Journal of Econometrics 249 (2025) 106002

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Predictive quantile regressions with persistent and heteroskedastic predictors: A powerful 2SLS testing approach☆

Matei Demetrescu [a], Paulo M.M. Rodrigues [b,*], A.M. Robert Taylor [c]

[a] Department of Statistics, TU Dortmund University, Germany
[b] Banco de Portugal and Nova School of Business and Economics, Portugal
[c] Essex Business School, University of Essex, United Kingdom

ARTICLE INFO

JEL classification:
C12
C22
G17

Keywords:
Predictive regression
Conditional quantile
Unknown persistence
Endogeneity
Time-varying volatility

ABSTRACT

We develop new tests for predictability at a given quantile, based on the Lagrange Multiplier [LM] principle, in the context of quantile regression [QR] models which allow for persistent and endogenous predictors driven by heteroskedastic errors. Of the extant predictive QR tests in the literature, only the moving blocks bootstrap implementation, due to Fan and Lee (2019) , of the Wald-type test of Lee (2016) can allow for conditionally heteroskedastic errors in the context of a QR model with persistent predictors. In common with all other tests in the literature these tests cannot, however, allow for unconditionally heteroskedastic behaviour in the errors. The LM-based approach we adopt in this paper is obtained from a simple auxiliary linear test regression which facilitates inference based on established instrumental variable methods. We demonstrate that, as a result, the tests we develop, based on either conventional or heteroskedasticity-consistent standard errors in the auxiliary regression, are robust under the null hypothesis of no predictability to conditional heteroskedasticity and to unconditional heteroskedasticity in the errors driving the predictors, with no need for bootstrap implementation. We also propose tests for joint predictability across a set of multiple distinct quantiles. Simulation results for both conditionally and unconditionally heteroskedastic errors highlight the superior finite sample properties of our proposed LM tests over the tests of Lee (2016) and Fan and Lee (2019) and the recent variable addition tests of Cai et al. (2023). An empirical application to the equity premium for the S&P 500 highlights the practical usefulness of our proposed tests, uncovering significant evidence of predictability in the left and right tails of the returns distribution for a number of predictors containing information on market or firm risk.

1. Introduction

Predictive regression is a widely used tool in applied finance and economics. A leading example concerns whether future stock returns can be predicted by current information. In this context, predictive regression methods have been extensively used in studies

☆ The authors thank the Co-Editor, Michael Jansson, an Associate Editor and two anonymous referees for their helpful and constructive comments. We also thank participants of the IAAE Annual conference 2023 at the BI Norwegian Business School, Oslo, Norway, and seminar participants at the University of the Balearic Islands, Palma de Maiorca, Spain for their comments. Demetrescu gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grant 531866675. Rodrigues thanks the Portuguese Science Foundation (FCT)for financial support through project PTDC/EGE-ECO/7493/2020, and (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209). Taylor gratefully acknowledges financial support provided by the U.K. Economic and Social Research Council under research grant ES/R00496X/1.

* Correspondence to: Economics and Research Department, Banco de Portugal, Av. Almirante Reis 71, 6th floor, 1150-012 Lisbon, Portugal.
E-mail address: pmrodrigues@bportugal.pt (P.M.M. Rodrigues).

https://doi.org/10.1016/j.jeconom.2025.106002
Received 7 March 2024; Received in revised form 27 March 2025; Accepted 30 March 2025
Available online 17 April 2025
0304-4076/

of mutual fund performance, tests of the conditional CAPM and studies of optimal asset allocation; see, for example, Paye and Timmermann (2006, pp. 274–275) and references therein.

Empirical evidence presented in, among others, Campbell and Yogo (2006) and Goyal and Welch (2003), suggests that many – though not all – of the predictors commonly considered are highly persistent with autoregressive roots close to unity, and that a strong negative correlation often exists between returns and the errors driving the predictors. Nelson and Kim (1993) and Stambaugh (1999) show that this causes a bias in standard ordinary least squares (OLS) coefficient estimates from the predictive regressions. As a result, a number of predictability tests have been developed in the literature which are designed to be asymptotically valid when the predictor is strongly persistent and endogenous; see, among others, Cavanagh et al. (1995), Campbell and Yogo (2006), Jansson and Moreira (2006), Maynard and Shimotsu (2009), Kostakis et al. (2015), Breitung and Demetrescu (2015), and Elliott et al. (2015).

The vast majority of the literature, typified by the references above, has focused on the question of whether the conditional mean depends on putative predictors or not. A recent strand of the predictability literature uses the quantile regression (QR) method of Koenker and Bassett (1978); see, among others, Lee (2016), Fan and Lee (2019), Ohno and Ando (2018), Meligkotsidou et al. (2014, 2021), Cai et al. (2023), Maynard et al. (2024) and Liu et al. (2024). This is an appealing concept in the context of return predictability as it allows us to examine whether the distribution of returns is predictable at non-central quantiles of the distribution, such as the shoulders and tails (see e.g. Meligkotsidou et al., 2014; Gungor and Luger, 2021 and Cederburg et al., 2022). Moreover, financial time series data often display heavy-tailed behaviour, and so one might expect potentially greater predictability at quantiles away from the median. Moreover, different predictor variables could potentially be simultaneously predictive at different quantiles of the distribution; for example, Gungor and Luger (2021) find that for excess returns on the S&P value-weighted stock market index, the default yield spread predicts the right tail of the distribution of returns while short-term interest rates are predictive for the centre of the return distribution.

It is well-documented in empirical finance that information from the tails of a return distribution plays a crucial role in measuring risk (see e.g. Nicolau et al., 2023). Bollerslev and Todorov (2014) find that the magnitude of the left jump tail associated with extreme market declines far exceeds that of the right jump tail corresponding to large market appreciations. They also find non-trivial predictable temporal dependencies in the tail index parameters characterising the decay in both tails. Kelly and Jiang (2014) reinforce the importance of time varying tail risks, and further argue that temporal variation in the tail parameters may help understand aggregate market returns as well as cross-sectional differences in average returns. Chevapatrakul et al. (2019) extend the work of Kelly and Jiang (2014), analysing asymmetries in the response of stock markets to tail risk. Using a QR approach they examine the effect of tail risk on the excess market returns at the different points of the return distribution. They argue that return predictability varies depending on where the return is located in the return distribution and that modelling risk and returns using QR yields significant improvement over conditional mean methods.

When predictors are persistent, predictive QR estimators are prone to similar bias and non-normality problems as occur with the OLS estimates from standard predictive regressions, discussed above; see, for example, Maynard et al. (2024). Consequently, analogous solutions are needed. Accordingly, for the case of testing for predictability at a single quantile, Maynard et al. (2024) follow the approach of Campbell and Yogo (2006) and use Bonferroni-based methods, under the assumption of stationary errors. More relevant to the present paper, Lee (2016) adapts the so-called extended instrumental variables (IVX) methodology developed for conditional mean predictability testing by Kostakis et al. (2015) to the QR setting. Using this approach, Lee (2016) develops Wald-type tests for predictability at a given single quantile which have standard limiting null distributions even with strongly persistent and endogenous predictors. Lee (2016) assumes that both the errors driving the predictors and the predictive regression errors are conditionally homoskedastic martingale difference [MD] sequences satisfying certain moment restrictions typical in the predictive regression literature, and with the predictive regression errors allowed to exhibit mild forms of heterogeneity in their conditional densities. Relatedly, and assuming essentially the same conditions on the errors as Lee (2016) and Cai et al. (2023) develop tests for predictability at a given quantile based on a weighted variable addition approach, expanding that discussed by Breitung and Demetrescu (2015) in the context of conditional mean predictability testing.

Fan and Lee (2019) investigate the behaviour of the Wald tests developed in Lee (2016) when the predictive regression errors are conditionally heteroskedastic. They show that the limiting null distribution of the IVX-based Wald-type tests in this case depends, in a rather intricate manner, on nuisance parameters arising from the conditional variances. As a result, they advocate using a moving blocks bootstrap (MBB) implementation of the Wald test and show that this delivers asymptotically pivotal inference under the null in the presence of conditional heteroskedasticity.

The tests of Lee (2016), Fan and Lee (2019) and Cai et al. (2023) all require the errors driving the predictors to be unconditionally homoskedastic. While it might in theory be possible to correct the asymptotic tests of Lee (2016) and Cai et al. (2023) to allow for unconditionally heteroskedastic errors, this is certainly not possible with the MBB approach of Fan and Lee (2019) as this resamples the data in such a way that the temporal ordering of the original data is not preserved in the bootstrap data and so will not, in general, be asymptotically valid in the presence of unconditionally heteroskedastic errors.[1] As we will show in our empirical application to S&P (excess) returns, while unconditional heteroskedasticity does not seem to be a great concern for many of the predictors traditionally used to test for conditional mean predictability in returns (of which no significant evidence is found for any

[1] A further test for predictability in quantile regressions, that also assumes unconditionally homoskedastic errors, has recently been proposed by Liu et al. (2024). However, this approach can only allow for at most one strongly persistent predictor in the predictive QR (see Remark 2 of Liu et al., 2024) and so is not directly comparable with the tests proposed in this paper or with the tests of Lee (2016), Fan and Lee (2019) and Cai et al. (2023). For this reason we will not discuss this test further.

predictor considered), it seems, perhaps unsurprisingly, to be an important data feature of many key measures of market risk for which, consistent with the earlier studies discussed above, our new tests yield significant predictability in the tails of the returns distribution.

In order to develop tests for predictability at a given quantile that are robust to strongly persistent endogenous predictors and to the presence of conditional and/or unconditional heteroskedasticity in the errors driving the predictors we will take a different route from Lee (2016), Fan and Lee (2019) and Cai et al. (2023) whilst still maintaining an approach based around the use of instrumental variable estimation. In particular, we will consider an approach based on applying the LM testing principle to the predictive QR. The LM test we propose is computationally simpler than the Wald test of Lee (2016), and can be obtained from a linear least-squares regression of the so-called generalised sign transform of the suitably centred dependent variable onto the set of putative lagged predictors and an intercept. As a result, conditional heteroskedasticity in the errors is considerably easier to deal with than it is in the testing framework of Lee (2016), requiring no bootstrap implementation. Our LM statistic will, however, suffer from the usual problem that its limiting null distribution would not be free of nuisance parameters if any of the regressors is strongly persistent and endogenous. To deal with this we estimate the linear regression by instrumental variable methods using the (over-identified) two stage least squares [2SLS] framework of Breitung and Demetrescu (2015) and Demetrescu et al. (2022). We demonstrate that this delivers statistics with pivotal χ^2 limiting null distributions regardless of the persistence and endogeneity strengths of the predictors. Indeed, this result holds even if the endogeneity correlation is time-varying.

As we will show in the finite sample simulations reported in Section 4, both the original Wald test of Lee (2016) and its MBB implementation display poor finite sample size control for some well-known models of conditional heteroskedasticity when the predictors are highly persistent. The size control of the variable addition tests of Cai et al. (2023) is overall better than the Wald-type tests but can still be rather oversized in the far tails of the distribution. In contrast, our LM-type tests display excellent size control even in the far tails. Moreover, we also find that the finite sample power of the Wald-type test of Lee (2016), its MBB implementation, and the variable addition test of Cai et al. (2023) are in general considerably lower than for the LM-based tests we propose.

The remainder of the paper is organised as follows. Section 2 details the heteroskedastic predictive QR model we consider, together with the assumptions needed for our analysis. In Section 3 we derive our LM-based heteroskedasticity-robust test for quantile predictability at a given quantile and establish its asymptotic null distribution. Here we also propose a joint test for predictability across a set of distinct quantiles. This would appear to be the first such multiple quantile test developed in the literature. A Monte Carlo comparison of the finite sample size and power properties of our proposed LM test with those of the test of Lee (2016), its MBB implementation and the variable addition test of Cai et al. (2023) is reported in Section 4. An empirical application to stock returns is reported in Section 5. Here we investigate quantile predictability for the equity premium using the (updated) database of predictors from Welch and Goyal (2008), which has been analysed widely in the conditional mean predictability literature, together with various measures of market risk in the spirit of Kelly and Jiang (2014). Section 6 concludes. An on-line supplementary appendix contains: a brief review of the QR predictability test of Lee (2016) and its MBB implementation developed in Fan and Lee (2019), providing corrections to implementation errors present in the MBB algorithm outlined in Fan and Lee (2019), and Cai et al. (2023); a proof of our main technical results; additional material relating to the empirical application in Section 5; additional Monte Carlo results referred to in Section 4; and some theoretical and Monte Carlo results pertaining to the cases of mixed persistence and cointegrated predictors.

2. The heteroskedastic predictive quantile regression model

Given a series y_t of stock returns, we follow Lee (2016), Fan and Lee (2019), and Cai et al. (2023), *inter alia*, and consider, for $\tau \in (0,1)$, the linear predictive QR model

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \alpha_\tau + \boldsymbol{\beta}_\tau' \boldsymbol{x}_{t-1}, \qquad t = 1, \dots, T,$$

where $Q_{y_t}(\tau|\mathcal{F}_{t-1})$ denotes the conditional τ-quantile of y_t, $\mathcal{F}_{t-1} = \{y_{t-1}, \boldsymbol{x}_{t-1}', y_{t-2}, \boldsymbol{x}_{t-2}', \dots\}$ denotes the natural filtration, and $\boldsymbol{x}_t := (x_{1,t}, \dots, x_{K,t})'$ is a K-vector of putative predictors with associated parameter vector $\boldsymbol{\beta}_\tau := (\beta_{1,\tau}, \dots, \beta_{K,\tau})'$. More generally, we denote the conditional τ-quantile of a series w_t by $Q_{w_t}(\tau|\mathcal{F}_{t-1})$, the marginal quantile of w_t by $Q_{w_t}(\tau)$, and the sample quantile of w_t (i.e. of the sample $\{w_1, \dots, w_T\}$) by $\hat{Q}_{w_t}(\tau)$. Below, we will make assumptions which ensure the uniqueness of the conditional quantile $Q_{y_t}(\tau|\mathcal{F}_{t-1})$ at any level τ, such that $\mathrm{P}\left[y_t \le Q_{y_t}(\tau|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}\right] = \tau$. For the present we will focus on conducting tests for predictability at a single quantile, τ. In Section 3.3 we will discuss joint tests for quantile predictability at a given set of distinct quantiles.

For the purposes of this paper, it will be convenient to express the model in terms of deviations from the conditional quantile, to obtain the equivalent multiple QR model with K regressors,

$$y_t = \alpha_\tau + \boldsymbol{\beta}_\tau' \boldsymbol{x}_{t-1} + u_{t,\tau}, \qquad t = 1, \dots, T, \tag{1}$$

where $u_{t,\tau} := y_t - Q_{y_t}(\tau|\mathcal{F}_{t-1})$ has zero conditional quantile of level $\tau \in (0,1)$ given $\mathcal{F}_{t-1}$; that is, $Q_{u_{t,\tau}}(\tau|\mathcal{F}_{t-1}) = 0$ and therefore the quantile regression error $u_{t,\tau}$ is not predictable at quantile τ. Our interest in this paper is in developing a test of the null hypothesis of no quantile-specific predictability, $H_0 : \boldsymbol{\beta}_\tau = \boldsymbol{0}$, against the two-sided alternative, $H_1 : \boldsymbol{\beta}_\tau \neq \boldsymbol{0}$ in the context of the predictive QR in (1).

We will make the following set of assumptions regarding the QR disturbances, $u_{t,\tau}$.

Assumption 1. The sequence of forecast errors, $u_{t,\tau}$, has zero conditional quantile and is strictly stationary and α (strongly) mixing, such that for some $r > 2$, $C > 0$ and $\kappa > 0$, the α-mixing coefficients satisfy $\alpha(m) \leq Cm^{-r/(r-2)-\kappa}$. Furthermore, the pdf of $u_{t,\tau}$ is bounded and absolutely continuous.

Remark 1. Assumption 1 is closely related to those parts of Assumption 3.1 of Fan and Lee (2019) which pertain to the QR disturbances, $u_{t,\tau}$. Fan and Lee (2019) additionally assume that the disturbances have a multiplicative structure: in their notation, $u_t = \sigma_t \epsilon_t$ with ϵ_t a zero-mean IID sequence of random variables and where σ_t^2 is the conditional variance of u_t given $\mathcal{F}_{t-1}$. Our setup also allows for such multiplicative structures, while not requiring the finiteness of conditional means and variances, for instance, that are imposed by Fan and Lee (2019). Indeed, we do not require finiteness to hold for any moments of the errors $u_{t,\tau}$. ◊

Remark 2. An implication of the strict stationarity assumption on $u_{t,\tau}$ made in Assumption 1 for a given τ, is that strict stationarity of $u_{t,\tau}$ must in fact hold for all $\tau \in (0,1)$. This is because, for any $\tau^* \neq \tau$, u_{t,τ^*} is given, by construction, as $u_{t,\tau} - Q_{u_{t,\tau}}(\tau^*|\mathcal{F}_{t-1})$, where the conditional quantile of $u_{t,\tau}$, $Q_{u_{t,\tau}}(\tau^*|\mathcal{F}_{t-1})$, must be strictly stationary itself, by virtue of the strict stationarity of $u_{t,\tau}$. ◊

The QR disturbances $u_{t,\tau}$ are measurable w.r.t. $\mathcal{F}_{t-1}$ and satisfy $\mathrm{P}\left[u_{t,\tau} \leq 0|\mathcal{F}_{t-1}\right] = \tau$. Then, denoting by ψ_τ the so-called generalised sign function,

$$\psi_\tau(u) := \begin{cases} \tau - 1 & u \leq 0 \\ \tau & u > 0, \end{cases} \tag{2}$$

it can easily be verified that the MD property, $\mathrm{E}(\psi_\tau\left(u_{t,\tau}\right)|\mathcal{F}_{t-1}) = 0$, holds on $s_{t,\tau} := \psi_\tau\left(u_{t,\tau}\right)$.

Remark 3. The MD property of $s_{t,\tau}$ characterises lack of quantile predictability of the errors $u_{t,\tau}$. To see this, notice that $\psi_\tau\left(u_{t,\tau}\right)$ represents the so-called generalised forecast errors for forecasting problems where the conditional quantile is the optimal forecast, and it is the MD property of generalised forecast errors that implies lack of predictability; see, for example, Granger (1999). ◊

The putative predictors are assumed to follow the additive components model,

$$\boldsymbol{x}_t = \boldsymbol{\mu}_x + \boldsymbol{\xi}_t, \tag{3}$$

where $\boldsymbol{\mu}_x \in \mathbb{R}^K$ is a vector of constants. The (zero-mean) stochastic component $\boldsymbol{\xi}_t := (\xi_{1,t}, \dots, \xi_{K,t})'$ is taken to have an autoregressive structure which allows one to control the degree of regressor persistence. We allow $\boldsymbol{x}_t$ to be either weakly or strongly persistent through the following assumption.

Assumption 2. Let

$$\boldsymbol{\xi}_t = \boldsymbol{\Gamma}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t, \qquad \boldsymbol{v}_t := (v_{1,t}, \dots, v_{K,t})', \tag{4}$$

where $\boldsymbol{\xi}_0$ is bounded in probability and exactly one of the following two conditions holds true:

1. **Weakly persistent predictors:** The predictors exhibit stable dynamics, i.e. $\left|\mathbf{I}_K - \lambda\boldsymbol{\Gamma}\right| = 0$ implies $|\lambda| > C$ for some $C > 1$, where $\mathbf{I}_K$ denotes the $K \times K$ identity matrix.
2. **Strongly persistent predictors:** The predictors exhibit near-integrated dynamics, i.e. $\boldsymbol{\Gamma} := \mathbf{I}_K - \frac{1}{T}\mathbf{C}$, where $\mathbf{C}$ is a real $K \times K$ matrix whose elements are fixed and finite.

Furthermore, let $\boldsymbol{v}_t \equiv \mathbf{B}(L)\tilde{\boldsymbol{v}}_t$ where $\tilde{\boldsymbol{v}}_t$ is zero-mean white noise, uncorrelated with $s_{\ell,\tau}$ for any $\ell < t$, and the lag polynomial $\mathbf{B}(L) := \sum_{j\geq 0}\mathbf{B}_j L^j$ has 1-summable coefficients with $\mathbf{B}(1) = \boldsymbol{\Omega}$ of full rank.

