Conditional Quantile Coverage: an Application to Growth-at-Risk

Valentina Corradi*, Jack Fosten** and Daniel Gutknecht***

*University of Surrey

**Kings College

***Goethe University

SIDE

3 December 2021

Background

(a) Accurate predictions of the conditional mean or a (set of) conditional moment(s) are often not sufficient when interest lies in the prediction of intervals or the entire conditional distribution.

(b) In the finance literature, where risk management requires tracking of the entire distribution of a portfolio or measuring certain distributional aspects such as Value-at-Risk, this has already been widely acknowledged and conditional quantile models are routinely used.

(c) Interval prediction has also gained momentum in the empirical macroeconomic literature. Growth at Risk (GaR) is the probability that growth, conditional on current macrofinancial conditions, will fall below a given threshold. Monitoring of GaR is now a standard toolkit of the International Monetary Fund macrofinancial surveillance toolkit, see e.g. Prasad et al. (2019). Growth in Stress (Gonzalez-Rivera et al 2019).

(d) Growth vulnerability: downside growth risk increases with deteriorating financial conditions, while upside growth risk is rather stable. This implies that lower quantile in the output distribution are more sensible to deteriorating financial condition.

(e) Adrian, Boyarchenko and Giannone (2019) introduce a measure of growth vulnerability based on the distance between unconditional and conditional density below and above the median. Brownless and Souza (2021) assess GaR out of sample predictive ability of different quantile models as well as GARCH models.

(f) Usefulness of GaR monitoring strongly relies on the validity of models used to predict conditional quantiles. The use of systematically biased quantile prediction will totally hamper the benefits of GaR surveillance.

(g) To evaluate the accuracy of these interval predictions, numerous tests for correct specification of parametric quantile models have already been proposed. Quantile predictive accuracy is then measured in terms of coverage.

(h) Christoffersen (1998) introduced tests for correct unconditional and conditional coverage probability of quantile models at a pre-specified level $\tau \in (0, 1)$

(i) Escanciano and Olmo (2010) considered an out of sample version for unconditional coverage test taking into account also the contribution of estimation error.

(j) Escanciano and Velasco (2010) developed a test for correct specification of dynamic conditional quantile models, uniformly over a compact subset of quantile ranks $\tau \in T \subset (0, 1)$. Horwath et al (2021) consider a test which is uniform over the conditioning set, for given quantile level. Also Li et al (2021) propose tests for superior ability holding uniformly over conditioning states.

(k) Giacomini and Komunjer (2010), Manzan (2015) evaluate quantile predictive accuracy using the quantile score principle of Gneiting and Raftery (2007). However, such a principle delivers an optimal pointwise prediction of the variable of interest under a check type loss function.

In This Paper

(i) This paper goes one step further and introduces novel tests for pairwise and multiple out of sample comparisons of parametric conditional quantile models.

(ii) Our tests are based on a comparison of possibly overlapping models in terms of their relative coverage probability, conditional on the union of the information sets of the candidate models.

(iii) The conditional coverage error is constructed as the difference between the implied conditional coverage of the model and the nominal level $\tau \in \mathcal{T}$.

(iv) The null hypothesis in the pair-wise comparison set-up is that both models have equal expected coverage error, for a given loss function, over a given interval.

(vi) Our test statistics are constructed by computing the mean over difference of losses from conditional coverage errors.

(vii) Two sources of estimation error. Parametric estimation of the conditional quantile model. Nonparametric estimation of the conditional coverage probability

(vi) Because of double estimation error, the limiting distribution have a rather complicated, non nuisance parameter free, covariance structure.

(vii) Common solution is the use of subsampling. In the GaR case, samples are too small for effectively use subsampling.

(viii) We suggest a novel wild bootstrap procedure and establish its first order validity

(ix) We finally consider a Reality Check set-up, in which we test whether some of competing models have more accurate coverage, over some interval, than the benchmark model.

Outline

SET-UP

LIMITING DISTRIBUTION

BOOTSTRAP VALIDITY

MULTIPLE MODELS and MULTIPLE INTERVALS

MONTE CARLO

EMPIRICAL ILLUSTRATION TO GaR

SET-UP

CONDITIONAL COVERAGE

Z_t is a random vector which contains a response variable of interest y_t and other observable predictors, $F_t = \sigma(Z_s; s \leq t)$. The conditional τ-quantile of y_{t+1} is defined as:

$$q(\tau|\mathcal{F}_t) = \inf\{s : F_{y_{t+1}|\mathcal{F}_t}(s|\mathcal{F}_t) \geq \tau\},$$

By construction, for any $\tau_L, \tau_U \in T$ with $\tau_L < \tau_U$, it holds that:

$$\begin{aligned} & C\left([\tau_L, \tau_U]\,; \mathcal{F}_t\right) \\ = \; & \Pr\left(q(\tau_L|\mathcal{F}_t) \leq y_{t+1} \leq q(\tau_U|\mathcal{F}_t)|\mathcal{F}_t\right) = \tau_U - \tau_L \end{aligned}$$

