A Quantile Regression Approach to Equity Premium Prediction

Loukia Meligkotsidou[a], Ekaterini Panopoulou[b],
Ioannis D.Vrontos[c]*, Spyridon D. Vrontos[b]

[a]Department of Mathematics, University of Athens, Athens, Greece
[b]Department of Statistics and Insurance Science, University of Piraeus, Piraeus, Greece
[c]Department of Statistics, Athens University of Economics and Business, Athens, Greece

Abstract

We propose a quantile regression approach to equity premium forecasting. Robust point forecasts are generated from a set of quantile forecasts, using both fixed and time-varying weighting schemes, thus exploiting the entire distributional information associated with each predictor. Further gains are achieved by incorporating the forecast combination methodology in our quantile regression setting. Our approach using a time-varying weighting scheme delivers statistically and economically significant out-of-sample forecasts relative to both the historical average benchmark and the combined predictive mean regression modeling approach.

JEL classification: G11; G12; C22; C53

Keywords: Equity premium; Forecast combination; Predictive quantile regression; Robust point forecasts; Time-varying weights.

*Corresponding author. Department of Statistics, Athens University of Economics and Business, Patission 76, GR-10434 Athens, Greece. Tel.: +30-210-8203927. Email: `vrontos@aueb.gr`

1. Introduction

Equity premium predictability has attracted the attention of both academics and practitioners in finance. Results are mixed, since different techniques, variables and time periods are employed in the related research.[1] The list of predictors is quite exhaustive and typically contains valuation ratios, various interest rates and spreads, distress indicators, inflation rates along with other macroeconomic variables, indicators of corporate activity, etc. The early contributions to equity premium predictability mainly focused on the in-sample predictive ability of the potential predictors and the development of proper econometric techniques for valid inference.[2] Lately, interest has turned to the out-of-sample performance of the candidate variables. Goyal and Welch (2008) show that their long list of predictors can not deliver consistently superior out-of-sample performance. The authors employ a variety of predictive regression models ranging from single variable ones to their 'kitchen sink' model that contains all the predictors simultaneously. Campbell and Thompson (2008) show that when imposing simple restrictions, suggested by economic theory, on predictive regressions' coefficients, the out-of-sample performance improves and market timing strategies can deliver profits to investors (see also Ferreira and Santa-Clara, 2011). More recently, Rapach et al. (2010) consider another approach for improving equity premium forecasts based on forecast combinations. The authors find that combinations of individual single variable predictive regression forecasts, which help reducing model uncertainty/parameter instability, significantly beat the historical average forecast. Finally, Ludvigson and Ng (2007) and Neely et al. (2011) adopt a diffusion index approach, which can conveniently track the key movements in a large set of predictors, and find evidence of improved equity premium forecasting ability.

It still remains an open question whether there is clear evidence of equity premium predictability, with the majority of studies conducted within a linear regression framework. However, recent contributions to the literature have pointed out that the relationship between returns and predictors is not linear and several approaches have been proposed

---

[1]Following the related literature, equity premium is proxied by excess returns.
[2]Rapach and Zhou (2012) offer a detailed review on the issue of equity return predictability.

to capture this non linearity. Markov-switching models are among the most popular models for forecasting stock returns (Guidolin and Timmermann, 2009; Henkel et al., 2011). Other well-known non-linear specifications include threshold models and neural nets (Franses and van Dijk, 2000; Terasvirta, 2006; White, 2006; Guidolin et al., 2009). Non or semi-parametric modeling represents another approach for approximating general functional forms for the relationship between expected returns and predictors (Chen and Hong, 2010; Ait-Sahalia and Brandt, 2001).

In this paper, we address the issue of non linearity between excess returns and predictive variables by considering predictive quantile regression models for equity premium forecasting. We argue that due to non-linearity and non-normality patterns, a linear approach might not be adequate for exploring the ability of various predictors to forecast the entire distribution of returns. Looking at just the conditional mean of the return series may 'hide' interesting characteristics as it can lead us to conclude that a predictor has poor predictive performance, while it is actually valuable for predicting the lower or/and the upper quantiles of returns. For example, the most popular variables in the returns prediction literature, namely the dividend-price ratio and the term spread, may capture different aspects of economic conditions. Furthermore, not only fluctuations of the business cycle induce a time-varying nature on mean predictive relationships, but also across quantiles, since there is no compeling theoretical reason for the slope coefficients to be constant across quantiles. To the extent that candidate predictor variables contain significant information for some parts of the return distribution, but not for the whole, a methodology that properly integrates this information would lead to additional benefits.

Since the seminal paper of Koenker and Bassett (1978), quantile regression models have attracted a vast amount of attention. Both theoretical and empirical research has been conducted in the area of quantile regression, including model extensions, new inferential procedures and numerous empirical applications; see, for example, Buchinsky (1994, 1995) and Yu et al. (2003) among others.[3] The paper more closely related to

---

[3] Applications in the field of finance include Bassett and Chen (2001), Engle and Manganelli (2004), Meligkotsidou et al. (2009), Chuang et al. (2009) and Baur et al. (2012).

the present paper is that of Cenesizoglu and Timmermann (2008) who employ a quantile regression approach to capture the predictive ability of a list of state variables for the distribution of stock returns. The authors find quantile-varying predictability both in-sample and out-of-sample which can be exploited in an asset allocation framework. In a follow-up paper, Cenesizoglu and Timmermann (2012) point out that return prediction models that allow for a time-varying return distribution lead to better estimates of the tails of the returns' distribution and suffer less from unanticipated outliers. Similar conclusions are reached by Pedersen (2010) who employs both univariate and multivariate quantile regressions to jointly model the distribution of stocks and bonds.

In this paper, we construct robust and accurate point forecasts of the equity premium from the quantile forecasts produced by a set of predictive quantile regressions, using both a fixed and a time-varying weighting scheme. We design two novel forecasting approaches which utilize distributional information, as well as information from a set of available predictors. The first approach initially constructs robust point forecasts from a set of quantile predictions all of which are based on the same predictive variable. Next, it combines the robust forecasts obtained from different predictors using several existing combination methods in order to produce a final point forecast. The second approach initially combines all the predictions of the same quantile obtained from different single predictor model specifications, in order to produce combined quantile forecasts. This is done via a number of forecast combination methods, developed in the present paper, which are appropriate for combining quantile forecasts. Then, our proposed approach constructs robust point forecasts by synthesizing the combined quantile predictions. For comparison purposes, we employ the updated Goyal and Welch (2008) dataset along with the standard linear regression predictive framework, as well as existing methods of combining individual forecasts from single predictor linear models. All different forecasts are evaluated against the benchmark constant equity premium using both statistical and economic evaluation criteria.

To anticipate our key results, we find considerable heterogeneity among the candidate variables, as far as their ability to predict the return distribution is concerned.

More importantly, no single predictor proves successful in forecasting the entire return distribution. Overall, superior predictive performance, both in statistical and economic evaluation terms, is achieved under the quantile regression approach as follows. First, a set of quantiles of the conditional distribution of returns are optimally predicted by combining information from different predictors using a quantile forecast combination method. Next, robust point forecasts of the equity premium are produced by synthesizing the quantile predictions using time-varying weighting schemes.

The remainder of the paper is organized as follows. Section 2 describes the econometric models considered in this study, including predictive mean and quantile regression models. Section 3 outlines our proposed methodology for robust estimation of the central location of the distribution of returns. Section 4 discusses various methods of combining forecasts from different model specifications in the context of standard mean regression and quantile regression. Our dataset and the framework for forecast evaluation is presented in Section 5, while our empirical results are reported in Section 6. Section 7 outlines the economic evaluation framework and presents the associated findings. Section 8 summarizes and concludes.

2. Predictive Regressions

In this section we present the predictive regression models we employ to forecast the equity premium, denoted by r_t, using a set of N predictive variables.

2.1. Quantile Regression Models

First we consider all possible predictive mean regression models with a single predictor of the form

$$r_{t+1} = \alpha_i + \beta_i x_{it} + \varepsilon_{t+1}, \ i = 1, \ldots, N, \tag{1}$$

where r_{t+1} is the observed excess return on a stock market index in excess of the risk-free interest rate at time $t+1$, x_{it} are the N observed predictors at time t, and the error terms ε_{t+1} are assumed to be independent with mean zero and variance σ^2. Equation (1) is the standard equity premium prediction model (see, for example, Rapach et al. 2010).

The above regression specification can only predict the mean and not the entire dis-

tribution of returns in the event that the joint distribution of r_{t+1} and x_{it} is not bivariate Gaussian and, therefore, their relationship is not linear. Following the literature on the non-linear relationship between returns and predictors (Guidolin and Timmermann, 2009; Guidolin et al., 2009; Chen and Hong, 2010; Henkel et al., 2011) we adopt a more sophisticated approach to equity premium forecasting by employing predictive quantile regression models (Koenker and Bassett, 1978; Buchinsky, 1998; Yu et al., 2003). Quantile regression estimators are more efficient and more robust than mean regression estimators in cases that non linearities and deviations from normality exist.