Remark 4. Assumption 2 allows the errors, $\boldsymbol{v}_t$, driving the predictors to exhibit (weak) serial dependence. Weak dependence in $\boldsymbol{v}_t$ is allowed for in a variety of ways in the literature. For instance, Fan and Lee (2019) adopt mixing conditions, while Lee (2016) and Cai et al. (2023) follow Kostakis et al. (2015) and work with linear processes driven by MD innovations (Lee, 2016 and Cai et al., 2023 both impose conditional homoskedasticity). More recently, Demetrescu et al. (2022, 2023a) maintain the linear process assumption with MD innovations but, importantly, additionally allow for unconditional heteroskedasticity. We follow Demetrescu et al. (2022, 2023a) in respect of the errors driving the predictors, cf. Assumptions 3 and 4 below, allowing us to draw on some of their theoretical results. ◊

Remark 5. Assumption 2 follows the bulk of the predictive regression literature in considering regressors that follow either stable (weakly dependent) processes (Assumption 2.1), or are near-integrated (Assumption 2.2). The latter allows for pure $I(1)$ predictors, locally stable predictors, and locally explosive predictors, without assuming knowledge of which of these holds. For clarity of exposition, and in common with Lee (2016) and Fan and Lee (2019), Assumption 2 imposes the condition that the predictors all belong to the same persistence class, such that they are either all weakly or all strongly persistent. In the strongly persistent case this does still allow for a mix of pure unit root, locally stationary, and locally explosive predictors. However, in section E of the supplementary appendix we show that, in common with the tests of Cai et al. (2023), our tests can also be validly used in cases where subsets of the predictors belong to different persistence classes (so that the set of predictors can contain a mix of (marginally) weakly and strongly persistent predictors). However, we also show that in the case where there are two or more strongly persistent predictors, these can be co-integrated such that linear combinations of the predictors exist that are weakly dependent, a scenario explicitly ruled out by Cai et al. (2023). Finite sample simulation results relating to the cases of mixed persistence and co-integrated predictors are also reported in section E of the supplementary appendix; these show that size of the LM-type tests developed in this paper remain well controlled in both of these scenarios. ◇

Remark 6. In the case of strongly persistent regressors, the near-unit root dynamics of the putative predictors, $\boldsymbol{x}_{t-1}$, is captured by the matrix of localisation coefficients, $\boldsymbol{\Gamma} := \mathbf{I}_K - \frac{1}{T}\mathbf{C}$. In contrast to Fan and Lee (2019), we do not restrict $\mathbf{C}$ to be a diagonal matrix. The assumption of a non-diagonal $\mathbf{C}$ implies that the normalised vector of predictors weakly converges to an unrestricted multivariate heteroskedastic Ornstein–Uhlenbeck [OU] processes (this would be a standard OU process in the case where $\boldsymbol{v}_t$ is unconditionally homoskedastic); see (7) below. To gauge the difference between the dynamics implied by diagonal and non-diagonal $\mathbf{C}$, suppose that the errors $\boldsymbol{v}_t$ driving the predictors have a diagonal long-run covariance matrix. In this case, a diagonal $\mathbf{C}$ implies that each predictor (when suitably normalised) weakly converges to a (marginal) heteroskedastic OU process (this reduces to a standard OU process in the case where $\boldsymbol{v}_t$ is unconditionally homoskedastic), where the marginal processes are necessarily independent. For non-diagonal C, and regardless of whether the long-run covariance matrix of $\boldsymbol{v}_t$ is diagonal or not, the marginal processes are not in general independent and do not possess a scalar heteroskedastic OU representation; they can, however, be represented by a combination of scalar (though not necessarily independent) heteroskedastic OU processes. As such, in the non-diagonal $\mathbf{C}$ case, the vector of predictors can potentially exhibit weak forms of equilibrium correction, whereby linear combinations of the predictors weakly converge (on normalisation) to a scalar heteroskedastic OU process whose non-centrality parameter is further from zero than at least one of predictors present in the linear combination. ◇

Assumption 2 allows for non-zero contemporaneous correlation between $\tilde{\boldsymbol{v}}_t$ and $s_{t,\tau}$. Recall that QR predictive regression endogeneity, while related to, is not identical to the concept of endogeneity in the case of conditional mean predictive regressions; see Lee (2016) and Maynard et al. (2024). In particular it is the contemporaneous correlation of $s_{t,\tau}$ and $\tilde{\boldsymbol{v}}_t$ that is relevant, and, just like in the case of conditional mean predictive regressions one may express this by means of a linear projection, viz., $s_{t,\tau} = \boldsymbol{\gamma}'_{t,\tau}\tilde{\boldsymbol{v}}_t + error$ at each time t, where $\boldsymbol{\gamma}_{t,\tau} := \left(\mathrm{Cov}\left(\tilde{\boldsymbol{v}}_t\right)\right)^{-1}\mathrm{Cov}\left(\tilde{\boldsymbol{v}}_t, s_{t,\tau}\right)$. This implies that endogeneity is not only quantile-specific, but may also be time-varying. It will be convenient in the following to reformulate this as a quantile-specific decomposition for $\tilde{\boldsymbol{v}}_t$,

$$\tilde{\boldsymbol{v}}_t = \mathbf{H}_{t,\tau}\boldsymbol{a}_{t,\tau} + \boldsymbol{h}_{t,\tau}s_{t,\tau}, \tag{5}$$

where $\boldsymbol{h}_{t,\tau} := (\tau(1-\tau))^{-1}\mathrm{Cov}\left(\tilde{\boldsymbol{v}}_t, s_{t,\tau}\right)$ and $\mathbf{H}_{t,\tau}$ satisfies $\mathbf{H}_{t,\tau}\mathbf{H}'_{t,\tau} = \mathrm{Cov}\left(\tilde{\boldsymbol{v}}_t - \boldsymbol{h}_{t,\tau}s_{t,\tau}\right)$, such that $\boldsymbol{a}_{t,\tau} = \mathbf{H}^{-1}_{t,\tau}\left(\tilde{\boldsymbol{v}}_t - \boldsymbol{h}_{t,\tau}s_{t,\tau}\right)$ is uncorrelated with $s_{t,\tau}$, and has identity covariance matrix, whenever the inverse exists. Notice that $\boldsymbol{a}_{t,\tau}$ is a white noise sequence under Assumptions 1 and 2. To match the setup of Demetrescu et al. (2022), we now state the set of conditions we require to hold on (5).

Assumption 3. For the decomposition $\tilde{\boldsymbol{v}}_t = \mathbf{H}_{t,\tau}\boldsymbol{a}_{t,\tau} + \boldsymbol{h}_{t,\tau}s_{t,\tau}$, where $\boldsymbol{a}_{t,\tau}$ is white noise with identity covariance matrix, and where $s_{t,\tau}$ and $\boldsymbol{a}_{t,\tau}$ are contemporaneously uncorrelated and satisfy Assumption 4 below, let $\mathbf{H}_{t,\tau} = \mathbf{H}_\tau(t/T)$ and $\boldsymbol{h}_{t,\tau} = \boldsymbol{h}_\tau(t/T)$, where $\mathbf{H}_\tau(\cdot)$ and $\boldsymbol{h}_\tau(\cdot)$ are, respectively, a matrix and vector of piecewise Lipschitz-continuous bounded functions on $(-\infty, 1]$, and where the matrix $[\mathbf{H}_\tau(\cdot); \boldsymbol{h}_\tau(\cdot)]$ is of full rank at all but a finite number of points.

Assumption 4. For some $\delta > 0$, let $\boldsymbol{a}_{t,\tau}$ be a uniformly $L_{4+\delta}$ bounded MD sequence with respect to the filtration $\mathcal{F}_{t-1}$. Furthermore, defining $\boldsymbol{\zeta}_t := \left((\tau(1-\tau))^{-1/2}s_{t,\tau}, \boldsymbol{a}'_{t,\tau}\right)'$, we assume that $\sup_t \mathrm{E}\left\|\mathrm{E}\left(\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_t - \mathbf{I}_{K+1}|\boldsymbol{\zeta}_{t-m}, \boldsymbol{\zeta}_{t-m-1}, \ldots\right)\right\| \to 0$, as $m \to \infty$. Finally, we assume that $\sup_{t\in\mathbb{Z}}\left\|\mathrm{E}\left(\left(\tilde{\boldsymbol{v}}_t\tilde{\boldsymbol{v}}'_t - \mathrm{E}\left(\tilde{\boldsymbol{v}}_t\tilde{\boldsymbol{v}}'_t\right)\right) \otimes \tilde{\boldsymbol{v}}_{t-j}\tilde{\boldsymbol{v}}'_{t-k}\right)\right\| \le C(jk)^{-1/2-\vartheta/2}$, for some $\vartheta > 0$.

Remark 7. Because $s_{t,\tau}$ is, by construction, bounded, Assumption 4 implies that $\left(\psi_\tau\left(u_{t,\tau}\right), \boldsymbol{a}'_{t,\tau}\right)'$ possesses the MD property and is uniformly L_4-bounded. The moment condition prevents, among other things, the innovations to the regressors from belonging to the infinite-variance class of processes. The QR errors $u_{t,\tau}$, on the other hand, need not even have finite expectation; cf. Remark 1. The decomposition of $\tilde{\boldsymbol{v}}_t$ in (5) allows one to combine these different moment properties in a convenient manner without imposing conditions directly on the joint pdf of $u_{t,\tau}$ and $\tilde{\boldsymbol{v}}_t$. ◇

Remark 8. In common with Assumption 3.1 of Fan and Lee (2019), Assumption 4 allows for the presence of conditional heteroskedasticity in the errors. In contrast, the assumptions adopted in Lee (2016) and Cai et al. (2023) impose conditional (and, hence, also unconditional) homoskedasticity on all of the error terms in the model. However, unlike Assumption 3.1 of Fan and Lee (2019), Assumption 3, additionally allows for unconditional heteroskedasticity in the errors driving the putative predictors. While, like Fan and Lee (2019), we impose finite moments of some order larger than four for the regressor innovations, the conditions we place on the short-run serial dependence in $\boldsymbol{v}_t$ (u_{tx} in the notation of Fan and Lee, 2019) are not directly comparable with the mixing conditions-based assumption from Fan and Lee (2019), as the mixing coefficients depend on the moments of u_{tx} whereas we require in Assumption 4 a summability condition on 4th order cross-product moments of the shocks to $\boldsymbol{v}_t$ paired with a 1-summability condition in Assumption 3 on the coefficients of the linear filter inducing serial correlation in $\boldsymbol{v}_t$. ◇

Remark 9. A nontrivial limitation imposed by our set of assumptions is that we do not allow for unconditional heteroskedasticity in the QR errors, $u_{t,\tau}$, as is clear from the fact that Assumption 1 requires them to be strictly stationary. Relaxing this assumption may be possible but would require the development of a rigorous analysis of adjusting using sample quantiles of non-stationary data, which is beyond the scope of this paper; cf. Remark 12. Certain forms of time-varying endogeneity (as captured by time-variation in $\boldsymbol{h}_\tau(\cdot)$) are nevertheless allowed for as they do not impact on the marginal strict stationarity of $u_{t,\tau}$. In order to explore the likely empirical performance of our tests in the case where the QR errors are unconditionally heteroskedastic, we include simulation results relating to this case in Tables D.11–D.14 of the supplementary appendix. ◊

Under Assumptions 1, 3 and 4, it follows from Boswijk et al. (2016, Lemma 1) that the following multivariate invariance principle applies,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor sT\rfloor}\begin{pmatrix} s_{t,\tau} \\ \boldsymbol{v}_t \end{pmatrix} \Rightarrow \begin{pmatrix} \sqrt{\tau(1-\tau)}W_\tau(s) \\ \boldsymbol{\Omega}\left(\int_0^s \mathbf{H}_\tau(r)\mathrm{d}\boldsymbol{W}(r) + \sqrt{\tau(1-\tau)}\int_0^s \boldsymbol{h}_\tau(r)\mathrm{d}W_\tau(r)\right)\end{pmatrix} =: \begin{pmatrix} \sqrt{\tau(1-\tau)}W_\tau(s) \\ \boldsymbol{M}(s)\end{pmatrix} \tag{6}$$

where "⇒" denotes weak convergence of the associated probability measures, and $\left(W_\tau(s), \boldsymbol{W}'(s)\right)'$ is a $(K+1)$-vector of independent standard Wiener processes.[2] Where time-invariant endogeneity is present, such that the vector of correlations between $\psi_\tau\left(u_{t,\tau}\right)$ and $\tilde{\boldsymbol{v}}_t$, and thus between $\psi_\tau\left(u_{t,\tau}\right)$ and $\boldsymbol{v}_t$, is constant and non-zero, then so $\boldsymbol{M}(s)$ and $W_\tau(s)$ will not be independent of one another. More generally, however, Assumption 3 allows these endogeneity correlations to be time-varying, since the marginal correlation of $\psi_\tau\left(u_{t,\tau}\right)$ and $\boldsymbol{v}_t$ is not assumed to be constant.

Under Assumption 2.2, where the predictors are strongly persistent, it follows from (6) that

$$\frac{1}{\sqrt{T}}\boldsymbol{\xi}_{\lfloor sT\rfloor} \Rightarrow \int_0^s e^{-\mathbf{C}(r-s)}\mathrm{d}\boldsymbol{M}(r) := \boldsymbol{J}_{\mathbf{C}}(r). \tag{7}$$

Notice, in the case where $\boldsymbol{v}_t$ is unconditionally homoskedastic, $\boldsymbol{M}(\cdot)$ is a (vector) Brownian motion process, such that $\boldsymbol{J}_{\mathbf{C}}(\cdot)$ is a (vector) OU process. More generally, we refer to $\boldsymbol{M}(\cdot)$ as a heteroskedastic Brownian motion and $\boldsymbol{J}_{\mathbf{C}}(\cdot)$ as a heteroskedastic OU process.

An implication of (7) is that, under Assumption 2.2, conventional statistics obtained from the QR in (1) will in general have limiting distributions that depend on $\mathbf{H}_\tau$ and $\boldsymbol{h}_\tau$. However, this dependence is not present in the weakly persistent case, where Assumption 2.1 holds, and so standard QR inference is recovered. In practice, the critical issue for inference purposes lies with the fact that the practitioner will not know whether the predictors are weakly or strongly persistent. In the next section we develop LM-type tests for quantile predictability that are robust to whether the predictors are weakly or strongly persistent and which allow for conditional and unconditional heteroskedasticity in the errors of the form specified in Assumptions 3 and 4.

3. LM tests for quantile predictability

We now introduce our proposed heteroskedasticity, and regressor persistence and endogeneity robust QR testing approach based on the LM principle. A considerable advantage of the LM approach over the tests developed in Lee (2016), Fan and Lee (2019) and Cai et al. (2023) (a brief review of these tests, including corrections to implementation errors present in the MBB algorithm outlined in Fan and Lee (2019), is provided in Part A of the supplementary appendix) is that it provides us with an auxiliary regression that is easy to robustify against the presence of conditional heteroskedasticity in the innovations (both the predictive regression errors and the errors driving the predictors) and to unconditional heteroskedasticity in the innovations driving the predictors. In contrast, the Wald-based tests require a numerically-intensive MBB implementation to accommodate conditional heteroskedasticity. Moreover, none of the Lee (2016), Fan and Lee (2019) or Cai et al. (2023) tests are valid, in general, if there is either unconditional heteroskedasticity in the innovations driving the predictors or if time-varying endogeneity is present.[3]

3.1. A quasi-likelihood argument

In order to develop a test of the null hypothesis of no quantile-specific predictability at a given quantile, τ, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$, against the two-sided alternative, $H_1 : \boldsymbol{\beta}_\tau \neq \mathbf{0}$ in the context of (1), first recall that QR can be motivated in a likelihood framework under the assumption that the QR disturbances $u_{t,\tau}$ follow an asymmetric Laplace distribution.[4] Assuming the scale parameter of the

---

[2] Observe from the quantile-specific decomposition in (5) that $\mathrm{Cov}\left(\tilde{\boldsymbol{v}}_t\right)$ may be expressed as $\mathbf{H}_{t,\tau}\mathbf{H}'_{t,\tau} + \tau(1-\tau)\boldsymbol{h}_{t,\tau}\boldsymbol{h}'_{t,\tau}$. Given the definition of $\boldsymbol{h}_{t,\tau}, \mathbf{H}_{t,\tau}$, this covariance matrix can be seen not to depend on τ, from which it follows that the (marginal) weak limit of the normalised partial sum of $\tilde{\boldsymbol{v}}_t$ (and, hence of $\boldsymbol{v}_t$, denoted by $\boldsymbol{M}(s)$ in (6)) also does not depend on τ.

[3] Indeed, Lee (2016, Section 3.3) advocates a specific implementation of the IVX Wald procedure when testing the specific null hypothesis of no predictability, based on running a standard QR of the (quantile-adjusted) dependent variable on the IVX instruments directly rather than a computationally more cumbersome IV QR. This is similar in spirit to the approach we outline in this section in that it conveniently exploits the null restriction. However, the resulting test statistic of no predictability is not motivated by the LM principle, and, as will be seen in the numerical results we subsequently present, inherits many of the disadvantages of the unrestricted Wald IVX QR approach, including lack of robustness to unconditional heteroskedasticity and poor finite sample size control at extreme quantiles.

[4] This property is shared by all members of the family of so-called tick-exponential distributions generalising the asymmetric Laplace distribution in this respect; see Komunjer (2005). We derive our statistics using the asymmetric Laplace distribution for ease of exposition.

asymmetric Laplace distribution characterising $u_{t,\tau}$ to be known, maximising the resulting log-likelihood reduces to the minimisation problem,

$$\left(\hat{\alpha}_\tau, \hat{\boldsymbol{\beta}}_\tau\right)' := \underset{\alpha,\boldsymbol{\beta}}{\arg\min} \left\{-\ell\left(\alpha, \boldsymbol{\beta}\right)\right\},$$

where the log-likelihood, $\ell(\alpha, \boldsymbol{\beta})$, is given by $\ell(\alpha, \boldsymbol{\beta}) = C_1 - C_2 \sum_{t=1}^{T} \rho_\tau\left(y_t - \alpha - \boldsymbol{\beta}'\boldsymbol{x}_{t-1}\right)$, in which $C_i, i = 1, 2$ are suitable constants, and ρ_τ is the usual quantile check function, $\rho_\tau(u) := u\psi_\tau(u)$, where $\psi_\tau(u)$ is as defined in (2).

Formally, the gradient of the log-likelihood is given by

$$\nabla\ell\left(\alpha, \boldsymbol{\beta}\right) := -C_2 \sum_{t=1}^{T} \begin{pmatrix} 1 \\ \boldsymbol{x}_{t-1} \end{pmatrix} \psi_\tau\left(y_t - \alpha - \boldsymbol{\beta}'\boldsymbol{x}_{t-1}\right). \tag{8}$$

Standard properties of the likelihood require this gradient to have zero expectation when evaluated at the true parameter values, α_τ and $\boldsymbol{\beta}_\tau$. Clearly, $\ell(\cdot,\cdot)$ is not differentiable due to the kink in the quantile check function ρ_τ at the origin. However, taking the zero expectation property of $\nabla\ell(\alpha_\tau, \boldsymbol{\beta}_\tau)$ at face-value will allow us to obtain a test statistic for our inferential problem.

