A Specification Test for Dynamic Conditional Distribution Models with Function-Valued Parameters

Victor Troster[1] and Dominik Wied[2]

[1] Department of Applied Economics, Universitat de les Illes Balears, Cra. de Valldemossa, km 7.5, Palma de Mallorca, 07122, Spain

[2] Institute for Econometrics and Statistics, University of Cologne, 50923 Köln, Germany

Abstract

This paper proposes a practical and consistent specification test of conditional distribution models for dependent data in a general setting. Our approach covers conditional distribution models indexed by function-valued parameters, allowing for a wide range of useful models for risk management and forecasting, such as the quantile autoregressive model, the CAViaR model, and the distributional regression model. The new specification test (i) is valid for general linear and nonlinear conditional quantile models under dependent data, (ii) allows for dynamic misspecification of the past information set, (iii) is consistent against fixed alternatives, and (iv) has nontrivial power against Pitman deviations from the null hypothesis. As the test statistic is non-pivotal, we propose and theoretically justify a subsampling approach to obtain valid inference. Finally, we illustrate the applicability of our approach by analyzing models of the returns distribution and Value-at-Risk (VaR) of two major stock indexes.

Keywords: Quantile regression; Distributional regression; Dynamic misspecification; Empirical processes; Subsampling.

JEL classification: C12; C22; C52.

1. Introduction

Many important economic and finance hypotheses are investigated through testing the specification of restrictions on the conditional distribution of a time series, such as conditional goodness-of-fit (Box and Pierce, 1970), conditional quantiles (Koenker and Machado, 1999), and distributional Granger non-causality (Taamouti et al., 2014). Rather than focusing only on a single part of the conditional distribution, the quantile regression model provides a more detailed analysis than the one obtained by a least squares regression. For example, the conditional quantile regression helps examine how a treatment affects the distribution of an outcome of interest; or it allows one to directly measure the market risk of financial institutions by estimating a particular quantile of future portfolio values, the Value at Risk (VaR). Therefore, an accurate specification of the whole distribution should consider the specification of the conditional quantile regression at all quantile levels.

This paper proposes a practical and consistent specification test of conditional distribution models for dependent data in a general setting. Our approach covers dynamic conditional distribution models indexed by function-valued parameters. The difference between our approach and that taken elsewhere is motivated within the framework of Corradi and Swanson (2006) and Rothe and Wied (2013). First, we generalize the approach of Rothe and Wied (2013) to testing the specification of dynamic conditional distribution models indexed by function-valued parameters in contexts with dependent data. Allowing the parameters to be function-valued is important for many empirical applications. For example, our approach covers the linear quantile autoregressive (QAR) of Koenker and Xiao (2006), which implies a linear structure for the inverse of the dynamic conditional distribution $F_{Y_t}^{-1}(\tau|\mathcal{F}_{t-1},\theta_0) = X_t'\theta_0(\tau)$ for some quantile $\tau \in (0,1)$, where $\mathcal{F}_t$ is the σ-field generated by $\{Y_s, s \leq t\}$, $X_t = \{Y_{t-1},\ldots,Y_{t-p}\} \in \mathcal{F}_{t-1}$, and $\theta_0(\cdot)$ is some vector-valued functional parameter. Our procedure also considers testing the specification of nonlinear quantile autoregressive models, such as the CAViaR model of Engle and Manganelli (2004), that directly measures the market risk of financial institutions by estimating a particular quantile of future portfolio values - the Value-at-Risk (VaR). In this way, our method allows for testing the validity of models for expected shortfall (ES) and VaR, thus complementing former backtests such as in Christoffersen and Pelletier (2004), Candelon et al. (2011), Wied et al. (2016), and Du and Escanciano (2016).

Besides, we extend the validity of Kolmogorov-type conditional distribution tests proposed by Corradi and Swanson (2006) to the context of dynamic conditional distribution models indexed by function-valued parameters. Rather than analysing models indexed by finite-dimensional parameters as in Corradi and Swanson (2006), we derive a test statistic for conditional distribution models indexed by function-valued parameters that is valid under dynamic misspecification of the past information set and parameter estimation error. We are unaware of a consistent specification test of conditional distribution models indexed by function-valued parameters under dependent data.

Our test statistic is a functional of the difference between the empirical distribution function and a restricted estimate implied by the dynamic conditional quantile or distributional regression model. Since its asymptotic distribution under general time series assumptions is non-pivotal, we propose and justify a subsampling scheme to obtain valid inference. We develop a test statistic that (i) allows for dynamic misspecification of the past information set, (ii) does not require the estimation of smoothing parameters or nuisance functions used in a martingale transformation as in Koenker and Xiao (2002) or in Bai (2003), and (iii) is consistent against all fixed alternatives. Besides, our test statistic has nontrivial power against Pitman local alternatives.

Zheng (1998), Koenker and Machado (1999), Bierens and Ginther (2001), Horowitz and Spokoiny (2002), and Koenker and Xiao (2002) have developed tests for the specification of quantile regression models for independent observations. However, these tests do not check for the validity of the quantile regression model itself, as they analyze only a single quantile, which is generally the median. Rothe and Wied (2013) and Escanciano and Goh (2014) proposed specification tests of a conditional quantile regression over a continuum of quantiles. Nevertheless, none of these tests is justified for dependent data, ruling out time series applications.

Koul and Stute (1999), He and Zhu (2003), and Whang (2006) proposed consistent specification tests of conditional quantile models for time series data. However, their approach is valid only for a single quantile. Escanciano and Velasco (2010) developed specification tests of dynamic quantile models over a continuum of quantiles. In contrast to the approach of Escanciano and Velasco (2010), our test does not require a martingale difference sequence assumption. As a result, our proposed approach allows for dynamic misspecification of the past information set. In addition, our method can also test the specification of models for the whole conditional distribution, like distributional regression models, while the framework Escanciano and Velasco (2010) analyzes only conditional quantile regression models.

An additional benefit of our approach is that it tests the specification of the distributional regression model introduced by Foresi and Peracchi (1995), where the conditional distribution is modeled through the application of a continuum of binary regressions to the data.

Chernozhukov et al. (2013) show that distributional regression models encompass the transformation/duration model of Cox (1972) as a special case, and it provides an alternative to the quantile regression model of Koenker and Bassett (1978). While the quantile regression model requires a smooth conditional distribution, the distributional regression does not need smoothness conditions, as the approximation is done pointwise in the threshold $y \in \mathbb{R}$, for the event that the variable Y_t exceeds this threshold y; hence it deals with continuous, discrete, or mixed variables without any other adjustment. Andersen et al. (2011) and Bollerslev et al. (2016) show that there are intraday stock price discontinuities, or jumps, during the active part of the trading day and the overnight close-to-open stock return, respectively. This has important implications to forecasting excess stock market returns and finding an optimal portfolio. Thus, the distributional regression approach uncovers stock price discontinuities and a higher-order multidimensional structure that are ignored by modeling only the first two moments of the conditional distribution. We are unaware of a specification test for the validity of distributional regression models under dependent data.

Koul and Stute (1999), Zheng (2000), Fan et al. (2006), Li and Tkacz (2006), Delgado and Stute (2008), Neumann and Paparoditis (2008), Li and Tkacz (2011), Bierens and Wang (2012), Chen and Hong (2014), Kheifets (2015), and Bierens and Wang (2017), among others, have also developed consistent specification tests for conditional distribution models for dependent data, but these methods cannot be applied to evaluate models indexed by function-valued parameters. In sum, our approach is a useful alternative to existing specification methods for dynamic conditional models under dependent data because it allows for models indexed by possibly function-valued parameters, covering the setups of Corradi and Swanson (2006), Escanciano and Velasco (2010), Chernozhukov et al. (2013), and Rothe and

Wied (2013) in a unified way.

The rest of the paper is organized as follows. In Section 2, we propose a test statistic for the null hypothesis of correct specification of dynamic conditional distribution models indexed by function-valued parameters under dependent data; we also derive the asymptotic limit distribution of our test statistic under the null and the alternative hypotheses. In addition, we show that our test statistic has nontrivial power against Pitman local alternatives. In Section 3, we theoretically justify the validity of the subsampling approach in our framework. Section 4 presents Monte Carlo simulation results. In Section 5, we present an application to two European stock indexes, showing that our approach is a flexible alternative to standard procedures in evaluating VaR and distributional regression models. Finally, Section 6 concludes the paper.

2. Test statistic and asymptotic theory

Suppose we observe a sample $\{(Y_t, X_t) \in \mathbb{R} \times \mathbb{R}^d, t = 1, \ldots, T\}$ of T observations from a strictly stationary process $\{Y_t, X_t\}_{t=-\infty}^{\infty}$ defined on a complete probability space $(\Omega, \mathcal{A}, P)$, with unknown conditional distribution function $F(y|x) = P(Y_t \leq y|X_t = x)$ and joint distribution $F_{YX}(y, x) = P(Y_t \leq y, X_t \leq x)$, where X_t may contain lags of Y_t and of other variables. Let $\mathcal{F}_{t-1}$ be a σ-algebra generated by a sequence of historically observed random vectors e.g., $\mathcal{F}_{t-1} = \sigma\{Z_{t-1}, Z_{t-2}, Z_{t-3}, \ldots\}$ for $Z_t := (Y_t, X_t')'$, including all "relevant" past information. We are interested in the distribution of Y_t given a finite dimensional vector of conditioning variables $X_t \in \mathbb{R}^d$, for $X_t \in \mathcal{F}_{t-1}$. If X_t does not include enough past information, then there may be dynamic misspecification (Corradi and Swanson, 2006).

In empirical applications, we are unaware *a priori* of the "relevant" past information set

$\mathcal{F}_{t-1}$, and determining how much information to include requires pre-testing. By allowing for dynamic misspecification, we avoid such pre-testing. Besides, the critical values obtained for specification tests under the correct specification given $\mathcal{F}_{t-1}$ are generally invalid when the conditional distribution is correctly specified given X_t, for $X_t \in \mathcal{F}_{t-1}$. Hereafter, we follow the notation of Corradi and Swanson (2006) and define $\mathcal{F}_{t-1}$ as the information set containing all relevant past information, such that for any set $\mathcal{F}^*_{t-1} \ni \mathcal{F}_{t-1}$, we have $Y_t|\mathcal{F}_{t-1} \overset{d}{=} Y_t|\mathcal{F}^*_{t-1}$, where $\overset{d}{=}$ denotes equality in distribution. Then, we define dynamic misspecification as the case when $X_t \in \mathcal{F}_{t-1}$ and $Y_t|X_t \overset{d}{\neq} Y_t|\mathcal{F}_{t-1}$, where $\overset{d}{\neq}$ denotes nonequality in distribution.

Let $F(y|x, \theta_0)$ be a parametric family of conditional distribution functions possibly indexed by a function-valued parameter $\theta(\cdot) \in \mathcal{B}(\mathcal{T}, \Theta)$, a class of mappings $\tau \mapsto \theta(\tau)$ such that $\theta(\tau) \in \Theta \subset \mathbb{R}^K$ for every $\tau \in \mathcal{T} \subset \mathbb{R}$. We characterize the null hypothesis of correct specification as follows:

$$H_0: \quad F(y|x) = F(y|x, \theta_0), \text{ for some } \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and all } (y, x) \in \mathcal{W}, \tag{2.1}$$

$$H_A: \quad F(y'|x') \neq F(y'|x', \theta), \text{ for all } \theta \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and some } (y', x') \in \mathcal{W}, \tag{2.2}$$

where $\mathcal{W}$ is the support of $W_t := (Y_t, X_t')'$. As the alternative hypothesis consists of all the possible deviations from the null, the null hypothesis is true if and only if there exists some $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$ such that $F(y'|x') - F(y'|x', \theta) = 0$ for some $(y', x') \in \mathcal{W}$. Let $F_X(x) = P(X_t \leq x)$ be the marginal distribution of X_t and $\mathbb{1}\{X_t \leq x\}$ be an indicator function of the event where X_t is less than or equal to x.