We consider single predictor quantile regression models of the form

$$ r_{t+1} = \alpha_i^{(\tau)} + \beta_i^{(\tau)} x_{it} + \varepsilon_{t+1}, \ i = 1, \dots, N, \tag{2} $$

where $\tau \in (0, 1)$ and the errors ε_{t+1} are assumed independent from an error distribution $g_\tau(\varepsilon)$ with the τth quantile equal to 0, i.e. $\int_{-\infty}^{0} g_\tau(\varepsilon)d\varepsilon = \tau$. Model (2) suggests that the τth quantile of r_{t+1} given x_{it} is $Q_\tau(r_{t+1}|x_{it}) = \alpha_i^{(\tau)} + \beta_i^{(\tau)} x_{it}$, where the intercept and the regression coefficients depend on τ. The $\beta_i^{(\tau)}$'s are likely to vary across τ's, revealing a larger amount of information about returns than the predictive mean regression model.

2.2. Inference on Predictive Regression Models

The predictive mean regression model can be estimated using the Ordinary Least Squares (OLS) method. Least squares estimation is based on the fact that the expectation of a random variable r with distribution function F arises as the point estimate of r corresponding to the quadratic loss function $\rho(u) = u^2$, i.e. it arises as the value of $\bar{r}$ which minimizes the expected loss

$$ E\rho(r - \bar{r}) = \int \rho(r - \bar{r})dF(r). $$

Therefore, the OLS estimators $\hat{\alpha}_i, \hat{\beta}_i$ of the parameters in the predictive mean regression models in (1) can be estimated by minimizing the sample estimate of the quadratic

5

expected loss, $\sum_{t=0}^{T-1} \left(r_{t+1} - \alpha_i - \beta_i x_{it} \right)^2$, with respect to α_i, β_i.[4] Then, the point forecast of the equity premium at time $t+1$, based on the ith model specification, is obtained as $\hat{r}_{i,t+1} = \hat{\alpha}_i + \hat{\beta}_i x_{it}$.

Similarly to the expectation of the random variable r, its τth quantile arises as the solution to a decision theoretic problem; that of obtaining the point estimate of r corresponding to the asymmetric linear loss function, usually referred to as the check function,

$$\rho_\tau(u) = u \left(\tau - I(u < 0) \right) = \frac{1}{2} \left[|u| + (2\tau - 1)u \right]. \tag{3}$$

More in detail, minimization of the expected loss

$$E\rho_\tau(r - \bar{r}^{(\tau)}) = \int \rho_\tau(r - \bar{r}^{(\tau)}) dF(r),$$

with respect to $\bar{r}^{(\tau)}$ leads to the τth quantile. In the symmetric case of the absolute loss function ($\tau = 1/2$) we obtain the median. Estimators of the parameters of the linear quantile regression models in (2), $\hat{\alpha}_i^{(\tau)}, \hat{\beta}_i^{(\tau)}$, can be obtained by minimizing the sum $\sum_{t=0}^{T-1} \rho_\tau \left(r_{t+1} - \alpha_i^{(\tau)} - \beta_i^{(\tau)} x_{it} \right)$, where the check function $\rho_\tau(u)$ has been given in (3). Then, the forecast of the τth quantile of the distribution of the equity premium at time $t+1$, based on the ith model specification, is obtained as $\hat{r}_{i,t+1}(\tau) = \hat{\alpha}_i^{(\tau)} + \hat{\beta}_i^{(\tau)} x_{it}$.

2.3. Forecasting Approaches based on Quantile Regression

In order to produce robust and accurate point forecasts of the equity premium based on quantile forecasts, we utilize two different sources of information. We consider distributional information, regarding how the relationship between the equity premium and a given predictive variable varies across the conditional quantiles of returns, as well as predictor information, regarding the different model specifications that can be used for forecasting. To take account of both sources of information we propose the following two novel forecasting approaches.

The first approach is designed to initially construct robust point forecasts of the equity

---

[4]The sample size T denotes any estimation sample employed in our recursive forecasting experiment. Details on the forecasting design are given in Section 4.

premium from a set of quantile forecasts based on a single predictor x_{it}. This is done by employing several fixed and time-varying weighting schemes (see Section 3). Then, these robust point forecasts are combined in order to reduce uncertainty risk associated with a single predictive variable. This is done by using various methods of combining predictor information (see Subsection 4.1), based on the Mean Squared Forecast Error (MSFE). We refer to this approach as the Combining Robust Forecasts approach (CRF).

The second approach is designed to construct point forecasts as follows. First, the quantile forecasts obtained from different single predictor model specifications are combined according to several combination schemes based on the asymmetric linear loss function (see Subsection 4.2). Then, robust point forecasts are obtained by synthesizing the above quantile forecasts, that is exploiting distributional information, based on the weighting schemes of Section 3. We refer to this second approach as the Quantile Forecasts Synthesis approach (QFS).

3. Robust Point Forecasts based on Regression Quantiles

In this section we consider the problem of constructing robust point forecasts of the equity premium based on quantile regression as an alternative to the standard approach which produces forecasts based on the predictive mean regression model. Robust point estimates of the central location of a distribution can be constructed as weighted averages of a set of quantile estimators employing mainly fixed weighting schemes. Relaxing the assumption of a constant weighting scheme seems to be a natural extension. A number of factors, such as changes in regulatory conditions, market sentiment, monetary policies, institutional framework or even changes in macroeconomic interrelations (Campbell and Cochrane, 1999; Menzly et al., 2004; Dangl and Halling, 2012) can motivate the employment of time-varying schemes in the generation of robust point forecasts.

3.1. Point Forecasts based on a Fixed Weighting Scheme

Robust point forecasts of the equity premium can be constructed as weighted averages of a set of quantile forecasts. First, we employ standard estimators with fixed, prespecified

weights of the form

$$\hat{r}_{i,t+1} = \sum_{\tau \in S} p_\tau \hat{r}_{i,t+1}(\tau), \quad \sum_{\tau \in S} p_\tau = 1,$$

where S denotes the set of quantiles considered. Here the weights represent probabilities attached to different quantile forecasts, suggesting how likely to predict the return at the next period each regression quantile is.

We consider Tukey's (1977) trimean and the Gastwirth (1966) three-quantile estimator given, respectively, by the following formulae

FW1: $\widehat{r}_{i,t+1} = 0.25\hat{r}_{i,t+1}(0.25) + 0.50\hat{r}_{i,t+1}(0.50) + 0.25\hat{r}_{i,t+1}(0.75)$

FW2: $\widehat{r}_{i,t+1} = 0.30\hat{r}_{i,t+1}(1/3) + 0.40\hat{r}_{i,t+1}(0.50) + 0.30\hat{r}_{i,t+1}(2/3).$

Furthermore, we use the alternative five-quantile estimator, suggested by Judge et al. (1988), which attaches more weight on extreme positive and negative events as follows

$$
\begin{aligned}
\text{FW3: } \widehat{r}_{i,t+1} \;=\;& 0.05\hat{r}_{i,t+1}(0.10) + 0.25\hat{r}_{i,t+1}(0.25) + 0.40\hat{r}_{i,t+1}(0.50) \\
+\;& 0.25\hat{r}_{i,t+1}(0.75) + 0.05\hat{r}_{i,t+1}(0.90).
\end{aligned}
$$

Finally, we consider a fourth estimator which combines information from a larger set of quantiles, i.e.

FW4: $\widehat{r}_{i,t+1} = 0.05\hat{r}_{i,t+1}(0.50) + 0.05 \sum_{\tau \in S} \hat{r}_{i,t+1}(\tau),$ where $S = \{0.05, 0.10, ..., 0.95\}.$

All the above fixed weighting schemes (FW1-FW4) provide estimators of the central location of the return distribution at time $t+1$. A subset of the above specifications has been employed by Taylor (2007) and Ma and Pohlman (2008) among others.

3.2. Point Forecasts based on a Time-varying Weighting Scheme

Time-varying weighting schemes are derived by some optimization procedure aiming at producing an empirical model that allows for economic changes over time and which is capable of determining the 'right' parameter values in time to help investors (Spiegel,

2008). The variable of interest, $r_{i,t+1}$, is predicted using an optimal linear combination $\mathbf{p}_t = [p_{\tau,t}]_{\tau \in S}$ of the quantile forecasts $\hat{r}_{i,t+1}(\tau)$ given by

$$\hat{r}_{i,t+1} = \sum_{\tau \in S} p_{\tau,t}\hat{r}_{i,t+1}(\tau), \quad \sum_{\tau \in S} p_{\tau,t} = 1.$$

