Essex Finance Centre
Working Paper Series

Working Paper No 86: 12-2023

“Predictive Quantile Regressions with Persistent and Heteroskedastic Predictors: A Powerful 2SLS Testing Approach”

“Matei Demetrescu, Paulo M. M. Rodrigues and A. M. Robert Taylor”

Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ
Web site: http://www.essex.ac.uk/ebs/

Predictive Quantile Regressions with Persistent and Heteroskedastic Predictors: A Powerful 2SLS Testing Approach*

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

[a] Department of Statistics, TU Dortmund University

[b] Banco de Portugal and Nova School of Business and Economics

[c] Essex Business School, University of Essex

September 13, 2023

Abstract

We develop new tests for predictability, based on the Lagrange Multiplier [LM] principle, in the context of quantile regression [QR] models which allow for persistent and endogenous predictors driven by conditionally and/or unconditionally 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 it cannot, however, allow for any form of 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. Tests are developed both for predictability at a single quantile, and also jointly over a set of quantiles. Simulation results highlight the superior finite sample size and power properties of our proposed LM tests over the tests of Lee (2016) and Fan and Lee (2019) for both conditionally and unconditionally heteroskedastic errors. 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.

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

JEL classifications: C12, C22, G17.

*The authors thank the 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 useful comments. 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 Economic and Social Research Council of the United Kingdom under research grant ES/R00496X/1. Correspondence to: Robert Taylor, Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: `robert.taylor@essex.ac.uk`

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 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 effects 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, Cenesizoglu and Timmermann (2008), Maynard et al. (2011), Lee (2016), Fan and Lee (2019), Ohno and Ando (2018), Meligkotsidou et al. (2014, 2021) and Cai et al. (2023). 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. Different predictor variables could also be simultaneously predictive but 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. 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 modeling risk and returns using QR yields significant improvement over conditional mean methods.

A further appealing use of quantile predictive regression methods can be found in the emerging literature on economic vulnerability. Specifically, QR are used to characterise the conditional relation between future GDP growth and current financial and economic conditions in order to develop a Growth at Risk (GaR) measure (Adrian et al., 2019). The GaR measure yields the probability of future real GDP growth falling below a pre-specified threshold. Essentially, the analysis first uses QR to establish the relation between future GDP growth and macrofinancial conditions, and then derives the growth distribution by fitting a parametric distribution using the estimated growth quantiles, as suggested by Adrian et al. (2019); see also, *inter alia*, Prasad et al. (2019) and Plagborg-Møller et al. (2020) for further applications and extensions. Because of the tractability of the GaR framework, it has also recently been extended for use in the development of risk measures associated to other macroeconomic indicators such as inflation (see, for example, López-Salido and Loria, 2020) and house prices (see, for example, Alter and Mahoney, 2021).

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. (2011). Consequently, analogous solutions are needed. Accordingly, Maynard et al. (2011) follow the approach of Campbell and Yogo (2006) and use Bonferroni-based methods, under the assumption of homoskedastic errors. Of most relevance 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 quantile predictability 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. In subsequent work, 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. As we will show in the finite sample simulations reported in section 4, this asymptotic property does not appear to translate well into finite samples, with the original Wald test and its MBB implementation both displaying poor finite sample size control for some well-known models of conditional heteroskedasticity when the predictors are highly persistent. We also find that the finite sample power of both the original Wald test of Lee (2016) and its MBB implementation can be very poor relative to the LM-based tests we propose in this paper. Finally, it is important to note that the MBB approach of Fan and Lee (2019) resamples 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.

In order to develop tests for quantile predictability 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) and Fan and Lee (2019) 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 centered dependent variable onto the set of putative lagged predictors and an intercept. The 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. A further advantage of this approach is that heteroskedasticity in the errors is considerably easier to deal with than in the Wald testing framework of Lee (2016), requiring no bootstrap implementation; we also explore the limits of the 2SLS approach in respect of heteroskedastic errors.

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 quantile predictability test and establish its asymptotic null distribution. 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) and its MBB implementation 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 Wald-type QR predictability test of Lee (2016) and its MBB implementation developed in Fan and Lee (2019) - here we also provide corrections to implementation errors present in the MBB algorithm outlined in Fan and Lee (2019); a proof of our main technical result; additional material relating to the empirical application in section 5, and additional Monte Carlo results.

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 the linear predictive QR model

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

where $Q_{y_t}(\tau|\mathcal{F}_{t-1})$ denotes the conditional τ-quantile, $\tau \in (0,1)$, of y_t, $\mathcal{F}_{t-1} = \{y_{t-1}, \boldsymbol{x}_{t-1}', y_{t-2}, \boldsymbol{x}_{t-2}', \ldots\}$ denotes the natural filtration, and $\boldsymbol{x}_t$ is a K-vector of putative predictors. 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 $\mathrm{P}\left[y_t \leq Q_{y_t}(\tau|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}\right] = \tau$. For the present we will focus on conducting tests for predictability at a single quantile, τ. In section 3.3 we will discuss analogous 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,

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

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, $Q_{u_{t,\tau}}(\tau|\mathcal{F}_{t-1}) = 0$. Our interest in this paper is in developing a test of the null hypothesis of no quantile specific predictability, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$, against the two-sided alternative, $H_1 : \boldsymbol{\beta}_\tau \neq \mathbf{0}$ 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, $u_{t,\tau}$, 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, $u_t = \sigma_t \varepsilon_t$ 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 $u_{t,\tau}$.* $\diamondsuit$

Remark 2 *An implication of the strict stationarity assumption on $u_{t,\tau}$ made in Assumption 1, 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 $u_{t,\tau}$.* $\diamondsuit$

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, $\mathrm{E}(\psi_\tau(u_{t,\tau})|\mathcal{F}_{t-1}) = 0$, holds on $s_{t,\tau} := \psi_\tau(u_{t,\tau})$.

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 $\psi_\tau(u_{t,\tau})$ 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).* $\diamondsuit$

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 $\boldsymbol{\xi}_t$ 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*

$$\boldsymbol{\xi}_t = \boldsymbol{\Gamma}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t, \tag{4}$$

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. $|\boldsymbol{I}_K - \lambda\boldsymbol{\Gamma}| = 0$ implies $|\lambda| > C$ for some $C > 1$, where $\boldsymbol{I}_K$ denotes the $K \times K$ identity matrix.*

2. ***Strongly persistent predictors:*** *The predictors exhibit near-integrated dynamics, i.e. $\boldsymbol{\Gamma} := \boldsymbol{I}_K - \frac{1}{T}\boldsymbol{C}$, where $\boldsymbol{C}$ 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) follows Kostakis et al. (2015) and works with linear processes driven by MD innovations (Lee, 2016 imposes conditional homoskedasticity). More recently, Demetrescu et al. (2022, 2023) maintain the linear process assumption with MD innovations but, importantly, additionally allow for unconditional heteroskedasticity. We follow Demetrescu et al. (2022, 2023) 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 or are near-integrated, the latter allowing for pure $I(1)$ predictors, locally stable predictors, and locally explosive predictors, without assuming knowledge of which of these holds. For clarity of exposition, however, we focus our attention on the case where the regressors all have the same degree of persistence, such that they* *are either all weakly or all strongly persistent.*[1] *Notice that in the strongly persistent case this still allows for a mix of pure unit root, locally stationary, and locally explosive predictors.* $\diamondsuit$

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). 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 $\boldsymbol{\gamma}_{t,\tau} := (\mathrm{Cov}(\tilde{\boldsymbol{v}}_t))^{-1}\mathrm{Cov}(\tilde{\boldsymbol{v}}_t, s_{t,\tau})$. 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 $\boldsymbol{h}_{t,\tau} := (\tau(1-\tau))^{-1}\mathrm{Cov}(\tilde{\boldsymbol{v}}_t, s_{t,\tau})$ and $\mathbf{H}_{t,\tau}$ satisfies $\mathbf{H}_{t,\tau}\mathbf{H}'_{t,\tau} = \mathrm{Cov}(\tilde{\boldsymbol{v}}_t - \boldsymbol{h}_{t,\tau}s_{t,\tau})$, such that $\boldsymbol{a}_{t,\tau} = \mathbf{H}^{-1}_{t,\tau}(\tilde{\boldsymbol{v}}_t - \boldsymbol{h}_{t,\tau}s_{t,\tau})$ 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 $\sup_t \mathrm{E}\left\|\mathrm{E}\left(\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_t - \boldsymbol{I}_{K+1}|\boldsymbol{\zeta}_{t-m}, \boldsymbol{\zeta}_{t-m-1}, \ldots\right)\right\| \to 0$, as $m \to \infty$. Finally, we assume that $\sup_{t\in\mathbb{Z}}\left\|\mathrm{E}\left((\tilde{\boldsymbol{v}}_t\tilde{\boldsymbol{v}}'_t - \mathrm{E}(\tilde{\boldsymbol{v}}_t\tilde{\boldsymbol{v}}'_t)) \otimes \tilde{\boldsymbol{v}}_{t-j}\tilde{\boldsymbol{v}}'_{t-k}\right)\right\| \le C(jk)^{-1/2-\vartheta/2}$, for some $\vartheta > 0$.*

Remark 6 *Because $s_{t,\tau}$ is, by construction, bounded, Assumption 4 implies that $\left(\psi_\tau(u_{t,\tau}), \boldsymbol{a}'_{t,\tau}\right)'$ 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] In the context of conventional conditional mean predictive regressions, Demetrescu et al. (2022, Supplementary Material, Section S.2.2) argue that having subsets of predictors belonging to different persistence classes (so that the set of predictors can contain a mix of (marginally) weakly and strongly persistent predictors) is permitted in the case where 2SLS estimation of the predictive regression is employed, and we conjecture that this is the case here too.

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 $\tilde{\boldsymbol{v}}_t$. $\diamondsuit$

Remark 7 *In common with Assumption 3.1 of Fan and Lee (2019), Assumption 4 allows for the presence of conditional heteroskedasticity in the errors. However, and in contrast to 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 $\boldsymbol{v}_t$.* $\diamondsuit$

Remark 8 *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 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.7–D.10 of the supplementary appendix.* $\diamondsuit$

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 "$\Rightarrow$" denotes weak convergence of the associated probability measures, and $(W_\tau(s), \boldsymbol{W}'(s))'$ is a $(K+1)$-vector of independent standard Wiener processes.[2] Where time-invariant endogeneity is present, such that the vector of correlations between $\psi_\tau(u_{t,\tau})$ and $\tilde{\boldsymbol{v}}_t$, and thus between $\psi_\tau(u_{t,\tau})$ 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.

[2] Observe from the quantile-specific decomposition in (5) that $\mathrm{Cov}(\tilde{\boldsymbol{v}}_t)$ may be expressed as $\mathbf{H}_{t,\tau}\mathbf{H}'_{t,\tau} + \tau(1-\tau)\boldsymbol{h}_{t,\tau}\boldsymbol{h}'_{t,\tau}$. This covariance matrix can be shown 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.

More generally, however, Assumption 3 allows these endogeneity correlations to be time-varying, since the marginal correlation of $\psi_\tau(u_{t,\tau})$ 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 $\frac{1}{\sqrt{T}}\boldsymbol{\xi}_{\lfloor sT \rfloor} \Rightarrow \int_0^s e^{-\mathbf{C}(r-s)}\mathrm{d}\boldsymbol{M}(r)$, and 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 Wald-based tests of Lee (2016) and Fan and Lee (2019) – a brief review of the Wald 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, with neither valid, in general, if there is either unconditional heteroskedasticity in the innovations driving the predictors or if time-varying endogeneity is present.

3.1 A Quasi-Likelihood Argument

In order to develop a test of the null hypothesis of no quantile specific predictability, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$, 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.[3] Assuming the scale parameter of the asymmetric Laplace

[3]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.

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

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

where the log-likelihood, $\ell\left(\alpha, \boldsymbol{\beta}\right)$, is given by $\ell\left(\alpha, \boldsymbol{\beta}\right) = C_1 - C_2 \sum_{t=1}^{T} \rho_\tau\left(y_t - \alpha - \boldsymbol{\beta}'\boldsymbol{x}_{t-1}\right)$, 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

$$\nabla\ell\left(\alpha, \boldsymbol{\beta}\right) := -C_2 \sum_{t=1}^{T} \begin{pmatrix} 1 \\ \boldsymbol{x}_{t-1} \end{pmatrix} \psi_\tau\left(y_t - \alpha - \boldsymbol{\beta}'\boldsymbol{x}_{t-1}\right). \tag{7}$$

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 $\boldsymbol{\beta}_\tau = \mathbf{0}$, the sample gradient in (7) 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, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$. This restricted estimator of α_τ, say $\hat{\alpha}_{\tau,0}$, is obtained from

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

In other words, $\hat{\alpha}_{\tau,0}$ 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

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

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 sample moment conditions.[4] 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 $\boldsymbol{\beta}_\tau = \mathbf{0}$ in the QR in (1) by testing the null hypothesis that $\boldsymbol{\delta}_\tau = \mathbf{0}$ in the auxiliary regression,

$$\tilde{s}_{t,\tau} = \boldsymbol{\delta}'_\tau \left(\boldsymbol{x}_{t-1} - \bar{\boldsymbol{x}} \right) + error. \tag{8}$$

Least-squares [LS] estimation of (8) 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.

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 (8). It would also be feasible to base our feasible LM tests on the just-

[4]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 $\sum_{t=1}^{T} \boldsymbol{z}_{I,t-1}\psi_\tau \left(y_t - \alpha_\tau - \boldsymbol{\beta}'_\tau \boldsymbol{x}_{t-1} \right) = \mathbf{0}$ to obtain estimators of the parameters α_τ and $\boldsymbol{\beta}_\tau$, and thence to test the null hypothesis, $\boldsymbol{\beta}_\tau = \mathbf{0}$.

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. However, we choose to focus on 2SLS-based inference because it is known from the results reported in Breitung and Demetrescu (2015) for conditional mean predictability testing, that 2SLS-based tests display superior power properties to methods that use only an IVX instrument in the case where the predictors are strongly persistent. The finite sample power simulations reported in section 4.1.2 below suggest that this property carries over into 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).

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 $\boldsymbol{x}_t$ in the case 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

$$\boldsymbol{z}_{I,t} := \sum_{j=0}^{t-1} \varrho^j \Delta \boldsymbol{x}_{t-j}, \quad \text{with} \quad \varrho := 1 - \frac{a}{T^\eta}, \tag{9}$$

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 (9) 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 predictablity 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,

$$z_{II,t,k} = \sin\left(\frac{\omega_k t}{T}\right), \text{ for } k = 1, \ldots, K, \tag{10}$$

where the ω_k, $k = 1, ..., K$, are distinct spectral frequencies. See Breitung and Demetrescu (2015) and Demetrescu et al. (2022) for further details, including a review of other possible valid choices of both the type-I and type-II instruments.

For this choice of instruments, defining $\boldsymbol{z}_{I,t}$ as in (9) and $\boldsymbol{z}_{II,t} := (z_{II,t,1}, ..., z_{II,t,K})'$, the 2SLS auxiliary estimator is then given by

$$\hat{\boldsymbol{\delta}}_\tau := \left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{C}_T, \tag{11}$$

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 (11), an LM-type test implemented with an Eicker-White heteroskedasticity-consistent (HC) covariance matrix estimation then obtains as

$$\mathcal{T}_\tau := \hat{\boldsymbol{\delta}}_\tau'\widehat{\mathrm{Cov}_{HC}}\left(\hat{\boldsymbol{\delta}}_\tau\right)^{-1}\hat{\boldsymbol{\delta}}_\tau, \tag{12}$$

where the HC covariance matrix estimator is given by

$$\widehat{\mathrm{Cov}_{HC}}\left(\hat{\boldsymbol{\delta}}_\tau\right) := \left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_T\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}, \tag{13}$$

with $\mathbf{D}_T := \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau}^2$. 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 9 *Although we have used Eicker-White HC standard errors in constructing our proposed test, $\mathcal{T}_\tau$ in (12), 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,*

$$\widehat{\mathrm{Cov}}\left(\hat{\boldsymbol{\delta}}_\tau\right) = \hat{\sigma}_{\tilde{s}_{t,\tau}}^2\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}, \text{ with } \hat{\sigma}_{\tilde{s}_{t,\tau}}^2 := \frac{1}{T}\sum_{t=1}^T \tilde{s}_{t,\tau}^2, \tag{14}$$

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 (12), 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.* $\diamondsuit$

Remark 10 *LM-type tests for the significance of subsets of the predictors can also be developed in the obvious way. For example, individual t-type ratios can be considered to test for the predictive power of individual regressors.* $\diamondsuit$

In Proposition 1, whose proof is provided in Part A of the supplementary appendix, we now establish the limiting null distribution of our proposed LM-type statistic, $\mathcal{T}_\tau$, in (12).

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 $T \to \infty$, $\mathcal{T}_\tau \xrightarrow{d} \chi^2_K$.*

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.* $\diamondsuit$

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 assumes 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 $\tilde{s}_{t,\tau}$.* $\diamondsuit$

3.3 Testing for Predictability at Multiple Quantiles

Thus far, in common with the tests developed in Lee (2016) and Fan and Lee (2019), 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 distinct quantile levels each corresponding to a distinct multiple QR of the form in (1) with intercept and slope parameters, α_{τ_j} and $\boldsymbol{\beta}_{\tau_j}$, $j = 1, ..., m$. Then construct,

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

and consider the multivariate auxiliary regression,

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

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 (8)) 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

$$\mathcal{T}_{\boldsymbol{\tau}} := \left(\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}\right)'\left(\widehat{\mathrm{Cov}_{HC}\left(\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}\right)}\right)^{-1}\mathrm{vec}\,\hat{\boldsymbol{\Upsilon}}_{\boldsymbol{\tau}}. \tag{15}$$

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 3–4 assumed to jointly hold on the quantile-specific decompositions for $\tilde{\boldsymbol{v}}_t$, of the generic form in (5), for $\tau_1, \ldots, \tau_m$. 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 13 *In the case where predictability holds at some quantile levels but not at others, then 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 also 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, $\mathcal{T}_{\tau_1}, ..., \mathcal{T}_{\tau_m}$.* ◇

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. The bulk of our results are reported in section 4.1, where we consider finite sample size and local power for a single predictor. Finite sample size simulations for the case of mutliple predictors are then subsequently discussed in section 4.2.

All simulations are preformed in MATLAB, versions R2018b and R2020a, 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 run at the nominal asymptotic 5% significance level, with 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

$$y_t = \alpha_\tau + \beta_\tau x_{t-1} + u_{t,\tau}, \; t = 1, ..., T \tag{16}$$

$$x_t = \mu_x + \xi_t, \;\; \xi_t = \rho \xi_{t-1} + v_t \tag{17}$$

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 (16). Without loss of generality we may 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 (12), implemented with Eicker-White standard errors, and the corresponding test discussed in Remark 9, denoted $\mathcal{T}_\tau^0$ in what follows, constructed using the conventional covariance estimator defined in (14). Both $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ are based on the instrument vector $\boldsymbol{z}_{t-1} := (z_{I,t-1}, z_{II,t-1})'$ with the type-II instrument, $z_{II,t-1}$, defined as in (10) with $k = 1$, and the type-I instrument, $z_{I,t-1}$, given by the IVX choice of Kostakis et al. (2015) defined as in (9), with $a = 1$ and $\eta = 0.95$. Except for the IVX instrument, $z_{I,t-1}$, all variables and instruments entering the estimated predictive regressions are demeaned, as described in section 3.

4.1.1 Empirical Size

This section reports empirical rejection frequencies for $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, IVX_{QR} and IVX_{QR}^{MBB} under the no predictability null hypothesis, $H_0 : \beta_\tau = 0$ in (16), for a range of DGPs featuring serial correlation in the errors driving the predictors and conditional and unconditional heterskedasticity 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 $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}$, where ϕ corresponds to the correlation between u_t and ε_t, which is set as $\phi = -0.95$.[5] 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,

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

[5]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.

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 $\sigma_{it}^2 = 1 + 0.9 e_{i,t-1}^2$, $i = 1,2$ 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 $(\xi_{it}, e_{it})' \sim i.i.d.\ N(0, diag(\sigma_\xi^2, 1))$, independent across i = 1,2. Results are reported for $(\lambda, \sigma_\xi)' = (0.951, 0.314)'$.

DGP3: **Unconditional Heteroskedasticity:** Denoting the time-varying unconditional covariance matrix of $(u_t, v_t)'$ as $\boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{ut}^2 & \phi\sigma_{ut}\sigma_{vt} \\ \phi\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2 \end{bmatrix}$, we allow for a one-time break in the variance of v_t. Specifically, $\sigma_{ut}^2 = 1$ and $\sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + \mathsf{b}\mathbb{I}(t > \lfloor \lambda T \rfloor)$, 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., $\sigma_{ut}^2 = \sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + \mathsf{b}\mathbb{I}(t > \lfloor \lambda T \rfloor)$, again with $\mathsf{b} = 4$ and $\mathsf{b} = 1/4$.

Remark 14 *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 8. Similarly, IVX_{QR}^{MBB} can be validly applied for DGP1, DGP2 and DGP3, but not DGP4, while IVX_{QR} 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 $c \in \{0, -2.5, -10, -0.5T\}$, covering a spectrum of exact unit root ($c = 0$), local-to-unity ($c = -2.5, -10$), and weakly stationary ($c = -0.5T$) predictors. Corresponding results for $T = 750$ can be found in Tables D.1 - D.6 of the supplementary appendix.

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} test is significantly over-sized when testing for predictability at the upper and lower tails of the distribution in the case where x_t is very 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 $c = -2.5$. These size distortions are worse, other things equal, the further out one goes into the tails, and 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, $\tau = 0.1$ and $\tau = 0.9$, it actually displays larger size distortions than IVX_{QR} away from the tails, $\tau = 0.3, 0.4., ..., 0.7$. From Table D.1, 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 $T = 750$. In contrast, for $c \leq -10$ both IVX_{QR} and IVX_{QR}^{MBB} show decent size control.

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 three GARCH parameter constellations considered. 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 $c = -2.5$), 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.2 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 and D.3 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, $c = 0$.

We next turn to the results for the unconditional heteroskedastic case. Tables 4 and D.4 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.5 and D.6 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. 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.7–D.10 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} and IVX_{QR}^{MBB} 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} and IVX_{QR}^{MBB} 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 $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, IVX_{QR} and IVX_{QR}^{MBB}. To that end, we simulate data from DGP (16)-(17) 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 $\pi = \theta = 0$. 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 (16) 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, $b = \{0, 1, ..., 26\}$. 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.

Figure 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.1 in the supplementary appendix, are almost identical to those in Figure 1. A number of features can be observed from these results. First, in terms of a comparison between the Wald-type IVX_{QR} and IVX_{QR}^{MBB} tests we see that the former is significantly more powerful than the latter for $c = 0$, and to a lesser extent for $c = -2.5$, while there is relatively little to choose between them on power for $c \leq -10$. It is of course important to recall, from the discussion surrounding Table 1 that IVX_{QR} (IVX_{QR}^{MBB}) is not well size controlled for $c = 0$ and $c = -2.5$ when $\tau = 0.2, 0, 8$ ($\tau = 0.3, 0.4., ..., 0.7$), leading to artificially inflated power in both cases. In contrast, the local power functions of $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ are almost identical throughout. Second, for strongly persistent predictors our proposed $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ tests can be seen to enjoy considerable power advantages over the IVX_{QR} and IVX_{QR}^{MBB} tests, despite the aforementioned lack of size control of the the latter two tests. 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 $c = -2.5$. 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% for $\tau = 0.5$; while IVX_{QR}^{MBB} has power of about 47% for $\tau = 0.2$ and $\tau = 0.8$, and 42% 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 $\tau = 0.5$; while IVX_{QR}^{MBB} has power of 36% for all three values of τ. Third, for all of the tests, power varies to some degree with τ; 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} and IVX_{QR}^{MBB}. Finally, for the weakly stationary case, $c = -0.5T$, the local power functions of all four tests are seen to be essentially indistinguishable.

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 that of the IVX_{QR} test.[6] For our analysis we generate our simulation data according to the DGP

$$y_t = \alpha_\tau + \boldsymbol{\beta}_\tau' \boldsymbol{x}_{t-1} + u_{t,\tau} \tag{18}$$

$$\boldsymbol{x}_t = \boldsymbol{\mu}_x + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t = \boldsymbol{\Gamma}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t \tag{19}$$

with $\boldsymbol{\xi}_0 = \mathbf{0}$ and, like in the $K = 1$ DGP in (16), $u_{t,\tau} = u_t - Q_{u_t}(\tau|\mathcal{F}_{t-1})$. The innovations are generated as,

$$u_t = \gamma f_t + a_t \tag{20}$$

$$\boldsymbol{v}_t = -\gamma f_t + \boldsymbol{e}_t \tag{21}$$

[6] 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}.

where f_t is a common factor to both innovations, which we generate as a standard normal variate, i.e., f_t is a $T \times 1$ vector of independent standard normal random variables, $\gamma = -0.95$ and $(a_t, \boldsymbol{e}_t')' \sim$ *i.i.d.* $N(\mathbf{0}, \boldsymbol{I}_{K+1})$. We set $\alpha_\tau = 0$ and $\boldsymbol{\mu}_x = \mathbf{0}$, without loss of generality. The predictors are generated setting $\boldsymbol{\Gamma} = \rho \boldsymbol{I}_K$, 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 (18).