The LM principle exploits the zero expectation of the gradient of the log-likelihood under correct specification. As the gradient should have zero expectation under the null when plugging in the null restriction $\boldsymbol{\beta}_\tau = \boldsymbol{0}$, the sample gradient in (8) should not be significantly different from zero when the null restriction (i.e. no predictability) holds true, and we construct such a test based on this observation. To this end, we will require an estimate of α_τ obtained under the null hypothesis, $H_0 : \boldsymbol{\beta}_\tau = \boldsymbol{0}$. This restricted estimator of α_τ, say $\hat{\alpha}_{\tau,0}$, is obtained from

$$\hat{\alpha}_{\tau,0} := \underset{\alpha|\boldsymbol{\beta}=\boldsymbol{0}}{\arg\min} \left\{-\ell\left(\alpha, \boldsymbol{\beta}\right)\right\} = \underset{\alpha}{\arg\min} \sum_{t=1}^{T} \rho_\tau\left(y_t - \alpha\right).$$

As observed in Koenker and Bassett (1978), the solution $\hat{\alpha}_{\tau,0}$ to this minimisation problem is the sample τ-quantile of y_t, $\hat{Q}_{y_t}(\tau)$. Next define the generalised sign transform of the sample-quantile adjusted series to be predicted, viz.,

$$\tilde{s}_{t,\tau} := \psi_\tau\left(y_t - \hat{\alpha}_{\tau,0}\right) = \psi_\tau\left(\tilde{y}_{t,\tau}\right).$$

Evaluated at the null of no predictability, the gradient of the negative log-likelihood then delivers

$$-\nabla\ell\left(\hat{Q}_{y_t}(\tau), \boldsymbol{0}\right) = C_2 \sum_{t=1}^{T} \begin{pmatrix} \tilde{s}_{t,\tau} \\ \boldsymbol{x}_{t-1}\tilde{s}_{t,\tau} \end{pmatrix},$$

which should not differ significantly from zero under the null hypothesis of no quantile predictability. It is known that $\tilde{s}_{t,\tau}$ averages to (almost) zero by construction, such that only the sample moments $\sum_{t=1}^{T} \boldsymbol{x}_{t-1}\tilde{s}_{t,\tau}$ are relevant for the testing problem and an LM test simply checks whether they are significantly different from zero.

To interpret the moment conditions derived from the LM principle, recall that the sample quantile is shift-equivariant. Consequently, under the null hypothesis $y_t = \alpha_\tau + u_{t,\tau}$, we have that

$$\tilde{s}_{t,\tau} = \psi_\tau\left(y_t - \hat{Q}_{y_t}(\tau)\right) = \psi_\tau\left(u_{t,\tau} - \hat{Q}_{u_{t,\tau}}(\tau)\right).$$

Furthermore, since the sample quantile is consistent for the τ-quantile of $u_{t,\tau}$, $Q_{u_{t,\tau}}(\tau)$, under Assumption 1 (see the proof of Proposition 1), we have that $\hat{Q}_{u_{t,\tau}}(\tau) \xrightarrow{p} 0$, by virtue of the fact that the τ-quantile of $u_{t,\tau}$, $Q_{u_{t,\tau}}(\tau)$, is zero by construction. Therefore, $\tilde{s}_{t,\tau} \approx s_{t,\tau} = \psi_\tau(u_{t,\tau})$ under the null hypothesis, where, at the same time, the conditional expectation of $\psi_\tau(u_{t,\tau})$ is zero. In other words, under the null, the generalised sign transform of the de-quantiled y_t is not predictable, implying that y_t is not predictable at quantile τ, and an LM test by design checks this using suitable (restricted) sample moment conditions.[5] Moreover, it is shown in the proof of Proposition 1 that $\tilde{s}_{t,\tau} = s_{t,\tau} - 1/T\sum_{t=1}^{T} s_{t,\tau} + O_p(T^{-1/2}) \approx s_{t,\tau} - \bar{s}_\tau$. Consequently, noting that the sample moments, $\sum_{t=1}^{T} \boldsymbol{x}_{t-1}\tilde{s}_{t,\tau}$, are in fact cross-product moments, we may equivalently test the null hypothesis of no predictability $\boldsymbol{\beta}_\tau = \boldsymbol{0}$ in the QR in (1) by testing the null hypothesis that $\boldsymbol{\delta}_\tau = \boldsymbol{0}$ in the least-squares [LS] auxiliary regression,

$$\tilde{s}_{t,\tau} = \boldsymbol{\delta}'_\tau\left(\boldsymbol{x}_{t-1} - \bar{\boldsymbol{x}}\right) + error \tag{9}$$

with $\bar{\boldsymbol{x}}$ denoting the sample average of $\boldsymbol{x}_{t-1}$. Of course, LS estimation of (9) will suffer from the same problems arising from endogeneity and uncertain predictor persistence, discussed in Section 1, as in the conditional mean predictive regression setting. We now address this issue, developing feasible LM-type tests.

[5] Notice that the Wald-type tests outlined in part A of the supplementary appendix are also based on moment conditions. These, however, are based on the *unrestricted* moment conditions $\sum_{t=1}^{T} \boldsymbol{z}_{I,t-1}\psi_\tau\left(y_t - \alpha_\tau - \boldsymbol{\beta}'_\tau\boldsymbol{x}_{t-1}\right) = \boldsymbol{0}$, with $\boldsymbol{z}_{I,t-1}$ denoting the IVX instruments defined in Section 3.2, to obtain estimators of the parameters α_τ and $\boldsymbol{\beta}_\tau$, and thence to test the null hypothesis, $\boldsymbol{\beta}_\tau = \boldsymbol{0}$.

3.2. A feasible LM-type test

A considerable advantage of working with a LS auxiliary regression is that we may draw on existing robust methods developed in the conditional mean predictive regression setting. In particular, we propose using the over-identified 2SLS inference based approach introduced by Breitung and Demetrescu (2015) and further developed by Demetrescu et al. (2022) to conduct inference for the auxiliary regression (9). It would also be feasible to base our feasible LM tests on the just-identified IVX estimation approach of Kostakis et al. (2015), as is done by Lee (2016) and Fan and Lee (2019) in their Wald-based quantile predictability tests. We choose to focus on 2SLS-based inference, incorporating both type-I and type-II instruments (defined below), because it is known from the results reported in Breitung and Demetrescu (2015) and Demetrescu et al. (2022), relating to conditional mean predictability testing, that 2SLS-based tests display superior power properties to methods that use only IVX (type-I) instruments when the candidate predictors are strongly persistent. This is because the type-II instruments considered in Demetrescu et al. (2022) are strong instruments in the case of strongly persistent regressors and dominate the corresponding IVX instrument asymptotically; the situation reverses for weakly persistent regressors, as here the IVX instrument is strong and asymptotically dominates the type-II instrument. The finite sample power simulations, reported in Section 4.1.2, uncover the same pattern in the QR predictability setting, with our proposed 2SLS-based tests considerably out-performing the just-identified IVX-based tests of Lee (2016) and Fan and Lee (2019) for strongly persistent regressors, but with no discernable power loss relative to these tests for weakly persistent regressors.

Following Breitung and Demetrescu (2015), we will therefore use 2SLS to estimate the regression of $\tilde{s}_{t,\tau}$ on $\boldsymbol{x}_{t-1}$, instrumenting $\boldsymbol{x}_{t-1}$ by the combination of so-called type-I and type-II instruments. The former are constructed to be of lower persistence than $\boldsymbol{x}_t$ where $\boldsymbol{x}_t$ is strongly persistent, while the latter are constructed to be exogenous with respect to the predictive regression error.

For the type-I instruments we will employ the mildly integrated IVX instrument of Kostakis et al. (2015) whereby a vector of instrumental variables, $\boldsymbol{z}_{I,t}$ for $\boldsymbol{x}_t$, is constructed as

$$\boldsymbol{z}_{I,t} := \sum_{j=0}^{t-1} \varrho^j \Delta \boldsymbol{x}_{t-j}, \quad \text{with} \quad \varrho := 1 - \frac{a}{T^\eta}, \tag{10}$$

for some $a > 0$ and $\eta \in (0,1)$, and initialised at $\boldsymbol{z}_{I,0} = 0$. The IVX scale and exponent parameters, a and η respectively, are tuning parameters set by the practitioner. Kostakis et al. (2015) recommend setting $a = 1$ and $\eta = 0.95$. Where $\boldsymbol{x}_t$ is near-integrated, satisfying Assumption 2.2, $\boldsymbol{z}_{I,t}$ in (10) is approximately a mildly integrated process and is therefore of lower persistence than $\boldsymbol{x}_t$. Moreover, where $\boldsymbol{x}_t$ is weakly dependent, satisfying Assumption 2.1, we have that $\boldsymbol{z}_{I,t} \approx \boldsymbol{x}_t$. As a result in the standard conditional mean predictive regression setting, Kostakis et al. (2015) demonstrate that the IVX full-sample estimator of the slope parameter in the predictive regression is asymptotically (mixed) Gaussian under the null hypothesis, regardless of whether the predictors are weakly or strongly persistent, and that consequently, the full-sample instrumental variable tests for the null of no mean predictability have standard limiting null distributions regardless of the degree of persistence or endogeneity of $\boldsymbol{x}_t$.

For the type-II instruments, we follow Breitung and Demetrescu (2015) and use,

$$z_{II,t,k} = \sin\left(\frac{\omega_k t}{T}\right), \text{ for } k = 1, \dots, K, \tag{11}$$

where the ω_k, $k = 1, \dots, K$, are distinct spectral frequencies.

These choices for the type-I and type-II instruments satisfy the general conditions for instrument validity stated in Assumptions 3 and 4, respectively, of Breitung and Demetrescu (2015), and also satisfy the somewhat weaker validity conditions given in Assumptions 4 and 5, respectively, of Demetrescu et al. (2022).[6] The validity of basing our tests on this choice of instruments is formally established in the proof of Proposition 1 given in the supplementary appendix.

For this choice of instruments, defining $\boldsymbol{z}_{I,t}$ as in (10) and $\boldsymbol{z}_{II,t} := (z_{II,t,1}, \dots, z_{II,t,K})'$, the 2SLS auxiliary estimator is then given by

$$\hat{\boldsymbol{\delta}}_\tau := \left(\mathbf{A}_T' \mathbf{B}_T^{-1} \mathbf{A}_T\right)^{-1} \mathbf{A}_T' \mathbf{B}_T^{-1} \mathbf{C}_T, \tag{12}$$

where $\mathbf{A}_T := \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1} \tilde{\boldsymbol{x}}_{t-1}'$, $\mathbf{B}_T := \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1} \tilde{\boldsymbol{z}}_{t-1}'$, $\mathbf{C}_T := \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1} \tilde{s}_{t,\tau}$, $\boldsymbol{z}_{t-1} := \left(\boldsymbol{z}_{I,t-1}', \boldsymbol{z}_{II,t-1}'\right)'$, and $\tilde{\boldsymbol{x}}_t = \boldsymbol{x}_t - \frac{1}{T}\sum_{s=1}^T \boldsymbol{x}_s$ as well as $\tilde{\boldsymbol{z}}_t = \boldsymbol{z}_t - \frac{1}{T}\sum_{s=1}^T \boldsymbol{z}_s$. Following Kostakis et al. (2015), we do not need to demean the IVX instrument vector $\boldsymbol{z}_{I,t-1}$ given that it is invariant to $\boldsymbol{\mu}_x$ by construction; for convenience we thus set $\tilde{\boldsymbol{z}}_{I,t-1} = \boldsymbol{z}_{I,t-1}$. Furthermore, as we will show below, $\tilde{s}_{t,\tau}$ is implicitly demeaned when removing the sample quantile prior to applying the generalised sign transform, and so demeaning them again in the auxiliary regression is also redundant.

Based on (12), an LM-type test for the null of no predictability at quantile τ, implemented with an Eicker–White heteroskedasticity-consistent (HC) covariance matrix estimation then obtains as

$$\mathcal{T}_\tau := \hat{\boldsymbol{\delta}}_\tau' \widehat{\mathrm{Cov}}_{HC}\left(\hat{\boldsymbol{\delta}}_\tau\right)^{-1} \hat{\boldsymbol{\delta}}_\tau, \tag{13}$$

[6] For simplicity we focus attention on these specific choices for the instruments because they are the choice of instruments recommended by Demetrescu et al. (2022). Like Demetrescu et al. (2022), we found in preliminary simulation experiments that these choices gave a good size-power trade-off and were somewhat superior on size control to the other examples of possible type-I and type-II instruments discussed in Breitung and Demetrescu (2015) and Demetrescu et al. (2022).

where the HC covariance matrix estimator is given by

$$\widehat{\mathrm{Cov}_{HC}}\left(\hat{\boldsymbol{\delta}}_\tau\right) := \left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_T\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}, \tag{14}$$

with $\mathbf{D}_T := \sum_{t=1}^T \tilde{\mathbf{z}}_{t-1}\tilde{\mathbf{z}}_{t-1}'\tilde{s}_{t,\tau}^2$. Observe that $\mathcal{T}_\tau = \mathbf{C}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_T\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{C}_T$.

Remark 10. Although we have used Eicker–White HC standard errors in constructing our proposed test, $\mathcal{T}_\tau$ in (13), the use of these is in fact not strictly necessary because $s_{t,\tau} = \psi_\tau(u_{t,\tau})$ is conditionally homoskedastic, given $\mathcal{F}_{t-1}$. The key to this result is that $s_{t,\tau}$ follows a conditional two-point distribution characterised only by τ, and so it is both conditionally and unconditionally homoskedastic. Given the conditional homoskedasticity of $s_{t,\tau}$, the standard covariance matrix estimator,

$$\widehat{\mathrm{Cov}}\left(\hat{\boldsymbol{\delta}}_\tau\right) = \hat{\sigma}_{\tilde{s}_{t,\tau}}^2\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}, \text{ with } \hat{\sigma}_{\tilde{s}_{t,\tau}}^2 := \frac{1}{T}\sum_{t=1}^T \tilde{s}_{t,\tau}^2, \tag{15}$$

is asymptotically equivalent to the HC estimator. For this reason we will not distinguish between versions of $\mathcal{T}_\tau$ constructed with the HC, as in (13), or with the standard covariance matrix estimator in place of the HC estimator when discussing the large sample properties of $\mathcal{T}_\tau$ in Proposition 1 below. This asymptotic equivalence does not, however, hold in finite samples or for the case of joint testing across multiple quantiles discussed in Section 3.3. ♢

In Proposition 1, whose proof is provided in Part B of the supplementary appendix, we now establish the limiting null distribution of our proposed LM-type statistic, $\mathcal{T}_\tau$, in (13).

Proposition 1. *Let the data $(y_t, \boldsymbol{x}_t')$ be generated according to (1) and let Assumptions 1–4 hold. Then under the null hypothesis, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$, it follows that, as $T \to \infty$, $\mathcal{T}_\tau \xrightarrow{d} \chi_K^2$.*

Remark 11. The result in Proposition 1 can be seen to hold regardless of whether the predictors are weakly or strongly persistent and as such implies that conventional critical values can be used without knowledge of the degree of persistence of the predictors. ♢

Remark 12. Recalling that $\tilde{s}_{t,\tau}$ is a generated regressand, it is necessary in establishing the limiting distribution of $\mathcal{T}_\tau$ to assess the impact of removing the sample quantile under the null hypothesis. It turns out that, under the conditions of Proposition 1, $\tilde{s}_{t,\tau} = \psi_\tau(u_{t,\tau}) - \frac{1}{T}\sum_{t=1}^T \psi_\tau(u_{t,\tau}) + R_{t,T}$, where the term $R_{t,T}$ can be controlled for in the relevant sums; see the proof of Proposition 1. This derivation relies on strict stationarity of $u_{t,\tau}$ which is therefore critical in establishing the result in Proposition 1. In particular, unconditional heteroskedasticity in the QR errors, $u_{t,\tau}$, would imply a different behaviour of the sample quantile (see Portnoy, 1991) compared to the strictly stationary case, which would in turn affect the behaviour of $\tilde{s}_{t,\tau}$. ♢

Remark 13. The $\mathcal{T}_\tau$ statistic in (13) is used to test the joint significance of all of the predictors $x_{1,t-1}, \ldots, x_{K,t-1}$. Analogous tests for the significance of (proper) subsets of the predictors can also be developed. To that end, define the $q \times K$ full row rank matrix $\mathbf{R}$ of constants defining q linearly independent restrictions on $\boldsymbol{\beta}$. The LM-type statistic for testing the null hypothesis $\mathbf{R}\boldsymbol{\beta} = \mathbf{0}$ is then given by $\mathcal{T}_\tau^{\mathbf{R}} := (\mathbf{R}\hat{\boldsymbol{\delta}}_\tau)'[\mathbf{R}\widehat{\mathrm{Cov}_{HC}}(\hat{\boldsymbol{\delta}}_\tau)\mathbf{R}']^{-1}(\mathbf{R}\hat{\boldsymbol{\delta}}_\tau)$. A (squared) t-type ratio for testing the significance of any given predictor, $x_{i,t-1}$, $i \in \{1, \ldots, K\}$, obtains by setting $\mathbf{R}$ equal to the $1\times k$ selection (row) vector whose ith element is equal to unity and all other elements are equal to zero. Under the conditions of Proposition 1, $\mathcal{T}_\tau^{\mathbf{R}}$ will have a χ_q^2 limiting null distribution. ♢

3.3. Testing for predictability at multiple quantiles

Thus far, in common with the tests developed in Lee (2016), Fan and Lee (2019) and Cai et al. (2023), we have focussed on tests of quantile predictability at a single user-specified quantile. In practice, however, researchers often run quantile predictive regressions across a range of distinct quantiles and it is possible that this could lead to multiple testing issues. Within the LS regression framework developed in Section 3.1 it is straightforward to develop size-controlled LM-type tests of joint predictability across multiple distinct quantiles. This reduces to testing zero restrictions in a system of seemingly unrelated equations. We now outline how this can be done and demonstrate that standard critical values again apply to the resulting statistic.

Consider a vector $\boldsymbol{\tau} := (\tau_1, \ldots, \tau_m)'$ of m (where m is finite) distinct quantile levels each corresponding to a distinct multiple QR of the form in (1) with intercept and slope parameters, α_{τ_j} and $\boldsymbol{\beta}_{\tau_j}$, $j = 1, \ldots, m$. Then construct,

$$\tilde{\mathbf{s}}_{t,\boldsymbol{\tau}} := \begin{pmatrix} \psi_{\tau_1}\left(y_t - \hat{Q}_{y_t}(\tau_1)\right) \\ \vdots \\ \psi_{\tau_m}\left(y_t - \hat{Q}_{y_t}(\tau_m)\right) \end{pmatrix} = \begin{pmatrix} \tilde{s}_{t,\tau_1} \\ \vdots \\ \tilde{s}_{t,\tau_m} \end{pmatrix},$$

and consider the multivariate auxiliary regression,

$$\tilde{\mathbf{s}}_{t,\boldsymbol{\tau}} = \boldsymbol{\Upsilon}_{\boldsymbol{\tau}}'\left(\boldsymbol{x}_{t-1} - \bar{\boldsymbol{x}}\right) + error,$$

where the $K \times m$ matrix of coefficients $\boldsymbol{\Upsilon}_{\boldsymbol{\tau}}$ contains the m quantile-specific coefficients $\boldsymbol{\delta}_{\tau_1}, \ldots, \boldsymbol{\delta}_{\tau_m}$ (from the m corresponding auxiliary regressions of the form given in (9)) columnwise, and is zero under the (joint) null hypothesis of no predictability at any of the quantiles $\tau_1, \ldots, \tau_m$.

Each equation of this system of seemingly unrelated regressions has the same set of regressors, so system GLS would be equivalent to equation-by-equation OLS estimation. As with the single quantile case considered in Section 3.2, our feasible tests will be based on equation-by-equation 2SLS estimation. Stacking the resulting vector of m individual equation estimates yields the system estimate $\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}} := \left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{C}_T$, where $\mathbf{C}_T = \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{s}}_{t,\boldsymbol{\tau}}'$.

While it would be tempting to base a multiple quantile LM-type test on the standard covariance matrix estimate, $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{s}}_{\boldsymbol{\tau}}} \otimes \left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}$, where $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{s}}_{\boldsymbol{\tau}}}$ is the sample covariance matrix of $\tilde{\boldsymbol{s}}_{t,\boldsymbol{\tau}}$, this would not be valid under our assumptions. In particular, our assumptions allow for the case where the conditional *covariance* between s_{t,τ_i} and s_{t,τ_j} is not constant for some $i \neq j$. We will therefore require a HC covariance matrix estimate, which we denote by $\widehat{\mathrm{Cov}_{HC}\left(\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}\right)}$, whose (i,j)th $K \times K$ block is given as, $\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_{T,ij}\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}$, with $\mathbf{D}_{T,ij} := \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau_i}\tilde{s}_{t,\tau_j}$.

