



## UvA-DARE (Digital Academic Repository)

### Penalized Averaging of Parametric and Non-Parametric Quantile Forecasts

De Gooijer, J.G.; Zerom, D.

**DOI**

[10.1515/jtse-2019-0021](https://doi.org/10.1515/jtse-2019-0021)

**Publication date**

2020

**Document Version**

Final published version

**Published in**

Journal of Time Series Econometrics

[Link to publication](#)

**Citation for published version (APA):**

De Gooijer, J. G., & Zerom, D. (2020). Penalized Averaging of Parametric and Non-Parametric Quantile Forecasts. *Journal of Time Series Econometrics*, 12(1), [20190021]. <https://doi.org/10.1515/jtse-2019-0021>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Jan C. De Gooijer<sup>1</sup> / Dawit Zerom<sup>2</sup>

# Penalized Averaging of Parametric and Non-Parametric Quantile Forecasts

<sup>1</sup> Amsterdam School of Economics, University of Amsterdam, Amsterdam 1018 WB, Netherlands, E-mail: j.g.degooijer@uva.nl

<sup>2</sup> Mihaylo College of Business and Economics, California State University Fullerton, Fullerton, CA, USA, E-mail: dzerom@fullerton.edu

## Abstract:

We propose a hybrid penalized averaging for combining parametric and non-parametric quantile forecasts when faced with a large number of predictors. This approach goes beyond the usual practice of combining conditional mean forecasts from parametric time series models with only a few predictors. The hybrid methodology adopts the adaptive LASSO regularization to simultaneously reduce predictor dimension and obtain quantile forecasts. Several recent empirical studies have considered a large set of macroeconomic predictors and technical indicators with the goal of forecasting the S&P 500 equity risk premium. To illustrate the merit of the proposed approach, we extend the mean-based equity premium forecasting into the conditional quantile context. The application offers three main findings. First, combining parametric and non-parametric approaches adds quantile forecast accuracy over and above the constituent methods. Second, a handful of macroeconomic predictors are found to have systematic forecasting power. Third, different predictors are identified as important when considering lower, central and upper quantiles of the equity premium distribution.

**Keywords:** large database, non-parametric, parametric, penalized averaging, quantile forecasting

**DOI:** 10.1515/jtse-2019-0021

## 1 Introduction

The increasing availability of a large database of time series variables across diverse disciplines of business and economics has motivated a new strand of forecasting research catered to data rich environments. For example, Ma, Fildes, and Huang (2016) use variable selection and model estimation by a multistage LASSO regression, followed by a scheme to generate out-of-sample forecasts in the field of inventory management. Jiang et al. (2018) obtain forecasts of a large set of Australian macroeconomic predictors using a wide variety of dimension reduction methods. Another forecasting study in high-dimensions is by Exterkate et al. (2011) who consider least angle regression, a shrinkage and variable selection method proposed by Efron et al. (2004), and multi-response sparse regression (Similä and Tikka 2006). The above examples are far from exhaustive, and the literature is growing rapidly.

Mimicking a similar trend in non-forecasting studies, most high-dimensional forecasting research focuses mainly on conditional mean forecasts (Garcia, Medeiros, and Vasconcelos 2017; Konzen and Ziegelmann 2016). Further, the modeling framework of existing high-dimensional studies is often based on the assumption that high-dimensional data come from a linear data generating process (DGP); see, e. g. Bonaccolto, Caporin, and Paterlini (2018) and Bayer (2018). In an attempt to provide modeling flexibility, few studies have introduced extensions. For instance, Bai and Ng (2008) and Giovannelli (2012) introduce nonlinear principal component methods within the conditional mean context. Recently, deviating from parametric modeling approaches, Chen et al. (2018) introduce non-parametric based high-dimensional averaging.

With the goal of offering yet another modeling option in the presence of high-dimensional predictors, we consider the case of conditional quantile forecasts. In particular, we further extend the non-parametric averaging method of De Gooijer and Zerom (2019). Similar to that of Chen et al. (2018), the approach of De Gooijer and Zerom (2019) rests on the idea of combining or averaging a large number of possibly misspecified non-parametric quantile forecasts to form an approximation of a target conditional quantile. In addition to incorporating possible nonlinearities in the effect of exogenous predictors, the non-parametric approach has robustness to distributional assumptions (such as normality) of the innovation process in fully specified parametric models.

Given our focus on quantile forecasting, it is also of practical interest to explore if the non-parametric averaging approach can be complemented by parametric models. To motivate such a hybrid approach, consider

Jan C. De Gooijer is the corresponding author.

© 2019 Walter de Gruyter GmbH, Berlin/Boston.

the case of conditional quantile estimation of financial data such as stock returns. In financial applications, conditional quantiles play an essential role in risk assessment. For example, evaluation of Value-at-Risk is a conditional quantile estimation problem. Although the literature on estimating conditional quantiles is large, the overwhelming majority of the approaches rely on the assumption that returns follow a fully specified conditional distribution (such as normality or  $t$ ). As a result, the estimation of conditional quantiles is equivalent to estimating conditional variance (volatility) of returns. The massive literature on GARCH models (volatility modeling) reflect the popularity of such an approach. In addition to their computational convenience, GARCH type models have proven highly successful in modeling financial data. For instance, one key practical appeal of GARCH family of models is that they can parsimoniously capture the persistent influence of long past shocks. In contrast, the averaging method of De Gooijer and Zerom (2019) only allows dependence up to certain autoregressive lags. In other words, the non-parametric method considers a truncated quantile auto-regression (possibly nonlinear) approximation of persistence. If the GARCH type behavior (i. e. infinite ARCH) is indeed an important feature, the averaging method will not completely capture it even with a large autoregressive term.

An important implication of the foregoing discussion is that non-parametric approaches do not necessarily include all parametric models as special cases. Thus, the role of non-parametric and parametric approaches can be complementary. Motivated by this observation, we propose augmenting the non-parametric approach by parametric models. To illustrate the merit of the proposed hybrid quantile averaging approach, we apply it to quantile forecasting of the S&P 500 equity risk premium using a large data set of predictors involving both macroeconomic variables and technical indicators. Our findings show that the hybrid averaging method provides more accurate quantile forecasts than several benchmark methods including the non-parametric averaging method of De Gooijer and Zerom (2019). These findings hold for the full out-of-sample forecasting period as well as for periods of expansion and contraction of the U.S. economy.

The remainder of the paper is organized as follows. Section 2 provides details of the proposed hybrid averaging method including various methodological and numerical issues of the penalization methods. Section 3 describes the empirical data and forecasting procedure. Then, in Section 4, we conduct an extensive out-of-sample forecasting experiment and compare the proposed hybrid quantile averaging method with six, non-hybrid, alternative models/methods. Finally, Section 5 concludes.

## 2 Methodology

We consider the set of observations  $\{Y_t\}_{t=1}^N$  obtained from a strictly stationary time series process  $\{Y_t, t \in \mathbb{Z}\}$  that depends on  $q_y \geq 1$  past values of  $Y_t$ , and on a  $q_z$ -dimensional vector  $\mathbf{Z}_t = (Z_{1,t}, \dots, Z_{q_z,t})'$  that consists of exogenous, possibly lagged, stationary time series. Let  $\mathbf{X}_t = (\mathbf{Y}'_{t-1}, \mathbf{Z}'_t)' \in \mathbb{R}^q$  where  $\mathbf{Y}'_{t-1} = (Y_{t-1}, \dots, Y_{t-q_y})'$ , and  $q = q_y + q_z$ . Given the observed data set  $\{(\mathbf{X}_t, Y_t)\}_{t=1}^N$ , our goal is to obtain the  $h$ -step ( $h \geq 1$ ) out-of-sample  $\tau$ th ( $0 < \tau < 1$ ) conditional quantile of  $Y_{N+h}$  given  $\mathbf{X}_t = \mathbf{X}_N$  which we denote by  $Q_{Y_{N+h}}(\tau | \mathbf{X}_N)$ . To this end, we define the associated process  $\{(\mathbf{X}_t, Y_t^*)\} \in \mathbb{R}^q \times \mathbb{R}$  where the components of the predictor vector  $\mathbf{X}_t = (X_{1,t}, \dots, X_{q,t})'$  are given by

$$X_{j,t} = \begin{cases} Z_{j,t}, & j = 1, \dots, q_z \\ Y_{t+(j-q_z-1)}, & j = q_z + 1, \dots, q \end{cases}$$

and

$$Y_t^* = Y_{t+h+q_y-1}, \quad t = 1, \dots, n, \quad (n = N - h - q_y + 1).$$

With this re-formulation, we obtain  $Q_{Y_{N+h}}(\tau | \mathbf{X}_N)$  directly via

$$Q_{Y_t^*}(\tau | \mathbf{X}_{N-q_y+1}) = \inf\{y: F_{Y_t^*}(y | \mathbf{X}_{N-q_y+1}) \geq \tau\}, \quad (1)$$

where  $F_{Y_t^*}(\cdot | \mathbf{x})$  is the conditional distribution of  $Y_t^*$  given  $\mathbf{X}_t = \mathbf{x}$ .

