(0.005)	(0.013)	(0.936)	(0.220)	(0.126)
	0.53%	1.27%	0.07%	1.62%	0.16%	0.44%
0.1	0.010	**0.015**	0.003	**-1.539**	0.340	**-0.212**
	(0.006)	(0.005)	(0.012)	(0.760)	(0.788)	(0.090)
	0.53%	0.98%	0.02%	1.14%	4.04%	0.85%
0.2	0.003	0.003	-0.001	-0.653	**0.412**	**-0.179**
	(0.005)	(0.005)	(0.009)	(0.548)	(0.129)	(0.057)
	0.07%	0.07%	0.00%	0.16%	0.73%	0.99%
0.3	0.006	0.003	-0.005	-0.443	0.164	**-0.203**
	(0.004)	(0.004)	(0.008)	(0.458)	(0.140)	(0.070)
	0.18%	0.11%	0.03%	0.08%	0.23%	0.80%
0.4	0.004	0.002	-0.009	-0.508	0.223	**-0.210**
	(0.005)	(0.005)	(0.007)	(0.423)	(0.131)	(0.054)
	0.09%	0.01%	0.12%	0.08%	0.36%	1.26%
0.5	0.003	0.002	-0.002	0.116	0.145	**-0.183**
	(0.004)	(0.003)	(0.006)	(0.337)	(0.106)	(0.053)
	0.08%	0.02%	0.02%	0.01%	0.19%	0.78%
0.6	**0.007**	0.003	0.005	0.443	0.183	**-0.118**
	(0.004)	(0.004)	(0.006)	(0.344)	(0.109)	(0.053)
	0.29%	0.11%	0.07%	0.14%	0.14%	0.46%
0.7	**0.009**	0.005	0.009	0.829	0.050	-0.092
	(0.004)	(0.004)	(0.007)	(0.449)	(0.129)	(0.064)
	0.50%	0.13%	0.17%	0.37%	0.04%	0.12%
0.8	**0.010**	0.004	0.012	**1.071**	0.108	-0.072
	(0.003)	(0.003)	(0.006)	(0.382)	(0.124)	(0.057)
	0.73%	0.18%	0.37%	0.73%	0.06%	0.11%
0.9	0.004	0.000	0.005	**1.756**	0.258	-0.031
	(0.006)	(0.005)	(0.011)	(0.475)	(0.179)	(0.077)
	0.13%	0.01%	0.03%	1.83%	0.45%	0.06%
0.95	0.010	-0.002	**0.025**	**2.293**	**0.465**	0.004
	(0.005)	(0.007)	(0.009)	(0.466)	(0.239)	(0.132)
	0.36%	0.09%	0.66%	2.78%	0.63%	0.00%

Note: This table shows the predictive quantile regression estimates for each considered predictor and quantile level. The numbers below the point estimates are the associated standard errors (in parentheses) and the R_p^2 goodness-of-fit measures (in percentages). Bold entries are instances where the conventional t-test appears significant at the 5% level.