In computing $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$, we used the instrument vector $\boldsymbol{z}_{t-1} := (\boldsymbol{z}'_{I,t-1}, \boldsymbol{z}'_{II,t-1})'$ with the K elements of the type-II instrument vector, $\boldsymbol{z}_{II,t-1}$, defined as in (10) for $k = 1, ..., K$, and the K elements of the type-I instrument vector, $\boldsymbol{z}_{I,t-1}$, given by the IVX defined as in (9), 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 as this was found to improve the finite sample properties of the tests.[7]

Table 5 reports empirical null rejection frequencies for $T = 250$ and for $K = \{2, 3, 4, 5\}$, for the $\mathcal{T}_\tau$, $\mathcal{T}_\tau^0$, and IVX_{QR} tests when $\boldsymbol{\beta}_\tau = \mathbf{0}$ in (18). Corresponding results for $T = 750$ can be found in Table D.11 in the supplementary appendix. The results show that $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ display decent size control throughout, albeit for $T = 250$ there is a small degree of over-size for $K = 4, 5$ for quantiles near the centre of the distribution when $c = 0$, although these distortions are largely eliminated for $T = 750$. In contrast, the empirical size properties of IVX_{QR} deteriorate significantly as K increases, most notably in the more extreme quantiles, $\tau < 0.3$ and $\tau > 0.7$, with over-sizing present in such cases for all values of c. For example, for $T = 250$ and $c = 0$ the IVX_{QR} test for $\tau = 0.1$ has size 13.6% and 16.5% for $K = 3$ and $K = 4$, respectively, while $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^0$ have sizes 4.6% and 4.8% respectively for $K = 3$, and 4.8% and 5.0% respectively for $K = 4$.

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), 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

[7] In the present context, this entails replacing $\mathbf{D}_T$ in (13) by $(\mathbf{D}_T - \boldsymbol{\Xi})$, where $\boldsymbol{\Xi} := T(\bar{\tilde{\boldsymbol{z}}}_{-1}\bar{\tilde{\boldsymbol{z}}}'_{-1})(\hat{\sigma}^2_{u_\tau} - \hat{\boldsymbol{\Omega}}_{\boldsymbol{v}u_\tau}\hat{\boldsymbol{\Omega}}_{\boldsymbol{vv}}\hat{\boldsymbol{\Omega}}'_{\boldsymbol{v}u_\tau})$, with $\bar{\tilde{\boldsymbol{z}}}_{-1} = T^{-1}\sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}$, $\tilde{\boldsymbol{z}}_{t-1} := (\boldsymbol{z}'_{I,t-1}, \tilde{\boldsymbol{z}}'_{II,t-1})'$, and $\tilde{\boldsymbol{z}}'_{II,t-1}$ is a vector of demeaned type II instruments. Note that the mean of $\tilde{\boldsymbol{z}}_{II,t-1}$ will be a zero vector, so that $(\bar{\tilde{\boldsymbol{z}}}_{-1}\bar{\tilde{\boldsymbol{z}}}'_{-1}) = \text{diag}\{(\bar{\boldsymbol{z}}_{I,-1}\bar{\boldsymbol{z}}'_{I,-1}), \mathbf{0}\}$. Furthermore, $\hat{\boldsymbol{\Omega}}_{\boldsymbol{vv}}$ and $\hat{\boldsymbol{\Omega}}_{\boldsymbol{v}u_\tau}$ 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 $\mathcal{T}_\tau^0$.

five different sources: (1) the (updated) monthly dataset used in Welch and Goyal (2008); (2) a time varying tail index which we compute from 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 of Lee (2016) and Fan and Lee (2019) 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 14). 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. 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 8) 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}, and IVX_{QR}^{MBB} quantile predictability tests to each of the twenty-two predictors at each of nine different quantiles, $\tau = 0.1, 0.2, ..., 0.9$, together with the joint test $\mathcal{T}_{\boldsymbol{\tau}}$ from section 3.3 across these nine quantiles. This 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 vix_t and $ACMTP10_t$. While IVX_{QR}, and IVX_{QR}^{MBB} 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 information on market or firm risk, or market uncertainty where the quantile predictablity 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 realized 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 the log earnings price ratio (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 $\tau = 0.1$ 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 and rv_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. Interestingly, $ACMTP10_t$ appears to display predictive power only in the centre and shoulders of the return distribution ($0.3 \leq \tau \leq 0.6$).

These empirical results in Table 7 highlight that for a number of the predictors considered, particularly those encompassing 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 quantile predictability for financial returns data. Our proposed tests can be applied to models with either a single or multiple putative predictors and can be used to test for no predictability at either a single quantile or jointly over a specified set of distinct quantiles. 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 Wald-type tests of Lee (2016), and their moving blocks bootstrap implementation developed in Fan and Lee (2019), our LM-type tests are also asymptotically valid in the presence of unconditional heteroskedasticity in the errors driving the predictors. Monte Carlo simulations suggest that, for strongly persistent predictors, our proposed tests display significantly more robust finite sample size properties coupled with markedly superior power profiles to the Wald-type tests and their bootstrap implementation. 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.

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019). Vulnerable growth. *American Economic Review 109*(4), 1263–1289.

Alter, A. and E. M. Mahoney (2021). Local house-price vulnerability: Evidence from the U.S. and Canada. *Journal of Housing Economics 54*, 101791.

Bollerslev, T., G. Tauchen, and H. Zhou (2009, 02). Expected stock returns and variance risk premia. *The Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2014). Time-varying jump tails. *Journal of Econometrics 183*(2), 168–180.

Boswijk, H. P., G. Cavaliere, A. Rahbek, and A. M. R. Taylor (2016). Inference on co-integration parameters in heteroskedastic vector autoregressions. *Journal of Econometrics 192*(1), 64–85.

Breitung, J. and M. Demetrescu (2015). Instrumental variable and variable addition based inference in predictive regressions. *Journal of Econometrics 187*(1), 358–375.

Cai, Z., H. Chen, and X. Liao (2023). A new robust inference for predictive quantile regression. *Journal of Econometrics 234*(1), 227–250.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*(1), 27–60.

Cavaliere, G. and A. Taylor (2008). Time-transformed unit root tests for models with non-stationary volatility. *Journal of Time Series Analysis 29*(2), 300–330.

Cavanagh, C. L., G. Elliott, and J. H. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory 11*(5), 1131–1147.

Cederburg, S., T. L. Johnson, and M. S. O'Doherty (2022). On the economic significance of stock return predictability. *Review of Finance 27*(2), 619–657.

Cenesizoglu, T. and A. Timmermann (2008). Is the distribution of stock returns predictable? SSRN working paper 1107185.

Chevapatrakul, T., Z. Xu, and K. Yao (2019). The impact of tail risk on stock market returns: The role of market sentiment. *International Review of Economics & Finance 59*, 289–301.

Demetrescu, M., I. Georgiev, P. M. M. Rodrigues, and A. M. R. Taylor (2022). Testing for episodic predictability in stock returns. *Journal of Econometrics 227*(1), 85–113.

Demetrescu, M., I. Georgiev, P. M. M. Rodrigues, and A. M. R. Taylor (2023). Extensions to IVX methods of inference for return predictability. *Journal of Econometrics forthcoming.*

Elliott, G., U. K. Müller, and M. W. Watson (2015). Nearly optimal tests when a nuisance parameter is present under the null hypothesis. *Econometrica 83*(2), 771–811.

Fan, R. and J. H. Lee (2019). Predictive quantile regressions under persistence and conditional heteroskedasticity. *Journal of Econometrics 213*(1), 261–280.

Goyal, A. and I. Welch (2003). Predicting the equity premium with dividend ratios. *Management Science 49*(5), 639–654.

Granger, C. W. J. (1999). Outline of forecast theory using generalized cost functions. *Spanish Economic Review 1*(2), 161–173.

Gungor, S. and R. Luger (2021). Exact inference in long-horizon predictive quantile regressions with an application to stock returns. *Journal of Financial Econometrics 19*(4), 746–788.

Jansson, M. and M. J. Moreira (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica 74*(3), 681–714.

Kelly, B. and H. Jiang (2014). Tail Risk and Asset Prices. *The Review of Financial Studies 27*(10), 2841–2871.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *Journal of Econometrics 128*(1), 137–164.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *Review of Financial Studies 28*(5), 1506–1553.

Lee, J. H. (2016). Predictive quantile regression with persistent covariates: IVX-QR approach. *Journal of Econometrics 192*(1), 105–118.

López-Salido, J. D. and F. Loria (2020). Inflation at Risk. Finance and Economics Discussion Series 2020-013, Board of Governors of the Federal Reserve System.

Maynard, A. and K. Shimotsu (2009). Covariance-based orthogonality tests for regressors with unknown persistence. *Econometric Theory 25*(1), 63–116.

Maynard, A., K. Shimotsu, and Y. Wang (2011). Inference in predictive quantile regressions. Unpublished manuscript.

Meligkotsidou, L., E. Panopoulou, I. D. Vrontos, and S. D. Vrontos (2014). A quantile regression approach to equity premium prediction. *Journal of Forecasting 33*(7), 558–576.

Meligkotsidou, L., E. Panopoulou, I. D. Vrontos, and S. D. Vrontos (2021). Out-of-sample equity premium prediction: A complete subset quantile regression approach. *The European Journal of Finance 27*(1-2), 110–135.

Montiel Olea, J. L. and M. Plagborg-Møller (2019). Simultaneous confidence bands: Theory, implementation, and an application to svars. *Journal of Applied Econometrics 34*(1), 1–17.

Nelson, C. R. and M. J. Kim (1993). Predictable stock returns: The role of small sample bias. *The Journal of Finance 48*(2), 641–661.

Ohno, S. and T. Ando (2018). Stock return predictability: A factor-augmented predictive regression system with shrinkage method. *Econometric Reviews 37*(1), 29–60.

Paye, B. S. and A. Timmermann (2006). Instability of return prediction models. *Journal of Empirical Finance 13*(3), 274–315.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020). When is growth at risk? *Brookings Papers on Economic Activity 2020*(Spring), 167–229.

Portnoy, S. (1991). Asymptotic behavior of regression quantiles in non-stationary, dependent cases. *Journal of Multivariate Analysis 38*(1), 100–113.

Prasad, A., P. Jeasakul, A. Alter, R. Lafarguette, A. X. Feng, S. A. Elekdag, and C. Wang (2019). Growth at Risk: Concept and application in IMF country surveillance. IMF Working Papers 2019/036, International Monetary Fund.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance 44*(5), 1115–1153.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics 54*(3), 375–421.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

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}, and IVX_{QR}^{MBB}. 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}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$(\pi,\theta) = (0,0)$				$(\pi,\theta) = (0.5,0)$				$(\pi,\theta) = (-0.5,0)$				$(\pi,\theta) = (0,0.5)$				$(\pi,\theta) = (0,-0.5)$			
0	0.1	0.051	0.055	0.085	0.055	0.051	0.058	0.094	0.057	0.049	0.055	0.094	0.055	0.047	0.049	0.080	0.053	0.043	0.050	0.084	0.051
	0.2	0.052	0.056	0.067	0.069	0.053	0.058	0.078	0.074	0.052	0.058	0.072	0.073	0.052	0.058	0.073	0.070	0.050	0.056	0.064	0.066
	0.3	0.059	0.060	0.077	0.093	0.057	0.059	0.072	0.082	0.055	0.060	0.066	0.077	0.052	0.055	0.065	0.078	0.053	0.057	0.057	0.069
	0.4	0.057	0.062	0.060	0.075	0.059	0.065	0.068	0.084	0.059	0.064	0.064	0.077	0.055	0.061	0.072	0.086	0.057	0.061	0.068	0.082
	0.5	0.059	0.064	0.054	0.073	0.061	0.065	0.066	0.080	0.058	0.062	0.061	0.080	0.058	0.061	0.073	0.088	0.057	0.060	0.064	0.086
	0.6	0.058	0.061	0.062	0.078	0.062	0.065	0.073	0.084	0.056	0.059	0.065	0.078	0.060	0.062	0.071	0.084	0.057	0.061	0.063	0.074
	0.7	0.056	0.057	0.075	0.092	0.053	0.057	0.062	0.078	0.053	0.055	0.064	0.075	0.054	0.057	0.063	0.074	0.051	0.056	0.055	0.065
	0.8	0.056	0.058	0.065	0.070	0.057	0.058	0.071	0.080	0.056	0.058	0.072	0.074	0.059	0.065	0.072	0.074	0.057	0.060	0.070	0.068
	0.9	0.052	0.053	0.085	0.054	0.053	0.053	0.091	0.051	0.051	0.052	0.092	0.057	0.052	0.050	0.086	0.056	0.049	0.047	0.079	0.054
-2.5	0.1	0.043	0.046	0.070	0.045	0.045	0.048	0.076	0.048	0.040	0.043	0.074	0.048	0.042	0.045	0.068	0.045	0.037	0.041	0.066	0.040
	0.2	0.046	0.050	0.052	0.053	0.047	0.049	0.060	0.061	0.048	0.050	0.058	0.059	0.043	0.046	0.055	0.057	0.041	0.044	0.051	0.053
	0.3	0.048	0.052	0.065	0.071	0.047	0.053	0.052	0.064	0.048	0.050	0.056	0.065	0.042	0.045	0.051	0.064	0.044	0.047	0.047	0.062
	0.4	0.054	0.057	0.048	0.062	0.051	0.054	0.053	0.067	0.054	0.057	0.052	0.064	0.047	0.049	0.053	0.068	0.044	0.047	0.054	0.066
	0.5	0.050	0.051	0.045	0.066	0.049	0.051	0.051	0.067	0.047	0.050	0.044	0.065	0.045	0.047	0.055	0.069	0.047	0.050	0.052	0.069
	0.6	0.050	0.052	0.051	0.067	0.049	0.053	0.049	0.068	0.045	0.048	0.050	0.066	0.047	0.049	0.050	0.064	0.048	0.051	0.047	0.060
	0.7	0.046	0.049	0.054	0.070	0.044	0.048	0.053	0.065	0.046	0.050	0.050	0.065	0.046	0.050	0.050	0.062	0.044	0.047	0.044	0.055
	0.8	0.044	0.050	0.066	0.057	0.048	0.052	0.063	0.061	0.044	0.049	0.060	0.061	0.043	0.050	0.052	0.055	0.045	0.049	0.050	0.054
	0.9	0.042	0.045	0.073	0.049	0.040	0.044	0.073	0.050	0.041	0.045	0.070	0.052	0.044	0.045	0.072	0.047	0.043	0.042	0.071	0.046
-10	0.1	0.043	0.046	0.064	0.042	0.043	0.047	0.061	0.043	0.044	0.045	0.062	0.040	0.044	0.048	0.060	0.039	0.043	0.047	0.066	0.039
	0.2	0.047	0.053	0.049	0.048	0.048	0.053	0.047	0.047	0.046	0.053	0.052	0.045	0.049	0.052	0.051	0.051	0.047	0.049	0.042	0.046
	0.3	0.051	0.054	0.048	0.058	0.051	0.053	0.043	0.053	0.049	0.053	0.046	0.051	0.056	0.056	0.044	0.053	0.050	0.053	0.043	0.051
	0.4	0.055	0.058	0.039	0.055	0.053	0.056	0.042	0.054	0.056	0.057	0.039	0.054	0.051	0.054	0.043	0.058	0.054	0.055	0.045	0.056
	0.5	0.050	0.051	0.038	0.056	0.050	0.053	0.041	0.058	0.049	0.050	0.035	0.055	0.053	0.054	0.046	0.061	0.048	0.050	0.044	0.055
	0.6	0.058	0.061	0.041	0.054	0.058	0.060	0.043	0.058	0.056	0.058	0.040	0.054	0.055	0.058	0.043	0.055	0.048	0.049	0.041	0.053
	0.7	0.053	0.057	0.044	0.058	0.050	0.054	0.044	0.053	0.057	0.058	0.039	0.051	0.050	0.053	0.044	0.046	0.046	0.049	0.044	0.050
	0.8	0.054	0.056	0.050	0.046	0.053	0.057	0.054	0.049	0.051	0.054	0.051	0.050	0.053	0.057	0.044	0.048	0.051	0.054	0.046	0.042
	0.9	0.048	0.055	0.064	0.038	0.048	0.056	0.069	0.043	0.048	0.053	0.064	0.043	0.048	0.053	0.068	0.041	0.045	0.050	0.062	0.039
-0.5T	0.1	0.048	0.053	0.068	0.037	0.047	0.050	0.066	0.039	0.050	0.051	0.067	0.039	0.040	0.047	0.063	0.035	0.044	0.046	0.058	0.033
	0.2	0.048	0.050	0.047	0.043	0.049	0.055	0.047	0.040	0.055	0.055	0.045	0.043	0.047	0.053	0.043	0.038	0.050	0.049	0.048	0.041
	0.3	0.051	0.055	0.036	0.041	0.052	0.056	0.041	0.046	0.048	0.049	0.042	0.039	0.049	0.051	0.041	0.044	0.049	0.054	0.046	0.041
	0.4	0.051	0.053	0.038	0.045	0.051	0.053	0.036	0.052	0.049	0.051	0.039	0.041	0.053	0.054	0.038	0.051	0.053	0.056	0.042	0.048
	0.5	0.047	0.049	0.036	0.041	0.049	0.050	0.036	0.043	0.049	0.051	0.034	0.042	0.056	0.058	0.041	0.051	0.054	0.056	0.040	0.045
	0.6	0.047	0.049	0.038	0.044	0.048	0.049	0.039	0.041	0.046	0.048	0.043	0.041	0.053	0.053	0.047	0.046	0.051	0.051	0.045	0.044
	0.7	0.052	0.054	0.041	0.047	0.046	0.049	0.038	0.046	0.054	0.055	0.043	0.042	0.049	0.053	0.043	0.047	0.054	0.057	0.047	0.043
	0.8	0.050	0.052	0.047	0.044	0.042	0.048	0.046	0.038	0.051	0.052	0.045	0.041	0.049	0.057	0.053	0.044	0.052	0.053	0.053	0.044
	0.9	0.046	0.052	0.064	0.037	0.045	0.053	0.061	0.036	0.046	0.054	0.069	0.037	0.048	0.055	0.062	0.037	0.047	0.054	0.065	0.038

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}, and IVX_{QR}^{MBB}. 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}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\theta_1 = 0.1$, $\theta_2 = 0.5$				$\theta_1 = 0.1$, $\theta_2 = 0.8$				$\theta_1 = 0.05$, $\theta_2 = 0.9$			
0	0.1	0.049	0.052	0.085	0.056	0.049	0.052	0.091	0.055	0.049	0.053	0.090	0.056
	0.2	0.056	0.061	0.070	0.071	0.055	0.062	0.071	0.072	0.053	0.060	0.070	0.072
	0.3	0.060	0.064	0.078	0.095	0.060	0.063	0.082	0.093	0.060	0.063	0.078	0.092
	0.4	0.059	0.063	0.055	0.077	0.058	0.064	0.061	0.080	0.056	0.062	0.055	0.079
	0.5	0.058	0.060	0.052	0.071	0.056	0.058	0.058	0.071	0.055	0.058	0.055	0.070
	0.6	0.060	0.064	0.055	0.073	0.061	0.064	0.062	0.073	0.056	0.061	0.060	0.071
	0.7	0.057	0.059	0.074	0.086	0.059	0.061	0.077	0.089	0.059	0.062	0.077	0.086
	0.8	0.056	0.060	0.071	0.064	0.055	0.061	0.071	0.068	0.054	0.059	0.072	0.067
	0.9	0.048	0.055	0.087	0.052	0.052	0.053	0.091	0.058	0.050	0.055	0.091	0.056
-2.5	0.1	0.039	0.044	0.073	0.046	0.040	0.044	0.081	0.046	0.039	0.042	0.076	0.048
	0.2	0.044	0.048	0.057	0.055	0.043	0.048	0.059	0.059	0.042	0.047	0.058	0.057
	0.3	0.046	0.047	0.057	0.071	0.045	0.047	0.060	0.068	0.045	0.047	0.061	0.075
	0.4	0.049	0.050	0.044	0.061	0.047	0.051	0.047	0.063	0.047	0.049	0.045	0.060
	0.5	0.046	0.049	0.042	0.056	0.047	0.049	0.046	0.059	0.045	0.048	0.045	0.060
	0.6	0.047	0.050	0.041	0.057	0.047	0.050	0.050	0.059	0.046	0.050	0.045	0.059
	0.7	0.047	0.051	0.059	0.071	0.048	0.053	0.059	0.068	0.045	0.051	0.059	0.066
	0.8	0.044	0.051	0.051	0.050	0.045	0.050	0.055	0.050	0.044	0.049	0.054	0.051
	0.9	0.044	0.046	0.068	0.046	0.042	0.048	0.074	0.047	0.043	0.046	0.070	0.048
-10	0.1	0.049	0.051	0.067	0.041	0.047	0.051	0.068	0.039	0.046	0.048	0.068	0.039
	0.2	0.053	0.054	0.047	0.045	0.054	0.057	0.050	0.046	0.053	0.056	0.050	0.047
	0.3	0.053	0.055	0.043	0.058	0.054	0.056	0.048	0.056	0.052	0.056	0.044	0.055
	0.4	0.048	0.050	0.038	0.047	0.052	0.054	0.040	0.048	0.048	0.051	0.041	0.046
	0.5	0.048	0.050	0.036	0.050	0.052	0.053	0.039	0.051	0.050	0.053	0.041	0.052
	0.6	0.051	0.052	0.038	0.051	0.048	0.050	0.042	0.050	0.051	0.054	0.039	0.050
	0.7	0.052	0.054	0.048	0.054	0.051	0.054	0.051	0.056	0.052	0.054	0.050	0.054
	0.8	0.050	0.054	0.045	0.043	0.048	0.052	0.049	0.044	0.050	0.053	0.048	0.043
	0.9	0.043	0.049	0.059	0.037	0.045	0.046	0.065	0.040	0.045	0.048	0.062	0.041
$-0.5T$	0.1	0.046	0.055	0.062	0.038	0.046	0.052	0.069	0.039	0.047	0.054	0.064	0.039
	0.2	0.051	0.052	0.046	0.044	0.051	0.055	0.051	0.043	0.051	0.053	0.050	0.042
	0.3	0.050	0.051	0.042	0.041	0.050	0.052	0.046	0.042	0.049	0.053	0.044	0.043
	0.4	0.050	0.052	0.042	0.045	0.050	0.052	0.047	0.047	0.052	0.053	0.045	0.044
	0.5	0.047	0.048	0.037	0.044	0.048	0.049	0.044	0.046	0.048	0.050	0.039	0.044
	0.6	0.050	0.052	0.041	0.046	0.051	0.053	0.046	0.047	0.051	0.053	0.042	0.047
	0.7	0.047	0.050	0.040	0.042	0.048	0.050	0.049	0.044	0.046	0.049	0.042	0.043
	0.8	0.045	0.051	0.043	0.043	0.046	0.050	0.052	0.043	0.046	0.052	0.048	0.041
	0.9	0.043	0.050	0.064	0.037	0.045	0.049	0.071	0.035	0.045	0.049	0.069	0.036

Table 3: Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker-White standard errors), $\mathcal{T}^0_\tau$ (conventional standard errors), IVX_{QR}, and IVX^{MBB}_{QR}. 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 $(u_t, v_t)' = [\sigma_{1t} \quad 0; 0 \quad \sigma_{2t}]\,\boldsymbol{\eta}_t$; $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{\Omega})$ with $\boldsymbol{\Omega} = [1 \quad -0.95; -0.95 \quad 1]$ and $\sigma^2_{it} = 1 + 0.9e^2_{i,t-1}$, $i = 1, 2$. 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 $(\xi_{it}, e_{it})' \sim i.i.d.\ N(0, \mathrm{diag}(\sigma^2_\xi, 1))$, independent across $i = 1, 2$, with results reported for $(\lambda, \sigma_\xi)' = (0.951, 0.314)'$.

