Inference in Predictive Quantile Regressions*

Alex Maynard†
University of Guelph

Katsumi Shimotsu‡
Hitotsubashi University

Yini Wang§
Queen's University

October 30, 2011

Abstract

This paper studies inference in predictive quantile regressions when the predictive regressor has a near-unit root. We derive nonstandard distributions for the quantile regression estimator and t-statistic in terms of functionals of diffusion processes. The critical values are found to depend on both the quantile of interest and the local-to-unity parameter, which is not consistently estimable. Based on these critical values, we propose a valid Bonferroni bounds test for quantile predictability with persistent regressors. We employ this new methodology to test the ability of many commonly employed and highly persistent regressors, such as the dividend yield, earnings price ratio, book to market ratio, term spread and T-bill rate, to predict the median, shoulders, and tails of the stock return distribution.

JEL Classification: C22, G1

Keywords: local-to-unity, quantile regression, Bonferroni method, predictability, stock return

*We thank Ying Chen, Chuan Goh, Bruce Hansen, Chrisian Gourieroux, Frank Kleibergen, Benoit Perron, Chi Wan, Zhijie Xiao, and conference participants at the Canadian Econometrics Study Group, the Midwest Econometrics Study Group, and the Canadian Economic Association for useful comments and discussion. We thank Tolga Cenesizoglu for providing the data. Maynard thanks the SSHRC for research funding. We take full responsibility for any remaining errors.

†Department of Economics, McKinnan Building, University of Guelph, Guelph, Ontario N1G 2W1, Canada. E-mail: maynarda@uoguelph.ca

‡Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. E-mail: shimotsu@econ.hit-u.ac.jp

§Department of Economics, Queen's University, Kingston, Ontario K7L 3N6, Canada. E-mail: wangyini@econ.queensu.ca

1 Introduction

In this paper, we develop asymptotic theory in the context of predictive quantile regressions with nearly nonstationary regressors. This has important empirical applications, such as testing the predictability of the stock return distribution. Beginning with influential work by Shiller (1984), Campbell and Shiller (1988a, 1988b), Fama and French (1988) and Hodrick (1992), there has been an extensive literature on predictive tests for mean returns, debating whether predictors such as dividend yield can forecast stock returns. This has implications, not only for the risk neutral market efficiency hypothesis, but also for portfolio analysis. Indeed, some of the most recent empirical work debates the ability of investors to use predictors, such as dividend yields, to create dynamic asset allocation strategies that outperform the market (Goyal and Welch 2008, Campbell and Thompson forthcoming). Similarly, asset allocation strategies employing the large and widely successful literature on volatility modelling, may improve the expected welfare of the risk averse investor.

While the most of the empirical literature has focused exclusively on predictive means or variances, outside of a very few special cases, such as the mean-variance portfolio models, the portfolio decision depends on the entire return distribution. Likewise, the tails of the distribution are of particular interest to risk managers and are also important to policy makers, who must consider worst case, as well as base-line, forecast scenarios. Recent empirical work by Cenesizoglu and Timmermann (2008) employs the quantile regression method introduced by Koenker and Bassett (1978), to extract a richer set of return predictions involving not just the center of the predictive density, but also the shoulders and tails. Their empirical results suggest that there is valuable information in the conditional quantiles that cannot be ascertained from the conditional mean and variance alone. For example, they find that a number of pre-determined predictors have an asymmetric effect on various quantiles in the return distribution. Likewise, they find that some predictors, which have little information for the center of the distribution, nonetheless have important implications for the tails.

One reason that the on-going debate over-predictive mean regression has lasted so long is that standard inference procedures can be unreliable when predictors are persistent, as is commonly the case in practice. Earlier work adopts standard t-tests and finds strong evidence of predictability (Shiller 1984, Campbell and Shiller 1988a, 1988b, Fama and French 1988, and Hodrick 1992). However, as Shiller (1984) points out, some regressors in the predictive regressions can be stochastic, so that the conventional t-tests are invalid. In particular, the predictability of the return with these predictors may be overstated. Mankiw and Shapiro (1986) investigates small sample properties of tests for rational expectation models. They find that ordinary tests of orthogonality reject too frequently when the predictive variables are nearly nonstationary. Other papers, such as Stambaugh (1986), Cavanagh et al. (1995) and Stambaugh (1999), also emphasize the poor small sample properties in predictive regressions with highly persistent regressors.

Conventional predictive tests of stock returns are not valid for the following reasons. Firstly, the predictor variables, such as dividend yields, dividend price and earning price ratios, are strongly autocorrelated. Secondly, although pre-determined, typically the predictor is not strictly exogenous, since its innovation is often highly correlated with the error term in the predictive regression. For example, with financial data, such as dividend price ratios and returns, it is reasonable to expect strong correlation between the regressor's innovation and the regression disturbance. This tends to inflate the t-statistic resulting in over-rejection. In linear predictive regressions, it is well-known that when the regressor is not strictly exogenous and has a largest autoregressive root close to one, the limiting distribution of the t-statistic will be nonstandard. In this case, the t-statistic is too large, and tests using the standard normal critical values will over-reject the null hypothesis of nonpredictability.

Much attention has been devoted to overcoming such size distortions in linear predictive regressions, resulting in a rich literature, with several general approaches having been explored. A number of alternatives to the standard regression based predictive testing approach have been proposed. Campbell and Dufour (1995 and 1997) propose the use of tests based on nonparametric sign and sign-rank statistics, with exact finite sample size. Maynard and Shimotsu (2009), suggest a semi-parametric covariance based predictive test that is asymptotically normal and allows for an alternative hypothesis with a stationary dependent variable, even when the predictor is near nonstationary. Wright (2000) and Lanne (2002) reinterpret the predictive test as a stationarity test and use this insight to provide conservative testing procedures.

A second strand of the literature applies finite sample corrections to the standard predictive t or F-test, often motivated under more tightly parameterized models. Stambaugh (1999) derives small-sample Bayesian posterior distributions for the regression parameters. Lewellen (2004) performs finite sample correction by calculating a joint significance level from a combination of the conditional and unconditional tests. Amihud and Hurvich (2004) and Amihud et al. (2004) obtain bias reductions using augmented regression methods. Likewise, resampling approaches have also been proposed to improve finite sample inference (Nelson and Kim 1993, Goetzmann and Jorion 1993, and Wolf, 2000).

A final approach that has proved productive, and which we adapt to the quantile context, is the use of an explicit local-to-unity specification for the predictor, in order to provide for improved estimation and inference in the standard predictive model. Cavanagh et al. (1995) propose corrected critical values based on a local-to-unity model with known values of the local to unity parameter. Since this parameter cannot be consistently estimated, they then propose feasible inference methods using bounds procedures to control size. Campbell and Yogo (2006) instead employ the local-to-unity based setting to correct the asymptotic bias of the predictive regression estimator. The resulting estimator is mixed normal and asymptotically efficient for known c. Since their correction depends on c, a refined bounds procedure is employed for feasible inference. Hjalmarsson (2007) notes that the Campbell and Yogo (2006) procedure can be interpreted as a local-to-unity version of the fully modified estimator of Phillips and Hansen (1990) and proposes a generalization. Jansson and Moreira (2006) derive a test that is both unbiased and conditionally optimal, without knowledge of the localization parameter, using Gaussian asymptotic power envelopes.

In contrast to this large literature devoted to proper inference techniques for predictive mean regression, we are aware of no theoretical work to date, that establishes valid econometric inference methods in quantile predictive regression with persistent regressors. As demonstrated by Cenesizoglu and Timmermann (2008), quantile predictive regressions have considerable empirical interest, yet our preliminary simulations indicate that standard tests based on predictive quantile regressions are subject to similar size distortion as linear predictive tests. In this paper, we develop proper inference methods for short-horizon predictive quantile regressions with nearly integrated regressors. We first derive the limit distribution of the quantile regression coefficients by generalizing results of Xiao (2009), who derives inference in a quantile regression with cointegrated time series, to the local-to-unity setting of Chan and Wei (1987), Chan (1988), Phillips (1987), Phillips (1988a, 1988b), and Nabeya and Sorensen (1994)[1]. We then provide a test of quantile predictability based on an asymptotically valid Bonferroni bounds methodology in the spirit of Cavanagh et al. (1995).

Our results contribute to a rapidly developing literature in quantile regression, in which there are substantial recent works in both theoretical and empirical areas. Recent developments of quantile methods for time series data include quantile autoregression (Koenker and Xiao 2006), unit root quantile autoregression (Koenker and Xiao 2004), quantile cointegration (Xiao, 2009), conditional quantile estimation for GARCH models (Xiao and Koenker 2009) and Copula-based quantile autoregression (Chen et al. 2009). Chernozhukov and Du (2006) and Chernozhukov (2010) develop the theory of conditional extremal quantiles with applications to value-at-risk and birth weights. Nonparametric quantile methods and their applications have also been developing rapidly (see, for example, Frolich and Melly 2008, Huber 2010, Jun et al. 2009 and Koenker 2010). Other recent contributions include Koenker (2008), who proposes censored quantile regression estimation, Chernozhukov et al. (2009), who derives finite sample inference for quantile regression models under minimal assumptions, and Chernozhukov and Belloni (2010), who study l_1-penalized quantile regression in high-dimensional sparse models.

The remainder of the paper is organized as follows. In Section 2, the framework of the problem is established and the asymptotic theory for predictive quantile regression under a local-to-unity specification is developed. In Section 3, the Bonferroni method used to correct the size distortion is proposed. In Section 4, results from our simulation study are reported. In Section 5, the techniques are applied to test the predictability of the stock return distribution using various pre-determined predictive regressors. Section 6 concludes the paper.

[1]These papers are all based on nearly nonstationary AR(1) models.

2 Theory

2.1 Model and assumptions

Denote $\mathcal{F}_t$ as information set up to time t. The standard mean prediction model is given by

$$y_t = \gamma_0 + \gamma_1 x_{t-1} + \varepsilon_{2t} \tag{1}$$

where $E[\varepsilon_{2t}|\mathcal{F}_{t-1}] = 0$, y_t is typically a financial return and x_t is a predictor, such as an earnings or dividend price ratio. This yields the mean prediction $E[y_t|\mathcal{F}_{t-1}] = \gamma_0 + \gamma_1 x_{t-1}$, but provides no information on other aspects of the predictive distribution without further assumption on the error term.

Let $F(\cdot)$ and $F_{t-1}(\cdot) = Pr(\varepsilon_{2t} < \cdot|\mathcal{F}_{t-1})$ denote the cumulative and conditional distributions of $\varepsilon_{2,t}$ and define the τth unconditional and conditional quantiles of ε_{2t} by $Q_{\varepsilon_{2t}}(\tau) = F^{-1}(\tau)$ and $Q_{\varepsilon_{2t}}(\tau|\mathcal{F}_{t-1}) = F_{t-1}^{-1}(\tau)$, respectively. With the added assumption that $F_{t-1}(\cdot) = F(\cdot)$, it is possible to use (1) to forecast quantiles of y_t employing

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \gamma_0 + Q_{\varepsilon_2}(\tau) + \gamma_1 x_{t-1} \tag{2}$$

for $\tau \in (0, 1)$, where $Q_{y_t}(\tau|\mathcal{F}_{t-1})$ is defined as the τth quantile of y_t conditional on $\mathcal{F}_{t-1}$. Nonetheless, the quantile forecasts in (2) may be viewed as somewhat restrictive since all quantiles share the same slope coefficient γ_1, allowing only for parallel shifts of the conditional quantile. It implies, for example, that the predictor x_{t-1} shifts the center, shoulders, and tails of the distribution all in the same direction and by exactly the same amount.

To obtain a more flexible predictive model, we relax the assumption that residual distribution in (1) is time invariant and allow x_{t-1} to impact not only the mean of y_t, but also the distribution of its error term. In particular, we model the conditional quantile of the error term as

$$Q_{\varepsilon_{2t}}(\tau|\mathcal{F}_{t-1}) = \gamma_0(\tau) - \gamma_0 + (\gamma_1(\tau) - \gamma_1)\, x_{t-1}, \tag{3}$$

in which the dependence of $\gamma_1(\tau)$ on τ allows the impact of x_{t-1} to vary across the quantiles of $\varepsilon_{2,t}$. Combining (2) and (3) gives the predictive quantile model for y_t:

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \gamma_0(\tau) + \gamma_1(\tau) x_{t-1}, \tag{4}$$

which provides a flexible specification allowing the effect of x_{t-1} to be heterogeneous across the quantiles of y_t.

It is evident from (2) that the standard linear mean prediction model (with constant residual distribution) is a special case of the quantile regression model in (4). As noted by Cenesizoglu and Timmermann (2008), the quantile predictive model also encompasses a number of other empirical models for financial returns. For example, if x_t is a variable with predictive content for volatility, such as squared returns or realized volatility, we may consider a model of the form

$$y_t = \gamma_0 + \gamma_1 x_{t-1} + (\delta_0 + \delta_1 x_{t-1})\varepsilon_{2,t} \tag{5}$$

where $F_{t-1}(\cdot) = F(\cdot)$. The predictive quantile for y_t then takes the form:

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \gamma_0(\tau) + \gamma_1(\tau)x_{t-1} \tag{6}$$

where $\gamma_0(\tau) = \gamma_0 + \delta_0 F^{-1}(\tau)$ and $\gamma_1(\tau) = \gamma_1 + \delta_1 F^{-1}(\tau)$ both depend non-trivially on τ.