$C\left([\tau_L, \tau_U]\,; \mathcal{F}_t\right)$ is the conditional coverage wrt information set $\mathcal{F}_t$

CONDITIONAL QUANTILE MODELS

Linear conditional quantile model:

$$q(\tau|\mathcal{F}_t) = \mathbf{Z}_t'\beta(\tau).$$

This model comprises the QAR model as discussed in Koenker and Xiao (2006)

$$y_{t+1} = \theta_0(U_{t+1}) + \theta_1(U_{t+1})y_t + \ldots + \theta_p(U_{t+1})y_{t-p+1}$$

or the location scale model:

$$y_{t+1} = \mathbf{Z}_t'\delta + (\mathbf{Z}_t'\gamma)\varepsilon_{t+1},$$

we have that $q(\tau|F_t) = Z_t'\beta(\tau)$ with $\beta(\tau) = \delta + \gamma q_\varepsilon(\tau)$ and $q_\varepsilon(\tau)$ denoting the τ unconditional quantile of the error term ε_{t+1}.

or nonlinear location scale models:

$$y_{t+1} = m\left(\mathbf{Z}_t, \theta_1\right) + \sigma(\mathbf{Z}_t, \theta_2)\varepsilon_{t+1}$$

$$q(\tau|\mathcal{F}_t) = m\left(\mathbf{Z}_t, \theta_1\right) + \sigma(\mathbf{Z}_t, \theta_2) q_\varepsilon(\tau).$$

INFORMATION SETS

$X_{j,t}$, $j = 1, \ldots, J$, denote the specific information set of each of the J candidate models, which consists of elements of $\{Z_s, s \leq t\}$, and let $q_j(\tau|X_{j,t})$ denote its τ level quantile conditional on information set $X_{j,t}$.

$\mathbf{X}_t = \{X_{1,t} \cup \ldots \cup X_{J,t}\}$ to denote the union of the information sets of all candidate models

Coverage for model j

$$\begin{aligned} & C_j\left([\tau_L, \tau_U]; \mathbf{X}_t\right) \\ = \; & \Pr\left(q_j(\tau_L|X_{j,t}) \leq y_{t+1} \leq q_j(\tau_U|X_{j,t})|\mathbf{X}_t\right) \end{aligned}$$

the conditional coverage error of model j:

$$\begin{aligned} & \mathcal{E}_j\left(\left[\tau_L, \tau_U\right]; \mathbf{X}_t\right) \\ = \; & C_j\left(\left[\tau_L, \tau_U\right]; \mathbf{X}_t\right) - \left(\tau_U - \tau_L\right). \end{aligned}$$

and given that we allow for (dynamically) misspecified models

$$\mathsf{E}\left(\mathcal{E}_j\left(\left[\tau_L, \tau_U\right]; \mathbf{X}_t\right)\right) \neq 0$$

for some or all models $j = 1, \ldots, J$ almost surely.

LIMITING DISTRIBUTION

For the pairwise comparison case, single interval

$H_0 : \mathsf{E}\left(L\left(\mathcal{E}_1\left(\left[\tau_L, \tau_U\right]; \mathbf{X}_t\right)\right) - L\left(\mathcal{E}_2\left(\left[\tau_L, \tau_U\right]; \mathbf{X}_t\right)\right)\right) = 0$

against its negation.

Three cases may arise under H_0:

CASE I:

$$\Pr\left(C_1([\tau_L, \tau_U]; \mathbf{X}_t) = C_2([\tau_L, \tau_U]; \mathbf{X}_t)\right) < 1.$$

The models are not overlapping

CASE II:

$$\Pr\left(C_1([\tau_L, \tau_U]; \mathbf{X}_t) = C_2([\tau_L, \tau_U]; \mathbf{X}_t)\right) = 1$$

and for $j = 1, 2$:

$$\Pr\left(C_j([\tau_L, \tau_U]; \mathbf{X}_t) = \tau_U - \tau_L\right) < 1.$$

The models are overlapping, but misspecified and so both have incorrect conditional coverage

CASE III: for $j = 1, 2$

$$\Pr\left(C_j([\tau_L, \tau_U]; \mathbf{X}_t) = \tau_U - \tau_L\right) = 1.$$

The models are overlapping and both have correct conditional coverage.

Let $T = R + P$. We use the first R observations to estimate the parametric quantile models, $\widehat{q}_{j,R}(\tau | X_{j,t})$ for $j = 1, 2$.