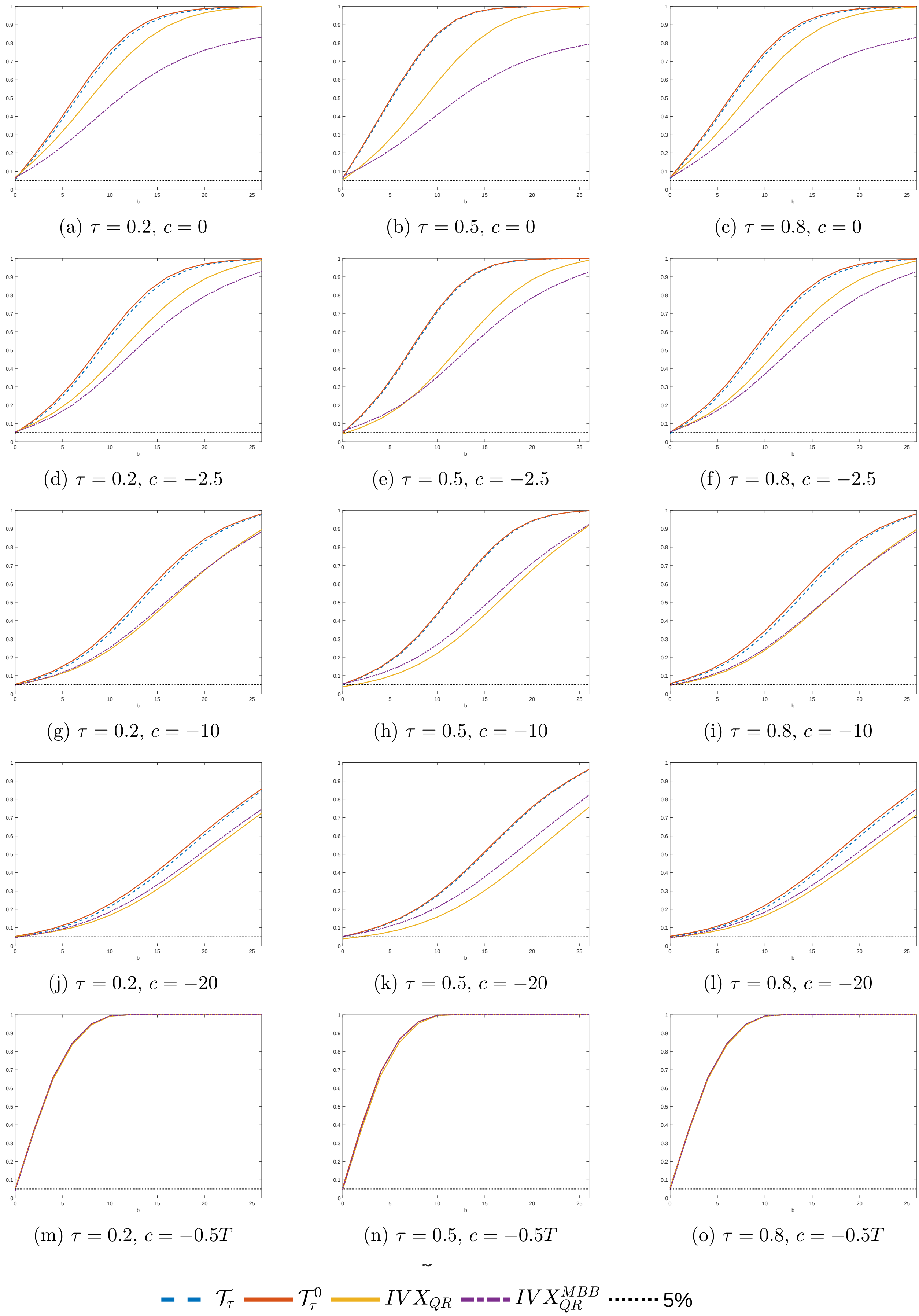
c	τ	$\mathcal{T}_\tau$	$\mathcal{T}^0_\tau$	IVX_{QR}	IVX^{MBB}_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}^0_\tau$	IVX_{QR}	IVX^{MBB}_{QR}
		$ARCH(1)$				SV			
0	0.1	0.046	0.045	0.103	0.053	0.048	0.054	0.096	0.058
	0.2	0.050	0.053	0.085	0.069	0.060	0.065	0.078	0.065
	0.3	0.053	0.055	0.078	0.074	0.060	0.062	0.067	0.078
	0.4	0.052	0.055	0.076	0.084	0.060	0.063	0.069	0.087
	0.5	0.049	0.051	0.075	0.079	0.055	0.058	0.070	0.089
	0.6	0.051	0.053	0.072	0.077	0.056	0.058	0.073	0.090
	0.7	0.048	0.051	0.077	0.072	0.060	0.065	0.069	0.082
	0.8	0.049	0.051	0.079	0.064	0.059	0.063	0.077	0.069
	0.9	0.043	0.046	0.102	0.051	0.054	0.054	0.103	0.058
-2.5	0.1	0.046	0.041	0.098	0.050	0.041	0.044	0.083	0.047
	0.2	0.043	0.045	0.072	0.060	0.045	0.052	0.063	0.054
	0.3	0.043	0.043	0.065	0.063	0.047	0.053	0.051	0.061
	0.4	0.044	0.046	0.062	0.065	0.049	0.050	0.053	0.067
	0.5	0.039	0.042	0.063	0.063	0.052	0.054	0.051	0.073
	0.6	0.043	0.045	0.065	0.060	0.056	0.059	0.056	0.071
	0.7	0.043	0.045	0.061	0.063	0.056	0.059	0.050	0.067
	0.8	0.042	0.041	0.071	0.057	0.053	0.056	0.058	0.061
	0.9	0.045	0.043	0.087	0.046	0.051	0.050	0.090	0.051
-10	0.1	0.056	0.058	0.102	0.045	0.047	0.055	0.073	0.043
	0.2	0.053	0.051	0.073	0.050	0.046	0.051	0.053	0.046
	0.3	0.051	0.053	0.068	0.055	0.045	0.048	0.041	0.052
	0.4	0.053	0.055	0.062	0.053	0.053	0.055	0.043	0.053
	0.5	0.044	0.045	0.062	0.051	0.054	0.057	0.042	0.059
	0.6	0.048	0.049	0.062	0.051	0.056	0.059	0.043	0.062
	0.7	0.049	0.051	0.068	0.049	0.055	0.058	0.042	0.051
	0.8	0.046	0.049	0.077	0.046	0.051	0.054	0.051	0.047
	0.9	0.047	0.045	0.099	0.044	0.048	0.050	0.076	0.040
$-0.5T$	0.1	0.044	0.049	0.194	0.043	0.047	0.050	0.066	0.041
	0.2	0.049	0.052	0.164	0.048	0.043	0.052	0.045	0.036
	0.3	0.049	0.052	0.159	0.051	0.052	0.053	0.040	0.046
	0.4	0.051	0.052	0.153	0.056	0.049	0.054	0.033	0.048
	0.5	0.048	0.049	0.145	0.053	0.050	0.052	0.035	0.042
	0.6	0.049	0.051	0.148	0.056	0.053	0.056	0.034	0.048
	0.7	0.047	0.052	0.151	0.049	0.053	0.055	0.036	0.048
	0.8	0.049	0.053	0.166	0.050	0.043	0.047	0.046	0.041
	0.9	0.038	0.049	0.189	0.043	0.043	0.049	0.075	0.038

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 $\mathbf{\Sigma}_t = \begin{bmatrix}\sigma_{ut}^2 & -0.95\sigma_{ut}\sigma_{vt}; & -0.95\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2\end{bmatrix}$ and $\sigma_{ut}^2 = 1$, $\sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + 1/4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.046	0.050	0.093	0.055	0.047	0.050	0.090	0.051	0.049	0.052	0.091	0.058
	0.2	0.051	0.054	0.074	0.065	0.056	0.058	0.073	0.061	0.054	0.059	0.074	0.066
	0.3	0.049	0.053	0.088	0.084	0.059	0.062	0.083	0.091	0.053	0.056	0.085	0.091
	0.4	0.052	0.057	0.092	0.098	0.057	0.063	0.081	0.097	0.054	0.059	0.072	0.090
	0.5	0.051	0.054	0.084	0.103	0.061	0.064	0.079	0.091	0.060	0.064	0.065	0.088
	0.6	0.059	0.062	0.085	0.104	0.064	0.068	0.089	0.098	0.063	0.066	0.075	0.091
	0.7	0.054	0.059	0.080	0.092	0.060	0.065	0.083	0.087	0.059	0.064	0.086	0.095
	0.8	0.057	0.061	0.083	0.070	0.061	0.064	0.068	0.069	0.060	0.066	0.067	0.065
	0.9	0.048	0.047	0.100	0.058	0.047	0.051	0.099	0.057	0.050	0.056	0.094	0.052
-2.5	0.1	0.043	0.047	0.068	0.039	0.037	0.038	0.068	0.042	0.041	0.048	0.072	0.049
	0.2	0.040	0.041	0.059	0.050	0.040	0.042	0.059	0.049	0.043	0.044	0.057	0.052
	0.3	0.039	0.041	0.054	0.058	0.042	0.044	0.055	0.060	0.042	0.044	0.060	0.066
	0.4	0.043	0.046	0.059	0.075	0.041	0.043	0.057	0.070	0.044	0.047	0.050	0.070
	0.5	0.043	0.045	0.054	0.069	0.042	0.044	0.051	0.063	0.043	0.045	0.049	0.067
	0.6	0.044	0.046	0.056	0.074	0.044	0.046	0.058	0.071	0.041	0.043	0.049	0.067
	0.7	0.046	0.050	0.056	0.069	0.046	0.048	0.053	0.066	0.045	0.048	0.056	0.067
	0.8	0.043	0.043	0.059	0.059	0.046	0.046	0.054	0.053	0.039	0.043	0.056	0.052
	0.9	0.040	0.042	0.073	0.049	0.040	0.042	0.072	0.044	0.040	0.043	0.071	0.047
-10	0.1	0.051	0.054	0.065	0.037	0.045	0.050	0.067	0.039	0.052	0.055	0.064	0.042
	0.2	0.047	0.049	0.052	0.045	0.047	0.051	0.045	0.045	0.048	0.053	0.046	0.046
	0.3	0.045	0.049	0.045	0.053	0.047	0.051	0.043	0.050	0.045	0.047	0.044	0.052
	0.4	0.051	0.053	0.045	0.062	0.048	0.050	0.037	0.056	0.052	0.055	0.039	0.058
	0.5	0.053	0.056	0.046	0.060	0.049	0.051	0.037	0.058	0.049	0.050	0.037	0.059
	0.6	0.057	0.059	0.046	0.065	0.052	0.052	0.041	0.054	0.050	0.051	0.039	0.056
	0.7	0.053	0.058	0.045	0.057	0.049	0.051	0.045	0.048	0.049	0.049	0.041	0.053
	0.8	0.049	0.052	0.047	0.051	0.051	0.055	0.043	0.046	0.044	0.049	0.045	0.046
	0.9	0.048	0.050	0.064	0.044	0.043	0.049	0.056	0.043	0.045	0.050	0.063	0.046
$-0.5T$	0.1	0.039	0.051	0.065	0.033	0.041	0.048	0.061	0.036	0.046	0.053	0.061	0.036
	0.2	0.048	0.051	0.046	0.039	0.045	0.049	0.048	0.039	0.048	0.046	0.046	0.038
	0.3	0.049	0.051	0.043	0.041	0.046	0.047	0.042	0.043	0.045	0.050	0.040	0.041
	0.4	0.049	0.050	0.036	0.046	0.050	0.053	0.035	0.042	0.045	0.048	0.034	0.043
	0.5	0.051	0.053	0.035	0.051	0.046	0.047	0.035	0.042	0.045	0.046	0.035	0.043
	0.6	0.054	0.057	0.041	0.048	0.051	0.052	0.033	0.044	0.049	0.049	0.036	0.040
	0.7	0.054	0.056	0.040	0.045	0.044	0.049	0.034	0.041	0.049	0.052	0.038	0.044
	0.8	0.052	0.054	0.046	0.046	0.045	0.051	0.046	0.043	0.049	0.052	0.047	0.043
	0.9	0.045	0.050	0.071	0.041	0.039	0.050	0.068	0.034	0.043	0.050	0.071	0.039

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 IVX_{QR}^{MBB}. 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 $\boldsymbol{v}_t = -\gamma f_t + \boldsymbol{e}_t$, where the common factor f_t is generated as a sequence of independent standard normals, $\gamma = -0.95$ and $(a_t, \boldsymbol{e}_t')' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{I}_{K+1})$. The predictors are generated for $\boldsymbol{\Gamma} = \rho \mathbf{I}_K$ with $\rho = 1 + c/T$ for $c = \{0, -2.5, -10, -0.5T\}$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$ $K=2$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$ $K=3$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$ $K=4$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$ $K=5$	IVX_{QR}
0	0.1	0.038	0.038	0.084	0.038	0.038	0.094	0.046	0.048	0.136	0.048	0.050	0.165
	0.2	0.037	0.037	0.049	0.049	0.051	0.064	0.054	0.056	0.083	0.055	0.061	0.105
	0.3	0.041	0.043	0.055	0.052	0.049	0.055	0.060	0.063	0.076	0.067	0.069	0.088
	0.4	0.046	0.046	0.046	0.051	0.051	0.052	0.066	0.067	0.064	0.065	0.067	0.072
	0.5	0.041	0.041	0.045	0.049	0.050	0.050	0.061	0.061	0.068	0.074	0.074	0.079
	0.6	0.043	0.042	0.046	0.054	0.054	0.055	0.058	0.058	0.068	0.070	0.073	0.077
	0.7	0.047	0.044	0.046	0.052	0.050	0.057	0.059	0.059	0.076	0.063	0.066	0.082
	0.8	0.041	0.041	0.053	0.050	0.051	0.069	0.055	0.057	0.080	0.061	0.067	0.102
	0.9	0.037	0.039	0.074	0.044	0.045	0.113	0.043	0.051	0.132	0.044	0.053	0.165
-2.5	0.1	0.048	0.046	0.079	0.039	0.045	0.106	0.043	0.049	0.133	0.036	0.051	0.155
	0.2	0.048	0.048	0.057	0.051	0.054	0.065	0.048	0.053	0.075	0.046	0.055	0.094
	0.3	0.050	0.050	0.052	0.055	0.054	0.053	0.054	0.053	0.065	0.055	0.058	0.073
	0.4	0.050	0.050	0.043	0.050	0.051	0.049	0.056	0.056	0.061	0.055	0.056	0.067
	0.5	0.046	0.047	0.041	0.050	0.050	0.049	0.055	0.055	0.055	0.063	0.063	0.070
	0.6	0.052	0.049	0.043	0.054	0.054	0.045	0.052	0.051	0.058	0.064	0.064	0.068
	0.7	0.047	0.045	0.051	0.047	0.049	0.052	0.053	0.052	0.059	0.056	0.058	0.082
	0.8	0.050	0.049	0.054	0.045	0.046	0.064	0.053	0.052	0.076	0.049	0.051	0.084
-10	0.2	0.055	0.052	0.052	0.062	0.062	0.060	0.049	0.054	0.070	0.052	0.052	0.085
	0.3	0.058	0.061	0.043	0.057	0.061	0.050	0.053	0.055	0.057	0.055	0.056	0.068
	0.4	0.058	0.059	0.045	0.062	0.064	0.046	0.054	0.055	0.050	0.062	0.063	0.066
	0.5	0.054	0.053	0.036	0.061	0.061	0.040	0.059	0.059	0.047	0.057	0.057	0.056
	0.6	0.053	0.054	0.040	0.061	0.057	0.047	0.059	0.057	0.052	0.058	0.058	0.060
	0.7	0.057	0.053	0.045	0.052	0.051	0.049	0.053	0.054	0.063	0.058	0.057	0.073
	0.8	0.057	0.057	0.056	0.057	0.059	0.065	0.054	0.052	0.077	0.054	0.057	0.085
$-0.5T$	0.1	0.050	0.051	0.078	0.048	0.048	0.099	0.048	0.046	0.125	0.041	0.045	0.158
	0.2	0.049	0.049	0.047	0.052	0.050	0.060	0.049	0.051	0.063	0.041	0.049	0.084
	0.3	0.059	0.060	0.048	0.054	0.054	0.044	0.048	0.050	0.054	0.048	0.048	0.067
	0.4	0.047	0.049	0.038	0.055	0.053	0.044	0.051	0.052	0.052	0.046	0.045	0.054
	0.5	0.054	0.054	0.038	0.047	0.047	0.038	0.042	0.042	0.043	0.047	0.046	0.050
	0.6	0.053	0.054	0.041	0.049	0.050	0.050	0.048	0.048	0.043	0.046	0.048	0.052
	0.7	0.058	0.059	0.046	0.051	0.051	0.047	0.045	0.048	0.052	0.045	0.048	0.062
	0.8	0.055	0.056	0.046	0.052	0.049	0.062	0.052	0.047	0.066	0.047	0.049	0.076
	0.9	0.057	0.055	0.079	0.048	0.046	0.099	0.045	0.043	0.125	0.037	0.044	0.160



(a) $\tau = 0.2$, $c = 0$ (b) $\tau = 0.5$, $c = 0$ (c) $\tau = 0.8$, $c = 0$

(d) $\tau = 0.2$, $c = -2.5$ (e) $\tau = 0.5$, $c = -2.5$ (f) $\tau = 0.8$, $c = -2.5$

(g) $\tau = 0.2$, $c = -10$ (h) $\tau = 0.5$, $c = -10$ (i) $\tau = 0.8$, $c = -10$

(j) $\tau = 0.2$, $c = -20$ (k) $\tau = 0.5$, $c = -20$ (l) $\tau = 0.8$, $c = -20$

(m) $\tau = 0.2$, $c = -0.5T$ (n) $\tau = 0.5$, $c = -0.5T$ (o) $\tau = 0.8$, $c = -0.5T$

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

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 labeled "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).

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.

							τ				
	$\mathcal{T}_\tau$		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
dp_t	0.531	$\mathcal{T}_\tau$	0.154	0.498	0.367	0.368	0.007	0.068	0.049	0.370	0.034
		$\mathcal{T}_\tau^0$	0.171	0.554	0.395	0.383	0.007	0.068	0.049	0.371	0.032
		IVX_{QR}	0.481	0.174	0.000	0.002	0.230	**2.365****	0.171	0.219	0.000
		IVX_{QR}^{MBB}	-0.019	0.007	0.000	0.001	-0.006	-0.017	-0.004	-0.005	0.000
dy_t	0.708	$\mathcal{T}_\tau$	0.486	0.928	0.720	0.577	0.008	0.022	0.092	0.589	0.244
		$\mathcal{T}_\tau^0$	0.576	1.070	0.796	0.607	0.008	0.022	0.091	0.605	0.257
		IVX_{QR}	0.255	1.455	**4.077*****	1.469	0.629	0.094	0.187	0.419	0.025
		IVX_{QR}^{MBB}	0.012	0.017	0.023	0.012	0.007	-0.003	-0.004	-0.006	-0.002
e/p_t	5.828	$\mathcal{T}_\tau$	**2.343****	1.861*	**1.966****	1.595	0.604	0.000	0.947	1.653*	**2.254****
		$\mathcal{T}_\tau^0$	**3.306*****	**1.975****	**2.034****	1.637	0.606	0.000	1.225	**2.465****	**4.513*****
		IVX_{QR}	**13.098*****	**6.989*****	**5.097*****	**6.702*****	0.643	0.004	**2.854*****	**2.456****	**6.877*****
		IVX_{QR}^{MBB}	0.051	0.026	0.019	0.018	0.005	0.000	-0.012	-0.011	-0.021
de_t	9.536	$\mathcal{T}_\tau$	1.849*	0.161	0.192	0.106	0.234	0.063	0.884	0.910	**2.391****
		$\mathcal{T}_\tau^0$	**2.237****	0.140	0.176	0.104	0.234	0.070	1.193	1.437	**5.237*****
		IVX_{QR}	**9.454*****	1.783*	1.047	0.033	0.528	0.000	0.785	1.504	**4.290*****
		IVX_{QR}^{MBB}	-0.040	-0.012	-0.008	-0.001	-0.004	0.000	0.005	0.008	0.016
$rvol_t$	**28.482*****	$\mathcal{T}_\tau$	**4.672*****	0.649	0.280	0.078	1.734*	**5.623*****	**11.708*****	**18.089*****	**13.076*****
		$\mathcal{T}_\tau^0$	**2.660*****	0.564	0.251	0.076	1.740*	**6.065*****	**14.019*****	**23.242*****	**17.187*****
		IVX_{QR}	**11.228*****	0.783	1.015	0.011	0.182	**5.527*****	**11.354*****	**17.851*****	**35.494*****
		IVX_{QR}^{MBB}	-0.282	-0.055	-0.053	-0.005	0.017	0.092	**0.132*****	**0.170*****	**0.249*****
bm_t	1.664	$\mathcal{T}_\tau$	0.018	1.217	0.257	0.057	0.222	0.167	0.456	0.534	0.099
		$\mathcal{T}_\tau^0$	0.022	1.474	0.282	0.058	0.222	0.169	0.492	0.632	0.122
		IVX_{QR}	0.261	1.477	0.008	0.026	0.269	0.581	**2.118****	0.282	0.000
		IVX_{QR}^{MBB}	-0.042	0.050	-0.004	-0.005	-0.016	-0.023	-0.044	-0.018	0.001
$ntis_t$	4.934	$\mathcal{T}_\tau$	**3.285*****	0.041	0.034	0.041	0.149	0.001	0.172	0.024	0.001
		$\mathcal{T}_\tau^0$	**3.721*****	0.046	0.036	0.042	0.149	0.001	0.156	0.021	0.001
		IVX_{QR}	**5.041*****	0.526	0.010	0.677	0.038	1.620	0.384	0.328	0.591
		IVX_{QR}^{MBB}	0.645	0.136	0.017	0.115	0.026	0.161	0.083	0.081	0.142
tbl_t	0.743	$\mathcal{T}_\tau$	0.026	0.494	0.737	0.565	1.221	1.878*	0.244	0.580	0.514
		$\mathcal{T}_\tau^0$	0.023	0.451	0.698	0.554	1.223	1.951*	0.247	0.590	0.516
		IVX_{QR}	0.001	1.174	0.354	0.572	0.395	0.004	1.247	1.933*	0.344
		IVX_{QR}^{MBB}	-0.000	-0.002	-0.001	-0.001	-0.001	0.000	-0.002	-0.002	-0.001
lty_t	0.932	$\mathcal{T}_\tau$	0.031	0.448	0.968	0.699	1.544	**2.381****	0.270	0.612	0.535
		$\mathcal{T}_\tau^0$	0.027	0.412	0.927	0.690	1.554	**2.452****	0.267	0.587	0.493
		IVX_{QR}	0.001	1.932*	**4.534*****	1.604	0.872	0.610	1.644	0.836	0.035
		IVX_{QR}^{MBB}	-0.000	-0.002	-0.003	-0.002	-0.001	-0.001	-0.002	-0.001	0.000
ltr_t	4.369	$\mathcal{T}_\tau$	0.924	0.043	0.031	0.061	0.001	0.379	1.170	1.191	**3.468*****
		$\mathcal{T}_\tau^0$	1.585	0.050	0.033	0.061	0.001	0.373	1.206	1.381	**5.424*****
		IVX_{QR}	**2.295****	0.584	0.017	0.191	0.000	0.667	**3.332*****	**3.731*****	**2.313****
		IVX_{QR}^{MBB}	-0.003	-0.001	0.000	0.000	0.000	0.001	0.001	0.002*	**0.002****
tms_t	4.215	$\mathcal{T}_\tau$	0.164	0.031	1.677*	0.906	1.353	**2.174****	0.419	0.436	0.217
		$\mathcal{T}_\tau^0$	0.191	0.031	1.776*	0.922	1.358	**2.072****	0.397	0.381	0.209
		IVX_{QR}	**2.832*****	0.883	**3.382*****	0.404	0.365	0.547	0.218	0.110	0.084
		IVX_{QR}^{MBB}	-0.008	-0.003	-0.004	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001