Table 8: Individual quantile predictability

	τ	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
d/p	LMC$^{\text{OLS}}$	0.22	0.11	0.63	0.21	0.42	0.37	0.10	0.10	**0.02**	0.55	0.18
	LMC$^{\text{LAD}}$	0.25	0.12	0.62	0.22	0.46	0.39	0.10	0.11	**0.02**	0.56	0.20
	MMC$^{\text{OLS}}$	0.37	0.25	0.80	0.36	0.62	0.61	0.15	0.10	0.05	0.68	0.20
	MMC$^{\text{LAD}}$	0.37	0.25	0.80	0.36	0.62	0.61	0.15	0.10	0.05	0.68	0.20
	t-test	0.17	0.09	0.58	0.17	0.39	0.42	**0.04**	**0.02**	**0.00**	0.47	0.07
	IVX-QR	**0.00**	**0.00**	**0.02**	**0.00**	0.18	0.75	0.49	0.13	**0.03**	**0.00**	**0.00**
e/p	LMC$^{\text{OLS}}$	**0.04**	**0.04**	0.59	0.45	0.63	0.58	0.36	0.23	0.28	0.95	0.62
	LMC$^{\text{LAD}}$	0.05	**0.02**	0.60	0.45	0.62	0.63	0.35	0.24	0.30	0.96	0.63
	MMC$^{\text{OLS}}$	0.09	0.05	0.70	0.56	0.78	0.74	0.50	0.35	0.37	1.00	0.74
	MMC$^{\text{LAD}}$	0.09	0.05	0.70	0.56	0.78	0.74	0.50	0.35	0.37	1.00	0.72
	t-test	**0.01**	**0.00**	0.53	0.43	0.61	0.64	0.40	0.22	0.25	0.93	0.79
	IVX-QR	0.50	0.61	0.33	0.09	0.77	0.40	0.80	0.90	0.92	0.40	0.94
btm	LMC$^{\text{OLS}}$	0.82	0.63	0.92	0.56	0.24	0.76	0.49	0.25	0.13	0.72	**0.04**
	LMC$^{\text{LAD}}$	0.83	0.68	0.96	0.66	0.29	0.77	0.54	0.36	0.14	0.76	**0.03**
	MMC$^{\text{OLS}}$	0.89	0.81	0.99	0.74	0.40	0.88	0.64	0.36	0.16	0.87	0.07
	MMC$^{\text{LAD}}$	0.85	0.81	0.97	0.74	0.40	0.88	0.64	0.36	0.16	0.87	0.07
	t-test	0.70	0.80	0.87	0.50	0.17	0.77	0.46	0.19	0.06	0.65	**0.00**
	IVX-QR	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	0.29	0.63	0.05	**0.02**	**0.00**	**0.00**
dfy	LMC$^{\text{OLS}}$	0.13	0.05	0.25	0.38	0.22	0.67	0.17	0.08	**0.02**	**0.01**	**0.01**
	LMC$^{\text{LAD}}$	0.11	0.06	0.23	0.36	0.19	0.67	0.15	0.07	**0.01**	**0.01**	**0.01**
	MMC$^{\text{OLS}}$	0.18	0.13	0.35	0.43	0.33	0.79	0.28	0.14	**0.04**	**0.02**	**0.02**
	MMC$^{\text{LAD}}$	0.18	0.13	0.35	0.43	0.33	0.76	0.28	0.14	**0.04**	**0.02**	**0.02**
	t-test	**0.05**	**0.04**	0.23	0.33	0.23	0.73	0.20	0.07	**0.01**	**0.00**	**0.00**
	IVX-QR	**0.00**	**0.03**	**0.00**	0.51	0.25	0.84	0.38	**0.01**	0.09	0.06	**0.05**
tms	LMC$^{\text{OLS}}$	0.48	0.14	**0.02**	0.28	0.10	0.15	0.11	0.73	0.37	0.21	0.13
	LMC$^{\text{LAD}}$	0.46	0.16	**0.02**	0.30	0.10	0.15	0.12	0.74	0.32	0.19	0.11
	MMC$^{\text{OLS}}$	0.62	0.28	**0.03**	0.32	0.18	0.27	0.22	0.77	0.45	0.26	0.15
	MMC$^{\text{LAD}}$	0.60	0.21	**0.02**	0.32	0.18	0.24	0.16	0.75	0.44	0.26	0.15
	t-test	0.45	0.13	**0.00**	0.24	0.09	0.17	0.09	0.70	0.38	0.15	**0.05**
	IVX-QR	0.63	0.41	0.62	0.84	0.26	0.52	0.27	0.19	0.14	0.18	0.49
tbl	LMC$^{\text{OLS}}$	0.30	**0.04**	**0.02**	**0.01**	**0.01**	**0.01**	0.07	0.19	0.25	0.76	0.97
	LMC$^{\text{LAD}}$	0.28	**0.02**	**0.01**	**0.02**	**0.01**	**0.01**	**0.03**	0.16	0.20	0.73	0.93
	MMC$^{\text{OLS}}$	0.36	0.10	**0.03**	**0.03**	**0.02**	**0.02**	0.08	0.28	0.28	0.86	1.00
	MMC$^{\text{LAD}}$	0.26	**0.03**	**0.01**	**0.02**	**0.01**	**0.01**	**0.04**	0.20	0.17	0.79	0.97
	t-test	0.23	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.03**	0.15	0.20	0.68	0.98
	IVX-QR	0.75	0.57	**0.05**	**0.02**	**0.01**	**0.01**	**0.01**	0.08	**0.02**	0.06	0.04

Note: This table shows the individual p-values of the proposed LMC and MMC tests along with the standard t-test and the IVX-QR test of Lee (2016). The entries in bold are instances of statistical significance at the 5% level.