Based on these quantities, our multiple-quantile LM-type predictability test statistic can be defined as

$$\mathcal{T}_{\boldsymbol{\tau}} := \left(\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}\right)'\left(\widehat{\mathrm{Cov}_{HC}\left(\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}\right)}\right)^{-1}\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}. \tag{16}$$

We conclude this section by detailing the limiting null distribution of $\mathcal{T}_{\boldsymbol{\tau}}$. The proof of this result is a straightforward extension of the result given in Proposition 1.

Proposition 2. *Let the conditions of Proposition 1 hold with Assumptions 1, 2 and 4 taken to hold jointly on the quantile-specific decompositions for $\tilde{\boldsymbol{v}}_t$ of the generic form in (5) across $\tau_1, \ldots, \tau_m$. Then under the null hypothesis, $H_0 : \boldsymbol{\beta}_{\tau_1} = \cdots = \boldsymbol{\beta}_{\tau_m} = \mathbf{0}$, it follows that, as $T \to \infty$, $\mathcal{T}_{\boldsymbol{\tau}} \xrightarrow{d} \chi^2_{mK}$.*

Remark 14. Under $H_0 : \boldsymbol{\beta}_{\tau_1} = \cdots = \boldsymbol{\beta}_{\tau_m} = \mathbf{0}$, the additional condition imposed by Proposition 2, requiring Assumption 1 to hold jointly across the m quantiles of interest, $\tau_1, \ldots, \tau_m$, implies that $Q_{y_t}(\tau_j|\mathcal{F}_{t-1}) = \alpha_{\tau_j}$ for all $\tau_j \in \{\tau_1, \ldots, \tau_m\}$. In other words, all of the relevant conditional quantiles of the dependent variable y_t are not predictable under the null. While this is the natural multiple-quantile null hypothesis to consider here, it imposes m null restrictions (rather than just one as in the single-quantile testing case) on the conditional distribution of y_t. One straightforward case where these restrictions are met, irrespective of the actual levels $\tau_1, \ldots, \tau_m$ of interest, is when y_t is an IID sequence. Beyond the IID case, there may exist DGPs for which y_t has different predictability behaviour at different quantiles, such that the conditions of Proposition 2 are violated. To see how, consider the series $y_t = w_t - Q_{w_t}(\tau_1|\mathcal{F}_{t-1})$ where w_t is strictly stationary and has unique conditional quantiles at any level. Then, the conditional τ_1-quantile of y_t is zero, and so we may set $\alpha_{\tau_1} = 0$ and $u_{t,\tau_1} = y_t$ to match Assumption 1 for the specific quantile level τ_1. Notwithstanding this, the conditional τ_2-quantile of y_t (for $\tau_1 \neq \tau_2$) is given, by virtue of the shift equivariance of quantiles, by $Q_{y_t}(\tau_2|\mathcal{F}_{t-1}) = Q_{w_t}(\tau_2|\mathcal{F}_{t-1}) - Q_{w_t}(\tau_1|\mathcal{F}_{t-1})$. This difference is only constant (and as such $u_{t,\tau_2} = y_t - \alpha_{\tau_2}$ is only compliant with our joint null hypothesis) when the two conditional quantiles of w_t are such that any time variation present in them is *parallel*, meaning that it cancels out on taking their difference. It is not, however, uncommon in quantile regression settings to have DGPs with non-parallel conditional quantile functions; see e.g. Koenker and Bassett (1978). The fact that the regularity conditions needed for the asymptotic validity of joint tests across multiple quantiles are more restrictive than for the single quantile case is intrinsic to the quantile regression framework and not a product of our specific choice of testing procedure. ◊

Remark 15. Where predictability holds at some quantile level(s) but not at others, individual (or joint) tests solely at the latter will be asymptotically size controlled. This is because each distinct quantile LM statistic is based on whether the generalised sign transform of the quantile-specific forecast errors are predictable or not, and this property is not affected by whether or not predictability holds at any other distinct quantile(s). This implies that should the joint predictability test $\mathcal{T}_{\boldsymbol{\tau}}$ reject, then one could identify at which of these quantiles predictability is most likely to hold by investigating the corresponding set of single quantile statistics, $\mathcal{T}_{\tau_1}, \ldots, \mathcal{T}_{\tau_m}$. ◊

Simulation experiments investigating the finite sample performance of the multiple quantile $\mathcal{T}_{\boldsymbol{\tau}}$ test across the $m = 11$ distinct quantile levels $\boldsymbol{\tau} = (0.05, 0.1, 0.2, \ldots, 0.9, 0.95)'$ are reported in Tables D.17 and D.18 in the supplementary appendix. These report the empirical rejection frequencies of $\mathcal{T}_{\boldsymbol{\tau}}$ for data generated under DGP1 and DGP3, respectively, outlined in Section 4.1 for $T = 250$ and $T = 750$, allowing the persistence parameter characterising the predictor to vary among $c \in \{5, 0, -2.5, -10, -0.5T\}$, and with slope parameters $\beta_{\tau_i} = \beta = b/T$ (except for $c = -0.5T$ where $\beta_{\tau_i} = \beta = b/\sqrt{T}$), $i = 1, \ldots, 11$, for $b = 0, 5, 10, 20, 50$. Overall, $\mathcal{T}_{\boldsymbol{\tau}}$ is seen to display good size control and strong empirical power properties for both of these DGPs.

4. Finite sample simulations

In this section we report results from a large set of Monte Carlo experiments investigating the finite sample properties of our proposed LM-type tests and, where relevant, comparing these with the IVX_{QR} test of Lee (2016) and the associated MBB test of Fan and Lee (2019), modified to correct for implementation errors present in the MBB algorithm outlined in Fan and Lee (2019) and detailed in Algorithm A.1 in the supplementary appendix, denoted IVX_{QR}^{MBB} in what follows. Comparison is also made with the recently developed weighted variable addition tests of Cai et al. (2023), which we denote by t_τ^w in the single predictor case and by Q_τ^w in the multiple predictor case. The bulk of our results are reported in Section 4.1 (the complete set of simulation results for all parameter constellations considered is provided in Part D of the supplementary appendix), where we consider finite sample size and local power for the case of a single predictor. Finite sample size simulations for the case of multiple predictors are discussed in Section 4.2.

All simulations are preformed in MATLAB, versions R2023a and R2024a, using the Mersenne Twister random number generator function and are based on 5000 Monte Carlo replications. In connection with the IVX_{QR}^{MBB} test, we use 499 bootstrap replications. All results pertain to two-sided tests for predictability at a given single quantile run at the nominal asymptotic 5% significance level (qualitatively similar results obtained for other conventional significance levels).

4.1. Single predictor ($K = 1$)

In this section we report finite sample size and power for the case of a single predictor, $K = 1$. Similarly to the simulation DGP used in Fan and Lee (2019, p.267), our simulated data are generated according to the predictive QR,

$$y_t = \alpha_\tau + \beta_\tau x_{t-1} + u_{t,\tau},\ t = 1, \dots, T \tag{17}$$

$$x_t = \mu_x + \xi_t,\ \ \xi_t = \rho\xi_{t-1} + v_t, \tag{18}$$

with $\xi_0 = 0$, where $u_{t,\tau} := u_t - Q_{u_t}(\tau|\mathcal{F}_{t-1})$ has zero conditional quantile by construction with $Q_{u_t}(\tau|\mathcal{F}_{t-1})$ computed as the inverse of the conditional CDF of u_t evaluated at τ, where u_t may be interpreted as the error term in the conditional mean predictive regression of y_t on x_{t-1}; specific examples for the disturbances u_t will be formulated in DGP1–DGP4 below. The parameters α_τ, β_τ, μ_x and $\rho = 1 + c/T$ are scalars, the latter characterising the degree of persistence of the predictor. All of the reported results pertain to tests of $H_0 : \beta_\tau = 0$ against $H_1 : \beta_\tau \neq 0$ in (17). Without loss of generality we set $\alpha_\tau = \mu_x = 0$, as all of the tests are based on de-quantiled data.

Results are reported for both $\mathcal{T}_\tau$, as defined in (13), implemented with Eicker–White standard errors, and the corresponding test discussed in Remark 10, denoted $\mathcal{T}_\tau^0$ in what follows, constructed using the conventional covariance estimator defined in (15). Both $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ are based on the instrument vector $z_t := \left(z_{I,t}, z_{II,t}\right)'$. Following the recommendations of Demetrescu et al. (2022) the type-II instrument used is $z_{II,t} = \sin\left(\frac{\pi t}{T}\right)$, and the type-I instrument, $z_{I,t}$, is given by the IVX choice of Kostakis et al. (2015) defined as in (10), with $a = 1$ and $\eta = 0.95$. Excepting the IVX instrument, $z_{I,t}$, all variables and instruments entering the estimated predictive regressions are demeaned, as described in Section 3. Results are reported for tests at each of the quantiles $\tau \in \{0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80.0.90, 0.95\}$.

4.1.1. Empirical size

This section reports empirical rejection frequencies for the $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w tests under the no predictability null hypothesis, $H_0 : \beta_\tau = 0$ in (17), for a range of DGPs featuring serial correlation in the errors driving the predictors and conditional and unconditional heteroskedasticity in the errors. Specifically, we report results for the following cases for the errors $(u_t,\ v_t)'$:

DGP1: **Homoskedastic iid and Serially Correlated Innovations:** In this DGP we set $v_t = \pi v_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$, with $(u_t,\ \varepsilon_t)'$ drawn from an i.i.d. bivariate Gaussian distribution with mean zero and unconditional covariance matrix $\mathbf{\Sigma} = \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}$, where ϕ corresponds to the correlation between u_t and ε_t, which is set as $\phi = -0.95$.[7] Results are reported for: (i) $\pi = \theta = 0$ ($v_t \sim iid$) ; (ii) $\pi = \pm 0.5$, $\theta = 0$ ($v_t \sim AR(1)$), and (iii) $\pi = 0$, $\theta = \pm 0.5$ ($v_t \sim MA(1)$).

DGP2: **Conditional Heteroskedasticity:** Here the innovations $(u_t, v_t)'$ are generated to exhibit time-varying conditional second-order moments according to the design,

$$(u_t, v_t)' = \begin{bmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{bmatrix} \boldsymbol{\eta}_t; \quad \mathrm{E}\left(\boldsymbol{\eta}_t\right) = \mathbf{0}, \quad \mathrm{E}\left(\boldsymbol{\eta}_t\boldsymbol{\eta}_t'\right) =: \boldsymbol{\Omega} = \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix},$$

where $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})'$ is drawn from an i.i.d. bivariate Gaussian distribution. Again we set $\phi = -0.95$. Three specific models of conditional heteroskedasticity are considered:

DGP2a - GARCH(1,1): The conditional variances $\{\sigma_{it}^2\}$ are driven by (normalised) stationary GARCH(1,1) processes $\sigma_{it}^2 = (1 - \theta_1 - \theta_2) + \theta_1 e_{i,t-1}^2 + \theta_2\sigma_{i,t-1}^2,\ i = 1, 2$, where $e_{1,t-1} \equiv u_{t-1}$ and $e_{2,t-1} \equiv v_{t-1}$, with $\theta_1, \theta_2 \geq 0$ and $\theta_1 + \theta_2 < 1$, such that $\mathrm{E}\left(u_t^2\right) = \mathrm{E}\left(v_t^2\right) = 1$. Results are reported for $(\theta_1, \theta_2) = \{(0.1, 0.5), (0.1, 0.8), (0.05, 0.9)\}$.

DGP2b - ARCH(1): Here, the conditional variances $\{\sigma_{it}^2\}$ are driven by a stationary ARCH(1) process $\sigma_{it}^2 = 1 + 0.9e_{i,t-1}^2$, $i = 1, 2$ with $e_{1,t-1} \equiv u_{t-1}$ and $e_{2,t-1} \equiv v_{t-1}$.

DGP2c - Stochastic Volatility: In this case, the innovations $(u_t, v_t)'$ are generated according to the first-order autoregressive stochastic volatility (ARSV) process, $u_t = e_{1t}\exp(h_{1t})$, $v_t = e_{2t}\exp(h_{2t})$, with $h_{it} = \lambda h_{i,t-1} + 0.5\xi_{it}$, where $(\xi_{it}, e_{it})' \sim i.i.d.\ N(0, \mathrm{diag}(\sigma_\xi^2, 1))$, independent across i = 1,2. Results are reported for $(\lambda, \sigma_\xi)' = (0.951, 0.314)'$.

[7] In predictive regression models for the equity premium employing valuation ratios as predictors (e.g. the dividend–price ratio, earnings–price ratio), the relevant innovation terms are strongly negatively correlated. Fan and Lee (2019, p.268) also use $\phi = -0.95$ in their simulation experiments pertaining to iid errors.

Table 1
Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker–White standard errors), $\mathcal{T}_\tau^0$ (conventional standard errors), IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w. Sample size $T = 250$. DGP1 (homoskedastic innovations): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, and $v_t = \pi v_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$, where $\beta_\tau = 0$, $\rho = 1 + c/T$, and $(u_t, \varepsilon_t)' \sim i.i.d.\ N(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad -0.95; \quad -0.95 \quad 1]$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w
		$(\pi,\theta) = (0,0)$					$(\pi,\theta) = (0.5,0)$					$(\pi,\theta) = (0,-0.5)$				
0	0.05	0.049	0.047	0.109	0.040	0.094	0.049	0.047	0.106	0.042	0.096	0.040	0.044	0.088	0.034	0.098
	0.1	0.051	0.055	0.085	0.055	0.080	0.051	0.058	0.094	0.057	0.069	0.043	0.050	0.084	0.051	0.081
	0.2	0.052	0.056	0.067	0.069	0.053	0.053	0.058	0.078	0.074	0.054	0.050	0.056	0.064	0.066	0.064
	0.3	0.059	0.060	0.077	0.093	0.047	0.057	0.059	0.072	0.082	0.048	0.053	0.057	0.057	0.069	0.057
	0.4	0.057	0.062	0.060	0.075	0.050	0.059	0.065	0.068	0.084	0.045	0.057	0.061	0.068	0.082	0.050
	0.5	0.059	0.064	0.054	0.073	0.048	0.061	0.065	0.066	0.080	0.043	0.057	0.060	0.064	0.086	0.054
	0.6	0.058	0.061	0.062	0.078	0.047	0.062	0.065	0.073	0.084	0.044	0.057	0.061	0.063	0.074	0.059
	0.7	0.056	0.057	0.075	0.092	0.053	0.053	0.057	0.062	0.078	0.046	0.051	0.056	0.055	0.065	0.061
	0.8	0.056	0.058	0.065	0.070	0.056	0.057	0.058	0.071	0.080	0.055	0.057	0.060	0.070	0.068	0.068
	0.9	0.052	0.053	0.085	0.054	0.073	0.053	0.053	0.091	0.051	0.071	0.049	0.047	0.079	0.054	0.080
	0.95	0.044	0.046	0.106	0.042	0.098	0.044	0.047	0.106	0.041	0.091	0.042	0.042	0.093	0.041	0.099
−10	0.05	0.044	0.042	0.080	0.032	0.094	0.044	0.045	0.083	0.032	0.092	0.039	0.038	0.079	0.027	0.094
	0.1	0.043	0.046	0.064	0.042	0.064	0.043	0.047	0.061	0.043	0.063	0.043	0.047	0.066	0.039	0.068
	0.2	0.047	0.053	0.049	0.048	0.050	0.048	0.053	0.047	0.047	0.046	0.047	0.049	0.042	0.046	0.050
	0.3	0.051	0.054	0.048	0.058	0.052	0.051	0.053	0.043	0.053	0.043	0.050	0.053	0.043	0.051	0.049
	0.4	0.055	0.058	0.039	0.055	0.044	0.053	0.056	0.042	0.054	0.039	0.054	0.055	0.045	0.056	0.048
	0.5	0.050	0.051	0.038	0.056	0.040	0.050	0.053	0.041	0.058	0.040	0.048	0.050	0.044	0.055	0.043
	0.6	0.058	0.061	0.041	0.054	0.047	0.058	0.060	0.043	0.058	0.043	0.048	0.049	0.041	0.053	0.049
	0.7	0.053	0.057	0.044	0.058	0.049	0.050	0.054	0.044	0.053	0.046	0.046	0.049	0.044	0.050	0.051
	0.8	0.054	0.056	0.050	0.046	0.054	0.053	0.057	0.054	0.049	0.052	0.051	0.054	0.046	0.042	0.059
	0.9	0.048	0.055	0.064	0.038	0.071	0.048	0.056	0.069	0.043	0.078	0.045	0.050	0.062	0.039	0.068
	0.95	0.041	0.046	0.088	0.037	0.090	0.045	0.048	0.092	0.034	0.097	0.045	0.053	0.084	0.035	0.100
$-0.5T$	0.05	0.036	0.049	0.085	0.032	0.097	0.038	0.049	0.085	0.031	0.103	0.041	0.051	0.082	0.035	0.103
	0.1	0.048	0.053	0.068	0.037	0.071	0.047	0.050	0.066	0.039	0.075	0.044	0.046	0.058	0.033	0.072
	0.2	0.048	0.050	0.047	0.043	0.053	0.049	0.055	0.047	0.040	0.055	0.050	0.049	0.048	0.041	0.052
	0.3	0.051	0.055	0.036	0.041	0.044	0.052	0.056	0.041	0.046	0.046	0.049	0.054	0.046	0.041	0.044
	0.4	0.051	0.053	0.038	0.045	0.044	0.051	0.053	0.036	0.052	0.044	0.053	0.056	0.042	0.048	0.043
	0.5	0.047	0.049	0.036	0.041	0.038	0.049	0.050	0.036	0.043	0.045	0.054	0.056	0.040	0.045	0.044
	0.6	0.047	0.049	0.038	0.044	0.044	0.048	0.049	0.039	0.041	0.048	0.051	0.051	0.045	0.044	0.048
	0.7	0.052	0.054	0.041	0.047	0.045	0.046	0.049	0.038	0.046	0.049	0.054	0.057	0.047	0.043	0.054
	0.8	0.050	0.052	0.047	0.044	0.053	0.042	0.048	0.046	0.038	0.056	0.052	0.053	0.053	0.044	0.054
	0.9	0.046	0.052	0.064	0.037	0.069	0.045	0.053	0.061	0.036	0.071	0.047	0.054	0.065	0.038	0.072
	0.95	0.039	0.048	0.089	0.032	0.105	0.042	0.052	0.090	0.033	0.092	0.043	0.050	0.084	0.031	0.103

DGP3: **Unconditional Heteroskedasticity:** Denoting the time-varying unconditional covariance matrix of $(u_t, v_t)'$ as $\mathbf{\Sigma}_t = \begin{bmatrix} \sigma_{ut}^2 & \phi\sigma_{ut}\sigma_{vt} \\ \phi\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2 \end{bmatrix}$, we allow for a one-time break in the variance of v_t. Specifically, $\sigma_{ut}^2 = 1$ and $\sigma_{vt}^2 = 1\mathbb{I}(t \le \lfloor \lambda T \rfloor) + \mathsf{b}\mathbb{I}(t > \lfloor \lambda T \rfloor)$, where $\mathbb{I}(\cdot)$ denotes the indicator function, and allow for an upward change ($\mathsf{b} = 4$) and for a downward change ($\mathsf{b} = 1/4$) in variance at break fractions $\lambda = \{1/3; 1/2; 2/3\}$. Again we set $\phi = -0.95$. We also consider a **DGP4** which allows for a change in variance in both u_t and v_t, viz., $\sigma_{ut}^2 = \sigma_{vt}^2 = 1\mathbb{I}(t \le \lfloor \lambda T \rfloor) + \mathsf{b}\mathbb{I}(t > \lfloor \lambda T \rfloor)$, again with $\mathsf{b} = 4$ and $\mathsf{b} = 1/4$.