To test H_0 in (2.1), we first restate our null hypothesis as an equality of unconditional distributions by integrating-up both sides of H_0 with respect to the marginal distribution of the conditioning variable $F_X(\cdot)$; see Theorem 16.10 (iii) in Billingsley (1995). In a time series context, Corradi and Swanson (2006) and Neumann and Paparoditis (2008) also applied this method to check for the correct specification of dynamic conditional distributions indexed by finite-dimensional parameters. Nevertheless, our null hypothesis tests the validity of a conditional distributional model indexed by function-valued parameters. As $F(y|x) = E(\mathbb{1}\{Y_t \leq y\}|X_t = x)$, we can restate the null hypothesis H_0 of (2.1) as follows:

$$\int F(y|x^*)\mathbb{1}\{x^* \leq x\}dF_X(x^*) = \int F(y|x^*, \theta_0)\mathbb{1}\{x^* \leq x\}dF_X(x^*),$$
$$\text{for some } \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and all } \ (y, x) \in \mathcal{W},$$

Then, our test statistic is the sample analog of $S(y, \theta)$:

$$S(y, \theta) := \int (F(y|x^*) - F(y|x^*, \theta))\, \mathbb{1}\{x^* \leq x\}dF_X(x^*). \tag{2.3}$$

Under the null hypothesis, we have $S(y, \theta_0) = 0$ for some $\theta_0 \in \mathcal{B}(\mathcal{T}, \Theta)$ and for all $(y, x) \in \mathcal{W}$, whereas $S(y, \theta) \neq 0$ for all $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$ and for some $(y, x) \in \mathcal{W}$ under the alternative hypothesis. The function $S(y, \theta_0)$ is a distance measure between the unconditional joint distribution function, $F_{YX}(y, x) := \int F(y|x^*)\mathbb{1}\{x^* \leq x\}dF_X(x^*)$, and the unconditional distribution function implied by the parametric model, $F(y, x, \theta_0) := \int F(y|x^*, \theta_0)\mathbb{1}\{x^* \leq x\}dF_X(x^*)$.

We estimate the conditional distribution via parametric quantile regressions to obtain our test statistic as we intend to test the specification of quantile regression models over a continuum of quantiles. Thus, we apply a generalized method of moments to estimate the conditional distribution function. Under the null hypothesis of (2.1), the functional parameter $\theta_0(\cdot)$ is identified for every $\tau \in \mathcal{T}$ through a moment condition. Thus, we assume that the function-valued parameter $\theta_0(\tau)$ solves the following moment condition for every $\tau \in \mathcal{T}$:

$$G(\theta, \tau) \equiv E\left[g(W_t, \theta, \tau)\right] = 0, \tag{2.4}$$

where $g(W_t, \theta, \tau) : \mathcal{W} \times \Theta \times \mathcal{T} \mapsto \mathbb{R}^K$ is a known uniformly integrable function. Thus, under H_0 of (2.1), any $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$ such that $F(y|x) = F(y|x, \theta)$, $\forall\ (y, x) \in \mathcal{W}$, also implies that $\theta(\tau) = \theta_0(\tau)$, $\forall\ \tau \in \mathcal{T}$. Therefore, the moment condition $G(\theta, \tau)$ in (2.4) uniquely defines the "true" function-valued parameter. We assume that there is a Z-estimator $\hat{\theta}_T$ that satisfies $\|\hat{G}(\hat{\theta}_T(\tau), \tau)\|^2 \leq \inf_{\theta \in \Theta} \|\hat{G}(\theta, \tau)\|^2 + o_p(T^{-1/2})$ uniformly over $\tau \in \mathcal{T}$, where $\hat{G}(\theta, \tau) \equiv (1/T)\sum_{t=1}^{T} g(W_t, \theta, \tau)$ is the sample analog of the moment condition (2.4), and $\|\cdot\|$ denotes the supremum norm. We also assume that the moment condition $G(\theta, \tau)$ is smooth with respect to both θ and τ described below. Under mild conditions (Chernozhukov et al., 2013), we have:

$$\sqrt{T}(\hat{\theta}_T(\cdot) - \theta_0(\cdot)) \Rightarrow -\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot), \cdot)) \text{ in } \ell^\infty(\mathcal{T}),$$

where "$\Rightarrow$" denotes weak convergence to a random element in the function space $\ell^\infty(\mathcal{T})$ (in the Hoffmann-Jørgensen sense) for the metric induced by $\|\cdot\|$, $\partial G(\theta_0, \tau)/\partial\theta := \dot{G}_{\theta_0,\tau}$, and $\tilde{\mathbb{H}}_2$ is a tight mean zero Gaussian process with covariance function defined in the Appendix (see Lemma A.2 in the Appendix). Our framework allows for many flexible models such as the quantile autoregressive model and the distributional regression model. For instance, the quantile autoregressive model of order p of Koenker and Xiao (2006) implies the inverse of the conditional distribution

$$F^{-1}(\tau|Y_{t-1},\ldots,Y_{t-p},\theta) = \theta_0(\tau) + \theta_1(\tau)Y_{t-1} + \ldots + \theta_p(\tau)Y_{t-p}$$
$$= X_t'\theta(\tau),$$

for every $\tau \in \mathcal{T} \subset (0,1)$, where $F^{-1}(\tau|Y_{t-1},\ldots,Y_{t-p},\theta) = Q_\tau(Y_t|Y_{t-1},\ldots,Y_{t-p},\theta)$ given a parametric conditional distribution implied by the QAR model, $F(y|X_t,\theta) = \int_{\mathcal{T}} \mathbb{1}\{y - X_t'\theta(\tau) \leq 0\}d\tau$, and $X_t = (1, Y_{t-1},\ldots,Y_{t-p})'$. We estimate the parameter vector $\theta_0(\tau)$ of the QAR model above as any solution $\hat{\theta}_T(\tau)$ to the problem

$$\hat{\theta}_T(\tau) := \arg\min_{\theta\in\Theta} \sum_{t=1}^{T} (Y_t - X_t'\theta)(\tau - \mathbb{1}\{Y_t - X_t'\theta \leq 0\}).$$

Then, our setup can test the specification of QAR models by setting $g(W_t,\theta,\tau) = (\mathbb{1}\{Y_t - X_t'\theta \leq 0\} - \tau)X_t$ and a moment condition $G(\theta,\tau) = E[(F(y|X_t,\theta) - \tau)X_t]$. Our framework also checks the validity of the distributional regression model proposed by Foresi and Peracchi (1995), where the conditional distribution function of Y_t given X_t is specified for a range of thresholds $\tau \in \mathcal{T} \subset \mathbb{R}$ as $F(\tau|X_t,\theta) = \Lambda(X_t'\theta(\tau))$, where $\Lambda(\cdot)$ is a known strictly increasing link function. We estimate the parameter vector $\theta_0(\tau)$ for every $\tau \in \mathcal{T} \subset \mathbb{R}$ as follows:

$$\hat{\theta}_T(\tau) := \arg\max_{\theta\in\Theta} \sum_{t=1}^{T} [\mathbb{1}\{Y_t \leq \tau\} \ln(\Lambda(X_t'\theta)) + (1 - \mathbb{1}\{Y_t \leq \tau\}) \ln(1 - \Lambda(X_t'\theta))],$$

which consists of successive logistic or probit regressions of $\mathbb{1}\{Y_t \leq \tau\}$ on X_t for every $\tau \in \mathcal{T}$. Then, the distributional regression model fits into our framework with $g(W_t,\theta,\tau) = (\Lambda(X_t'\theta)(1 - \Lambda(X_t'\theta)))^{-1}(\Lambda(X_t'\theta) - \mathbb{1}\{Y_t \leq \tau\})\lambda(\Lambda(X_t'\theta))X_t$ and $G(\theta,\tau) = E[g(W_t,\theta,\tau)]$, where $\lambda(\cdot)$ is the derivative of $\Lambda(\cdot)$.

We estimate the parametric conditional distribution function $F(y|x, \theta_0)$ by the plug-in estimate $F(y|x, \hat{\theta}_T)$ built on an estimate $\hat{\theta}_T$ of θ_0. Let $\hat{Z}_T(y, x) = (1/T) \sum_{t=1}^{T} \mathbb{1}\{Y_t \leq y\}\mathbb{1}\{X_t \leq x\}$ be the empirical joint distribution of $\{Y_t, X_t\}_{t=1}^{T}$, and let $\hat{F}_T(y, x, \hat{\theta}_T) = (1/T) \sum_{t=1}^{T} F(y|X_t, \hat{\theta}_T)\mathbb{1}\{X_t \leq x\}$ be the semi-parametric analog of $F(y, x, \theta_0)$. Introducing these estimators into the definition of $S(y, \theta)$ in (2.3), we propose a test statistic based on the following discrepancy:

$$S_T(y, x) = (1/T) \sum_{t=1}^{T} \Big(\mathbb{1}\{Y_t \leq y\}\mathbb{1}\{X_t \leq x\} - F(y|X_t, \hat{\theta}_T)\mathbb{1}\{X_t \leq x\}\Big). \tag{2.5}$$

Under the null hypothesis in (2.1), both $\hat{Z}_T(y, x)$ and $\hat{F}_T(y, x, \hat{\theta}_T)$ are consistent for $F_{YX}(y, x)$ and $F(y, x, \theta_0)$, respectively, and we expect that $S_T(y, x)$ is approximately zero. Conversely, $F(y|x, \hat{\theta}_T)$ and $F(y|x)$ differ on a set with positive probability under the alternative hypothesis (2.2) so that the test statistic $S_T(y, x)$ diverges. Thus, based on $S_T(y, x)$ in (2.5), we apply a Cramér-von Mises-type test statistics for testing the null hypothesis in (2.1) against the alternative hypothesis in (2.2):

$$S_T^{CM} = \int_{\mathcal{W}} \left\| \sqrt{T} S_T(y, x) \right\|^2 d\hat{Z}_T(y, x). \tag{2.6}$$

It is also possible to apply a Kolmogorov-Smirnov-type functional norm of the test statistic as $S_T^{KS} = \sqrt{T} \sup_{(y,x)\in\mathcal{W}} \|S_T(y, x)\|$. However, unreported simulations suggest that the S_T^{CM} test statistic has better finite-sample size and power properties than the ones provided by the Kolmogorov-Smirnov-type test statistic.