The weights, $\mathbf{p}_t$, are estimated recursively using a holdout out-of-sample period continuously updated by one observation at each step. Optimal estimates of the weights are obtained by minimizing the mean squared forecast errors, $E_t(r_{t+1} - \hat{r}_{i,t+1})^2$, under an appropriate set of constraints. Our optimization procedure is the analogue of the constrained Granger and Ramanathan (1984) method for quantile regression forecasts (see also Timmermann, 2006; Hansen, 2008; Hsiao and Wan, 2012). Specifically, we employ constrained least squares using the quantile forecasts as regressors in lieu of a standard set of predictors. The time-varying weights on the quantile forecasts bear an interesting relationship to the portfolio weight constraints in finance. In this sense we constrain the weights to be non-negative, sum to one and not to exceed certain lower and upper bounds in order to reduce the weights' volatility and stabilize forecasts. In our empirical application, we employ three time-varying specifications which may be viewed as the time-varying counterparts of our FW1-FW3 schemes.[5] More specifically, FW1 with time-varying coefficients becomes

$$\text{TVW1: } \widehat{r}_{i,t+1} = p_{0.25,t}\hat{r}_{i,t+1}(0.25) + p_{0.50,t}\hat{r}_{i,t+1}(0.50) + p_{0.75,t}\hat{r}_{i,t+1}(0.75),$$

where $p_{\tau,t}, \tau \in S = \{0.25, 0.50, 0.75\}$ are estimated by the optimization procedure

$$\mathbf{p}_t = \arg\min_{\mathbf{p}_t} E[r_{t+1} - (p_{0.25,t}\hat{r}_{i,t+1}(0.25) + p_{0.50,t}\hat{r}_{i,t+1}(0.50) + p_{0.75,t}\hat{r}_{i,t+1}(0.75))]^2$$

$$s.t. \ p_{0.25,t} + p_{0.50,t} + p_{0.75,t} = 1, 0.20 \le p_{0.25,t} \le 0.40,$$

$$0.40 \le p_{0.50,t} \le 0.60, 0.20 \le p_{0.75,t} \le 0.40.$$

---

[5]Since our methodology requires a holdout out-of-sample period during which the optimal linear combination $\mathbf{p}_t$ is estimated, a fourth specification based on FW4 is not employed due to the increased parameter space.

Similarly, the FW2 scheme with time-varying coefficients becomes

$$\text{TVW2: } \widehat{r}_{i,t+1} = p_{1/3,t}\hat{r}_{i,t+1}(1/3) + p_{0.5,t}\hat{r}_{i,t+1}(0.50) + p_{2/3,t}\hat{r}_{i,t+1}(2/3),$$

where $p_{\tau,t}, \tau \in S = \{1/3, 0.50, 2/3\}$ are estimated by the following optimization procedure

$$\mathbf{p}_t = \arg\min_{\mathbf{p}_t} E[r_{t+1} - (p_{1/3,t}\hat{r}_{i,t+1}(1/3) + p_{0.5,t}\hat{r}_{i,t+1}(0.50) + p_{2/3,t}\hat{r}_{i,t+1}(2/3))]^2 \quad (4)$$

$$s.t. \ p_{1/3,t} + p_{0.50,t} + p_{2/3,t} = 1, 0.15 \leq p_{1/3,t} \leq 0.45,$$

$$0.30 \leq p_{0.5,t} \leq 0.50, 0.15 \leq p_{2/3,t} \leq 0.45.$$

Finally, the FW3 scheme with time-varying coefficients becomes

$$\begin{aligned}\text{TVW3: } \widehat{r}_{i,t+1} &= p_{0.10,t}\hat{r}_{i,t+1}(0.10) + p_{0.25,t}\hat{r}_{i,t+1}(0.25) + p_{0.5,t}\hat{r}_{i,t+1}(0.50) \\ &+ p_{0.75,t}\hat{r}_{i,t+1}(0.75) + p_{0.90,t}\hat{r}_{i,t+1}(0.90),\end{aligned}$$

where $p_{\tau,t}, \tau \in S = \{0.10, 0.25, 0.50, 0.75, 0.90\}$ are estimated by the following optimization procedure

$$\mathbf{p}_t = \arg\min_{\mathbf{p}_t} E[r_{t+1} - (p_{0.10,t}\hat{r}_{i,t+1}(0.10) + p_{0.25,t}\hat{r}_{i,t+1}(0.25) +$$

$$+p_{0.5,t}\hat{r}_{i,t+1}(0.5) + p_{0.75,t}\hat{r}_{i,t+1}(0.75) + p_{0.90,t}\hat{r}_{i,t+1}(0.90))]^2$$

$$s.t. \ p_{0.10,t} + p_{0.25,t} + p_{0.50,t} + p_{0.75,t} + p_{0.90,t} = 1$$

$$0.00 \leq p_{0.10,t} \leq 0.10, 0.15 \leq p_{0.25,t} \leq 0.35,$$

$$0.40 \leq p_{0.50,t} \leq 0.60, 0.15 \leq p_{0.75,t} \leq 0.35, 0.00 \leq, p_{0.90,t} \leq 0.10.$$

4. Forecast Combination

Since Bates and Granger's (1969) seminal contribution, it has been known that combining individual models' forecasts can reduce uncertainty risk associated with a single predictive model and display superior predictive ability (see also Hendry and Clements, 2004). In the context of equity premium predictability, Rapach et al. (2010) show that

combination forecasts of individual predictive models can consistently beat the benchmark. The design of our forecast experiment is identical to the one employed by Goyal and Welch (2008) and Rapach et al. (2010). We generate out-of-sample forecasts of the equity premium using a recursive (expanding) window. More specifically, we divide the total sample of T observations into an in-sample portion of the first K observations and an out-of-sample portion of $P = T - K$ observations used for forecasting. The estimation window is continuously updated following a recursive scheme, by adding one observation to the estimation sample at each step. As such, the coefficients in any predictive model employed are re-estimated after each step of the recursion. Proceeding in this way through the end of the out-of-sample period, we generate a series of P out-of-sample forecasts for the equity premium $\{\hat{r}_{i,t+1}\}_{t=K}^{T-1}$. This experiment simulates the situation of a forecaster in real time, since she employs data as soon as they become available.

Following Rapach et al. (2010), we consider various combining methods, ranging from simple averaging schemes to more advanced ones, based on both the single predictor model specifications of Section 2 and the robust point forecasts of Section 3. In order to produce combined quantile forecasts we need to develop appropriate combining methods based on the asymmetric linear loss function (Equation 3). In the following subsections we outline the combining methods employed in this study. In Subsection 4.1 we present the existing combining methods that are used for producing combined forecasts based on single predictor mean forecasts or robust point forecasts, while in Subsection 4.2 we introduce the respective combining methods that are appropriate for producing combined quantile forecasts.

4.1. Combination Methods for Central Location Forecasting

The combination forecasts of r_{t+1}, denoted by $\hat{r}_{t+1}^{(C)}$, are weighted averages of the N single predictor individual forecasts, $\hat{r}_{i,t+1}$, $i = 1, \ldots, N$, of the form $\hat{r}_{t+1}^{(C)} = \sum\limits_{i=1}^{N} w_{i,t}^{(C)} \hat{r}_{i,t+1}$, where $w_{i,t}^{(C)}, i = 1, ..., N$, are the *a priori* combining weights at time t. Some of the combining methods described below require a holdout out-of-sample period during which the combining weights are estimated. The first P_0 out-of-sample observations are employed as the initial holdout period. In this respect, we compute combination forecasts over the

post-holdout out-of-sample period, leaving us with a total of $T - (K + P_0) = P - P_0$ forecasts available for evaluation.

The simplest combining scheme is the one that attaches equal weights to all individual models, i.e. $w_{i,t}^{(C)} = 1/N$, for $i = 1, ..., N$, called the Mean combining scheme. The next schemes we employ are the Trimmed Mean and Median ones. The Trimmed Mean combination scheme sets $w_{i,t}^{(C)} = 0$ for the smallest and largest forecasts and $w_{i,t}^{(C)} = 1/(N-2)$ for the remaining ones, while the Median combination scheme employs the median of the $\{\hat{r}_{i,t+1}\}_{i=1}^{N}$ forecasts.

The second class of combining methods we consider, proposed by Stock and Watson (2004), suggests forming weights based on the historical performance of the individual models over the holdout out-of-sample period. Specifically, their Discount Mean Squared Forecast Error (DMSFE) combining method suggests forming weights as follows

$$w_{i,t}^{(C)} = m_{i,t}^{-1} / \sum_{j=1}^{N} m_{j,t}^{-1}, \ m_{i,t} = \sum_{s=K}^{t-1} \psi^{t-1-s} (r_{s+1} - \hat{r}_{i,s+1})^2,$$

where ψ is a discount factor which attaches more weight on the recent forecasting accuracy of the individual models in the cases where $\psi \in (0, 1)$. The values of ψ we consider are 1.0 and 0.9. When ψ equals one, there is no discounting and the combination scheme coincides with the optimal combination forecast of Bates and Granger (1969) in the case of uncorrelated forecasts.

The third class of combining methods, namely the Cluster combining method, was introduced by Aiolfi and Timmermann (2006). In order to create the Cluster combining forecasts, we form L clusters of forecasts of equal size based on the MSFE performance. Each combination forecast is the average of the individual model forecasts in the best performing cluster. This procedure begins over the initial holdout out-of-sample period and goes through the end of the available out-of-sample period using a rolling window. In our analysis, we consider $L = 2, 3$.

Next, the Principal Components combining method of Chan et al. (1999) and Stock and Watson (2004) is considered. In this case, a combination forecast is based on the

fitted n principal components of the uncentered second moment matrix of the individual model forecasts, $\widehat{F}_{1,s+1}, ..., \widehat{F}_{n,s+1}$ for $s = K, ..., t-1$. The OLS estimates of $\varphi_1, ..., \varphi_n$ of the following regression

$$r_{s+1} = \varphi_1 \widehat{F}_{1,s+1} + ... + \varphi_n \widehat{F}_{n,s+1} + \nu_{s+1}$$

can be thought of as the individual combining weights of the principal components. In order to select the number n of principal components we employ the IC_{p3} information criterion developed by Bai and Ng (2002) and set the maximum number of factors to 5.