Remark 16. Noting that (un)conditional homoskedasticity in u_t implies (un)conditional homoskedasticity in $u_{t,\tau}$ and that (un)conditional heteroskedasticity in u_t implies (un)conditional heteroskedasticity in $u_{t,\tau}$, it can be observed that for each of DGP1, DGP2 and DGP3, $s_{t,\tau} = \psi(u_{t,\tau})$ satisfies Assumption 3 and so our LM-type tests may be validly applied with the $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ statistics both having χ_1^2 limiting null distributions; cf. Proposition 1. In contrast, DGP4 violates Assumption 3; cf. Remark 9. Similarly, IVX_{QR}^{MBB} can be validly applied for DGP1, DGP2 and DGP3, but not DGP4, while IVX_{QR} and t_τ^w can only be validly applied for DGP1. ◊

Tables 1, 2, 3, and 4 report results for DGP1, DGP2a, DGP2b and DGP2c, and DGP3, respectively, in each case for a sample of size $T = 250$. We vary the persistence parameter characterising the predictor among $c \in \{0, -10, -0.5T\}$, covering a spectrum of exact unit root ($c = 0$), local-to-unity ($c = -10$), and weakly stationary ($c = -0.5T$) predictors. Additional results for $T = 250$ for the cases of locally explosive ($c = 5$) and local-to-unity ($c = -2.5$) predictors, together with some of the ARMA, GARCH and unconditionally heteroskedastic models, together with results for $T = 750$ for all cases, can be found in Tables D.1–D.10 in Part D of the supplementary appendix. The results for $c = 5$ are qualitatively similar to the results discussed in the main text below, except that $\mathcal{T}_\tau$ is somewhat oversized at extreme quantiles for the smaller sample size, $T = 250$, although it is still not as oversized as the IVX_{QR} and t_τ^w tests in this case. This pattern is not observed in either $\mathcal{T}_\tau^0$ or IVX_{QR}^{MBB}, however, whose empirical sizes remain close to the nominal level when $c = 5$.

Table 2

Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker–White standard errors), $\mathcal{T}_\tau^0$ (conventional standard errors), IVX_{QR}, IVX_{QR}^{MBB}, and t_τ^w. Sample size $T = 250$. DGP2a (GARCH(1,1)): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, where $\beta_\tau = 0$, $\rho = 1 - c/T$, and $(u_t, v_t)' = [\sigma_{1t} \quad 0; 0 \quad \sigma_{2t}]\, \boldsymbol{\eta}_t$; $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{\Omega})$ with $\boldsymbol{\Omega} = [1 \quad -0.95; \quad -0.95 \quad 1]$ and $\sigma_{it}^2 = \theta_0 + \theta_1 e_{i,t-1}^2 + \theta_2 \sigma_{i,t-1}^2$, $i = 1, 2$, with $\theta_0 = 1 - \theta_1 - \theta_2$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w
		$\theta_1 = 0.1,\ \theta_2 = 0.5$					$\theta_1 = 0.05,\ \theta_2 = 0.9$				
0	0.05	0.049	0.049	0.105	0.040	0.092	0.047	0.048	0.105	0.037	0.092
	0.1	0.049	0.052	0.085	0.056	0.066	0.049	0.053	0.090	0.056	0.059
	0.2	0.056	0.061	0.070	0.071	0.055	0.053	0.060	0.070	0.072	0.051
	0.3	0.060	0.064	0.078	0.095	0.049	0.060	0.063	0.078	0.092	0.041
	0.4	0.059	0.063	0.055	0.077	0.049	0.056	0.062	0.055	0.079	0.048
	0.5	0.058	0.060	0.052	0.071	0.045	0.055	0.058	0.055	0.070	0.044
	0.6	0.060	0.064	0.055	0.073	0.049	0.056	0.061	0.060	0.071	0.045
	0.7	0.057	0.059	0.074	0.086	0.056	0.059	0.062	0.077	0.086	0.052
	0.8	0.056	0.060	0.071	0.064	0.051	0.054	0.059	0.072	0.067	0.049
	0.9	0.048	0.055	0.087	0.052	0.067	0.050	0.055	0.091	0.056	0.067
	0.95	0.046	0.045	0.110	0.038	0.095	0.047	0.046	0.110	0.038	0.087
−10	0.05	0.044	0.044	0.089	0.033	0.092	0.044	0.044	0.089	0.033	0.085
	0.1	0.049	0.051	0.067	0.041	0.066	0.046	0.048	0.068	0.039	0.064
	0.2	0.053	0.054	0.047	0.045	0.052	0.053	0.056	0.050	0.047	0.048
	0.3	0.053	0.055	0.043	0.058	0.045	0.052	0.056	0.044	0.055	0.048
	0.4	0.048	0.050	0.038	0.047	0.050	0.048	0.051	0.041	0.046	0.046
	0.5	0.048	0.050	0.036	0.050	0.042	0.050	0.053	0.041	0.052	0.043
	0.6	0.051	0.052	0.038	0.051	0.047	0.051	0.054	0.039	0.050	0.043
	0.7	0.052	0.054	0.048	0.054	0.048	0.052	0.054	0.050	0.054	0.043
	0.8	0.050	0.054	0.045	0.043	0.049	0.050	0.053	0.048	0.043	0.050
	0.9	0.043	0.049	0.059	0.037	0.070	0.045	0.048	0.062	0.041	0.069
	0.95	0.044	0.051	0.093	0.037	0.091	0.044	0.050	0.094	0.036	0.085
−0.5T	0.05	0.040	0.050	0.112	0.035	0.098	0.040	0.049	0.099	0.033	0.099
	0.1	0.046	0.055	0.062	0.038	0.076	0.047	0.054	0.064	0.039	0.070
	0.2	0.051	0.052	0.046	0.044	0.055	0.051	0.053	0.050	0.042	0.052
	0.3	0.050	0.051	0.042	0.041	0.052	0.049	0.053	0.044	0.043	0.053
	0.4	0.050	0.052	0.042	0.045	0.049	0.052	0.053	0.045	0.044	0.049
	0.5	0.047	0.048	0.037	0.044	0.047	0.048	0.050	0.039	0.044	0.052
	0.6	0.050	0.052	0.041	0.046	0.054	0.051	0.053	0.042	0.047	0.051
	0.7	0.047	0.050	0.040	0.042	0.051	0.046	0.049	0.042	0.043	0.053
	0.8	0.045	0.051	0.043	0.043	0.063	0.046	0.052	0.048	0.041	0.066
	0.9	0.043	0.050	0.064	0.037	0.081	0.045	0.049	0.069	0.036	0.077
	0.95	0.040	0.050	0.106	0.036	0.101	0.041	0.048	0.100	0.033	0.098

Consider first the results for DGP1 in Table 1. It can be seen that $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ show very good size control for the case of conditionally homoskedastic innovations, regardless of the strength of persistence of the predictor and regardless of whether the predictor's innovations are serially uncorrelated or not. In contrast, the IVX_{QR} and t_τ^w tests both display significant over-sizing when testing for predictability at the upper and lower tails of the distribution in the case where x_t is strongly persistent. These size distortions are most clearly seen for the pure unit root case ($c = 0$), but are also present, albeit to a lesser extent, for $c = -2.5$ (for the latter see results in Table D.1 in the supplementary appendix). These size distortions are worse, other things equal, the further out one goes into the tails (see results for $\tau = 0.05$ and 0.95) and also tend to worsen where the predictor's innovations are serially correlated. Although the MBB implementation of the test, IVX_{QR}^{MBB}, does a good job at controlling the over-size seen in IVX_{QR} for $c = 0$ at the extreme quantiles, $\tau = 0.05, 0.1, 0.9$ and 0.95, it actually displays larger size distortions than IVX_{QR} away from the tails, $\tau = 0.3, 0.4., \ldots, 0.7$. From Table D.2, it is seen that these patterns of over-sizing in IVX_{QR} and IVX_{QR}^{MBB} for very strongly persistent predictors are also present for $T = 750$, while the size properties of t_τ^w are significantly improved.

Consider next the results in Table 2 for DGP2a relating to the case of GARCH innovations. It can be seen from the results that our proposed $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ tests both display very good size control across the all GARCH parameter constellations considered. The empirical size properties of the t_τ^w test are again decent, except in the extreme tails where some significant oversizing is again seen. In contrast, and as expected, this is not the case for the IVX_{QR} test which can be significantly over-sized, again most notably where the degree of persistence of x_t is strongest ($c = 0$ and $c = -2.5$, for the latter see results in Table D.3 of the supplementary appendix), when testing for predictability at the upper and lower tails of the distribution. Also as expected, this over-sizing is seen from the results in Table D.4 to remain for $T = 750$. It is also noteworthy that the MBB test, IVX_{QR}^{MBB} is not entirely successful in eliminating the over-size caused by GARCH innovations, but in contrast to IVX_{QR}, and consistent with the results in Table 1, tends to be over-sized when testing for predictability away from the tails of the distribution. The qualitative conclusions drawn from the results for the case of ARCH and SV innovations in Tables 3, D.5 and D.6 are very similar to those for the GARCH cases, albeit IVX_{QR} displays significant over-size in the ARCH case regardless of the persistence degree of x_t. Again the IVX_{QR}^{MBB} test does not control size well away from the tails of the distribution for either ARCH or SV innovations in the pure unit root case, $c = 0$. The t_τ^w test is over-sized in the tails regardless of whether ARCH or SV innovations are considered. Moreover, in the case of weakly persistent

Table 3

Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker–White standard errors), $\mathcal{T}_\tau^0$ (conventional standard errors), IVX_{QR}, IVX_{QR}^{MBB}, and t_τ^w. Sample size $T = 250$. Left panel, DGP2b (ARCH(1)): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, where $\beta_\tau = 0$, $\rho = 1 - c/T$, and $(u_t, v_t)' = \left[\sigma_{1t}\quad 0; 0\quad \sigma_{2t}\right] \boldsymbol{\eta}_t$; $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{\Omega})$ with $\boldsymbol{\Omega} = [1\quad -0.95; -0.95\quad 1]$ and $\sigma_{it}^2 = 1 + 0.9e_{i,t-1}^2$, $i = 1, 2$. Right panel, DGP2c (Stochastic Volatility [SV]): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, where $\beta_\tau = 0$, $\rho = 1 - c/T$, and $(u_t, v_t)'$ follow from a first-order AR stochastic volatility process as $(u_t = e_{1t}\exp(h_{1t}), v_t = e_{2t}\exp(h_{2t}))'$, and $h_{it} = \lambda h_{i,t-1} + 0.5\xi_{it}$ with $(\xi_{it}, e_{it})' \sim i.i.d.\ N(0, \mathrm{diag}(\sigma_\xi^2, 1))$, independent across $i = 1, 2$, with results reported for $(\lambda, \sigma_\xi)' = (0.951, 0.314)'$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w
		$ARCH(1)$					SV				
0	0.05	0.045	0.048	0.116	0.033	0.102	0.046	0.048	0.141	0.038	0.133
	0.1	0.046	0.045	0.104	0.053	0.067	0.048	0.054	0.096	0.058	0.096
	0.2	0.050	0.053	0.088	0.069	0.057	0.060	0.065	0.078	0.065	0.074
	0.3	0.053	0.055	0.089	0.074	0.054	0.060	0.062	0.067	0.078	0.064
	0.4	0.052	0.055	0.094	0.084	0.051	0.060	0.063	0.069	0.087	0.064
	0.5	0.049	0.051	0.100	0.079	0.048	0.055	0.058	0.070	0.089	0.057
	0.6	0.051	0.053	0.103	0.077	0.052	0.056	0.058	0.073	0.090	0.057
	0.7	0.048	0.051	0.098	0.072	0.054	0.060	0.065	0.069	0.082	0.066
	0.8	0.049	0.051	0.095	0.064	0.055	0.059	0.063	0.077	0.069	0.075
	0.9	0.043	0.046	0.100	0.051	0.072	0.054	0.054	0.103	0.058	0.097
	0.95	0.052	0.050	0.124	0.038	0.105	0.049	0.047	0.131	0.041	0.146
−10	0.05	0.048	0.043	0.133	0.027	0.103	0.051	0.050	0.115	0.038	0.128
	0.1	0.056	0.058	0.109	0.045	0.071	0.047	0.055	0.073	0.043	0.088
	0.2	0.053	0.051	0.092	0.050	0.055	0.046	0.051	0.053	0.046	0.065
	0.3	0.051	0.053	0.088	0.055	0.053	0.045	0.048	0.041	0.052	0.054
	0.4	0.053	0.055	0.086	0.053	0.046	0.053	0.055	0.043	0.053	0.049
	0.5	0.044	0.045	0.088	0.051	0.049	0.054	0.057	0.042	0.059	0.052
	0.6	0.048	0.049	0.088	0.051	0.051	0.056	0.059	0.043	0.062	0.052
	0.7	0.049	0.051	0.086	0.049	0.049	0.055	0.058	0.042	0.051	0.056
	0.8	0.046	0.049	0.090	0.046	0.051	0.051	0.054	0.051	0.047	0.064
	0.9	0.047	0.045	0.113	0.044	0.078	0.048	0.050	0.076	0.040	0.087
	0.95	0.057	0.047	0.141	0.037	0.109	0.055	0.051	0.116	0.034	0.135
$-0.5T$	0.05	0.041	0.039	0.333	0.046	0.232	0.038	0.049	0.114	0.037	0.125
	0.1	0.044	0.049	0.292	0.043	0.185	0.047	0.050	0.066	0.041	0.073
	0.2	0.049	0.052	0.275	0.048	0.150	0.043	0.052	0.045	0.036	0.047
	0.3	0.049	0.052	0.269	0.051	0.145	0.052	0.053	0.040	0.046	0.043
	0.4	0.051	0.052	0.257	0.056	0.137	0.049	0.054	0.033	0.048	0.038
	0.5	0.048	0.049	0.262	0.053	0.136	0.050	0.052	0.035	0.042	0.037
	0.6	0.049	0.051	0.259	0.056	0.144	0.053	0.056	0.034	0.048	0.035
	0.7	0.047	0.052	0.265	0.049	0.147	0.053	0.055	0.036	0.048	0.039
	0.8	0.049	0.053	0.278	0.050	0.156	0.043	0.047	0.046	0.041	0.050
	0.9	0.038	0.049	0.303	0.043	0.187	0.043	0.049	0.075	0.038	0.072
	0.95	0.038	0.039	0.334	0.043	0.226	0.042	0.054	0.113	0.036	0.120

predictors ($c = -0.5T$) and ARCH innovations t_τ^w is very significantly over-sized at all of the quantiles considered. These patterns of size distortion in t_τ^w are also observed in the larger sample ($T = 750$).

We next turn to the results for the unconditional heteroskedastic case. Tables 4, D.7 and D.8 present the results for DGP3 with $\mathsf{b} = 1/4$ where we observe a downward change in the unconditional variance of v_t, with u_t kept unconditionally homoskedastic. The corresponding results for an upward change, $\mathsf{b} = 4$, can be found in Tables D.9 and D.10 of the supplementary appendix. It can be seen from these results that $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ again both display very good size control throughout. The t_τ^w test displays over-sizing in the more extreme quantiles when $T = 250$, but size seems to improve for $T = 750$. In contrast, and as expected, neither IVX_{QR} nor IVX_{QR}^{MBB} is size controlled with the largest deviations from the nominal level again seen in the pure unit root case, with the distortions somewhat worse for $\mathsf{b} = 1/4$ than for $\mathsf{b} = 4$, other things equal, with the distortions often increased for $T = 750$ *vis-à-vis* $T = 250$. Finally, Tables D.11–D.14 in the supplementary appendix report corresponding results for DGP4 where we observe a contemporaneous one-time break of equal magnitude in the unconditional variances of u_t and v_t. Recall that none of the tests have been shown to be valid in this case. The findings for DGP4 with $\mathsf{b} = 1/4$ are similar to those for DGP3 with $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ displaying decent size control throughout, but with IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w badly over-sized in many cases, again most notably for $c = 0$. For DGP4 with $\mathsf{b} = 4$ the size distortions are somewhat worse than those seen for $\mathsf{b} = 1/4$ for all of the tests, with some over-sizing (though lower than is seen with the IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w tests) also seen for the $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ tests when $c = 0$, and which persists for $T = 750$.

4.1.2. Finite sample local power

We next evaluate the relative finite sample local power properties of the $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w tests. To that end, we simulate data from DGP (17)–(18) under a variety of local alternatives. To keep the set of results to a manageable level we only report results for the serially uncorrelated and homoskedastic case for $(u_t, v_t)'$ given by DGP1 from Section 4.1.1 with $\pi = \theta = 0$ and $T = 250$, in the main text (Fig. 1). Additional power results relating to DGP1 with autocorrelated errors, DGP2a and DGP3 are given

Table 4
Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker–White standard errors), $\mathcal{T}_\tau^0$ (conventional standard errors), IVX_{QR}, and IVX_{QR}^{MBB}. Sample size $T = 250$. DGP 3 (Unconditional Heteroskedasticity): $y_t = \beta_\tau x_{t-1} + u_{t\tau}, x_t = \rho x_{t-1} + v_t$ where $\beta_\tau = 0$, $\rho = 1 + c/T$ and $(u_t, v_t)' \sim i.i.d.\ N(\mathbf{0}, \mathbf{\Sigma_t})$, with $\mathbf{\Sigma}_t = [\sigma_{ut}^2 \quad -0.95\sigma_{ut}\sigma_{vt}; \quad -0.95\sigma_{ut}\sigma_{vt} \quad \sigma_{vt}^2]$ and $\sigma_{ut}^2 = 1$, $\sigma_{vt}^2 = 1\mathbb{I}(t \le \lfloor \lambda T \rfloor) + 1/4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	t_τ^w
		$\lambda = 1/3$					$\lambda = 2/3$				
0	0.05	0.064	0.060	0.109	0.042	0.095	0.065	0.063	0.108	0.042	0.092
	0.1	0.046	0.050	0.093	0.055	0.078	0.049	0.052	0.091	0.058	0.076
	0.2	0.051	0.054	0.074	0.065	0.061	0.054	0.059	0.074	0.066	0.064
	0.3	0.049	0.053	0.088	0.084	0.059	0.053	0.056	0.085	0.091	0.055
	0.4	0.052	0.057	0.092	0.098	0.053	0.054	0.059	0.072	0.090	0.049
	0.5	0.051	0.054	0.084	0.103	0.057	0.060	0.064	0.065	0.088	0.052
	0.6	0.059	0.062	0.085	0.104	0.061	0.063	0.066	0.075	0.091	0.054
	0.7	0.054	0.059	0.080	0.092	0.052	0.059	0.064	0.086	0.095	0.053
	0.8	0.057	0.061	0.083	0.070	0.062	0.060	0.066	0.067	0.065	0.060
	0.9	0.048	0.047	0.100	0.058	0.079	0.050	0.056	0.094	0.052	0.082
	0.95	0.066	0.065	0.118	0.038	0.104	0.067	0.062	0.110	0.038	0.107
−10	0.05	0.052	0.050	0.092	0.036	0.097	0.056	0.056	0.089	0.038	0.101
	0.1	0.051	0.054	0.065	0.037	0.070	0.052	0.055	0.064	0.042	0.071
	0.2	0.047	0.049	0.052	0.045	0.050	0.048	0.053	0.046	0.046	0.051
	0.3	0.045	0.049	0.045	0.053	0.048	0.045	0.047	0.044	0.052	0.048
	0.4	0.051	0.053	0.045	0.062	0.045	0.052	0.055	0.039	0.058	0.045
	0.5	0.053	0.056	0.046	0.060	0.041	0.049	0.050	0.037	0.059	0.039
	0.6	0.057	0.059	0.046	0.065	0.045	0.050	0.051	0.039	0.056	0.041
	0.7	0.053	0.058	0.045	0.057	0.048	0.049	0.049	0.041	0.053	0.045
	0.8	0.049	0.052	0.047	0.051	0.053	0.044	0.049	0.045	0.046	0.051
	0.9	0.048	0.050	0.064	0.044	0.070	0.045	0.050	0.063	0.046	0.069
	0.95	0.053	0.049	0.080	0.029	0.094	0.052	0.052	0.087	0.032	0.091
$-0.5T$	0.05	0.035	0.045	0.080	0.033	0.090	0.042	0.049	0.088	0.029	0.104
	0.1	0.039	0.051	0.065	0.033	0.068	0.046	0.053	0.061	0.036	0.066
	0.2	0.048	0.051	0.046	0.039	0.049	0.048	0.046	0.046	0.038	0.050
	0.3	0.049	0.051	0.043	0.041	0.048	0.045	0.050	0.040	0.041	0.044
	0.4	0.049	0.050	0.036	0.046	0.042	0.045	0.048	0.034	0.043	0.038
	0.5	0.051	0.053	0.035	0.051	0.046	0.045	0.046	0.035	0.043	0.042
	0.6	0.054	0.057	0.041	0.048	0.043	0.049	0.049	0.036	0.040	0.043
	0.7	0.054	0.056	0.040	0.045	0.047	0.049	0.052	0.038	0.044	0.040
	0.8	0.052	0.054	0.046	0.046	0.051	0.049	0.052	0.047	0.043	0.051
	0.9	0.045	0.050	0.071	0.041	0.074	0.043	0.050	0.071	0.039	0.075
	0.95	0.035	0.053	0.089	0.034	0.101	0.038	0.052	0.089	0.034	0.107

in Figures D.3–D.4 in the supplementary appendix. In terms of the relative power rankings of the tests, these results are in general qualitatively similar to those discussed below for the serially uncorrelated and homoskedastic case. One clear exception, however, is the case of GARCH errors [DGP2a] where the power of the t_τ^w test, relative to the other tests, is very significantly diminished.