2.1. Asymptotic null distribution and power

This subsection derives the asymptotic distribution of our test statistic S_T^{CM} in (2.6) under the null and alternative hypothesis. Our test statistic S_T^{CM} in (2.6) is based on empirical processes indexed by a class of functions $\ell^\infty(\mathcal{H})$, which is the class of real-valued functions that are uniformly bounded on $\mathcal{H}$, with $\mathcal{H} := \mathcal{W} \times \mathcal{T}$, equipped with the supremum norm $\|\cdot\|_{\ell^\infty(\mathcal{H})}$. To simplify notation, let $\|\cdot\|$ denote the supremum norm. All limits are taken as $T \to \infty$. To establish the asymptotic validity of our test statistic, we maintain the following assumption:

Assumption 1. (a) The class of functions $\mathcal{G} = \{W_t \mapsto g(W_t, \theta, \tau) : \theta \in \Theta, \tau \in \mathcal{T}\}$ is a VC-subgraph class of measurable functions with envelope $\mathbb{F}$ satisfying $E|\mathbb{F}|^p < \infty$, for some $2 < p < \infty$;

(b) $\{(Y_t, X_t) \in \mathbb{R} \times \mathbb{R}^d, t = 1, \ldots, T\}$ is a strictly stationary β-mixing sequence whose mixing coefficient is of order $O(T^{-b})$ for some $b > p/(p-2)$ and some $2 < p < \infty$. Besides, $E\|X_t\|^2 < \infty \ \forall t \geq 1$;

(c) For each $\tau \in \mathcal{T}$, the moment condition $G(\theta, \tau)$ in (2.4) has a unique zero at $\theta_0(\tau) \in \mathcal{B}(\mathcal{T}, \Theta)$, where $\mathcal{B}(\mathcal{T}, \Theta)$ is family of uniformly bounded functions from $\mathcal{T}$ to $\Theta \subset \mathbb{R}^K$, for an arbitrary subset Θ of $\mathbb{R}^K$ and a compact set $\mathcal{T}$ of some metric space;

(d) The conditional distribution function $F(y|x, \theta)$ has a density function $f(y|x, \theta)$ that is continuous, bounded, and bounded away from zero uniformly over the quantiles of interest $\tau \in \mathcal{T}$, almost surely. In addition, $F(y|x, \theta)$ has uniformly bounded second-order derivatives with respect to $y \in \mathbb{R}$ a.s., and the distribution of X_t is absolutely continuous with respect to Lebesgue measure $\forall t \geq 1$. Besides, $\partial G(\theta_0, \tau)/\partial\theta := \dot{G}_{\theta_0,\tau}$ is nonsingular uniformly over $\tau \in \mathcal{T}$.

Assumption 1 provides standard conditions for quantile auto-regressive and distributional regression models under dependent data. Assumption 1 (a) is necessary to establish the weak convergence of an empirical process indexed by functions, see e.g. Arcones and Yu (1994) and Radulović (1996). Assumption 1 (b) is needed to restrict the dependence of $\{Y_t, X_t\}$ and holds for many econometric models in practice, including autoregressive moving average (ARMA) and generalized autoregressive conditional heteroscedasticity (GARCH) processes under mild additional assumptions. We assume β-mixing dependence because it allows for decoupling and yields exponential inequalities, rather than imposing stringent conditions on the entropy numbers of $\mathcal{G}$ and on the rate of decay for the α-mixing coefficients as in Andrews and Pollard (1994). Thus, the parts (a)-(b) of Assumption 1 guarantee the stochastic equicontinuity of the empirical process $\sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ (see Arcones and Yu, 1994).

The parts (c)-(d) of Assumption 1 are necessary to establish a functional central limit theorem for the Z-estimator process $\tau \mapsto \sqrt{T}(\hat{\theta}_T(\tau) - \theta_0(\tau))$, addressing the asymptotic continuity of the empirical processes of function-valued parameters (see e.g. Chernozhukov et al., 2013; Rothe and Wied, 2013). Assumption 1.(c) ensures that the moment condition $G(\theta, \tau)$ in (2.4) uniquely defines the "true" function-valued parameter. Therefore, θ_0 is still well-defined as the solution to the moment condition $G(\theta, \tau)$ in (2.4) under the alternative hypothesis in (2.2), and it can be regarded as a pseudo-true value of the function-valued parameter as in Rothe and Wied (2013).

Besides, Assumption 1 (d) is a smoothness condition required to deliver a functional delta-method for the subsampling of our test statistic. In what follows, "$\xrightarrow{d}$" denotes convergence in distribution. The following result delivers the limit distribution of the proposed test statistic S_T^{CM} in (2.6) under the null and the alternative hypotheses.

Proposition 1. If Assumption 1 is satisfied then

(a) Under H_0 in (2.1), $S_T^{CM} \xrightarrow{d} \int \|\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x)\|^2 \, dF_{YX}(y,x)$, where $(\mathbb{H}_1, \mathbb{H}_2)$ are Gaussian processes with zero mean and covariance function defined in the Appendix;

(b) Under H_A in (2.2), $\lim_{T\to\infty} P(S_T^{CM} > C) = 1$ for all fixed constants $C > 0$.

Proposition 1 shows that the asymptotic null distribution of S_T^{CM} is a functional of a bivariate mean zero Gaussian process $(\mathbb{H}_1, \mathbb{H}_2)$ defined in the Appendix. By Proposition 1, we reject the null hypothesis H_0 whenever S_T^{CM} is significantly large.

2.2. Local power of the test statistic

Now we analyze the asymptotic power of S_T^{CM} in (2.6) against a sequence of Pitman local alternatives converging to the null hypothesis at the rate $\sqrt{T}$. We need to establish that S_T^{CM} in (2.6) has nontrivial power against local alternatives to ensure that it is asymptotically locally unbiased. Let $J(y|x)$ be a conditional distribution function such that $J(y|x) \neq F(y|x,\theta)$ for all $\theta \in \mathcal{B}(\mathcal{T},\Theta)$ and some $(y,x) \in \mathcal{W}$. For any constant $0 < \delta \leq \sqrt{T}$, we define the following sequence of local alternative conditional distribution functions of Y_t given X_t:

$$H_{A,T} : F_T\left(y|x\right) = \left(1 - \frac{\delta}{\sqrt{T}}\right) F\left(y|x,\theta\right) + \left(\frac{\delta}{\sqrt{T}}\right) J(y|x), \tag{2.7}$$

where $F(y|x,\theta) = F(y|x,\theta_0)$ for some $\theta_0 \in \mathcal{B}(\mathcal{T},\Theta)$ and all $(y,x) \in \mathcal{W}$. To ensure nontrivial local power of our test statistic, we make the following assumption:

Assumption 2. Under the local alternative in (2.7), the conditional distribution implies a sequence of distribution functions $F_T^A(y,x) = \int F_T(y|x^*)\mathbb{1}\{x^* \leq x\} dF_X(x^*)$ that is contiguous to the distribution function $F(y,x,\theta_0) = \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \leq x\} dF_X(x^*)$ implied by $F(y|x,\theta_0)$.

Assumption 2 is standard in the study of the asymptotic power under a sequence of

Pitman local alternatives (see Andrews, 1997; Escanciano and Velasco, 2010; Rothe and Wied, 2013). We define the moment conditions $G_J(\theta, \tau) := E_J[g(W_t, \theta, \tau)]$ and $G_F(\theta, \tau) := E_F[g(W_t, \theta, \tau)]$, where $E_J[\cdot]$ and $E_F[\cdot]$ denote expectation with respect to $J \equiv J(y|x)$ and $F \equiv F(y|x, \theta)$, respectively, in (2.7). Let $\theta_0(\cdot)$ and $\theta_0^*(\cdot)$ be solutions to the moment conditions $G_F(\theta_0, \tau) = 0$ and $G_J(\theta_0^*, \tau) = 0$ for every $\tau \in \mathcal{T}$, respectively. The following result delivers the asymptotic distribution of our test statistic S_T^{CM} in (2.6) under a sequence of local alternatives satisfying (2.7).

Proposition 2. Under the local alternative $H_{A,T}$ in (2.7) and Assumptions 1-2,

$$S_T^{CM} \xrightarrow{d} \int \|\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x) + \Delta(y,x)\|^2 \, dF_{YX}(y,x),$$

where $\Delta(y,x) = \delta \int (J(y|x^*) - F(y|x^*, \theta_0) + \dot{F}(y|x^*, \theta_0)[h]) \mathbb{1}\{x^* \leq x\} dF_X(x^*)$, and h is the function $h(\tau) = (\partial G_F(\theta_0, \tau)/\partial\theta)^{-1} G_J(\theta_0, \tau)$.

Let $c(1-\alpha) \equiv \inf\{s : P(S_T^{CM} \leq s) \geq 1-\alpha\}$. By Anderson's Lemma (see Corollary 2 of Anderson, 1955), since $\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x)$ has mean zero, $\forall (y,x) \in \mathcal{W}$, under $H_{A,T}$ in (2.7) we have that

$$\begin{aligned} &P\left(\int \|\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x) + \Delta(y,x)\|^2 dF_{YX}(y,x) > c(1-\alpha)\right) \\ \geq &P\left(\int \|\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x)\|^2 dF_{YX}(y,x) > c(1-\alpha)\right) = \alpha, \end{aligned}$$

where equality holds when $\Delta(y,x) = 0$ (see e.g. Theorem 4 in Andrews, 1997). Then, Proposition 2 implies that our test statistic S_T^{CM} is asymptotically unbiased against a sequence of local alternatives $H_{A,T}$ in (2.7).

In most time series applications, it is more useful to test the specification of a dynamic conditional quantile or distribution model over a range of the conditional distribution rather than in the entire distribution. For instance, when testing the validity of ES or VaR models, it is more important to test the specification of quantile regression model over a range of quantiles of the tail of the conditional distribution, (τ_*, τ^*) with $0 < \tau_* < \tau^* < 1$. Let $\overline{\mathcal{W}}$ be the subset of $W_t = (Y_t, X_t)'$ implied by the inverse of the conditional distribution $Y_t|X_t$ at the quantiles τ_* and τ^*, i.e., $\overline{\mathcal{W}} = \{W_t : F_{Y_t}^{-1}(\tau_*|X_t) \leq Y_t \leq F_{Y_t}^{-1}(\tau^*|X_t)\}$. Then, we can implement our test on a subset of the distribution by modifying our null hypothesis in (2.1) as follows:

$$H_0: \quad F(y|x) = F(y|x, \theta_0), \text{ for some } \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and all } (y, x) \in \overline{\mathcal{W}}, \tag{2.8}$$

$$H_A: \quad F(y'|x') \neq F(y'|x', \theta), \text{ for all } \theta \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and some } (y', x') \in \overline{\mathcal{W}}. \tag{2.9}$$

Assumptions 1 and 2 still hold under the subset $\overline{\mathcal{W}}$ so that we can apply our test statistic. Besides, our test statistic in (2.5) should be calculated only over the subset $\overline{\mathcal{W}}$. It is also possible to use a censored quantile regression estimator of Powell (1986) if one wants to allow for a "wrong" quantile regression model outside the subset $\overline{\mathcal{W}}$.

3. Subsampling approximation

As the asymptotic distribution of the test statistic S_T^{CM} in (2.6) under H_0 depends on the data-generating process, we propose a subsampling approach to tabulate critical values. If there were no dynamic misspecification under H_0 in (2.1), we could apply a parametric bootstrap method on $F(y|x, \hat{\theta}_T)$ to obtain critical values. However, if the past information set is dynamically misspecified, $X_t \in \mathcal{F}_{t-1}$, a parametric bootstrap based on resampling values from $F(y|x, \hat{\theta}_T)$ does not guarantee that the long-run variance of the bootstrapped statistic adequately mimics the long-variance of the statistic of the original sample (Corradi and Swanson, 2006); then, the bootstrapped critical values may be asymptotically invalid. Thus, we propose and theoretically justify a subsampling approach.

Subsampling is a resampling method that provides asymptotically valid critical values under general forms of data dependence (Politis et al., 1999). It is a resampling procedure that considers the parameter estimation error effect, is suitable for linear and nonlinear quantile and distributional regression models, allows for dynamic misspecification of the past information set, and is computationally fast. Besides, unreported simulations show that the subsampling method has better finite-sample size and power than a block bootstrap version of the test statistic S_T^{CM} in (2.6). Chernozhukov and Fernández-Val (2005), Whang (2006), and Escanciano and Velasco (2010) developed subsampling methods for specification testing of quantile regression models under dependent data. We extend these approaches by allowing for models indexed by function-valued parameters on the whole conditional distribution.