However, even for moderate dimension  $q$ , non-parametrically estimating  $Q_{Y_t^*}(\tau | \mathbf{X}_{N-q_y+1})$  is very challenging due to the curse of dimensionality. Therefore, extending Chen et al. (2018) into the conditional quantile context, De Gooijer and Zerom (2019) suggested approximating  $Q_{Y_t^*}(\tau | \mathbf{x})$  by a linear combination (averaging) of marginal conditional quantiles  $\theta_j(\tau | x_j) = \inf\{w: F_{Y_t^*}(w | X_{j,t} = x_j) \geq \tau\}$ , i. e.

$$\tilde{Q}_{Y_t^*}(\tau | \mathbf{x}) = \gamma_0(\tau) + \sum_{j=1}^q \gamma_{0,j}(\tau) \hat{\theta}_{0,j}(\tau | x_j), \quad (2)$$

where  $\hat{\theta}_{0,j}(\tau|x_j)$  is some nonparametric estimate of  $\theta_j(\tau|x_j)$  and  $\gamma_{0,j}(\tau)$  are the weights depending on the quantile level  $\tau$ . The key idea of this approach is that the non-parametrically estimated marginal conditional quantile functions carry information about the target  $Q_{Y_i^*}(\tau|\mathbf{x})$ . At the same time, all marginal functions are unlikely to be equally informative. The averaging will assign higher weights to better (more informative) candidate functions. The conditional quantile approximation (2) is non-parametric because we do not impose a parametric structure about the candidate marginal quantile functions  $\theta_j(\tau|x_j)$ .

## 2.1 A Hybrid Approach

We assume that  $m$  candidate parametric models are available with the corresponding estimated marginal quantile functions  $\{\hat{\theta}_{i,1}(\tau|x_{1,t}), \dots, \hat{\theta}_{i,q}(\tau|x_{q,t})\}$  ( $i = 1, \dots, m$ ). For each parametric model  $i$ , we define a hybrid extension of  $\tilde{Q}_{Y_i^*}(\tau|\mathbf{x})$  by

$$\tilde{Q}_{Y_{i,t}^*}^{(H)}(\tau|\mathbf{x}) = \underbrace{\gamma_0(\tau) + \sum_{j=1}^q \gamma_{0,j}(\tau)\hat{\theta}_{0,j}(\tau|x_j)}_{\text{non-parametric}} + \underbrace{\sum_{j=1}^{q^*} \gamma_{i,j}(\tau)\hat{\theta}_{i,j}(\tau|x_j)}_{\text{parametric}}, \quad (i = 1, \dots, m), \quad (3)$$

where  $\{\gamma_{0,j}(\tau)\}_{j=1}^q$  and  $\{\gamma_{i,j}(\tau)\}_{j=1}^{q^*}$  are sets of weights depending on the quantile level  $\tau$ , which we summarize by the set  $\{\gamma_{i,u}^{(H)}(\tau)\}_{u=1}^{q+q^*}$  ( $i = 1, \dots, m$ ). In the equity premium quantile forecasting application (see Section 3),  $q^* = q_z$ . Let the  $i$ th design matrix  $\mathbf{X}_t$  consists of the set of  $q + q^*$  potential predictors:

$$\{\hat{\theta}_{0,1}(\tau|X_{1,t}), \dots, \hat{\theta}_{0,q}(\tau|X_{q,t}), \hat{\theta}_{i,1}(\tau|X_{1,t}), \dots, \hat{\theta}_{i,q^*}(\tau|X_{q^*,t})\};$$

its  $u$ th component will be denoted by  $\hat{\theta}_{i,u}^{(H)}(\tau|X_{u,t})$  ( $u = 1, \dots, q + q^*$ ). So for the  $i$ th model the  $h$ -step ahead hybrid quantile forecast of  $Y_{N+h}$  is given by

$$\tilde{Q}_{Y_{i,t}^*}^{(H)}(\tau|\mathbf{X}_{N-q_y+1}) = \hat{\gamma}_0(\tau) + \sum_{u \in \widehat{\mathcal{M}}_{i,\tau}^{(H)}(h)} \hat{\gamma}_{i,u}^{(H)}(\tau)\hat{\theta}_{i,u}^{(H)}(\tau|X_{u,N-q_y+1}), \quad (4)$$

where  $\widehat{\mathcal{M}}_{i,\tau}^{(H)}(h)$  is an estimate of the set of selected nonzero quantile predictors  $\mathcal{M}_{i,\tau}^{(H)}(h) = \{u: \theta_{i,u}(\tau) \neq 0\}$  associated with  $\tilde{Q}_{Y_{i,t}^*}^{(H)}(\tau|\mathbf{X}_{N-q_y+1})$ , and  $\hat{\gamma}_{i,u}^{(H)}(\tau)$  are the corresponding hybrid quantile regression estimates, or weights. The key idea is to use a set of marginal conditional quantiles as ‘‘pilot’’ forecasts of  $Y_{t+h}$  and later discard irrelevant predictors via appropriate penalization.

## 2.2 Obtaining $\widehat{\mathcal{M}}_{i,\tau}^{(H)}(h)$

Let  $\hat{\boldsymbol{\theta}}_i^{(H)}(\tau|\mathbf{X}_t) = (\hat{\theta}_{i,1}^{(H)}(\tau|X_{1,t}), \dots, \hat{\theta}_{i,q+q^*}^{(H)}(\tau|X_{q+q^*,t}))'$  denote a  $(q + q^*) \times 1$  vector of estimates of the combined marginal quantiles at quantile level  $\tau$  using the  $i$ th parametric model. These estimates and the approximation  $\tilde{Q}_{Y_{i,t}^*}(\tau|\mathbf{x})$  jointly serve as a basis to obtain sparse penalized quantile regression estimators of the  $(q + q^*)$ -dimensional vector of hybrid prediction weights  $\boldsymbol{\gamma}_i^{(H)}(\tau) = (\gamma_{i,1}^{(H)}(\tau), \dots, \gamma_{i,q+q^*}^{(H)}(\tau))'$  ( $i = 1, \dots, m$ ). To this end, we consider the following weighted  $L_1$ -penalized quantile estimator

$$\hat{\boldsymbol{\gamma}}_i^{(H)}(\tau) = \arg \min_{\boldsymbol{\gamma}_i^{(H)}(\tau)} \left\{ n^{-1} \sum_{t=1}^n (\rho_\tau(Y_t^* - \gamma_0^{(H)}(\tau) - \hat{\boldsymbol{\theta}}_i^{(H)}(\tau|\mathbf{X}_t))' \boldsymbol{\gamma}_i^{(H)}(\tau)) + \lambda_n \|\mathbf{w}_i \odot \boldsymbol{\gamma}_i^{(H)}(\tau)\|_1 \right\}, \quad (5)$$

where  $\rho_\tau(z) = \{\tau - I(z < 0)\}z$  is the quantile check function,  $I(\cdot)$  the indicator function,  $\lambda_n > 0$  is a tuning parameter, depending on  $n$ ,  $\mathbf{w}_i = (w_{i,1}, \dots, w_{i,q+q^*})'$  is a vector of nonnegative weights, and  $\|\mathbf{w}_i \odot \boldsymbol{\gamma}_i^{(H)}(\tau)\|_1$  with  $\odot$  the Hadamard (or direct) product of two vectors. Given the minimizer of eq. (5), we define the set of relevant predictors at quantile level  $\tau$ , forecast horizon  $h$ , and using parametric model  $i$ , by

$$\widehat{\mathcal{M}}_{i,\tau}^{(H)}(h) = \{u: \hat{\gamma}_{i,u}^{(H)}(\tau) \neq 0\}, \quad (6)$$

where  $\hat{\gamma}_{i,u}^{(H)}(\tau)$  ( $u = 1, \dots, q + q^*$ ) is an estimate of the  $u$ th component of  $\boldsymbol{\gamma}_i^{(H)}(\tau)$ .