We next consider the data generating process for the predictor. Since the majority of predictors employed in practice are highly persistent, we model the regressor x_t as a near unit root process. In particular, following Cavanagh et. al. (1995), we assume that the predictor x_t is a finite order autoregressive process:

$$x_t = \alpha_0 + v_t, \qquad (1-\alpha_1 L)b(L)v_t = \varepsilon_{1t} \tag{7}$$

where $b(L) = \sum_{i=0}^{k} b_i L^i, b_0 = 1$. Assume that the roots of $b(L)$ are fixed and less than 1 in absolute value. Equation (7) can be rewritten as:

$$\Delta x_t = \beta_0 + \beta_1 x_{t-1} + \zeta(L)\Delta x_{t-1} + \varepsilon_{1t} \tag{8}$$

where $\beta_0 = (1-\alpha_1)b(1)\alpha_0$, $\beta_1 = (\alpha_1 - 1)b(1)$, and $\zeta(L) = -\sum_{j=1}^{k} L^{-1}[b_j - (1-\alpha_1)\sum_{i=j}^{k} b_i]L^j$. In this Augmented Dicky-Fuller representation, it is straightforward that x_t follows an $AR(k+1)$ process with largest root α_1. Of particular interest in this process is the local-to-unity specification, $\alpha_1 = 1 + \frac{c}{T}$, where $c \leq 0$, and T is the sample size. When $c = 0$, this generalizes to a unit root process. The mean reverting case $c < 0$ provides a useful large sample approximation for the case in which the largest root is very close to, but still less than one. A number of prior studies have used this framework to model predictors such as earnings and dividend price ratios which are highly persistent, but a priori stationary on economic grounds.

Assumption 1 *$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a martingale difference sequence with $E(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}) = \Sigma$, with typical element σ_{ij}.*

Assumption 2 *The cumulative distribution function of ε_{2t}, $F(\cdot)$, has a continuous density function $f(\cdot)$, which is positive on $\{\varepsilon_{2t} : 0 < F(\varepsilon_{2t}) < 1\}$.*

Assumption 3 *The conditional distribution function $F_{t-1}(\cdot) = Pr(\varepsilon_{2t} < \cdot|\mathcal{F}_{t-1})$ has derivative $f_{t-1}(\cdot)$ a.s., and $f_{t-1}(s_n)$ is uniformly integrable for any sequence $s_n \to F^{-1}(\tau)$, and $E[f_{t-1}^{\xi}(F^{-1}(\tau))] < \infty$ for some $\xi > 1$.*

Xiao (2009) makes the same assumptions on the distribution functions of the error term.

The standard quantile regression coefficient estimates are given by

$$(\hat{\gamma_0}(\tau), \hat{\gamma_1}(\tau)) = \arg \min_{(\gamma_0,\gamma_1)\in R^2} \sum_{t=1}^{T} \rho_\tau(y_t - \gamma_0 - \gamma_1 x_{t-1}). \tag{9}$$

$\rho_\tau(\cdot)$ is the asymmetric absolute deviation loss function defined by $\rho_\tau(u) = u\psi_\tau(u)$, where $\psi_\tau(u) = \tau - I(u < 0)$ (Koenker and Bassett 1978). For the median $\tau = 0.5$, $\rho_\tau(u) = 1/2|u|$ is used for Laplaces median regression function. In the proceeding subsection, we derive the asymptotic behavior of these estimates under the local-to-unity specification for x_t.

2.2 Theoretical results

Define $\delta = corr(\varepsilon_{1t}, \varepsilon_{2t})$. Let $B = (B_1, B_2)'$ denote a two-dimensional Brownian motion with covariance matrix $\bar{\Sigma}$, where $\bar{\sigma}_{11} = \bar{\sigma}_{22} = 1$ and $\bar{\sigma}_{12} = \bar{\sigma}_{21} = \delta$. We have

$$T^{-\frac{1}{2}} \sum_{t=1}^{[Tr]} \begin{bmatrix} b^{-1}(L)\varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \rightarrow_d \begin{bmatrix} \omega B_1(r) \\ \sigma_{22}^{\frac{1}{2}} B_2(r) \end{bmatrix} = BM(0, \Omega) \tag{10}$$

where $\Omega = \begin{bmatrix} \omega^2 & \omega\sigma_{22}^{\frac{1}{2}}\delta \\ \omega\sigma_{22}^{\frac{1}{2}}\delta & \sigma_{22} \end{bmatrix}$ with $\omega = \frac{\sigma_{11}^{\frac{1}{2}}}{b(1)}$.

Let $Z = (Z_1, Z_2)'$ be a two-dimensional Brownian motion such that $Z_1(r) = \omega B_1(r)$ and $Z_2(r) = \sigma_{22}^{\frac{1}{2}} B_2(r)$. Also, denote by J_c the diffusion process defined by $dJ_c(s) = cJ_c(s)ds + dB_1(s)$ with initial condition $J_c(0) = 0$. We distinguish demeaned variables by superscript μ. For example, $x_t^\mu = x_t - (T-1)^{-1}\sum_{t=2}^{T} x_{t-1}$ and $J_c^\mu(s) = J_c(s) - \int_0^1 J_c(r)dr$. In light of Phillips (1987), under $\alpha_1 = 1 + \frac{c}{T}$,

$$\begin{aligned} T^{-\frac{1}{2}} x_{[Tr]}^\mu &\rightarrow_d \omega J_c^\mu \\ T^{-2}\sum (x_{t-1}^\mu)^2 &\rightarrow_d \int (\omega J_c^\mu)^2 dr \\ T^{-1}\sum x_{t-1}^\mu b^{-1}(L)\varepsilon_{1t} &\rightarrow_d \int \omega J_c^\mu dZ_1 \end{aligned}$$

where $[\cdot]$ denotes the greatest lesser integer function.

The following standard result (see, for example, Stock 1991 and references within) gives the asymptotic of the t-statistic associated with the largest root of the autoregressive model for x_t in (8), when this root is modelled local-to-unity. The result is presented here without proof.

Proposition 1 *The asymptotic representation of the standard t-statistic used to test* $H_0 : \beta_1 = 0$

in (8) is given by

$$t_{\beta_1} \to_d \frac{c}{b(1)} \left[\int (J_c^\mu)^2 dr\right]^{\frac{1}{2}} + \frac{\int J_c^\mu dB_1}{\left[\int (J_c^\mu)^2 dr\right]^{\frac{1}{2}}}. \tag{11}$$

Defining

$$\varepsilon_{2t\tau} = \varepsilon_{2t} - F_{t-1}^{-1}(\tau) = y_t - \gamma_0(\tau) - \gamma_1(\tau) x_{t-1} \tag{12}$$

and $Q_{\varepsilon_{2t\tau}}(\tau|\mathcal{F}_{t-1})$ as the τth quantile of $\varepsilon_{2t\tau}$ conditional on $\mathcal{F}_{t-1}$, we may rewrite (4) as

$$y_t = \gamma_0(\tau) + \gamma_1(\tau) x_{t-1} + \varepsilon_{2t\tau} = \gamma(\tau)' z_{t-1} + \varepsilon_{2t\tau}, \tag{13}$$

where $Q_{\varepsilon_{2t\tau}}(\tau|\mathcal{F}_{t-1}) = 0$. Since $\psi_\tau(\varepsilon_{2t\tau}) = \tau - I(\varepsilon_{2t} < F^{-1}(\tau))$, we have $E\left[\psi_\tau(\varepsilon_{2t\tau})\,|\mathcal{F}_{t-1}\right] = 0$, and the variance of the indicator function $I(\cdot)$ is $\tau(1-\tau)$. The following preliminary convergence result, which is comparable to Assumption A of Xiao (2009) and stated without proof, will be needed for the subsequent analysis.

Lemma 1

$$T^{-\frac{1}{2}} \sum_{t=1}^{[Tr]} \begin{bmatrix} b^{-1}(L)\varepsilon_{1t} \\ \psi_\tau(\varepsilon_{2t\tau}) \end{bmatrix} \to_d \begin{bmatrix} Z_1(r) \\ Z_\psi(r) \end{bmatrix} = BM(0, \Omega_\tau) \tag{14}$$

where $\Omega_\tau = \begin{bmatrix} \omega^2 & \omega\sqrt{\tau(1-\tau)}\delta \\ \omega\sqrt{\tau(1-\tau)}\delta & \tau(1-\tau) \end{bmatrix}$.

It follows that $T^{-1}\sum x_{t-1}^\mu \psi_\tau(\varepsilon_{2t\tau}) \to_d \int \omega J_c^\mu dZ_\psi$.

The following result provides the limiting distribution of the predictive quantile regression estimator in (13).

Proposition 2

$$T(\hat{\gamma}_1(\tau) - \gamma_1(\tau)) \;\to_d\; \frac{1}{f(F^{-1}(\tau))} \left[\int (\omega J_c^\mu)^2\right]^{-1} \left[\int \omega J_c^\mu dZ_\psi\right] \tag{15}$$

$$D_T(\hat{\gamma}(\tau) - \gamma(\tau)) \;\to_d\; \frac{1}{f(F^{-1}(\tau))} \left[\int \bar{J}_c \bar{J}_c'\right]^{-1} \left[\int \bar{J}_c dZ_\psi\right] \tag{16}$$

where $D_T = \begin{bmatrix} T^{\frac{1}{2}} & 0 \\ 0 & T \end{bmatrix}$, *and* $\bar{J}_c = (1, \omega J_c)'$.

The distribution is nonstandard. In the case when $c = 0$, it specializes the result of the quantile cointegrating regression (Xiao, 2009, Theorem 1) to the case of predictive regression. The extension to $c < 0$ is new to the best of our knowledge. As in the case of cointegrating regression, some further insight into the bias can be gained from the projection of Z_ψ on to B_1, which yields the orthogonal decomposition (see Phillips, 1989, pp. 30-31) $Z_\psi = \sqrt{\tau(1-\tau)}\left[\omega^{-1}\delta Z_1(r) + \sqrt{1-\delta^2} Z_{\psi\cdot 1}\right] = \sqrt{\tau(1-\tau)}\left[\delta B_1(r) + \sqrt{1-\delta^2} Z_{\psi\cdot 1}\right]$, where $Z_{\psi\cdot 1} = BM(1)$

and is independent of Z_1. Using this decomposition to substitute for dZ_ψ, we may re-express the numerator of (2) as

$$\frac{1}{f(F^{-1}(\tau))}\left[\delta\sqrt{\tau(1-\tau)}\left(\int(\omega^2 J_c^\mu)^2\right)^{-1}\int \omega J_c^\mu dB_1 + \omega_{\psi\cdot 1}\left(\int(\omega^2 J_c^\mu)^2\right)^{-1}\int \omega J_c^\mu dZ_{\psi\cdot 1}\right]. \tag{17}$$

where we define $\omega_{\psi\cdot 1}^2 = (1-\delta^2)\tau(1-\tau)$. The stochastic integral inside the brackets is the local-to-unity generalization of the (de-meaned) Dickey-Fuller distribution and contributes a downward (upward) second order bias to the estimate of $\hat{\gamma}_1(\tau)$ for $\delta > 0$ ($\delta < 0$). It is analogous to what Phillips and Hansen (1990) refer to as the serial correlation term in the context of cointegration. It is evident that the extent of the bias of the bias depends on both δ and on c, as is also true in the linear predictive regression. The second term is in brackets is mixed normal, and normal conditional on $\mathcal{F}_1 = \sigma\left(B_1(r), 0 \leq r \leq 1\right)$, due to the independence of J_c and $z_{\psi\cdot 1}$. As in the case of predictive regression, there is no endogeneity term due to the assumption that $\varepsilon_{2,t}$ is a martingale difference sequence. The distribution of the estimator depends on τ both directly and through $Z_{\psi\cdot 1}$, which is itself a function of τ.

The standard error of $\hat{\gamma}_1(\tau)$ is given by $\sqrt{\frac{1}{f(\widehat{F^{-1}(\tau)})^2}[\sum(x_{t-1}^\mu)^2]^{-1}}$, where $f(\widehat{F^{-1}(\tau)})$ is a consistent estimator of $f(F^{-1}(\tau))$. The following proposition provides the null limiting distribution of the standard t-statistic in the predictive quantile regression.

Proposition 3 *The asymptotic representation of the t-statistic to test $H_0 : \gamma_1(\tau) = 0$*[2] *for a given τ is*

$$t_{\gamma_1}(\tau) \quad \to_d \quad \frac{\int J_c^\mu dZ_\psi}{[\int(J_c^\mu)^2]^{\frac{1}{2}}} \equiv \sqrt{\tau(1-\tau)}\left[\delta\frac{\int J_c^\mu dB_1}{[\int(J_c^\mu)^2]^{\frac{1}{2}}} + \sqrt{1-\delta^2}z\right] \tag{18}$$

where z is a standard normal random variable that is independent of (B_1, J_c).

The asymptotic distribution of $t_{\gamma_1}(\tau)$ depends on (c, δ, τ). The dependence on c and δ is similar to that in the linear regression case, while dependence on τ is new.

3 Inference

In equation (8), if $\alpha_1 < 1$ and is fixed, the unit root test will reject with probability of one asymptotically. Since x_t is stationary in this case, the quantile regressions of equation (1) are well-behaved in large samples. It would be sensible to test $H_0 : \gamma_1(\tau) = 0$ using standard normal critical values. If $\alpha_1 = 1$, which corresponds to $c = 0$, then x_t has a unit root. When $c < 0$ and α_1 is large, x_t has a near unit root. In the stock return predictability example, many

[2]Note that the null hypothesis can be generalized to $\gamma_1(\tau) = \gamma_{1,0}$.

predictors, such as the dividend price ratio, have large autoregressive roots that are very close to one. Hence, it is appropriate to model x_t as a near unit root process. As Proposition 3 shows, in this case the limiting distribution of $t_{\gamma_1}(\tau)$ is nonstandard and dependent on the nuisance parameter c. Moreover, the asymptotic distribution of $t_{\gamma_1}(\tau)$ also depends on δ. If the two error terms in equation (8) and (1) are uncorrelated, $t_{\gamma_1}(\tau)$ will have a standard normal asymptotic distribution. However, when the two error terms are correlated, the limiting distribution of the quantile regression coefficient is no longer standard normal, which causes size distortion of the predictability test. Thus, standard normal critical values are not reliable for the test of interest, since they tend to over-reject the null. This is a problem in practice, because financial data such as prices and dividends do not satisfy strict exogeneity. The over-rejection is especially severe when the residual cross-correlation is large.