Subsampling consists of simulating the test statistic S_T^{CM} in (2.6) on small subsamples of size $b << T$ to calculate the critical values of the test. The choice of the subsample size has a significant effect on the result for finite samples (Politis et al., 1999; Sakov and Bickel, 2000). Following Sakov and Bickel (2000), we employ subsamples of sizes $b = [kT^{2/5}]$, where $[\cdot]$ is the floor function and k is a constant.

We propose the following algorithm for computing a subsampling realization of our test statistic S_T^{CM} in (2.6):

1. We generate $B = T - b + 1$ subsamples of size b of the form $\{W_j, \ldots, W_{j+b-1}\} \equiv \{(Y_j, X_j), \ldots, (Y_{j+b-1}, X_{j+b-1})\}$ without replacement from the sample $W_t = (Y_t, X_t)$, with

$b << T$. At each subsample $j << B$, we construct the subsampling version of S_T^{CM} as:

$$S_{j,b}^{CM} = \int \left\| \sqrt{b} S_{j,b}(y, x) \right\|^2 d\hat{Z}_{j,b}(y, x),$$

where $S_{j,b}(y, x)$ is the test statistic of (2.5) computed in the j-th subsample of size b, and $\hat{Z}_{j,b}(y, x) = (1/b) \sum_{t=j}^{j+b-1} \mathbb{1}\{Y_t \leq y\} \mathbb{1}\{X_t \leq x\}$. We calculate the recentered subsampling test statistic as follows:

$$S_{j,b}^{CM*} := S_{j,b}^{CM} - (b/T) S_T^{CM}; \tag{3.1}$$

2. We estimate the conditional distribution of our test statistic in (2.5), $V(s) = P(S_T^{CM} \leq s)$, by $V^*(s) = (1/B) \sum_{j=1}^{B} 1(S_{j,b} \leq s)$. Then, given a significance level $\alpha \in (0, 1)$, the $(1-\alpha)$-th quantile of $V^*(s)$ is the subsampling critical value, $c_T^{CM*}(1-\alpha) \equiv \inf\{s : V^*(s) \geq 1-\alpha\}$. Our test rejects H_0 in (2.1) if $S_T^{CM} > c_T^{CM*}(1 - \alpha)$.

We recenter the subsampling test statistic S_T^{CM*} to ensure a better power performance, following the approaches of Chernozhukov and Fernández-Val (2005), Linton et al. (2005), and Escanciano and Velasco (2010). Another possible choice of the recentering is $\int \| \sqrt{T}(S_{j,b}^{CM} - (b/T)^{1/2} S_T^{CM}) \|^2 d\hat{Z}_{j,b}(y, x)$, but it has delivered worse finite-sample size and power than $S_{j,b}^{CM*}$ in (3.1) in our simulations, and thus we overlook it in the paper.

The following proposition shows that the subsampling test statistic S_T^{CM*} is asymptotically valid and consistent. Besides, it has nontrivial power against local alternatives $H_{A,T}$ in (2.7).

Proposition 3. Under Assumptions 1-2, if the subsample size satisfies $b/T \to 0$ and $b \to \infty$ as $T \to \infty$, then:

(a) Under H_0 in (2.1), $\lim_{T\to\infty} P(S_T^{CM} > c_T^{CM*}(1-\alpha)) = \alpha$;

(b) Under H_A in (2.2), $\lim_{T\to\infty} P(S_T^{CM} > c_T^{CM*}(1-\alpha)) = 1$;

(c) Under the local alternative $H_{A,T}$ in (2.7), $\lim_{T\to\infty} P(S_T^{CM} > c_T^{CM*}(1-\alpha)) \geq \alpha$, where equality holds when the shift function $\Delta(y,x)$ defined in Proposition 2 is trivial.

Proposition 3 illustrates that the subsampling test has asymptotically correct size since $S_{j,b}^{CM*} \xrightarrow{d} \int \|\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x)\|^2 dF_{YX}(y,x)$ as $b \to \infty$ under H_0 in (2.1). Besides, S_T^{CM} diverges (in probability) faster than $c_T^{CM*}(1-\alpha)$ under H_A in (2.2) to ∞ since $\lim_{T\to\infty}(T/b) > 1$ (see Whang, 2006). This ensures the consistency of our subsampling test statistic.

Our subsampling test statistic S_T^{CM*} is also valid for testing the specification of a dynamic conditional quantile or distribution model over a range of the conditional distribution, as in (2.8). In this case, we need to modify the subsampling procedure. In the first step, We generate $B = T - b + 1$ subsamples of size b without replacement only from the subset $\overline{\mathcal{W}}$, rather than from the entire sample W_t. In the second step, we estimate the conditional distribution of our test statistic over these B subsamples, $V^*(s) = (1/B)\sum_{j=1}^{B} 1(S_{j,b} \leq s)$, to obtain the subsampling critical values.

4. Finite-sample performance

To examine the finite-sample performance of our test statistic, we perform simulation experiments with data generating processes (DGPs) under the null and the alternative hypothesis.

The data are generated from the processes below.

$$
\begin{aligned}
&\text{Size.1:} && Y_t = 0.3Y_{t-1} + u_t, \\
&\text{Size.1\_B}: && Y_t = \Phi^{-1}(U_t) + (0.2 + 0.1U_t)\, Y_{t-1},\ \ U_t \overset{i.i.d.}{\sim} U(0,1), \\
&\text{Size.2:} && Y_t = 0.3Y_{t-1} - 0.3Y_{t-2} + u_t, \\
&\text{Power.1:} && \begin{cases} Y_t = 1 + 0.6Y_{t-1} + u_t, & \text{if } Y_{t-1} \le 1, \\ Y_t = 1 - 0.5Y_{t-1} + u_t, & \text{if } Y_{t-1} \ge 1, \end{cases} \\
&\text{Power.2:} && Y_t = 0.8Y_{t-1}u_{t-1} + u_t, \\
&\text{Power.3:} && Y_t = 0.8u_{t-1}^2 + u_t, \\
&\text{Power.4:} && Y_t = 0.05Y_{t-1} + h_t w_t,\ \ h_t^2 = 0.05 + 0.1Y_{t-1}^2 + 0.88h_{t-1}^2, \\
&\text{Power.5:} && Y_t = \mathbb{1}\{Y_{t-1} > 0\} - \mathbb{1}\{Y_{t-1} < 0\} + \sigma u_t,\ \ \sigma = 0.43, \\
&\text{LocalAlt.1:} && Y_t = 0.3Y_{t-1} + cX_{t-1}^2 + u_t,\ X_t = 0.5X_{t-1} + \varepsilon_t,
\end{aligned}
$$

where u_t and ε_t follow an i.i.d process with distribution $\mathcal{N}(0,1)$, and $\{w_t\}$ is a standardized sequence of Student-t innovations with five degrees of freedom. We test the null hypothesis that the conditional quantiles of Y_t follow an AR(1) process:

$$
H_0 : F_{Y_t}^{-1}(\tau | y_{t-1}, \theta_0(\tau)) = \alpha + \beta y_{t-1} + \Phi_u^{-1}(\tau),\ \text{a.s.},
$$

where $\Phi_u^{-1}(\tau)$ is the τ-quantile of the standard Normal error distribution. Size.1 checks the size performance of our test statistic. The QAR(1) model correctly specifies the conditional distribution in Size.1. In addition, Size.1_B verifies the size performance of our test under a stationary QAR(1) model $Q_\tau(Y_t|Y_{t-1}) = \Phi^{-1}(\tau) + (0.2 + 0.1\tau)Y_{t-1}$ of Koenker and Xiao

(2006), in which the linear QAR(1) model has a correct specification, when the influence of Y_{t-1} on the conditional quantile function depends on τ. We allow for dynamic misspecification in Size.2, as $F(y|Y_{t-1}, \theta_0) \neq F(y|Y_{t-1}, Y_{t-2}, \theta_0^*)$ with $\theta_0 \neq \theta_0^*$. DGPs Power.1-Power.5 and LocalAlt.1 evaluate the finite-sample power performance of our test statistic. They have been considered by Hong and Lee (2003) and Escanciano and Velasco (2010).

In these experiments, rejection arises because of misspecification of the conditional distribution model. Power.2 and Power.3 are second-order stationary processes, though they are not invertible (Granger and Andersen, 1978). Power.4 is a common nonlinear model used in the time series literature, the $AR(1) - GARCH(1,1) - t5$ model. Power.5 is the SIGN model analyzed in Granger and Teräsvirta (1999), which is a first-order nonlinear autoregressive process with the same autocorrelation function as an AR(1) process. Finally, DGP LocalAlt.1 analyzes the small-sample performance of our test against Pitman deviations from the null hypothesis that holds when $c = 0$. We consider three different values of $c = \{0.10, 0.20, 0.30\}$.

Our test statistic also verifies the validity of the distributional regression model proposed by Foresi and Peracchi (1995), where the conditional distribution function of Y_t is modeled through a family of binary response models for the event that Y_t exceeds some threshold $y \in \mathbb{R}$ as follows:

$$H_0^{DR} : F(y|x) = \Lambda \left(x'\theta(\tau) \right), \text{for some } \theta(\tau) \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and all } (y, x) \in \mathcal{W}, \tag{4.1}$$

where $\Lambda(\cdot)$ is a known strictly increasing link function (e.g., the logistic or standard normal distribution), and $\theta(\cdot)$ is a function-valued parameter taking values in $\mathcal{B}(\mathcal{T}, \Theta)$. We test the null hypothesis of correct specification of a distributional regression model (4.1) conditioning Y_t on $X_t = Y_{t-1}$, where $\Lambda(\cdot)$ is specified as a standard normal distribution function. Nevertheless, we do not employ the DGP Size.1_B for the distributional regression because it is correctly specified only under a linear QAR(1) model.

For all the experiments, we consider the empirical rejection frequencies for 5% nominal level tests with different sample sizes ($T = 200$ and 300). We use an equally spaced grid of 100 quantiles $\mathcal{T}_n \subset \mathcal{T}$, for $\mathcal{T} = [0.01, 0.99]$, to calculate the test statistics. We perform $1,000$ Monte Carlo repetitions for three different subsample sizes $b = [kT^{2/5}]$, where $[\cdot]$ is the floor function, and $k = \{2, 3, 4\}$. In all the replications, we generated and discarded 200 pre-sample data values. Except for the distributional regression specification test, we compare our results with the test proposed by Escanciano and Velasco (2010) (EV henceforth), based on

$$EV := \int\int \left|(\mathbb{1}(Y_t - m(X_t, \hat{\theta}_T(\tau)) \leq 0) - \tau_j) \exp(i\boldsymbol{x}'X_t)\right|^2 dW(\boldsymbol{x}) d\Phi(\tau), \qquad (4.2)$$

where W and Φ are some integrating measures on $\mathcal{W}$ and $\mathcal{T}$ respectively, $i = \sqrt{-1}$ is the imaginary root, $\exp(i\boldsymbol{x}'X_t)$ is a weighting function with $\boldsymbol{x} \in \mathbb{R}^d$, and $m(X_t, \hat{\theta}_T(\tau))$ is the estimated parametric QAR(1) model for each τ-quantile, for $\tau \in \mathcal{T}$. The critical values of the test (4.2) are obtained by subsampling. We generate $T - b - 1$ subsamples of size b and calculate a subsampling EV statistic in (4.2) for each Monte Carlo replication. We also apply the EV test for three different subsample sizes $b = [kT^{2/5}]$, with $k = \{2, 3, 4\}$.