Table 7 (Cont.): 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								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
dfy_t	**31.016*****	$\mathcal{T}_\tau$	**5.825*****	1.791*	0.775	0.824	0.144	0.002	1.364	**2.901*****	**3.937*****
		$\mathcal{T}_\tau^0$	**16.729*****	**2.719*****	0.880	0.859	0.144	0.002	1.505	**4.181*****	**6.989*****
		IVX_{QR}	**15.276*****	**12.136*****	**15.635*****	**4.055*****	0.866	0.301	**4.391*****	**10.704*****	**5.643*****
		IVX_{QR}^{MBB}	**-0.053****	-0.033*	-0.033	-0.013	-0.006	-0.003	0.013	0.020	**0.019****
dfr_t	6.099	$\mathcal{T}_\tau$	1.102	1.438	**2.453****	0.459	0.647	0.178	0.001	0.159	1.667*
		$\mathcal{T}_\tau^0$	**2.604*****	**2.052****	**2.816*****	0.484	0.648	0.176	0.001	0.209	**3.043*****
		IVX_{QR}	**3.073*****	**6.712*****	**3.564*****	0.896	0.343	0.245	0.047	0.211	**3.309*****
		IVX_{QR}^{MBB}	0.006	0.005	0.003	0.001	0.001	0.001	0.000	-0.001	-0.004
$infl_t$	1.332	$\mathcal{T}_\tau$	0.965	0.393	0.313	0.135	0.091	0.083	0.009	0.100	0.128
		$\mathcal{T}_\tau^0$	1.400	0.451	0.311	0.135	0.091	0.082	0.009	0.100	0.127
		IVX_{QR}	**4.255*****	0.837	0.707	0.103	0.010	0.000	0.000	0.166	0.090
		IVX_{QR}^{MBB}	0.028	0.009	-0.007	-0.002	-0.001	0.000	0.000	-0.003	-0.003
vix_t	**42.357*****	$\mathcal{T}_\tau$	**5.919*****	**2.201****	0.216	0.087	**2.397****	**6.193*****	**15.880*****	**28.333*****	**20.665*****
		$\mathcal{T}_\tau^0$	**6.992*****	**2.248****	0.203	0.083	**2.410****	**6.647*****	**18.902*****	**43.260*****	**39.311*****
		IVX_{QR}	**11.749*****	**3.407*****	0.044	0.707	1.579	**8.525*****	**20.524*****	**24.512*****	**60.127*****
		IVX_{QR}^{MBB}	-0.046	-0.019	-0.002	0.006	0.009	0.019	0.029	**0.031****	**0.046*****
$skew_t$	7.251	$\mathcal{T}_\tau$	**2.667*****	1.956*	**2.056****	**1.962****	1.326	**3.564*****	0.145	0.030	0.063
		$\mathcal{T}_\tau^0$	**3.310*****	**2.266****	**2.081****	1.940*	1.331	**3.779*****	0.156	0.034	0.067
		IVX_{QR}	0.599	**5.440*****	**5.700*****	1.168	**2.387****	0.990	0.025	0.055	0.000
		IVX_{QR}^{MBB}	0.066	0.106	0.099	0.038	0.050	0.032	0.006	0.010	0.001
$TailIndex_t$	12.991	$\mathcal{T}_\tau$	1.721*	1.116	0.653	0.133	1.191	1.867*	**3.180*****	**8.401*****	**10.040*****
		$\mathcal{T}_\tau^0$	1.549	1.112	0.651	0.136	1.197	1.808*	**2.961*****	**6.810*****	**7.645*****
		IVX_{QR}	**8.520*****	**4.578*****	**5.275*****	0.003	1.728*	**4.255*****	**10.121*****	**18.466*****	**41.178*****
		IVX_{QR}^{MBB}	-0.045	**-0.023****	-0.021	0.001	0.010	0.015	0.024	0.035	0.055*
$ACMY10_t$	0.930	$\mathcal{T}_\tau$	0.030	0.459	0.976	0.706	1.558	**2.385****	0.276	0.622	0.541
		$\mathcal{T}_\tau^0$	0.025	0.422	0.935	0.698	1.570	**2.457****	0.273	0.598	0.500
		IVX_{QR}	0.001	**1.981****	**4.609*****	1.489	0.989	0.532	1.285	0.702	0.035
		IVX_{QR}^{MBB}	-0.000	-0.002	-0.003	-0.002	-0.001	-0.001	-0.001	-0.001	0.000
$ACMTP10_t$	2.435	$\mathcal{T}_\tau$	0.704	1.284	**2.939*****	1.755*	**2.612*****	**3.594*****	0.502	0.768	0.456
		$\mathcal{T}_\tau^0$	0.583	1.105	**2.686*****	1.711*	**2.636*****	**3.733*****	0.498	0.764	0.478
		IVX_{QR}	**5.129*****	0.928	**5.795*****	1.127	0.686	1.742*	0.778	0.281	0.046
		IVX_{QR}^{MBB}	-0.014	-0.004*	-0.007*	-0.002	-0.002*	-0.003*	-0.002	-0.001	-0.001
vrp_t	4.021	$\mathcal{T}_\tau$	0.156	0.231	0.219	0.423	0.795	1.378	1.751*	1.632	0.396
		$\mathcal{T}_\tau^0$	0.305	0.271	0.208	0.388	0.796	1.654*	**2.753*****	**3.944*****	1.901*
		IVX_{QR}	0.457	**4.105*****	**7.882*****	**17.798*****	**25.351*****	**33.974*****	**42.053*****	**42.019*****	**10.547*****
		IVX_{QR}^{MBB}	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.000
iv_t	**48.963*****	$\mathcal{T}_\tau$	**6.919*****	**4.692*****	0.722	0.055	0.644	**3.909*****	**12.674*****	**22.910*****	**18.416*****
		$\mathcal{T}_\tau^0$	**16.380*****	**6.293*****	0.771	0.054	0.646	**4.027*****	**14.697*****	**36.040*****	**51.128*****
		IVX_{QR}	**43.378*****	**11.850*****	**10.057*****	0.102	1.581	**28.958*****	**53.129*****	**68.802*****	**92.882*****
		IVX_{QR}^{MBB}	**-0.001*****	**-0.000****	-0.000	0.000	0.000	0.000	**0.000****	**0.001*****	**0.001*****
rv_t	**76.700*****	$\mathcal{T}_\tau$	**3.362*****	**2.935*****	0.835	0.070	0.001	0.319	**1.995****	**3.910*****	**3.724*****
		$\mathcal{T}_\tau^0$	**11.437*****	**5.022*****	0.970	0.070	0.001	0.353	**2.854*****	**8.795*****	**17.249*****
		IVX_{QR}	**78.129*****	**20.046*****	**4.811*****	**6.820*****	0.016	**8.508*****	**9.923*****	**44.421*****	**83.715*****
		IVX_{QR}^{MBB}	**-0.001****	**-0.000****	-0.000	-0.000	0.000	0.000	0.000	**0.000****	**0.000*****

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 (15) applied over the nine quantiles ($\tau = 0.1, ..., 0.9$) 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 999 bootstrap replications. Associated MBB confidence intervals for β_τ are reported in Table C.1 in Part C of the supplementary appendix.

On-Line Supplementary Appendix

to

"Predictive Quantile Regressions with Persistent and Heteroskedastic Predictors: A Powerful 2SLS Testing Approach"

by

Matei Demetrescu, Paulo M. M. Rodrigues and Robert Taylor

Summary of Contents

Part A of this supplement briefly reviews the Wald-type QR predictability test of Lee (2016) and its MBB implementation developed in Fan and Lee (2019). Here we also correct implementation errors present in the MBB algorithm outlined in Fan and Lee (2019) and use Monte Carlo simulations to highlight the size distortions that can arise when these implementation errors are not corrected. **Part B** of this supplement contains a proof of Proposition 1. **Part C** contains data descriptions for the time series variables used in the empirical application in section 5, together with the MBB confidence intervals for β_τ referred to in the notes to Table 7 (Cont.). **Part D** contains the additional Monte Carlo simulations results referred to in section 4.

Part A - A Review of Wald Tests for Quantile Predictability

In this section we briefly review the Wald-type tests for quantile predictability proposed in Lee (2016), together with the MBB implementation of these tests developed in Fan and Lee (2019), designed to allow for conditional heteroskedasticity in the errors. In the context of the latter we correct two implementation errors present in the MBB algorithm outlined in Fan and Lee (2019, p.268).

In order to develop a test of the null hypothesis of no quantile specific predictability, $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$, against the two-sided alternative, $H_1 : \boldsymbol{\beta}_\tau \neq \mathbf{0}$ in the context of (1), Lee (2016) first *de-quantiles* the data, under the restriction of the null hypothesis, to obtain $\tilde{y}_{t,\tau} := y_t - \hat{Q}_{y_t}(\tau)$, where $\hat{Q}_{y_t}(\tau)$ denotes the sample (marginal) τ-quantile of y_t defined as

$$\hat{Q}_{y_t}(\tau) := \arg\min_q \sum_{t=1}^{T} \rho_\tau (y_t - q), \tag{A.1}$$

where ρ_τ is the usual quantile check function, $\rho_\tau(u) := u\psi_\tau(u)$, and where $\psi_\tau(u)$ is as defined in (2). His proposed Wald-type test, given in Proposition 3.2 of Lee (2016, p.110), and denoted IVX_{QR} in what follows, is then based on regressing $\tilde{y}_{t,\tau}$ on $\boldsymbol{x}_{t-1}$ using instrumental variable [IV] estimation. Specifically, in order to develop tests which are robust to both the persistence and degree of endogeneity of the predictors, Lee (2016) employs the IVX estimation method of Kostakis et al. (2015) using the vector of instrumental variables, $\boldsymbol{z}_{I,t}$ defined exactly as in (9).

In order to test $H_0 : \boldsymbol{\beta}_\tau = \mathbf{0}$ against the two-sided alternative, $H_1 : \boldsymbol{\beta}_\tau \neq \mathbf{0}$, in the context of the quantile predictive regression in (1), Lee (2016) and Fan and Lee (2019) base their tests around the IVX_{QR} estimator of $\boldsymbol{\beta}_\tau$ in an auxiliary QR of $\tilde{y}_{t,\tau}$ on $\boldsymbol{z}_{I,t-1}$ without intercept; that is,

$$\hat{\boldsymbol{\beta}}_\tau^{ivx} := \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{T} \rho_\tau \left(\tilde{y}_{t,\tau} - \boldsymbol{\beta}'\boldsymbol{z}_{I,t-1}\right), \tag{A.2}$$

where ρ_τ is as defined in the context of (A.1). Although the limiting distribution of $\hat{\boldsymbol{\beta}}_\tau^{ivx}$ is asymptotically (mixed) Gaussian, Fan and Lee (2019) show that the associated IVX_{QR} tests proposed in Lee (2016) have limiting null distributions that depend, in a complicated way, on nuisance parameters arising from the conditional variances which they argue renders the conventional solution of correcting the statistics using estimates of the nuisance parameters involved impractical. As a result, Fan and Lee (2019, p. 268) propose a MBB implementation of the IVX_{QR} tests, which they demonstrate to be first-order asymptotically valid under conditional heteroskedasticity.

As discussed in Remarks A.1 and A.2 below, it would appear that there are implementation errors in the MBB algorithm detailed in Fan and Lee (2019, p.268). We correct these in Algorithm A.1 below, which we state in full for completeness. The simulations results presented in section 4 for the MBB implementation of the IVX_{QR} tests are based on Algorithm A.1. In line with Fan and Lee (2019, p.268), the MBB algorithms we detail below are implemented for the case of a scalar predictor; cf. Footnote 6.

Algorithm A.1 (MBB IVX_{QR}) *:*

Step 1: *Given data (y_t, x_t) for $t = 1, ..., T$, select the IVX tuning parameters a and η according to the rule outlined in Lee (2016, pp.110-111) and construct the instrument $z_{I,t}$.*

Step 2: *Set the block length to $b = \lceil T^{1/4} \rceil$ where $\lceil a \rceil$ denotes the least integer that is greater or equal to a. Let $m = \lceil T/b \rceil$. Randomly sample m data blocks from $(\mathcal{B}(1), \ldots, \mathcal{B}(q))$, $q = T - b + 1$, where $\mathcal{B}(t) = (\boldsymbol{w}_t, \boldsymbol{w}_{t+1}, \ldots, \boldsymbol{w}_{t+b-1})$ with $\boldsymbol{w}_t = (y_t, z_{I,t-1})$. Denote the resulting resampled data as $(\boldsymbol{w}_1^*, \ldots, \boldsymbol{w}_\ell^*)$, where $\ell = mb$ and $\boldsymbol{w}_t^* = (y_t^*, z_{I,t-1}^*)$.*

Step 3: *Compute $\tilde{y}_{t,\tau}^* := y_t^* - \hat{Q}_{y_t^*}(\tau)$ where $\hat{Q}_{y_t^*}(\tau)$ is obtained analogously as in (A.1), and estimate*

$$\hat{\beta}_\tau^{ivx,*} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{\ell} \rho_\tau \left(\tilde{y}_{t,\tau}^* - \beta z_{I,t-1}^* \right). \tag{A.3}$$

Step 4: *Repeat steps 2 and 3 N times to obtain $\hat{\beta}_\tau^{ivx,*(1)}, \ldots, \hat{\beta}_\tau^{ivx,*(N)}$.*

Step 5: *The $(1-\alpha)100\%$ percentile interval is $\left[\hat{G}^{-1}(\alpha/2), \hat{G}^{-1}(1-\alpha/2)\right]$, where $\hat{G}$ is the empirical CDF of $\hat{\beta}_\tau^{ivx,*}$ from step 4.*

Remark A.1 *Unfortunately, the MBB IVX_{QR} tests, as prescribed in Fan and Lee (2019, p.268), are not invariant to the intercept term, α_τ, in the QR in (1). The reason for this is as follows. In the MBB implementation outlined in Fan and Lee (2019) (note that this coincides with the $IVX_{QR}^{MBB_1}$ procedure outlined in Algorithm A.2 below) the bootstrap data are obtained by resampling blockwise from the pairs $(y_t, \tilde{z}_t)$ (where $\tilde{z}_t$ is the IVX instrument in the notation of Fan and Lee, 2019, Section 4), and so the resulting bootstrap dependent variable, y_t^*, will depend on α_τ. However, the bootstrap statistics in Fan and Lee (2019, p.266) are computed without adjusting y_t^* for a non-zero quantile, even though the data used in their original statistic is adjusted; see their equations (11) and (12), where $y_{t\tau}$ in their equation (11) is seen from equation (3.1) of Lee (2016) to be the original y_t de-quantiled by subtracting $Q_{y_t}(\tau)$, the sample τ-quantile of y_t. The resulting bootstrap statistics will therefore not be size controlled, except in the case where $\alpha_\tau = 0$. Numerical illustrations of this lack of invariance to α_τ can be found in Table A.1 below. A pivotal MBB can easily be obtained, however, by de-quantiling y_t^* before computing the bootstrap statistics in equation (12) of Fan and Lee (2019), as is done in Step 3 of Algorithm A.1 above.* $\diamondsuit$

Remark A.2 *Additionally, it is important to observe that in Step 2 of Algorithm A.1 the data blocks are built from* $\boldsymbol{w}_t = (y_t, z_{I,t-1})$ *and not from* $\boldsymbol{w}_t = (y_t, z_{I,t})$ *as is suggested in Fan and Lee (2019, p.266). The latter will generate a mismatch between the observations of* y_t^* *and* $z_{I,t}^*$ *when the latter is lagged to compute the predictive regression. Given that the block length* b *is likely to be small in practical applications (e.g.* $b = 4$ *and* $b = 6$ *for* $T = 250$ *and* $T = 750$*, respectively), this mismatch has the potential to generate significant finite sample size distortions.* ♢

Remark A.3 *The MBB scheme described in Fan and Lee (2019, p.268) resamples the original data in such a way that the temporal ordering of the original data is not preserved in the bootstrap data. Consequently, this implementation would be expected to fail to control size in the presence of unconditionally heteroskedastic errors and/or time-varying endogeneity correlations (simulation results for the former, reported in Tables D.7–D.10 of the supplementary appendix, confirm this). In the conditional mean predictive regression setting, Demetrescu et al. (2022, 2023) adopt a wild bootstrap re-sampling device to address this issue, but it is unclear how a wild bootstrap could be implemented in the setup of Fan and Lee (2019).* ♢

We conclude this part of the appendix by evaluating the alternative MBB schemes discussed in Remarks A.1 and A.2, including the original MBB algorithm proposed by Fan and Lee (2019):

Algorithm A.2 (MBB IVX_{QR}) *:*

Step 1: Given data (y_t, x_t) *for* $t = 1, ..., T$ *choose* (a, η) *using the rule in Lee (2016) (in the notation of Lee (2016) the tuning parameters are* (C_z, δ)*) and construct the instrument* $z_{I,t}$*.*

Step 2: Set the block length to $b = \lceil T^{1/4} \rceil$ *where* $\lceil a \rceil$ *denotes the least integer that is greater or equal to* a*. Let* $m = \lceil T/b \rceil$*. Randomly sample* m *data blocks from* $(\mathcal{B}(1), \dots, \mathcal{B}(q))$*,* $q = T - b + 1$*, where* $\mathcal{B}(t) = (\boldsymbol{w}_t, \boldsymbol{w}_{t+1}, \dots, \boldsymbol{w}_{t+b-1})$*. We consider two possibilities for* $\boldsymbol{w}_t$*: (i)* $\boldsymbol{w}_t = (y_t, z_{I,t})$*; or (ii)* $\boldsymbol{w}_t = (y_t, z_{I,t-1})$*. Denote the resulting resampled data under case (i) by* $(\boldsymbol{w}_1^*, \dots, \boldsymbol{w}_\ell^*)$ *where* $\boldsymbol{w}_t^* = (y_t^*, z_{I,t}^*)$*, and under case (ii) by* $(\boldsymbol{w}_1^{**}, \dots, \boldsymbol{w}_\ell^{**})$ *where* $\boldsymbol{w}_t^{**} = (y_t^{**}, z_{I,t-1}^{**})$*, where, in each case,* $\ell = mb$*.*

Step 3: Either: based on $\boldsymbol{w}_t^*$*, estimate,*

$$\hat{\beta}_{k,\tau}^{ivx,*} = \arg\min_{\beta} \sum_{t=2}^{\ell} \rho_\tau \left(\tilde{y}_{t,\tau}^* - \beta z_{I,t-1}^* \right) \qquad k = 1, 2 \tag{A.4}$$

where, for the case of $k = 1$, $\tilde{y}^*_{t,\tau} = y^*_t$, *and for the case of* $k = 2$, $\tilde{y}^*_{t,\tau} := y^*_t - \hat{Q}_{y^*_t}(\tau)$, *with* $\hat{Q}_{y^*_t}(\tau)$ *obtained analogously as in (A.1);*

Or: based on $\boldsymbol{w}^{**}_t$, *estimate,*

$$\hat{\beta}^{ivx,*}_{k,\tau} = \arg\min_{\beta} \sum_{t=1}^{\ell} \rho_\tau \left(\tilde{y}^{**}_{t,\tau} - \beta z^{**}_{I,t-1} \right) \qquad k = 3, 4 \tag{A.5}$$

where, for the case of $k = 3$, $\tilde{y}^{**}_{t,\tau} = y^{**}_t$, *and for the case of* $k = 4$, $\tilde{y}^{**}_{t,\tau} := y^{**}_t - \hat{Q}_{y^{**}_t}(\tau)$, *with* $\hat{Q}_{y^{**}_t}(\tau)$ *obtained analogously as in (A.1).*

Observe that the main difference between $z^*_{I,t-1}$ *and* $z^{**}_{I,t-1}$, *is that the former corresponds to the lag of the bootstrapped* $z^*_{I,t}$ *from Step 2, whereas* $z^{**}_{I,t-1}$ *is directly bootstrapped from* $z_{I,t-1}$ *in Step 2. For this reason, the summation index in* (A.4) *needs to start at* $t = 2$, *rather than* $t = 1$ *as in* (A.5).

Step 4: Repeat steps 2 and 3 N *times to obtain* $\hat{\beta}^{ivx,*(1)}_{k,\tau}, \ldots, \hat{\beta}^{ivx,*(N)}_{k,\tau}$.

Step 5: The respective $(1-\alpha)100\%$ *percentile intervals are computed as* $\left[\hat{G}^{-1}(\alpha/2), \hat{G}^{-1}(1-\alpha/2)\right]$, *where* $\hat{G}$ *is the empirical CDF of* $\hat{\beta}^{ivx,*}_\tau$ *from step 4.*

For convenience in what follows, we will refer to the statistics resulting from the four MBB schemes described in Algorithm A.2 above as $IVX^{MBB_k}_{QR}, k = 1, ..., 4$, respectively. Notice that $IVX^{MBB_1}_{QR}$ corresponds to the MBB tests computed as described in the algorithm of Fan and Lee (2019, p.268), while $IVX^{MBB_4}_{QR}$ corresponds to the MBB test outlined in Algorithm A.1. To evaluate the invariance to nonzero intercepts at the different quantiles of the four MBB algorithms, we consider the DGP in (16)-(17) with homoskedastic and serially uncorrelated innovations, but we now set $\alpha_\tau = 1$ in (16). The results are presented in Table A.1 below. Comparing the results for $IVX^{MBB_1}_{QR}$ in Table A.1 with the corresponding results in Table 1 and Table D.1, where $\alpha_\tau = 0$, the lack of invariance of $IVX^{MBB_1}_{QR}$ to α_τ is very clearly seen (the same sets of random numbers were used across the two sets of simulation experiments, such that the only difference between the two sets of simulated data is whether $\alpha_\tau = 0$, as with the results in Table 1 and Table D.1, or $\alpha_\tau = 1$, as with the results in Table A.1).

Table A.1: Empirical null rejection frequencies at 5% significance level of MBB QR based predictability tests $IVX_{QR}^{MBB_k}$, $k = 1, ..., 4$, for sample sizes $T = 250$ and $T = 750$. DGP: $y_t = \alpha_\tau + \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, with $\alpha_\tau = 1$, $\beta_\tau = 0$, $\rho = 1 + c/T$ for $c = \{0, -2.5, -10, -0.5T\}$, and $(u_t, v_t)' \sim i.i.d.\ N(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad -0.95; \quad -0.95 \quad 1]$.

c	τ	$IVX_{QR}^{MBB_1}$	$IVX_{QR}^{MBB_2}$	$IVX_{QR}^{MBB_3}$	$IVX_{QR}^{MBB_4}$	$IVX_{QR}^{MBB_1}$	$IVX_{QR}^{MBB_2}$	$IVX_{QR}^{MBB_3}$	$IVX_{QR}^{MBB_4}$
		$T = 250$				$T = 750$			
0	0.1	0.665	0.007	0.665	0.055	0.759	0.027	0.759	0.066
	0.2	0.604	0.013	0.609	0.069	0.688	0.031	0.688	0.071
	0.3	0.597	0.021	0.602	0.093	0.675	0.037	0.677	0.080
	0.4	0.523	0.018	0.527	0.075	0.639	0.031	0.641	0.070
	0.5	0.515	0.018	0.523	0.073	0.629	0.032	0.630	0.071
	0.6	0.517	0.019	0.525	0.078	0.634	0.033	0.635	0.071
	0.7	0.586	0.020	0.587	0.092	0.665	0.038	0.666	0.081
	0.8	0.591	0.014	0.599	0.070	0.674	0.028	0.677	0.073
	0.9	0.627	0.009	0.638	0.054	0.743	0.028	0.742	0.066
-2.5	0.1	0.356	0.005	0.366	0.045	0.518	0.020	0.518	0.055
	0.2	0.281	0.012	0.292	0.053	0.396	0.022	0.401	0.055
	0.3	0.284	0.017	0.292	0.071	0.380	0.028	0.387	0.062
	0.4	0.174	0.014	0.183	0.062	0.315	0.028	0.320	0.060
	0.5	0.164	0.013	0.180	0.066	0.311	0.027	0.318	0.062
	0.6	0.169	0.014	0.186	0.067	0.313	0.024	0.315	0.060
	0.7	0.272	0.015	0.295	0.070	0.365	0.027	0.372	0.063
	0.8	0.268	0.010	0.283	0.057	0.382	0.021	0.389	0.057
	0.9	0.314	0.006	0.340	0.049	0.484	0.019	0.490	0.054
-10	0.1	0.100	0.006	0.113	0.042	0.213	0.017	0.220	0.045
	0.2	0.073	0.009	0.086	0.048	0.110	0.016	0.116	0.049
	0.3	0.095	0.010	0.107	0.058	0.112	0.021	0.117	0.057
	0.4	0.019	0.011	0.028	0.055	0.056	0.024	0.061	0.055
	0.5	0.014	0.012	0.026	0.056	0.057	0.021	0.067	0.058
	0.6	0.017	0.010	0.033	0.054	0.056	0.022	0.062	0.056
	0.7	0.087	0.011	0.100	0.058	0.110	0.023	0.114	0.054
	0.8	0.073	0.008	0.092	0.046	0.106	0.019	0.116	0.053
	0.9	0.082	0.006	0.108	0.038	0.194	0.018	0.207	0.046
-0.5T	0.1	0.003	0.006	0.006	0.037	0.003	0.016	0.008	0.047
	0.2	0.004	0.007	0.008	0.043	0.002	0.015	0.004	0.041
	0.3	0.004	0.009	0.008	0.041	0.001	0.016	0.003	0.042
	0.4	0.001	0.007	0.004	0.045	0.000	0.018	0.001	0.045
	0.5	0.001	0.007	0.006	0.041	0.000	0.015	0.003	0.042
	0.6	0.001	0.007	0.005	0.044	0.000	0.016	0.002	0.040
	0.7	0.006	0.007	0.016	0.047	0.002	0.015	0.007	0.042
	0.8	0.004	0.008	0.015	0.044	0.003	0.016	0.008	0.048
	0.9	0.005	0.005	0.024	0.037	0.004	0.015	0.016	0.044

Part B - Proof of Proposition 1

In this section we provide a proof of Proposition 1. We first state and prove a preparatory lemma which will be needed for the purposes of proving Proposition 1.

Lemma A.1 *Let the conditions of Proposition 1 hold. Then, for all $t = 1, \ldots, T$, it holds that*

$$\tilde{s}_{t,\tau} = s_{t,\tau} - \bar{s}_\tau + R_{t,T}$$

in which $\bar{s}_\tau := \frac{1}{T}\sum_{t=1}^{T} s_{t,\tau}$, and $R_{t,T} = O_p(T^{-1/2})$, uniformly in t.