### 2.3 Choice of Penalty

When  $\mathbf{w} = \mathbf{1}_{q+q^*}$ , a vector of ones, the last term of eq. (5) becomes the LASSO (least absolute shrinkage and selection operator) penalty function  $p_{\lambda_n}(|\gamma_{i,u}^{(H)}(\tau)|) = \lambda_n |\gamma_{i,u}^{(H)}(\tau)|$  ( $u = 1, \dots, q + q^*; i = 1, \dots, m$ ). Note that LASSO assigns the same penalty to each coefficient  $\gamma_{i,u}^{(H)}(\tau)$ , regardless of whether predictor effects are relevant or irrelevant. This may result in estimation inefficiency and model selection inconsistency (Wang and Leng 2008). ada-LASSO, on the other hand, uses different weights for penalizing irrelevant predictors. In the empirical part of this study, we adopt ada-LASSO with  $\widehat{\gamma}_{i,u}^{(mi)}(\tau) = \widehat{\gamma}_{i,u}^{(LASSO)}(\tau)$ .<sup>1</sup> In order to optimize eq. (5) with ada-LASSO, we use a proximal alternating direction method of multipliers algorithm contained in the R-FHDQR package.<sup>2</sup> As shown by Gu et al. (2018), this algorithm drastically reduces the computational burden as compared to other algorithms for penalized quantile regression in high-dimensions.

### 2.4 Numerical Issues and Choice of $\lambda_n$

Medeiros and Mendes (2014, 2016) studied the asymptotic properties of ada-LASSO when the errors of a linear regression model are non-Gaussian and may be conditionally heteroskedastic. These authors derived asymptotic properties of sign consistency for ada-LASSO. Moreover, the ada-LASSO estimator has oracle properties; see also Audrino and Camponovo (2017). Recently, Han and Tsay (2019) investigated the properties of the LASSO estimators of a linear regression model in the presence of serial dependence (AR models) in both the covariate vector and the errors. In particular, these authors provide model selection consistency of LASSO estimators under certain conditions on the tuning parameter  $\lambda_n$ .

To select the tuning parameter  $\lambda_n$ , we consider the following prediction-based criterion. For a given  $\lambda_n$ , let  $\widehat{\boldsymbol{\gamma}}_{\lambda_n}^{(H)}(\tau)$  be the  $(q + q^*)$ -dimensional vector of penalized estimates, and  $|\widehat{\mathcal{M}}_{\lambda_n, \tau}^{(H)}(h)|$  be the corresponding number of non-zero estimates. Then, we choose  $\lambda_n$  as the value that minimizes the following high-dimensional BIC criterion

$$\begin{aligned} \text{QBIC}_{\tau}(\lambda_n) &= \log \left( \sum_{t=1}^n \rho_{\tau}(Y_t^* - \widehat{\boldsymbol{\gamma}}_{\lambda_n, 0}^{(H)}(\tau) - \widehat{\boldsymbol{\theta}}^{(H)}(\tau | \mathbf{X}_t))' \widehat{\boldsymbol{\gamma}}_{\lambda_n}^{(H)}(\tau) \right) \\ &\quad + |\widehat{\mathcal{M}}_{\lambda_n, \tau}^{(H)}(h)| \frac{\log(q+q^*) \log(\log(n))}{2n}. \end{aligned} \quad (7)$$

Using simulations, Sherwood and Wang (2016) report that  $\text{QBIC}_{\tau}$  can be effective in terms of finite-sample prediction accuracy and predictor selection. De Gooijer and Zerom (2019, Sec. 3.4) show results of a robustness check on  $\lambda_n$  for the LASSO penalty in a time series penalized averaging framework.

## 3 Models, Data, and Forecasting Procedure

### 3.1 Parametric Models

In Section 4, we evaluate the ability of the hybrid conditional quantile approach in forecasting the one-step ahead ( $h = 1$ ) risk premium of the monthly S&P 500 index, denoted by  $R_t$ , using a large set of potential predictors  $Z_{j,t}$ . In particular, we adopt the following parametric ( $m = 4$ ) DGPs.

1. Time-varying mean (TVM) model with constant volatility:

$$R_{t+1} = \beta_{0,j} + \beta_{1,j} Z_{j,t} + \varepsilon_{j,t+1}, \quad \varepsilon_{j,t+1} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2). \quad (8)$$

2. Constant, or prevailing mean (PM) model with EGARCHZ volatility (PM-EGARCHZ):

$$\begin{cases} R_{t+1} = \beta_{0,j} + \varepsilon_{j,t+1}, & \varepsilon_{j,t+1} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{j,t+1}^2), \\ \log(\sigma_{j,t+1}^2) = \delta_{0,j} + \delta_{1,j} Z_{j,t} + \delta_{2,j} \log(\sigma_{j,t}^2) + \delta_{3,j} \left| \frac{\varepsilon_{j,t}}{\sigma_{j,t}} \right| + \delta_{4,j} \frac{\varepsilon_{j,t}}{\sigma_{j,t}}. \end{cases} \quad (9)$$

3. Time-varying mean model with EGARCH volatility (TVM-EGARCH):

$$\begin{cases} R_{t+1} = \beta_{0,j} + \beta_{1,j} Z_{j,t} + \varepsilon_{j,t+1}, & \varepsilon_{j,t+1} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{j,t+1}^2), \\ \log(\sigma_{j,t+1}^2) = \delta_{0,j} + \delta_{2,j} \log(\sigma_{j,t}^2) + \delta_{3,j} \left| \frac{\varepsilon_{j,t}}{\sigma_{j,t}} \right| + \delta_{4,j} \frac{\varepsilon_{j,t}}{\sigma_{j,t}}. \end{cases} \quad (10)$$

#### 4. Time varying mean model with EGARCHZ volatility (TVM-EGARCHZ):

$$\begin{cases} R_{t+1} = \beta_{0,j} + \beta_{1,j}Z_{j,t} + \varepsilon_{j,t+1}, & \varepsilon_{j,t+1} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{j,t+1}^2), \\ \log(\sigma_{j,t+1}^2) = \delta_{0,j} + \delta_{1,j}Z_{j,t} + \delta_{2,j} \log(\sigma_{j,t}^2) + \delta_{3,j} \left| \frac{\varepsilon_{j,t}}{\sigma_{j,t}} \right| + \delta_{4,j} \frac{\varepsilon_{j,t}}{\sigma_{j,t}}. \end{cases} \quad (11)$$

Models (8)–(11) were analyzed by Cenesizoglu and Timmermann (2012) in the context of conditional mean prediction. Clearly, models (8)–(10) are all nested within model (11). However, too many model parameters can lead to efficiency losses in the out-of-sample forecasting performance due to estimation uncertainty and model misspecification. So, we gain flexibility by considering a wider set of models than just using model (11). There is a large literature in finance exploring the predictability of the conditional mean and/or variance of stock returns; see, e. g. Rapach and Zhou (2013) for a review.

### 3.2 Data

To help obtain the one-step-ahead quantile forecasts for  $R_t$ , we consider a total of 28 exogenous predictors  $Z_{j,t}$  ( $j = 1, \dots, 28$ ) consisting of 14 macroeconomic predictors and 14 technical indicators covering the time period 1951:01–2016:12 (792 observations). The macroeconomic predictors are summarized as follows: dividend-price ratio (DP); dividend yield (DY); earnings-price ratio (EP); dividend-payout ratio (DE); equity risk premium volatility (RVOL); book-to-market ratio (BM); net equity expansion (NTIS); treasury bill rate (TBL); long-term yield (LTY); long-term return (LTR); term spread (TMS); default yield spread (DFY); default return spread (DFR); and inflation (INFL). These predictors have been the subject of conditional mean forecasting studies by Welch and Goyal (2008), Cenesizoglu and Timmermann (2012), and Pedersen (2015), among others. Neely et al. (2014) apply single-variable mean regressions to predict  $R_t$  using the above data set. Lima and Meng (2017) and Meligkotsidou et al. (2014, 2019) consider the 14 macroeconomic predictors within the context of quantile prediction.<sup>3</sup>

The 14 technical indicators fall into three categories:

- A moving average (MA) rule that generates a buy (sell) signal  $S_{i,t}$  with  $S_{i,t} = 1$  if  $\text{MA}_{s,t} \geq \text{MA}_{\ell,t}$ ; 0 otherwise. Here  $\text{MA}_{j,t} = (1/j) \sum_{i=0}^{j-1} P_{t-i}$  for  $j = s, \ell$  with  $P_t$  the level of the stock price index, and  $s(\ell)$  the length of the short (long) MA ( $s < \ell$ ). The resulting MA indicator is denoted by  $\text{MA}(s, \ell)$  with  $s = 1, 2, 3$  and  $\ell = 9, 12$ .
- A momentum (MOM) rule where  $S_{i,t} = 1$  if  $P_t \geq P_{t-m}$ ; 0 otherwise. The momentum indicator is denoted by  $\text{MOM}(m)$  with  $m = 9, 12$ .
- A trading volume rule using MAs of “on-balance” volumes (OBVs) defined as  $\text{OBV}_t = \sum_{k=1}^t \text{VOL}_k D_k$  where  $\text{VOL}_k$  is a measure of the trading volume during period  $k$  and  $D_k$  is a binary variable that takes a value of 1 if  $P_k - P_{k-1} \leq 0$  and  $-1$  otherwise. Then  $S_{i,t} = 1$  if  $\text{MA}_{s,t}^{\text{OBV}} \geq \text{MA}_{\ell,t}^{\text{OBV}}$ ; 0 otherwise. Here  $\text{MA}_{j,t}^{\text{OBV}} = (1/j) \sum_{i=0}^{j-1} \text{OBV}_{t-i}$  for  $j = s, \ell$ . The corresponding indicator is denoted by  $\text{VOL}(s, \ell)$  with  $s = 1, 2, 3$  and  $\ell = 9, 12$ .