Then, we need to construct nonparametrically, the conditional coverage,

$$\begin{aligned} & \widehat{C}_{j,R,P}(\tau;\mathbf{X}_t) \\ = \; & \frac{1}{Ph^d}\sum_{s=R}^{T-1} 1\left\{y_{s+1} \leq \widehat{q}_{j,R}(\tau|X_{j,t})\right\} \\ & \times \frac{1}{\widehat{f}_X(\mathbf{X}_t)}\mathbf{K}\left(\frac{\mathbf{X}_s-\mathbf{X}_t}{h}\right) \end{aligned}$$

Let $\widetilde{C}_j(\tau;\mathbf{X}_t) = F\left(\widehat{q}_j(\tau|X_{j,t})\right)$ and $C_j(\tau;\mathbf{X}_t) = F\left(q_j(\tau|X_{j,t})\right)$ then

$$\begin{aligned} & \widehat{C}_{j,R,P}(\tau;\mathbf{X}_t) - C_j(\tau;\mathbf{X}_t) \\ = \; & \left(\widehat{C}_{j,R,P}(\tau;\mathbf{X}_t) - \widetilde{C}_j(\tau;\mathbf{X}_t)\right) \\ & + \left(\widetilde{C}_{j,R,P}(\tau;\mathbf{X}_t) - C_j(\tau;\mathbf{X}_t)\right) \end{aligned}$$

The first term captures nonparameteric conditional coverage estimation error. The second term captures parameteric conditional quantile estimation error. The second term is negligible if $P/R \to 0$.

The statistics is

$$\begin{aligned}
&\widehat{S}_{P,RP,R}(\tau_L,\tau_U) \\
= \; &\frac{1}{P^{1/2}}\sum_{t=R}^{T-1}\Big(L\left(\widehat{\mathcal{E}}_{1,R,P}\left([\tau_L,\tau_U]\,;\mathbf{X}_t\right)\right) \\
&-L\left(\widehat{\mathcal{E}}_{2,R,P}\left([\tau_L,\tau_U]\,;\mathbf{X}_t\right)\right)\Big)
\end{aligned}$$

THEOREM 1: Let Assumptions A.1-A.6 hold. If as $P,R\to\infty,\ P/R\to\pi,\ 0<\pi<\infty,\ Ph^{2r}\to 0$ and $Ph^d/\ln(P)\to\infty$,

(i) Under H_0, in CASE I:

$$\widehat{S}_{P,R}(\tau_L,\tau_U)\xrightarrow{d}G(\tau_L,\tau_U)$$

where G is a zero mean normal with covariance $\Omega\left(\tau_L,\tau_U\right)$.

(ii) Under H_0, in CASE II:

$$\widehat{S}_{P,R}(\tau_L, \tau_U) \xrightarrow{d} \widetilde{G}(\tau_L, \tau_U)$$

where $\widetilde{G}$ is a Gaussian process with variance kernel $\widetilde{\Omega}(\tau_L, \tau_U)$

(iii) Under H_0, in Case III:

$$\widehat{S}_{P,R}(\tau_L, \tau_U) = O_p\left(\frac{1}{R^{1/2}}\right) = O_p\left(\frac{1}{P^{1/2}}\right),$$

since P and R grow at the same rate

(iv) Under H_A, there exists $\varepsilon > 0$ such that

$$\lim_{P,R\to\infty} \Pr\left(\widehat{S}_{P,R} > \varepsilon\right) = 1.$$

In Case I, the asymptotic variance takes into account loss differential, nonparametric and parametric estimation error. In Case II, only quantile parametric estimation error matters. In Case III the statistis is degenerate. Since $L(0) = \nabla L(0) = 0$, only the second term in a Taylor expansion is non zero, and so in Case III,

$$
\begin{aligned}
& \widehat{S}_{P,R}(\tau) \\
\approx\ & \frac{1}{P^{1/2}} \sum_{j=R}^{T-1} \Big(\Big(\widehat{C}_{1,P}\left(\widehat{\boldsymbol{\psi}}_{1,R}(\tau); \mathbf{X}_j \right) - C_1\left(\boldsymbol{\psi}_1^{\dagger}(\tau); \mathbf{X}_j \right) \Big)^2 \\
& - \left(\widehat{C}_{2,P}\left(\widehat{\boldsymbol{\psi}}_{2,R}(\tau); \mathbf{X}_j \right) - C_2\left(\boldsymbol{\psi}_2^{\dagger}(\tau); \mathbf{X}_j \right) \right)^2 \Big)
\end{aligned}
$$

BOOTSTRAP CRITICAL VALUES

We suggest a wild bootstrap procedure in the spirit of the conditional p-value approach of Hansen (1996). Inoue (2001) extended Hansen to the time series case, Corradi and Swanson (2002) extend Inoue to parameter estimation error , here we extend CS to nonparametric estimation error as well. Importantly, critical values based on our procedure are first order asymptotically valid, regardless of whether we are under CASE I, II or III.