We again set $\phi = -0.95$ and consider five values of the persistence parameter, c, associated with x_t; specifically, $c = \{0, -2.5, -10, -20, -0.5T\}$. The slope parameter $\beta_{1\tau}$ in (17) is set to be local-to-zero; that is, for $c = \{0, -2.5, -10, -20\}$, where x_{t-1} is strongly persistent, we set $\beta_\tau = b/T$, while for the weakly dependent predictor case, $c = -0.5T$, we set $\beta_\tau = b/\sqrt{T}$. In both cases results are reported for the range of Pitman drift values, $b = \{0, 1, \ldots, 26\}$. Again to keep the amount of results manageable, we consider tests at quantiles $\tau = \{0.2, 0.5, 0.8\}$. All other aspects of the simulation design are as described previously.

Fig. 1 graphs the simulated finite sample local power curves for the four tests for $T = 250$. The corresponding curves for $T = 750$, reported in Figure D.2 in the supplementary appendix, are almost identical to those in Fig. 1. Additional results for the extreme quantiles $\tau = 0.05, 0.95$, and for the case of a locally explosive predictor ($c = 5$) are also reported in Figures D.1 and D.2 of the supplementary appendix; in this locally explosive case all of the tests have very steep power curves, though the tests proposed in this paper do still enjoy a power advantage over the other tests considered, most notably for $\tau = 0.05, 0.95$.

Consider first the strongly persistent predictor case. Here a comparison between the Wald-type IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w tests shows IVX_{QR} to be significantly more powerful than the latter two for $c = 0$, and to a lesser extent for $c = -2.5$. For $c \le -10$ there is relatively little to choose between IVX_{QR} and IVX_{QR}^{MBB} on power with both of these tests significantly more powerful than t_τ^w. It is of course important to recall, from the discussion surrounding Table 1 that IVX_{QR} and t_τ^w (IVX_{QR}^{MBB}) are not well size controlled for $c = 0$ and $c = -2.5$ when $\tau = 0.05, 0.1, 0.9, 0.95$ ($\tau = 0.3, 0.4., \ldots, 0.7$), leading to artificially inflated power in these cases. The local power functions of $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ are almost identical to each other throughout the cases considered and are seen to both enjoy considerable power advantages over the t_τ^w, IVX_{QR} and IVX_{QR}^{MBB} tests (despite the aforementioned lack of size control in some settings for these three tests) across all of the settings considered. The power advantage of $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ over IVX_{QR}^{MBB} is greatest in the pure unit root case, $c = 0$, while their advantage over IVX_{QR} is greatest for $c = -2.5$. The power advantage of $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ over t_τ^w is of a fairly similar magnitude regardless of c. For example, for $b = 10$ and $c = 0$: $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ both have power of about 79% for $\tau = 0.2$ and $\tau = 0.8$, and about 92% for $\tau = 0.5$; IVX_{QR} has power of just over 65% for $\tau = 0.2$ and $\tau = 0.8$, and 61%

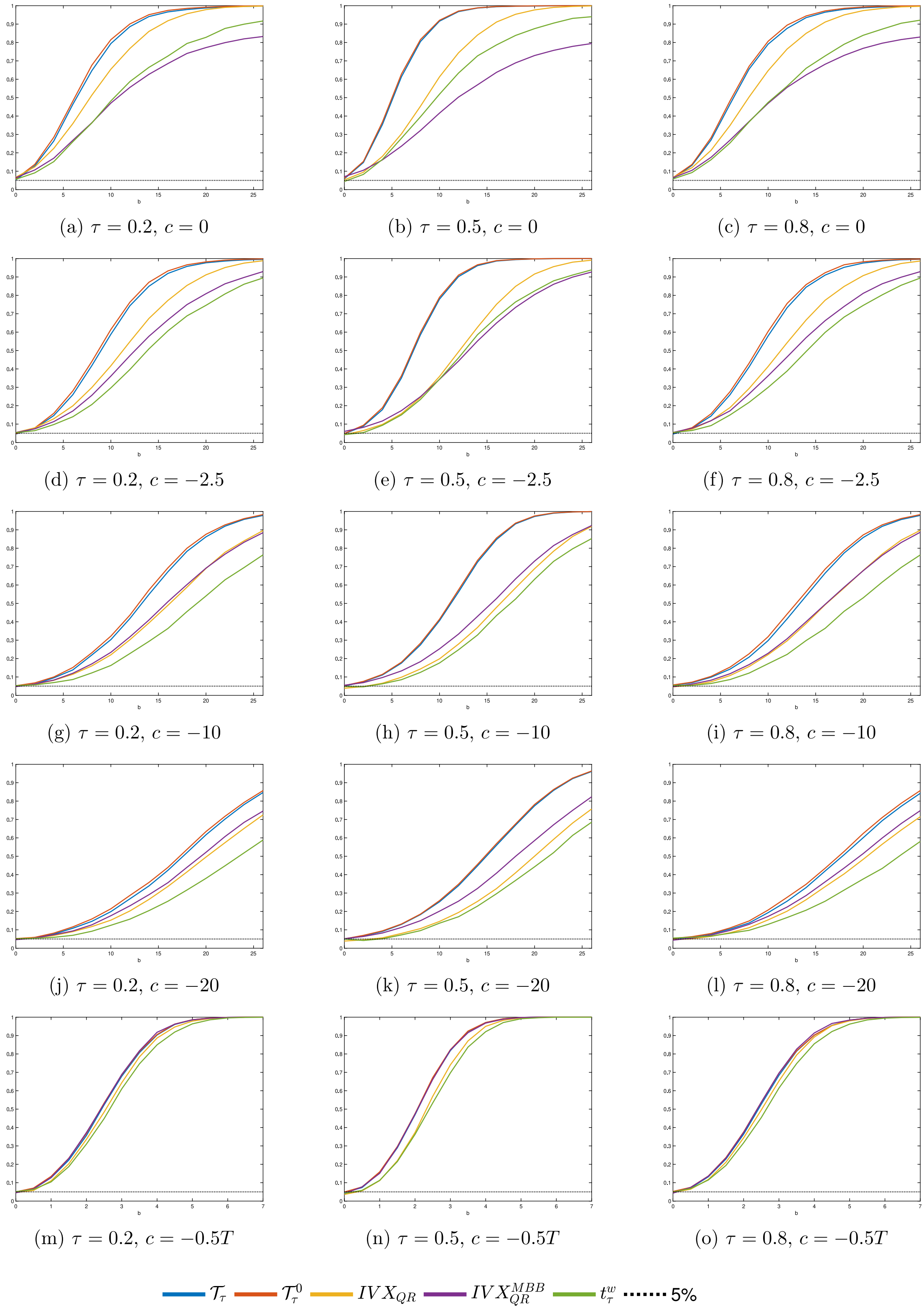

Fig. 1. Finite sample local power of predictability tests for $T = 250$. DGP1 (homoskedastic and serially uncorrelated innovations): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, $\rho = 1 + c/T$, $\beta_\tau = b_\tau/T$ (except for $c = -0.5T$ where $\beta_\tau = b_\tau/\sqrt{T}$), $b_\tau = \{0, 1, \ldots, 26\}$, and $(u_t, v_t)' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} = [1 \quad -0.95; \quad -0.95 \quad 1]$.

for $\tau = 0.5$; IVX_{QR}^{MBB} has power of about 47% for $\tau = 0.2$ and $\tau = 0.8$, and 42% for $\tau = 0.5$, and t_{τ}^{w} has power of about 48% for $\tau = 0.2$ and $\tau = 0.8$, and 52% for $\tau = 0.5$. For $b = 10$ and $c = -2.5$, $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$ both have power of about 58% for $\tau = 0.2$ and $\tau = 0.8$, and 78% for $\tau = 0.5$; IVX_{QR} has power of 41% for $\tau = 0.2$ and $\tau = 0.8$, and 36% for $\tau = 0.5$; IVX_{QR}^{MBB} has power of 36% for all three values of τ; and t_{τ}^{w} has power of about 30% for $\tau = 0.2$ and $\tau = 0.8$, and 34% for $\tau = 0.5$. For the case of a weakly stationary predictor, $c = -0.5T$, the local power functions of the $\mathcal{T}_{\tau}$, $\mathcal{T}_{\tau}^{0}$ and IVX_{QR}^{MBB} tests are very similar to each other, and dominate the power functions of the IVX_{QR} and t_{τ}^{w} tests. This is most clearly seen for $\tau = 0.5$.

For all of the tests, power varies to some degree with τ. For example, in the case of a strongly persistent predictor power is slightly higher for $\tau = 0.5$ *vis-à-vis* $\tau = 0.2$ and $\tau = 0.8$, for given values of c and b for $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$, with the converse tending to hold for IVX_{QR}, IVX_{QR}^{MBB} and t_{τ}^{w}.

4.2. Multiple predictors ($K > 1$)

We conclude this section by briefly investigating the finite sample size performance of our proposed $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$ tests in multiple predictor ($K > 1$) settings. We also compare their performance with those of the IVX_{QR} and Q_{τ}^{w} tests.[8] We generate our simulation data according to the DGP

$$y_t = \alpha_\tau + \boldsymbol{\beta}_\tau' \boldsymbol{x}_{t-1} + u_{t,\tau} \tag{19}$$

$$\boldsymbol{x}_t = \boldsymbol{\mu}_x + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t = \boldsymbol{\Gamma} \boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t, \tag{20}$$

with $\boldsymbol{\xi}_0 = \mathbf{0}$ and, like in the $K = 1$ DGP in (17), $u_{t,\tau} = u_t - Q_{u_t}(\tau|\mathcal{F}_{t-1})$. The innovations are generated as,

$$u_t = \gamma f_t + a_t \tag{21}$$

$$\boldsymbol{v}_t = -\gamma f_t \mathbf{1}_K + \boldsymbol{e}_t, \tag{22}$$

where $\{f_t\}$ is a common factor to both innovations, which we generate as a sequence of independent standard normals, $\mathbf{1}_K$ denotes the $K \times 1$ vector of ones, $\gamma = -4.5$, and $(a_t, \boldsymbol{e}_t')' \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{K+1})$.[9] We set $\alpha_\tau = 0$ and $\boldsymbol{\mu}_x = \mathbf{0}$, without loss of generality. The predictors are generated setting $\boldsymbol{\Gamma} = \rho \mathbf{I}_K$, with $\rho = 1 + c/T$ for $c = \{0, -2.5, -10, -0.5T\}$. All of the reported results pertain to tests of $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$ against $H_1 : \boldsymbol{\beta}_\tau \neq \mathbf{0}$ in (19). Results are again reported for tests at each of the quantiles $\tau = \{0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80.0.90, 0.95\}$.

In computing $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$, we used the instrument vector $\boldsymbol{z}_t := (\boldsymbol{z}_{I,t}', \boldsymbol{z}_{II,t}')'$ with the K elements of the type-II instrument vector, $\boldsymbol{z}_{II,t}$, defined as in (11) setting $\omega_k = (2k-1)\pi$, for $k = 1, \ldots, K$, and the K elements of the type-I instrument vector, $\boldsymbol{z}_{I,t}$, given by the IVX defined as in (10), with $a = 1$ and $\eta = 0.95$. A finite-sample correction factor of the form considered in Kostakis et al. (2015, p.1515) was employed which improved the finite sample properties of the tests.[10]

Table 5 reports empirical null rejection frequencies for $T = 250$ and for $c = \{0, -10, -0.5T\}$ and $K = \{2, 3, 5\}$, for the $\mathcal{T}_{\tau}$, $\mathcal{T}_{\tau}^{0}$, IVX_{QR} and Q_{τ}^{w} tests when $\boldsymbol{\beta}_\tau = \mathbf{0}$ in (19). Tables D.15 and D.16 of Part D of the supplementary appendix present the full set of results for sample sizes $T = 250$ and $T = 750$, when $c = \{0, -2.5, -10, -0.5T\}$ and $K = \{2, 3, 4, 5\}$. The results in Table 5 show that the $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$ tests display reasonable size control throughout, albeit both tests display a degree of over-sizing for $K \geq 3$ for quantiles near the centre of the distribution, most notably where $c = 0$. This over-sizing is not, however, present for $c = -0.5T$. In contrast, the empirical size properties of the IVX_{QR} and Q_{τ}^{w} tests tend to be better behaved in the centre of the distribution, but deteriorate very significantly as K increases for quantiles away from the centre of the distribution, $\tau < 0.3$ and $\tau > 0.7$, with over-sizing present in such cases for all values of c. For example, for $T = 250$, $c = 0$ and $K = 3$, the IVX_{QR} test for $\tau = 0.05$ has size 19.5% and Q_{τ}^{w} has size 17.6%, while $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$ have sizes 2.1% and 4.4%, respectively.

5. Empirical application to S&P 500 returns

In light of the discussion in Section 1, we will apply the LM-based quantile predictability tests developed in Section 3.2, along with the extant Wald-type tests of Lee (2016) and Fan and Lee (2019), and the variable addition tests of Cai et al. (2023) to analyse the predictive content of a number of predictors believed to convey information on the dynamics of firm or market risk for (excess) stock returns. For comparison purposes with the conditional mean predictability testing literature, we will also apply the tests to the variables in the Goyal–Welch dataset (Welch and Goyal, 2008). Specifically, we analyse data from five different sources: (1) the (updated) monthly dataset used in Welch and Goyal (2008); (2) a time varying tail index which we compute from

8 Fan and Lee (2019) outline their MBB procedure for the case of a single predictor ($K = 1$). It is unclear how one should implement a multiple predictor version of their MBB algorithm, as this would require setting up confidence regions in $\mathbb{R}^K$; see also Montiel Olea and Plagborg-Møller (2019) for a recent overview of multiple simultaneous confidence intervals. For this reason we do not report a MBB implementation of IVX_{QR}.

9 For (21)–(22), $Corr(u_t, v_{it}) = -\gamma^2/(1+\gamma^2)$, for each of $i = 1, \ldots, K$. We accordingly set $\gamma = -4.5$ so that the correlation between u_t and each element of $\boldsymbol{v}_t$ is approximately -0.95. We thank a referee for this observation.

10 In the present context, this entails replacing $\mathbf{D}_T$ in (14) by $(\mathbf{D}_T - \Xi)$, where $\Xi := T(\bar{\tilde{\mathbf{z}}}_{-1}\bar{\tilde{\mathbf{z}}}_{-1}')(\hat{\sigma}_{u_\tau}^2 - \hat{\boldsymbol{\Omega}}_{vu_\tau}\hat{\boldsymbol{\Omega}}_{vv}\hat{\boldsymbol{\Omega}}_{vu_\tau}')$, with $\bar{\tilde{\mathbf{z}}}_{-1} = T^{-1}\sum_{t=1}^{T}\tilde{\mathbf{z}}_{t-1}$, $\tilde{\mathbf{z}}_{t-1} := (\boldsymbol{z}_{I,t-1}', \bar{\mathbf{z}}_{II,t-1}')'$, and $\bar{\mathbf{z}}_{II,t-1}'$ is a vector of demeaned type II instruments. Note that the mean of $\bar{z}_{II,t-1}$ will be a zero vector, so that $(\bar{\tilde{\mathbf{z}}}_{-1}\bar{\tilde{\mathbf{z}}}_{-1}') = \mathrm{diag}\{(\bar{z}_{I,-1}\bar{\mathbf{z}}_{I,-1}'), \mathbf{0}\}$. Furthermore, $\hat{\boldsymbol{\Omega}}_{vv}$ and $\hat{\boldsymbol{\Omega}}_{vu_\tau}$ are the estimates of the long-run variance of $\boldsymbol{v}_t$, and of the long-run covariance between $u_{t,\tau}$ and $\boldsymbol{v}_t$, respectively, which where computed using a Bartlett kernel with bandwidth $T^{1/3}$; a discussion on the practical choice of these estimators is provided in Kostakis et al. (2015, p.1513 and 1524). The use of this correction factor does not alter any of the large sample results previously stated for $\mathcal{T}_{\tau}$ and $\mathcal{T}_{\tau}^{0}$.

Table 5
Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker–White standard errors), $\mathcal{T}_\tau^0$ (conventional standard errors), IVX_{QR}, and Q_τ^w. Sample size $T = 250$. Multiple Predictors DGP: $y_t = \boldsymbol{\beta}_\tau' \boldsymbol{x}_{t-1} + u_{t\tau}$, $\boldsymbol{x}_t = \boldsymbol{\Gamma} \boldsymbol{x}_{t-1} + \boldsymbol{v}_t$, with $\boldsymbol{\beta}_\tau = \mathbf{0}$, and where $u_t = \gamma f_t + a_t$ and $\boldsymbol{v}_t = -\gamma f_t \mathbf{1}_K + \boldsymbol{e}_t$, where the common factor f_t is generated as a sequence of independent standard normals, $\mathbf{1}_K$ is a $K \times 1$ vector of ones, $\gamma = -4.5$ and $(a_t, \boldsymbol{e}_t')' \sim i.i.d.\ N(\mathbf{0}, \mathbf{I}_{K+1})$. The predictors are generated for $\boldsymbol{\Gamma} = \rho \mathbf{I}_K$ with $\rho = 1 + c/T$ for $c = \{0, -10, -0.5T\}$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	Q_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	Q_τ^w	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	Q_τ^w
		$K = 2$				$K = 3$				$K = 5$			
0	0.05	0.025	0.040	0.143	0.128	0.021	0.044	0.195	0.176	0.012	0.063	0.339	0.290
	0.1	0.034	0.040	0.107	0.079	0.033	0.052	0.136	0.107	0.049	0.077	0.215	0.155
	0.2	0.040	0.047	0.071	0.055	0.055	0.066	0.080	0.058	0.078	0.107	0.123	0.091
	0.3	0.050	0.054	0.068	0.051	0.076	0.087	0.076	0.050	0.101	0.124	0.098	0.073
	0.4	0.061	0.067	0.060	0.042	0.076	0.086	0.073	0.052	0.115	0.134	0.085	0.060
	0.5	0.063	0.065	0.065	0.045	0.080	0.088	0.063	0.052	0.121	0.143	0.084	0.058
	0.6	0.059	0.062	0.056	0.046	0.078	0.084	0.067	0.051	0.120	0.140	0.085	0.064
	0.7	0.057	0.062	0.065	0.044	0.074	0.085	0.074	0.050	0.104	0.128	0.103	0.074
	0.8	0.047	0.051	0.078	0.059	0.059	0.071	0.077	0.067	0.084	0.109	0.126	0.083
	0.9	0.036	0.042	0.105	0.084	0.044	0.063	0.147	0.107	0.049	0.077	0.218	0.159
	0.95	0.025	0.034	0.152	0.121	0.023	0.049	0.222	0.178	0.006	0.062	0.357	0.298
−10	0.05	0.049	0.058	0.128	0.131	0.046	0.055	0.183	0.183	0.026	0.057	0.304	0.300
	0.1	0.053	0.062	0.082	0.076	0.051	0.059	0.112	0.103	0.039	0.062	0.163	0.147
	0.2	0.064	0.068	0.060	0.050	0.057	0.063	0.070	0.062	0.055	0.068	0.096	0.083
	0.3	0.059	0.065	0.046	0.047	0.058	0.063	0.049	0.048	0.060	0.072	0.072	0.072
	0.4	0.066	0.069	0.037	0.046	0.059	0.067	0.042	0.046	0.062	0.074	0.063	0.060
	0.5	0.067	0.073	0.043	0.046	0.060	0.064	0.038	0.046	0.068	0.080	0.061	0.059
	0.6	0.058	0.063	0.037	0.039	0.063	0.067	0.047	0.041	0.064	0.074	0.062	0.056
	0.7	0.061	0.063	0.044	0.045	0.061	0.067	0.051	0.053	0.062	0.070	0.068	0.073
	0.8	0.060	0.065	0.053	0.051	0.060	0.065	0.066	0.057	0.055	0.072	0.096	0.091
	0.9	0.049	0.057	0.080	0.076	0.047	0.063	0.111	0.101	0.041	0.060	0.163	0.157
	0.95	0.045	0.054	0.133	0.130	0.037	0.046	0.191	0.174	0.020	0.051	0.326	0.305
$-0.5T$	0.05	0.041	0.050	0.143	0.136	0.034	0.049	0.191	0.182	0.026	0.042	0.306	0.304
	0.1	0.049	0.048	0.084	0.084	0.040	0.049	0.102	0.108	0.032	0.048	0.159	0.144
	0.2	0.052	0.054	0.049	0.052	0.044	0.054	0.063	0.062	0.036	0.048	0.079	0.080
	0.3	0.046	0.050	0.043	0.051	0.043	0.048	0.051	0.051	0.033	0.043	0.062	0.061
	0.4	0.049	0.054	0.040	0.044	0.049	0.051	0.039	0.043	0.034	0.042	0.049	0.054
	0.5	0.057	0.060	0.038	0.044	0.043	0.050	0.045	0.038	0.036	0.043	0.050	0.050
	0.6	0.050	0.054	0.038	0.043	0.046	0.047	0.041	0.041	0.041	0.048	0.053	0.051
	0.7	0.050	0.053	0.046	0.049	0.050	0.056	0.052	0.049	0.038	0.050	0.067	0.064
	0.8	0.052	0.055	0.051	0.055	0.045	0.050	0.057	0.067	0.034	0.048	0.074	0.069
	0.9	0.045	0.048	0.086	0.088	0.040	0.045	0.105	0.107	0.029	0.043	0.166	0.149
	0.95	0.041	0.050	0.129	0.139	0.031	0.043	0.186	0.190	0.023	0.041	0.316	0.314

7951 firms' asset prices for all active stocks listed in NASDAQ and NYSE, using the approach developed in Kelly and Jiang (2014); (3) the log of the VIX and of the SKEW indices, obtained from the Chicago Board Options Exchange (https://www.cboe.com); (4) the 10 year treasury term premium and yield, obtained from the Federal Reserve Bank of New York (https://www.newyorkfed.org/research/data_indicators/term-premia-tabs#/overview); and (5) the variance risk premium of Bollerslev et al. (2009), obtained from Hao Zhou's website (https://sites.google.com/site/haozhouspersonalhomepage). All data are monthly and cover the period 1990:01–2021:12 ($T = 384$). Detailed data descriptions are given in Part C of the supplementary appendix.