The monthly macroeconomic predictors are available from Amit Goyal’s website at <http://www.hec.unil.ch/agoyal/>. The technical indicators can be constructed using Matlab code made available with the paper of Neely et al. (2014). The S&P 500 volume data, used to compute the monthly VOL indicators, can be downloaded from Yahoo Finance.

### 3.3 Forecasting Procedure

We generate one-step ahead quantile forecasts, i. e.  $R_{n+1}$ , with a recursive window scheme. Our first forecast origin/base is 1965:12 and hence the first sample covers the period  $t \in [1951:01, 1964:12]$  ( $n = 168$ ). By recursive, we mean that our combined conditional quantile procedure is repeated by extending the forecast origin, one observation at a time, up to and including 2016:11. Thus, our last time series sample covers the period 1951:01–2016:11 ( $n = 791$ ). With such a recursive design, our forecasts cover the period 1966:01–2016:12 and a total of  $F = 624$  conditional quantile forecasts are generated.

In Section 4, we report recursive prediction results for the following seven models/methods. Note that forecasts are for each quantile level  $\tau \in \{0.1, 0.5, 0.9\}$ . Methods 1–5 are logically connected in the sense that they are based on the concept of averaging or combining forecasts. For example, Method 1 allocates all the weight to one predictor. Methods 2–5 assign weights (equal or varying) to all the predictors. Methods 6 and 7 are popular approaches used in the quantile regression literature. Unlike Methods 6 and 7, penalized averaging does not require additivity.

1. Benchmark Model (BM). For each parametric model  $i \in \{1, 2, 3, 4\}$  (corresponding to (8)–(11)) and each time period  $t \in \{1951:01\text{--}2016:11\}$ , we select the predictor from the set  $Z_{j,t}$  ( $j \in \{1, \dots, 28\}$ ) that produces the lowest quantile forecast error. This results in  $F$  total forecast errors, and the average of these  $F$  values will be reported as a measure of accuracy, see Section 3.4. In this way, we will have four forecast benchmarks, denoted by BM- $i$  ( $i = 1, \dots, 4$ ).
2. Equal weighting (EW). For each parametric model  $i \in \{1, 2, 3, 4\}$  and each time period  $t \in \{1951:01\text{--}2016:11\}$ , we obtain quantile forecasts by equally averaging the quantile forecasts obtained from the 28 predictors. In this way, we will have four forecast equally weighted methods, denoted by EW- $i$  ( $i = 1, \dots, 4$ ). Unlike the benchmark model, all predictors contribute equally to quantile forecasts regardless their relative forecast accuracy.
3. Penalized Averaging (PA). Forecasts are based on weighted averaging similar to the EW approach, but the weights are data driven obtained by penalized quantile averaging. In terms of the proposed hybrid approach, see eq. (3), we only consider the second component (or the parametric part). Thus, PA is a special case of the hybrid method. In this way, we will have four methods, denoted by PA- $i$  ( $i = 1, \dots, 4$ ).
4. Penalized Averaging of Nonparametric quantiles (PA-NP). This approach is also a special case of the hybrid method, see eq. (3), where only the non-parametric component is considered. This approach was proposed in De Gooijer and Zerom (2019).
5. Hybrid (H). This is the hybrid approach introduced in this paper (3) where parametric and non-parametric approaches are combined. In this way, there are four hybrid methods, denoted by H- $i$  ( $i = 1, \dots, 4$ ), and 56 potential predictors.
6. Penalized Linear Quantile Regression (P-Lin-QR). This approach is well known in the literature. It does not allow for nonlinearities in the predictors.
7. Additive QR (Ad-QR). This approach is a generalization of (6) where each additively entered predictor has a non-parametric (possibly nonlinear) effect. Unlike the approaches (3), (4), (5) and (6), we do not conduct predictor selection for the Ad-QR. Instead, we use the predictor selection results of approach (4) and implement a low dimensional additive models using the R-gamboost package. To by-pass predictor selection, we set the hyper parameter (called  $m_{stop}$ ) at 5,000.

### 3.4 Evaluating Predicted Quantiles

Let  $e_t^{(\cdot)} = Y_{N+1} - \widehat{Q}_{Y_t^*}^{(\cdot)}(\tau | \mathbf{X}_{N-q_y+1})$  ( $t = 1, \dots, n$ ) be the one-step ahead out-of-sample quantile prediction error (QPE) at quantile level  $\tau$ , where the superscript  $(\cdot)$  refers either to the one of the seven forecast approaches. In this paper, we evaluate the accuracy of the quantile forecasts using an average of the check loss function of QPE values, i. e.

$$\widehat{L}_\tau^{(\cdot)} = \frac{1}{n} \sum_{t=1}^n \rho_\tau(e_t^{(\cdot)}), \quad (12)$$

which is an estimate of the expected loss  $L_\tau^{(\cdot)} = \mathbb{E}[Y_{N+1} - Q_{Y_t^*}^{(\cdot)}(\tau | \mathbf{X}_{N-q_y+1})]$ .  $\widehat{L}_\tau^{(\cdot)}$  weights the difference between the observed value  $Y_{N+1}$  and the forecasted quantile  $\widehat{Q}_{Y_t^*}^{(\cdot)}(\tau | \mathbf{X}_{N-q_y+1})$  by  $(1 - \tau)$  when  $Y_{N+1}$  is lower than the  $\tau$ th quantile, and by  $\tau$  when  $Y_{N+1}$  exceeds the quantile. Hence, eq. (12) is a natural way to evaluate quantile forecasts; see, e. g. Giacomini and Komunjer (2005).

We also make pairwise comparisons between each of the four BM forecast errors  $e_{t,i}^{(BM)} = Y_{N+1} - \widehat{Q}_{\tau,i}^{(BM)}$  ( $i = 1, \dots, 4; N \equiv n = 168, \dots$ ) and the quantile forecast errors obtained from the hybrid method. In particular, we assess their differences via a Diebold–Mariano (DM) type test statistic. To this end, let  $e_t^{(A)}$  and  $e_t^{(B)}$  be the associated one-step ahead QPEs for the pair  $(A, B)$  of methods. Then, for fixed  $\tau$ , the null- and alternative hypotheses of interest are, respectively,

$$\mathbb{H}_0: \mu_\tau \equiv \mathbb{E}[\rho_\tau(e_t^{(A)}) - \rho_\tau(e_t^{(B)})] = 0 \text{ and } \mathbb{H}_1: \mathbb{E}[\rho_\tau(e_t^{(A)}) - \rho_\tau(e_t^{(B)})] < 0. \quad (13)$$

The corresponding loss differential is defined as

$$D_{t,\tau}(A, B) = \{\tau - I(e_t^{(A)} < 0)\}e_t^{(A)} - \{\tau - I(e_t^{(B)} < 0)\}e_t^{(B)}.$$

The null hypothesis that method  $A$  produces as accurate forecasts as method  $B$  can be tested using the test statistic

$$\bar{D}_\tau(A, B) / \sqrt{\text{Var}(\bar{D}_\tau(A, B))} \sim \mathcal{N}(0, 1), \quad (14)$$

where  $\bar{D}_\tau(A, B)$  is the average over  $t$  of  $D_{t, \tau}(A, B)$ . Under the alternative hypothesis, we specify a one-sided test (right tail), so that rejection of the null indicates that method  $A$  is more accurate than method  $B$ . For  $h = 1$ , a consistent estimator of  $\text{Var}(\bar{D}_\tau(A, B))$  is given by the sample variance of  $D_{t, \tau}(A, B)$ . For  $h > 1$ , one may use the Newey–West estimator for the variance.<sup>4</sup>

## 4 Forecast Results

Throughout this section, we present predictor selection and forecasting results for the full sample and for periods of contraction and expansion of the U.S. business cycle, as dated by the U.S. National Bureau of Economic Research. This provides insight in the relative strength of the proposed combined method in forecasting conditional quantiles of  $R_t$  during each time period. For the full sample period  $F = 624$ , and  $F = 534$  (90) for the expansion (contraction) time periods.