Therefore, this paper proposes a Bonferroni bounds method in the spirit of Cavanagh et al. (1995), which results in asymptotically valid test. The procedure has two steps. Because the local-to-unity parameter c is a nuisance parameter that cannot be consistently estimated, we first derive a $100(1-\eta_1)\%$ confidence interval for c, and denote it by $CI_c(\eta_1)$. This first-stage confidence interval is computed by inverting a unit root test statistic as illustrated in Stock (1991). In the second stage, based on the previously derived asymptotic distribution, we compute the critical value for $t_{\gamma_1}(\tau)$ of size η_2 at each point of c in the first-stage confidence interval. Then, the Bonferroni bounds $[C_l(\eta_1,\eta_2), C_u(\eta_1,\eta_2)]$ gives the asymptotically valid critical values for $t_{\gamma_1}(\tau)$, where

$$\begin{aligned} \text{lower bound:} \qquad C_l(\eta_1,\eta_2) &= \min_{c\in CI_c(\eta_1)} C_{t_{\gamma_1},c,\frac{\eta_2}{2}} \\ \text{upper bound:} \qquad C_u(\eta_1,\eta_2) &= \max_{c\in CI_c(\eta_1)} C_{t_{\gamma_1},c,1-\frac{\eta_2}{2}}, \end{aligned}$$

and $C_{t_{\gamma_1},c,\frac{\eta_2}{2}}$ is the $100\frac{\eta_2}{2}\%$ quantile of the limiting distribution of $t_{\gamma_1}(\tau)$ for a given δ. Tests based on the Bonferroni bounds are conservative with size less than or equal to η, for $\eta = \eta_1 + \eta_2$.

Although the Bonferroni interval is conservative, the size of the test can be adjusted by selecting appropriate (η_1, η_2). More specifically, one can set $\eta_2 = \eta$, and choose η_1 such that the asymptotic size equals to η. In our Monte Carlo experiments, we consider $\eta = 10\%$ as the size of the test, and choose η_1 under $T = 1000$ such that the size is close to the benchmark case where c is known.

4 Simulation Study

Our simulation study focuses on the demeaned case in which a regression contains a constant term (without a time trend term). The results are based on asymptotic results from 2000 Monte Carlo replications with sample size T for each set of parameters (c, δ, τ), where c ranges from 0 to $-T$, δ ranges from 0 to 0.95, and τ ranges from 0.05 to 0.95. In each replication, x_t is generated according to equation (8). We set $b_j = 0$ for $j > 1$. Under the local-to-unity specification, we have $x_t = (1 + \frac{c}{T})x_{t-1} + \varepsilon_{1t}$. y_t is generated by equation (1) under the null $H_0 : \gamma_1 = 0$. We set $\gamma_0 = 0$ so that $y_t = \varepsilon_{2t}$. ε_{1t} and ε_{2t} are both $N(0, 1)$ with correlation coefficient δ. The level of the predictability test is 10%.

As discussed in previous sections, conventional t-tests from the predictive quantile regression over-reject the null, since with nearly integrated predictors the asymptotic distribution of t-statistics is nonstandard. As a preliminary experiment, we investigate the size distortion problem associated with the t-test based on standard normal critical values, ± 1.65. As Table 1 shows, when $\delta = 0, 0.25$, the test is conservative for $\tau = 0.2$ to 0.8 with large sample size such as $T = 1000$. From Table 2, even with small sample size as $T = 200$, the rejection rates for $\tau = 0.2$ to 0.8 are still close to 10%. That is, when there is no or small correlation between ε_{1t} and ε_{2t}, there is no or very little over-rejection for inner quantiles. As the correlation coefficient increases, over-rejection becomes more severe. For $\delta = 0.50, 0.75, 0.95$, the null is over-rejected for all quantiles, even with $T = 1000$. Particularly, when the correlation is as large as 0.95, rejection rates can be over 25%. Similar over-rejection problem is illustrated for linear predictive regression in Mankiw and Shapiro (1986) and Cavanagh et al. (1995): with $T = 500, 200, 100$, conventional t-tests can lead to substantial size distortion, especially for large correlation coefficients.

To control the size distortion, we employ asymptotically valid critical values according to the asymptotic representation stated in Section 2 for a number of parameter sets. Since the analytic critical values cannot be obtained directly from the distribution, we calculate the simulated critical values under a large sample size, and store them in computerized lookup tables. The simulation is based on 49999 Monte Carlo replications with $T = 1000$. Again, x_t is simulated under the local-to-unity specification $\alpha_1 = 1 + \frac{c}{T}$, and y_t is simulated under the null of no predictability. Since the limiting distribution of $\gamma_1(\tau)$ depends on c, δ and τ, 42 percentiles from the empirical distribution of $t_{\gamma_1}(\tau)$ are calculated on a grid of 281 values of c, from -100 to 9.5, with the grid most dense on $[-15, 5]$, for all δ and τ.

To investigate the accuracy of the simulated critical values, we calculate the rejection frequencies under a set of known c with $T = 1000, 200$ using the critical values from the lookup table. As shown in Table 3, when $T = 1000$, the rejection frequencies are close to 10% by construction, and we take this case as the benchmark. With $T = 200$, Table 4 shows that the over-rejection problem is lessened. Although for the outer quantiles moderate size distortion still exists with small sample size, the rejection rates are significantly reduced compared to those from the standard t-tests. Thus, for example, if we consider the predictability of the tails in the return distribution, it is useful to perform size correction. For the inner quantiles, when δ is small, the rejection rates are sometimes slightly larger than those from the conventional t-tests; when δ is large, the rejection rates are smaller than those from the conventional t-tests. Hence, size correction is not beneficial when there is none or small correlation between the error terms, but it is necessary when the correlation is large.

Since c is a nuisance parameter, which is generally unknown in practice, we next follow the two-step procedure described in the previous section. First, we run the OLS regression according to the ADF representation as in equation (8) to obtain $\hat{t}_{\beta_1}$. Then we construct the $100(1-\eta_1)\%$ confidence interval for c by inverting the t-statistic (Stock 1991). Next, we run the quantile regression (13) and obtain $\hat{t}_{\gamma_1}(\tau)$ for each τ. Then we compute the second stage critical values, that is, the $100(1-\eta_2)\%$ Bonferroni bounds for $\hat{t}_{\gamma_1}(\tau)$ over the confidence region of c from stage one. The rejection frequencies are calculated according to the following rejection rule: if $\hat{t}_{\gamma_1}(\tau)$ lies outside its Bonferroni interval, we reject the null; otherwise, we fail to reject.

The size of the test is adjusted by choosing η_1 and η_2 according to the previous section so that the test will not be too conservative. Given $\eta_2 = 10\%$, η_1 is selected with $T = 1000$ such that the rejection rates are close to the benchmark case. For $\delta = 0.25, 0.50, 0.75, 0.95$, this leads us to select $\eta_1 = 0.7, 0.5, 0.4, 0.3$, respectively. Table 5 to 6 summarize the results from performing the size-adjusted Bonferroni method with large and small sample sizes. More precisely, under the level of $\eta_2 = 10\%$, we plug in η_1 calculated above for each of value of δ, so that the size from Bonferroni correction is not too conservative. With large sample size T=1000, the rejection rates of the null, $\gamma_1(\tau) = 0$, are between 0.0800 to 0.1185, which are close to those from the benchmark case due to size adjustment. With smaller sample size, such as $T = 200$, there is negligible size distortion for the inner quantiles, and moderate over-rejection for the outer quantiles, which is a finite sample problem[3]. Compared to the experiments with known c, when performing the Bonferroni procedure, the rejection rates in Table 6 are generally no larger than those from Table 4. Therefore, although using the Bonferroni bounds may not fully correct the size distortion for small samples, it is asymptotically valid and can effectively solve the over-rejection problem with large samples. Note that when $\delta = 0$, there is no need for size adjustment. Instead, simulated i.i.d. critical values are used in the tests. Similarly to producing the lookup table, we compute the simulated i.i.d. critical values based on 49999 Monte Carlo replications with $T = 1000$, and $c = -T$, and store them in a lookup table. Consequently, as the first panel in Table 5 illustrates, the rejection frequencies range from 0.0845 to 0.1150 with $T = 1000$. Again, for smaller sample sizes, the finite sample problem is more significant.

The Bonferroni technique is also evaluated by simulations assuming the residual correlation is unknown, with the data generated under true $\delta = 0.5$ and 0.9. δ is estimated by $\hat{\delta} = corr(\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t})$, where $\hat{\varepsilon}_{1t}$ is the residual from regressing x_t on a constant and its own lag, and $\hat{\varepsilon}_{2t}$ is the residual from regressing y_t on a constant and x_{t-1}. The Bonferroni critical values are calculated by interpolation with respect to $\hat{\delta}$. The results are reported in Table 7 and 8, which shows the rejection frequencies are close to those from cases with known δ. For $T = 1000$, the rejection frequencies are close to 10%, the level of the test. The test is conservative for c of

[3]When $c = 0$, the normal approximation is poor in extreme quantiles. See, for example, Chernozhukov (2005) for the case in which x_t is $I(0)$.

smaller absolute values. For $T = 200$, the rejection frequencies are close to 10% for the central part of the distribution. There is some moderate size distortion for the outer quantiles. The over-rejection problem is more severe for the extreme quantiles.

5 Empirical Study

This paper applies the Bonferroni technique to test the predictability at different points in the stock return distribution. The univariate quantile regressions of the return on 16 pre-determined predictors are analyzed by Cenesizoglu and Timmermann (2008) using the data from Goyal and Welch (2008). This paper will use the same data set that comprises monthly observations over the period from February 1871 to December 2005, with the shortest series from May 1937 to December 2002. As in Cenesizoglu and Timmermann (2008), the stock return is measured by the S&P500 index including dividends. The 16 predictor variables, including dividend price ratio, dividend yield and earnings price ratio, are listed in Table 9.

To examine the predictor variables, first we run OLS regression of each predictor on its own lags as in equation (8). As shown in Table 9, the lag length is selected by the Bayes information criterion (BIC) with a maximum of twelve lags. In addition, the confidence interval for the local-to-unity parameter c of each predictor is calculated based on Stock (1991). From Table 9, the largest autoregressive root α_1 is close to but smaller than one in many cases. Some predictors such as dividend price ratio, dividend yield and T-bill rate have upper bounds of the 95% interval slightly larger than one. Default return spread, long term rate of return, stock variance, and inflation are close to stationary. Besides, the correlation coefficient δ is estimated by $\hat{\delta} = corr(\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t})$, where $\hat{\varepsilon}_{1t}$ is the residual from the ADF regression of equation (8), and $\hat{\varepsilon}_{2t}$ is the residual from regressing the stock return on the predictor as in equation (1). As Table 9 shows, $\hat{\delta}$ is negative for most predictors. Moreover, $\hat{\varepsilon}_{1t}$ is strongly correlated with $\hat{\varepsilon}_{2t}$ for dividend price ratio, earnings price ratio, smoothed earnings price ratio and book to market ratio. In contrast, for default return spread, long term rate of return and inflation, the cross-correlations are positive and small.

To test the predictability of different parts in the return distribution, we then run quantile regression of the return on each predictor for eleven quantiles. The quantile coefficients are estimated for each quantile. The magnitude of the coefficient estimates are the same as those from Cenesizoglu and Timmermann (2008), where they have discussed the asymmetric effects of the predictive variables on different positions in the return distribution. For our predictive tests, the relevant Bonferroni critical values are interpolated from the lookup tables with respect to $\hat{\delta}$. The test results from using these Bonferroni bounds are reported in Table 10.

The first row of each panel in Table 10 contains the quantile regression slope coefficient estimates for eleven quantiles in the stock return distribution. The second row shows the corresponding t-statistics. At 10% and 5%[4] significance level, we find evidence of predictability at various points in the stock return distribution for most predictor variables. Some variables have significant effects on the outer quantiles of the return. For example, with default yield spread we find evidence of predictability in the tails and shoulders of the return distribution, but no evidence for the inner quantiles such as $\tau = 0.4, 0.5, 0.6, 0.7$ at 5% significance level. Interestingly, the quantile coefficient estimates are negative and significant for the left tail, but positive and significant for the right tail. Similar evidence can be found with stock variance, which also has predictive content in the tails and shoulders, but not the center of the return distribution. The asymmetric effects of stock variance on the return distribution is significant at least at 10% level. For smoothed earnings price ratio and book to market ratio, there is only evidence of predictability in the two tails of the return distribution. For default return spread, the null is only rejected at $\tau = 0.05$, and for net equity expansion, in the left tail where $\tau = 0.05, 0.1, 0.2$. Other variables have more effects on the central part of the distribution. For the long term yield we reject the null in the center and right shoulder. For T-bill rate we reject when $\tau \geq 0.4$, and for inflation we reject when $\tau \geq 0.3$. At 5% level, among all these predictors, inflation can predict most quantiles of the return distribution. T-bill rate, long term yield, default yield spread and stock variance also have considerable predictive content compared with the other variables. On the other hand, at 5% level, dividend price ratio, dividend yield and earnings price ratio fail to predict any quantile in the return distribution. However, at 10% level, the proposed test shows that dividend price ratio has predictive content in the median.

The p-values in the third row of each panel in Table 10 are from the standard t-tests. In many cases, our critical values give the same inferences as the standard normal critical values do. However, there are some discrepancies between the proposed test and conventional t-test. For example, at 10% significance level, dividend price ratio, earnings price ratio, smoothed earnings price ratio, and book to market ratio, due to the large residual cross-correlations, conventional t-tests over-reject for some quantiles. This corresponds to our simulation results that using the Bonferroni procedure with the simulated critical values is beneficial when there is large residual cross-correlation. At 5% significance level, besides smoothed earnings price ratio, and book to market ratio, for cross sectional premium, stock variance, dividend earnings ratio, standard t-tests also over-reject in some parts of the return distribution. On the contrary, our tests over-reject for long term yield and long term rate of return, where $\hat{\delta}$ is small, which also corresponds to the simulation results.