5.1. Preliminary analysis

In Table 6 we provide some preliminary analysis for the variables used in this section. We start by reporting the 95% confidence interval (CI) for the largest autoregressive root (defined here as ρ) for the twenty-two potential predictor time series considered. These CIs are computed following the approach of Campbell and Yogo (2006, Section 3.2, pp. 35–37). The autoregressive order for each predictor variable is selected using the Bayesian Information Criteria (BIC) with the maximum lag length determined using the rule of Schwert (1989). Consistent with the earlier findings of Goyal and Welch (2003), most of the Goyal–Welch predictors appear to be very strongly persistent; indeed for seven of these predictors the CI includes $\rho = 1$, such that an exact unit root cannot be rejected. Of all of the series considered, 14 have a lower bound on the CI for ρ that is greater than 0.8, suggesting most of the predictors are strongly persistent. Only ltr_t, dfr_t and vrp_t appear to be of weaker persistence, each with an upper CI bound smaller than 0.5. Recall from the simulation results reported in Section 4.1.1, that the Wald-type tests can be significantly over-sized for predictors with a dominant autoregressive root very close to unity.

An important feature of the LM type tests we propose, not shared by the Wald-type tests, is their robustness to unconditional heteroskedasticity in the predictor (see Remark 16). In connection with this, Table 6 reports the results of applying the four stationary volatility tests ($\mathcal{H}_{KS}$, $\mathcal{H}_R$, $\mathcal{H}_{CvM}$ and $\mathcal{H}_{AD}$, using a Bartlett long run variance estimator with lag truncation parameter 4) proposed by Cavaliere and Taylor (2008, pp. 311–312) to our data. These tests are valid regardless of whether the variable is weakly or strongly persistent, and test the null hypothesis of unconditional homoksedasticity against the alternative of unconditional heteroskedasticity.

Table 6
Predictor persistence, tests for stationary volatility in the log of the equity premium, ep_t, and in the twenty-two predictors considered, and conditional mean predictability test results.

	95% CI for ρ	$\mathcal{H}_{KS}$	$\mathcal{H}_R$	$\mathcal{H}_{CVM}$	$\mathcal{H}_{AD}$	t_β^{2*}
ep_t	–	0.192	0.305	0.002	0.027	–
dp_t	[0.994, 1.012]	0.537	0.891	0.069	0.421	0.219
dy_t	[0.993, 1.012]	0.549	0.919	0.071	0.428	0.366
e/p_t	[0.915, 0.983]	0.856	1.406	0.179	0.813	0.287
de_t	[0.826, 0.929]	1.047	1.684*	0.280	1.268	0.000
$rvol_t$	[0.937, 0.994]	0.710	0.951	0.088	0.582	1.811
bm_t	[0.984, 1.010]	0.664	1.234	0.121	0.631	0.151
$ntis_t$	[0.962, 1.005]	0.757	1.307	0.158	0.713	1.083
tbl_t	[0.992, 1.012]	0.800	1.262	0.080	0.458	0.342
lty_t	[0.997, 1.013]	0.606	0.840	0.073	0.429	0.395
ltr_t	[0.144, 0.392]	0.908	0.965	0.183	0.870	0.134
tms_t	[0.949, 1.000]	1.044	1.159	0.272	1.409	0.507
dfy_t	[0.847, 0.942]	0.972	1.267	0.237	1.103	0.462
dfr_t	[0.227, 0.463]	**1.635****	1.737*	**0.646****	**3.007****	0.283
$infl_t$	[0.547, 0.726]	1.003	1.436	0.297	1.425	0.020
vix_t	[0.772, 0.892]	**1.361****	1.389	**0.625****	**3.146****	1.493
$skew_t$	[0.943, 0.997]	1.163	1.338	**0.614****	**3.311****	1.551
$TailIndex_t$	[0.684, 0.830]	**1.553****	1.654*	**0.742****	**3.414****	0.400
$ACMY10_t$	[0.996, 1.012]	0.698	0.888	0.092	0.495	0.403
$ACMTP10_t$	[0.984, 1.011]	0.738	1.187	0.117	0.540	1.500
vrp_t	[0.141, 0.389]	0.980	1.008	0.315	2.056*	0.058
iv_t	[0.787, 0.903]	1.042	1.042	0.269	1.583	0.181
rv_t	[0.282, 0.510]	0.882	0.946	0.303	1.883	0.006

Notes: (i) Bold entries denote statistical significance at the 5% level (or stricter), while *, ** and *** indicate significant outcomes at the 10%, 5% and 1% significance levels, respectively. (ii) The column labelled "95% CI for ρ" refers to the 95% confidence interval for the largest autoregressive root (ρ) of the predictor variable. (iii) The critical values for the stationary volatility tests, $\mathcal{H}_{KS}$, $\mathcal{H}_R$, $\mathcal{H}_{CVM}$ and $\mathcal{H}_{AD}$, are respectively 1.64, 2.01, 0.743, 3.85 for a 1% significance level; 1.36, 1.75, 0.461, 2.492 for a 5% significance level and 1.22, 1.62, 0.347, 1.933 for a 10% significance level; see Cavaliere and Taylor (2008, Section 6, pp.311-315). (iv) The final column, headed t_β^{2*}, reports the full sample fixed regressor wild bootstrap based test results for conventional conditional mean predictability proposed by Demetrescu et al. (2022).

The tests were applied to the fitted residuals from the BIC-selected autoregressive model estimated for each variable. With the exception of dfr_t, for the Goyal–Welch predictors none of the tests are able to reject the null hypothesis at any conventional significance level. However, among the additional predictors measuring market risk and firm dynamics, statistically significant evidence of unconditional heteroskedasticity at the 5% level is found for vix_t, $skew_t$, and $TailIndex_t$ variables. Again, we would therefore expect the Wald-type tests to be potentially unreliable for these variables. Finally, in the case of the equity premium, ep_t, none of the tests find evidence of unconditional heteroskedasticity at any conventional significance level, suggesting that our assumption that the QR errors are unconditionally homoskedastic (see Remark 9) is supported by the data.

The last column of Table 6 reports, for each possible predictor, the outcome of the conventional IV-combination test, denoted t_β^2, for conditional mean predictability proposed by Demetrescu et al. (2022). The results provided are for the full sample statistics using Eicker–White standard errors. The t_β^2 statistic, like the LM quantile predictability tests developed in this paper, is based on a combination of the IVX instrument, $z_{I,t-1}$ and the sine instrument, $z_{II,t-1}$. Statistical significance is determined based on fixed regressor wild bootstrap p-values computed according to Demetrescu et al. (2022, Algorithm 1, p.99) with 999 bootstrap replications. For the period under analysis, the results in Table 6 show that t_β^2 does not yield evidence of predictability for the conditional mean of the equity premium at any conventional significance level for any of the predictors considered.

5.2. Quantile predictability results

The finding of no predictability for any of the predictors from t_β^2 in Table 6 only relates to the nature of possible predictive relations present during "normal" times; it says nothing about whether predictive relations hold when the return is far from the mean. To that end, we next test for predictability at different parts of the return distribution, applying the $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w quantile predictability tests to each of the 22 predictors at each of the 11 different quantiles, $\tau = \{0.05, 0.1, 0.2, \ldots, 0.8, 0.9, 0.95\}$, together with the joint test $\mathcal{T}_\tau$ from Section 3.3 across these 11 quantiles. The choice of these quantiles allows us to track the complete return distribution (left-tail, centre and right-tail) to better characterise the predictive potential of the predictors considered.

The results in Table 7, consistent with the results for conditional mean predictability in Table 6, uncover relatively little evidence of predictability at the median. In particular, the size-controlled $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ tests signal the existence of median predictability at the 5% level (or stricter) only for $rvol_t$ and vix_t. While IVX_{QR}, IVX_{QR}^{MBB} and t_τ^w give a small number of additional rejections in favour of median predictability, these are for predictors where significant evidence of unconditional heteroskedasticity is found, and so may be attributable to size control issues with these tests.

The most interesting feature of the results in Table 7 is the amount of statistically significant evidence the tests uncover of predictability at the shoulders or tails of the equity premium distribution. Moreover, it is precisely those predictors which contain

Table 7
Equity premium predictability: Univariate Predictor Performance. This table reports results from short-run (1-month ahead) predictive regressions of CRSP month-end values index returns. The first 14 rows report the test results for the predictors in Welch and Goyal (2008), followed by results for the log of the VIX (vix_t), the log of SKEW ($skew$), the tail index, the 10 year treasury term premium and yield, and the variance risk premium.

	$\mathcal{T}_\tau$		τ										
			0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
dp_t	3.242	$\mathcal{T}_\tau$	0.53	0.686	0.275	0.244	0.393	0.054	0.007	0.002	0.287	0.048	0.420
		$\mathcal{T}_\tau^0$	0.051	0.806	0.304	0.266	0.409	0.054	0.007	0.002	0.291	0.049	0.360
		IVX_{QR}	0.740	0.504	0.492	0.022	0.362	0.071	1.492	1.282	0.069	0.002	2.816
		IVX_{QR}^{MBB}	−0.025*	−0.019*	0.012*	0.002	**0.008****	−0.003	**−0.014*****	**−0.013*****	**−0.003****	0.001	**0.025*****
		t_τ^w	1.383	1.519	0.561	0.608	1.008	0.652	0.166	−0.416	−0.160	−0.226	1.863*
dy_t	3.754	$\mathcal{T}_\tau$	0.286	1.248	0.612	0.495	0.650	0.152	0.000	0.010	0.458	0.013	0.132
		$\mathcal{T}_\tau^0$	0.312	1.556	0.701	0.553	0.686	0.153	0.000	0.010	0.477	0.015	0.126
		IVX_{QR}	0.224	0.552	1.855*	**3.701*****	**2.990*****	0.896	0.299	0.003	0.054	0.005	1.243
		IVX_{QR}^{MBB}	−0.012	**0.017*****	**0.019*****	**0.022*****	**0.017*****	**0.009*****	**0.004*****	0.001*	**−0.002*****	−0.001*	0.016
		t_τ^w	1.583	1.379	0.496	1.419	1.146	0.898	0.157	−0.415	−0.154	−0.121	1.675*
e/p_t	12.477	$\mathcal{T}_\tau$	**2.733*****	**2.516****	1.920*	**2.072****	1.579	0.474	0.051	0.734	1.386	1.796*	**2.198****
		$\mathcal{T}_\tau^0$	**5.497*****	**3.594*****	**2.029****	**2.176****	1.621	0.475	0.054	0.946	**2.060****	**3.647*****	**7.644*****
		IVX_{QR}	**3.656*****	**11.372*****	**7.853*****	**5.270*****	**5.120*****	0.988	0.070	**2.366****	**2.540****	**6.851*****	1.249
		IVX_{QR}^{MBB}	**0.034*****	**0.048*****	**0.028*****	**0.019*****	**0.016*****	**0.007****	0.002	**−0.010*****	**−0.012*****	**−0.021*****	**−0.012****
		t_τ^w	**2.747*****	**3.509*****	**2.490****	1.336	1.392	0.435	0.273	−0.404	−0.167	**−2.218****	−0.388
de_t	11.757	$\mathcal{T}_\tau$	**3.319*****	1.153	0.342	0.314	0.088	0.008	0.007	0.956	0.812	**2.282****	**3.571*****
		$\mathcal{T}_\tau^0$	**6.374*****	1.379	0.300	0.287	0.086	0.008	0.008	1.287	1.275	**5.135*****	**15.010*****
		IVX_{QR}	**4.154*****	**7.939*****	1.599	0.820	0.525	0.548	0.014	0.772	**2.074****	**3.815*****	0.943
		IVX_{QR}^{MBB}	−0.030	−0.037	−0.011	−0.007	−0.005	−0.004	−0.001	0.005	0.009	0.015	0.009
		t_τ^w	−1.120	−1.865*	−0.625	−0.844	−0.019	0.023	0.738	1.538	1.648*	**1.986****	0.736
$rvol_t$	**37.202*****	$\mathcal{T}_\tau$	**4.913*****	**5.098*****	0.620	0.723	0.042	**2.313****	**4.474*****	**11.131*****	**18.980*****	**13.817*****	**8.777*****
		$\mathcal{T}_\tau^0$	**3.082*****	**2.912*****	0.535	0.668	0.040	**2.325****	**4.787*****	**13.253*****	**24.534*****	**18.052*****	**16.017*****
		IVX_{QR}	**11.457*****	**10.781*****	0.870	1.290	0.017	0.473	**5.471*****	**11.111*****	**17.085*****	**39.005*****	**25.404*****
		IVX_{QR}^{MBB}	−0.367*	−0.272	−0.059	−0.059	−0.006	0.028	**0.091*****	**0.131*****	**0.167*****	**0.252*****	**0.265*****
		t_τ^w	**−2.167****	**−3.366*****	−0.953	−1.281	0.081	1.233	**2.038****	**3.126*****	**4.210*****	**6.551*****	**4.020*****
bm_t	3.477	$\mathcal{T}_\tau$	0.039	0.262	0.912	0.131	0.080	0.029	0.014	0.302	0.380	0.005	0.253
		$\mathcal{T}_\tau^0$	0.042	0.331	1.093	0.145	0.083	0.029	0.014	0.321	0.456	0.006	0.295
		IVX_{QR}	0.530	0.225	1.750*	0.000	0.548	0.576	0.351	1.171	0.425	0.002	**2.577*****
		IVX_{QR}^{MBB}	−0.064	−0.039	**0.055*****	0.000	−0.024	−0.023	−0.017	−0.033	**−0.020****	0.002	0.087
		t_τ^w	0.690	1.361	1.296	1.883*	0.763	−0.001	0.136	−1.048	−1.133	−0.744	1.105
$ntis_t$	12.314	$\mathcal{T}_\tau$	**3.786*****	**3.997*****	0.000	0.367	0.040	0.005	0.029	0.241	0.010	0.174	0.394
		$\mathcal{T}_\tau^0$	**5.464*****	**4.331*****	0.000	0.383	0.041	0.005	0.027	0.219	0.008	0.141	0.319
		IVX_{QR}	**3.366*****	**5.347*****	0.586	0.000	0.433	0.011	1.313	0.572	0.796	0.788	0.066
		IVX_{QR}^{MBB}	0.593	0.643	0.145	0.000	0.092	0.013	0.144	0.099	0.127	0.172	−0.059
		t_τ^w	**2.088****	**2.341****	−0.259	−0.205	0.357	−0.028	0.278	0.501	0.037	0.151	−0.527
tbl_t	7.064	$\mathcal{T}_\tau$	0.949	0.435	0.814	1.114	0.578	0.888	1.049	0.176	0.489	0.011	0.296
		$\mathcal{T}_\tau^0$	0.766	0.394	0.749	1.079	0.567	0.890	1.084	0.178	0.503	0.011	0.311
		IVX_{QR}	0.563	0.010	1.238	0.325	0.790	0.463	0.013	1.235	1.292	0.329	0.725
		IVX_{QR}^{MBB}	0.002	0.000	−0.002	−0.001	−0.001	−0.001	0.000	−0.002	−0.002	−0.001	−0.002
		t_τ^w	1.435	1.026	−0.500	0.218	0.014	−0.513	−0.173	−0.843	−0.220	−0.578	−0.518
lty_t	7.718	$\mathcal{T}_\tau$	1.078	0.549	0.787	1.359	0.684	1.054	1.468	0.183	0.557	0.009	0.262
		$\mathcal{T}_\tau^0$	0.707	0.450	0.740	1.336	0.676	1.060	1.507	0.182	0.539	0.008	0.250
		IVX_{QR}	1.380	0.000	1.820*	**4.330*****	0.919	0.659	0.907	**1.979****	0.590	0.002	0.465
		IVX_{QR}^{MBB}	0.004*	0.000	**−0.002****	**−0.003*****	−0.001	−0.001	−0.001	**−0.002*****	−0.001	0.000	−0.002
		t_τ^w	1.413	1.652	0.301	0,204	0.794	0.017	−0.026	−0.729	0.072	−0.512	−0.251
ltr_t	5.445	$\mathcal{T}_\tau$	1.025	0.841	0.044	0.000	0.025	0.343	0.488	1.080	1.188	**2.518****	**2.784*****
		$\mathcal{T}_\tau^0$	**2.440****	1.435	0.050	0.000	0.025	0.344	0.483	1.109	1.377	**3.522*****	5.264
		IVX_{QR}	0.498	**2.155****	0.236	0.000	0.160	0.032	0.840	**3.457*****	**3.036*****	**3.135*****	**11.364*****
		IVX_{QR}^{MBB}	−0.002	−0.003	−0.001	0.000	0.000	0.000	0.001	0.001	0.001	0.002	0.004
		t_τ^w	−1.095	−0.481	0.090	0.629	0.260	0.098	0.847	1.119	1.039	0.888	1.679*
tms_t	8.618	$\mathcal{T}_\tau$	0.254	0.046	0.110	1.270	0.727	0.628	**2.890*****	0.292	0.784	0.314	0.004
		$\mathcal{T}_\tau^0$	0.362	0.053	0.115	1.340	0.740	0.629	**2.750*****	0.275	0.665	0.286	0.004
		IVX_{QR}	0.217	0.045	1.032	**2.877*****	0.479	0.515	0.673	0.253	0.136	0.135	0.000
		IVX_{QR}^{MBB}	−0.003	−0.001	−0.003	−0.004	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	0.000
		t_τ^w	1.078	1.114	0.134	−1.316	−0.692	−0.092	−0.274	0.123	0.587	0.091	0.231