4.2. Combination Methods for Quantile Forecasting

The DMSFE, Cluster and Principal Components combining methods have been designed in the framework of standard linear regression, in order to construct forecasts that exploit the entire set of predictive variables. The combining weights, $w_{i,t}^{(C)}$, are computed based on the MSFE, that is on a quadratic loss function that measures how close to the realized excess returns the individual forecasts are. These methods are appropriate within the framework of the CRF approach since, according to this approach, several robust point forecasts are first obtained from different single predictor quantile regressions and then these point forecasts are combined in order to exploit information from the available set of predictors. However, these combining schemes are not appropriate for combining predictor information within the QFS approach since variable information is now combined in the context of forecasting several quantiles of returns rather than producing point forecasts. In this case, the MSFE is no longer suitable for measuring the performance of the produced forecasts and has to be replaced by a metric based on the asymmetric linear loss function.[6]

Below we describe how we modify the existing combining methods in order to produce quantile forecasts that exploit variable information. These modified combining methods are new to the forecast combination literature and can be considered as a further contribution of the present study. The combined quantile forecasts, $\hat{r}_{t+1}^{(C)}(\tau)$, are weighted

---

[6]We are grateful to an anonymous referee for pointing this out.

averages of the form $\hat{r}_{t+1}^{(C)}(\tau) = \sum\limits_{i=1}^{N} w_{i,t}^{(C)} \hat{r}_{i,t+1}(\tau)$, where the combining weights, $w_{i,t}^{(C)}$, have to be computed based on the check function (3).

First, we introduce the Discount Asymmetric Loss Forecast Error (DALFE) combining method which suggests forming weights as follows

$$w_{i,t}^{(C)} = m_{i,t}^{-1} / \sum_{j=1}^{N} m_{j,t}^{-1}, \ m_{i,t} = \sum_{s=K}^{t-1} \psi^{t-1-s} \rho_\tau (r_{s+1} - \widehat{r}_{i,s+1}(\tau)),$$

where $\psi \in (0,1)$ is a discount factor. The combining weights are computed based on the historical performance of the individual quantile regression models over the holdout out-of-sample period and ψ is set equal to 0.9 and 1.

We also modify the Cluster combining method by forming L clusters of forecasts based on their performance as measured by the asymmetric loss forecast error. The Asymmetric Loss Cluster (AL Cluster) combination forecast is the average of the individual quantile forecasts in the best performing cluster which contains the forecasts with the lower expected asymmetric loss values. We consider forming $L = 2, 3$ clusters.

Next, we introduce the Asymmetric Loss Principal Components method (AL Principal Components) under which the combination of forecasts is based on the fitted, n, principal components of the uncentered second moment matrix of the individual quantile forecasts, $\widehat{F}_{1,s+1}^{(\tau)}, ..., \widehat{F}_{n,s+1}^{(\tau)}$, where the combination weights are computed by minimizing the sum

$$\sum_{s=K}^{t-1} \rho_\tau (r_{s+1} - \varphi_1 \widehat{F}_{1,s+1}^{(\tau)} - ... - \varphi_n \widehat{F}_{n,s+1}^{(\tau)}).$$

The IC_{p3} information criterion is used to select the number n of principal components.

Finally, we put forward two combining methods under which optimal quantile forecasts, $\hat{r}_{t+1}^{(C)}(\tau)$, are obtained by minimizing an objective function based on the asymmetric linear loss. More in detail, we first consider the following optimization scheme, which is

an analogue of the lasso quantile regression

$$\mathbf{w}_t = \arg\min_{\mathbf{w}_t} \sum_t \rho_\tau \left(r_{t+1} - \sum_{i=1}^N w_i \hat{r}_{i,t+1}(\tau) \right)$$

$$s.t. \ \sum_{i=1}^N w_i = 1, \ \sum_{i=1}^N |w_i| \leq \delta_1,$$

where the parameter δ_1 is used as a control for the amount of shrinkage. We refer to this combination quantile forecast as Asymmetric Loss Lasso (AL Lasso). We also consider the Asymmetric Loss Ridge (AL Ridge) optimization scheme which is an analogue of the ridge quantile regression

$$\mathbf{w}_t = \arg\min_{\mathbf{w}_t} \sum_t \rho_\tau \left(r_{t+1} - \sum_{i=1}^N w_i \hat{r}_{i,t+1}(\tau) \right)$$

$$s.t. \ \sum_{i=1}^N w_i = 1, \ \sum_{i=1}^N w_i^2 \leq \delta_2,$$

where the parameter δ_2 is used as a control for the amount of shrinkage. In our study, the parameters δ_1, δ_2 are set equal to 1.4 and 0.4, respectively.[7]

5. Data and forecast evaluation

The data we employ are from Goyal and Welch (2008) who provide a detailed description of transformations and datasources.[8] The equity premium is calculated as the difference of the continuously compounded S&P500 returns, including dividends, and the Treasury Bill rate. As already mentioned, following the line of work of Goyal and Welch (2008), Rapach et al. (2010) and Ferreira and Santa-Clara (2011), out-of-sample forecasts of the equity premium are generated by continuously updating the estimation window, i.e. following a recursive (expanding) window. Our forecasting experiment is conducted on a quarterly basis and data span 1947:1 to 2010:4. Our out-of-sample forecast evaluation

---

[7]The above two optimization schemes can be written equivalently using the L_1 norm for the lasso quantile regression and the L_2 norm for the ridge quantile regression in the objective function. More details on lasso regression can be found in Tibsirani (1996), on lasso quantile regression in Wu and Liu (2009) and on ridge regression can be found in Hastie et al. (2009).

[8]The data are available at http://www.hec.unil.ch/agoyal/. We thank Prof. Goyal for making them available to us.

period corresponds to the 'long' one analyzed by Goyal and Welch (2008) and Rapach et al. (2010) covering the period 1965:1-2010:4.[9] The 15 economic variables employed in our analysis are related to stock-market characteristics, interest rates and broad macroeconomic indicators. With respect to stock market characteristics, we employ the following variables.

- Dividend–price ratio (log), D/P: Difference between the log of dividends paid on the S&P 500 index and the log of stock prices (S&P 500 index), where dividends are measured using a one-year moving sum.

- Dividend yield (log), D/Y : Difference between the log of dividends and the log of lagged stock prices.

- Earnings–price ratio (log), E/P: Difference between the log of earnings on the S&P 500 index and the log of stock prices, where earnings are measured using a one-year moving sum.

- Dividend–payout ratio (log), D/E: Difference between the log of dividends and the log of earnings.

- Stock variance, SVAR: Sum of squared daily returns on the S&P 500 index.

- Book-to-market ratio, B/M: Ratio of book value to market value for the Dow Jones Industrial Average.

- Net equity expansion, NTIS: Ratio of twelve-month moving sums of net issues by NYSE-listed stocks to total end-of-year market capitalization of NYSE stocks.

Turning to interest-rate related variables, we employ six variables ranging from short-term government rates to long-term government and corporate bond yields and returns along with their spreads as follows.

---

[9]Please note that the out-of-sample period refers to the period used to evaluate the out-of-sample forecasts. We use the ten years 1955:1 to 1964:4 (40 quarters) before the start of the out-of-sample evaluation period as the initial holdout out-of-sample period, required for both constructing our time-varying robust forecasts and for several forecast combination schemes.

- Treasury bill rate, TBL: Interest rate on a three-month Treasury bill (secondary market).

- Long-term yield, LTY: Long-term government bond yield.

- Long-term return, LTR: Return on long-term government bonds.

- Term spread, TMS: Difference between the long-term yield and the Treasury bill rate.

- Default yield spread, DFY: Difference between BAA- and AAA-rated corporate bond yields.

- Default return spread, DFR: Difference between long-term corporate bond and long-term government bond returns.

To capture the overall macroeconomic environment, we employ the inflation rate and the investment-to-capital ratio defined as follows.

- Inflation, INFL: Calculated from the CPI (all urban consumers).

- Investment-to-capital ratio, I/K: Ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the entire economy (Cochrane, 1991).

The natural benchmark forecasting model is the historical mean or prevailing mean (PM) model, according to which the forecast of the equity premium coincides with the estimate, $\widehat{\alpha}_i$, in the linear regression model (1) when no predictor is included. As a measure of forecast accuracy we employ the ratio $\frac{MSFE_i}{MSFE_{PM}}$, where $MSFE_i$ is the Mean Square Forecast Error associated with each of our competing models and specifications and $MSFE_{PM}$ is the respective value for the PM model, both computed over the out-of-sample period. Values lower than 1 are associated with superior forecasting ability of our proposed model/specification.