Furthermore, our empirical study has some similarities to that conducted by Cenesizoglu and Timmermann (2008). In their paper, the Bonferroni p-values are used as a "summary measure" in a joint test across all quantiles considered under the null that a given predictor does not predict any of the quantiles in the return distribution. According to their findings,

[4]When the level of the predictability test, η_2, is 5%, for $\delta = 0.25, 0.50, 0.75, 0.95$, we choose $\eta_1 = 0.40, 0.30, 0.25, 0.20$, respectively.

eight of the predictors considered are significant at 5% level, while the others are insignificant. In particular, the p-values are one for dividend price ratio and dividend yield, which indicates no predictability of any points in the return distribution using either one of the two variables. Unlike their approach, the Bonferroni methodology we proposed is used for a different purpose. Instead of testing the significance of each predictor across all quantiles jointly, our paper adopts a Bonferroni method in the local-to-unity context and tests whether the quantile regression coefficient is zero for each individual τ. Although, at 10% level, we find evidence of predictability in the median of the return distribution for dividend price ratio, at 5% level, the proposed test method confirms the insignificance of dividend price ratio and dividend yield, which is consistent with results from Cenesizoglu and Timmermann (2008). In addition, as mentioned before, our empirical results also show that earnings price ratio has no predictive content of the return distribution.

6 Conclusion

This paper develops inference in predictive quantile regressions with a nearly nonstationary regressor. The predictive setting of interest is a quantile regression of a dependent variable y_t on a lagged regressor x_{t-1} which is possibly strongly autocorrelated and not strictly exogenous. We derive the limiting distributions of the quantile regression coefficient and its corresponding t-statistic under the local-to-unity specification that $\alpha = 1 + \frac{c}{T}$. Both of these distributions depend on the nuisance parameter c and the residual cross-correlation coefficient δ. According to the asymptotic representation of the test statistic, we create computerized lookup tables consisting of simulated critical values for a number of (c, δ) pairs. We also present an asymptotically valid Bonferroni bounds methodology to test whether x_{t-1} has predictive content of the τth quantile of the conditional distribution of y_t using the critical values from our look-up tables. The simulation results from 2000 Monte Carlo replications show that it is worthwhile to adopt the Bonferroni procedure to correct the size distortion especially when there is considerable residual cross-correlation. Furthermore, despite the substantial literature on the predictability of the mean of stock return, few research has studied the predictability of the entire return distribution. Thus, we apply the asymptotic theory and Bonferroni methodology to test the predictability at different points in the return distribution using the 16 pre-determined predictor variables that have been investigated by some previous studies. From the test results at 5% significance level, except for dividend price ratio, dividend yield and earnings price ratio, all these variables have some predictive component of certain quantiles of the return distribution.

References

Amihud, Yakov and Clifford M. Hurvich (2004). Predictive regressions: A reduced-bias estimation method. *Journal of Financial and Quantitative Analysis* 39, 813-841.

Amihud, Yakov, Clifford M. Hurvich and Yi Wang (2004). Hypothesis testing in predictive regressions. Mimeographed, New York University.

Campbell, B. and J.-M. Dufour (1995). Exact nonparametric orthogonality and random walk tests. *The Review of Economics and Statistics* 77, 1-16.

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

Campbell, John Y. and Robert J. Shiller (1988a). The dividend price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195-227.

Campbell, John Y. and Robert J. Shiller (1988b). Stock prices, earnings, and expected dividends. *Journal of Finance* 43, 661-76.

Campbell, J. Y. and S. B. Thompson (forthcoming). Predicting the equity premium out of sample: Can anything beat the historical average?. *Review of Financial Studies*.

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

Cavanagh, Christopher, Graham Elliott and James H. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory* 11, 1131-1147.

Cenesizoglu, Tolga and Allan Timmermann (2008). Is the distribution of stock returns predictable? Working paper, HEC Montreal and University of California at San Diego.

Chan, N. H. and C. Z. Wei (1987). Asymptotic inference for nearly nonstationary ar(1) processes. *Annals of Statistics* 15(3), 1050-1063.

Chan, Ngai Hang (1988). The parameter inference for nearly nonstationary time series. *Journal of the American Statistical Association* 83(403), 857-862.

Chen, Xiaohong, Roger Koenker, and Zhijie Xiao (2009). Copula-based quantile autoregression. *Econometrics Journal* 12(1), 50-67.

Chernozhukov, Victor (2005). Extremal quantile regression. *The Annals of Statistics* 33, 806-839.

Chernozhukov, Victor (2010). Inference for extremal conditional quantile models. *Review of Economic Studies* forthcoming.

Chernozhukov, Victor and Alexandre Belloni (2010). l_1-Penalized quantile regression in high dimensional sparse models. *Annals of Statistics* forthcoming.

Chernozhukov, Victor and Songzi Du (2006). Extremal quantiles and value-at-risk. MIT Department of Economics Working Paper No. 07-01.

Chernozhukov, Victor, Christian Hansen and Michael Jansson (2009). Finite sample inference for quantile regression models. *Journal of Econometrics* 152(2), 93-103.

Chuang, Chia-Chang, Chung-Ming Kuan and Hsin-yi Lin (2009). Causality in quantiles and dynamic stock return-volume relations. *Journal of Banking & Finance* 33, 1351-1360.

Fama, Eugene and Kenneth French (1988). Permanent and temporary components of stock prices. *Journal of Political Economy* 96, 246-273.

Frlich, Markus and Blaise Melly (2008). Quantile treatment effects in the regression discontinuity design. IZA Discussion Papers No. 3638.

Goetzmann, W. and P. Jorion (1993). Testing the predictive power of dividend yields. *Journal of Finance* 48(2), 663-670.

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

Hjalmarsson, Erik (2007). Fully modified estimation with nearly integrated regressors. *Finance Research Letters* 4, 92–94.

Hodrick, R. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5, 357-368.

Huber, Martin (2010). Testing for covariate balance using nonparametric quantile regression and resampling methods. University of St. Gallen Department of Economics working paper series 2010 No. 2010-18.

Jansson, Michael and Marcelo J. Moreira (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica* 74, 681-714.

Jun, Sung Jae, Joris Pinkse and Haiqing Xu (2009). Tighter bounds in triangular systems. Preprint.

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

Koenker, R. (2008). Censored Quantile Regression Redux. *Journal of Statistical Software* 27(6).

Koenker, R. (2010). Additive models for quantile regression: model selection and confidence bandaids. Preprint.

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

Koenker, Roger and and Zhijie Xiao (2004). Unit root quantile autoregression inference. *Journal of the American Statistical Association* 99, 775-787.

Koenker, Roger and and Zhijie Xiao (2006). Quantile autoregression. *Journal of the American Statistical Association* 101, 980-990.

Lanne, Markku (2002). Testing the predictability of stock returns. *The Review of Economics and Statistics* 84(3), 407-415.

Lewellen, Jonathan (2004). Predicting returns with financial ratios. *Journal of Financial Economics* 74(2), 209-235.

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

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

Nabeya, Seiji and Bent E. Srensen (1994). Asymptotic distributions of the least-squares estimators and test statistics in the near unit root model with non-zero initial value and local drift and trend. *Econometric Theory* 10, 937-966.

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

Phillips, P.C.B. (1987). Toward a unified asymptotic theory for autoregression. *Biometrika* 74, 535-547.

Phillips, Peter P.C.B. 1989. Partially identified econometric models. *Econometric Theory* 5, 181-240.

Phillips, Peter C. B. and Bruce E. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Shiller, Robert (1984). Stock prices and social dynamics. Brookings papers on Economic Activity 2, 457-498.

Stambaugh, R.F. (1986). Bias in regressions with lagged stochastic regressors. Center for Research in Security Prices Working Paper 156, University of Chicago.

Stambaugh, Robert F. (1999). Predictive regressions. *Journal of Financial Economics* 54, 375-421.

Stock, J.H. (1991). Confidence intervals for the largest autoregressive root in U.S. economic time series. *Journal of Monetary Economics* 28 (3), 435-460.

Wolf, M. (2000). Stock returns and dividend yields revisited: A new way to look at an old problem. *Journal of Business and Economic Statistics* 18(1), 18-30.

Wright, Jonathan (2000). Confidence sets for cointegrating coefficients based on stationarity tests. *Journal of Business and Economic Statistics* 18(2), 211-222.

Xiao, Zhijie (2009). Quantile cointegrating regression. *Journal of Econometrics* 150, 248-260.

Xiao, Zhijie and Roger Koenker (2009). Conditional quantile estimation for GARCH models.

Boston College Working Papers in Economics Number 725.

A Proofs

A.1 Proof of Proposition 2

This proof follows the procedure in Xiao (2009). Denote $\hat{u} = D_T(\hat{\gamma}(\tau) - \gamma(\tau))$.

$$\begin{aligned} \rho_\tau(y_t - \hat{\gamma}(\tau)' z_{t-1}) &= \rho_\tau(\gamma(\tau)' z_{t-1} + \varepsilon_{2t\tau} - \hat{\gamma}(\tau)' z_{t-1}) \\ &= \rho_\tau(\varepsilon_{2t\tau} - (\hat{\gamma}(\tau)' - \gamma(\tau)') z_{t-1}) \\ &= \rho_\tau(\varepsilon_{2t\tau} - (D_T^{-1}\hat{u})' z_{t-1}) \end{aligned}$$

Thus, the minimization problem can be rewritten as $\min_u G_T(u)$, where the objective function $G_T(u) = \sum_{t=1}^T [\rho_\tau(\varepsilon_{2t\tau} - (D_T^{-1}u)' z_{t-1}) - \rho_\tau(\varepsilon_{2t\tau})]$. From Xiao (2009) for any $u \neq 0$,

$$\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + (u - v) I(0 > u > v) - I(0 < u < v).$$

Then we have

$$\begin{aligned} G_T(u) &= -\sum_{t=1}^T (D_T^{-1}u)' z_{t-1} \psi_\tau(\varepsilon_{2t\tau}) \\ &\quad + \sum_{t=1}^T (\varepsilon_{2t\tau} - (D_T^{-1}u)' z_{t-1})[I(0 > \varepsilon_{2t\tau} > (D_T^{-1}u)' z_{t-1}) - I(0 < \varepsilon_{2t\tau} < (D_T^{-1}u)' z_{t-1})] \\ &= -\sum_{t=1}^T (D_T^{-1}u)' z_{t-1} \psi_\tau(\varepsilon_{2t\tau}) \\ &\quad + \sum_{t=1}^T ((D_T^{-1}u)' z_{t-1} - \varepsilon_{2t\tau})[I(0 < \varepsilon_{2t\tau} < (D_T^{-1}u)' z_{t-1}) - I(0 > \varepsilon_{2t\tau} > (D_T^{-1}u)' z_{t-1})] \end{aligned}$$

For the first term, since $T^{-1} \sum x_{t-1} \psi_\tau(\varepsilon_{2t\tau}) \to_d \int \omega J_c dZ_\psi$,

$$D_T^{-1} \sum_{t=1}^T z_{t-1} \psi_\tau(\varepsilon_{2t\tau}) = \begin{bmatrix} T^{-\frac{1}{2}} \sum_{t=1}^T \psi_\tau(\varepsilon_{2t\tau}) \\ T^{-1} \sum_{t=1}^T x_{t-1} \psi_\tau(\varepsilon_{2t\tau}) \end{bmatrix} \to_d \begin{bmatrix} \int dZ_\psi \\ \int \omega J_c dZ_\psi \end{bmatrix}.$$

For the second term, we start with $W_T(u) = \sum_{t=1}^T (u' D_T^{-1} z_{t-1} - \varepsilon_{2t\tau}) I(0 < \varepsilon_{2t\tau} < u' D_T^{-1} z_{t-1})$. Denote $\xi_t(u) = (u' D_T^{-1} z_{t-1} - \varepsilon_{2t\tau}) I(0 < \varepsilon_{2t\tau} < u' D_T^{-1} z_{t-1})$. Define $W_{Tm}(u) = \sum_{t=1}^T \xi_{tm}(u)$, where $\xi_{tm}(u) = (u' D_T^{-1} z_{t-1} - \varepsilon_{2t\tau}) I(0 < \varepsilon_{2t\tau} < u' D_T^{-1} z_{t-1}) I(u' D_T^{-1} z_{t-1} \leq m)$ is the truncation of $u' D_T^{-1} z_{t-1}$ at some finite number $m > 0$. Denote $\bar{\xi}_{tm}(u) = E[\xi_{tm}(u) | \mathcal{F}_{t-2}]$ and $\bar{W}_{Tm}(u) =$