### 4.1 Results

#### 4.1.1 Predictor Selection

Table 1 shows predictor selection frequencies of five models/methods for the full sample period. To ease interpretation, we only report those predictors which are selected in at least 5% of the times for the three quantile levels  $\tau$ .



Table 1: Frequency of predictors which are selected in at least 5% of the times. Full prediction sample.

Method	DP	DY	EP	DE	RVOL	BM	NTIS	TBL	LTY	LTR	TMS	DFY	DFR	INFL	MA (1,9)	MA (2,12)	MA (3,9)	MOM (9)	MOM (12)	VOL (2,9)	VOL (1,12)	VOL (2,12)	VOL (3,9)	
<b>Quantile level <math>\tau = 0.1</math></b>																								
BM-1					10.3		17.0	10.1	12.8	27.7														
BM-2					13.3		8.7	17.1	8.7	6.1		10.7	5.9						6.7					
BM-3					21.5	6.1	15.1	5.8	18.1	14.1			5.8											
BM-4					6.4		11.2	18.6	19.1	15.7		6.6	6.4											
PA-1		12.9		10.9	14.8	8.6	6.6	5.7	5.1	11.3	10.2													
PA-2	34.0		5.5		10.1	5.8	6.6		5.1	5.0	12.5													
PA-3	6.1	14.9		8.2	12.6	13.7	6.4		5.4	8.4			5.8											
PA-4		17.3			9.0	14.1	15.3		6.9	10.0		10.7	5.4											
PA-NP		5.3						10.7		12.7		7.3	7.5					5.8	8.9			7.9		
P-Lin- QR		7.0						7.2	9.3	9.1	6.4	8.6	7.4					9.2				5.8		
H-1		6.4*						8.0*		9.6*		7.4*	7.1*						9.9*			6.0*		
H-2		5.7*			5.1				7.8*	8.2*	5.0	7.2	5.1*						8.7*			5.4*		
H-3		5.1*						7.1*	7.6*	8.8*		6.5*	5.6*						8.1*			6.1*		
H-4		5.8*			7.0			7.2*	5.4*	8.0*		7.1	5.0*						8.6*			5.4*		
<b>Quantile level <math>\tau = 0.5</math></b>																								
BM-1					10.1		17.6	10.6	13.1	26.3														
BM-2				8.3			34.8	9.9	9.5			26.0												
BM-3					16.5	5.3	25.8		11.1	23.7														
BM-4					5.6		19.2	13.8	14.4	25.0														
PA-1		17.0			7.2			8.3	30.6	9.7									7.0			8.6		
PA-2																								

Table 1: (continued)

Method	DP	DY	EP	DE	RVOL	BM	NTIS	TBL	LTY	LTR	TMS	DFY	DFR	INFL	MA (1,9)	MA (2,12)	MA (3,9)	MOM (9)	MOM (12)	VOL (1,12)	VOL (2,9)	VOL (2,12)	VOL (3,9)	
<b>Quantile level <math>\tau = 0.5</math></b>																								
PA-3				44.0	9.2	5.4	8.0	9.9													6.8			
PA-4					44.7			10.1	13.5								8.0					7.5		
PA-NP				12.4	18.1	6.7	21.3	11.7	6.0											15.7				
P-Lin-QR					16.6	12.8	21.8	15.7	9.8			5.7	6.0											
H-1				5.4*	9.0*	10.7*	10.4*	6.3,												8.6*				
								5.9*																
H-2				6.1*	10.2*	11.7*	11.1*	13.1*	5.7*											9.2*				
H-3				5.9*	5.4*	10.0*	10.8*	6.2*				6.6								8.9*				
H-4				5.8*	9.4*	10.7*	10.9*	6.2*				6.7								8.9*				
<b>Quantile level <math>\tau = 0.9</math></b>																								
BM-1	8.0		11.7			7.0	8.8		20.3															
BM-2			19.7	25.5		12.3		17.6	12.8															
BM-3			29.2	12.4		34.0						15.6												
BM-4			27.8	21.4			29.2																	
PA-NP	7.6					11.3	10.6	5.7		5.7	5.1													
P-Lin-QR	6.1		6.3			11.5	8.0	7.6		8.3	5.3	9.8											7.2	
H-1	6.3*					14.5*	9.7*	7.4		9.7*	6.1*												8.9*	
H-2	6.5*					15.1	10.7*			10.0*	8.8*		13.2										9.3*	
H-3	6.6*					14.8,	10.3*			8.5	10.6*	8.0*											7.0*	
						8.6*																		
H-4	6.6*					15.4	11.2*			10.1*	8.7*		13.1										9.4*	

Notes: (i) BM = Benchmark Model, PA = Penalized Averaging, PA-NP = Penalized Averaging with NP marginal quantiles, P-Lin-QR = Penalized linear model with quantile regression, H = Hybrid; (ii) \* denotes NP marginal predictors selected by the hybrid method.

Five important observations follow from Table 1. First, it is clear that only a small subset of the total number of macroeconomic predictors plays a role in the selection of the final set of predictors at all quantile levels. In particular this applies to the set of predictors {RVOL, NTIS, TBL, LTR, TMS}. Second, only a few technical indicators are selected as a part of the overall set of predictors. Overall, MA(3,9), MOM(12), VOL(1,12), and VOL(3,9) contribute to the forecasting performance of the hybrid quantile averaging approach. But as can be seen from the selection frequencies, their contribution is relatively low. These results differ from the study by Neely et al. (2014), who showed that technical indicators provide complementary information to conditional mean forecasting. Third, we find considerable differences in the selection of predictors at the three quantile levels. For instance, INFL contributes to the forecasting performance of all models/methods at  $\tau = 0.1$  while at  $\tau = 0.9$  the selection frequency of this variable is below 5%. Also the macroeconomic predictor DP is selected frequently at  $\tau = 0.9$  by the hybrid approach and hardly anywhere else by other models/methods. Fourth, very few macroeconomic predictors have more than 10% selection scores. Some exceptions are DE (44.0%) with method PA-3, RVOL (44.7%) with method PA-4, and TBL (34.8%) with method BM-2. Lastly, nonparametric marginal predictors are frequently selected by the hybrid approach, as denoted by the superscript \*. Contributory predictors obtained from parametric conditional quantiles are RVOL, DFY and DFR at  $\tau = 0.1$ , LTR and DFR at  $\tau = 0.5$ , and RVOL, TBL, LTY, and DFR at  $\tau = 0.9$ .

#### 4.1.2 Forecasting Performance

Table 2 presents ratios of thick loss functions,  $\rho_\tau(e_t^{(\text{Method})})/\rho_\tau(e_t^{(\text{BM-}i)})$  ( $i = 1, \dots, 4$ ), averaged over  $F = 624$  predictions, i. e. the full prediction sample. Values less (greater) than one indicate that a particular method is more (less) accurate than a particular benchmark model (BM).

**Table 2:** Comparing conditional quantile averaging methods. The entries are ratios of thick loss functions,  $\rho_\tau(e_t^{(\text{Method})})/\rho_\tau(e_t^{(\text{BM-}i)})$ , averaged over  $F = 624$  predictions ( $i = 1, \dots, 4$ ). Embolded entries show the lowest ratios for each  $\tau$  value and each BM model.