Proof of Lemma A.1

Observe first that

$$\tilde{s}_{t,\tau} - s_{t,\tau} = \begin{cases} 1 & u_{t,\tau} \leq 0 \text{ and } \hat{Q}_{u_{t,\tau}}(\tau) < u_{t,\tau} \\ -1 & u_{t,\tau} > 0 \text{ and } \hat{Q}_{u_{t,\tau}}(\tau) > u_{t,\tau} \\ 0 & \text{otherwise,} \end{cases}$$

that is,

$$\begin{aligned} \tilde{s}_{t,\tau} - s_{t,\tau} &= \mathbb{I}(u_{t,\tau} \leq 0)\, \mathbb{I}\left(\hat{Q}_{u_{t,\tau}}(\tau) < u_{t,\tau}\right) - \mathbb{I}(u_{t,\tau} > 0)\, \mathbb{I}\left(\hat{Q}_{u_{t,\tau}}(\tau) > u_{t,\tau}\right) \\ &=: \phi_{t,T}. \end{aligned}$$

Noting that $Q_{u_{t,\tau}}(\tau) = 0$, and using the fact that $\hat{Q}_{u_{t,\tau}}(\tau)$ is $\sqrt{T}$ consistent for $Q_{u_{t,\tau}}(\tau)$ (which can be established by a straightforward adaptation of the proof of Lemma 4 of De Jong et al., 2007 covering the case of the median), it therefore holds that $\mathrm{P}(\phi_{t,T} \neq 0) = O\left(T^{-1/2}\right)$, uniformly in t. Moreover, because $|\phi_{t,T}|$ can only take the values 0 or 1, $\mathrm{E}(\phi_{t,T})$ exists and is itself of order $T^{-1/2}$.

Next write

$$\tilde{s}_{t,\tau} = s_{t,\tau} + \mathrm{E}(\phi_{t,T}) + (\phi_{t,T} - \mathrm{E}(\phi_{t,T})), \tag{B.1}$$

where we note that $\mathrm{E}(\phi_{t,T})$ is time-invariant under Assumption 1, which it will be recalled imposes the condition that $u_{t,\tau}$ is strictly stationary, implying that

$$\sum_{t=1}^{T} \tilde{s}_{t,\tau} - T\,\mathrm{E}(\phi_{t,T}) = \sum_{t=1}^{T} s_{t,\tau} + \sum_{t=1}^{T} (\phi_{t,T} - \mathrm{E}(\phi_{t,T}))$$

and

$$T\,\mathrm{E}(\phi_{t,T}) = \sum_{t=1}^{T} \tilde{s}_{t,\tau} - \sum_{t=1}^{T} s_{t,\tau} - \sum_{t=1}^{T} (\phi_{t,T} - \mathrm{E}(\phi_{t,T})).$$

It is not difficult to show that $\sum_{t=1}^{T} \tilde{s}_{t,\tau} = O_p(1)$ since $\hat{Q}_{u_{t,\tau}}(\tau)$ is the sample quantile of $u_{t,\tau}$, and we therefore obtain that

$$\mathrm{E}\left(\phi_{t,T}\right) = -\bar{s}_\tau - \frac{1}{T}\sum_{t=1}^{T}\left(\phi_{t,T} - \mathrm{E}\left(\phi_{t,T}\right)\right) + \bar{\tilde{s}}_\tau$$

where $\bar{\tilde{s}}_\tau := 1/T\sum_{t=1}^{T} \tilde{s}_{t,\tau}$ is of order $O(T^{-1})$.

Plugging this into (B.1) yields,

$$\begin{aligned}\tilde{s}_{t,\tau} \quad &= \quad s_{t,\tau} - \bar{s}_\tau + \phi_{t,T} - \frac{1}{T}\sum_{t=1}^{T}\left(\phi_{t,T} - \mathrm{E}\left(\phi_{t,T}\right)\right) + \bar{\tilde{s}}_\tau \\ &=: \quad s_{t,\tau} - \bar{s}_\tau + R_{t,T},\end{aligned}$$

where, observing that $\mathrm{P}\left(\phi_{t,T} \neq 0\right) = O\left(T^{-1/2}\right)$, uniformly in t, and that $|\phi_{t,T}| \leq 1$, $R_{t,T}$ is easily shown to be uniformly bounded and $O_p(T^{-1/2})$, uniformly in t, as required.

Proof of Proposition 1

The proof builds on the structure and arguments used in the proof of Proposition 1 of Demetrescu et al. (2022). In order to examine the asymptotic behaviour of our LM-type statistic,

$$\mathcal{T}_\tau = \mathbf{C}_T'\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_T\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{C}_T,$$

where $\mathbf{A}_T$, $\mathbf{B}_T$, $\mathbf{C}_T$ and $\mathbf{D}_T$ are defined in the main text, denote by $\mathcal{T}_\tau^\dagger$ the analogous statistic obtained by replacing $\tilde{s}_{t,\tau}$ by $s_{t,\tau}$; viz,

$$\mathcal{T}_\tau^\dagger := \mathbf{C}_T^{\dagger\prime}\mathbf{B}_T^{-1}\mathbf{A}_T\left(\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{D}_T^\dagger\mathbf{B}_T^{-1}\mathbf{A}_T\right)^{-1}\mathbf{A}_T'\mathbf{B}_T^{-1}\mathbf{C}_T^\dagger$$

with $\mathbf{C}_T^\dagger := \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}s_{t,\tau}$ and $\mathbf{D}_T^\dagger := \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'s_{t,\tau}^2$, analogously to $\mathbf{C}_T = \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}\tilde{s}_{t,\tau}$ and $\mathbf{D}_T = \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau}^2$ in the definition of $\mathcal{T}_\tau$.

Observing that $s_{t,\tau}$, $\tilde{\boldsymbol{z}}_{t-1}$ and $\boldsymbol{x}_{t-1}$ are easily verified to obey multivariate versions of the assumptions of Lemma 1 of Demetrescu et al. (2022) (cf. section S.2.2 of their on-line supplementary appendix), their Corollary 2 then applies and the limiting null chi-square distribution follows for $\mathcal{T}_\tau^\dagger$; see Remark 17 in Demetrescu et al. (2022). To complete the proof it therefore suffices to show that $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^\dagger$ are asymptotically equivalent under the null hypothesis.

In doing so, we can reduce the proof of Proposition 1 in Demetrescu et al. (2022) to the full-sample case (see their Lemma S.6), and let the scaling matrix $\mathbf{W}_T$ be defined as

1. $\mathbf{W}_T := \begin{pmatrix} \frac{1}{T^{1/2+\eta/2}}\mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \frac{1}{T^{1/2}}\mathbf{I}_K \end{pmatrix}$ under strong persistence and

2. $\mathbf{W}_T := \begin{pmatrix} \frac{1}{T^{1/2}}\mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \frac{1}{T^{1/2}}\mathbf{I}_K \end{pmatrix}$ under weak persistence.

To establish asymptotic equivalence of $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^\dagger$ we need only compare the behaviour of the suitably normalised $\mathbf{C}_T^\dagger$ and $\mathbf{C}_T$, and of the suitably normalised $\mathbf{D}_T^\dagger$ and $\mathbf{D}_T$.

Examine first $\mathbf{W}_T\boldsymbol{C}_T := \mathbf{W}_T \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}\tilde{s}_{t,\tau}$, where we distinguish between the type-I and type-II instruments under the two different types of persistence in the following.

Consider the strongly persistent case. Here we have that

$$\mathbf{W}_T \sum_{t=1}^T \tilde{\boldsymbol{z}}_{t-1}\tilde{s}_{t,\tau} = \begin{pmatrix} \frac{1}{T^{1/2+\eta/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}\tilde{s}_{t,\tau} \\ \frac{1}{T^{1/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{II,t-1}\tilde{s}_{t,\tau} \end{pmatrix}. \tag{B.2}$$

Examine the first element of (B.2). Recall from Demetrescu et al. (2022) that it suffices to show that the first element of this vector is bounded in probability, because, under strong persistence, it is multiplied by a zero-limit weight appearing in the probability limit of the suitably normalized matrix $\boldsymbol{A}_T$ (which thus plays the role of a selector). This can indeed be seen to be the case, on noting, appealing to Lemma A.1, that

$$\frac{1}{T^{1/2+\eta/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}\tilde{s}_{t,\tau} = \frac{1}{T^{1/2+\eta/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}s_{t,\tau} - \frac{\bar{s}_\tau}{T^{1/2+\eta/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1} + \frac{1}{T^{1/2+\eta/2}} \sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}R_{t,T}$$

where the three terms on the right hand side are all bounded in probability or vanishing. The boundedness of the first term follows from Demetrescu et al. (2022), while for the second it follows because

$$\sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1} = O_p\left(T^{1/2+\eta}\right) \quad \text{and} \quad \bar{s}_\tau = O_p\left(T^{-1/2}\right),$$

and for the third term it follows because

$$\begin{aligned} \left|\sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}R_{t,T}\right| &\leq \sup_t |R_{t,T}| \left|\sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}\right| \leq \sup_t |R_{t,T}| \sqrt{T\sum_{t=1}^T \tilde{\boldsymbol{z}}_{I,t-1}^2} \\ &= O_p\left(T^{-1/2}\sqrt{T}T^{1/2+\eta/2}\right); \end{aligned}$$

see Lemma A.1. Turning to a consideration of the second element of the vector in (B.2), define

$\tilde{S}_{t,\tau} := \sum_{j=1}^{t} \tilde{s}_{j,\tau}$, and use the partial summation formula to conclude that

$$\begin{aligned}\sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{s}_{t,\tau} &= \tilde{\boldsymbol{z}}_{II,T-1}\tilde{S}_{T,\tau} - \sum_{t=2}^{T} \tilde{S}_{t,\tau}\Delta\tilde{\boldsymbol{z}}_{II,t} \\ &= \tilde{\boldsymbol{z}}_{II,T-1}S_{T,\tau} - \sum_{t=2}^{T} S_{t,\tau}\Delta\tilde{\boldsymbol{z}}_{II,t} \\ &\quad +\tilde{\boldsymbol{z}}_{II,T-1}\left(\tilde{S}_{T,\tau} - S_{T,\tau}\right) - \sum_{t=2}^{T}\left(\tilde{S}_{T,\tau} - S_{T,\tau}\right)\Delta\tilde{\boldsymbol{z}}_{II,t}\end{aligned}$$

where $S_{t,\tau} := \sum_{j=1}^{t} (s_{j,\tau} - \bar{s}_{\tau})$. De Jong et al. (2007, Lemma 3) implies that $\tilde{S}_{T,\tau} - S_{T,\tau} = o_p\left(T^{1/2}\right)$. Therefore, since $\Delta\tilde{\boldsymbol{z}}_{II,t} = O\left(T^{-1}\right)$ due to the Lipschitz-type property of the type-II instruments, and since $\tilde{\boldsymbol{z}}_{II,t}$ is demeaned such that $\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{II,t-1} = 0$, we have that

$$\begin{aligned}\sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{s}_{t,\tau} &= \tilde{\boldsymbol{z}}_{II,T-1}S_{T,\tau} - \sum_{t=2}^{T} S_{t,\tau}\Delta\tilde{\boldsymbol{z}}_{II,t} + o_p\left(T^{1/2}\right) = \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{II,t-1}\left(s_{t,\tau} - \bar{s}\right) + o_p\left(T^{1/2}\right) \\ &= \sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{II,t-1}s_{t,\tau} + o_p\left(T^{1/2}\right),\end{aligned}$$

as required. Noting that type-II instruments do not depend on the predictors, this result can be seen to hold irrespective of the persistence of $\boldsymbol{x}_t$.

Consider next the weakly persistent case. Here we have that

$$\mathbf{W}_T\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}\tilde{s}_{t,\tau} = \begin{pmatrix} \frac{1}{T^{1/2}}\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{I,t-1}\tilde{s}_{t,\tau} \\ \frac{1}{T^{1/2}}\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{II,t-1}\tilde{s}_{t,\tau}\end{pmatrix}$$

where the result that

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{I,t-1}\tilde{s}_{t,\tau} = \frac{1}{T^{1/2}}\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{I,t-1}s_{t,\tau} + o_p(1)$$

can be established using the arguments given in Section 1.4 of the online supplement to Lee (2016). The result for $\frac{1}{T^{1/2}}\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{II,t-1}\tilde{s}_{t,\tau}$ has already been derived above, where it was noted that this result holds regardless of whether $\boldsymbol{x}_t$ is strongly or weakly persistent.

To complete the proof, it follows from Lemma A.1 that replacing $\tilde{s}_{t,\tau}$ by $s_{t,\tau}$ in

$$\mathbf{W}_T\mathbf{D}_T\mathbf{W}_T = \mathbf{W}_T\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau}^2\mathbf{W}_T$$

to obtain $\mathbf{W}_T\mathbf{D}_T^{\dagger}\mathbf{W}_T$, makes no difference asymptotically, regardless of whether $\boldsymbol{x}_t$ is strongly or weakly persistent. Concretely, we have that

$$\mathbf{W}_T \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau}^2 \mathbf{W}_T = \begin{pmatrix} \frac{1}{T^{1+\eta}} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{I,t-1}\tilde{\boldsymbol{z}}_{I,t-1}'\tilde{s}_{t,\tau}^2 & \frac{1}{T^{1+\eta/2}} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{I,t-1}\tilde{\boldsymbol{z}}_{II,t-1}'\tilde{s}_{t,\tau}^2 \\ \frac{1}{T^{1+\eta/2}} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{\boldsymbol{z}}_{I,t-1}'\tilde{s}_{t,\tau}^2 & \frac{1}{T} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{\boldsymbol{z}}_{II,t-1}'\tilde{s}_{t,\tau}^2 \end{pmatrix}$$

under strong persistence, and

$$\mathbf{W}_T \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{t-1}\tilde{\boldsymbol{z}}_{t-1}'\tilde{s}_{t,\tau}^2 \mathbf{W}_T = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{I,t-1}\tilde{\boldsymbol{z}}_{I,t-1}'\tilde{s}_{t,\tau}^2 & \frac{1}{T} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{I,t-1}\tilde{\boldsymbol{z}}_{II,t-1}'\tilde{s}_{t,\tau}^2 \\ \frac{1}{T} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{\boldsymbol{z}}_{I,t-1}'\tilde{s}_{t,\tau}^2 & \frac{1}{T} \sum_{t=1}^{T} \tilde{\boldsymbol{z}}_{II,t-1}\tilde{\boldsymbol{z}}_{II,t-1}'\tilde{s}_{t,\tau}^2 \end{pmatrix}$$

under weak persistence. Appealing to the Cauchy-Schwarz inequality, it then only suffices to establish that the diagonal elements of $\mathbf{W}_T\mathbf{D}_T\mathbf{W}_T - \mathbf{W}_T\mathbf{D}_T^{\dagger}\mathbf{W}_T$ vanish in the limit, which is because $\bar{s}_\tau = O_p\left(T^{-1/2}\right)$ and $R_{t,T} = O_p\left(T^{-1/2}\right)$, where the latter holds uniformly in t regardless of the actual regressor persistence (cf. Lemma A.1), implying that $\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2$ must vanish uniformly in t. We then have for the diagonal entry corresponding to kth type-II instrument that

$$\begin{aligned} \left|\frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,II,t-1}^2\tilde{s}_{t,\tau}^2 - \frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,II,t-1}^2 s_{t,\tau}^2\right| &\leq \sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,II,t-1}^2 \\ &= O_p\left(\sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\right) = o_p(1), \end{aligned}$$

and, for the diagonal entry corresponding to kth type-II instrument, that under strong persistence,

$$\begin{aligned} \left|\frac{1}{T^{1+\eta}}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2\tilde{s}_{t,\tau}^2 - \frac{1}{T^{1+\eta}}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2 s_{t,\tau}^2\right| &\leq \sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\frac{1}{T^{1+\eta}}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2 \\ &= O_p\left(\sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\right) = o_p(1) \end{aligned}$$

while, under weak persistence,

$$\begin{aligned} \left|\frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2\tilde{s}_{t,\tau}^2 - \frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2 s_{t,\tau}^2\right| &\leq \sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\frac{1}{T}\sum_{t=1}^{T}\tilde{z}_{k,I,t-1}^2 \\ &= O_p\left(\sup_t\left|\tilde{s}_{t,\tau}^2 - s_{t,\tau}^2\right|\right) = o_p(1) \end{aligned}$$

where the magnitude orders of the sums of squared instruments may be obtained from Lemmas S.2 and S.3 of Demetrescu et al. (2022).

Collecting these results together we obtain that $\mathcal{T}_\tau$ and $\mathcal{T}_\tau^{\dagger}$ are asymptotically equivalent to one another under the null hypothesis, with this result holding regardless of whether $\boldsymbol{x}_t$ is strongly or weakly persistent. The stated result then follows immediately from the asymptotic equivalence of

$\mathcal{T}_\tau$ and $\mathcal{T}_\tau^\dagger$ under the null hypothesis, demonstrated above.

Part C - Detailed Data Descriptions and IVX_{QR}^{MBB} CIs

- **Goyal and Welch's data set**

  This dataset consists of observations on the equity premium for the S&P Composite index, calculated using CRSP's month-end values, together with 14 putative predictors, taken from the updated monthly dataset used in Welch and Goyal (2008) available on Amit Goyal's website (https://sites.google.com/view/agoyal145). The equity premium is the variable to be predicted and is defined as the log return on the value-weighted CRSP stock market index minus the log return on the risk-free Treasury bill: $ep_t := \log(1+R_{m,t}) - \log(1+R_{f,t})$, where $R_{m,t}$ is the CRSP return and $R_{f,t}$ is the Treasury bill return (Goyal and Welch, 2003). The variables are in log form (as in Goyal and Welch, 2003) and each of the predictors is lagged one period. Specifically, the predictors used are the log dividend price ratio (dp_t), the log dividend yield (dy_t), the log earnings price ratio (e/p_t), log dividend payout ratio (de_t), the equity risk premium volatility, $rvol_t$, the book to market ratio (bm_t), the net equity expansion ($ntis_t$), treasury bill rate (tbl_t), the long-term government bond yield (lty_t), the long-term government bond rate of return (ltr_t), the term spread (tms_t), the default yield spread (dfy_t), the default return spread (dfr_t), and inflation ($infl_t$).

- **A market tail-risk measure**

  Following Kelly and Jiang (2014), we will consider a market tail risk measure. Computing a market tail risk index is a challenging task, (see e.g. Andersen et al., 2015a,b, and Christoffersen et al., 2020).[8] Kelly and Jiang (2014) propose a panel estimation approach that captures common variation in the tail risk of individual firms. Their argument for the use of this approach is that if firm-level tail distributions possess common dynamics, then cross-sections of crash events can be used to estimate the common component of their stocks' tail risk at each point in time. The advantage of their approach is that it can be straightforwardly applied to low frequency data and is not subject to the data limitation of single time series where a long span of data is necessary. To briefly describe their method, suppose that the time t left-tail distribution of stock $i's$ returns is defined as the set of returns that fall below some extreme negative threshold w_t,, and it is assumed that the lower tail of stock $i's$

[8]According to Kelly and Jiang (2014) there are two current approaches to measuring tail risk dynamics for stock returns: one is based on option price data, and the other on high frequency returns data. Regarding the first see, *inter alia*, Bakshi et al. (2003), who study risk-neutral skewness and kurtosis, Bollerslev et al. (2009), who examine how the variance risk premium relates to the equity premium, and Backus et al. (2011) and Gao et al. (2019), who infer disaster risk premia from options (see also additional work on the tail index computed from option prices by Torben G. Andersen and Viktor Todorov at https://tailindex.com/volatilityindex.html). Regarding the second approach see, for example, Bollerslev and Todorov (2011), Almeida et al. (2017) and Weller (2018).

returns behaves as,

$$P(r_{i,t+1} < r | r_{i,t+1} < w_t \text{ and } \mathcal{F}_t) = \left(\frac{r}{w_t}\right)^{-a_i/\lambda_t}, \tag{C.1}$$

where $r < w_t < 0$ and $\mathcal{F}_t$ is the information set available up to time t. Since $r < w_t < 0$, $r/w_t > 1$ and $a_i/\lambda_t > 0$, this ensures that $0 \leq \left(\frac{r}{w_t}\right)^{-a_i/\lambda_t} \leq 1$. Kelly and Jiang (2014) refer to λ_t as the "tail risk" at time t. Notice that high values of λ_t correspond to heavy tailed distributions, implying high probabilities of extreme returns.

The identifying assumption for the estimation of $1/\lambda_t$ is that the tail risk of individual assets includes a common component which, according to Kelly and Jiang (2014), can be computed at each point in time from a cross-section of extreme returns for individual assets. λ_t is estimated based on the tail index estimator of Hill (1975) applied to the time t cross-section of firms' returns. For estimation of the tail index λ_t in our application, we use the conditional maximum likelihood estimator proposed by Hill, i.e.,

$$\lambda_t^{Hill} = \frac{1}{K_t}\sum_{k=1}^{K_t} ln\frac{R_{kt}}{w_t}, \tag{C.2}$$

where R_{kt} is the kth daily return that falls below w_t, the fifth percentile of the cross-section, during month t and K_t is the total number of returns that exceed w_t within month t (for further details on the approach see Kelly, 2014, Kelly and Jiang, 2014 and Nicolau et al., 2023). Notably, the tail exponent λ_t is highly persistent - Kelly and Jiang report a monthly AR(1) coefficient for the resulting tail index series of 0.927. In what follows, we refer to the tail index estimated in this way as $TailIndex_t$.

- **VIX and SKEW Indexes**

  We also consider the log of the CBOE Volatility Index (vix) and the log of the CBOE SKEW index ($skew$). The vix is a real-time market index that represents the market's expectations for volatility over the coming 30 days which is used as a way to gauge market sentiment; that is, it is an "investor fear gauge"; cf. Bollerslev et al. (2015, p.113). The $skew$ measures the potential risk in financial markets and, like the vix, is also used as a proxy for investor sentiment and volatility. It is calculated using S&P 500 options that measure tail risk (returns two or more standard deviations from the mean) in S&P 500 returns over the next 30 days; for details see https://cdn.cboe.com/resources/indices/documents/SKEWwhitepaperjan2011.pdf.

- **Treasury Term Premia and Yields**

  We also consider the 10 year treasury term premia and yields ($ACMTP10_t$ and $ACMTY10_t$), obtained from the Federal Reserve Bank of New York (https://www.newyorkfed.org/research). The term premium corresponds to the compensation that investors require for bearing the risk that interest rates may change over the life of the bond. Since the term premium is not directly observable, it must be estimated, most often from financial and macroeconomic variables. The data we use are based on the approach of Adrian et al. (2013), who developed a statistical model to describe the joint evolution of the treasury yields and term premia across time and maturities. We use end-of-month observations.

- **Variance Risk Premium**

  Finally, we also consider the variance risk premium (VRP_t), which is a short-term predictor extracted from the option market (Bollerslev et al., 2009). According to Zhou (2018) the variance risk premium, provides significant predictability for equity returns, bond returns, forward premiums, and credit spreads. The documented return predictability peaks around one month, and then dies out as the forecasting horizon increases. In fact, the conventional variance risk premium, defined as the difference between the squared Volatility Index (iv_t) and the realised return variance (rv_t), is a quasi-variance risk premium because it includes the pure variance risk premium and higher moment premium components.