In order to compare the information content in our proposed models/specifications relevant to the benchmark PM model, we use encompassing tests. The notion of forecast

encompassing was developed by Granger and Newbold (1973) and Chong and Hendry (1986) through the formation of composite forecasts as weighted averages of the forecasts of two competing models.[10] Specifically, consider forming a composite forecast, $\hat{r}_{c,t+1}$, as a convex combination of model A forecasts, $\hat{r}_{A,t+1}$, and the ones of model B, $\hat{r}_{B,t+1}$, in an optimal way so that $\hat{r}_{c,t+1} = \lambda_A \hat{r}_{A,t+1} + \lambda_B \hat{r}_{B,t+1}, \lambda_A + \lambda_B = 1$. If the optimal weight attached to model A forecasts is zero ($\lambda_A = 0$), then model B forecasts encompass model A forecasts in the sense that model B contains a significantly larger amount of information than that already contained in model A. Harvey et al. (1998) developed the encompassing test, denoted as $ENC - T$, based on the approach of Diebold and Mariano (1995) to test the null hypothesis that $\lambda_A = 0$, against the alternative hypothesis that $\lambda_A > 0$. Let $u_{A,t+1} = r_{t+1} - \hat{r}_{A,t+1}$, $u_{B,t+1} = r_{t+1} - \hat{r}_{B,t+1}$ denote the forecast errors of the competing models A and B, respectively and define $d_{t+1} = (u_{B,t+1} - u_{A,t+1})u_{B,t+1}$. The $ENC - T$ statistic is given by

$$ENC - T = \sqrt{(P - P_0)}\frac{\overline{d}}{\sqrt{\widehat{Var}(d)}},$$

where $\overline{d}$ is the sample mean, $\widehat{Var}(d)$ is the sample-variance of $\{d_{s+1}\}_{s=K+P_0}^{T-1}$ and $P - P_0$ is the length of the out-of-sample evaluation window.[11] The $ENC - T$ statistic is asymptotically distributed as a standard normal variate under the null hypothesis. To improve finite sample performance, Harvey et al. (1998) recommend employing the Student's t distribution with $P - P_0 - 1$ degrees of freedom. To render a model as superior in forecasting ability, one also needs to test whether model A forecasts encompass model B forecasts ($\lambda_B = 0$) by employing the $ENC - T$ statistic based on $d_{t+1} = (u_{A,t+1} - u_{B,t+1})u_{A,t+1}$. When both null hypotheses are rejected, then the competing models contain discrete information about the future and an optimal convex ($\lambda_A, \lambda_B \in (0,1)$) combination forecast can be formed. In the event that none of the hypotheses of interest is rejected, both models contain similar information and the competing models

---

[10]See also Clements and Hendry (1998).

[11]For forecast horizons greater than one, an estimate of the long-run variance should be employed.

are equivalent in terms of forecasting ability. When one of the null hypotheses is rejected, then the respective model forecasts dominate the forecasts of the competing model.

6. Empirical Results

6.1. A motivating illustration

Before presenting our empirical results, we provide an illustration on the sources of potential benefits of our proposed methodology. The aim of this exercise is to assess the predictive ability of the individual predictor variables, $x_{i,t}$, to forecast the τth quantile. To this end, we generate forecasts employing a single predictor at a time, $\widehat{r}_{i,t+1}^{(\tau)} = \widehat{\alpha}_i^{(\tau)} + \widehat{\beta}_i^{(\tau)} x_{it}$, $i = 1,...,N$, and calculate the expected asymmetric loss, $\sum_t \rho_\tau \left(r_{t+1} - \widehat{r}_{i,t+1}^{(\tau)} \right)$, associated with each model specification. Then we calculate the expected loss associated with the quantile forecasts, $\widehat{r}_{t+1}^{(\tau)} = \widehat{\alpha}^{(\tau)}$, obtained from the Prevailing Quantile (PQ) model, i.e. the model that contains only a constant. This prevailing quantile model serves as a benchmark in the same fashion as the historical average (prevailing mean) serves as a benchmark in typical predictive mean regressions. Table 1, Panel A illustrates our findings with highlighted (in grey) cells suggesting superior predictive ability, i.e. lower out-of-sample values of the expected asymmetric loss. Overall, we observe considerable heterogeneity among the candidate variables as far as their ability to predict the return distribution is concerned. For example, the D/P and D/Y variables display predictive ability for the 10th and 15th quantile, but mainly for the central and some right-tail quantiles of the distribution of returns, i.e. from the 45th to the 80th quantiles. On the other hand, DFR, INFL and I/K are valuable predictors for the left-tail and central quantiles of the return distribution. Finally, D/E, SVAR and DFY help predicting some upper quantiles and TBL the 30th to 45th quantiles. It is apparent that no single predictor proves successful in predicting the entire distribution of returns.

[TABLE 1 AROUND HERE]

We now examine whether combining the information of different predictors in order to predict each quantile enhances our ability to forecast the quantiles of the return distribution. For this purpose, we employ a variety of simple combination methods, such as the

Mean, Median and Trimmed Mean combination strategies, as well as the proposed new combination methods that are based on the asymmetric loss function and are appropriate for combining quantile forecasts, i.e. the DALFE, AL Cluster, AL Principal Components, AL Lasso and AL Ridge combination methods described in Subsection 4.2. The potential predictive ability of the combining schemes considered is outlined in Table 1, Panel B. Our results suggest that these combination methods contain substantial information for the future return distribution. The Mean, Trimmed Mean, DALFE and AL Ridge methods cover the full range of the distribution, while the Median and the AL Cluster methods are successful in all parts of the distribution, with the exception of the 90th and the 5th quantile, respectively. The AL Principal Components combining method does not outperform the PQ model in terms of predictive ability except for the 30th and 40th quantile. Finally, the AL Lasso method is superior to the PQ model at forecasting the left part of the return distribution and some right-tail quantiles.

6.2. Out-of-sample performance of predictive regressions

In this subsection, we conduct an out-of-sample forecasting exercise with the aim to present and discuss the results of the proposed forecasting approaches, i.e. the CRF and QFS approaches. For reasons of comparison we also present results of the combined mean regression forecasts.

6.2.1. Performance of Mean Regression Forecasts

Table 2 reports the out-of-sample performance of both the single predictor mean regression forecasts and the combined forecasts obtained using the combination methods of Subsection 4.1. We refer to this forecasting approach as the Combining Mean Forecasts approach (CMF). In particular, Table 2 presents the MSFE ratios of each of the individual predictive regression models relative to the historical average benchmark model for the out-of-sample period 1965:1-2010:4. Values lower than 1 indicate superior forecasting performance of the predictive models with respect to the historical average forecast. We observe that only four out of the 15 individual predictors, namely D/P, D/Y, DFR and I/K, have lower than one MSFE ratios, indicating superior predictive ability.

Next, we test the statistical significance of the out-of-sample performance of the forecasts of the various competing models with respect to the PM forecasts using the encompassing test. Forecast encompassing provides a means for comparing the information content in different forecasts. In Table 2, λ_A denotes the parameter associated with the test which examines whether the PM forecasts encompass the forecasts taken from the individual predictive models, while λ_B denotes the parameter associated with the test that examines whether the individual predictive model forecasts encompass the PM ones. We observe that for the D/P, D/Y and I/K predictors we reject the null hypothesis that PM forecasts encompass the respective individual predictor's forecasts, and as such these variables contain useful forecasting information beyond the information already contained in the PM model. On the other hand, we may notice that the D/P, D/Y and I/K forecasts encompass the PM forecasts, therefore the PM forecasts do not contain any useful information. Based on these results the D/P, D/Y and I/K forecasts dominate the forecasts of the PM model, while the PM forecasts dominate the D/E, B/M, NTIS, LTY, LTR and DFY forecasts. Turning to the CMF approach, our findings suggest that all the combining schemes (except for the Principal Components method) produce lower than unity MSFE ratios, indicating that the combining methods' forecasts have superior predictive ability. The encompassing test confirms the statistical significance of our forecasts obtained from the combining methods (with the exception of the Principal Components method). Overall the results of Table 2 are in line with the findings of Rapach et al. (2010) who found that the D/P, D/Y and I/K predictors have significant forecasting ability, and that the combination methods outperform the individual predictive regression models.

[TABLE 2 AROUND HERE]

6.2.2. Evaluation of the Combining Robust Forecasts approach

We turn our attention to the out-of-sample performance of the robust point forecasts obtained by using fixed weighting (FW) and time-varying weighting (TVW) schemes, based on single predictor quantile regression models. Furthermore, we evaluate the predictive ability of the forecasts obtained by the CRF approach. Table 3 reports the MSFE

ratios and the results of the encompassing test for the single predictor robust point forecasts and for the CRF approach forecasts, relative to the historical average (PM) benchmark model. Based on Panel A of Table 3, which reports the performance of the robust point forecasts formed by the fixed weighting schemes on individual predictive quantile models, we observe that only three predictors, namely D/P, D/Y and I/K, have MSFE ratios below unity for all weighting schemes FW1-FW4, while DFR seems to have some forecasting ability under the FW4 scheme. Turning to the encompassing test results, the null hypothesis that the PM forecasts encompass the D/P, D/Y and I/K robust point forecasts is rejected for all weighting schemes (except for I/K under FW3 scheme), indicating that the robust forecasts contain useful information. On the other hand, the robust point forecasts encompass the PM forecasts, that is the PM forecasts do not contain any useful information. These results are similar in spirit with those of the individual predictive mean regression model, and indicate superior forecasting ability of these three predictors over the historical average using different fixed weighting schemes. Note, however, that the MSFE ratios of the robust point forecasts are lower than those of the individual mean regression of D/P and D/Y for all weighting schemes, and of I/K for FW1, FW2 and FW4, indicating some improvement over the mean regression approach.