$\sum_{t=1}^{T} \bar{\xi}_{tm}(u)$, where $\mathcal{F}_{t-2}$ is information set up to time $t-1$, and $z_{t-1} \in \mathcal{F}_{t-2}$. Thus, we have

$$\begin{aligned}
\bar{W}_{Tm}(u) &= \sum_{t=1}^{T} E[(u'D_T^{-1}z_{t-1} - \varepsilon_{2t\tau})I(0 < \varepsilon_{2t\tau} < u'D_T^{-1}z_{t-1})I(u'D_T^{-1}z_{t-1} \le m)|\mathcal{F}_{t-2}] \\
&= \sum_{t=1}^{T} E[(u'D_T^{-1}z_{t-1} + F^{-1}(\tau) - \varepsilon_{2t})I(F^{-1}(\tau) < \varepsilon_{2t} < u'D_T^{-1}z_{t-1} + F^{-1}(\tau))I(u'D_T^{-1}z_{t-1} \le m)|\mathcal{F}_{t-2}] \\
&= \sum_{t=1}^{T} \int_{F^{-1}(\tau)}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} [\int_{r}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} ds]f_{t-1}(r)dr \\
&= \sum_{t=1}^{T} \int_{F^{-1}(\tau)}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} [\int_{F^{-1}}^{s} f_{t-1}(r)dr]ds \\
&= \sum_{t=1}^{T} \int_{F^{-1}(\tau)}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} [s - F^{-1}(\tau)][\frac{F_{t-1}(s) - F_{t-1}(F^{-1}(\tau))}{s - F^{-1}(\tau)}]ds \\
&= \sum_{t=1}^{T} \int_{F^{-1}(\tau)}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} [s - F^{-1}(\tau)]f_{t-1}(F^{-1}(\tau))ds + op(1) \\
&= \sum_{t=1}^{T} f_{t-1}(F^{-1}(\tau)) \left[\frac{[s - F^{-1}(\tau)]^2}{2} |_{F^{-1}(\tau)}^{[u'D_T^{-1}z_{t-1}+F^{-1}(\tau)]I(u'D_T^{-1}z_{t-1}\le m)} \right] + op(1) \\
&= \frac{1}{2} \sum_{t=1}^{T} f_{t-1}(F^{-1}(\tau))[u'D_T^{-1}z_{t-1}]^2 I(u'D_T^{-1}z_{t-1} \le m) + op(1) \\
&= \frac{1}{2T} \sum_{t=1}^{T} f_{t-1}(F^{-1}(\tau))u'[\sqrt{T}D_T^{-1}z_{t-1}z_{t-1}'D_T^{-1}\sqrt{T}]uI(u'D_T^{-1}z_{t-1} \le m) + op(1)
\end{aligned}$$

By Assumption 3 and stationarity of $\{f_{t-1}(F^{-1}(\tau))\}$, for some $\epsilon > 0$

$$\sup_{0\le r\le 1} |\frac{1}{T^{1-\epsilon}} \sum_{t=1}^{[Tr]}[f_{t-1}(F^{-1}(\tau)) - f(F^{-1}(\tau))]| \to_d 0.$$

Therefore, we have

$$\begin{aligned}
\bar{W}_{Tm}(u) &\to_d \frac{1}{2} f(F^{-1}(\tau))u'[\int B_z B_z']I(0 < u'B_z \le m)u \\
&:= \eta_m
\end{aligned}$$

where $B_z = (1, \omega J_c)'$.
Since $(u'D_T^{-1}z_{t-1})I(0 \leq u'D_T^{-1}z_{t-1} \leq m) \to_p 0$ uniformly in t,

$$\begin{aligned}\sum_{t=1}^{T} E[\xi_{tm}(u)^2|\mathcal{F}_{t-2}] &\leq \max\{(u'D_T^{-1}z_{t-1})I(0 \leq u'D_T^{-1}z_{t-1} \leq m)\} \times \sum \bar{\xi}_{tm}(u) \\ &\to_p 0.\end{aligned}$$

Thus, $\sum_{t=1}^{T}[\xi_{tm}(u) - \bar{\xi}_{tm}(u)] \to_p 0$, where $\xi_{tm}(u) - \bar{\xi}_{tm}(u)$ is a MDS.
By the Asymptotic Equivalence lemma, the limiting distribution of $\sum_{t=1}^{T} \xi_{tm}(u)$ is the same as that of $\sum_{t=1}^{T} \bar{\xi}_{tm}(u)$, i.e. $W_{Tm}(u) \to_d \eta_m$. As $m \to \infty$, $\eta_m \to_d \frac{1}{2}f(F^{-1}(\tau))u'[\int B_z B_z']uI(u'B_z > 0) = \eta$.
We have $\lim_{m\to\infty} \limsup_{n\to\infty} Pr[|W_T(u) - W_{Tm}(u)| \geq \epsilon] = 0$, since

$$\begin{aligned}Pr[|W_T(u) - W_{Tm}(u)| \geq 0] &= Pr[\sum_{t=1}^{T}(u'D_T^{-1}z_{t-1} - \varepsilon_{2t\tau})I(0 < \varepsilon_{2t\tau} < u'D_T^{-1}z_{t-1})I(u'D_T^{-1}z_{t-1} > m)] \\ &\leq Pr[\bigcup_t \{u'D_T^{-1}z_{t-1} > m\}] \\ &= Pr[\max_t \{u'D_T^{-1}z_{t-1}\} > m] \\ \lim_{m\to\infty} Pr[\sup_{0\leq r\leq 1} u'B_z(r) > m] &= 0.\end{aligned}$$

Also, $W_T \to_d \eta$, i.e.

$$\sum_{t=1}^{T}(u'D_T^{-1}z_{t-1} - \varepsilon_{2t\tau})I(0 < \varepsilon_{2t\tau} < u'D_T^{-1}z_{t-1}) \to_d \frac{1}{2}f(F^{-1}(\tau))u'[\int B_z B_z']uI(u'B_z > 0).$$

Similarly,

$$\sum_{t=1}^{T}(u'D_T^{-1}z_{t-1} - \varepsilon_{2t\tau})I(0 > \varepsilon_{2t\tau} > u'D_T^{-1}z_{t-1}) \to_d \frac{1}{2}f(F^{-1}(\tau))u'[\int B_z B_z']uI(u'B_z < 0).$$

Thus, the second term of the objective function converges to $\frac{1}{2}f(F^{-1}(\tau))u'[\int B_z B_z']u$ in distribution. The limiting distribution of the objective function is as follows,

$$G_T(u) \to_d -u' \begin{bmatrix} \int dZ_\psi \\ \int \omega J_c dZ_\psi \end{bmatrix} + \frac{1}{2}f(F^{-1}(\tau))u' \begin{bmatrix} 1 & \int \omega J_c \\ \int \omega J_c & \int (\omega J_c)^2 \end{bmatrix} u := G(u).$$

Consider the first order condition of the minimization problem of $G(u)$,

$$\frac{\partial G(u)}{\partial u'} = -\begin{bmatrix} \int dZ_\psi \\ \int \omega J_c dZ_\psi \end{bmatrix} + f(F^{-1}(\tau)) \begin{bmatrix} 1 & \int \omega J_c \\ \int \omega J_c & \int (\omega J_c)^2 \end{bmatrix} u = 0,$$

the solution of which is the following

$$u = \frac{1}{f(F^{-1}(\tau))} \begin{bmatrix} 1 & \int \omega J_c \\ \int \omega J_c & \int (\omega J_c)^2 \end{bmatrix}^{-1} \begin{bmatrix} \int dZ_\psi \\ \int \omega J_c dZ_\psi \end{bmatrix}.$$

$G_T(u)$ is minimized at $\hat{u} = D_T(\hat{\gamma}(\tau) - \gamma(\tau))$, and $\hat{u} \rightarrow_d u$ (Knight 1989, Xiao 2009), i.e.

$$D_T(\hat{\gamma}(\tau) - \gamma(\tau)) \rightarrow_d \frac{1}{f(F^{-1}(\tau))} \begin{bmatrix} 1 & \int \omega J_c \\ \int \omega J_c & \int (\omega J_c)^2 \end{bmatrix}^{-1} \begin{bmatrix} \int dZ_\psi \\ \int \omega J_c dZ_\psi \end{bmatrix}.$$

A.2 Proof of Proposition 3

We have $\psi_\tau(\varepsilon_{2t\tau}) = \tau - I(\varepsilon_{2t\tau} < 0) = \tau - I(\varepsilon_{2t} < F^{-1}(\tau))$, and $E[\psi_\tau(\varepsilon_{2t\tau})] = 0, var[\psi_\tau(\varepsilon_{2t\tau})] = \tau(1-\tau)$. Then, we have $\Omega_\tau = \begin{bmatrix} \omega^2 & \omega\sqrt{\tau(1-\tau)}\delta \\ \omega\sqrt{\tau(1-\tau)}\delta & \tau(1-\tau) \end{bmatrix}$.

Thus,

$$\begin{aligned} Z_\psi &= \sqrt{\tau(1-\tau)}B_\psi \quad \text{where } B_\psi \text{ is a standard BM} \\ &= \sqrt{\tau(1-\tau)}[\delta B_1 + \sqrt{1-\delta^2}\tilde{B}_\psi] \quad \text{where } \tilde{B}_\psi \text{ is independent of } B_1. \end{aligned}$$

Therefore,

$$\begin{aligned} t_{\gamma_1}(\tau) &\rightarrow_d \frac{\int J_c^\mu dZ_{2\tau}}{[\int (J_c^\mu)^2]^{\frac{1}{2}}} \\ &= \frac{\sqrt{\tau(1-\tau)} \int J_c^\mu d(\delta B_1 + \sqrt{1-\delta^2}\tilde{B}_\psi)}{[\int (J_c^\mu)^2]^{\frac{1}{2}}} \\ &= \sqrt{\tau(1-\tau)} \left[\delta \frac{\int J_c^\mu dB_1}{[\int (J_c^\mu)^2]^{\frac{1}{2}}} + \sqrt{1-\delta^2} \frac{\int J_c^\mu d\tilde{B}_\psi}{[\int (J_c^\mu)^2]^{\frac{1}{2}}} \right] \\ &= \sqrt{\tau(1-\tau)} \left[\delta \frac{\int J_c^\mu dB_1}{[\int (J_c^\mu)^2]^{\frac{1}{2}}} + \sqrt{1-\delta^2} z \right] \end{aligned}$$

where z is a standard normal random variable that is independent of (B_1, J_c).

B Tables

Table 1: Standard t-tests: Finite-sample size (T = 1000)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.1240	0.1055	0.0885	0.0900	0.0915	0.0870	0.0915	0.0930	0.1005	0.1055	0.1235
- 1	0.1305	0.1125	0.0855	0.0950	0.0910	0.0915	0.0950	0.0910	0.0940	0.1050	0.1255
- 5	0.1370	0.1085	0.0990	0.0915	0.0855	0.0835	0.0820	0.0910	0.0955	0.1100	0.1225
-10	0.1405	0.1055	0.1015	0.1000	0.0885	0.0880	0.0840	0.0865	0.0925	0.1175	0.1215
-20	0.1340	0.1130	0.0980	0.0905	0.0940	0.0880	0.0835	0.0950	0.0880	0.1140	0.1230
-30	0.1335	0.1150	0.1015	0.0900	0.0915	0.0955	0.0850	0.0900	0.0910	0.1105	0.1245
- T	0.1555	0.1065	0.0965	0.1020	0.0875	0.0935	0.0940	0.0935	0.1100	0.1175	0.1310
						$\delta = 0.25$					
0	0.1350	0.1130	0.1065	0.0980	0.1070	0.0975	0.0975	0.0975	0.1070	0.1225	0.1200
- 1	0.1350	0.1200	0.0975	0.0980	0.0975	0.0965	0.1045	0.1040	0.1055	0.1245	0.1265
- 5	0.1370	0.1050	0.0970	0.0975	0.0950	0.0940	0.0925	0.0990	0.1015	0.1190	0.1155
-10	0.1470	0.1155	0.1010	0.0940	0.0980	0.0940	0.0905	0.0950	0.0955	0.1185	0.1245
-20	0.1385	0.1155	0.0995	0.0925	0.0905	0.0945	0.0900	0.0930	0.0920	0.1210	0.1220
-30	0.1300	0.1090	0.1005	0.0975	0.0945	0.0995	0.0855	0.0935	0.0895	0.1110	0.1265
- T	0.1450	0.1055	0.1025	0.1000	0.0915	0.1025	0.0955	0.1005	0.1075	0.1070	0.1310
						$\delta = 0.50$					
0	0.1440	0.1320	0.1355	0.1365	0.1340	0.1345	0.1320	0.1315	0.1370	0.1405	0.1410
- 1	0.1445	0.1275	0.1215	0.1320	0.1210	0.1185	0.1285	0.1215	0.1205	0.1395	0.1325
- 5	0.1395	0.1140	0.0990	0.1090	0.1095	0.1035	0.1105	0.1045	0.1140	0.1195	0.1190
-10	0.1490	0.1125	0.1010	0.1030	0.0995	0.0985	0.1005	0.0980	0.1080	0.1190	0.1180
-20	0.1525	0.1205	0.1005	0.0975	0.0935	0.0980	0.0960	0.0950	0.1045	0.1215	0.1215
-30	0.1335	0.1160	0.1015	0.0955	0.0940	0.0865	0.0915	0.0965	0.1030	0.1135	0.1220
- T	0.1420	0.1075	0.1020	0.0960	0.0900	0.0945	0.0940	0.1010	0.1110	0.1040	0.1220
						$\delta = 0.75$					
0	0.1765	0.1725	0.1945	0.2095	0.2120	0.1915	0.1855	0.1975	0.1790	0.1790	0.1675
- 1	0.1535	0.1565	0.1630	0.1685	0.1725	0.1650	0.1615	0.1665	0.1650	0.1670	0.1595
- 5	0.1435	0.1305	0.1190	0.1310	0.1240	0.1180	0.1305	0.1155	0.1325	0.1330	0.1445
-10	0.1415	0.1195	0.1055	0.1150	0.1030	0.1095	0.1030	0.1045	0.1250	0.1185	0.1325
-20	0.1430	0.1175	0.1060	0.1010	0.0960	0.0940	0.0985	0.1025	0.1105	0.1280	0.1235
-30	0.1445	0.1160	0.0995	0.1040	0.0935	0.0950	0.0975	0.0965	0.1055	0.1200	0.1245
- T	0.1455	0.1115	0.1075	0.0970	0.0900	0.0940	0.0890	0.1025	0.1080	0.1140	0.1120
						$\delta = 0.95$					
0	0.2030	0.2190	0.2615	0.2765	0.2825	0.2725	0.2610	0.2645	0.2420	0.2075	0.1895
- 1	0.1750	0.1985	0.2180	0.2260	0.2300	0.2245	0.2200	0.2265	0.2175	0.1940	0.1845
- 5	0.1560	0.1465	0.1540	0.1470	0.1380	0.1370	0.1370	0.1480	0.1550	0.1420	0.1545
-10	0.1455	0.1305	0.1280	0.1245	0.1040	0.1100	0.1145	0.1215	0.1310	0.1240	0.1410
-20	0.1370	0.1235	0.1080	0.1035	0.1020	0.0995	0.1005	0.1060	0.1120	0.1155	0.1335
-30	0.1355	0.1225	0.1015	0.1055	0.1025	0.1035	0.1005	0.1030	0.1040	0.1135	0.1330
- T	0.1380	0.1130	0.1025	0.0990	0.0990	0.0915	0.0890	0.1040	0.1115	0.1170	0.1235

Notes: The asymptotic results are based on 2000 Monte Carlo replications with sample size T. Simulated data is generated according to: $x_t = (1 + \frac{c}{T})x_{t-1} + \varepsilon_{1t}$ and $y_t = \varepsilon_{2t}$, where ε_{1t} and ε_{2t} are $N(0,1)$ with correlation coefficient δ. The level of the test is 10%.