Table C.1: IVX_{QR}^{MBB} confidence intervals

	τ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
dp_t	[-0.0182, 0.0061]	[-0.0052, 0.0128]	[-0.0049, 0.0104]	[-0.0034, 0.0067]	[-0.0033, 0.0078]	[-0.0041, 0.0082]	[-0.0003, 0.0010]	[-0.0013, 0.0059]	[-0.0020, 0.0076]
dy_t	[-0.0111, 0.0050]	[-0.0020, 0.0053]	[-0.0019, 0.0055]	[-0.0017, 0.0049]	[-0.0020, 0.0045]	[-0.0010, 0.0043]	[-0.0005, 0.0009]	[-0.0005, 0.0005]	[-0.0003, 0.0007]
e/p_t	[-0.0096, 0.0091]	[-0.0010, 0.0083]	[-0.0016, 0.0069]	[-0.0015, 0.0067]	[-0.0014, 0.0054]	[-0.0017, 0.0049]	[-0.0041, 0.0041]	[-0.0074, 0.0019]	[-0.0012, 0.0003]
de_t	[-0.0481, 0.0137]	[-0.0375, 0.0105]	[-0.0264, 0.0077]	[-0.0185, 0.0108]	[-0.0138, 0.0129]	[-0.0101, 0.0128]	[-0.0128, 0.0144]	[-0.0031, 0.0190]	[-0.0000, 0.0162]
$rvol_t$	[-0.3093, 0.0393]	[-0.1178, 0.0134]	[-0.1188, 0.0506]	[-0.0485, 0.0196]	[-0.0121, 0.0389]	[-0.0008, 0.0606]	**[0.0020, 0.0797]*****	**[0.0076, 0.1084]*****	**[0.0046, 0.1192]*****
bm_t	[-0.0510, 0.0807]	[-0.0044, 0.0203]	[-0.0399, 0.0536]	[-0.0374, 0.0286]	[-0.0461, 0.0187]	[-0.0566, 0.0052]	[-0.0651, 0.0147]	[-0.0525, 0.0071]	[-0.0580, 0.0540]
$ntis_t$	[-0.1260, 0.8995]	[-0.2416, 0.7430]	[-0.3104, 0.4457]	[-0.0973, 0.2184]	[-0.0725, 0.2959]	[-0.0430, 0.2605]	[-0.1147, 0.3557]	[-0.1802, 0.4764]	[-0.4197, 0.6673]
tbl_t	[-0.0027, 0.0035]	[-0.0030, 0.0006]	[-0.0031, 0.0011]	[-0.0021, 0.0009]	[-0.0019, 0.0007]	[-0.0026, 0.0014]	[-0.0028, 0.0017]	[-0.0040, 0.0002]	[-0.0052, 0.0004]
lty_t	[-0.0016, 0.0019]	[-0.0014, 0.0001]	[-0.0021, 0.0001]	[-0.0017, 0.0002]	[-0.0011, 0.0001]	[-0.0013, 0.0000]	[-0.0012, 0.0005]	[-0.0010, 0.0002]	[-0.0017, 0.0004]
ltr_t	[-0.0046, 0.0021]	[-0.0030, 0.0022]	[-0.0024, 0.0015]	[-0.0014, 0.0016]	[-0.0006, 0.0009]	[-0.0003, 0.0019]	[-0.0003, 0.0025]	[0.0000, 0.0029]*	**[0.0003, 0.0045]****
tms_t	[-0.0118, 0.0048]	[-0.0092, 0.0008]	[-0.0038, 0.0004]	[-0.0038, 0.0006]	[-0.0014, 0.0001]	[-0.0021, 0.0000]	[-0.0028, 0.0003]	[-0.0018, 0.0011]	[-0.0035, 0.0011]
dfy_t	**[-0.0588, -0.0022]****	[-0.0357, -0.0000]*	[-0.0336, 0.0045]	[-0.0283, 0.0026]	[-0.0221, 0.0063]	[-0.0148, 0.0082]	[-0.0078, 0.0146]	[-0.0000, 0.0082]	[0.0000, 0.0123]**
dfr_t	[-0.0044, 0.0127]	[-0.0006, 0.0082]	[0.0000, 0.0082]	[-0.0006, 0.0046]	[-0.0012, 0.0036]	[-0.0020, 0.0027]	[-0.0037, 0.0031]	[-0.0045, 0.0026]	[-0.0064, 0.0003]
$infl_t$	[-0.0079, 0.0406]	[-0.0080, 0.0187]	[-0.0139, 0.0070]	[-0.0168, 0.0058]	[-0.0149, 0.0045]	[-0.0156, 0.0065]	[-0.0165, 0.0078]	[-0.0142, 0.0100]	[-0.0257, 0.0113]
vix_t	[-0.0173, 0.0058]	[-0.0108, 0.0013]	[-0.0068, 0.0055]	[-0.0046, 0.0040]	[-0.0030, 0.0035]	[-0.0005, 0.0008]	[0.0000, 0.0021]	**[0.0000, 0.0042]****	**[0.0000, 0.0049]*****
$skew_t$	[-0.0009, 0.0017]	[-0.0008, 0.0001]	[-0.0010, 0.0002]	[-0.0007, 0.0002]	[-0.0005, 0.0000]	[-0.0005, 0.0001]	[-0.0034, 0.0017]	[-0.0027, 0.0007]	[-0.0054,0.0015]
$TailIndex_t$	[-0.0088, 0.0120]	[-0.0132, 0.0003]	**[-0.0035, -0.0000]****	[-0.0069, 0.0038]	[-0.0030, 0.0046]	[-0.0005, 0.0007]	[-0.0022, 0.0071]	[-0.0002, 0.0147]	[0.0002, 0.0213]*
$ACMY10_t$	[-0.0018, 0.0020]	[-0.0018, 0.0000]	[-0.0022, 0.0000]	[-0.0018, 0.0003]	[-0.0014, 0.0001]	[-0.0014, 0.0001]	[-0.0013, 0.0005]	[-0.0011, 0.0002]	[-0.0018, 0.0003]
$ACMTP10_t$	[-0.0197, 0.0035]	[-0.0105, -0.0002]*	[-0.0091, -0.0006]*	[-0.0065, 0.0001]	[-0.0034, -0.0000]*	[-0.0032, -0.0001]*	[-0.0039, 0.0010]	[-0.0027, 0.0013]	[-0.0045, 0.0021]
vrp_t	[-0.0004, 0.0002]	[-0.0003, 0.0003]	[-0.0003, 0.0006]	[-0.0003, 0.0005]	[-0.0003, 0.0006]	[-0.0002, 0.0007]	[-0.0002, 0.0008]	[-0.0002, 0.0006]	[-0.0002,0.0006]
iv_t	**[-0.0012, -0.0001]*****	**[-0.0007, -0.0001]****	[-0.0005, 0.0002]	[-0.0003, 0.0003]	[-0.0000, 0.0003]	[-0.0000, 0.0004]	**[0.0001, 0.0005]****	**[0.0002, 0.0006]*****	**[0.0003, 0.0008]*****
rv_t	**[-0.0012, -0.0001]****	**[-0.0006, -0.0000]****	[-0.0005, 0.0002]	[-0.0002, 0.0002]	[-0.0002, 0.0002]	[-0.0002, 0.0002]	[-0.0001, 0.0005]	**[0.0000, 0.0006]****	**[0.0001, 0.0009]*****

Note: Bold entries denote confidence intervals (CIs) at the 95% confidence level (or stricter) that do not include the null hypothesis $H_0 : \beta_\tau = 0$, while entries with *, ** and *** indicate CIs that do not include the null hypothesis $H_0 : \beta_\tau = 0$ at the 90%, 95% and 99% confidence levels, respectively. All other entries correspond to cases where the 90% CI includes $H_0 : \beta_\tau = 0$.

Part D - Monte Carlo Simulations: Additional Results

- Results for DGP1 when $T = 750$ are detailed in Table D.1.
- Results for DGP2a when $T = 750$ are reported in Table D.2.
- Results for DGP2b and DGP2c when $T = 750$ are reported in Table D.3.
- Results for DGP3 pertaining to a downward change in variance ($\mathsf{b} = 1/4$) when $T = 750$ are given in Table D.4, while Tables D.5 and D.6 report results for an upward change ($\mathsf{b} = 4$) for samples sizes $T = 250$ and $T = 750$, respectively.
- Results for DGP4 pertaining to a downward change ($\mathsf{b} = 1/4$) in variance are given in Tables D.7 and D.8, for $T = 250$ and $T = 750$, respectively. Corresponding results for an upward change ($\mathsf{b} = 4$) are reported in Tables D.9 and D.10, for $T = 250$ and $T = 750$, respectively.
- Results for DGP5 when $T = 750$ are reported in Table D.11.
- Local power results for $T = 750$ are graphed in Figure D.1.

Table D.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}, and IVX_{QR}^{MBB}. Sample size $T = 750$. 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}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
			$(\pi,\theta)=(0,0)$				$(\pi,\theta)=(0.5,0)$				$(\pi,\theta)=(-0.5,0)$				$(\pi,\theta)=(0,0.5)$				$(\pi,\theta)=(0,-0.5)$		
0	0.1	0.048	0.050	0.078	0.066	0.047	0.049	0.082	0.072	0.048	0.049	0.079	0.070	0.048	0.047	0.074	0.066	0.049	0.046	0.077	0.069
	0.2	0.055	0.056	0.066	0.071	0.056	0.057	0.086	0.081	0.054	0.056	0.079	0.083	0.049	0.049	0.069	0.074	0.047	0.047	0.070	0.079
	0.3	0.060	0.063	0.069	0.080	0.060	0.062	0.063	0.071	0.061	0.062	0.066	0.072	0.052	0.053	0.060	0.067	0.051	0.051	0.060	0.070
	0.4	0.058	0.058	0.060	0.070	0.059	0.059	0.065	0.077	0.058	0.059	0.064	0.075	0.050	0.053	0.069	0.081	0.053	0.055	0.066	0.076
	0.5	0.063	0.065	0.061	0.071	0.062	0.064	0.066	0.075	0.065	0.066	0.066	0.079	0.058	0.059	0.076	0.089	0.059	0.060	0.079	0.093
	0.6	0.066	0.068	0.063	0.071	0.065	0.066	0.066	0.075	0.065	0.066	0.065	0.074	0.053	0.055	0.067	0.081	0.055	0.056	0.071	0.084
	0.7	0.066	0.068	0.072	0.081	0.065	0.067	0.072	0.076	0.064	0.066	0.070	0.077	0.061	0.062	0.066	0.070	0.057	0.059	0.062	0.072
	0.8	0.060	0.061	0.069	0.073	0.059	0.059	0.083	0.087	0.060	0.060	0.087	0.083	0.054	0.057	0.079	0.075	0.056	0.057	0.078	0.076
	0.9	0.049	0.051	0.082	0.066	0.050	0.051	0.083	0.068	0.050	0.050	0.087	0.072	0.050	0.049	0.081	0.060	0.047	0.049	0.082	0.061
-2.5	0.1	0.042	0.044	0.068	0.055	0.041	0.045	0.068	0.055	0.043	0.044	0.069	0.053	0.040	0.039	0.059	0.052	0.041	0.041	0.058	0.052
	0.2	0.048	0.049	0.056	0.055	0.049	0.049	0.061	0.060	0.048	0.049	0.056	0.062	0.043	0.043	0.054	0.062	0.042	0.042	0.054	0.059
	0.3	0.050	0.049	0.051	0.062	0.050	0.049	0.049	0.059	0.048	0.048	0.049	0.061	0.045	0.044	0.053	0.057	0.043	0.044	0.049	0.059
	0.4	0.049	0.049	0.049	0.060	0.049	0.050	0.052	0.062	0.046	0.047	0.050	0.061	0.046	0.046	0.054	0.061	0.044	0.045	0.053	0.064
	0.5	0.051	0.051	0.051	0.062	0.050	0.051	0.051	0.060	0.051	0.052	0.051	0.062	0.047	0.048	0.057	0.069	0.046	0.046	0.059	0.067
	0.6	0.050	0.051	0.050	0.060	0.049	0.050	0.050	0.062	0.048	0.050	0.050	0.064	0.046	0.046	0.050	0.060	0.044	0.046	0.054	0.066
	0.7	0.052	0.054	0.057	0.063	0.052	0.052	0.058	0.058	0.054	0.054	0.056	0.060	0.047	0.050	0.051	0.056	0.052	0.053	0.052	0.058
	0.8	0.050	0.051	0.056	0.057	0.053	0.054	0.063	0.065	0.051	0.052	0.060	0.066	0.047	0.049	0.058	0.062	0.046	0.049	0.055	0.059
	0.9	0.044	0.044	0.067	0.054	0.044	0.043	0.070	0.054	0.045	0.044	0.072	0.055	0.041	0.041	0.064	0.053	0.042	0.043	0.063	0.048
-10	0.1	0.046	0.047	0.056	0.045	0.048	0.050	0.060	0.048	0.046	0.048	0.063	0.044	0.048	0.049	0.061	0.048	0.046	0.048	0.058	0.046
	0.2	0.048	0.050	0.045	0.049	0.047	0.050	0.051	0.050	0.049	0.052	0.053	0.050	0.047	0.048	0.051	0.053	0.047	0.051	0.049	0.052
	0.3	0.051	0.052	0.042	0.057	0.049	0.051	0.041	0.052	0.052	0.054	0.043	0.053	0.047	0.048	0.046	0.051	0.048	0.050	0.046	0.051
	0.4	0.050	0.052	0.046	0.055	0.051	0.052	0.047	0.055	0.049	0.049	0.046	0.055	0.049	0.050	0.051	0.055	0.048	0.049	0.048	0.054
	0.5	0.053	0.054	0.047	0.058	0.054	0.055	0.048	0.056	0.051	0.052	0.045	0.058	0.051	0.051	0.047	0.056	0.049	0.050	0.047	0.058
	0.6	0.049	0.051	0.046	0.056	0.049	0.050	0.042	0.055	0.050	0.050	0.045	0.057	0.048	0.047	0.045	0.053	0.049	0.050	0.044	0.057
	0.7	0.054	0.055	0.048	0.054	0.053	0.053	0.048	0.052	0.052	0.053	0.044	0.053	0.050	0.051	0.043	0.048	0.051	0.052	0.042	0.047
	0.8	0.051	0.051	0.052	0.053	0.052	0.050	0.054	0.057	0.052	0.052	0.056	0.057	0.051	0.053	0.048	0.050	0.051	0.051	0.052	0.051
	0.9	0.049	0.050	0.059	0.046	0.048	0.047	0.055	0.050	0.049	0.050	0.060	0.046	0.044	0.045	0.057	0.044	0.045	0.047	0.057	0.046
-0.5T	0.1	0.051	0.050	0.055	0.047	0.053	0.055	0.054	0.048	0.048	0.048	0.057	0.040	0.050	0.052	0.059	0.046	0.048	0.046	0.052	0.040
	0.2	0.044	0.043	0.047	0.041	0.046	0.048	0.047	0.045	0.043	0.043	0.045	0.039	0.052	0.054	0.049	0.046	0.052	0.053	0.045	0.043
	0.3	0.046	0.045	0.040	0.042	0.049	0.049	0.046	0.048	0.048	0.048	0.038	0.041	0.051	0.051	0.041	0.049	0.054	0.055	0.043	0.043
	0.4	0.045	0.045	0.039	0.045	0.053	0.052	0.043	0.048	0.048	0.050	0.038	0.044	0.054	0.054	0.043	0.052	0.048	0.050	0.038	0.046
	0.5	0.045	0.045	0.040	0.042	0.051	0.052	0.041	0.047	0.045	0.045	0.041	0.041	0.051	0.052	0.043	0.049	0.045	0.045	0.039	0.040
	0.6	0.047	0.048	0.037	0.040	0.048	0.048	0.042	0.046	0.043	0.044	0.035	0.039	0.047	0.047	0.037	0.042	0.044	0.045	0.039	0.041
	0.7	0.045	0.047	0.042	0.042	0.049	0.050	0.044	0.044	0.047	0.047	0.045	0.040	0.046	0.046	0.039	0.042	0.044	0.047	0.045	0.039
	0.8	0.052	0.053	0.044	0.048	0.049	0.053	0.048	0.046	0.051	0.052	0.045	0.047	0.047	0.047	0.045	0.044	0.049	0.049	0.048	0.043
	0.9	0.051	0.050	0.064	0.044	0.055	0.055	0.063	0.045	0.045	0.047	0.055	0.041	0.054	0.056	0.060	0.043	0.050	0.053	0.059	0.042

Table D.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}, and IVX_{QR}^{MBB}. Sample size $T = 750$. 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 $(u_t, v_t)' = [\sigma_{1t} \quad 0; 0 \quad \sigma_{2t}]\,\boldsymbol{\eta}_t$; $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})' \sim i.i.d.\ N(\mathbf{0}, \mathbf{\Omega})$ with $\mathbf{\Omega} = [1 \quad -0.95; \quad -0.95 \quad 1]$ 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}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\theta_1 = 0.1$, $\theta_2 = 0.5$				$\theta_1 = 0.1$, $\theta_2 = 0.8$				$\theta_1 = 0.05$, $\theta_2 = 0.9$			
0	0.1	0.051	0.051	0.079	0.063	0.052	0.051	0.078	0.063	0.052	0.049	0.080	0.064
	0.2	0.056	0.056	0.062	0.068	0.056	0.056	0.063	0.069	0.054	0.054	0.064	0.068
	0.3	0.063	0.063	0.074	0.088	0.063	0.066	0.082	0.097	0.063	0.065	0.077	0.090
	0.4	0.064	0.065	0.057	0.072	0.062	0.064	0.061	0.071	0.064	0.064	0.059	0.071
	0.5	0.064	0.065	0.054	0.063	0.062	0.064	0.058	0.064	0.063	0.064	0.056	0.065
	0.6	0.062	0.063	0.060	0.074	0.065	0.067	0.061	0.073	0.063	0.064	0.060	0.074
	0.7	0.062	0.064	0.071	0.089	0.063	0.066	0.081	0.097	0.062	0.064	0.073	0.090
	0.8	0.056	0.057	0.061	0.063	0.056	0.057	0.064	0.067	0.057	0.058	0.063	0.064
	0.9	0.051	0.051	0.078	0.064	0.050	0.050	0.079	0.064	0.050	0.049	0.080	0.063
-2.5	0.1	0.043	0.045	0.069	0.051	0.046	0.044	0.068	0.055	0.044	0.043	0.069	0.054
	0.2	0.048	0.048	0.052	0.053	0.047	0.047	0.057	0.056	0.045	0.047	0.056	0.055
	0.3	0.049	0.049	0.054	0.067	0.048	0.049	0.064	0.070	0.050	0.051	0.058	0.069
	0.4	0.048	0.050	0.046	0.056	0.049	0.050	0.049	0.058	0.048	0.049	0.047	0.059
	0.5	0.049	0.049	0.042	0.051	0.050	0.050	0.050	0.053	0.051	0.052	0.048	0.054
	0.6	0.048	0.049	0.049	0.062	0.049	0.050	0.051	0.060	0.050	0.051	0.048	0.063
	0.7	0.050	0.051	0.054	0.065	0.048	0.050	0.066	0.069	0.049	0.051	0.060	0.065
	0.8	0.048	0.046	0.053	0.059	0.046	0.046	0.054	0.057	0.044	0.048	0.053	0.057
	0.9	0.042	0.044	0.070	0.054	0.042	0.042	0.070	0.054	0.043	0.044	0.071	0.056
-10	0.1	0.050	0.050	0.059	0.051	0.051	0.051	0.059	0.051	0.051	0.052	0.059	0.047
	0.2	0.048	0.050	0.050	0.049	0.050	0.051	0.053	0.049	0.050	0.049	0.052	0.049
	0.3	0.055	0.057	0.041	0.055	0.055	0.058	0.050	0.059	0.055	0.056	0.044	0.055
	0.4	0.050	0.052	0.039	0.050	0.051	0.052	0.045	0.051	0.049	0.049	0.040	0.051
	0.5	0.047	0.047	0.043	0.048	0.047	0.048	0.046	0.051	0.046	0.047	0.044	0.053
	0.6	0.056	0.057	0.047	0.057	0.055	0.056	0.046	0.058	0.054	0.056	0.049	0.059
	0.7	0.052	0.053	0.048	0.056	0.052	0.053	0.050	0.063	0.048	0.050	0.049	0.057
	0.8	0.052	0.053	0.047	0.053	0.051	0.050	0.053	0.052	0.052	0.051	0.052	0.048
	0.9	0.044	0.046	0.060	0.044	0.048	0.048	0.060	0.045	0.045	0.047	0.059	0.046
$-0.5T$	0.1	0.043	0.045	0.051	0.037	0.043	0.045	0.057	0.037	0.044	0.045	0.056	0.037
	0.2	0.046	0.045	0.046	0.042	0.046	0.048	0.050	0.040	0.047	0.049	0.046	0.039
	0.3	0.052	0.052	0.043	0.047	0.052	0.053	0.049	0.049	0.052	0.053	0.045	0.047
	0.4	0.054	0.055	0.041	0.050	0.051	0.051	0.048	0.050	0.054	0.053	0.044	0.048
	0.5	0.053	0.054	0.044	0.050	0.051	0.052	0.049	0.049	0.052	0.053	0.047	0.048
	0.6	0.051	0.051	0.046	0.047	0.052	0.052	0.052	0.050	0.052	0.052	0.048	0.049
	0.7	0.052	0.053	0.044	0.050	0.050	0.050	0.050	0.048	0.051	0.051	0.045	0.045
	0.8	0.050	0.052	0.048	0.041	0.051	0.052	0.054	0.042	0.050	0.050	0.049	0.043
	0.9	0.050	0.053	0.059	0.045	0.051	0.053	0.065	0.046	0.050	0.050	0.061	0.043

Table D.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}, and IVX_{QR}^{MBB}. Sample size $T = 750$. 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 $(u_t, v_t)' = [\sigma_{1t} \quad 0; 0 \quad \sigma_{2t}]\,\boldsymbol{\eta}_t$; $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{\Omega})$ with $\boldsymbol{\Omega} = [1 \quad -0.95; -0.95 \quad 1]$ 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}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		ARCH(1)				SV			
0	0.1	0.046	0.046	0.082	0.058	0.045	0.045	0.084	0.055
	0.2	0.048	0.047	0.074	0.071	0.054	0.054	0.069	0.070
	0.3	0.052	0.051	0.075	0.078	0.061	0.062	0.062	0.074
	0.4	0.046	0.046	0.068	0.071	0.061	0.063	0.071	0.086
	0.5	0.049	0.051	0.064	0.074	0.056	0.057	0.075	0.090
	0.6	0.047	0.049	0.067	0.074	0.055	0.057	0.063	0.079
	0.7	0.051	0.052	0.071	0.074	0.057	0.058	0.061	0.065
	0.8	0.046	0.046	0.072	0.070	0.056	0.055	0.068	0.071
	0.9	0.041	0.041	0.082	0.056	0.048	0.050	0.089	0.064
-2.5	0.1	0.042	0.043	0.075	0.052	0.034	0.035	0.071	0.050
	0.2	0.041	0.042	0.059	0.058	0.043	0.047	0.059	0.059
	0.3	0.047	0.049	0.064	0.061	0.048	0.047	0.050	0.056
	0.4	0.044	0.045	0.061	0.061	0.052	0.052	0.052	0.066
	0.5	0.043	0.044	0.057	0.058	0.046	0.047	0.055	0.068
	0.6	0.047	0.047	0.055	0.063	0.043	0.044	0.047	0.060
	0.7	0.043	0.045	0.058	0.065	0.044	0.043	0.044	0.055
	0.8	0.040	0.040	0.063	0.056	0.045	0.047	0.059	0.061
	0.9	0.042	0.042	0.074	0.051	0.042	0.043	0.075	0.051
-10	0.1	0.048	0.049	0.072	0.048	0.042	0.044	0.063	0.044
	0.2	0.047	0.046	0.063	0.048	0.046	0.047	0.052	0.054
	0.3	0.051	0.054	0.059	0.055	0.052	0.053	0.047	0.055
	0.4	0.050	0.050	0.052	0.057	0.052	0.053	0.042	0.057
	0.5	0.050	0.050	0.052	0.052	0.051	0.052	0.043	0.055
	0.6	0.053	0.053	0.054	0.056	0.047	0.049	0.039	0.049
	0.7	0.052	0.053	0.056	0.057	0.047	0.048	0.042	0.050
	0.8	0.047	0.050	0.059	0.052	0.050	0.050	0.050	0.051
	0.9	0.051	0.052	0.073	0.045	0.048	0.048	0.065	0.044
$-0.5T$	0.1	0.043	0.048	0.215	0.045	0.050	0.054	0.065	0.047
	0.2	0.042	0.045	0.189	0.046	0.052	0.052	0.050	0.048
	0.3	0.047	0.049	0.200	0.052	0.053	0.054	0.046	0.047
	0.4	0.050	0.051	0.190	0.054	0.054	0.054	0.038	0.049
	0.5	0.055	0.055	0.195	0.054	0.050	0.050	0.039	0.048
	0.6	0.056	0.055	0.200	0.058	0.053	0.054	0.039	0.049
	0.7	0.050	0.051	0.196	0.051	0.050	0.052	0.040	0.051
	0.8	0.050	0.051	0.203	0.050	0.052	0.056	0.044	0.047
	0.9	0.043	0.051	0.213	0.045	0.046	0.047	0.064	0.040