Table 1 reports the rejection frequencies of the S_T^{CM} test in (2.6) for sample sizes $T = 200$ and $T = 300$. The empirical size of the S_T^{CM} test is generally close to the nominal level under the null hypothesis, whether there is dynamic misspecification (Size.2) or not (Size.1). Moreover, the empirical size of the S_T^{CM} test is close to 5% when the influence of Y_{t-1} on the conditional quantile function depends on τ (Size.1_B). For small values of k and T, the finite-sample size is considerably lower than 5% because the subsample size needs to converge to infinity as $T \to \infty$. This is clear because the subsampling test statistics are draws from the original DGP, and we approximate the critical values with these draws. Conversely, the EV test of Escanciano and Velasco (2010) presents some size distortions for a small sample size of $T = 200$, although it delivers an approximately correct size when $T = 300$. The S_T^{CM} and the EV tests have a similar power, except for DGP Power.4 where the S_T^{CM} test outperforms the EV test. These findings are robust to different sample sizes and subsample choices. In addition, Table 1 shows that the S_T^{CM} and the EV tests display nontrivial power against Pitman deviations for the null hypothesis. Under DGP LocalAlt.1, our test has a better finite-sample power for $k = \{3, 4\}$, while it is outperformed when $k = 2$.

Table 2 presents the empirical rejection frequencies for the distributional regression specification of the S_T^{CM} test for sample sizes $T = 200$ and $T = 300$. The finite-sample size of the S_T^{CM} test is close to the nominal level for a distributional regression model. Besides, the S_T^{CM} allows for dynamic misspecification (Size.2). These results are robust to different sample sizes and subsample choices. Our test statistic is also powerful against misspecifications in the distributional regression model, as the power for testing H_0^{DR} in (4.1) is high for all power DGPs and for both sample sizes. Further, the S_T^{CM} test also presents nontrivial power against local deviations from the null hypothesis for the distributional regression specification. In sum, our results illustrate that the S_T^{CM} test has good finite-sample properties.

5. Empirical application: Value-at-Risk and distributional regression models

Many papers have proposed methods to check the specification of models for Value-at-Risk (VaR). The outcome of a VaR model determines the multiplication factors for market risk capital requirements of all regulated financial institutions. An inaccurate VaR model provides an underestimated multiplicative factor that delivers an insufficient reserve of capital risk. Therefore, the specification of VaR models is crucial for risk managers, regulators, and financial institutions.

Since the VaR is a quantile of the portfolio returns, conditional on past information, and as the distribution of portfolio returns evolves over time, it is challenging to model time-varying conditional quantiles. An accurate VaR model satisfies $P(Y_t \leq -VaR_t|\mathcal{F}_{t-1}) = \tau$, for a portfolio return series Y_t, a past information set $\mathcal{F}_{t-1}$, and a quantile $\tau \in (0,1)$. The conditional quantile regression approach specifies a conditional VaR model using only the relevant past information that influences the quantiles of interest (Chernozhukov and Umantsev, 2001; Engle and Manganelli, 2004; Escanciano and Olmo, 2010).

To highlight the applicability of our approach, we test different specifications of conditional quantile regression models for estimating the VaR of stock returns. We estimate the VaR of the returns of two major stock indexes, the Frankfurt Dax Index (DAX) and the London FTSE-100 Index (FTSE-100). The DAX and the FTSE-100 daily stock indexes are two representatives of the data for which linear and non-linear quantile regression models have been widely used, see e.g. Escanciano and Velasco (2010), Iqbal and Mukherjee (2012), and Jeon and Taylor (2013).

Our dataset consists of 2,981 daily observations (from January 2003 to June 2014) on Y_t, the one-day returns, and X_t, the lagged returns $(Y_{t-1}, \ldots, Y_{t-p})$. Table 3 presents the summary statistics of the series. Both log-returns series are highly leptokurtic and present autocorrelation. Figure 1 displays the daily log-return series of the two series. It shows that both log-return series display calm as well as volatile periods and single outlying log-return observations.

PLEASE INSERT FIGURE 1 HERE

We estimate a conditional quantile autoregressive model for the VaR, i.e., the CAViaR model of Engle and Manganelli (2004). As the CAViaR is an auto-regressive model of the quantile, it avoids distributional assumptions. Besides, it is intuitively appealing since the financial returns usually display volatility clustering (Jeon and Taylor, 2013). The CAViaR model has also obtained more success than other models in predicting the VaR risk measure for various periods (Bao et al., 2006; Yu et al., 2010). We test the hypothesis H_0: the VaR of the log-return Y_t follows a CAViaR specification. We test the four CAViaR models introduced by Engle and Manganelli (2004):

Symmetric absolute value (SAV) CAViaR:

$$Q_t(\tau) = \beta_1 + \beta_2 Q_{t-1}(\tau) + \beta_3 |Y_{t-1}|.$$

Asymmetric slope (AS) CAViaR:

$$Q_t(\tau) = \beta_1 + \beta_2 Q_{t-1}(\tau) + \beta_3 (Y_{t-1})^+ + \beta_4 (Y_{t-1})^-.$$

Indirect GARCH(1,1) CAViaR:

$$Q_t(\tau) = (1 - 2\mathbb{1}\{\tau < 0.5\}) \left(\beta_1 + \beta_2 Q_{t-1}(\tau)^2 + \beta_3 Y_{t-1}^2\right)^{\frac{1}{2}}.$$

Adaptive CAViaR:

$$Q_t(\tau) = Q_{t-1}(\tau) + \beta_1 \left\{\left[1 + \exp\left(G\left(Y_{t-1} - Q_{t-1}(\tau)\right)\right)\right]^{-1} - \tau\right\},$$

where $Q_t(\tau)$ is the τ-quantile of Y_t conditional on X_t, $F_{Y_t}^{-1}(\tau|X_t)$, $(Y_{t-1})^+ = \max(Y_{t-1}, 0)$, $(Y_{t-1})^- = -\min(Y_{t-1}, 0)$, and G is some positive finite number. We evaluate the correct specification of the previous VaR models for an equally spaced grid of 10 quantiles $\{\tau_i\}_{i=1}^{i=10}$ from $\tau_1 = 0.01$ to $\tau_{10} = 0.10$. Thus, we can cover the VaR backtesting region suggested by the Basel Committee on Banking Supervision (1996). We apply our S_T^{CM} test in (2.6) using the subsampling method. As we use only the observations between $\tau_1 = 0.01$ to $\tau_{10} = 0.10$, we have a sample size of $T = 298$. We employ the subsamples of sizes $b = [kT^{2/5}]$, where $[\cdot]$ is the floor function., for $k = \{2, 3, 4\}$. For comparison purposes, we apply the EV test with the same subsample sizes.

Figure 2 displays the VaR forecasts and violations of the Asymmetric Slope (AS) and Symmetric Absolute Value (SAV) CAViaR models for $\tau = 0.05$. There seems to be no difference in the 5%-VaR-forecast of the AS and SAV models. Besides, VaR-violations occur more frequently during high volatile periods.

PLEASE INSERT FIGURE 2 HERE

Table 4 reports the p-values of the specification tests of the CAViaR models. For the DAX series, our S_T^{CM} test rejects the specifications of all CAViaR models at the 5% significance level. These results are robust to three different subsample choices. Conversely, the EV test of Escanciano and Velasco (2010) does not reject the AS and the Adaptive specification at the 5% significance level. For the FTSE-100 series, the S_T^{CM} test also rejects the specifications of all CAViaR models at the 5% significance level; conversely, the EV test does not reject the SAV, AS, and Adaptive specifications at the 5% significance level. Therefore, our results indicate that CAViaR models are not able to fit the tails of these stock returns series appropriately. Our approach is useful for risk managers and financial institutions to detect models that underestimate risk, and thus calculate correct capital risk requirements.

We present an additional application where we test the specification of dynamic distributional regression models for the tails of the DAX and FTSE-100 daily log-returns, as in equation (4.1). We test the hypothesis H_0^{DR} in (4.1) for four different lag specifications of the distributional regression model. We specify $\Lambda(\cdot)$ as a standard normal distribution function, but other link functions can be applied since the distributional regression model is flexible. Moreover, Chernozhukov et al. (2013) show that the choice of the link function is irrelevant for a sufficiently rich information set X_t.

Table 5 reports the p-values of the specification tests of the distributional regression models. For the DAX series, our S_T^{CM} test rejects all four proposed specifications of a distributional regression AR model at the 5% significance level. These results are robust to three different subsample sizes. For the FTSE-100 series, the S_T^{CM} test does not reject the AR(4) specification of distributional regression models at the 5% significance level.

For comparison purposes, we also test the specification of distributional regression models on the whole distribution of the log returns of DAX and FTSE-100 indices. We test the hypothesis H_0^{DR} in (4.1) for an equally spaced grid of 100 quantiles. We also specify $\Lambda(\cdot)$ as a standard normal distribution function. Table 6 displays the p-values of the specification tests of distributional regression models over the entire distribution. The S_T^{CM} test fails to reject the AR(4) specification of distributional regression models for the whole distribution of the DAX log-returns at the 5% significance level. In line with the results of Table 5, the S_T^{CM} test also does not reject the AR(4) specification for the FTSE-100 log-returns on the entire distribution.

6. Conclusions

In this paper, we present a practical and consistent specification test of conditional quantile and distributional regression models in a general setting. Our approach covers conditional distribution models indexed by function-valued parameters, allowing for a wide range of applications in economics and finance, such as the quantile autoregressive, the CAViaR, and the distributional regression models. Our proposed test statistic has the correct asymptotic size, is consistent against fixed alternatives, and has nontrivial power against Pitman deviations from the null hypothesis. In addition, our approach has the correct asymptotic size under dynamic misspecification of the past information set.

As the proposed test statistic is non-pivotal, we propose and theoretically justify a subsampling approach to tabulate critical values. Finite-sample experiments suggest that our proposed test has good finite-sample size and power. Besides, based on the evaluation of linear and nonlinear models for the returns distribution and VaR of two major European stock indexes, we highlight the ability of our test to detect possibly misspecified CAViaR and distributional regression models. Therefore, our method is useful for risk managers and financial institutions to obtain the correct multiplicative factors for their market risk capital requirements.

A possible direction for future work is to extend our approach to the class of multivariate models. Further research could focus on specification tests for vector autoregressive and multivariate linear and non-linear quantile models, extending the procedures of Francq and Raïssi (2007), Escanciano et al. (2013), Chen and Hong (2014), and Kheifets (2018).

References

Andersen, T. G., Bollerslev, T., Huang, X. (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics* 160:176–189.

Anderson, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proceedings of the American Mathematical Society* 6:170–176.

Andrews, D. W. K. (1997). A conditional Kolmogorov test. *Econometrica* 65:1097–1128.

Andrews, D. W. K., Pollard, D. (1994). An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review* 62:119–132.

Arcones, M. A., Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* 7:47–71.

Bai, J. (2003). Testing parametric conditional distributions of dynamic models. *Review of Economics and Statistics* 85:531–549.

Bao, Y., Lee, T.-H., Saltoglu, B. (2006). Evaluating predictive performance of value-at-risk models in emerging markets: A reality check. *Journal of Forecasting* 25:101–128.

Basel Committee on Banking Supervision (1996). Amendment to the capital accord to incorporate market risks. Technical report, Bank for International Settlements.

Bierens, H. J., Ginther, D. K. (2001). Integrated conditional moment testing of quantile regression models. *Empirical Economics* 26:307–324.

Bierens, H. J., Wang, L. (2012). Integrated conditional moment tests for parametric conditional distributions. *Econometric Theory* 28:328–362.

Bierens, H. J., Wang, L. (2017). Weighted simulated integrated conditional moment tests for parametric conditional distributions of stationary time series processes. *Econometric Reviews* 36:103–135.

Billingsley, P. (1995). *Probability and Measure*. Hoboken, NJ: Wiley.

Bollerslev, T., Li, S. Z., Todorov, V. (2016). Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns. *Journal of Financial Economics* 120:464–490.

Box, G. E. P., Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* 65:1509–1526.