Table 2: Standard t-tests: Finite-sample size (T = 200)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.1640	0.1335	0.0995	0.0910	0.0945	0.0820	0.0880	0.1005	0.1040	0.1245	0.1630
- 1	0.1615	0.1305	0.1000	0.0865	0.0870	0.0765	0.0885	0.0955	0.1020	0.1125	0.1600
- 5	0.1625	0.1425	0.1085	0.0915	0.0965	0.0840	0.0835	0.0905	0.1055	0.1230	0.1680
-10	0.1710	0.1325	0.1060	0.0965	0.0940	0.0855	0.0885	0.0980	0.0970	0.1270	0.1680
-20	0.1685	0.1285	0.1025	0.1030	0.0975	0.0790	0.0830	0.0830	0.0955	0.1265	0.1700
-30	0.1665	0.1335	0.1110	0.1000	0.0930	0.0800	0.0875	0.0865	0.0860	0.1245	0.1785
- T	0.1550	0.1325	0.1140	0.0985	0.0880	0.0880	0.0850	0.1005	0.1085	0.1320	0.1875
						$\delta = 0.25$					
0	0.1710	0.1440	0.1145	0.1045	0.1045	0.0980	0.1065	0.1060	0.1050	0.1305	0.1750
- 1	0.1710	0.1460	0.1130	0.0955	0.1020	0.0885	0.0860	0.0965	0.1040	0.1240	0.1715
- 5	0.1635	0.1405	0.1060	0.0900	0.0975	0.0945	0.0865	0.0910	0.1120	0.1305	0.1770
-10	0.1660	0.1380	0.1090	0.0965	0.0990	0.0920	0.0895	0.0950	0.1055	0.1325	0.1645
-20	0.1580	0.1320	0.1070	0.1070	0.0960	0.0910	0.0865	0.0815	0.0970	0.1275	0.1625
-30	0.1625	0.1350	0.1180	0.1005	0.0930	0.0865	0.0940	0.0915	0.0905	0.1245	0.1755
- T	0.1630	0.1355	0.1060	0.1020	0.0900	0.0825	0.0855	0.1000	0.1100	0.1355	0.1895
						$\delta = 0.50$					
0	0.1775	0.1670	0.1500	0.1415	0.1425	0.1315	0.1345	0.1345	0.1400	0.1535	0.1880
- 1	0.1780	0.1565	0.1470	0.1220	0.1295	0.1115	0.1150	0.1190	0.1220	0.1475	0.1805
- 5	0.1615	0.1455	0.1170	0.1075	0.1045	0.1000	0.0970	0.1000	0.1100	0.1325	0.1740
-10	0.1685	0.1385	0.1185	0.1165	0.1035	0.1025	0.1020	0.0950	0.0965	0.1320	0.1595
-20	0.1660	0.1400	0.1130	0.1065	0.0935	0.0955	0.0890	0.0920	0.0975	0.1230	0.1625
-30	0.1640	0.1375	0.1085	0.1030	0.0990	0.0880	0.0910	0.0920	0.0930	0.1275	0.1665
- T	0.1545	0.1305	0.1095	0.0970	0.0880	0.0820	0.0855	0.1050	0.1170	0.1375	0.1810
						$\delta = 0.75$					
0	0.2155	0.2010	0.2085	0.2000	0.2085	0.1970	0.1895	0.1965	0.1995	0.1865	0.2075
- 1	0.2015	0.1835	0.1810	0.1855	0.1840	0.1675	0.1595	0.1675	0.1690	0.1755	0.1975
- 5	0.1685	0.1600	0.1435	0.1355	0.1270	0.1255	0.1175	0.1140	0.1270	0.1445	0.1670
-10	0.1625	0.1500	0.1190	0.1240	0.1135	0.1175	0.1110	0.1085	0.1075	0.1310	0.1730
-20	0.1595	0.1365	0.1145	0.1110	0.1000	0.1050	0.0985	0.0980	0.1035	0.1260	0.1560
-30	0.1580	0.1320	0.1170	0.1070	0.0900	0.0940	0.0985	0.0965	0.0935	0.1220	0.1650
- T	0.1715	0.1220	0.1030	0.0920	0.0955	0.0775	0.0920	0.1055	0.1160	0.1330	0.1850
						$\delta = 0.95$					
0	0.2475	0.2395	0.2505	0.2600	0.2645	0.2610	0.2630	0.2590	0.2470	0.2325	0.2355
- 1	0.2140	0.2140	0.2155	0.2320	0.2310	0.2265	0.2200	0.2150	0.2210	0.1985	0.2190
- 5	0.1810	0.1700	0.1520	0.1585	0.1515	0.1500	0.1480	0.1495	0.1355	0.1515	0.1765
-10	0.1790	0.1545	0.1305	0.1305	0.1240	0.1240	0.1250	0.1175	0.1125	0.1335	0.1635
-20	0.1690	0.1405	0.1200	0.1090	0.1010	0.1050	0.0955	0.1015	0.1030	0.1335	0.1655
-30	0.1670	0.1460	0.1175	0.0990	0.0910	0.0960	0.0950	0.0940	0.1015	0.1295	0.1690
- T	0.1625	0.1305	0.0950	0.0880	0.0900	0.0815	0.0950	0.1075	0.1140	0.1345	0.1885

Table 3: Known c: Finite-sample size (T = 1000)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.0950	0.0950	0.0890	0.0955	0.1050	0.0935	0.0955	0.0980	0.1000	0.0925	0.0845
- 1	0.0980	0.1025	0.0895	0.0965	0.1005	0.0995	0.1015	0.0945	0.0940	0.0910	0.0920
- 5	0.1005	0.0980	0.1030	0.0955	0.0995	0.0910	0.0910	0.0945	0.0955	0.0930	0.0895
-10	0.1075	0.0950	0.1025	0.1035	0.0950	0.0975	0.0880	0.0925	0.0925	0.1020	0.0910
-20	0.1010	0.1015	0.1000	0.0955	0.1035	0.0940	0.0885	0.0985	0.0875	0.0980	0.0910
-30	0.0955	0.1045	0.1025	0.0940	0.0985	0.1060	0.0890	0.0930	0.0915	0.0960	0.0945
- T	0.1150	0.0950	0.0995	0.1080	0.1010	0.1040	0.1005	0.0985	0.1095	0.0985	0.0960
						$\delta = 0.25$					
0	0.0975	0.0985	0.0860	0.0920	0.1075	0.0950	0.1030	0.0975	0.1035	0.1070	0.0930
- 1	0.1055	0.1010	0.0925	0.1010	0.1065	0.1000	0.1025	0.1010	0.0970	0.1035	0.0970
- 5	0.1050	0.0955	0.0975	0.0995	0.1080	0.0965	0.0965	0.1030	0.1000	0.1050	0.0825
-10	0.1130	0.1000	0.1030	0.1005	0.1060	0.0915	0.1030	0.0990	0.0930	0.1050	0.0915
-20	0.1045	0.1005	0.0995	0.0965	0.1005	0.0985	0.0955	0.1025	0.0945	0.1050	0.0915
-30	0.1005	0.0975	0.0950	0.0995	0.1030	0.1070	0.0970	0.1025	0.0915	0.0990	0.0925
- T	0.1185	0.0905	0.1050	0.1045	0.0990	0.1070	0.1070	0.1045	0.1075	0.0975	0.1025
						$\delta = 0.50$					
0	0.0995	0.1020	0.0905	0.0950	0.1025	0.1010	0.1030	0.0945	0.0930	0.1020	0.0955
- 1	0.1000	0.0970	0.0900	0.0975	0.1090	0.1005	0.0970	0.0955	0.1010	0.1035	0.0960
- 5	0.1005	0.0950	0.1005	0.1050	0.1085	0.0950	0.1050	0.0975	0.1025	0.1065	0.0880
-10	0.1115	0.0985	0.1020	0.1030	0.1065	0.0975	0.1050	0.1015	0.0990	0.1070	0.0875
-20	0.1100	0.1000	0.0910	0.1015	0.1060	0.1015	0.1000	0.1000	0.0985	0.1045	0.0940
-30	0.1000	0.1025	0.0945	0.0980	0.1020	0.1000	0.0960	0.0985	0.0970	0.0975	0.0965
- T	0.1140	0.0940	0.1015	0.1060	0.1025	0.1015	0.1040	0.1070	0.1105	0.0935	0.0930
						$\delta = 0.75$					
0	0.1025	0.1055	0.0925	0.0975	0.1055	0.1055	0.1045	0.0925	0.0975	0.1015	0.1005
- 1	0.0995	0.1085	0.0935	0.1070	0.1090	0.0950	0.1015	0.0930	0.1075	0.1110	0.0975
- 5	0.1025	0.0945	0.1005	0.1035	0.0995	0.0995	0.1010	0.1025	0.1025	0.1050	0.0960
-10	0.1000	0.1000	0.1015	0.1035	0.1025	0.0920	0.0980	0.1030	0.1055	0.1075	0.0915
-20	0.1125	0.1020	0.0900	0.1055	0.1010	0.0960	0.0980	0.1020	0.1000	0.1040	0.0870
-30	0.1045	0.1040	0.0915	0.0985	0.0955	0.0980	0.1000	0.1035	0.1000	0.1015	0.0955
- T	0.1125	0.0990	0.1080	0.1040	0.0975	0.0950	0.1040	0.1125	0.1110	0.1035	0.0930
						$\delta = 0.95$					
0	0.1040	0.1040	0.0990	0.1060	0.0985	0.1085	0.0965	0.0915	0.0925	0.0935	0.1020
- 1	0.0960	0.1070	0.0915	0.1050	0.0965	0.0900	0.1015	0.0940	0.0980	0.1100	0.0995
- 5	0.1005	0.0940	0.0915	0.1040	0.0955	0.0925	0.0905	0.0925	0.0980	0.1060	0.1120
-10	0.0985	0.1045	0.0900	0.1070	0.0955	0.0875	0.0915	0.0945	0.1030	0.0990	0.1025
-20	0.1015	0.1010	0.0915	0.0950	0.0960	0.0880	0.0935	0.0955	0.0950	0.1025	0.1060
-30	0.0980	0.1085	0.0930	0.0995	0.0990	0.0880	0.0965	0.0985	0.0930	0.0940	0.1040
- T	0.1095	0.1055	0.0985	0.1025	0.1090	0.1030	0.0985	0.1160	0.1130	0.1070	0.1010

Table 4: Known c: Finite-sample size (T = 200)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.1265	0.1185	0.1020	0.0960	0.1035	0.0910	0.0915	0.1055	0.1045	0.1105	0.1280
- 1	0.1275	0.1160	0.1025	0.0900	0.0955	0.0840	0.0925	0.0975	0.1025	0.0995	0.1230
- 5	0.1220	0.1250	0.1120	0.0955	0.1015	0.0885	0.0900	0.0935	0.1045	0.1045	0.1345
-10	0.1375	0.1175	0.1070	0.0995	0.1040	0.0945	0.0935	0.1000	0.0970	0.1075	0.1315
-20	0.1295	0.1190	0.1045	0.1100	0.1055	0.0910	0.0890	0.0870	0.0950	0.1140	0.1305
-30	0.1315	0.1175	0.1140	0.1030	0.1000	0.0895	0.0925	0.0925	0.0865	0.1105	0.1370
- T	0.1185	0.1195	0.1180	0.1050	0.0980	0.0960	0.0895	0.1045	0.1070	0.1185	0.1490
						$\delta = 0.25$					
0	0.1330	0.1095	0.0980	0.0970	0.0965	0.0955	0.0995	0.1100	0.1015	0.1175	0.1350
- 1	0.1215	0.1170	0.1055	0.0905	0.0955	0.0860	0.0850	0.1005	0.1035	0.1065	0.1350
- 5	0.1305	0.1160	0.1030	0.0935	0.1025	0.0935	0.0895	0.1005	0.1110	0.1145	0.1415
-10	0.1290	0.1280	0.1065	0.1005	0.1030	0.1020	0.0965	0.0980	0.1055	0.1170	0.1350
-20	0.1245	0.1130	0.1080	0.1135	0.1060	0.0955	0.0960	0.0850	0.0975	0.1175	0.1350
-30	0.1270	0.1180	0.1145	0.1040	0.1000	0.0945	0.0995	0.0940	0.0875	0.1130	0.1410
- T	0.1300	0.1235	0.1070	0.1005	0.1000	0.0895	0.0935	0.1060	0.1110	0.1215	0.1515
						$\delta = 0.50$					
0	0.1295	0.1180	0.1095	0.0955	0.0990	0.0945	0.0875	0.0970	0.1065	0.1220	0.1295
- 1	0.1255	0.1140	0.1120	0.0945	0.1005	0.0825	0.0850	0.1055	0.1040	0.1195	0.1310
- 5	0.1200	0.1140	0.0985	0.0995	0.0990	0.0950	0.0935	0.1020	0.1090	0.1160	0.1385
-10	0.1290	0.1190	0.1130	0.1090	0.1005	0.1080	0.1000	0.0970	0.1040	0.1165	0.1285
-20	0.1330	0.1170	0.1050	0.1045	0.1030	0.0985	0.0980	0.1010	0.0975	0.1105	0.1330
-30	0.1255	0.1155	0.1075	0.1025	0.1045	0.0940	0.0980	0.1025	0.0895	0.1150	0.1375
- T	0.1285	0.1120	0.1060	0.1005	0.0960	0.0945	0.0985	0.1150	0.1180	0.1210	0.1530
						$\delta = 0.75$					
0	0.1320	0.1175	0.1210	0.1140	0.1095	0.0935	0.0930	0.1035	0.1105	0.1225	0.1330
- 1	0.1325	0.1235	0.1120	0.1090	0.1085	0.0980	0.0865	0.1020	0.1000	0.1255	0.1340
- 5	0.1270	0.1210	0.1105	0.1085	0.1085	0.1035	0.0975	0.1040	0.1045	0.1130	0.1240
-10	0.1240	0.1195	0.1075	0.1050	0.1010	0.1085	0.1000	0.1035	0.1030	0.1175	0.1360
-20	0.1200	0.1120	0.1040	0.1020	0.1045	0.1065	0.0970	0.1005	0.0970	0.1140	0.1265
-30	0.1270	0.1130	0.1110	0.1075	0.1010	0.1015	0.0945	0.1030	0.0920	0.1125	0.1425
- T	0.1335	0.1130	0.1025	0.1030	0.1055	0.0900	0.1035	0.1155	0.1185	0.1225	0.1535
						$\delta = 0.95$					
0	0.1270	0.1200	0.1260	0.1140	0.1025	0.1015	0.1030	0.1045	0.1170	0.1275	0.1380
- 1	0.1300	0.1325	0.1185	0.1130	0.1230	0.1060	0.1030	0.1050	0.1130	0.1215	0.1320
- 5	0.1240	0.1265	0.1070	0.1130	0.1115	0.1145	0.1070	0.1065	0.1070	0.1105	0.1235
-10	0.1170	0.1155	0.1090	0.1080	0.1180	0.1075	0.1010	0.0995	0.0945	0.1145	0.1295
-20	0.1310	0.1110	0.1115	0.1020	0.0970	0.1030	0.0940	0.0980	0.0920	0.1145	0.1335
-30	0.1295	0.1190	0.1030	0.0945	0.0985	0.0995	0.0875	0.0975	0.0955	0.1145	0.1340
- T	0.1325	0.1225	0.0905	0.1025	0.1045	0.0905	0.1095	0.1190	0.1190	0.1235	0.1555