Table D.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 $T = 750$. 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 $\mathbf{\Sigma}_t = \begin{bmatrix}\sigma_{ut}^2 & -0.95\sigma_{ut}\sigma_{vt}; & -0.95\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2\end{bmatrix}$ and $\sigma_{ut}^2 = 1$, $\sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor\lambda T\rfloor) + 1/4\mathbb{I}(t > \lfloor\lambda T\rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.047	0.048	0.088	0.060	0.057	0.057	0.083	0.057	0.049	0.050	0.084	0.061
	0.2	0.052	0.053	0.071	0.064	0.059	0.060	0.070	0.069	0.053	0.054	0.068	0.065
	0.3	0.057	0.060	0.088	0.091	0.060	0.060	0.088	0.097	0.055	0.059	0.098	0.107
	0.4	0.062	0.063	0.108	0.115	0.066	0.068	0.085	0.099	0.057	0.058	0.070	0.078
	0.5	0.065	0.067	0.088	0.089	0.065	0.067	0.071	0.085	0.061	0.062	0.062	0.073
	0.6	0.063	0.064	0.105	0.113	0.062	0.062	0.091	0.100	0.059	0.060	0.076	0.083
	0.7	0.065	0.065	0.093	0.099	0.059	0.059	0.099	0.101	0.061	0.060	0.102	0.108
	0.8	0.055	0.057	0.077	0.075	0.055	0.057	0.074	0.066	0.055	0.057	0.071	0.068
	0.9	0.048	0.051	0.093	0.066	0.050	0.050	0.087	0.066	0.049	0.048	0.078	0.062
-2.5	0.1	0.041	0.039	0.066	0.049	0.038	0.038	0.058	0.045	0.037	0.039	0.067	0.051
	0.2	0.042	0.043	0.050	0.049	0.043	0.044	0.052	0.057	0.039	0.038	0.050	0.055
	0.3	0.043	0.043	0.057	0.062	0.042	0.043	0.059	0.066	0.038	0.040	0.064	0.071
	0.4	0.042	0.043	0.072	0.075	0.040	0.041	0.055	0.069	0.039	0.038	0.047	0.064
	0.5	0.044	0.045	0.059	0.067	0.043	0.044	0.053	0.065	0.041	0.041	0.047	0.057
	0.6	0.045	0.047	0.063	0.079	0.042	0.042	0.060	0.075	0.044	0.043	0.051	0.060
	0.7	0.043	0.044	0.060	0.071	0.043	0.043	0.064	0.069	0.045	0.046	0.068	0.076
	0.8	0.044	0.044	0.055	0.060	0.040	0.042	0.060	0.057	0.043	0.044	0.058	0.053
	0.9	0.044	0.042	0.073	0.055	0.043	0.041	0.070	0.052	0.040	0.042	0.065	0.050
-10	0.1	0.045	0.047	0.059	0.045	0.044	0.048	0.055	0.043	0.049	0.052	0.058	0.047
	0.2	0.047	0.049	0.047	0.049	0.049	0.051	0.047	0.052	0.049	0.049	0.048	0.053
	0.3	0.049	0.052	0.046	0.054	0.049	0.051	0.047	0.055	0.049	0.050	0.051	0.063
	0.4	0.052	0.052	0.049	0.060	0.050	0.050	0.040	0.054	0.050	0.050	0.041	0.055
	0.5	0.047	0.047	0.050	0.063	0.051	0.052	0.041	0.053	0.051	0.051	0.041	0.050
	0.6	0.051	0.051	0.047	0.062	0.053	0.052	0.046	0.057	0.047	0.047	0.041	0.052
	0.7	0.052	0.051	0.050	0.057	0.050	0.051	0.050	0.061	0.056	0.054	0.049	0.057
	0.8	0.051	0.052	0.055	0.053	0.052	0.051	0.050	0.048	0.050	0.052	0.052	0.048
	0.9	0.055	0.059	0.063	0.053	0.052	0.056	0.058	0.049	0.050	0.053	0.062	0.045
$-0.5T$	0.1	0.050	0.053	0.058	0.043	0.049	0.051	0.054	0.044	0.048	0.051	0.060	0.040
	0.2	0.053	0.053	0.053	0.048	0.048	0.049	0.047	0.044	0.054	0.054	0.046	0.049
	0.3	0.046	0.046	0.044	0.045	0.052	0.052	0.043	0.045	0.054	0.056	0.045	0.047
	0.4	0.047	0.048	0.041	0.045	0.049	0.049	0.037	0.043	0.051	0.051	0.042	0.049
	0.5	0.048	0.049	0.042	0.046	0.048	0.048	0.040	0.044	0.047	0.048	0.040	0.044
	0.6	0.050	0.051	0.043	0.046	0.054	0.053	0.045	0.051	0.051	0.051	0.041	0.047
	0.7	0.050	0.050	0.045	0.044	0.050	0.050	0.047	0.050	0.052	0.052	0.046	0.046
	0.8	0.046	0.047	0.045	0.040	0.052	0.055	0.050	0.046	0.049	0.052	0.049	0.043
	0.9	0.047	0.049	0.055	0.041	0.045	0.049	0.055	0.039	0.045	0.047	0.054	0.036

Table D.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 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 $\mathbf{\Sigma}_t = \left[\sigma_{ut}^2 \quad -0.95\sigma_{ut}\sigma_{vt}; \quad -0.95\sigma_{ut}\sigma_{vt} \quad \sigma_{vt}^2\right]$ and $\sigma_{ut}^2 = 1$, $\sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + 4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.052	0.053	0.078	0.056	0.051	0.049	0.080	0.051	0.053	0.050	0.085	0.062
	0.2	0.053	0.056	0.055	0.062	0.051	0.052	0.058	0.064	0.052	0.057	0.064	0.068
	0.3	0.053	0.057	0.061	0.076	0.048	0.050	0.065	0.077	0.052	0.053	0.065	0.073
	0.4	0.056	0.059	0.056	0.084	0.054	0.059	0.061	0.093	0.056	0.058	0.067	0.095
	0.5	0.055	0.058	0.049	0.079	0.059	0.062	0.057	0.087	0.057	0.058	0.062	0.092
	0.6	0.062	0.066	0.057	0.086	0.055	0.059	0.066	0.088	0.051	0.055	0.067	0.088
	0.7	0.059	0.063	0.070	0.089	0.057	0.060	0.062	0.082	0.057	0.061	0.057	0.079
	0.8	0.056	0.062	0.059	0.066	0.056	0.060	0.061	0.067	0.056	0.057	0.065	0.062
	0.9	0.044	0.051	0.076	0.056	0.051	0.050	0.083	0.056	0.049	0.050	0.083	0.055
-2.5	0.1	0.051	0.051	0.071	0.046	0.047	0.045	0.070	0.044	0.048	0.048	0.074	0.049
	0.2	0.046	0.051	0.053	0.054	0.045	0.046	0.049	0.052	0.047	0.050	0.055	0.056
	0.3	0.043	0.049	0.051	0.063	0.044	0.046	0.050	0.064	0.041	0.044	0.050	0.059
	0.4	0.045	0.047	0.049	0.069	0.045	0.047	0.051	0.068	0.046	0.049	0.053	0.073
	0.5	0.049	0.052	0.043	0.068	0.048	0.051	0.047	0.075	0.045	0.048	0.047	0.070
	0.6	0.048	0.052	0.045	0.069	0.044	0.048	0.051	0.069	0.044	0.047	0.048	0.064
	0.7	0.051	0.053	0.051	0.071	0.046	0.050	0.047	0.068	0.043	0.046	0.046	0.063
	0.8	0.043	0.048	0.046	0.057	0.050	0.053	0.051	0.054	0.042	0.045	0.051	0.053
	0.9	0.040	0.045	0.066	0.045	0.051	0.048	0.072	0.047	0.045	0.045	0.073	0.048
-10	0.1	0.048	0.051	0.063	0.040	0.047	0.047	0.065	0.037	0.053	0.057	0.071	0.047
	0.2	0.045	0.048	0.047	0.049	0.045	0.048	0.047	0.051	0.048	0.049	0.049	0.048
	0.3	0.042	0.047	0.044	0.054	0.050	0.051	0.047	0.056	0.043	0.045	0.046	0.054
	0.4	0.048	0.049	0.041	0.064	0.047	0.049	0.044	0.059	0.050	0.054	0.042	0.060
	0.5	0.049	0.051	0.040	0.055	0.050	0.051	0.044	0.063	0.047	0.049	0.038	0.059
	0.6	0.050	0.053	0.043	0.057	0.046	0.047	0.045	0.057	0.042	0.045	0.043	0.054
	0.7	0.049	0.052	0.045	0.058	0.047	0.050	0.044	0.055	0.045	0.048	0.041	0.052
	0.8	0.050	0.051	0.045	0.051	0.049	0.054	0.048	0.050	0.044	0.045	0.049	0.048
	0.9	0.046	0.049	0.060	0.044	0.048	0.050	0.062	0.048	0.047	0.049	0.063	0.043
$-0.5T$	0.1	0.040	0.049	0.059	0.032	0.041	0.048	0.066	0.038	0.046	0.054	0.059	0.035
	0.2	0.050	0.052	0.047	0.043	0.050	0.049	0.047	0.044	0.047	0.051	0.047	0.041
	0.3	0.051	0.052	0.041	0.043	0.048	0.053	0.042	0.043	0.045	0.047	0.036	0.040
	0.4	0.050	0.051	0.036	0.044	0.047	0.049	0.034	0.043	0.045	0.048	0.034	0.043
	0.5	0.046	0.047	0.038	0.045	0.043	0.046	0.039	0.040	0.045	0.047	0.035	0.042
	0.6	0.054	0.055	0.038	0.045	0.051	0.053	0.037	0.045	0.051	0.053	0.033	0.045
	0.7	0.054	0.057	0.043	0.046	0.051	0.052	0.037	0.043	0.048	0.050	0.040	0.046
	0.8	0.050	0.052	0.046	0.049	0.052	0.055	0.045	0.043	0.054	0.058	0.048	0.045
	0.9	0.051	0.057	0.073	0.043	0.045	0.050	0.064	0.037	0.046	0.056	0.072	0.037

Table D.6: 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 $T = 750$. 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 $\mathbf{\Sigma}_t = \left[\sigma_{ut}^2 \quad -0.95\sigma_{ut}\sigma_{vt}; \quad -0.95\sigma_{ut}\sigma_{vt} \quad \sigma_{vt}^2\right]$ and $\sigma_{ut}^2 = 1$, $\sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + 4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.051	0.054	0.074	0.058	0.052	0.054	0.073	0.059	0.046	0.049	0.074	0.061
	0.2	0.058	0.058	0.056	0.064	0.054	0.057	0.058	0.069	0.048	0.050	0.066	0.069
	0.3	0.062	0.064	0.087	0.101	0.057	0.057	0.070	0.090	0.052	0.053	0.071	0.080
	0.4	0.062	0.062	0.061	0.075	0.061	0.061	0.076	0.090	0.053	0.054	0.084	0.095
	0.5	0.065	0.065	0.057	0.072	0.064	0.065	0.066	0.078	0.052	0.053	0.068	0.083
	0.6	0.057	0.058	0.060	0.077	0.061	0.063	0.071	0.085	0.052	0.052	0.078	0.093
	0.7	0.058	0.058	0.075	0.093	0.060	0.059	0.074	0.089	0.053	0.054	0.062	0.079
	0.8	0.057	0.059	0.062	0.067	0.051	0.052	0.061	0.064	0.048	0.047	0.064	0.067
	0.9	0.052	0.052	0.065	0.061	0.050	0.049	0.071	0.060	0.044	0.045	0.069	0.061
-2.5	0.1	0.046	0.047	0.064	0.049	0.046	0.046	0.060	0.048	0.043	0.045	0.066	0.049
	0.2	0.051	0.053	0.052	0.054	0.050	0.051	0.049	0.055	0.044	0.045	0.054	0.060
	0.3	0.052	0.052	0.063	0.073	0.050	0.053	0.058	0.068	0.047	0.047	0.057	0.066
	0.4	0.055	0.056	0.052	0.060	0.048	0.048	0.057	0.070	0.046	0.046	0.059	0.070
	0.5	0.051	0.051	0.049	0.063	0.053	0.053	0.049	0.062	0.041	0.042	0.053	0.069
	0.6	0.049	0.050	0.050	0.064	0.051	0.052	0.055	0.068	0.044	0.043	0.055	0.072
	0.7	0.049	0.051	0.058	0.072	0.047	0.048	0.056	0.067	0.043	0.043	0.053	0.065
	0.8	0.050	0.052	0.054	0.057	0.051	0.051	0.055	0.057	0.043	0.044	0.050	0.054
	0.9	0.046	0.045	0.057	0.054	0.047	0.046	0.056	0.052	0.043	0.045	0.061	0.056
-10	0.1	0.047	0.046	0.054	0.045	0.045	0.045	0.055	0.046	0.047	0.048	0.059	0.048
	0.2	0.046	0.047	0.048	0.049	0.050	0.051	0.046	0.053	0.049	0.050	0.046	0.055
	0.3	0.049	0.051	0.049	0.058	0.051	0.052	0.044	0.057	0.051	0.051	0.046	0.054
	0.4	0.050	0.050	0.045	0.056	0.045	0.047	0.043	0.054	0.050	0.050	0.047	0.063
	0.5	0.051	0.051	0.044	0.053	0.049	0.051	0.041	0.053	0.049	0.050	0.047	0.059
	0.6	0.049	0.049	0.043	0.058	0.051	0.053	0.046	0.057	0.047	0.048	0.048	0.062
	0.7	0.049	0.049	0.051	0.062	0.047	0.047	0.046	0.057	0.046	0.047	0.044	0.055
	0.8	0.050	0.051	0.051	0.054	0.047	0.046	0.050	0.052	0.046	0.045	0.050	0.052
	0.9	0.049	0.049	0.059	0.047	0.051	0.051	0.058	0.052	0.051	0.052	0.056	0.051
$-0.5T$	0.1	0.049	0.048	0.057	0.042	0.050	0.053	0.057	0.046	0.048	0.053	0.058	0.043
	0.2	0.050	0.052	0.050	0.047	0.053	0.054	0.049	0.050	0.052	0.053	0.053	0.047
	0.3	0.051	0.052	0.043	0.051	0.049	0.050	0.043	0.047	0.056	0.057	0.049	0.053
	0.4	0.048	0.047	0.037	0.046	0.047	0.048	0.038	0.044	0.051	0.051	0.044	0.052
	0.5	0.046	0.047	0.041	0.044	0.048	0.048	0.042	0.044	0.050	0.050	0.046	0.047
	0.6	0.050	0.051	0.037	0.048	0.047	0.048	0.038	0.044	0.051	0.051	0.040	0.052
	0.7	0.046	0.048	0.036	0.047	0.050	0.050	0.043	0.047	0.052	0.052	0.042	0.046
	0.8	0.042	0.045	0.046	0.038	0.049	0.051	0.047	0.043	0.051	0.050	0.047	0.046
	0.9	0.047	0.050	0.051	0.039	0.047	0.049	0.055	0.039	0.042	0.045	0.055	0.034

Table D.7: Empirical null rejection frequencies at 5% significance level of the QR based predictability tests $\mathcal{T}_\tau$ (Eicker-White standard errors), $\mathcal{T}^0_\tau$ (conventional standard errors), IVX_{QR}, and IVX^{MBB}_{QR}. Sample size $T = 250$. DGP 4 (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 $\mathbf{\Sigma}_t = \left[\sigma^2_{ut} \quad -0.95\sigma_{ut}\sigma_{vt}; \quad -0.95\sigma_{ut}\sigma_{vt} \quad \sigma^2_{vt}\right]$ and $\sigma^2_{ut} = \sigma^2_{vt} = 1\mathbb{I}(t \le \lfloor \lambda T \rfloor) + 1/4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}^0_\tau$	IVX_{QR}	IVX^{MBB}_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}^0_\tau$	IVX_{QR}	IVX^{MBB}_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}^0_\tau$	IVX_{QR}	IVX^{MBB}_{QR}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.054	0.055	0.170	0.063	0.055	0.057	0.160	0.056	0.060	0.065	0.144	0.064
	0.2	0.062	0.067	0.150	0.068	0.071	0.074	0.143	0.063	0.064	0.068	0.121	0.065
	0.3	0.062	0.067	0.162	0.087	0.073	0.080	0.152	0.094	0.063	0.066	0.140	0.095
	0.4	0.065	0.070	0.174	0.106	0.074	0.079	0.150	0.103	0.069	0.074	0.116	0.092
	0.5	0.064	0.067	0.162	0.109	0.070	0.075	0.142	0.093	0.067	0.072	0.112	0.086
	0.6	0.073	0.075	0.175	0.110	0.079	0.083	0.156	0.103	0.075	0.078	0.119	0.095
	0.7	0.067	0.071	0.165	0.093	0.076	0.082	0.155	0.091	0.071	0.077	0.129	0.098
	0.8	0.067	0.070	0.149	0.072	0.075	0.077	0.139	0.071	0.072	0.077	0.122	0.068
	0.9	0.055	0.058	0.178	0.064	0.058	0.059	0.159	0.058	0.057	0.064	0.142	0.062
-2.5	0.1	0.043	0.045	0.119	0.044	0.039	0.040	0.115	0.045	0.042	0.050	0.111	0.051
	0.2	0.041	0.044	0.097	0.055	0.045	0.047	0.098	0.050	0.041	0.045	0.088	0.054
	0.3	0.044	0.046	0.092	0.065	0.045	0.049	0.094	0.067	0.043	0.047	0.097	0.068
	0.4	0.043	0.046	0.096	0.075	0.044	0.046	0.095	0.069	0.048	0.051	0.083	0.068
	0.5	0.046	0.048	0.095	0.070	0.045	0.048	0.093	0.064	0.047	0.049	0.075	0.063
	0.6	0.047	0.049	0.098	0.077	0.046	0.049	0.092	0.072	0.045	0.046	0.077	0.067
	0.7	0.047	0.050	0.093	0.069	0.047	0.050	0.096	0.067	0.044	0.049	0.080	0.069
	0.8	0.044	0.048	0.098	0.060	0.046	0.046	0.095	0.056	0.046	0.047	0.083	0.054
	0.9	0.038	0.043	0.119	0.051	0.041	0.044	0.120	0.048	0.042	0.044	0.111	0.051
-10	0.1	0.052	0.051	0.107	0.042	0.045	0.049	0.107	0.043	0.052	0.056	0.103	0.045
	0.2	0.047	0.052	0.085	0.046	0.047	0.052	0.085	0.044	0.048	0.053	0.075	0.048
	0.3	0.049	0.051	0.075	0.055	0.049	0.051	0.082	0.050	0.047	0.048	0.073	0.054
	0.4	0.053	0.054	0.075	0.062	0.046	0.048	0.074	0.057	0.051	0.054	0.071	0.058
	0.5	0.051	0.053	0.072	0.062	0.047	0.050	0.076	0.055	0.048	0.050	0.061	0.056
	0.6	0.053	0.056	0.081	0.063	0.053	0.054	0.076	0.055	0.051	0.053	0.066	0.055
	0.7	0.055	0.060	0.081	0.058	0.050	0.053	0.078	0.051	0.048	0.050	0.068	0.055
	0.8	0.051	0.052	0.085	0.055	0.053	0.054	0.087	0.048	0.046	0.051	0.076	0.044
	0.9	0.048	0.049	0.105	0.047	0.044	0.049	0.111	0.045	0.046	0.052	0.105	0.047
$-0.5T$	0.1	0.041	0.051	0.130	0.041	0.044	0.051	0.124	0.040	0.045	0.053	0.110	0.040
	0.2	0.050	0.053	0.095	0.040	0.043	0.046	0.098	0.041	0.049	0.048	0.081	0.042
	0.3	0.051	0.054	0.094	0.046	0.046	0.048	0.089	0.046	0.046	0.049	0.077	0.044
	0.4	0.048	0.052	0.090	0.048	0.048	0.051	0.080	0.044	0.047	0.049	0.068	0.043
	0.5	0.051	0.052	0.088	0.051	0.042	0.044	0.082	0.043	0.043	0.045	0.069	0.045
	0.6	0.057	0.058	0.088	0.051	0.049	0.050	0.081	0.045	0.050	0.051	0.071	0.043
	0.7	0.053	0.056	0.094	0.049	0.044	0.049	0.085	0.046	0.046	0.050	0.075	0.047
	0.8	0.052	0.053	0.102	0.049	0.050	0.053	0.093	0.045	0.051	0.053	0.086	0.045
	0.9	0.046	0.053	0.133	0.044	0.040	0.049	0.134	0.039	0.043	0.048	0.114	0.041

Table D.8: 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 $T = 750$. DGP 4 (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 $\mathbf{\Sigma}_t = \begin{bmatrix}\sigma_{ut}^2 & -0.95\sigma_{ut}\sigma_{vt}; & -0.95\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2\end{bmatrix}$ and $\sigma_{ut}^2 = \sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + 1/4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.055	0.055	0.155	0.062	0.066	0.067	0.150	0.062	0.059	0.058	0.135	0.062
	0.2	0.057	0.061	0.140	0.064	0.074	0.074	0.134	0.068	0.064	0.062	0.118	0.067
	0.3	0.068	0.069	0.172	0.096	0.074	0.075	0.158	0.097	0.066	0.069	0.150	0.117
	0.4	0.075	0.076	0.192	0.121	0.080	0.082	0.154	0.104	0.067	0.069	0.114	0.081
	0.5	0.077	0.079	0.171	0.094	0.080	0.082	0.139	0.080	0.073	0.073	0.104	0.070
	0.6	0.076	0.077	0.195	0.124	0.075	0.075	0.162	0.106	0.068	0.069	0.116	0.079
	0.7	0.079	0.078	0.178	0.103	0.075	0.075	0.169	0.113	0.072	0.072	0.156	0.117
	0.8	0.069	0.070	0.157	0.074	0.069	0.070	0.140	0.070	0.069	0.069	0.118	0.067
	0.9	0.058	0.058	0.171	0.071	0.060	0.060	0.160	0.069	0.057	0.058	0.136	0.060
-2.5	0.1	0.044	0.043	0.103	0.049	0.041	0.043	0.107	0.048	0.038	0.040	0.107	0.055
	0.2	0.044	0.044	0.088	0.049	0.045	0.044	0.092	0.055	0.043	0.044	0.085	0.057
	0.3	0.043	0.045	0.093	0.064	0.045	0.046	0.099	0.065	0.039	0.040	0.100	0.073
	0.4	0.044	0.044	0.116	0.078	0.044	0.045	0.098	0.069	0.041	0.042	0.080	0.064
	0.5	0.044	0.044	0.102	0.067	0.044	0.045	0.097	0.067	0.043	0.043	0.079	0.057
	0.6	0.047	0.049	0.111	0.077	0.043	0.044	0.103	0.068	0.045	0.044	0.080	0.060
	0.7	0.044	0.045	0.102	0.071	0.045	0.045	0.107	0.070	0.047	0.047	0.099	0.073
	0.8	0.045	0.047	0.100	0.059	0.042	0.043	0.104	0.057	0.044	0.046	0.091	0.055
	0.9	0.048	0.047	0.123	0.060	0.047	0.045	0.117	0.055	0.042	0.044	0.112	0.052
-10	0.1	0.047	0.048	0.099	0.048	0.044	0.047	0.104	0.046	0.050	0.052	0.098	0.048
	0.2	0.048	0.048	0.089	0.049	0.049	0.050	0.088	0.053	0.049	0.049	0.076	0.054
	0.3	0.050	0.052	0.083	0.054	0.049	0.051	0.088	0.055	0.052	0.052	0.087	0.064
	0.4	0.051	0.053	0.088	0.062	0.052	0.053	0.079	0.054	0.049	0.050	0.067	0.054
	0.5	0.048	0.049	0.086	0.061	0.053	0.054	0.081	0.055	0.052	0.053	0.073	0.052
	0.6	0.053	0.054	0.085	0.061	0.052	0.052	0.084	0.055	0.048	0.048	0.073	0.054
	0.7	0.053	0.053	0.087	0.060	0.050	0.050	0.094	0.057	0.055	0.055	0.081	0.059
	0.8	0.053	0.054	0.098	0.055	0.053	0.052	0.097	0.053	0.053	0.054	0.088	0.049
	0.9	0.056	0.059	0.111	0.054	0.054	0.056	0.110	0.052	0.050	0.051	0.103	0.048
$-0.5T$	0.1	0.053	0.055	0.128	0.047	0.050	0.052	0.119	0.046	0.048	0.049	0.113	0.046
	0.2	0.054	0.055	0.110	0.053	0.048	0.049	0.097	0.046	0.054	0.054	0.092	0.053
	0.3	0.045	0.045	0.102	0.047	0.050	0.050	0.100	0.047	0.053	0.054	0.086	0.050
	0.4	0.049	0.049	0.094	0.044	0.047	0.047	0.094	0.047	0.053	0.053	0.085	0.049
	0.5	0.050	0.051	0.099	0.048	0.046	0.047	0.091	0.045	0.048	0.049	0.082	0.045
	0.6	0.052	0.052	0.098	0.045	0.054	0.053	0.094	0.052	0.051	0.051	0.078	0.047
	0.7	0.050	0.050	0.098	0.046	0.049	0.051	0.102	0.051	0.052	0.053	0.092	0.049
	0.8	0.044	0.046	0.109	0.041	0.052	0.054	0.113	0.048	0.049	0.049	0.098	0.048
	0.9	0.047	0.052	0.123	0.043	0.043	0.046	0.130	0.042	0.046	0.049	0.111	0.039

Table D.9: 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 4 (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 $\sigma_{ut}^2 = \sigma_{vt}^2 = 1\mathbb{I}(t \leq \lfloor \lambda T \rfloor) + 4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