[TABLE 3 AROUND HERE]

Panel B of Table 3 presents the performance of the CRF approach, where the robust point forecasts based on the fixed weighting schemes are combined to reduce uncertainty risk associated with a single predictor. Almost all of the combining methods, except for the Principal Components and in some cases the Cluster 3 method, provide MSFE ratios below unity and, hence, their forecasts dominate the PM forecast. A comparison of the different combination techniques suggests that the DMSFE methods rank first followed by the mean combination method, since they generally provide lower MSFE ratios. Among the four fixed weighting schemes, the FW4 scheme produces, in most of the cases, lower MSFE ratios indicating improved predictive performance, probably due to the fact that it utilizes distributional information obtained from a finer grid of return quantiles.

The results presented in Table 3 (Panel C) concern the out-of-sample performance of the robust point forecasts with time-varying weights (TVW1-TVW3) based on single predictor quantile regression models. The values of the MSFE ratio indicate that four predictors, namely D/P, D/Y, DFR and I/K, display superior forecasting ability irrespective of the weighting scheme employed (TVW1-TVW3). Moreover, the INFL (under TVW1 and TVW3 scheme) and the NTIS (under TVW3 scheme) predictors have lower than unity MSFE ratios. However, based on both encompassing tests we observe that only the D/P, D/Y and DFR individual robust point forecasts dominate the PM forecasts for all weighting schemes, and therefore contain valuable information. The time-varying weighting approach suggests that an additional predictor, the DFR, may contain valuable out-of-sample information, compared to the fixed weighting approach and the predictive mean regression model. The improved out-of-sample performance of the robust point forecasts using time-varying weights over the predictive mean regression model is also apparent since most of the MSFE ratios for the individual predictors are lower than the corresponding MSFE ratios of the predictive mean models (Table 2).

The most striking result can be drawn from panel D of Table 3, which reports the results of the CRF approach under the time-varying weighting schemes (TVW1-TVW3). The MSFE ratios in this case are all below unity, ranging from 0.976 for the Median combination method using TVW2 to 0.963 for the Mean combination method using TVW3.[12] Moreover, all the MSFE ratios for the CRF approach based on time-varying weights are lower than the corresponding MSFE ratios of both the CMF (Table 2) and the fixed weighting CRF approach (Table 3, Panel B). The encompassing tests suggest that the CRF forecasts dominate the PM model's forecasts.

6.2.3. Evaluation of the Quantile Forecasts Synthesis approach

Next, we present and discuss the results of the second forecasting approach introduced in our study, i.e. the QFS approach. The purpose of this analysis is to evaluate the forecasts obtained by first utilizing the predictor information to produce combined quantile

---

[12]Since the time-varying weighting schemes require a holdout out-of-sample period, they can only be used together with combining methods that do not require a holdout period.

forecasts from different individual predictive models and then synthesizing this distributional information through robust forecasting weighting schemes. This procedure aims to provide optimal forecasts of each part of the return distribution, by appropriately combining individual quantile forecasts, and then to construct robust point forecasts of the equity premium synthesizing the combined quantile forecasts. Table 4 presents the out-of-sample performance of the QFS robust point forecasts obtained under fixed (FW1-FW4) and time-varying weighting schemes (TVW1-TVW3). The results of Panel A suggest that the QFS forecasts, with the exception of the AL Principal Components combination method, based on the fixed weighting schemes provide MSFE ratios below unity, indicating superior performance relative to the historical average benchmark. A comparison of the different combining methods reveals that the AL Ridge method ranks first followed by the DALFE, the Mean and the AL Cluster 2, since they generally provide lower MSFE ratios. It is interesting to observe that more promising results in favor of the proposed QFS approach arise from the use of time-varying weighting schemes (TVW1-TVW3). More importantly, the QFS-TVW approach generates MSFE ratios below unity and in many cases the lowest ones among the different forecasting approaches considered in our analysis (see Table 4. Panel B). The results of Table 4 suggest that the best out-of-sample performance is obtained by applying the QFS approach using time-varying weights to the quantile forecasts obtained by the Mean combination method.

[TABLE 4 AROUND HERE]

Up to now our analysis has shown that the proposed forecasting methods based on quantile regression (i.e. the CRF and QFS approaches) using time-varying weighting provide superior forecasts, i.e. lower MSFE ratios than the standard CMF approach. Below, we present and discuss a more formal comparison of the CMF approach with the two alternative approaches proposed in this paper, via a series of encompassing tests. To this end we compare all pairs of forecasts obtained by the CMF, the CRF (under the TVW1, TVW2 and TVW3 weighting schemes) and the QFS (under the TVW1, TVW2 and TVW3 weighting schemes) approaches using pairwise encompassing tests.

The results of these tests are shown in Table 5. The comparison of the CMF with the CRF shows that the CMF forecasts are dominated by the CRF forecasts under the first weighting scheme, if either the Mean or the Trimmed Mean combining method is used, and under the third weighting scheme, if the Median combining method is used. On the other hand, the CMF forecasts are dominated by the QFS forecasts under both the first and the second weighting schemes for all combining methods considered. Moreover, it is interesting to note that the CMF forecasts do not prove more accurate than any of the proposed forecasting approaches based on quantile regression. Finally, the comparison of the two robust forecasting approaches with each other shows that the QFS forecasts are superior to the CRF forecasts for the Mean and Median combination methods under the third time-varying weighting scheme.

[TABLE 5 AROUND HERE]

6.2.4. Multiple Encompassing Tests and an Amalgam Forecast

We now consider an amalgamation of the approaches considered so far, namely the CMF, CRF and QFS approaches.[13] First, we employ the multiple forecast encompassing tests of Harvey and Newbold (2000) in order to check whether potential benefits can arise from combining the three approaches. Next we form equally weighted composite forecasts based on the three approaches and test their statistical significance.

Harvey and Newbold (2000) extend the pairwise encompassing tests developed by Harvey et al. (1998) to compare three or more forecasts. Consider forming a composite forecast, $\hat{r}_{c,t+1}$, as a combination of the forecasts of the predictive mean regressions (CMF), the combination of robust point forecasts (CRF) and the robust point forecasts deduced from the combined quantile (QFS) in an optimal way so that $\hat{r}_{c,t+1}$ = $\lambda_{CMF}\hat{r}_{CMF,t+1} + \lambda_{CRF}\hat{r}_{CRF,t+1} + \lambda_{QFS}\hat{r}_{QFS,t+1}$, where $\lambda_{CMF} + \lambda_{CRF} + \lambda_{QFS} = 1$. If $\lambda_{CMF} = 1$, and $\lambda_{CRF} = \lambda_{QFS} = 0$, the CMF forecasts encompass the CRF and QFS ones, as the CRF and QFS forecasts do not contain information useful for forecasting the equity premium other than that already employed in the linear model. In a similar

---

[13]The term 'amalgamation' is employed by Rapach and Strauss (2012) when considering combining three different econometric approaches to forecast US state employment growth.

manner, we can test whether the CRF model encompasses QFS and CMF and whether the QFS model encompasses the CMF and the CRF model. Harvey and Newbold develop two test statistics, namely the F-test statistic and the MS^* statistic, to test the null hypothesis of multiple forecast encompassing.[14] The authors show that the F-test exhibits significant size distortions in small and moderate samples with non-normal errors, while the MS^* test exhibits good size and power properties in moderately large samples. In order to gain a more thorough understanding on the relationship between the rival models, we must employ each one of the models as the reference model and conduct the test as many times as the models considered. Failure to reject the null hypothesis does not necessarily imply that the reference model is strictly dominant to the competing forecasts. Rather, the forecasts may be highly correlated, in which case a combination of nearly identical or similar forecasts cannot improve upon any individual forecast. On the other hand, rejection of the null hypothesis in the encompassing test suggests that the reference model's forecasts can be improved by combining them with the forecasts of the rival model.

Given the abundance of the models we considered so far, we only report multiple forecast encompassing tests for the models employed in the pairwise encompassing tests (Table 5). Table 6 (columns 2-4) reports the respective MS^* test statistics. Overall, non-rejections of the null dominate our findings pointing to similarities in forecasting ability of our competing models and possibly non gains from considering forming composite forecasts. More in detail, the only case that the MS^* test rejects the null of multiple encompassing is when the Mean combining scheme is employed and the robust point forecasts are generated by the TVW3 scheme. Forming composite forecasts of the three approaches considered can help us gain more insight on the nature of our forecasts. Given that our experiment should be real time, we do not estimate the weights in forming our composite forecasts, rather we attach a weight of 1/3 on each of our competing models. Table 6 (column 5) reports the MSFE ratio of our amalgam forecasts along with the related encompassing tests (columns 6-7). Overall, the MSFE ratio ranges from 0.964

---

[14]To save space, we do not report the explicit formulae of the tests.

for the amalgam forecast formed on the basis of Mean combining schemes and TVW1 robust forecasts to 0.983 for the forecasts formed based on the Median combining schemes and TVW3. More importantly, all amalgam forecasts dominate the benchmark forecasts of the historical average as indicated by the encompassing tests. However, no amalgam forecast proves more accurate than the QFS and/or CRF forecasts lending support to the superiority of our proposed approaches. Even in the case that the MS^* test pointed to benefits in combining, namely the Mean combining scheme with the robust point forecasts generated by TVW3, the amalgam forecast is superior to the CMF forecasts but not superior to the CRF or QFS ones.[15]

[TABLE 6 AROUND HERE]

In what follows, we evaluate the economic significance of our proposed specifications against the benchmark historical average.