Candelon, B., Colletaz, G., Hurlin, C., Tokpavi, S. (2011). Backtesting Value-at-Risk: A GMM duration-based test. *Journal of Financial Econometrics* 9:314–343.

Chen, B., Hong, Y. (2014). A unified approach to validating univariate and multivariate conditional distribution models in time series. *Journal of Econometrics* 178:22–44.

Chernozhukov, V., Fernández-Val, I. (2005). Subsampling inference on quantile regression processes. *Sankhyā: The Indian Journal of Statistics* 67:253–276.

Chernozhukov, V., Fernández-Val, I., Melly, B. (2013). Inference on counterfactual distributions. *Econometrica* 81:2205–2268.

Chernozhukov, V., Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics* 26:271–292.

Christoffersen, P., Pelletier, D. (2004). Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics* 2:84–108.

Corradi, V., Swanson, N. R. (2006). Bootstrap conditional distribution tests in the presence of dynamic misspecification. *Journal of Econometrics* 133:779–806.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B* 34:187–220.

Delgado, M. A., Stute, W. (2008). Distribution-free specification tests of conditional models. *Journal of Econometrics* 143:37–55.

Du, Z., Escanciano, J. C. (2016). Backtesting expected shortfall: Accounting for tail risk. *Management Science* 63:940–958.

Dudley, R. M. (1978). Central limit theorems for empirical measures. *The Annals of Probability* 6:899–929.

Engle, R., Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* 22:367–381.

Escanciano, J. C., Goh, S. C. (2014). Specification analysis of linear quantile models. *Journal of Econometrics* 178:495–507.

Escanciano, J. C., Lobato, I. N., Zhu, L. (2013). Automatic specification testing for vector autoregressions and multivariate nonlinear time series models. *Journal of Business and Economic Statistics* 31:426–437.

Escanciano, J. C., Olmo, J. (2010). Backtesting parametric value-at-risk with estimation risk. *Journal of Business and Economic Statistics* 28:36–51.

Escanciano, J. C., Velasco, C. (2010). Specification tests of parametric dynamic conditional quantiles. *Journal of Econometrics* 159:209–221.

Fan, Y., Li, Q., Min, I. (2006). A nonparametric bootstrap test of conditional distributions. *Econometric Theory* 22:587–613.

Foresi, S., Peracchi, F. (1995). The conditional distribution of excess returns: An empirical analysis. *Journal of the American Statistical Association* 90:451–466.

Francq, C., Raïssi, H. (2007). Multivariate portmanteau test for autoregressive models with uncorrelated but nonindependent errors. *Journal of Time Series Analysis* 28:454–470.

Granger, C. W. J., Andersen, A. P. (1978). *An Introduction to Bilinear Time Series Models.* Vandenhoeck and Ruprecht Göttingen.

Granger, C. W. J., Teräsvirta, T. (1999). A simple nonlinear time series model with misleading linear properties. *Economics Letters* 62:161–165.

He, X., Zhu, L.-X. (2003). A lack-of-fit test for quantile regression. *Journal of the American Statistical Association* 98:1013–1022.

Hong, Y., Lee, T.-H. (2003). Diagnostic checking for adequacy of nonlinear time series models. *Econometric Theory* 19:1065–1121.

Horowitz, J. L., Spokoiny, V. G. (2002). An adaptive, rate-optimal test of linearity for median regression models. *Journal of the American Statistical Association* 97:822–835.

Iqbal, F., Mukherjee, K. (2012). A study of value-at-risk based on M-estimators of the conditional heteroscedastic models. *Journal of Forecasting* 31:377–390.

Jeon, J., Taylor, J. W. (2013). Using CAViaR models with implied volatility for value-at-risk estimation. *Journal of Forecasting* 32:62–74.

Kheifets, I. L. (2015). Specification tests for nonlinear dynamic models. *The Econometrics Journal* 18:67–94.

Kheifets, I. L. (2018). Multivariate specification tests based on a dynamic Rosenblatt transform. *Computational Statistics & Data Analysis* 124:1–14.

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

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

Koenker, R., Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica* 70:1583–1612.

Koenker, R., Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association* 101:980–990.

Kosorok, M. R. (2007). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Science and Business Media.

Koul, H. L., Stute, W. (1999). Nonparametric model checks for time series. *The Annals of Statistics* 27:204–236.

Li, F., Tkacz, G. (2006). A consistent bootstrap test for conditional density functions with time-series data. *Journal of Econometrics* 133:863–886.

Li, F., Tkacz, G. (2011). A consistent test for multivariate conditional distributions. *Econometric Reviews* 30:251–273.

Linton, O., Maasoumi, E., Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *The Review of Economic Studies* 72:735–765.

Neumann, M. H., Paparoditis, E. (2008). Goodness-of-fit tests for Markovian time series models: Central limit theory and bootstrap approximations. *Bernoulli* 14:14–46.

Politis, D., Romano, J., Wolf, M. (1999). *Subsampling.* New York: Springer-Verlak.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

Powell, J. L. (1986). Censored regression quantiles. *Journal of Econometrics* 32:143–155.

Radulović, D. (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Processes and Their Applications* 65:259–279.

Rothe, C., Wied, D. (2013). Misspecification testing in a class of conditional distributional models. *Journal of the American Statistical Association* 108:314–324.

Sakov, A., Bickel, P. (2000). An Edgeworth expansion for the m out of n bootstrapped median. *Statistics and Probability Letters* 49:217–223.

Taamouti, A., Bouezmarni, T., El Ghouch, A. (2014). Nonparametric estimation and inference for conditional density based Granger causality measures. *Journal of Econometrics* 180:251–264.

Van der Vaart, A., Wellner, J. (2000). *Weak Convergence and Empirical Processes: With Applications to Statistics.* New York: Springer.

Whang, Y.-J. (2006). Consistent specification testing for quantile regression models. In: D. Corbae, S. N. Durlauf, and B. E. Hansen (Eds.)*Frontiers of Analysis and Applied Research: Essays in Honor of Peter C. B. Phillips.* Cambridge University Press, pp. 288–310.

Wied, D., Weiß, G. N. F., Ziggel, D. (2016). Evaluating Value-at-Risk forecasts: A new set of multivariate backtests. *Journal of Banking & Finance* 72:121–132.

Yu, P. L., Li, W. K., Jin, S. (2010). On some models for value-at-risk. *Econometric Reviews* 29:622–641.

Zheng, J. X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* 14:123–138.

Zheng, J. X. (2000). A consistent test of conditional parametric distributions. *Econometric Theory* 16:667–691.

Table 1. Empirical rejection frequencies

	S_T^{CM}			EV		
	$k = 2$	$k = 3$	$k = 4$	$k = 2$	$k = 3$	$k = 4$
			$T = 200$			
Size:						
Size.1	0.038	0.064	0.065	0.086	0.096	0.078
Size.1_B	0.031	0.053	0.071	0.079	0.078	0.068
Size.2	0.040	0.052	0.075	0.074	0.070	0.044
Power:						
Power.1	0.995	0.995	1.000	0.996	0.996	0.994
Power.2	1.000	1.000	1.000	1.000	1.000	1.000
Power.3	1.000	1.000	1.000	1.000	1.000	0.994
Power.4	0.450	0.490	0.490	0.212	0.204	0.230
Power.5	0.867	0.890	0.870	0.868	0.868	0.854
Local Alternatives:						
LocalAlt.1 ($c = 0.10$)	0.062	0.080	0.114	0.080	0.065	0.066
LocalAlt.1 ($c = 0.20$)	0.067	0.100	0.133	0.106	0.088	0.071
LocalAlt.1 ($c = 0.30$)	0.139	0.164	0.200	0.171	0.149	0.129
			$T = 300$			
Size:						
Size.1	0.043	0.044	0.060	0.070	0.065	0.059
Size.1_B	0.035	0.051	0.054	0.077	0.069	0.064
Size.2	0.036	0.052	0.048	0.038	0.042	0.052
Power:						
Power.1	1.000	1.000	1.000	1.000	1.000	1.000
Power.2	1.000	1.000	1.000	1.000	1.000	1.000
Power.3	1.000	1.000	1.000	1.000	1.000	1.000
Power.4	0.720	0.676	0.680	0.296	0.302	0.322
Power.5	0.956	0.950	0.967	0.958	0.958	0.946
Local Alternatives:						
LocalAlt.1 ($c = 0.10$)	0.060	0.104	0.130	0.073	0.065	0.063
LocalAlt.1 ($c = 0.20$)	0.097	0.135	0.143	0.124	0.098	0.084
LocalAlt.1 ($c = 0.30$)	0.192	0.250	0.231	0.244	0.202	0.181

Notes: S_T^{CM} denotes our test statistic, and EV is the subsampling specification test of Escanciano and Velasco (2010), with subsamples of sizes $b = [kT^{2/5}]$, where $[\cdot]$ is the floor function. We perform 1,000 Monte Carlo repetitions, and we use an equally spaced grid of 100 quantiles $\mathcal{T}_n$, $\mathcal{T}_n \subset [0.01, 0.99]$, to calculate the test statistics.

Table 2. Empirical rejection frequencies: Distributional regression specification

	S_T^{CM}		
	$k=2$	$k=3$	$k=4$
	$T=200$		
Size:			
Size.1	0.018	0.044	0.070
Size.2	0.023	0.040	0.076
Power:			
Power.1	1.000	1.000	1.000
Power.2	1.000	1.000	1.000
Power.3	1.000	1.000	1.000
Power.4	0.460	0.440	0.600
Power.5	0.720	0.700	0.840
Local Alternatives:			
LocalAlt.1 ($c=0.10$)	0.040	0.052	0.085
LocalAlt.1 ($c=0.20$)	0.044	0.061	0.107
LocalAlt.1 ($c=0.30$)	0.060	0.068	0.133
	$T=300$		
Size:			
Size.1	0.032	0.046	0.055
Size.2	0.028	0.043	0.060
Power:			
Power.1	1.000	1.000	1.000
Power.2	1.000	1.000	1.000
Power.3	1.000	1.000	1.000
Power.4	0.750	0.740	0.780
Power.5	0.900	0.880	0.940
Local Alternatives:			
LocalAlt.1 ($c=0.10$)	0.043	0.058	0.075
LocalAlt.1 ($c=0.20$)	0.054	0.090	0.117
LocalAlt.1 ($c=0.30$)	0.095	0.100	0.138

Notes: S_T^{CM} denotes our test statistic with subsamples of sizes $b=[kT^{2/5}]$, where $[\cdot]$ is the floor function. We test the null hypothesis H_0^{DR} of a correct specification of a distributional regression model specified in (4.1). We perform 1,000 Monte Carlo repetitions.

Table 3. Summary statistics: DAX and FTSE-100 daily log-returns

	DAX	FTSE-100
Mean	0.02	0.01
Std. Dev.	0.61	0.51
Median	0.03	0.01
Skewness	0.01	-0.12
Kurtosis	9.14	11.71
Minimum	-3.23	-4.02
Maximum	4.69	4.08
Autocorrelation	-0.01	-0.06
LB(10)	21.34	62.35

Notes: The Autocorrelation is the first-order autocorrelation coefficient, and $LB(10)$ denotes the Ljung-Box Q-statistic of order 10.