The wild bootstrap statistic mirror the linear expansion of the original statistic $\widehat{S}_{P,R}(\tau)$, i.e.

$$\begin{aligned} & \widehat{S}_{P,R}(\tau) \\ = \; & \frac{1}{P^{1/2}} \sum_{t=R}^{T-1} 1\{\mathbf{X}_i \in \mathcal{X}\} \left(\left(A_{1,t}(\tau) - A_{2,t}(\tau) \right) \right. \\ & \left. + \left(B_{1,t}(\tau) - B_{2,t}(\tau) \right) + \left(D_{1,t}(\tau) - D_{2,t}(\tau) \right) \right) + o_p(1) \end{aligned}$$

where $A_{1,t} - A_{2,t}(\tau)$ capture the "true" difference in coverage error, $B_{1,t}(\tau) - B_{2,t}(\tau)$ mimics contribution of nonparametric estimation error and $D_{1,t}(\tau) - D_{2,t}(\tau)$ contribution of parametric quantile estimation error, i.e.

$$A_{1,t}(\tau) = L\left(C_1\left(\boldsymbol{\psi}_1^{\dagger}(\tau); \mathbf{X}_t\right) - \tau\right)$$

$$
\begin{aligned}
& B_{1,t}(\tau) \\
= \; & \nabla L\left(C_1\left(\boldsymbol{\psi}_1^{\dagger}(\tau);\mathbf{X}_t\right)-\tau\right) \\
& \left(1\{y_{t+1}\le q_\tau(\boldsymbol{\psi}_1^{\dagger}(\tau))\}-F_{t+1}(q_\tau(\boldsymbol{\psi}_1^{\dagger}(\tau))|\mathbf{X}_{1,t})\right)
\end{aligned}
$$

$$
\begin{aligned}
& D_{1,t}(\tau) \\
= \; & \nabla L\left(C_1\left(\boldsymbol{\psi}_1^{\dagger}(\tau);\mathbf{X}_t\right)-\tau\right)\Lambda_{1,P,T}(\tau)\,H_{1,P,T}^{-1}(\tau) \\
& X_{1,t}\left(1\{y_{t+1}\le q_\tau(\boldsymbol{\psi}_1^{\dagger}(\tau);X_{1,t})\}-\tau)\right)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& \widehat{S}_{P,R}^{*}(\tau) \\
= \; & \frac{1}{P^{1/2}}\sum_{t=R}^{T-l_P-1}\varepsilon_t\sum_{i=t}^{t+l_P}\left(1\{\mathbf{X}_i\in\mathcal{X}\}\left(\left(\widehat{A}_{1,t}(\tau)-\widehat{A}_{2,t}(\tau)\right)\right.\right. \\
& \left.\left.+\left(\widehat{B}_{1,t}(\tau)-\widehat{B}_{1,t}(\tau)\right)\right)\right) \\
& +\frac{P^{1/2}}{R}\sum_{t=1}^{R-l_R-1}\eta_t\sum_{i=t}^{t+l_R}\widehat{\mu}_1(\tau)\left(\widehat{D}_{1,t}(\tau)-\widehat{D}_{2,t}(\tau)\right).
\end{aligned}
$$

Theorem 2: Let Assumption A.1-A.6. hold. Also, as $P, R, B \to \infty$, $l_R, l_P \to \infty$, $l_P/P^{1/2} \to 0$ and $l_R/R^{1/2} \to 0$. Then:

(i) Under H_0, in CASE I and CASE II with $P/R \to \pi > 0$, as $P, R, B \to \infty$

$$\begin{aligned} & \lim \Pr\left(c^*_{B,P,R}(\alpha/2) < \widehat{S}_{P,R}(\tau) < c^*_{B,P,R}(1-\alpha/2)\right) \\ = \ & 1-\alpha \end{aligned}$$

(ii) Under H_0, in CASE III:

$$\begin{aligned} & \lim \Pr\left(c^*_{B,P,R}(\alpha/2) < \widehat{S}_{P,R}(\tau) < c^*_{B,P,R}(1-\alpha/2)\right) \\ = \ & 1 \end{aligned}$$

(iii) Under H_A:

$$\begin{aligned} & \lim \Pr\left(c^*_{B,P,R}(\alpha/2) < \widehat{S}_{P,R}(\tau) < c^*_{B,P,R}(1-\alpha/2)\right) \\ = \ & 0 \end{aligned}$$

Theorem 2 establishes the first order validity of the block bootstrap critical values. In particular, we have a test of size α in CASE I and II, and a test of size not larger than α in CASE III. The key is that $\widehat{S}_{P,R}(\tau)$ goes to zero faster than $\widehat{S}^{*}_{P,R}(\tau)$.