Table 9: Joint quantile predictability

	$\text{LMC}^{\text{OLS}}_{\min}$	$\text{LMC}^{\text{OLS}}_{\times}$	$\text{LMC}^{\text{LAD}}_{\min}$	$\text{LMC}^{\text{LAD}}_{\times}$	$\text{MMC}^{\text{OLS}}_{\min}$	$\text{MMC}^{\text{OLS}}_{\times}$	$\text{MMC}^{\text{LAD}}_{\min}$	$\text{MMC}^{\text{LAD}}_{\times}$
Panel A: Left tail								
d/p	0.25	0.20	0.28	0.25	0.79	0.78	0.79	0.78
e/p	0.06	**0.04**	**0.03**	**0.03**	0.12	0.10	0.12	0.10
btm	0.98	0.97	0.97	0.97	1.00	1.00	1.00	1.00
dfy	0.18	0.10	0.18	0.10	0.25	0.14	0.25	0.14
tms	**0.04**	0.05	0.05	0.07	0.06	0.09	0.06	0.09
tbl	0.06	0.06	**0.01**	**0.01**	0.07	0.06	**0.04**	**0.04**
Panel B: Centre								
d/p	0.17	0.15	0.16	0.15	0.64	0.72	0.64	0.72
e/p	0.43	0.40	0.46	0.41	0.83	0.81	0.83	0.81
btm	0.48	0.33	0.62	0.51	0.83	0.80	0.83	0.80
dfy	0.19	0.12	0.20	0.18	0.35	0.30	0.27	0.27
tms	0.31	0.15	0.28	0.14	0.40	0.33	0.36	0.24
tbl	**0.01**	**0.01**	**0.01**	**0.01**	**0.02**	**0.02**	**0.02**	**0.01**
Panel C: Right tail								
d/p	0.05	0.08	0.05	0.09	0.16	0.27	0.16	0.27
e/p	0.57	0.72	0.61	0.73	0.73	0.84	0.73	0.84
btm	0.06	0.07	0.07	0.10	0.16	0.22	0.16	0.22
dfy	**0.01**	**0.01**	**0.01**	**0.01**	**0.02**	**0.02**	**0.02**	**0.02**
tms	0.23	0.16	0.24	0.18	0.26	0.25	0.25	0.21
tbl	0.45	0.55	0.38	0.53	0.61	0.73	0.51	0.64
Panel D: All quantiles								
d/p	0.12	0.11	0.13	0.11	0.53	0.72	0.53	0.72
e/p	0.12	0.19	0.08	0.20	0.30	0.55	0.30	0.55
btm	0.15	0.27	0.18	0.38	0.39	0.80	0.39	0.80
dfy	**0.01**	**0.03**	**0.01**	**0.01**	**0.03**	**0.04**	**0.03**	**0.04**
tms	0.10	0.08	0.11	0.09	0.09	0.14	0.09	0.14
tbl	**0.03**	**0.01**	**0.01**	**0.01**	0.07	**0.03**	0.05	**0.02**

Note: This table shows the p-values of the proposed LMC and MMC tests for joint quantile predictability. Panel A is for quantile levels $\{0.05, 0.1, 0.2\}$; Panel B is for quantile levels $\{0.3, 0.4, 0.5, 0.6, 0.7\}$; Panel C is for quantile levels $\{0.8, 0.9, 0.95\}$; and Panel D is for the union of all these quantile levels. The predictors of the excess stock returns are the dividend-price ratio (d/p), earnings-price ratio (e/p), book-to-market ratio (btm), default yield (dfy), term spread (tms), and short rate (tbl). The entries in bold are instances of statistical significance at the overall 5% level.