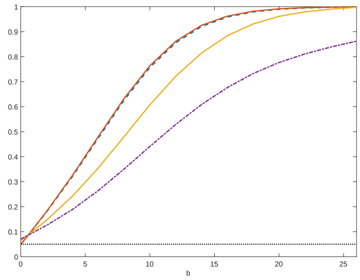
c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.052	0.054	0.101	0.058	0.053	0.053	0.125	0.055	0.054	0.051	0.143	0.062
	0.2	0.052	0.055	0.078	0.064	0.050	0.053	0.102	0.067	0.051	0.055	0.104	0.068
	0.3	0.055	0.057	0.080	0.077	0.046	0.050	0.102	0.076	0.051	0.052	0.106	0.077
	0.4	0.056	0.059	0.071	0.082	0.055	0.059	0.096	0.085	0.051	0.054	0.104	0.090
	0.5	0.056	0.058	0.069	0.079	0.059	0.062	0.088	0.087	0.052	0.055	0.100	0.088
	0.6	0.059	0.063	0.073	0.086	0.053	0.057	0.101	0.086	0.049	0.053	0.109	0.089
	0.7	0.062	0.066	0.086	0.091	0.052	0.055	0.107	0.080	0.054	0.058	0.105	0.078
	0.8	0.058	0.062	0.079	0.066	0.057	0.058	0.097	0.069	0.054	0.057	0.107	0.065
	0.9	0.048	0.051	0.106	0.060	0.051	0.050	0.135	0.064	0.048	0.050	0.139	0.061
-2.5	0.1	0.055	0.053	0.111	0.051	0.051	0.049	0.130	0.051	0.051	0.050	0.133	0.055
	0.2	0.048	0.053	0.085	0.056	0.049	0.051	0.105	0.057	0.047	0.052	0.108	0.055
	0.3	0.048	0.051	0.080	0.064	0.047	0.048	0.097	0.066	0.044	0.046	0.100	0.064
	0.4	0.049	0.050	0.072	0.068	0.049	0.052	0.094	0.073	0.048	0.050	0.099	0.071
	0.5	0.054	0.056	0.072	0.071	0.051	0.054	0.093	0.077	0.047	0.049	0.096	0.073
	0.6	0.052	0.055	0.075	0.069	0.047	0.050	0.103	0.069	0.045	0.048	0.096	0.070
	0.7	0.053	0.056	0.085	0.077	0.048	0.051	0.094	0.072	0.046	0.048	0.097	0.066
	0.8	0.048	0.051	0.086	0.062	0.053	0.056	0.101	0.061	0.045	0.048	0.101	0.055
	0.9	0.043	0.047	0.110	0.051	0.050	0.049	0.135	0.054	0.044	0.047	0.129	0.052
-10	0.1	0.047	0.052	0.120	0.046	0.048	0.048	0.134	0.044	0.052	0.057	0.130	0.050
	0.2	0.047	0.051	0.090	0.053	0.047	0.050	0.105	0.054	0.049	0.053	0.107	0.055
	0.3	0.046	0.050	0.081	0.060	0.053	0.052	0.097	0.060	0.044	0.048	0.097	0.059
	0.4	0.049	0.050	0.078	0.065	0.046	0.047	0.093	0.064	0.050	0.053	0.088	0.065
	0.5	0.051	0.053	0.077	0.061	0.053	0.055	0.094	0.064	0.050	0.052	0.090	0.063
	0.6	0.052	0.053	0.079	0.061	0.047	0.048	0.092	0.058	0.046	0.048	0.089	0.059
	0.7	0.052	0.055	0.087	0.060	0.049	0.052	0.098	0.059	0.044	0.050	0.091	0.057
	0.8	0.052	0.052	0.088	0.056	0.052	0.055	0.101	0.051	0.047	0.048	0.096	0.052
	0.9	0.050	0.050	0.115	0.049	0.051	0.048	0.129	0.052	0.047	0.051	0.130	0.046
$-0.5T$	0.1	0.039	0.047	0.119	0.036	0.043	0.049	0.134	0.039	0.045	0.056	0.126	0.042
	0.2	0.050	0.055	0.086	0.045	0.050	0.051	0.101	0.047	0.047	0.050	0.110	0.046
	0.3	0.048	0.050	0.084	0.046	0.048	0.052	0.086	0.044	0.046	0.047	0.095	0.044
	0.4	0.048	0.049	0.076	0.050	0.047	0.049	0.084	0.046	0.045	0.048	0.084	0.045
	0.5	0.047	0.048	0.073	0.047	0.044	0.047	0.085	0.045	0.047	0.048	0.085	0.045
	0.6	0.053	0.055	0.078	0.049	0.049	0.051	0.081	0.047	0.048	0.050	0.090	0.048
	0.7	0.053	0.055	0.082	0.049	0.050	0.053	0.090	0.050	0.045	0.047	0.095	0.047
	0.8	0.051	0.054	0.093	0.051	0.049	0.053	0.102	0.047	0.051	0.056	0.107	0.047
	0.9	0.054	0.058	0.129	0.046	0.044	0.050	0.137	0.040	0.046	0.057	0.141	0.040

Table D.10: 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 $T = 750$. DGP 4 (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 $\mathbf{\Sigma}_t = \begin{bmatrix}\sigma_{ut}^2 & -0.95\sigma_{ut}\sigma_{vt}; & -0.95\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2\end{bmatrix}$ and $\sigma_{ut}^2 = \sigma_{vt}^2 = 1\mathbb{I}(t \le \lfloor \lambda T \rfloor) + 4\mathbb{I}(t > \lfloor \lambda T \rfloor)$.

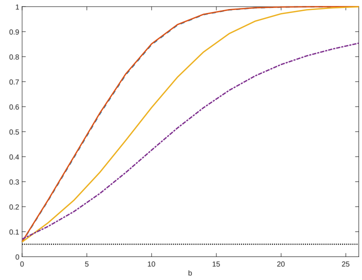
c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	IVX_{QR}^{MBB}
		$\lambda = 1/3$				$\lambda = 1/2$				$\lambda = 2/3$			
0	0.1	0.055	0.056	0.103	0.062	0.052	0.052	0.122	0.060	0.049	0.048	0.130	0.060
	0.2	0.058	0.058	0.079	0.065	0.055	0.057	0.101	0.069	0.050	0.051	0.108	0.069
	0.3	0.061	0.063	0.105	0.090	0.058	0.060	0.113	0.085	0.052	0.052	0.118	0.084
	0.4	0.059	0.060	0.081	0.072	0.060	0.062	0.110	0.088	0.053	0.055	0.128	0.098
	0.5	0.060	0.061	0.075	0.072	0.064	0.064	0.105	0.075	0.051	0.052	0.109	0.085
	0.6	0.059	0.061	0.080	0.075	0.059	0.060	0.107	0.083	0.053	0.054	0.126	0.095
	0.7	0.057	0.057	0.096	0.084	0.059	0.060	0.117	0.084	0.052	0.052	0.116	0.079
	0.8	0.060	0.059	0.085	0.068	0.052	0.053	0.102	0.071	0.048	0.047	0.112	0.068
	0.9	0.054	0.051	0.101	0.062	0.050	0.050	0.128	0.059	0.046	0.047	0.124	0.067
-2.5	0.1	0.049	0.051	0.113	0.050	0.049	0.048	0.124	0.053	0.046	0.047	0.130	0.054
	0.2	0.054	0.054	0.084	0.054	0.053	0.055	0.099	0.061	0.047	0.049	0.106	0.061
	0.3	0.053	0.053	0.094	0.071	0.054	0.055	0.106	0.070	0.052	0.051	0.109	0.066
	0.4	0.058	0.059	0.085	0.059	0.054	0.056	0.109	0.071	0.047	0.048	0.114	0.076
	0.5	0.052	0.053	0.080	0.064	0.058	0.059	0.102	0.065	0.044	0.044	0.106	0.067
	0.6	0.052	0.053	0.087	0.062	0.053	0.053	0.105	0.067	0.047	0.047	0.108	0.072
	0.7	0.049	0.051	0.092	0.067	0.052	0.053	0.103	0.067	0.048	0.048	0.111	0.064
	0.8	0.053	0.053	0.096	0.060	0.054	0.055	0.106	0.060	0.048	0.048	0.105	0.059
	0.9	0.047	0.046	0.109	0.056	0.047	0.047	0.134	0.055	0.043	0.046	0.123	0.061
-10	0.1	0.048	0.050	0.114	0.045	0.047	0.047	0.123	0.051	0.051	0.051	0.125	0.052
	0.2	0.047	0.048	0.093	0.053	0.053	0.054	0.100	0.057	0.053	0.055	0.105	0.057
	0.3	0.053	0.054	0.092	0.061	0.053	0.054	0.102	0.062	0.054	0.054	0.102	0.058
	0.4	0.052	0.053	0.089	0.060	0.049	0.050	0.104	0.058	0.050	0.050	0.101	0.064
	0.5	0.052	0.052	0.085	0.057	0.053	0.054	0.098	0.052	0.051	0.051	0.098	0.057
	0.6	0.049	0.049	0.087	0.059	0.053	0.054	0.101	0.057	0.050	0.051	0.105	0.060
	0.7	0.051	0.051	0.092	0.064	0.048	0.049	0.102	0.060	0.047	0.048	0.104	0.060
	0.8	0.052	0.052	0.092	0.054	0.050	0.050	0.109	0.057	0.047	0.049	0.104	0.051
	0.9	0.051	0.051	0.119	0.051	0.051	0.051	0.130	0.053	0.051	0.053	0.122	0.050
$-0.5T$	0.1	0.049	0.048	0.113	0.045	0.048	0.050	0.127	0.044	0.049	0.053	0.123	0.045
	0.2	0.051	0.052	0.101	0.050	0.050	0.053	0.112	0.050	0.052	0.052	0.107	0.051
	0.3	0.051	0.051	0.088	0.051	0.049	0.049	0.095	0.053	0.055	0.056	0.104	0.054
	0.4	0.049	0.047	0.078	0.048	0.048	0.049	0.094	0.047	0.050	0.051	0.099	0.053
	0.5	0.044	0.045	0.083	0.044	0.045	0.045	0.095	0.046	0.049	0.050	0.102	0.052
	0.6	0.050	0.050	0.085	0.049	0.050	0.050	0.098	0.048	0.048	0.049	0.096	0.051
	0.7	0.049	0.049	0.076	0.047	0.050	0.050	0.089	0.047	0.050	0.050	0.099	0.050
	0.8	0.044	0.046	0.089	0.040	0.049	0.051	0.108	0.047	0.053	0.052	0.111	0.047
	0.9	0.047	0.052	0.115	0.043	0.046	0.050	0.126	0.043	0.042	0.044	0.121	0.036

Table D.11: 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 $T = 750$. 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 $\boldsymbol{v}_t = -\gamma f_t + \boldsymbol{e}_t$, where the common factor f_t is generated as a sequence of independent standard normals, $\gamma = -0.95$ and $(a_t, \boldsymbol{e}_t')' \sim i.i.d.\ N(\mathbf{0}, \boldsymbol{I}_{K+1})$. The predictors are generated for $\boldsymbol{\Gamma} = \rho \mathbf{I}_K$ with $\rho = 1 + c/T$ for $c = \{0, -2.5, -10, -0.5T\}$.

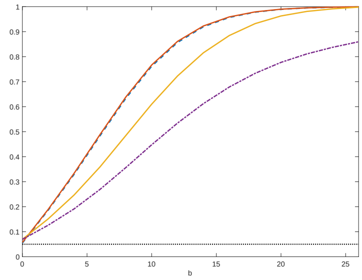
c	τ	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}	$\mathcal{T}_\tau$	$\mathcal{T}_\tau^0$	IVX_{QR}
			$K=2$			$K=3$			$K=4$			$K=5$	
0	0.1	0.039	0.039	0.062	0.038	0.038	0.081	0.042	0.046	0.090	0.048	0.049	0.111
	0.2	0.033	0.033	0.047	0.042	0.044	0.059	0.049	0.048	0.067	0.057	0.059	0.080
	0.3	0.038	0.039	0.046	0.047	0.047	0.059	0.054	0.054	0.065	0.055	0.056	0.072
	0.4	0.040	0.039	0.048	0.046	0.046	0.056	0.055	0.055	0.058	0.060	0.058	0.065
	0.5	0.036	0.036	0.053	0.048	0.048	0.056	0.057	0.057	0.058	0.066	0.066	0.071
	0.6	0.039	0.039	0.047	0.047	0.048	0.052	0.057	0.057	0.059	0.065	0.065	0.064
	0.7	0.043	0.042	0.052	0.045	0.045	0.057	0.049	0.050	0.060	0.057	0.056	0.068
	0.8	0.036	0.035	0.053	0.044	0.045	0.058	0.043	0.044	0.064	0.050	0.049	0.079
	0.9	0.036	0.033	0.067	0.036	0.038	0.080	0.046	0.045	0.098	0.045	0.047	0.109
-2.5	0.1	0.043	0.042	0.058	0.043	0.043	0.078	0.043	0.044	0.087	0.041	0.047	0.107
	0.2	0.051	0.051	0.050	0.044	0.046	0.055	0.051	0.051	0.068	0.050	0.050	0.078
	0.3	0.041	0.042	0.047	0.058	0.056	0.057	0.044	0.043	0.057	0.057	0.054	0.063
	0.4	0.048	0.048	0.042	0.047	0.047	0.057	0.045	0.045	0.049	0.055	0.056	0.065
	0.5	0.056	0.056	0.049	0.046	0.046	0.048	0.049	0.049	0.054	0.050	0.050	0.053
	0.6	0.050	0.051	0.047	0.052	0.051	0.050	0.045	0.044	0.057	0.060	0.059	0.065
	0.7	0.045	0.045	0.049	0.054	0.051	0.060	0.046	0.046	0.064	0.049	0.050	0.065
	0.8	0.044	0.044	0.047	0.049	0.048	0.055	0.049	0.047	0.066	0.047	0.046	0.067
	0.9	0.043	0.043	0.062	0.044	0.042	0.076	0.039	0.041	0.083	0.042	0.042	0.104
-10	0.1	0.050	0.056	0.062	0.053	0.052	0.073	0.051	0.050	0.089	0.044	0.049	0.091
	0.2	0.055	0.053	0.051	0.056	0.059	0.055	0.052	0.050	0.057	0.049	0.051	0.061
	0.3	0.058	0.058	0.046	0.056	0.058	0.049	0.053	0.053	0.049	0.057	0.056	0.057
	0.4	0.056	0.059	0.043	0.060	0.060	0.042	0.052	0.051	0.050	0.054	0.056	0.056
	0.5	0.058	0.058	0.042	0.054	0.054	0.045	0.063	0.063	0.052	0.053	0.053	0.050
	0.6	0.054	0.055	0.043	0.054	0.053	0.043	0.057	0.056	0.051	0.067	0.067	0.054
	0.7	0.052	0.056	0.043	0.053	0.053	0.045	0.062	0.064	0.057	0.051	0.053	0.058
	0.8	0.051	0.051	0.047	0.053	0.051	0.054	0.051	0.047	0.059	0.051	0.050	0.071
	0.9	0.055	0.052	0.060	0.052	0.051	0.073	0.048	0.050	0.083	0.047	0.045	0.098
$-0.5T$	0.1	0.052	0.052	0.059	0.049	0.045	0.069	0.046	0.048	0.073	0.043	0.046	0.096
	0.2	0.053	0.053	0.052	0.050	0.049	0.050	0.045	0.042	0.054	0.047	0.046	0.059
	0.3	0.055	0.056	0.047	0.051	0.050	0.050	0.050	0.050	0.049	0.043	0.045	0.047
	0.4	0.053	0.053	0.041	0.052	0.052	0.042	0.051	0.052	0.046	0.047	0.048	0.049
	0.5	0.053	0.053	0.042	0.049	0.049	0.039	0.048	0.048	0.044	0.045	0.045	0.048
	0.6	0.053	0.052	0.041	0.054	0.055	0.045	0.045	0.044	0.044	0.050	0.051	0.044
	0.7	0.053	0.053	0.042	0.045	0.046	0.044	0.044	0.047	0.052	0.046	0.048	0.048
	0.8	0.056	0.055	0.047	0.058	0.060	0.057	0.047	0.053	0.055	0.049	0.049	0.057
	0.9	0.051	0.050	0.065	0.048	0.050	0.072	0.048	0.049	0.082	0.044	0.046	0.095



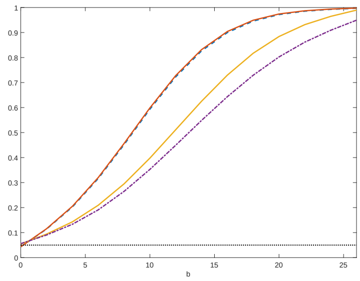
(a) $\tau = 0.2$, $c = 0$



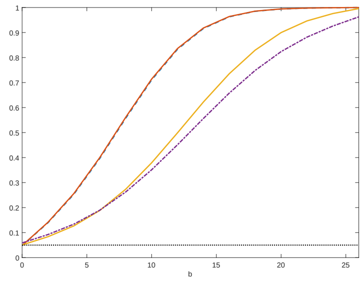
(b) $\tau = 0.5$, $c = 0$



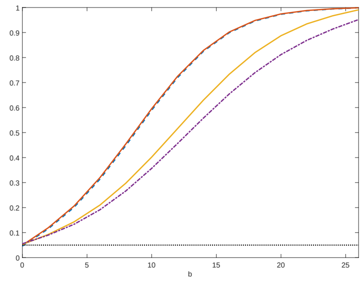
(c) $\tau = 0.8$, $c = 0$



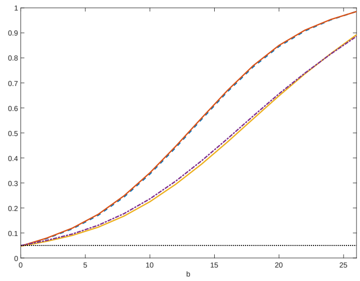
(d) $\tau = 0.2$, $c = -2.5$



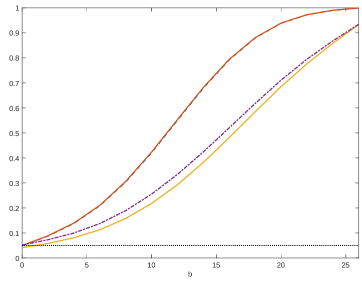
(e) $\tau = 0.5$, $c = -2.5$



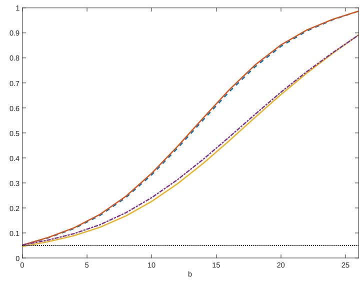
(f) $\tau = 0.8$, $c = -2.5$



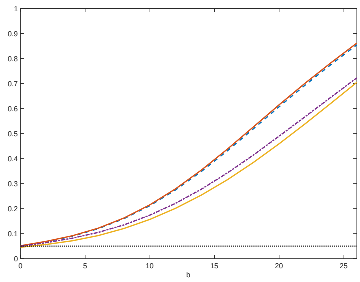
(g) $\tau = 0.2$, $c = -10$



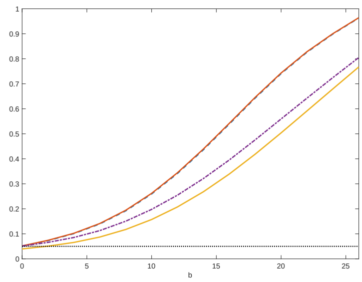
(h) $\tau = 0.5$, $c = -10$



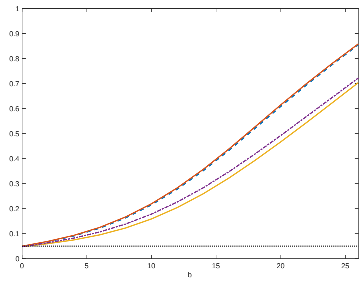
(i) $\tau = 0.8$, $c = -10$



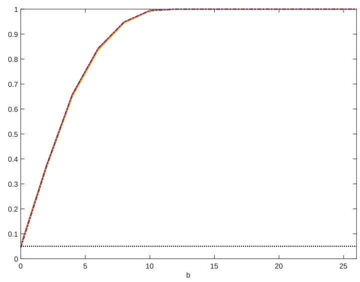
(j) $\tau = 0.2$, $c = -20$



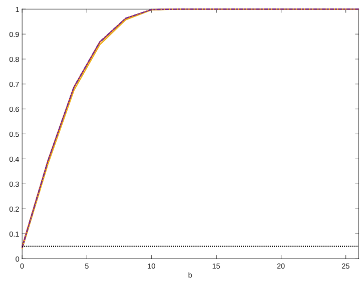
(k) $\tau = 0.5$, $c = -20$



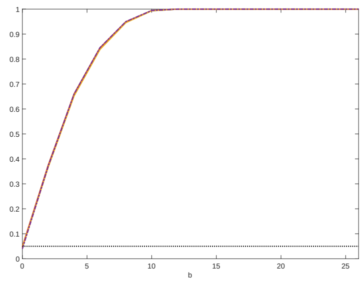
(l) $\tau = 0.8$, $c = -20$



(m) $\tau = 0.2$, $c = -0.5T$



(n) $\tau = 0.5$, $c = -0.5T$



(o) $\tau = 0.8$, $c = -0.5T$

$\mathcal{T}_\tau$ $\mathcal{T}_\tau^0$ IVX_{QR} IVX_{QR}^{MBB} 5%

Figure D.1: Finite sample local power of predictability tests for $T = 750$. DGP1 (homoskedastic and serially uncorrelated innovations): $y_t = \beta_\tau x_{t-1} + u_{t\tau}$, $x_t = \rho x_{t-1} + v_t$, $\rho = 1 + c/T$, $v_t = \varepsilon_t$ and $\beta_\tau = b_\tau/T$ (except for $c = -0.5T$ where $\beta_\tau = b_\tau/\sqrt{T}$), $b_\tau = \{0, 1, ..., 26\}$, and $(u_t, \varepsilon_t)' \sim$ $i.i.d.\ N(\mathbf{0}, \mathbf{\Sigma})$, with $\mathbf{\Sigma} = [1 \quad -0.95; \quad -0.95 \quad 1]$.

References

Adrian, T., R. K. Crump, and E. Moench (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics 110*(1), 110–138.

Almeida, C., K. Ardison, R. Garcia, and J. Vicente (2017, 03). Nonparametric Tail Risk, Stock Returns, and the Macroeconomy. *Journal of Financial Econometrics 15*(3), 333–376.

Andersen, T. G., N. Fusari, and V. Todorov (2015a). Parametric inference and dynamic state recovery from option panels. *Econometrica 83*(3), 1081–1145.

Andersen, T. G., N. Fusari, and V. Todorov (2015b). The risk premia embedded in index options. *Journal of Financial Economics 117*(3), 558–584.

Backus, D., M. Chernov, and I. Martin (2011). Disasters implied by equity index options. *The Journal of Finance 66*(6), 1969–2012.

Bakshi, G., N. Kapadia, and D. Madan (2003, 06). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies 16*(1), 101–143.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2011). Tails, fears, and risk premia. *The Journal of Finance 66*(6), 2165–2211.

Bollerslev, T., V. Todorov, and L. Xu (2015). Tail risk premia and return predictability. *Journal of Financial Economics 118*(1), 113–134.

Christoffersen, P., B. Feunou, Y. Jeon, and C. Ornthanalai (2020, 12). Time-Varying Crash Risk Embedded in Index Options: The Role of Stock Market Liquidity. *Review of Finance 25*(4), 1261–1298.

De Jong, R. M., C. Amsler, and P. Schmidt (2007). A robust version of the KPSS test based on indicators. *Journal of Econometrics 137*(2), 311–333.

Demetrescu, M., I. Georgiev, P. M. M. Rodrigues, and A. M. R. Taylor (2022). Testing for episodic predictability in stock returns. *Journal of Econometrics 227*(1), 85–113.

Fan, R. and J. H. Lee (2019). Predictive quantile regressions under persistence and conditional heteroskedasticity. *Journal of Econometrics 213*(1), 261–280.

Gao, G. P., X. Lu, and Z. Song (2019). Tail risk concerns everywhere. *Management Science 65*(7), 3111–3130.

Goyal, A. and I. Welch (2003). Predicting the equity premium with dividend ratios. *Management Science 49*(5), 639–654.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics 3*, 1163–1174.

Kelly, B. (2014). The dynamic power law model. *Extremes 17*, 557–583.

Kelly, B. and H. Jiang (2014). Tail Risk and Asset Prices. *The Review of Financial Studies 27*(10), 2841–2871.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *Review of Financial Studies 28*(5), 1506–1553.

Lee, J. H. (2016). Predictive quantile regression with persistent covariates: IVX-QR approach. *Journal of Econometrics 192*(1), 105–118.

Nicolau, J., P. M. M. Rodrigues, and M. Z. Stoykov (2023). Tail index estimation in the presence of covariates: Stock returns' tail risk dynamics. *Journal of Econometrics 235*(2), 2266–2284.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Weller, B. M. (2018, 12). Measuring Tail Risks at High Frequency. *The Review of Financial Studies 32*(9), 3571–3616.

Zhou, H. (2018). Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. *Annual Review of Financial Economics 10*(1), 481–497.