MULTIPLE MODELS AND INTERVALS

The null hypothesis is that none of the competing models has smaller conditional coverage error at any of the intervals considered. The alternative is that at least one competitor outperforms the benchmark on at least one interval. That is:

$$
\begin{aligned}
H_0^{RC} \quad : \quad & \max_{j=2,\ldots,J} \max_{i=1,\ldots,m} \mathsf{E}\left(L\left(\mathcal{E}_1\left(\left[\tau_{i,L}, \tau_{i,U}\right]; \mathbf{X}_t^j\right)\right)\right. \\
& \left. - L\left(\mathcal{E}_j\left(\left[\tau_{i,L}, \tau_{i,U}\right]; \mathbf{X}_t^j\right)\right)\right) 1\left\{\mathbf{X}_t^j \in \mathcal{X}\right\} \\
\leq \quad & 0
\end{aligned}
$$

versus it negation. Now,

$$H_0^{RC} = \cap_{j=2}^{J} \cap_{i=1}^{m} H_{0,i,j}^{RC}$$

and:

$$H_A^{RC} = \cup_{j=2}^{J} \cup_{i=1}^{m} H_{0,i,j}^{RC,c}$$

with $H_{0,j,i}^{RC,c}$ denoting the complement of $H_{0,j,i}^{RC}$. The number of moment weak inequalities is given by the product of the number of competing models times the number of intervals, $(J-1) \times m$. We have three cases under the null

CASE I-M: For at least one interval $i \in \{1, ..., m\}$ and one model $j \in \{2, ..., J\}$:

$$\Pr\Big(C_1([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j) = C_j([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j)\Big) < 1.$$

CASE II-M: For all $i = 1, ..., m$ and $j = 2, ..., J$:

$$\Pr\left(C_1([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j) = C_j([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j)\right) = 1,$$

but for at least one interval $i \in \{1, ..., m\}$ and one model $j \in \{2, ..., J\}$:

$$\Pr\left(C_j([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j) = \tau_{iU} - \tau_{iL}\right) < 1.$$

CASE III-M: For all $i = 1, ..., m$ and $j = 2, ..., J$

$$\Pr\left(C_j([\tau_{iL}, \tau_{iU}]; \mathbf{X}_t^j) = \tau_{iU} - \tau_{iL}\right) = 1.$$

The statistic reads as:

$$\widehat{S}_{P,R}^{\mathsf{max,max}} = \sum_{j=2}^{J} \sum_{i=1}^{m} \left(\max\left\{0, \widehat{S}_{P,R}([\tau_{iL}, \tau_{iU}]; j)\right\}\right)^2$$

Theorem 4: Let Assumptions A.1-A.6 hold. Then:

(i) Under H_0^{RC}, in CASE I-M:

$$\widehat{S}_{P,R}^{\mathsf{max,max}} \xrightarrow{d} \sum_{k=1}^{m(J-1)} \left(\max\{0, Z_k\}\right)^2,$$

where Z_k is the k-th element of a $m(J-1)$-dimensional zero mean normal random vector with variance covariance matrix equal to V

(ii) Under H_0^{RC}, in CASE II-M with $P/R \rightarrow \pi > 0$:

$$\widehat{S}_{P,R}^{\mathsf{max,max}} \xrightarrow{d} \sum_{k=1}^{m(J-1)} \left(\mathsf{max}\left\{0, \widetilde{Z}_k\right\}\right)^2,$$

where $\widetilde{Z}_k$ is the k-th element of a $m(J-1)$-dimensional zero mean normal random vector with variance covariance matrix equal to $\widetilde{V}$).

(iii) Under H_0^{RC}, in CASE III-M:

$$\widehat{S}_{P,R}^{\mathsf{max,max}} = O_p\left(\frac{1}{R^{1/2}}\right)$$

(iv) Under H_A^{RC}, there exists $\varepsilon > 0$ such that:

$$\lim_{P,R\to\infty} \Pr\left(\frac{1}{P^{1/2}}\widehat{S}_{P,R}^{\text{max,max}} > \varepsilon\right) = 1.$$

We now introduce the wild bootstrap counterpart of $\widehat{S}_{P,R}^{\text{max,max}}$ for $k = (j-2)m + i,\ j = 1, ..., J,\ i = 1, ..., m$ let $\widehat{S}^*_{P,R,k}$ be the analog of $\widehat{S}^*_{P,R}$ when comparing model 1 and j over interval i.

$$\begin{pmatrix} \widehat{S}^*_{P,R,1} \\ \vdots \\ \widehat{S}^*_{P,R,m(J-1)} \end{pmatrix} = \begin{pmatrix} \widehat{S}^*_{P,R}([\tau_{1L}, \tau_{1U}]; 2) \\ \vdots \\ \widehat{S}^*_{P,R}([\tau_{mL}, \tau_{mU}]; J-1) \end{pmatrix}.$$

By using the same draws of ε_t and η_t across intervals and models, the correlation present in the data is fully preserved.