Table 4. Specification tests: p-values

	S_T^{CM}			EV		
	$b = 49$	$b = 73$	$b = 98$	$b = 49$	$b = 73$	$b = 98$
DAX						
SAV CAViaR	0.001	0.001	0.001	0.048	0.051	0.042
AS CAViaR	0.001	0.001	0.001	0.272	0.166	0.184
Indirect GARCH(1,1) CAViaR	0.001	0.001	0.001	0.001	0.001	0.001
Adaptive CAViaR	0.001	0.001	0.001	0.206	0.172	0.170
FTSE-100						
SAV CAViaR	0.001	0.001	0.001	0.777	0.633	0.601
AS CAViaR	0.001	0.001	0.001	0.134	0.100	0.083
Indirect GARCH(1,1) CAViaR	0.001	0.001	0.001	0.001	0.001	0.001
Adaptive CAViaR	0.001	0.001	0.001	0.083	0.062	0.055

Notes: S_T^{CM} denotes our test statistic and EV is the subsampling specification test of Escanciano and Velasco (2010) with subsamples of sizes $b = \{49, 73, 98\}$. We test the null hypothesis H_0 of the correct specification of the CAViaR model. We use an equally spaced grid of 10 tail quantiles $\mathcal{T}_n$, $\mathcal{T}_n \subset [0.01, 0.10]$, to calculate the test statistics.

Table 5. Specification tests p-values: Distributional regression models for the lower tail

	S_T^{CM} $(b = 49)$	S_T^{CM} $(b = 73)$	S_T^{CM} $(b = 98)$
DAX			
AR(1)	0.001	0.001	0.001
AR(2)	0.001	0.001	0.001
AR(3)	0.001	0.001	0.001
AR(4)	0.012	0.001	0.001
FTSE-100			
AR(1)	0.001	0.001	0.001
AR(2)	0.001	0.001	0.001
AR(3)	0.001	0.001	0.001
AR(4)	0.728	0.478	0.358

Notes: S_T^{CM} denotes our proposed test statistic with subsamples of sizes $b = \{49, 73, 98\}$. We test the null hypothesis H_0^{DR} of a correct specification of a distributional regression model specified in (4.1) for the lower tail of the distributions of the DAX and FTSE-100 daily log-returns. We use an equally spaced grid of 10 tail quantiles $\mathcal{T}_n$, $\mathcal{T}_n \subset [0.01, 0.10]$, to calculate the test statistics.

Table 6. Specification tests p-values: Distributional regression models for the whole distribution

	S_T^{CM} $(b = 49)$	S_T^{CM} $(b = 73)$	S_T^{CM} $(b = 98)$
DAX			
AR(1)	0.001	0.001	0.001
AR(2)	0.001	0.003	0.006
AR(3)	0.020	0.013	0.014
AR(4)	0.423	0.271	0.199
FTSE-100			
AR(1)	0.001	0.001	0.001
AR(2)	0.001	0.001	0.001
AR(3)	0.001	0.001	0.001
AR(4)	0.294	0.167	0.141

Notes: S_T^{CM} denotes our proposed test statistic subsamples of sizes $b = \{49, 73, 98\}$. We test the null hypothesis H_0^{DR} of a correct specification of a distributional regression model specified in (4.1) for the whole distribution of the DAX and FTSE-100 daily log-returns. We use an equally spaced grid of 100 quantiles $\mathcal{T}_n$, $\mathcal{T}_n \subset [0.01, 0.99]$, to calculate the test statistics.

Appendix

A.1. Preliminary Results

In this subsection, we provide preliminary results used in the proofs of the propositions. Let $\mathcal{G}$ be a permissible class of functions in such a way that the following holds: (a) $\mathcal{T}$ is a Suslin metric space (a Hausdorff topological space that is the continuous image of a Polish space) with Borel σ-field $\mathcal{B}(\mathcal{T}, \Theta)$, and (b) $g(\cdot, \cdot, \cdot)$ is a $\mathcal{B}(\mathcal{T}, \Theta)$-measurable function from $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^K$ to $\mathbb{R}$ (see Kosorok, 2007, Section 11.6). Let $E_Q g = \int g(W_t, \theta, \tau) dQ(W_t, \theta, \tau)$, for $g \in \mathcal{G}$, with $\mathcal{G} := \{W_t \mapsto g(W_t, \theta, \tau) : \theta \in \Theta, \tau \in \mathcal{T}\}$. We assume that the $\mathcal{G}$ class of functions forms a so-called Vapnik-Chervonenkis subgraph (VC-subgraph) class of functions (see Dudley, 1978). The VC-subgraph class is an extension of the class of indicator functions and is useful for most statistical applications (Arcones and Yu, 1994; Radulović, 1996). If $\mathcal{G}$ is a VC-subgraph class, then for any given $1 \leq p < \infty$, there are constants a and b satisfying

$$N(\varepsilon, \mathcal{G}, \|\cdot\|) \leq a \left(\frac{(E_Q |\mathbb{F}|^p)^{1/p}}{\varepsilon} \right)^b,$$

for all $\varepsilon > 0$ and all probability measures Q, with $E_Q |\mathbb{F}|^p < \infty$, where $N(\varepsilon, \mathcal{G}, \|\cdot\|)$ is the covering number of $\mathcal{G}$ with respect to $\|\cdot\|$, i.e., the minimal number of $L_2(Q)$-balls of radius ε needed to cover $\mathcal{G}$, where a $L_2(Q)$-ball of radius ε around a function $g \in L_2(Q)$ is the set $\{h \in L_2(Q) : \|h - g\| < \varepsilon\}$ (see Pollard, 1984). Moreover, the class of functions $\mathcal{G}$ has a finite and integrable envelope function $\mathbb{F} := \sup_{g \in \mathcal{G}} |g(W_t, \theta, \tau)|$, and it can be covered by a finite number of elements, not necessarily in $\mathcal{G}$.

The following result establishes a central limit theorem for strong mixing processes for the empirical distribution, $\hat{Z}_T(y, x)$, under the null and the alternative hypothesis.

Lemma A.1. If Assumption 1 holds, under H_0 of (2.1) or H_A of (2.2),

$$v_T(y,x) := \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x)) \Rightarrow \mathbb{H}_1(y,x), \text{ in } \ell^\infty(\mathcal{W}),$$

where $\mathbb{H}_1$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$ with covariance function

$$\mathrm{Cov}(\mathbb{H}_1(y,x), \mathbb{H}_1(y',x')) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(\mathbb{1}\{Y_0 \le y\}\mathbb{1}\{X_0 \le x\}, \mathbb{1}\{Y_k \le y'\}\mathbb{1}\{X_k \le x'\}).$$

Proof: Parts (a) and (b) of Assumption 1 imply the conditions (2.3) and (2.4) in Theorem 2.1 in Arcones and Yu (1994), respectively. Then the results follow from a direct application of Theorem 2.1 in Arcones and Yu (1994). $\square$

The following result establishes a functional delta method for the empirical analog $\hat{G}(\theta,\tau)$ of the moment conditions in (2.4) and for a consistent estimator of the function-valued parameter $\hat{\theta}_T(\cdot)$.

Lemma A.2. Let $v_T(y,x) := \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ be the empirical process of Lemma A.1 and define the empirical process $r_T(\theta,\tau) := \sqrt{T}(\hat{G}(\theta,\tau) - G(\theta,\tau))$. If Assumption 1 is satisfied, under H_0 of (2.1) or H_A of (2.2),

$$\begin{aligned}
(v_T(y,x), r_T(\theta,\tau)) &\Rightarrow (\mathbb{H}_1(y,x), \tilde{\mathbb{H}}_2(\theta,\tau)), \text{ in } \ell^\infty(\mathcal{W}\times\Theta\times\mathcal{T}),\\
\sqrt{T}(\hat{\theta}_T(\cdot) - \theta_0(\cdot)) &\Rightarrow -\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot),\cdot)) \text{ in } \ell^\infty(\mathcal{T}),
\end{aligned}$$

where $\tilde{\mathbb{H}}_2$ is a tight mean zero Gaussian process with covariance function

$$\mathrm{Cov}(\tilde{\mathbb{H}}_2(\theta,\tau), \tilde{\mathbb{H}}_2(\theta',\tau')) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(g(W_0,\theta,\tau), g(W_k,\theta',\tau')).$$

Proof: By Lemma E.1 in Chernozhukov et al. (2013), Assumption 1 implies that (a) the inverse of $G(\cdot,\tau)$ defined as $G^{-1}(x,\tau) := \{\theta \in \Theta : G(\theta,\tau) = x\}$ is continuous at $x = 0$

uniformly in $\tau \in \mathcal{T}$ with respect to the Hausdorff distance, (b) there exists $\dot{G}_{\theta_0,\tau}$ such that

$$\lim_{t\to 0} \sup_{\tau\in\mathcal{T},\|h\|=1} |t^{-1}(G(\theta_0(\tau)+th,\tau) - G(\theta_0(\tau),\tau)) - \dot{G}_{\theta_0,\tau}h| = 0,$$

where $\inf_{\tau\in\mathcal{T}} \inf_{\|h\|=1} \|\dot{G}_{\theta_0,\tau}h\| > 0$, (c) the maps $\tau \mapsto \theta_0(\tau)$ and $\tau \mapsto \dot{G}_{\theta_0,\tau}$ are continuous, and (d) the mapping $\tau \mapsto \theta_0(\tau)$ is continuously differentiable. Under the previous conditions, Lemma E.2 in Chernozhukov et al. (2013) holds, and the process $r_T(\theta,\tau)$ weakly converges to $\tilde{\mathbb{H}}_2(\theta,\tau)$ in $\ell^\infty(\Theta\times\mathcal{T})$ and the map $\theta \mapsto G(\theta,\cdot)$ is Hadamard differentiable at θ_0 with continuously invertible derivative $\dot{G}_{\theta_0,\cdot}$. By Hadamard differentiability of the map $\theta \mapsto G(\theta,\cdot)$, it follows the weak convergence of the process $\sqrt{T}(\hat{\theta}_T(\cdot)-\theta_0(\cdot))$ in $\ell^\infty(\mathcal{T})$.

□

Lemma A.3. Let $v_T(y,x) := \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ be the empirical process of Lemma A.1 and define the empirical process $v_T^{\theta_0}(y,x) := \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F(y,x,\theta_0))$. If Assumption 1 holds, under H_0 of (2.1) or H_A of (2.2),

$$(v_T(y,x), v_T^{\theta_0}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x)) \text{ in } \ell^\infty(\mathcal{W}\times\mathcal{W}),$$

where $\mathbb{H}_2$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$.