Table 5: Bonferroni correction: Finite-sample size (T = 1000)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.0950	0.0950	0.0890	0.0955	0.1050	0.0935	0.0955	0.0980	0.1000	0.0925	0.0845
- 1	0.0980	0.1025	0.0895	0.0965	0.1005	0.0995	0.1015	0.0945	0.0940	0.0910	0.0920
- 5	0.1005	0.0980	0.1030	0.0955	0.0995	0.0910	0.0910	0.0945	0.0955	0.0930	0.0895
-10	0.1075	0.0950	0.1025	0.1035	0.0950	0.0975	0.0880	0.0925	0.0925	0.1020	0.0910
-20	0.1010	0.1015	0.1000	0.0955	0.1035	0.0940	0.0885	0.0985	0.0875	0.0980	0.0910
-30	0.0955	0.1045	0.1025	0.0940	0.0985	0.1060	0.0890	0.0930	0.0915	0.0960	0.0945
- T	0.1150	0.0950	0.0995	0.1080	0.1010	0.1040	0.1005	0.0985	0.1095	0.0985	0.0960
						$\delta = 0.25$					
0	0.0945	0.0970	0.0870	0.0900	0.1030	0.0955	0.0995	0.0960	0.1010	0.1005	0.0900
- 1	0.1060	0.1020	0.0910	0.0965	0.1035	0.0985	0.0955	0.1005	0.0935	0.1035	0.0945
- 5	0.1050	0.0965	0.0995	0.1025	0.1115	0.0935	0.0970	0.1025	0.0990	0.1075	0.0820
-10	0.1135	0.0965	0.1035	0.1020	0.1055	0.0945	0.1045	0.1000	0.0920	0.1035	0.0900
-20	0.1045	0.1025	0.0980	0.0945	0.1010	0.1005	0.0975	0.1020	0.0930	0.1020	0.0900
-30	0.1000	0.0960	0.0955	0.1000	0.1030	0.1070	0.0965	0.1015	0.0920	0.0990	0.0925
- T	0.1185	0.0905	0.1050	0.1045	0.0990	0.1070	0.1070	0.1045	0.1075	0.0975	0.1025
						$\delta = 0.50$					
0	0.0960	0.0925	0.0860	0.0865	0.0990	0.0920	0.0985	0.0910	0.0875	0.0940	0.0900
- 1	0.0990	0.0930	0.0860	0.0935	0.0980	0.0905	0.0965	0.0915	0.0870	0.0960	0.0845
- 5	0.0980	0.0930	0.0965	0.1020	0.1045	0.0945	0.1070	0.0980	0.0990	0.1015	0.0855
-10	0.1080	0.0965	0.0995	0.1025	0.1085	0.0975	0.1000	0.1030	0.1020	0.1035	0.0870
-20	0.1095	0.0980	0.0920	0.1015	0.1070	0.1025	0.0975	0.0980	0.0975	0.1030	0.0940
-30	0.0985	0.1025	0.0935	0.0975	0.1025	0.1000	0.0955	0.1000	0.0975	0.0980	0.0965
- T	0.1140	0.0940	0.1015	0.1060	0.1025	0.1015	0.1040	0.1070	0.1105	0.0935	0.0930
						$\delta = 0.75$					
0	0.0905	0.0985	0.0900	0.0890	0.1000	0.1000	0.0930	0.0900	0.0830	0.0945	0.0915
- 1	0.0890	0.0900	0.0890	0.0915	0.1005	0.0855	0.0980	0.0840	0.0925	0.1010	0.0835
- 5	0.0980	0.0880	0.0990	0.1005	0.1000	0.1050	0.1060	0.1010	0.1065	0.0970	0.0925
-10	0.0980	0.0940	0.0995	0.1080	0.1075	0.0955	0.1005	0.1045	0.1095	0.1055	0.0870
-20	0.1070	0.1015	0.0920	0.1055	0.1040	0.0955	0.1000	0.1005	0.1025	0.1035	0.0875
-30	0.1040	0.1045	0.0935	0.1025	0.0985	0.0995	0.1025	0.1045	0.1005	0.1010	0.0940
- T	0.1125	0.0990	0.1080	0.1040	0.0975	0.0950	0.1040	0.1125	0.1110	0.1035	0.0930
						$\delta = 0.95$					
0	0.0920	0.0950	0.0930	0.0965	0.0945	0.0955	0.0895	0.0860	0.0870	0.0885	0.0890
- 1	0.0850	0.0875	0.0800	0.0920	0.0850	0.0820	0.0820	0.0855	0.0845	0.0955	0.0860
- 5	0.0880	0.0855	0.0930	0.1025	0.1000	0.1055	0.0985	0.0990	0.1075	0.1020	0.0960
-10	0.0935	0.1000	0.0900	0.1085	0.1005	0.0930	0.0950	0.0955	0.1100	0.0980	0.0960
-20	0.0970	0.0995	0.0915	0.0975	0.1015	0.0915	0.0940	0.0975	0.0970	0.1020	0.0990
-30	0.0980	0.1040	0.0935	0.1015	0.1000	0.0920	0.1000	0.1010	0.0955	0.0950	0.1005
- T	0.1095	0.1055	0.0985	0.1025	0.1090	0.1030	0.0985	0.1160	0.1130	0.1070	0.1010

Notes: The level of the predictability test, η_2, is 10%. For $\delta = 0.25, 0.50, 0.75, 0.95$, $\eta_1 = 0.7, 0.5, 0.4, 0.3$, respectively. Rejection frequencies for $\delta = 0$ are calculated using simulated i.i.d. critical values.

Table 6: Bonferroni correction: Finite-sample size (T = 200)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
						$\delta = 0$					
0	0.1265	0.1185	0.1020	0.0960	0.1035	0.0910	0.0915	0.1055	0.1045	0.1105	0.1280
- 1	0.1275	0.1160	0.1025	0.0900	0.0955	0.0840	0.0925	0.0975	0.1025	0.0995	0.1230
- 5	0.1220	0.1250	0.1120	0.0955	0.1015	0.0885	0.0900	0.0935	0.1045	0.1045	0.1345
-10	0.1375	0.1175	0.1070	0.0995	0.1040	0.0945	0.0935	0.1000	0.0970	0.1075	0.1315
-20	0.1295	0.1190	0.1045	0.1100	0.1055	0.0910	0.0890	0.0870	0.0950	0.1140	0.1305
-30	0.1315	0.1175	0.1140	0.1030	0.1000	0.0895	0.0925	0.0925	0.0865	0.1105	0.1370
- T	0.1185	0.1195	0.1180	0.1050	0.0980	0.0960	0.0895	0.1045	0.1070	0.1185	0.1490
						$\delta = 0.25$					
0	0.1295	0.1095	0.0980	0.0975	0.1000	0.0915	0.0995	0.1035	0.1005	0.1125	0.1325
- 1	0.1210	0.1190	0.1045	0.0935	0.0930	0.0825	0.0825	0.0985	0.0970	0.1055	0.1315
- 5	0.1265	0.1190	0.1010	0.0945	0.1005	0.0900	0.0890	0.0990	0.1045	0.1115	0.1395
-10	0.1295	0.1260	0.1055	0.1005	0.1020	0.1020	0.0945	0.0960	0.1035	0.1180	0.1345
-20	0.1235	0.1130	0.1070	0.1145	0.1050	0.0960	0.0960	0.0845	0.0975	0.1165	0.1340
-30	0.1265	0.1165	0.1150	0.1045	0.1015	0.0920	0.0995	0.0940	0.0875	0.1085	0.1390
- T	0.1300	0.1235	0.1070	0.1005	0.1000	0.0895	0.0935	0.1060	0.1110	0.1215	0.1515
						$\delta = 0.50$					
0	0.1220	0.1110	0.1090	0.0920	0.0950	0.0920	0.0825	0.0990	0.1010	0.1105	0.1205
- 1	0.1160	0.1025	0.1080	0.0935	0.0965	0.0755	0.0790	0.0930	0.0915	0.1080	0.1215
- 5	0.1160	0.1125	0.0965	0.0970	0.0990	0.0910	0.0915	0.0965	0.1030	0.1090	0.1265
-10	0.1260	0.1155	0.1125	0.1090	0.1050	0.1040	0.0985	0.0940	0.1005	0.1135	0.1210
-20	0.1330	0.1170	0.1065	0.1070	0.1035	0.0995	0.0980	0.1010	0.0960	0.1100	0.1310
-30	0.1245	0.1150	0.1075	0.1020	0.1025	0.0915	0.1005	0.1000	0.0890	0.1120	0.1360
- T	0.1285	0.1120	0.1060	0.1005	0.0960	0.0945	0.0985	0.1150	0.1180	0.1210	0.1530
						$\delta = 0.75$					
0	0.1165	0.1115	0.1135	0.1015	0.1035	0.0900	0.0835	0.0960	0.1065	0.1130	0.1220
- 1	0.1205	0.1180	0.1055	0.0960	0.1030	0.0870	0.0890	0.0910	0.0970	0.1085	0.1185
- 5	0.1130	0.1130	0.1115	0.1130	0.1125	0.1105	0.1060	0.1050	0.1050	0.1065	0.1125
-10	0.1195	0.1160	0.1115	0.1075	0.1025	0.1165	0.0995	0.1080	0.1015	0.1155	0.1290
-20	0.1170	0.1085	0.1050	0.1040	0.1040	0.1085	0.0995	0.1020	0.0995	0.1135	0.1250
-30	0.1245	0.1150	0.1120	0.1090	0.1010	0.1000	0.0995	0.1030	0.0930	0.1120	0.1415
- T	0.1335	0.1130	0.1025	0.1030	0.1055	0.0900	0.1035	0.1155	0.1185	0.1225	0.1535
						$\delta = 0.95$					
0	0.1180	0.1095	0.1030	0.0940	0.1005	0.1000	0.0965	0.0995	0.0990	0.1050	0.1165
- 1	0.1110	0.1100	0.0985	0.1015	0.1035	0.0980	0.0910	0.0945	0.0950	0.0980	0.1135
- 5	0.1080	0.1190	0.1100	0.1150	0.1195	0.1235	0.1130	0.1125	0.1075	0.1035	0.1140
-10	0.1145	0.1130	0.1095	0.1090	0.1205	0.1120	0.1100	0.1115	0.1030	0.1105	0.1190
-20	0.1280	0.1115	0.1110	0.1045	0.1010	0.1065	0.0985	0.1020	0.0930	0.1145	0.1295
-30	0.1250	0.1210	0.1065	0.0965	0.1045	0.1015	0.0885	0.1025	0.0980	0.1185	0.1335
- T	0.1325	0.1225	0.0905	0.1025	0.1045	0.0905	0.1095	0.1190	0.1190	0.1235	0.1555

Table 7: Bonferroni correction: Finite-sample size with estimated δ (T = 1000)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
					$\delta = 0.50$						
0	0.0950	0.0915	0.0875	0.0885	0.1005	0.0935	0.0975	0.0920	0.0860	0.0925	0.0895
- 1	0.0980	0.0945	0.0865	0.0940	0.1000	0.0890	0.0975	0.0910	0.0875	0.0970	0.0825
- 5	0.0995	0.0910	0.0975	0.1025	0.1055	0.0940	0.1070	0.0950	0.1010	0.1010	0.0855
-10	0.1070	0.0965	0.0995	0.1025	0.1085	0.0965	0.1015	0.1045	0.1010	0.1025	0.0870
-20	0.1095	0.0980	0.0915	0.1010	0.1055	0.1030	0.0975	0.0975	0.0975	0.1030	0.0940
-30	0.0990	0.1025	0.0935	0.0975	0.1025	0.0990	0.0960	0.1015	0.0975	0.0975	0.0955
- T	0.1140	0.0940	0.1010	0.1055	0.1025	0.1010	0.1045	0.1080	0.1105	0.0935	0.0930
					$\delta = 0.90$						
0	0.0895	0.0965	0.0925	0.0950	0.0970	0.1005	0.0845	0.0865	0.0860	0.0920	0.0900
- 1	0.0920	0.0900	0.0875	0.0905	0.0915	0.0865	0.0870	0.0830	0.0880	0.0935	0.0855
- 5	0.0925	0.0915	0.0880	0.1050	0.0980	0.1040	0.1015	0.0925	0.1085	0.1045	0.0940
-10	0.0935	0.0980	0.0935	0.1075	0.1020	0.0965	0.0980	0.0955	0.1075	0.1020	0.0905
-20	0.0980	0.1020	0.0905	0.1030	0.1030	0.0930	0.0950	0.0925	0.0990	0.1045	0.0960
-30	0.0980	0.1045	0.0940	0.1035	0.1025	0.0910	0.0980	0.0975	0.0950	0.1020	0.0990
- T	0.1085	0.1030	0.1020	0.1060	0.1125	0.1005	0.1050	0.1125	0.1020	0.1075	0.0910

Notes: δ is estimated by $\hat{\delta} = corr(\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t})$, where $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ are the residuals. The relevant Bonferroni critical values are calculated by interpolation with respect to $\hat{\delta}$.