The bootstrap statistic reads as,

$$\widehat{S}^{*\,\max,\max}_{P,R} = \sum_{k=1}^{m(J-1)} \left(\max\left\{0, \widehat{S}^{*}_{P,R,k} 1\left\{\widehat{S}_{P,R,k} \geq -\widehat{v}_{kk,P,R}\kappa_P\right\}\right\}\right)^2,$$

with $\widehat{v}^2_{kk,P,R}$ being an estimator of the variance of $\widehat{S}_{P,R,k}$, $k_P \to \infty$ as $P \to \infty$. When $\widehat{S}_{P,R,k} < -\widehat{v}_{k,k}\kappa_P$, it holds that

$\max\left\{0, \widehat{S}^{*}_{P,R,1} 1\left\{\widehat{S}_{P,R,k} \geq -\widehat{v}_{kk,P,R}\kappa_P\right\}\right\} = 0$, so that sufficiently negative moment conditions do not contribute to the bootstrap statistic.

Theorem 5: Let Assumptions A.1-A.6 hold. Also, as $P, R, B \to \infty$, $l_R, l_P \to \infty$, $l_P/P^{1/2} \to 0$ and $l_R/R^{1/2} \to 0$, $\frac{\kappa_P}{\log \log P} \to \infty$ and $\kappa_P/P^{1/2} \to 0$. Then:

(i) (a) Under H_0^{RC}, in CASE I-M $P/R \to \pi > 0$:

$$\lim_{B,R,P\to\infty} \Pr\left(\widehat{S}_{P,R}^{\max,\max} \geq c_{B,R,P,1-\alpha}^{*\max,\max}\right) \leq \alpha.$$

(b) if in addition for some interval/model the two models have equal coverage error,

$$\lim_{B,R,P\to\infty} \Pr\left(\widehat{S}_{P,R}^{\max,\max} \geq c_{B,R,P,1-\alpha}^{*\max,\max}\right) = \alpha.$$

(ii) Under H_0^{RC}, in CASE II-M:

$$\lim_{B,R,P\to\infty} \Pr\left(\widehat{S}_{P,R}^{\max,\max} \geq c_{B,R,P,1-\alpha}^{*\max,\max}\right) = \alpha.$$

(iii) Under H_0^{RC}, in CASE III-M:

$$\lim_{B,R,P\to\infty} \Pr\left(\widehat{S}_{P,R}^{\max,\max} \geq c_{B,R,P,1-\alpha}^{*\max,\max}\right) = 0$$

(iv) Under H_A^{RC}:

$$\lim_{B,R,P\to\infty} \Pr\left(\widehat{S}_{P,R}^{\max,\max} \geq c_{B,R,P,1-\alpha}^{*\max,\max}\right) = 1.$$

MONTE CARLO SIMULATION

DGP

$$y_{t+1} = \beta_1 X_{1,t} + \beta_2 X_{2,t} + \left|1 + \beta_3(X_{1,t} + X_{2,t})\right| e_{t+1}$$

where $\beta_3 = 0$ gives a simple linear model with error term e_{t+1} whereas $\beta_3 \neq 0$ allows the conditional variance to be nonlinear in $X_{1,t}$ and $X_{2,t}$, quantile regression give raise to mis-specification.

$X_{j,t} = \rho X_{j,t-1} + v_{j,t}$ for $j = 1, 2$, and $v_{j,t}$ are generated as i.i.d.$N(0, 1 - \rho^2)$. For $j = 1, 2$ we write this as:

$$q_\tau(\boldsymbol{\beta}_j^\dagger(\tau); X_{j,t}) = \beta_{0j}^\dagger(\tau) + \beta_{1j}^\dagger(\tau) X_{j,t}$$

where $\beta_j^{\dagger}(\tau) = \left(\beta_{0j}^{\dagger}(\tau), \beta_{1j}^{\dagger}(\tau)\right)'$.

DGP1: $(\beta_1, \beta_2, \beta_3) = (1, 1, 0)$–CASE 1

DGP2: $(\beta_1, \beta_2, \beta_3) = (0, 0, 1)$–CASE II

DGP3: $(\beta_1, \beta_2, \beta_3) = (0, 0, 0)$ CASE III

DGP4: $(\beta_1, \beta_2, \beta_3) = (0, 1, 0)$ POWER

Quadratic Loss, $T = 240, 480, 960$, $P = R$, $\tau = 0.1, 0.2, 0.3$

EMPIRICAL ILLUSTRATION

We apply our test to evaluate the out-of-sample specification of the recent GaR framework of Adrian et al. (2019). We look at lower part of the distribution, $\tau = 0.1, 0.2, 0.3$. We want to predict quantile of GDP/IP using QR.