7. Economic evaluation

As Campbell and Thompson (2008) and Rapach et al. (2010) suggest, even small statistical gains in predictability can give an economically meaningful degree of return predictability that could result in increased portfolio returns for a mean-variance investor that maximizes expected utility. Within this stylized asset allocation framework, this utility-based approach, initiated by West et al. (1993), has been extensively employed in the literature as a measure for ranking the performance of competing models in a way that captures the trade-off between risk and return (Fleming et al., 2001; Marquering and Verbeek, 2004; Della Corte et al., 2009, 2010; Wachter and Warusawitharana, 2009).

7.1. The framework for measuring economic value

Consider a risk-averse investor who constructs a dynamically rebalanced portfolio consisting of the risk-free asset and one risky asset. Her portfolio choice problem is how to allocate wealth between the safe (risk-free Treasury Bill) and the risky asset (stock market), while the only source of risk stems from the uncertainty over the future path of

---

[15]This is probably due to the equal weighting scheme in forming the amalgam forecast. Having estimated weights might have proven beneficial in this case.

the stock market. Since only one risky asset is involved, this approach could be thought of as a standard exercise of market timing in the stock market. In a mean-variance framework, the solution to the maximization problem of the investor yields the following weight (w_t) on the risky asset

$$w_t = \frac{E_t(r_{t+1})}{\gamma Var_t(r_{t+1})} = \frac{\widehat{r}_{i,t+1}}{\gamma Var_t(r_{t+1})},$$

where E_t and Var_t denote the conditional expectation and variance operators, r_{t+1} is the equity premium and γ is the Relative Risk Aversion (RRA) coefficient that controls the investor's appetite for risk (Campbell and Viceira, 2002; Campbell and Thompson, 2008; Rapach et al., 2010). The conditional expectation $E_t(r_{t+1})$ of each model is given by the optimal forecast from the specific model, $\widehat{r}_{i,t+1}$, and the variance, $Var_t(r_{t+1})$ is calculated using a ten-year rolling window of quarterly returns.[16] In this way, the optimal weights vary with the degree the conditional mean varies, i.e. the forecast each model/ specification gives.[17] In this setting the optimally constructed portfolio gross return over the out-of-sample period, $R_{p,t+1}$, is equal to

$$R_{p,t+1} = w_t \cdot r_{t+1} + R_{f,t},$$

where $R_{f,t} = 1 + r_{f,t}$ denotes the gross return on the risk-free asset from period t to $t+1$.[18] Over the forecast evaluation period the investor with initial wealth of W_o realizes an average utility of

$$\overline{U} = \frac{W_o}{(P - P_0)} \left[\sum_{t=0}^{P-P_0-1} (R_{p,t+1}) - \frac{\gamma}{2} \sum_{t=0}^{P-P_0-1} (R_{p,t+1} - \overline{R}_p)^2 \right], \tag{5}$$

---

[16] Under the assumption of constant return volatility, the variance of stock returns can be estimated using the sample variance computed from a recursive window of historical returns. To allow for a time-varying variance one may use a rolling window of historical returns or a rolling window of a GARCH type model (Campbell and Thomson, 2008; Rapach et al., 2010; Rapach and Zhou, 2012).

[17] Alternatively, one could make use of information about the entire distribution provided by the quantile regression predictive models.

[18] We constrain the portfolio weight on the risky asset to lie between 0% and 150% each month, i.e. $0 \le w_t \le 1.5$.

where $R_{p,t+1}$ is the gross return on her portfolio at time $t+1$. At any point in time, the investor prefers the predictive model that yields the highest average realized utility.[19]

Given that a better model requires less wealth to attain a given level of $\overline{U}$ than an alternative model, a risk-averse investor will be willing to pay to have access to this superior model which would be subject to management fees as opposed to the simple PM model. In the event that the superior model is one of our proposed i specifications, the investor would pay a performance fee to switch from the portfolio constructed based on the historical average to the i specification. This performance fee, denoted by Φ, is the fraction of the wealth which, when subtracted from the i proposed portfolio returns, equates the average utilities of the competing models. In our set-up, the performance fee is calculated as the difference between the realized utilities as follows

$$\Phi = \Delta\overline{U} = \overline{U^i} - \overline{U^{PM}}, \tag{6}$$

where $\overline{U^i}$ is the average realized utility over the out-of-sample period of any of our competing models/ specifications (i) and $\overline{U^{PM}}$ is the respective value for the prevailing mean (PM) model. If our proposed model does not contain any economic value, the performance fee is negative ($\Phi \leq 0$), while positive values of the performance fee suggest superior predictive ability against the PM benchmark.

As a complement to the performance fee measure, we also employ the manipulation-proof performance measure proposed by Goetzmann et al. (2007). This measure can be interpreted as a portfolio's premium return after adjusting for risk and it remedies potential caveats associated with the popular Sharpe ratio, such as the effect of non normality (Jondeau and Rockinger, 2006), the underestimation of the performance of dynamic strategies (Marquering and Verbeek, 2004; Han, 2006) and the choice of utility function (Della Corte et al., 2012). This measure is defined as

$$M(R_p) = \frac{1}{1-\gamma} \ln \left\{ \frac{1}{(P-P_0)} \sum_{t=0}^{P-P_0-1} \left(\frac{R_{p,t+1}}{R_{f,t}} \right)^{1-\gamma} \right\}.$$

---

[19]We standardize the investor problem by assuming $W_o = 1$.

Intuitively the portfolio has the same score (ranking) as does a risk free asset whose continuously compounded return exceeds the interest rate by $M(R_p)$. The difference, Θ, between the $M(R_p)s$ of competing models, calculated as follows, is employed to assess the most valuable model

$$\Theta = M(R_p)^i - M(R_p)^{PM}. \tag{7}$$

Both Φ and Θ are reported in annualized basis points.

7.2. Empirical evidence on the economic value of predictive regressions

We assume that the investor dynamically rebalances her portfolio (updates the weights) quarterly over the out-of-sample period employing the respective recursive forecasts for all the models/specifications under consideration. The forecasts' precision normally increases as more information (data) becomes available. Similarly to Section 6, the out-of-sample period of evaluation is 1965:1-2010:4 and the benchmark strategy against which we evaluate our forecasts is the PM model. For every model/specification we calculate the performance fee associated with each strategy calculated from Equation (6) and the manipulation-proof performance measure (Equation 7). Following Campbell and Thompson (2008) and Rapach et al. (2010) we set RRA (γ) equal to 3.

We begin our analysis with the economic evaluation of the CMF approach. Table 7 (Panel A) reports the respective figures. Our results suggest that, irrespective of the method employed, an investor enjoys utility gains ranging from 145 (Median) to 321 (DMSFE(0.9)). Quite interestingly, while the Principal Components method is not statistically superior to the benchmark model, its employment can generate profits to an investor amounting to 236 bps. The combining methods with the highest ability to time the market are the DMSFE ones followed by the Mean and Trimmed Mean. The ranking of our combining methods remains almost unchanged when we employ the manipulation-proof performance measure (Θ).

[TABLE 7 AROUND HERE]

Next, we turn our attention to the economic performance of the fixed weighting CRF approach. Our results, reported in Table 7 (Panel B), may be summarized as follows.

First, the economic performance of our CRF approach is nearly as good as the performance of the CMF approach. Overall, similarly to the mean forecasts, our results suggest that an investor that employs the CRF approach will always generate positive abnormal returns. The lowest utility gains are observed in the Median method ranging from 18 bps to 108 bps while the highest utility gains are attained by the Principal Components, Cluster 2 and DMSFE(0.9). Second, a comparison of the four weighting schemes reveals that FW4 which aggregates information of quantiles over a finer grid, provides the investor with more utility gains and the highest performance fee of 275 bps is achieved when the investor employs FW4 with DMSFE(0.9). Finally, comparison of the different methods using Θ is consistent with the results obtained from the performance fees.

Turning to our time-varying CRF approach (Table 7, Panel C), we observe high positive risk-adjusted abnormal returns employing either Φ or Θ. More in detail, our findings indicate that the time-varying CRF approach outperforms both the CMF and the fixed weighting CRF approaches. The utility gains range from 159 bps (TVW2 Median) to 395 bps (TVW3 Mean), which is the highest value of the utility gain attained so far. Similarly, the highest value of the manipulation-proof performance measure is found for the same model (TVW3 Mean).