Method	$\tau = 0.1$				$\tau = 0.5$				$\tau = 0.9$			
	BM-1	BM-2	BM-3	BM-4	BM-1	BM-2	BM-3	BM-4	BM-1	BM-2	BM-3	BM-4
EW-1	0.968	1.022	1.018	1.013	0.981	0.990	0.974	0.976	0.970	0.983	0.979	0.954
EW-2	0.927	0.979	0.975	0.970	0.992	1.001	0.985	0.987	0.919	0.931	0.927	0.903
EW-3	0.927	0.978	0.974	0.969	0.983	0.992	0.976	0.978	0.909	0.920	0.917	0.893
EW-4	0.920	0.971	0.967	0.962	0.983	0.991	0.976	0.978	0.911	0.923	0.919	0.896
PA-1	0.865	0.913	0.909	0.905	0.934	0.942	0.928	0.929	0.903	0.914	0.911	0.888
PA-2	0.861	0.909	0.905	0.900	0.945	0.954	0.939	0.941	0.882	0.893	0.890	0.867
PA-3	0.877	0.925	0.921	0.917	0.941	0.949	0.934	0.936	0.874	0.885	0.882	0.859
PA-4	0.867	0.915	0.911	0.906	0.940	0.948	0.933	0.935	0.888	0.900	0.896	0.873
PA-NP	0.873	0.921	0.918	0.913	0.913	0.921	0.906	0.908	0.906	0.918	0.914	0.891
P-Lin-QR	1.185	1.251	1.246	1.239	0.959	0.968	0.952	0.954	0.936	0.948	0.945	0.920
Ad-QR	0.836	0.883	0.879	0.875	0.909	0.917	0.903	0.905	0.940	0.952	0.948	0.924
H-1	0.790	0.833	0.830	0.826	<b>0.896</b>	0.904	<b>0.890</b>	0.892	0.834	0.844	0.841	0.820
H-2	<b>0.775</b>	<b>0.818</b>	<b>0.815</b>	0.811	0.898	0.906	0.892	0.893	<b>0.818</b>	<b>0.828</b>	<b>0.825</b>	<b>0.804</b>
H-3	0.792	0.836	0.832	0.828	0.897	0.905	<b>0.890</b>	0.892	0.825	0.836	0.832	0.811
H-4	<b>0.775</b>	<b>0.818</b>	<b>0.815</b>	<b>0.810</b>	<b>0.896</b>	0.904	<b>0.890</b>	<b>0.891</b>	0.824	0.834	0.831	0.810

**Notes:** (i) EW = Equal Weighting, BM = Benchmark Model, PA = Penalized Averaging; PA-NP = Penalized Averaging with NP marginal quantiles, P-Lin-QR = Penalized linear model with quantile regression and Ad-QR = Additive quantile regression; (ii) For the purpose of this table, the entries are measured up to three decimal figures such that more than one ratio of thick loss functions may have the lowest value.

The hybrid approach has the best forecasting results relative to the BM models and across all values of  $\tau$ . Indeed, without any exception, the gains are considerable. Also these gains are present over all (semi-)parametric

models and methods. The ability of the hybrid approach to capture nonlinear effects (e. g. volatility) in a simple way is a likely reason for this success.

Two other results in Table 2 are noteworthy. First, the EW-2–EW-4 quantile forecasts are better than the forecasts obtained from BM-1 to BM-4 at all quantile levels. This is interesting, since the EW combination approach uses fixed (non-random) weights of conditional quantile forecasts across all time periods while the BM approach uses random weights obtained by choosing the best predictor at each time period. Second, the Ad-QR approach has lower thick loss function values than the P-Lin-QR approach across almost all BM models and  $\tau$  values. This is another indication that allowing for nonlinearities in a quantile forecasting approach can lead to systematic improvements.

The use of the test statistic  $\bar{D}_\tau$  to the pair (H- $i$ , BM- $i$ ) ( $i = 1, \dots, 4$ ) resulted in very small  $p$ -values at all quantile levels  $\tau$ . Thus, indicating that there are no benefits in using the BM approach. We also computed  $\bar{D}_\tau$  for pairwise comparisons between the hybrid approach and all other models/methods.<sup>5</sup> In all cases, the  $p$ -values do not exceed the 5 % nominal significance level. Hence, in summary, we conclude that the hybrid averaging approach yields statistically more accurate quantile forecasts than the (semi-)parametric models/methods.

## 4.2 Expansion and Contraction

Table 3, presents model selection frequencies of the hybrid approach for the expansion (Panel A) and contraction (Panel B) period. As expected from the results presented in Table 1, the list of selected predictors no longer includes the predictors EP, BM, LTY, MOM(9), and VOL(2,12). Panel A indicates that the set of macroeconomic predictors {DY, DE, RVOL, NTIS, TBL, LTR, TMS} have considerable forecasting power at all quantile levels. On the other hand, Panel B indicates that during periods of recession (contraction), the set of macroeconomic predictors {LTR, TMS, DFY, INFL} contribute to the forecasting performance of the hybrid quantile averaging approach. So, except for LTR, there is no overlap between both sets, given the 5 % selection threshold. Further, we see from Table 3 that only a few technical indicators contribute to the prediction performance of the hybrid approach. Overall, the list of selected predictors in Table 3 is not markedly different from the one given in Table 1, and the selection frequencies are about the same. It is also evident from Table 3 that the conditional parametric quantile estimates of the predictors RVOL, LTR, DFR, and INFL all contribute to the forecasting power of the hybrid approach.

Table 4 presents averaged ratios of the thick loss functions  $\rho_\tau(e_t^{(\text{Method})})/\rho_\tau(e_t^{(\text{BM-}i)})$  ( $i = 1, \dots, 4$ ) for the periods of expansion and contraction. For the expansion period (Panel A), the emboldened entries indicate that the hybrid quantile approach has higher predictive power than all other methods/models irrespective of the quantile level  $\tau$ . These findings are in line with the earlier evidence for the full sample period. Turning to the contraction period (Panel B), we see similar improvements of the hybrid quantile approach in the case of  $\tau = 0.1$  and  $\tau = 0.9$ . For  $\tau = 0.5$ , it appears that the prediction performance of Ad-QR is slightly “better” than the four HW methods. So the hybrid forecasting results are quite robust in the above two sub-sample periods.

**Table 3:** Frequency of predictors which are selected in at least 5 % of the times. Expansion and contraction period.

$\tau$	Method	DP	DY	DE	RVOL	NTIS	TBL	LTR	TMS	DFY	DFR	INFL	MA(1,9)	MA(2,12)	MA(3,9)	MOM(12)	VOL(1,12)	VOL(2,9)	VOL(3,9)	
<b>Panel A: Expansion</b>																				
0.1	H-1		6.6*				7.8		9.4*			7.3*			6.4*	9.8*			6.3*	
	H-2		6.0*		5.2			7.8*	8.1*	5.0	7.2	5.0*			6.4*	8.6*			5.7*	
	H-3		5.2*		9.8		6.7	7.2*	8.7*			6.5*			8.1*	9.8*			6.4*	
	H-4		6.0*		7.1		7.0*	5.3*	7.9*						6.6*	8.6*			5.7*	
0.5	H-1			5.2*	9.5*	10.8*	10.6*	6.7,						6.4*	5.0*		8.7*			
								5.6*												
	H-2		5.9*		10.3*	11.7*	11.3*	13.1*	5.4*					8.5*	5.8*		9.3*			
	H-3		5.7*		5.5*	10.0*	11.1*	7.0,						8.2*			9.9*			
							5.8*													
H-4		5.6*		9.9*	10.7*	11.1*	7.1,						8.2*			9.0*				
							5.8*													
0.9	H-1		6.1*		14.1*		10.1*	7.1	10.0*	6.1*									9.2*	
	H-2		6.5*		14.9		11.2*		10.4*	8.2*	12.8								9.7*	
	H-3		6.6*		4.7,		10.8*	8.6	11.2*	8.2*									7.3*	
							8.0*													
H-4		6.6*		15.2		11.6*		10.5*	9.0*	12.8									9.8*	
<b>Panel B: Contraction</b>																				
0.1	H-1		5.4*					9.6*	5.8*	10.5*					5.5*	10.6*				
	H-2						6.0*	6.0*	8.1*	9.0*	7.1				5.4*	9.1*				
	H-3		5.2*		10.4		9.8*	9.8*	9.4*	9.6*					8.0*	10.4*				
	H-4				6.9		8.9*	8.9*	6.1*	8.6*	7.1				6.4*	9.1*				
0.5	H-1			5.4*				9.6*	5.8*	10.5*					5.5*	10.6*				
	H-2						6.0*	6.0*	8.1*	9.0*	7.1				5.5*	9.1*				
	H-3				10.4		9.8*	9.8*	9.4*	9.6*					6.4*	9.1*				
	H-4				6.9		8.9*	8.9*	6.0*	8.6*	7.1				6.4*	9.1*				
0.9	H-1			5.4*				9.6*	5.8*	10.5*					5.5*	10.6*				
	H-2						6.0*	6.0*	8.1*	9.0*					5.5*	9.1*				
	H-3				10.4		9.8*	9.8*	9.4*	9.6*					6.4*	9.1*				
	H-4				6.9		8.9*	8.9*	6.0*	8.6*	7.1				6.4*	9.1*				
0.9	H-1			5.4*				9.6*	5.8*	10.5*					5.5*	10.6*				
	H-2						6.0*	6.0*	8.1*	9.0*					5.4*	9.1*				
	H-3				10.4		9.8*	9.8*	9.4*	9.6*					8.0*	10.4*				
	H-4				6.9		8.9*	8.9*	6.0*	8.6*	7.1				6.4*	9.1*				
0.9	H-1			5.4*				9.6*	5.8*	10.5*					5.5*	10.6*				
	H-2						6.0*	6.0*	8.1*	9.0*					5.4*	9.1*				
	H-3				10.4		9.8*	9.8*	9.4*	9.6*					8.0*	10.4*				
	H-4				6.9		8.9*	8.9*	6.0*	8.6*	7.1				6.4*	9.1*				

Notes: (i) H = Hybrid; (ii) \* denotes NP marginal predictors selected by the hybrid method.