Proof: By Lemma A.2, $\sqrt{T}(\hat{\theta}_T(\cdot)-\theta_0(\cdot)) \Rightarrow -\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot),\cdot))$ in $\ell^\infty(\mathcal{T})$, where $\tilde{\mathbb{H}}_2$ is a tight mean zero Gaussian process. Similarly to Lemma A.1, under H_0 of (2.1) or H_A of (2.2), if parts (a)-(b) of Assumption 1 hold, then $\sqrt{T}(\hat{F}_X(x^*) - F_X(x^*))$ weakly converges to a tight mean zero Gaussian process. Now, let the measurable functions $\Gamma : \mathcal{W} \mapsto [0,1]$ be defined by $(y,x) \mapsto \Gamma(y,x)$, and the bounded maps $\Pi : \mathcal{H} \mapsto \mathbb{R}$ be defined by $f \mapsto \int f d\Pi$. Then it follows from Lemma D.1 in Chernozhukov et al. (2013) that the mapping $(\Gamma,\Pi) \mapsto \int \Gamma(\cdot,x)d\Pi(x)$, with $\Gamma(\cdot,x) = \mathbb{1}\{x^* \leq x\}F(\cdot|x)$ and $\Pi = F_X(\cdot)$, is well defined and Hadamard differentiable at (Γ, Π). Under H_0 of (2.1) or H_A of (2.2), we can write $\hat{F}_T(y, x, \hat{\theta}_T) = \int F(y|x^*, \hat{\theta}_T)\mathbb{1}\{x^* \leq x\}d\hat{F}_X(x^*)$ and $F(y, x, \theta_0) = \int F(y|x^*)\mathbb{1}\{x^* \leq x\}dF_X(x^*)$. Then, by the functional delta method from Lemma B.1 of Chernozhukov et al. (2013), it follows that

$$\begin{aligned}
\sqrt{T}(\hat{F}_T(y, x, \hat{\theta}_T) - F(y, x, \theta_0)) &= \int \sqrt{T}\left[F(y|x^*, \hat{\theta}_T) - F(y|x^*)\right]\mathbb{1}\{x^* \leq x\}dF_X(x^*) \\
&+ \int F(y|x^*)\mathbb{1}\{x^* \leq x\}\sqrt{T}d\left(\hat{F}_X(x^*) - F_X(x^*)\right) + o_p(1).
\end{aligned}$$

Using the same arguments of the Proof of Lemma A.2, we can show that the map $\theta \mapsto F(\cdot|\cdot, \theta(\cdot))$ is Hadamard differentiable. Thus, we apply the functional delta method, for fixed y and x, as follows:

$$\begin{aligned}
\sqrt{T}\left(F(y|x, \hat{\theta}_T) - F(y|x)\right) &\Rightarrow -\dot{F}(y|x, \theta_0)(-\dot{G}^{-1}_{\theta_0, \cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot), \cdot))) \\
&:= \mathbb{H}^*_2(y, x) \text{ in } \ell^\infty(\mathcal{W}).
\end{aligned}$$

Given the Hadamard differentiability of the mapping $(\Gamma, \Pi) \mapsto \int \Gamma(\cdot, x)d\Pi(x)$, the result follows from an application of the functional delta method, where the Gaussian process $\mathbb{H}_2$ is given by

$$\mathbb{H}_2(y, x) := \int \mathbb{H}^*_2(y, x^*)\mathbb{1}\{x^* \leq x\}dF_X(x^*) + \int F(y|x^*)\mathbb{1}\{x^* \leq x\}d\mathbb{H}_1(\infty, x^*),$$

where $\mathbb{H}_1$ is the tight mean zero Gaussian process defined in Lemma A.1. $\square$

Lemma A.4. Under the local alternatives $H_{A,T}$ in (2.7) and Assumptions 1-2, let $F^A_T(y, x) = \int F_T(y|x^*)\mathbb{1}\{x^* \leq x\}dF_X(x^*)$ and $G_{F_T}(\theta, \tau) = E_{F_T}[g(W_t, \theta, \tau)]$, then:

$$\begin{pmatrix} \sqrt{T}\left(\hat{Z}_T(y,x) - F_T^A(y,x)\right) \\ \sqrt{T}\left(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau)\right) \end{pmatrix} \Rightarrow \begin{pmatrix} \mathbb{H}_1(y,x) \\ \tilde{\mathbb{H}}_2(\theta,\tau) \end{pmatrix} \text{ in } \ell^\infty(\mathcal{W}\times\Theta\times\mathcal{T}),$$

where $(\mathbb{H}_1, \tilde{\mathbb{H}}_2)$ are the tight mean zero Gaussian processes defined in Lemmas A.1-A.2.

Proof: Under Assumption 2, $F_T^A(y,x)$ is contiguous to $F(y,x,\theta_0)$, then under the sequence of local alternatives $H_{A,T}$ in (2.7) and Assumptions 1-2, $F_T(y|x)$ of (2.7) is a linear combination of two measures that are VC-subgraph class with a p-integrable envelope, for some $2 < p < \infty$. From an application of Lemma 2.8.7 in Van der Vaart and Wellner (2000), we have that $(\sqrt{T}(\hat{Z}_T(y,x) - F_T^A(y,x)), \sqrt{T}(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau))) \Rightarrow (\mathbb{H}_1(y,x), \tilde{\mathbb{H}}_2(\theta,\tau))$ in $\ell^\infty(\mathcal{W}\times\Theta\times\mathcal{T})$.
$\square$

A.2. Proofs of the Propositions

Proof of Proposition 1: To prove part (a), we consider the empirical processes $v_T(y,x) = \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ and $v_T^{\theta_0}(y,x) = \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F(y,x,\theta_0))$ defined in Lemma A.1 and in Lemma A.3, respectively. Under H_0 in (2.1), we have that $F_{YX}(y,x) \equiv F(y,x,\theta_0)$. Then:

$$\begin{aligned}
S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T)\right)^2 d\hat{Z}_T(y,x) \\
&= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm F_{YX}(y,x)\right)^2 d\hat{Z}_T(y,x) \\
&= \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 d\hat{Z}_T(y,x) \\
&= \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 dF_{YX}(y,x) \\
&+ \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 d(\hat{Z}_T(y,x) - F_{YX}(y,x)).
\end{aligned}$$

By Lemma A.1, $\sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x)) \Rightarrow \mathbb{H}_1(y,x)$, where $\mathbb{H}_1(y,x)$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$. Then, it follows that

$$S_T^{CM} = \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 dF_{YX}(y,x) + o_p(1).$$

By Lemma A.3, $(v_T(y,x), v_T^{\theta_0}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x))$ in $\ell^\infty(\mathcal{W}\times\mathcal{W})$, where $\mathbb{H}_2(y,x)$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$. Then, the result follows by an application of the continuous mapping theorem.

To prove part (b), under the alternative hypothesis H_A of (2.2), $F_{YX}(y,x) \neq F(y,x,\theta)$ for some $(y,x) \in \mathcal{W}$ and for all $\theta \in \mathcal{B}(\mathcal{T},\Theta)$, and $v_T^{\theta_0}(y,x)$ becomes $v_T^{\theta}(y,x) = \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F_T(y,x,\theta))$. Then,

$$\begin{aligned} S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm F_{YX}(y,x) \pm F(y,x,\theta)\right)^2 d\hat{Z}_T(y,x) \\ &= \int \left(v_T(y,x) - v_T^{\theta}(y,x) + \sqrt{T}\left(F_{YX}(y,x) - F(y,x,\theta)\right)\right)^2 dF_{YX}(y,x) + o_P(1). \end{aligned}$$

As a corollary of Lemma A.3, $(v_T(y,x), v_T^{\theta}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x))$ in $\ell^\infty(\mathcal{W}\times\mathcal{W})$. Therefore, for all fixed constants $C > 0$, we have $\lim_{T\to\infty} P(S_T^{CM} > C) = 1$, and the result follows. $\square$

Proof of Proposition 2: Under the local alternative $H_{A,T}$ in (2.7), consider the empirical processes

$$\begin{aligned} v_T^1(y,x) &= \sqrt{T}\left(\hat{Z}_T(y,x) - \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \leq x\} dF_X(x^*)\right), \text{ and} \\ r_T^1(\theta,\tau) &= \sqrt{T}(\hat{G}(\theta,\tau) - E_F[g(W_t,\theta,\tau)]), \end{aligned}$$

where $G_F(\theta,\tau) := E_F[g(W_t,\theta,\tau)]$, with $E_F[\cdot]$ defined as the expectation w.r.t. $F = F(y|x,\theta_0)$ in (2.7). Then

$$\begin{aligned}
v_T^1(y,x) &= \sqrt{T}\bigg(\hat{Z}_T(y,x) - \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \leq x\}dF_X(x^*)\bigg) \\
&= \sqrt{T}\hat{Z}_T(y,x) \\
&- \sqrt{T}\int\left(F_T(y|x^*) + \frac{\delta}{\sqrt{T}}\left[F(y|x^*,\theta_0) - J(y|x^*)\right]\right)\mathbb{1}\{x^* \leq x\}dF_X(x^*) \\
&= \sqrt{T}\left(\hat{Z}_T(y,x) - F_T^A(y,x)\right) + \delta\int[J(y|x^*) - F(y|x^*,\theta_0)]\mathbb{1}\{x^* \leq x\}dF_X(x^*).
\end{aligned}$$

Thus, it follows from Lemma A.4 that

$$v_T^1(y,x) \Rightarrow \mathbb{H}_1(y,x) + \delta\int[J(y|x^*) - F(y|x^*,\theta_0)]\mathbb{1}\{x^* \leq x\}dF_X(x^*),$$

where $\mathbb{H}_1$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$ defined in Lemma A.1. Then

$$\begin{aligned}
r_T^1(\theta,\tau) &= \sqrt{T}\left(\hat{G}(\theta,\tau) - E_F[g(W_t,\theta,\tau)]\right) \\
&= \sqrt{T}\left(\hat{G}(\theta,\tau) - (E_{F_T}[g(W_t,\theta,\tau)] + \delta E_F[g(W_t,\theta,\tau)] - \delta E_J[g(W_t,\theta,\tau)])\right) \\
&= \sqrt{T}\left(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau) + \delta\left(E_J[g(W_t,\theta,\tau)] - E_F[g(W_t,\theta,\tau)]\right)\right),
\end{aligned}$$

where $G_J(\theta,\tau) := E_J[g(W_t,\theta,\tau)]$, with $E_J[\cdot]$ defined as the expectation w.r.t. $J = J(y|x)$ in (2.7). We define the empirical process $v_T^{1\theta_0}(y,x)$ as follows:

$$v_T^{1\theta_0}(y,x) = \sqrt{T}\bigg(\int F(y|x^*,\hat{\theta}_T)\mathbb{1}\{x^* \leq x\}d\hat{F}_X(x^*) - \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \leq x\}dF_X(x^*)\bigg).$$

By Lemmas A.3-A.4,

$$\begin{pmatrix} v_T^{1\theta_0}(y,x) \\ \\ r_T^1(\theta,\tau) \end{pmatrix} \Rightarrow \begin{pmatrix} \mathbb{H}_2(y,x) + \delta \int \dot{F}(y|x^*)[h]\mathbb{1}\{x^* \leq x\} dF_X(x^*) \\ \\ \tilde{\mathbb{H}}_2(\theta,\tau) + \delta \left(E_J[g(W_t,\theta,\tau)] - E_F[g(W_t,\theta,\tau)]\right) \end{pmatrix},$$

with $h(\tau) = (\partial G_F(\theta_0,\tau)/\partial\theta)^{-1} G_J(\theta_0,\tau)$, and where $(\mathbb{H}_2, \tilde{\mathbb{H}}_2)$ are the tight mean zero Gaussian processes described in Lemmas A.2-A.3. Therefore, under $H_{A,T}$ in (2.7),

$$\begin{aligned} S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \leq x\} dF_X(x^*)\right)^2 d\hat{Z}_T(y,x) \\ &= \int \left(v_T^1(y,x) - v_T^{1\theta_0}(y,x)\right)^2 d\hat{Z}_T(y,x) \\ &= \int (v_T^1(y,x) - v_T^{1\theta_0}(y,x))^2 dF_{YX}(y,x) + o_p(1), \end{aligned}$$

Then,

$$S_T^{CM} \xrightarrow{d} \int \left(\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x) + \Delta(y,x)\right)^2 dF_{YX}(y,x),$$

with $\Delta(y,x) = \delta \int (J(y|x^*) - F(y|x^*,\theta_0) + \dot{F}(y|x^*,\theta_0)[h])\mathbb{1}\{x^* \leq x\} dF_X(x^*)$, and h is the function $h(\tau) = (\partial G_F(\theta_0,\tau)/\partial\theta)^{-1} G_J(\theta_0,\tau)$. $\square$

Proof of Proposition 3: Assumption 1 implies Assumptions 1-2 of Whang (2006). Then, parts (a) and (b) follow from an application of Theorems 2 and 3 of Whang (2006) using the convergence results of our Proposition 1. Further, Assumption 2 implies Assumption 2* of Whang (2006). Therefore, part (c) follows the same steps of Theorem 5 of Whang (2006) using the convergence results of our Proposition 1. $\square$