Data 1971-2019, quarterly for GDP and monthly for IP. $T = 196$ (quarterly) and $T = 588$ (monthly). $P = R$

Benchmark model is QAR(1)+NFCI (National Financial Condition Indicator). Candidates are

QAR(1)+SV stock market and volatility, from Brownless and Sousa (2021)

QAR(1)+GF Global Real Economic Activity Factor

QAR(1)+TS Term Spread 10yrs-1yes treasury rate

QAR(1)+HP House Prices

Table 1: Rejection Rates: Pairwise - Single Quantile Level

$T = 240$		DGP1		DGP2		DGP3		DGP4	
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	$\tau = 0.1$	0.0605	0.0330	0.0345	0.0160	0.0170	0.0055	0.1096	0.0535
	$\tau = 0.2$	0.0740	0.0355	0.0235	0.0090	0.0105	0.0045	0.2726	0.1336
	$\tau = 0.3$	0.0755	0.0360	0.0200	0.0080	0.0140	0.0055	0.5418	0.3122
$l = 2$	$\tau = 0.1$	0.0545	0.0230	0.0265	0.0130	0.0175	0.0040	0.1076	0.0430
	$\tau = 0.2$	0.0600	0.0295	0.0220	0.0110	0.0130	0.0070	0.2766	0.1276
	$\tau = 0.3$	0.0660	0.0330	0.0205	0.0080	0.0140	0.0055	0.4602	0.2841
$l = 5$	$\tau = 0.1$	0.0555	0.0215	0.0330	0.0145	0.0165	0.0030	0.1086	0.0450
	$\tau = 0.2$	0.0540	0.0265	0.0215	0.0095	0.0125	0.0055	0.2221	0.1116
	$\tau = 0.3$	0.0550	0.0255	0.0175	0.0065	0.0125	0.0040	0.3912	0.2016
$T = 480$		DGP1		DGP2		DGP3		DGP4	
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	$\tau = 0.1$	0.0780	0.0395	0.0265	0.0120	0.0100	0.0025	0.2756	0.1746
	$\tau = 0.2$	0.0815	0.0415	0.0315	0.0135	0.0185	0.0050	0.7809	0.6338
	$\tau = 0.3$	0.0735	0.0340	0.0230	0.0105	0.0105	0.0050	0.9700	0.9325
$l = 2$	$\tau = 0.1$	0.0570	0.0355	0.0260	0.0110	0.0085	0.0010	0.2391	0.1341
	$\tau = 0.2$	0.0765	0.0380	0.0260	0.0130	0.0150	0.0055	0.7129	0.4942
	$\tau = 0.3$	0.0660	0.0315	0.0225	0.0115	0.0135	0.0055	0.9625	0.9125
$l = 5$	$\tau = 0.1$	0.0540	0.0190	0.0290	0.0115	0.0075	0.0020	0.2226	0.1196
	$\tau = 0.2$	0.0670	0.0330	0.0245	0.0100	0.0185	0.0070	0.6033	0.4112
	$\tau = 0.3$	0.0495	0.0215	0.0220	0.0075	0.0105	0.0045	0.9515	0.8199
$T = 960$		DGP1		DGP2		DGP3		DGP4	
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	$\tau = 0.1$	0.0850	0.0410	0.0340	0.0175	0.0115	0.0020	0.6443	0.4817
	$\tau = 0.2$	0.1011	0.0560	0.0375	0.0140	0.0115	0.0035	0.9915	0.9685
	$\tau = 0.3$	0.1031	0.0540	0.0310	0.0125	0.0100	0.0025	1.0000	1.0000
$l = 2$	$\tau = 0.1$	0.0640	0.0270	0.0300	0.0130	0.0130	0.0025	0.5958	0.3962
	$\tau = 0.2$	0.0895	0.0495	0.0365	0.0155	0.0140	0.0045	0.9875	0.9425
	$\tau = 0.3$	0.0885	0.0520	0.0385	0.0110	0.0085	0.0020	1.0000	1.0000
$l = 5$	$\tau = 0.1$	0.0660	0.0295	0.0280	0.0125	0.0130	0.0020	0.5138	0.3177
	$\tau = 0.2$	0.0880	0.0485	0.0400	0.0155	0.0145	0.0045	0.9790	0.9340
	$\tau = 0.3$	0.0855	0.0475	0.0390	0.0150	0.0105	0.0030	1.0000	0.9995