**Table 4:** Comparing conditional quantile averaging methods for the expansion and contraction period. The entries are ratios of thick loss functions,  $\rho_\tau(e_i^{\text{Method}})/\rho_\tau(e_i^{\text{BM-1}})$ , averaged over  $F = 534$  (expansion) and  $F = 90$  (contraction) predictions ( $i = 1, \dots, 4$ ). Embolded entries show the lowest ratios for each  $\tau$  value, each BM model, and each time period.

Method	$\tau = 0.1$				$\tau = 0.5$				$\tau = 0.9$			
	BM-1	BM-2	BM-3	BM-4	BM-1	BM-2	BM-3	BM-4	BM-1	BM-2	BM-3	BM-4
<b>Panel A: Expansion</b>												
EW-1	0.999	1.005	1.000	1.000	0.985	0.995	0.983	0.982	0.955	0.973	0.962	0.942
EW-2	0.988	0.993	0.988	0.986	0.992	1.002	0.990	0.989	0.910	0.928	0.917	0.898
EW-3	0.987	0.993	0.988	0.986	0.986	0.996	0.984	0.983	0.901	0.919	0.907	0.889
EW-4	0.986	0.991	0.986	0.985	0.986	0.996	0.984	0.983	0.905	0.923	0.912	0.893
PA-1	0.945	0.950	0.945	0.943	0.936	0.945	0.934	0.934	0.880	0.897	0.886	0.867
PA-2	0.925	0.930	0.925	0.923	0.946	0.955	0.944	0.943	0.874	0.891	0.880	0.862
PA-3	0.931	0.936	0.930	0.929	0.942	0.951	0.940	0.940	0.868	0.885	0.874	0.855
PA-4	0.939	0.944	0.939	0.937	0.941	0.950	0.939	0.938	0.869	0.886	0.875	0.857
PA-NP	0.955	0.961	0.956	0.954	0.929	0.938	0.927	0.926	0.907	0.924	0.913	0.894
P-Lin-QR	1.375	1.382	1.375	1.373	0.958	0.968	0.956	0.955	0.926	0.944	0.932	0.913
Ad-QR	0.948	0.953	0.948	0.946	0.945	0.954	0.943	0.942	0.965	0.984	0.972	0.952
H-1	0.885	0.890	0.885	0.884	<b>0.917</b>	<b>0.926</b>	<b>0.915</b>	<b>0.914</b>	0.838	0.854	0.844	0.826
H-2	<b>0.874</b>	<b>0.878</b>	<b>0.874</b>	<b>0.872</b>	0.918	0.927	0.916	0.916	<b>0.819</b>	<b>0.835</b>	<b>0.824</b>	<b>0.807</b>
H-3	0.892	0.896	0.892	0.890	<b>0.917</b>	<b>0.926</b>	<b>0.915</b>	<b>0.914</b>	0.826	0.842	0.832	0.814
H-4	0.877	0.881	0.877	0.875	<b>0.917</b>	<b>0.926</b>	<b>0.915</b>	<b>0.914</b>	0.825	0.841	0.830	0.813
<b>Panel B: Contraction</b>												
EW-1	0.887	1.076	1.075	1.059	0.969	0.974	0.946	0.956	1.030	1.018	1.048	1.002
EW-2	0.770	0.935	0.934	0.920	0.992	0.997	0.968	0.979	0.951	0.940	0.967	0.925
EW-3	0.766	0.930	0.929	0.915	0.974	0.979	0.951	0.962	0.937	0.926	0.953	0.911
EW-4	0.747	0.907	0.907	0.892	0.970	0.975	0.947	0.958	0.933	0.923	0.950	0.908
PA-1	0.656	0.797	0.796	0.784	0.927	0.932	0.906	0.915	0.991	0.980	1.009	0.965
PA-2	0.694	0.842	0.842	0.828	0.944	0.949	0.922	0.932	0.913	0.903	0.929	0.888
PA-3	0.734	0.891	0.890	0.876	0.938	0.943	0.916	0.926	0.897	0.887	0.913	0.873
PA-4	0.678	0.823	0.822	0.810	0.938	0.942	0.915	0.925	0.962	0.951	0.979	0.936
PA-NP	0.656	0.797	0.796	0.784	0.858	0.863	0.838	0.847	0.902	0.892	0.918	0.878
P-Lin-QR	0.685	0.832	0.831	0.818	0.963	0.968	0.940	0.951	0.975	0.964	0.992	0.948
Ad-QR	0.538	0.653	0.652	0.642	<b>0.817</b>	<b>0.821</b>	<b>0.797</b>	<b>0.806</b>	0.835	0.826	0.850	0.813
H-1	0.539	0.654	0.653	0.643	0.829	0.832	0.809	0.818	0.818	0.808	0.832	0.800
H-2	0.517	0.627	0.627	0.617	0.829	0.834	0.810	0.819	<b>0.815</b>	<b>0.805</b>	<b>0.829</b>	<b>0.793</b>
H-3	0.529	0.643	0.642	0.632	0.830	0.834	0.810	0.819	0.821	0.812	0.836	0.799
H-4	<b>0.508</b>	<b>0.616</b>	<b>0.616</b>	<b>0.606</b>	0.826	0.830	0.806	0.815	0.820	0.810	0.834	0.800

**Notes:** (i) EW = Equal Weighting, BM = Benchmark Model, PA = Penalized Averaging; PA-NP = Penalized Averaging with NP marginal quantiles, P-Lin-QR = Penalized linear model with quantile regression and Ad-QR = Additive quantile regression; (ii) For the purpose of this table, the entries are measured up to three decimal figures such that more than one ratio of thick loss functions may have the lowest value.

## 5 Conclusions

We proposed a hybrid approach to combine relevant information from parametric and semiparametric quantile forecasts in high dimensions. It rests on the idea of combining, or averaging misspecified candidate parametric and nonparametric quantile forecasts, which in this study are the marginal quantiles, to form an approximation of the true combined conditional quantile. One advantage of the approach is that the weights of the combined quantile forecasts are unknown as opposed to methods with known, fixed, combinations of predictors. Indeed, we find overwhelming empirical evidence that in terms of quantile forecasting performance our hybrid approach works well in identifying relevant predictors from a large set of macroeconomic predictors and technical indicators and, more importantly, results in improved combined out-of-sample forecasts over (semi-)parametric models/methods. We also provided some insights as to where the gain of the hybrid method comes from. Further, we have seen that very different forecast results emerge in the right tail of the conditional quantile distribution. These results are mainly based on the prediction performance of the set of 14 macroeconomic variables, while there is hardly any evidence supporting the quantile predictive performance of the 14 technical indicators.

These empirical results raise questions for further research. In particular, it may be of interest to gain further insight in the forecasting performance of the hybrid approach via a controlled Monte Carlo experiment. One may also explore the possibility of making economic gains from utilizing information in the tails of the return distribution via application of the hybrid approach. Extending the paper to multi-step ahead conditional quantile forecasting within the current hybrid framework is another topic worth investigating.

## Acknowledgements

We would like to thank two anonymous referees for their helpful comments and suggestions.

## Notes

- 1 Konzen and Ziegelmann (2016) study some alternative forms of LASSO-type penalties for predictor selection and conditional mean forecasting.
- 2 Available at: <http://users.stat.umn.edu/~zouxx019/ftplib/code/fhdqr/>. The computer code is running on R-3.4.4.
- 3 These authors approximate the conditional mean via composite quantiles which is a topic well studied in statistics; see, e. g. Kong and Xia (2014). This naive approach of obtaining predictions is by construction restricted to linear specifications as quantile regressions do not capture nonlinearities in the DGP. For this reason, and many differences in forecast design, a comparison between our forecasting results in Section 4.1 and the empirical results obtained by these authors is not possible.
- 4 Note, we consider a non-nested pairwise forecast comparison. By allocating only a relatively small fraction of our observations to the forecast period, we are able to side-step issues related to the effects of in-sample parameter estimation uncertainty on the distribution of the DM test statistic; see West (2006) for a discussion.
- 5 Results are available upon request.