(*continued on next page*)

Table 7 (*continued*).

dfy_t	20.052**	$\mathcal{T}_\tau$	5.004***	5.537***	1.752*	0.496	0.554	0.000	0.001	1.202	3.429***	3.530***	3.828***
		$\mathcal{T}_\tau^0$	28.504***	15.924***	2.625***	0.559	0.576	0.000	0.001	1.330	4.984***	6.406***	12.498***
		IVX_{QR}	9.993***	15.222***	12.192***	14.220***	4.193***	1.072	0.224	4.165***	8.939***	4.838***	2.370**
		IVX_{QR}^{MBB}	−0.051*	−0.052*	−0.033	−0.032	−0.014	−0.006	−0.003	0.013**	0.018***	0.018***	0.013**
		t_τ^w	−2.420**	−2.620***	−1.333	−1.541	−0.471	−0.004	0.207	1.571	2.073**	1.680*	0.426
dfr_t	6.138	$\mathcal{T}_\tau$	1.377	1.090	1.390	1.352	0.625	0.352	0.316	0.000	0.149	1.723*	1.827*
		$\mathcal{T}_\tau^0$	5.379***	2.573***	1.959*	1.479	0.661	0.353	0.311	0.000	0.195	3.201***	5.903***
		IVX_{QR}	2.152**	1.578	6.739***	3.021***	0.853	0.264	0.585	0.000	0.384	3.030***	6.155***
		IVX_{QR}^{MBB}	0.006	0.004	0.005	0.003	0.001	0.001	0.001	0.000	−0.001	−0.004	−0.005
		t_τ^w	1.168	0.745	2.375**	1.910*	0.891	0.887	0.443	−0.378	−0.940	−0.744	−2.022**
$infl_t$	4.927	$\mathcal{T}_\tau$	0.880	1.014	0.324	0.170	0.175	0.498	0.041	0.002	0.153	0.020	0.019
		$\mathcal{T}_\tau^0$	1.651*	1.479	0.367	0.170	0.175	0.500	0.040	0.002	0.155	0.019	0.019
		IVX_{QR}	0.583	4.024***	0.782	0.516	0.269	0.077	0.000	0.000	0.041	0.004	0.343
		IVX_{QR}^{MBB}	0.015	0.028	0.009	−0.006	−0.004	−0.002	0.000	0.000	−0.002	−0.001	−0.007
		t_τ^w	0.933	2.226 **	0.734	−0.536	−0.198	−0.563	−0.239	−0.540	−0.824	−0.603	−0.343
vix_t	48.456***	$\mathcal{T}_\tau$	6.161***	6.487***	2.180**	0.143	0.093	2.812***	5.141***	14.074***	29.787***	21.306***	12.921***
		$\mathcal{T}_\tau^0$	12.640***	7.667***	2.205**	0.134	0.089	2.830***	5.478***	16.626***	45.736***	42.082***	34.296***
		IVX_{QR}	15.106***	11.857***	2.739***	0.000	0.652	1.693	7.279***	17.463***	28.051***	64.535***	127.824***
		IVX_{QR}^{MBB}	−0.063***	−0.046***	−0.017***	0.000	0.006**	0.009***	0.017***	0.026***	0.033***	0.047***	0.061***
		t_τ^w	−3.714***	−3.407***	−2.443**	−1.574	−0.714	0.582	2.825***	2.634***	3.663***	5.833***	5.017***
$skew_t$	9.137	$\mathcal{T}_\tau$	1.992**	0.791	2.310**	1.433	1.606	1.529	3.453***	0.243	0.116	0.014	0.104
		$\mathcal{T}_\tau^0$	2.272**	1.105	2.692***	1.453	1.586	1.536	3.663***	0.263	0.127	0.015	0.107
		IVX_{QR}	4.413***	0.528	5.405	3.909	1.006	2.118	0.745	0.172	0.035	0.000	0.008
		IVX_{QR}^{MBB}	0.238***	0.061***	0.107***	0.096***	0.035***	0.046***	0.032***	0.016***	0.008***	0.000	−0.007
		t_τ^w	−1.680*	0.905	5.961***	5.317***	4.273***	3.346***	2.756***	2.127**	0.797	0.459	0.263
$TailIndex_t$	20.146**	$\mathcal{T}_\tau$	3.217***	2.356**	0.881	0.203	0.036	1.023	1.755*	3.224***	8.500***	7.269***	12.903***
		$\mathcal{T}_\tau^0$	2.600***	2.009**	0.889	0.203	0.037	1.028	1.703*	3.002***	6.871***	5.813***	9.265***
		IVX_{QR}	5.614***	8.394***	3.695***	5.530***	0.001	1.683*	2.802***	10.213***	20.609***	40.635***	39.708***
		IVX_{QR}^{MBB}	−0.060***	−0.044***	−0.020***	−0.022***	0.000	0.010***	0.012***	0.024***	0.036***	0.053***	0.056***
		t_τ^w	−2.470**	−2.536**	−0.699	−0.620	−0.539	0.839	1.094	2.022 **	1.497	2.416**	3.121***
$ACMY10_t$	7.731	$\mathcal{T}_\tau$	1.067	0.539	0.803	1.370	0.693	1.065	1.471	0.189	0.567	0.008	0.266
		$\mathcal{T}_\tau^0$	0.704	0.443	0.756	1.347	0.685	1.072	1.511	0.188	0.550	0.008	0.255
		IVX_{QR}	1.331	0.000	2.120**	4.316***	1.618	0.654	0.892	1.950*	0.497	0.002	0.902
		IVX_{QR}^{MBB}	0.004	0.000	−0.002*	−0.003***	−0.002*	−0.001	−0.001	−0.002***	−0.001	0.000	−0.003
		t_τ^w	1.395	1.206	0.343	0.192	0.672	0.065	−0.031	−0.739	0.030	−0.511	−0.195
$ACMTP10_t$	7.948	$\mathcal{T}_\tau$	0.078	0.073	1.887*	2.979***	1.671*	1.689*	3.078***	0.398	0.889	0.002	0.101
		$\mathcal{T}_\tau^0$	0.057	0.060	1.655*	2.767***	1.632	1.699*	3.193***	0.394	0.873	0.002	0.113
		IVX_{QR}	1.140	4.897***	1.067	5.586***	0.515	0.886	1.906*	0.724	0.247	0.001	0.000
		IVX_{QR}^{MBB}	−0.008	−0.014	−0.004	−0.007	−0.002	−0.002	−0.003*	−0.002	−0.001	0.000	0.000
		t_τ^w	1.095	0.244	−0.113	−2.036	−0.529	−0.340	−0.631	−0.632	−0.264	−0.042	−0.678
vrp_t	14.710	$\mathcal{T}_\tau$	0.449	0.148	0.214	0.144	0.399	0.657	1.285	1.633	1.903*	0.531	0.006
		$\mathcal{T}_\tau^0$	1.672*	0.289	0.249	0.137	0.366	0.657	1.534	2.568***	4.616***	2.629***	0.059
		IVX_{QR}	0.008	0.487	4.174***	8.019***	16.928***	26.725***	34.546***	41.630***	42.801***	10.163***	0.166
		IVX_{QR}^{MBB}	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
		t_τ^w	1.485	1.918	0.271	1.290	0.520	0.282	0.157	0.090	0.423	−0.764	−0.581
iv_t	48.603***	$\mathcal{T}_\tau$	6.875***	7.049***	4.473***	0.560	0.034	1.658*	3.504***	11.592***	24.268***	18.466***	11.879***
		$\mathcal{T}_\tau^0$	31.749***	16.647***	5.944***	0.599	0.034	1.664*	3.603***	13.446***	38.181***	52.686***	58.524***
		IVX_{QR}	51.800***	43.884***	9.619***	9.690***	0.273	2.722***	28.601***	51.513***	60.586***	92.774***	82.369***
		IVX_{QR}^{MBB}	−0.001	−0.001*	0.000**	0.000	0.000	0.000	0.000***	0.000***	0.001***	0.001***	0.001***
		t_τ^w	−1.451	−1.351	−0.466	0.335	−0.545	0.162	1.153	1.405	0.826	1.432	1.090
rv_t	32.108***	$\mathcal{T}_\tau$	3.579***	3.370***	2.810***	0.626	0.084	0.156	0.277	1.814*	3.926***	3.491***	2.828***
		$\mathcal{T}_\tau^0$	25.619***	11.463***	4.744***	0.727	0.083	0.156	0.305	2.593***	8.830***	16.526***	28.331***
		IVX_{QR}	80.414***	65.813***	18.896***	4.722***	8.729***	0.002	6.916***	10.255***	43.030***	66.221***	80.569***
		IVX_{QR}^{MBB}	−0.001	−0.001	0.000*	0.000	0.000	0.000	0.000	0.000	0.000**	0.000***	0.001***
		t_τ^w	−1.791	−1.454	−0.166	−0.138	−0.628	−0.007	0.483	0.600	0.688	1.399	0.869

Notes: (i) Bold entries denote statistical significance at the 5% level (or stricter), while *, ** and *** indicate significant outcomes at the 10%, 5% and 1% significance levels, respectively. (ii) The entries in the column headed $\mathcal{T}_\tau$ are the results of the multiple-quantile LM-type test in (16) applied over the eleven quantiles ($\tau = 0.05, 0.1, \ldots, 0.9, 0.95$) considered in the empirical analysis. (iii) The entries for IVX_{QR}^{MBB} correspond to the estimates of the slope parameter, β_τ, in (1), computed as outlined in Fan and Lee (2019, p.264, Eq. (11)), with significance determined via the MBB based confidence intervals outlined in Algorithm A.1 in Part A of the supplementary appendix, using 1999 bootstrap replications. Associated MBB confidence intervals for β_τ are reported in Table C.1 in Part C of the supplementary appendix.

information on market or firm risk, or market uncertainty where the quantile predictability tests find evidence of predictability at outer quantiles. Specifically: e/p_t, which is a function of the perceived risk of a firm with the effect showing up in the cost of equity (i.e., a firm with a higher cost of equity will trade at a lower multiple of earnings than a similar firm with a lower cost of equity); $rvol_t$ is used to analyse broad or specific market risk environments, and provides a perspective on how the market has reacted to historical events; the default measures dfy_t and dfr_t correspond to the difference between BAA and AAA-rated corporate bond yields, and the difference between long-term corporate bond and long-term government bond returns, respectively; vix_t represents the market's expectation for the relative strength of near-term price change of the S&P 500; $skew_t$ is a proxy for investor sentiment and volatility, and measures potential risk in financial markets; $TailIndex_t$ is a measure of time-varying tail risk; $ACMTP10_t$ measures short-term Treasury yields term risk; and the variance risk premium vrp_t, which corresponds to the difference between implied volatility (iv_t) and realised volatility (rv_t), is a measure that captures aggregate market risk aversion.

Looking at these in turn for our size-controlled $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ tests, for e/p_t the tests find evidence of predictability in the tails and shoulders of the return distribution, but not for the inner quantiles, i.e., $\tau = (0.4, 0.5, 0.6, 0.7)$. The equity risk premium volatility ($rvol_t$), the tail index ($TailIndex_t$) and the log CBOE volatility index (vix_t), all display significance in the left tail (the first two predictors for $\tau \leq 0.1$ and the third for $\tau \leq 0.2$). Moreover, $rvol_t$ and vix_t also display significance for $\tau \geq 0.5$ and $TailIndex_t$ for $\tau > 0.5$. Similar results to those obtained for the for vix_t are observed for iv_t. The default yield spread, dfy_t, (and to a certain extend the default return spread, dfr_t), displays predictive content in the tails ($\tau \leq 0.2$ and $\tau \geq 0.8$), but not in the centre of the return distribution.

These empirical results in Table 7 highlight that for a number of the predictors considered, particularly those including information pertaining to market or firm risk and market uncertainty, while exhibiting limited predictive power for the central moments of the returns distribution, they do yield significant predictability in the left and right tails of the returns distribution, in line with the findings of Meligkotsidou et al. (2014). The relevance of these risk proxy measures varies depending on whether the market sentiment is bullish (i.e., when excess return is highly positive — right-tail) or bearish (i.e., when the excess return is highly negative — left-tail).

6. Conclusions

We have developed simple yet powerful tests, based on the LM principle, of the null hypothesis of no predictability at a user-chosen quantile for financial returns data. Our proposed tests can be applied to models with either a single or multiple putative predictors. The tests are based on statistics computed using simple linear 2SLS estimation, without the need to find external instruments. We have shown that our proposed statistics possess pivotal standard limiting null distributions regardless of whether the predictors are weakly or strongly persistent and regardless of the degree of endogeneity of the predictors. They are also valid, without the need for bootstrap implementation, in the presence of conditional heteroskedasticity in the quantile predictive regression errors and/or the errors driving the predictor. Moreover, and in contrast to the extant Wald-type tests of Lee (2016), their moving blocks bootstrap implementation developed in Fan and Lee (2019), and the variable addition tests of Cai et al. (2023) our LM-type tests are also asymptotically valid in the presence of unconditional heteroskedasticity in the errors driving the predictors. We have also developed analogous joint tests over a specified set of multiple distinct quantiles. Monte Carlo simulations suggest that, for strongly persistent predictors, our proposed tests display more robust finite sample size properties coupled with markedly superior power profiles to the extant tests. An empirical application looking at the predictive performance of twenty-two predictors for the equity premium for the S&P 500 found no evidence of conditional mean predictability, but uncovered significant evidence of predictability in the left and right tails of the returns distribution, particularly for predictors containing information on market or firm risk, or market uncertainty.

Our focus in this paper has been on short-horizon predictability testing. However, we close by noting that the methods we propose in this paper should straightforwardly extend to the long-horizon setting analysed by, *inter alia*, Kostakis et al. (2015), Xu (2020), Kostakis et al. (2023) and Demetrescu et al. (2023b), and the references therein. However, it is important to note that both Kostakis et al. (2023) and Demetrescu et al. (2023b) caution against the use of long-horizon tests. In particular, Kostakis et al. (2023, p.24) conclude by arguing that " ... our study questions the incremental value of using long-horizon predictive regressions, for the additional reason that correctly sized test statistics are less, not more powerful as the horizon increases. Hence, we argue that it is preferable to conduct inference relying on the actual data generating process, rather than resorting to long-horizon predictive regressions".

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2025.106002.

References

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Rev. Financ. Stud. 22 (11), 4463–4492.
Bollerslev, T., Todorov, V., 2014. Time-varying jump tails. J. Econometrics 183 (2), 168–180.
Boswijk, H.P., Cavaliere, G., Rahbek, A., Taylor, A.M.R., 2016. Inference on co-integration parameters in heteroskedastic vector autoregressions. J. Econometrics 192 (1), 64–85.
Breitung, J., Demetrescu, M., 2015. Instrumental variable and variable addition based inference in predictive regressions. J. Econometrics 187 (1), 358–375.
Cai, Z., Chen, H., Liao, X., 2023. A new robust inference for predictive quantile regression. J. Econometrics 234 (1), 227–250.
Campbell, J.Y., Yogo, M., 2006. Efficient tests of stock return predictability. J. Financ. Econ. 81 (1), 27–60.
Cavaliere, G., Taylor, A.M.R., 2008. Time-transformed unit root tests for models with non-stationary volatility. J. Time Series Anal. 29 (2), 300–330.
Cavanagh, C.L., Elliott, G., Stock, J.H., 1995. Inference in models with nearly integrated regressors. Econometric Theory 11 (5), 1131–1147.
Cederburg, S., Johnson, T.L., O'Doherty, M.S., 2022. On the economic significance of stock return predictability. Rev. Financ. 27 (2), 619–657.
Chevapatrakul, T., Xu, Z., Yao, K., 2019. The impact of tail risk on stock market returns: The role of market sentiment. Int. Rev. Econ. Financ. 59, 289–301.
Demetrescu, M., Georgiev, I., Rodrigues, P.M.M., Taylor, A.M.R., 2022. Testing for episodic predictability in stock returns. J. Econometrics 227 (1), 85–113.
Demetrescu, M., Georgiev, I., Rodrigues, P.M.M., Taylor, A.M.R., 2023a. Extensions to IVX methods of inference for return predictability. J. Econometrics 237 (2, Part C), 105271.
Demetrescu, M., Rodrigues, P.M.M., Taylor, A.M.R., 2023b. Transformed regression-based long-horizon predictability tests. J. Econometrics 237 (2, Part C), 105316.
Elliott, G., Müller, U.K., Watson, M.W., 2015. Nearly optimal tests when a nuisance parameter is present under the null hypothesis. Econometrica 83 (2), 771–811.
Fan, R., Lee, J.H., 2019. Predictive quantile regressions under persistence and conditional heteroskedasticity. J. Econometrics 213 (1), 261–280.
Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Manag. Sci. 49 (5), 639–654.
Granger, C.W.J., 1999. Outline of forecast theory using generalized cost functions. Span. Econ. Rev. 1 (2), 161–173.
Gungor, S., Luger, R., 2021. Exact inference in long-horizon predictive quantile regressions with an application to stock returns. J. Financ. Econ. 19 (4), 746–788.
Jansson, M., Moreira, M.J., 2006. Optimal inference in regression models with nearly integrated regressors. Econometrica 74 (3), 681–714.
Kelly, B., Jiang, H., 2014. Tail risk and asset prices. Rev. Financ. Stud. 27 (10), 2841–2871.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46 (1), 33–50.
Komunjer, I., 2005. Quasi-maximum likelihood estimation for conditional quantiles. J. Econometrics 128 (1), 137–164.
Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2015. Robust econometric inference for stock return predictability. Rev. Financ. Stud. 28 (5), 1506–1553.
Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2023. Taking stock of long-horizon predictability tests: Are factor returns predictable? J. Econometrics 237 (2, Part C), 105380.
Lee, J.H., 2016. Predictive quantile regression with persistent covariates: IVX-QR approach. J. Econometrics 192 (1), 105–118.
Liu, X., Long, W., Peng, L., Yang, B., 2024. A unified inference for predictive quantile regression. J. Amer. Statist. Assoc. 119 (546), 1526–1540.
Maynard, A., Shimotsu, K., 2009. Covariance-based orthogonality tests for regressors with unknown persistence. Econometric Theory 25 (1), 63–116.
Maynard, A., Shimotsu, K., Kuriyama, N., 2024. Inference in predictive quantile regressions. J. Econometrics 245 (1–2), 105875.
Meligkotsidou, L., Panopoulou, E., Vrontos, I.D., Vrontos, S.D., 2014. A quantile regression approach to equity premium prediction. J. Forecast. 33 (7), 558–576.
Meligkotsidou, L., Panopoulou, E., Vrontos, I.D., Vrontos, S.D., 2021. Out-of-sample equity premium prediction: A complete subset quantile regression approach. Eur. J. Financ. 27 (1–2), 110–135.
Montiel Olea, J.L., Plagborg-Møller, M., 2019. Simultaneous confidence bands: Theory, implementation, and an application to SVARs. J. Appl. Econometrics 34 (1), 1–17.
Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: The role of small sample bias. J. Financ. 48 (2), 641–661.
Nicolau, J., Rodrigues, P.M.M., Stoykov, M.Z., 2023. Tail index estimation in the presence of covariates: Stock returns' tail risk dynamics. J. Econometrics 235 (2), 2266–2284.
Ohno, S., Ando, T., 2018. Stock return predictability: A factor-augmented predictive regression system with shrinkage method. Econometric Rev. 37 (1), 29–60.
Paye, B.S., Timmermann, A., 2006. Instability of return prediction models. J. Empir. Financ. 13 (3), 274–315.
Portnoy, S., 1991. Asymptotic behavior of regression quantiles in non-stationary, dependent cases. J. Multivariate Anal. 38 (1), 100–113.
Schwert, G.W., 1989. Why does stock market volatility change over time? J. Financ. 44 (5), 1115–1153.
Stambaugh, R.F., 1999. Predictive regressions. J. Financ. Econ. 54 (3), 375–421.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21 (4), 1455–1508.
Xu, K.-L., 2020. Testing for multiple-horizon predictability: Direct regression based versus implication based. Rev. Financ. Stud. 33 (9), 4403–4443.