Table 2: Rejection Rates: Pairwise - Multiple Quantile Levels

$T = 240$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.0670	0.0340	0.0280	0.0105	0.0115	0.0040	0.6083	0.3892
$l = 2$	0.0555	0.0265	0.0300	0.0115	0.0125	0.0045	0.5593	0.3542
$l = 5$	0.0550	0.0265	0.0230	0.0105	0.0130	0.0050	0.5768	0.3437
$T = 480$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.0975	0.0555	0.0270	0.0095	0.0110	0.0035	0.9775	0.9505
$l = 2$	0.0805	0.0415	0.0245	0.0090	0.0130	0.0050	0.9780	0.9265
$l = 5$	0.0675	0.0335	0.0240	0.0090	0.0130	0.0070	0.9665	0.9075
$T = 960$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.1086	0.0565	0.0435	0.0215	0.0100	0.0025	1.0000	1.0000
$l = 2$	0.0890	0.0455	0.0410	0.0190	0.0145	0.0025	1.0000	1.0000
$l = 5$	0.0915	0.0495	0.0470	0.0240	0.0145	0.0030	1.0000	1.0000

Notes: The cases of DGP1 through DGP4 correspond to $(\beta_1, \beta_2, \beta_3)$ equal to $(1, 1, 0)$, $(0, 0, 1)$, $(0, 0, 0)$ and $(0, 1, 0)$ in (**??**). In this pairwise model set-up we have the non-degenerate CASE I under DGP1, the overlapping misspecified CASE II under DGP2, the strictly overlapping CASE III under DGP3 and we are under the alternative for DGP4.

Table 3: Rejection Rates: Multiple Models - Single Quantile Level - Benchmark $X_{1,t}$

$T = 240$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.2561	0.1351	0.0155	0.0050	0.0045	0.0000	0.5633	0.3492
$l = 2$	0.2176	0.1151	0.0110	0.0045	0.0040	0.0005	0.5933	0.3497
$l = 5$	0.1946	0.0955	0.0120	0.0050	0.0045	0.0005	0.5203	0.3047
$T = 480$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.7794	0.6078	0.0160	0.0035	0.0045	0.0010	0.9680	0.9160
$l = 2$	0.8129	0.5988	0.0130	0.0045	0.0040	0.0010	0.9760	0.9245
$l = 5$	0.7399	0.5198	0.0130	0.0030	0.0055	0.0010	0.9730	0.8829
$T = 960$	DGP1		DGP2		DGP3		DGP4	
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$
$l = 1$	0.9985	0.9945	0.0320	0.0105	0.0030	0.0000	1.0000	1.0000
$l = 2$	0.9975	0.9835	0.0275	0.0105	0.0030	0.0000	1.0000	1.0000
$l = 5$	0.9975	0.9800	0.0310	0.0110	0.0025	0.0000	1.0000	1.0000

Notes: The cases of DGP1 through DGP4 correspond to $(\beta_1, \beta_2, \beta_3)$ equal to $(1, 1, 0)$, $(0, 0, 1)$, $(0, 0, 0)$ and $(0, 1, 0)$ in (**??**). In this multiple model set-up with $X_{1,t}$ in the benchmark model and $X_{2,t}$ and $\mathbf{X}_t$ being used in the others, the benchmark is worse than the $\mathbf{X}_t$ model under DGP1, worse than both models under DGP4, and overlapping under DGP2 and DGP3.

Table 4: QAR(1)+NFCI vs. QAR(1)+SV - Pairwise Comparison - Single Quantile Level

		Real GDP Growth				IP Growth			
		$s_Q = 1$		$s_Q = 4$		$s_M = 3$		$s_M = 12$	
		Stat	p-value	Stat	p-value	Stat	p-value	Stat	p-value
$l = 1$	$\tau = 0.1$	0.5955	0.4052	-0.0157	0.9075	0.6611	0.2711	0.0333	0.8904
	$\tau = 0.2$	0.7311	0.3662	-0.0033	0.9705	1.4144	0.0420	0.5089	0.4752
	$\tau = 0.3$	1.9363	0.1871	0.2444	0.7764	1.2296	0.0200	0.3686	0.7194
$l = 2$	$\tau = 0.1$	-	0.3472	-	0.9375	-	0.2811	-	0.9275
	$\tau = 0.2$	-	0.3442	-	0.9775	-	0.0510	-	0.5423
	$\tau = 0.3$	-	0.1961	-	0.8014	-	0.0240	-	0.7504
$l = 5$	$\tau = 0.1$	-	0.3612	-	0.9375	-	0.3022	-	0.8774
	$\tau = 0.2$	-	0.3002	-	0.9755	-	0.0570	-	0.5773
	$\tau = 0.3$	-	0.2241	-	0.8114	-	0.0460	-	0.7764

Table 5: Multiple model, multiple quantile test

	Real GDP Growth				IP Growth			
	$s_Q = 1$		$s_Q = 4$		$s_M = 3$		$s_M = 12$	
	Stat	p-value	Stat	p-value	Stat	p-value	Stat	p-value
$l = 1$	6.4694	0.3412	0.0597	0.9490	8.9396	0.0260	0.4783	0.6193
$l = 2$	-	0.3332	-	0.9620	-	0.0265	-	0.6353
$l = 5$	-	0.3432	-	0.9535	-	0.0410	-	0.6728