## References

- Audrino, F., and L. Camponovo. 2017. "Oracle Properties, Bias Correction, and Bootstrap Inference for Adaptive Lasso for Time Series M-estimators." *Journal of Time Series Analysis* 39: 111–28. <http://doi.org/10.1111/jtsa.12270>.
- Bai, J., and S. Ng. 2008. "Forecasting Economic Time Series Using Targeted Predictors." *Journal of Econometrics* 146: 304–17. <http://doi.org/10.1016/j.jeconom.2008.08.010>.
- Bayer, S. 2018. "Combining Value-at-Risk Forecasts Using Penalized Quantile Regressions." *Econometrics and Statistics* 8: 56–77. <http://doi.org/10.1016/j.ecosta.2017.08.001>.
- Bonaccolto, G., M. Caporin, and S. Paterlini. 2018. "Asset Allocation Strategies Based on Penalized Quantile Regression." *Computational Management Science* 15: 1–32. <http://doi.org/10.1007/s10287-017-0288-3>.
- Cenesizoglu, T., and A. Timmermann. 2012. "Do Return Prediction Models add Economic Value?" *Journal of Banking & Finance* 36: 2974–87. <http://doi.org/10.1016/j.jbankfin.2012.06.008>.
- Chen, J., D. Li, O. Linton, and Z. Lu. 2018. "Semiparametric Ultra-high Dimensional Model Averaging of Nonlinear Dynamic Time Series." *Journal of the American Statistical Association* 113: 919–32. <http://doi.org/10.1080/01621459.2017.1302339>.
- De Gooijer, J. G., and D. Zerom. 2019. "Semiparametric Quantile Averaging in the Presence of High-Dimensional Predictors." *International Journal of Forecasting* 35: 891–909. <http://doi.org/10.1016/j.ijforecast.2018.10.009>.
- Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani. 2004. "Least Angle Regression." *The Annals of Statistics* 32: 407–99. <http://doi.org/10.1214/009053604000000067>.

- Exterkate, P., D. Van Dijk, C. Heij, and P. J. F. Groenen. 2011. "Forecasting the Yield Curve in a Data-Rich Environment Using the Factor-Augmented Nelson–Siegel Model." *Journal of Forecasting* 32: 193–215. <http://doi.org/10.1002/for.1258>.
- Garcia, M. G. P., M. C. Medeiros, G. F. R. Vasconcelos. 2017. "Real-Time Inflation Forecasting with High-Dimensional Models: The Case of Brazil." *International Journal of Forecasting* 33: 697–93. <http://doi.org/10.1016/j.ijforecast.2017.02.002>.
- Giacomini, R., and I. Komunjer. 2005. "Evaluation and Combination of Conditional Quantile Forecasts." *Journal of Business and Economic Statistics* 23: 416–31. <http://doi.org/10.1198/073500105000000018>.
- Giovannelli, A. 2012. "Nonlinear Forecasting Using Large Datasets: Evidence on US and Euro Area Economies?" CEIS Tor Vergata, Research Paper Series, Vol. 10, Issue 13, No. 255.
- Gu, Y., J. Fan, L. Kong, S. Ma, and H. Zou. 2018. "ADMM for High-dimensional Sparse Penalized Quantile Regression." *Technometrics* 60: 319–31. <http://doi.org/10.1080/00401706.2017.1345703>.
- Han, Y., and R. S. Tsay. 2019. "High-Dimensional Linear Regression for Dependent Data with Applications to Now casting." *Statistica Sinica* (forthcoming). <http://doi.org/10.5705/ss.202018.0044>.
- Jiang, B., G. Athanasopoulos, R. J. Hyndman, A. Panagiotelis, and F. Vahid. 2018. "Macroeconomic Forecasting for Australia Using a Large Number of Predictors." Working paper 17/02, Department of Econometrics and Business Statistics, Monash Business School.
- Kong, E., and Y. Xia. 2014. "An Adaptive Composite Quantile Approach to Dimension Reduction." *The Annals of Statistics* 42: 1657–88. <http://doi.org/10.1214/14-aos1242>.
- Konzen, E., and F. A. Ziegelmann. 2016. "LASSO-type Penalties for Covariate Selection and Forecasting in Time Series." *Journal of Forecasting* 35: 592–612. <http://doi.org/10.1002/for.2403>.
- Lee, E. R., H. Noh, and B. U. Park. 2014. "Model Selection via Bayesian Information Criterion for Quantile Regression Models." *Journal of the American Statistical Association* 109: 216–29. <http://doi.org/10.1080/01621459.2013.836975>.
- Lima, L. R., and F. Meng. 2017. "Out-of-Sample Return Predictability: A Quantile Combination Approach." *Journal of Applied Econometrics* 32: 877–95. <http://doi.org/10.1002/jae.2549>.
- Ma, S., R. Fildes, and T. Huang. 2016. "Demand Forecasting with High Dimensional Data: The Case of SKU Retail Sales Forecasting with Intra- and Inter-Category Promotional Information." *European Journal of Operational Research* 249: 245–57. <http://doi.org/10.1016/j.ejor.2015.08.029>.
- Medeiros, M. C., and E. F. Mendes. 2014. "Penalized Estimation of Semi-Parametric Additive Time-Series Models." In *Essays in Nonlinear Time Series Econometrics*, edited by N. Haldrup, M. Meitz, and P. Saikkonen. Oxford Scholarship Online. pp. 215–237. <http://doi.org/10.1093/acprof:oso/9780199679959.003.0009>.
- Medeiros, M. C., and E. F. Mendes. 2016. " $\ell_1$ -Regularization of High-Dimensional Time-Series Models with Non-Gaussian and Heteroskedastic Errors." *Journal of Econometrics* 191: 255–71. <http://doi.org/10.1016/j.jeconom.2015.10.011>.
- Meligkotsidou, L., E. Panopoulou, I. Vrontos, and S. Vrontos. 2014. "A Quantile Regression Approach to Equity Premium Prediction." *Journal of Forecasting* 33: 558–76. <http://doi.org/10.1002/for.2312>.
- Meligkotsidou, L., E. Panopoulou, I. Vrontos, and S. Vrontos. 2019. "Quantile Forecast Combinations in Realised Volatility Prediction." *Journal of the Operational Research Society* (forthcoming). <http://doi.org/10.1080/01605682.2018.1489354>.
- Neely, C. J., D. E. Rapach, J. Tu, and G. Zhou. 2014. "Forecasting the Equity Risk Premium: The Role of Technical Indicators." *Management Science* 60: 1772–91. <http://doi.org/10.1287/mnsc.2013.1838>.
- Pedersen, T. Q. 2015. "Predictable Return Distributions." *Journal of Forecasting* 34: 114–32. <http://doi.org/10.1002/for.2323>.
- Rapach, D. E., and G. Zhou. 2013. "Forecasting Stock Returns." In *Handbook of Economic Forecasting*, edited by G. Elliott, and A. Timmermann, Vol. 2, 328–383, Part A. North-Holland: Elsevier. <http://doi.org/10.1016/B978-0-444-53683-9.00006-2>.
- Sherwood, B., and L. Wang. 2016. "Partially Linear Additive Quantile Regression in Ultra-High Dimension." *The Annals of Statistics* 44: 288–317. <http://doi.org/10.1214/15-aos1367>.
- Similä, T., and J. Tikka. 2006. "Common Subset Selection of Inputs in Multiresponse Regression." In *Proceedings of the IEEE International Joint Conference on Neural Networks*, Vancouver, Canada, 1908–1915. <http://doi.org/10.1109/ijcnn.2006.246933>.
- Wang, H., and C. Leng. 2008. "A Note on Adaptive Group Lasso." *Computational Statistics and Data Analysis* 52: 5277–86. <http://doi.org/10.1016/j.csda.2008.05.006>.
- Welch, I., and A. Goyal. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21: 1455–508. <http://doi.org/10.1016/1093/rfs/hhmo14>.
- West, K. D. 2006. "Forecast Evaluation." In *Handbook of Economic Forecasting*, edited by G. Elliott, C. W. J. Granger, and A. Timmermann, Vol. 1, 99–134. North-Holland: Elsevier. [http://doi.org/10.1016/s1574-0706\(05\)01003-7](http://doi.org/10.1016/s1574-0706(05)01003-7).