Table 8: Bonferroni correction: Finite-sample size with estimated δ (T = 200)

c	$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
					$\delta = 0.50$						
0	0.1245	0.1115	0.1090	0.0920	0.0975	0.0910	0.0865	0.1005	0.1005	0.1145	0.1240
- 1	0.1180	0.1065	0.1090	0.0925	0.0965	0.0750	0.0800	0.0930	0.0935	0.1095	0.1245
- 5	0.1175	0.1135	0.0990	0.0940	0.0970	0.0890	0.0915	0.0985	0.1055	0.1080	0.1270
-10	0.1255	0.1150	0.1120	0.1085	0.1020	0.1030	0.0985	0.0950	0.1020	0.1135	0.1210
-20	0.1320	0.1180	0.1065	0.1070	0.1040	0.1010	0.0945	0.1000	0.0970	0.1105	0.1305
-30	0.1250	0.1155	0.1075	0.1020	0.1040	0.0925	0.1010	0.1005	0.0905	0.1150	0.1350
- T	0.1305	0.1115	0.1065	0.1025	0.0970	0.0945	0.0975	0.1155	0.1185	0.1215	0.1540
					$\delta = 0.90$						
0	0.1120	0.1105	0.1035	0.1035	0.1030	0.0995	0.0920	0.0940	0.1050	0.1060	0.1160
- 1	0.1145	0.1135	0.1000	0.1050	0.1050	0.0940	0.0920	0.0975	0.0995	0.0995	0.1175
- 5	0.1130	0.1150	0.1095	0.1165	0.1240	0.1185	0.1195	0.1150	0.1090	0.1105	0.1155
-10	0.1125	0.1080	0.1115	0.1090	0.1175	0.1145	0.1080	0.1090	0.0970	0.1135	0.1205
-20	0.1195	0.1120	0.1055	0.0995	0.1000	0.1080	0.0980	0.0995	0.0960	0.1110	0.1255
-30	0.1225	0.1195	0.1085	0.1000	0.0945	0.1070	0.0935	0.0975	0.0960	0.1145	0.1370
- T	0.1380	0.1215	0.0960	0.1020	0.1035	0.0915	0.1050	0.1185	0.1195	0.1265	0.1550

Table 9: Predictor variables

Notation	Variable name	Sample period	Lag length	95% CI on c	$\hat{\delta}$
d/p	Dividend price ratio	02/1871-12/2005	2	[-20.2719, 2.3421]	-0.9645
d/y	Dividend yield	02/1871-12/2005	2	[-20.4209, 2.3135]	-0.1142
e/p	Earnings price ratio	02/1871-12/2005	2	[-, -8.3450]	-0.8987
e10/p	Smoothed earnings price ratio	12/1880-12/2005	7	[-32.9945 -3.9005]	-0.8634
b/m	Book to market ratio	03/1921-12/2005	4	[-19.9380, 2.4088]	-0.8243
tbl	T-bill rate	02/1920-12/2005	3	[-18.9895, 2.6330]	-0.0756
lty	Long term yield	01/1919-12/2005	4	[-7.0319, 4.3407]	-0.1481
tms	Term spread	02/1920-12/2005	2	[-, -29.0576]	-0.0043
dfy	Default yield spread	01/1919-12/2005	4	[-28.7040, 0.5644]	-0.2428
dfr	Default return spread	01/1926-12/2005	1	[-, -]	0.0889
csp	Cross sectional premium	05/1937-12/2002	1	[-25.1078, 1.3438]	-0.0463
ltr	Long term rate of return	01/1926-12/2005	1	[-, -]	0.1329
svar	Stock variance	02/1885-12/2005	8	[-, -]	-0.2966
d/e	Dividend payout ratio	02/1871-12/2005	3	[-, -27.3195]	-0.0238
ntis	Net equity expansion	12/1926-12/2005	1	[-, -13.1848]	-0.0785
infl	Inflation	02/1913-12/2005	4	[-, -]	0.0240

Notes: The 16 predictor variables are analyzed by Cenesizoglu and Timmermann (2008) using the data from Goyal and Welch (2008). The lag length is selected by the BIC with a maximum of twelve lags. The confidence intervals for c are calculated based on Stock (1991) with simulations for $-38 \leq c \leq 6$. [-,-] meaning empty confidence interval results from $c < -38$, which implies stationarity. $\hat{\delta} = corr(\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t})$ denotes the residual cross-correlation estimate.

Table 10: Quantile coefficient estimates

		$\tau = 0.05$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
d/p	$\hat{\gamma}_1(\tau)$	-0.696	-0.313	-0.245	0.043	-0.234	-0.461**	-0.321	-0.036	-0.038	0.059	0.670
	$t_{\hat{\gamma}_1(\tau)}$	-1.141	-0.614	-0.681	0.127	-0.685	-1.377	-0.919	-0.098	-0.105	0.172	1.748
	p-value	0.254	0.539	0.496	0.899	0.493	0.169	0.358	0.922	0.916	0.863	0.081
d/y	$\hat{\gamma}_1(\tau)$	-0.316	0.155	-0.153	0.124	-0.078	-0.407	-0.254	0.043	-0.037	0.056	0.122
	$t_{\hat{\gamma}_1(\tau)}$	-0.512	0.277	-0.417	0.365	-0.225	-1.195	-0.724	0.118	-0.100	0.164	0.302
	p-value	0.608	0.782	0.677	0.715	0.822	0.232	0.469	0.906	0.920	0.870	0.763
e/p	$\hat{\gamma}_1(\tau)$	1.363	0.962	-0.060	0.280	0.380	-0.023	0.413	0.681	0.222	-0.008	-0.261
	$t_{\hat{\gamma}_1(\tau)}$	1.946	1.736	-0.146	0.714	0.975	-0.058	1.074	1.803	0.572	-0.019	-0.463
	p-value	0.052	0.083	0.884	0.475	0.329	0.954	0.284	0.071	0.568	0.985	0.644
e10/p	$\hat{\gamma}_1(\tau)$	-0.383*	-0.112	-0.059	0.015	0.033	0.005	0.114	0.161*	0.159**	0.281*	0.419*
	$t_{\hat{\gamma}_1(\tau)}$	-2.328	-0.936	-0.666	0.217	0.481	0.073	1.658	2.463	2.235	2.938	3.373
	p-value	0.020	0.349	0.505	0.828	0.631	0.942	0.097	0.014	0.025	0.003	0.001
b/m	$\hat{\gamma}_1(\tau)$	-4.762*	-2.280**	-1.034	-0.898	-0.711	-0.607	0.595	1.097	1.629	2.135	4.929*
	$t_{\hat{\gamma}_1(\tau)}$	-2.236	-1.628	-1.099	-1.021	-0.856	-0.749	0.742	1.424	2.021	2.106	4.260
	p-value	0.025	0.104	0.272	0.308	0.392	0.454	0.458	0.155	0.043	0.035	0.000
tbl	$\hat{\gamma}_1(\tau)$	0.115	0.014	-0.043	-0.104**	-0.146*	-0.139*	-0.135*	-0.158*	-0.127*	-0.190*	-0.225*
	$t_{\hat{\gamma}_1(\tau)}$	1.256	0.173	-0.619	-1.630	-2.428	-2.259	-2.072	-2.377	-2.055	-3.265	-3.004
	p-value	0.209	0.863	0.536	0.103	0.015	0.024	0.038	0.017	0.040	0.001	0.003
lty	$\hat{\gamma}_1(\tau)$	0.192	0.100	0.009	-0.086	-0.119*	-0.114*	-0.139*	-0.123*	-0.120*	-0.180*	-0.099
	$t_{\hat{\gamma}_1(\tau)}$	1.454	1.045	0.127	-1.338	-1.931	-1.832	-2.216	-1.927	-1.903	-2.689	-0.955
	p-value	0.146	0.296	0.899	0.181	0.053	0.067	0.027	0.054	0.057	0.007	0.340
tms	$\hat{\gamma}_1(\tau)$	-0.059	0.150	0.342*	0.126	0.201	0.126	0.091	0.081	0.172	0.511*	0.752*
	$t_{\hat{\gamma}_1(\tau)}$	-0.187	0.648	2.110	0.893	1.587	1.003	0.696	0.572	1.136	2.846	3.533
	p-value	0.852	0.517	0.035	0.372	0.113	0.316	0.486	0.567	0.256	0.004	0.000
dfy	$\hat{\gamma}_1(\tau)$	-4.145*	-2.946*	-1.475*	-0.942*	-0.518	0.039	0.502	0.661**	1.323*	2.108*	3.465*
	$t_{\hat{\gamma}_1(\tau)}$	-7.088	-4.850	-2.812	-2.207	-1.400	0.108	1.410	1.840	2.873	4.483	5.007
	p-value	0.000	0.000	0.005	0.027	0.162	0.914	0.159	0.066	0.004	0.000	0.000
dfr	$\hat{\gamma}_1(\tau)$	0.713*	0.174	0.270	0.139	0.082	0.093	-0.079	0.072	0.121	0.205	0.178
	$t_{\hat{\gamma}_1(\tau)}$	4.340	0.727	1.227	0.646	0.366	0.407	-0.376	0.363	0.641	0.967	0.711
	p-value	0.000	0.467	0.220	0.518	0.714	0.684	0.707	0.717	0.521	0.334	0.477
csp	$\hat{\gamma}_1(\tau)$	3.807*	2.391**	0.489	2.121*	2.065*	1.196	1.435**	2.298*	1.634**	2.022*	1.402
	$t_{\hat{\gamma}_1(\tau)}$	2.952	1.962	0.451	2.268	2.380	1.407	1.657	2.675	1.962	2.378	1.406
	p-value	0.003	0.050	0.652	0.023	0.017	0.160	0.098	0.008	0.050	0.017	0.160
ltr	$\hat{\gamma}_1(\tau)$	-0.174	0.079	0.034	-0.023	0.006	0.047	0.140*	0.098	0.068	0.180	0.250*
	$t_{\hat{\gamma}_1(\tau)}$	-0.784	0.706	0.341	-0.261	0.065	0.576	1.922	1.352	0.902	1.557	1.819
	p-value	0.433	0.480	0.733	0.794	0.948	0.565	0.055	0.176	0.367	0.120	0.069
svar	$\hat{\gamma}_1(\tau)$	-10.261**	-4.142**	-2.663*	-2.264*	-1.267*	-0.614	0.249	2.180*	2.745*	6.195**	8.517*
	$t_{\hat{\gamma}_1(\tau)}$	-2.035	-2.029	-3.778	-5.498	-3.838	-0.576	0.261	2.172	4.156	1.927	11.924
	p-value	0.042	0.043	0.000	0.000	0.000	0.565	0.794	0.030	0.000	0.054	0.000
d/e	$\hat{\gamma}_1(\tau)$	-0.040*	-0.028*	-0.009	-0.007	-0.009**	-0.014*	-0.014*	-0.014*	-0.004	0.004	0.020**
	$t_{\hat{\gamma}_1(\tau)}$	-3.971	-3.120	-1.540	-1.444	-1.810	-2.916	-2.699	-2.606	-0.649	0.638	1.976
	p-value	0.000	0.002	0.124	0.149	0.070	0.004	0.007	0.009	0.517	0.523	0.048
ntis	$\hat{\gamma}_1(\tau)$	-0.466*	-0.465*	-0.340*	-0.171	-0.153	-0.111	-0.076	-0.068	-0.034	-0.017	0.005
	$t_{\hat{\gamma}_1(\tau)}$	-3.435	-3.087	-2.392	-1.274	-1.285	-0.963	-0.723	-0.732	-0.390	-0.192	0.053
	p-value	0.001	0.002	0.017	0.203	0.199	0.335	0.470	0.464	0.697	0.848	0.958
infl	$\hat{\gamma}_1(\tau)$	0.312	-0.367	-0.493	-1.020*	-1.100*	-0.935*	-0.946*	-0.706*	-0.604*	-0.824*	-1.116*
	$t_{\hat{\gamma}_1(\tau)}$	1.056	-1.297	-1.590	-4.103	-4.699	-3.660	-3.458	-2.413	-2.045	-2.222	-2.467
	p-value	0.291	0.195	0.112	0.000	0.000	0.000	0.001	0.016	0.041	0.026	0.014

Notes: The first row of each panel contains the quantile coefficient estimates. The coefficient estimates for d/p, d/y, e/p and b/m are multiplied by 100. * denotes significance at 5% level and ** denotes significance at 10% level using simulated critical values from the look-up tables. The second row contains the corresponding t-statistics. The p-values in the third row are calculated under standard normal distribution.