Finally, Table 8 addresses the issue of the economic evaluation of both the fixed weighted and time-varying weighted QFS approach along with the amalgam forecasts. The overall picture that emerges confirms the robustness of our proposed methodology. More in detail, the performance fee that an investor would be willing to pay to utilize our proposed models (with the exception of the FW-Median combining method) ranges from 158 bps for the Trimmed Mean QFS-FW1 to 425 bps for QFS-TVW1 and the Mean combining method. When considering the fixed weighting schemes (Table 8, Panel A), the best performance is achieved by AL Lasso (QFS-FW1), AL Cluster 3 (QFS-FW2), AL Ridge (QFS-FW3 and QFS-FW4). Turning to the risk-adjusted abnormal return Θ, our findings remain broadly unchanged with a few exceptions such as the negative Θ of -1.8 bps for the Median QFS-FW2. More importantly, when an investor employs any of the QFS-TVW models, she can enjoy benefits ranging from 243 bps to 425 bps (Table

8, Panel B). Superior performance is achieved by the QFS-TVW1 scheme, irrespective of the combining method employed. Comparing our QFS to CRF time-varying approaches, we find that when employing either TVW1 or TVW2, QFS is to be preferred, while the opposite holds for TVW3. It is interesting to note that the amalgam forecasts attain a satisfactory performance ranging from a fee of 236 bps to 375 bps, with the exception of TVW3 and the Median combining scheme (Table 8, Panel C). Similar findings pertain when considering the manipulation proof performance measure, which is in the range of 98 bps to 458 bps.

[TABLE 8 AROUND HERE]

8. Conclusions

In this study we investigate whether there is evidence of out-of-sample predictive ability of various economic variables for the equity premium. We develop a novel forecasting approach based on predictive quantile regression models which produces robust and accurate point forecasts of the equity premium from a set of quantile forecasts, by using fixed and time-varying weighting schemes. To take into account the findings of recent academic studies which suggest that forecast combinations improve the out-of-sample equity premium prediction, we propose utilizing a variety of combination methods based on quantile forecasts. Thus, in our analysis, the crucial issue under consideration is to examine whether the framework that considers two different sources of information, i.e. distributional information and predictor information, is able to deliver more accurate out-of-sample forecasts of the equity premium.

The usefulness of the proposed forecasting approach stems from the highly complex and dynamic nature of equity returns. Our approach is able to capture the non-linear relationship of returns with predictors and to identify potential differences in the ability of predictors to forecast various quantiles of returns. For example, our analysis suggests that predictors with superior predictive ability for lower or/and upper quantiles of returns exist. Thus, the quantile regression approach is able to uncover interesting distributional information and, from an economic perspective, to incorporate meaningful

32

business conditions information.

Our study contributes to the growing empirical literature on equity premium predictability by putting forward the following novel forecasting approach. First, we recommend combining individual forecasts from different single predictor quantile regressions, thus incorporating information from various economic variables in order to produce accurate quantile predictions. Then, we propose constructing robust point forecasts of the equity premium by adopting a time-varying weighting scheme which combines these quantile predictions, thus incorporating information from the entire distribution of returns. Our findings suggest that the predictive ability of the proposed approach has substantial statistical and economic value over the standard predictive modeling approaches.

References

Ait-Sahalia, Y and M.W. Brandt. "Variable selection for portfolio choice." Journal of Finance, 56 (2001), 1297–1351.

Aiolfi, M., and A. Timmermann. "Persistence in Forecasting Performance and Conditional Combination Strategies." Journal of Econometrics, 135 (2006), 31–53.

Bai, J., and S. Ng . "Determining the Number of Factors in Approximate Factor Models." Econometrica, 70 (2002), 191–221.

Bassett, W. G., and H-L. Chen. "Portfolio style: Return-based attribution using quantile regression." Empirical Economics, 26 (2001), 293-305.

Bates, J. M., and C.W.J. Granger. "The combination of forecasts." Operational Research Quarterly, 20 (1969), 451–468.

Baur, D. G.; T. Dimpfl; and R. Jung. "Stock return autocorrelations revisited: A quantile regression approach." Journal of Empirical Finance, 19, 2 (2012), 254-265.

Buchinsky, M. "Changes in U.S. Wage Structure 1963-1987: An application of Quantile Regression." Econometrica, 62 (1994), 405-458.

Buchinsky, M. "Quantile Regression Box-Cox Transformation model, and the U.S.

wage structure, 1963-1987." Journal of Econometrics, 65 (1995), 109-154.

Buchinsky, M. "Recent Advances in Quantile Regression Models: A Practical Guidline for Empirical Research." Journal of Human Resources, 33 (1998), 88-126.

Campbell, J. Y., and J. H. Cochrane. "By force of habit: A consumption-based explanation of aggregate stock market behavior." The Journal of Political Economy, 107 (1999), 205–251.

Campbell, J. Y., and S. B. Thompson. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?." Review of Financial Studies, 21 (2008), 1509–31.

Campbell, J.Y., and L. Viceira. "Strategic Asset Allocation." Oxford University Press, Oxford (2002).

Cenesizoglu, T., and A. Timmermann. "Is the distribution of stock returns predictable?." Available online at SSRN http://ssrn.com/abstract=1107185, (2008).

Cenesizoglu, T., and A. Timmermann. "Do Return Prediction Models Add Economic Value?." Journal of Banking and Finance, 36, 11 (2012), 2974-2987.

Chan, Y. L.; J. H. Stock; and M. W. Watson. "A dynamic factor model framework for forecast combination." Spanish Economic Review, 1 (1999), 21–121.

Chen, Q. and Y. Hong. "Predictability of equity returns over different time horizons: a nonparametric approach." Manuscript, Cornell University, (2010).

Chong, Y.Y., and D.F. Hendry. "Econometric evaluation of linear macroeconomic models." Review of Economic Studies, 53 (1986), 671-690.

Chuang, C.-C.; C.-M. Kuan; and H.-Y. Lin. "Causality in quantiles and dynamic stock return-volume relations." Journal of Banking and Finance, 33, 7 (2009), 1351-1360.

Clements, M. P., and D. F. Hendry. "Forecasting Economic Time Series." Cambridge University Press, Cambridge (1998).

Cochrane, J. H. "Production-Based Asset Pricing and the Link between Stock Returns and Economic Fluctuations." Journal of Finance, 46 (1991), 209–37.

Dangl, T., and M. Halling. "Predictive regressions with time-varying coefficients." Journal of Financial Economics, 106, 1 (2012), 157-181.

Della Corte, P.; L. Sarno; and I. Tsiakas. "An Economic Evaluation of Empirical Exchange Rate Models." Review of Financial Studies, 22, 9 (2009), 3491-3530.

Della Corte, P.; L. Sarno; and G. Valente. "A Century of Equity Premium Predictability and Consumption-Wealth Ratios: An International Perspective." Journal of Empirical Finance, 17 (2010), 313-331.

Della Corte, P.; L. Sarno; and G. Sestieri. "The Predictive Information Content of External Imbalances for Exchange Rate Returns: How Much is it Worth?." Review of Economics and Statistics, 94, 1 (2012), 100-115.

Diebold, F. X., and R. S. Mariano. "Comparing Predictive Accuracy." Journal of Business and Economic Statistics, 13 (1995), 253–63.

Engle, R. F., and S. Manganelli. "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles." Journal of Business and Economic Statistics, 22 (2004), 367-381.

Ferreira, M.I., and P. Santa-Clara. "Forecasting stock market returns: the sum of the parts is more than the whole." Journal of Financial Economics, 100 (2011), 514–537.

Fleming, J.; C. Kirby; and B. Ostdiek. "The Economic Value of Volatility Timing." Journal of Finance, 56 (2001), 329-352.

Franses, P.H. and D. van Dijk. "Nonlinear Time Series Models in Empirical Finance". Cambridge University Press, Cambridge, United Kingdom, (2000).

Gastwirth, J.L. "On robust procedures." Journal of the American Statistical Association, 61 (1966), 929-948.

Goetzmann, W.; J. Ingersoll; M. Spiegel; and I. Welch. "Portfolio Performance Manipulation and Manipulation-proof Performance Measures." Review of Financial Studies, 20 (2007), 1503-1546.

Goyal, A. and I. Welch. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." Review of Financial Studies, 21 (2008), 1455–508.

Granger, C. W. J., and P. Newbold . "Some Comments on the Evaluation of Economic Forecasts." Applied Economics, 5 (1973), 35-47.

Granger, C. W. J., and R. Ramanathan. "Improved methods of combining forecasts."

Journal of Forecasting, 3 (1984), 197-204.

Guidolin, M.; S. Hyde; D. McMillan and S. Ono. "Non-linear predictability in stock and bond returns: When and where is it exploitable?" International Journal of Forecasting, 25 (2009), 373–399.

Guidolin, M. and A. Timmermann. "Forecasts of US short-term interest rates: A flexible forecast combination approach." Journal of Econometrics, 150 (2009), 297–311.

Han, Y. "Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model." Review of Financial Studies, 19 (2006), 237-271.

Hansen, B. "Least-squares forecast averaging." Journal of Econometrics, 146 (2008), 342-350.

Harvey, D. I.; S. J. Leybourne; and P. Newbold. "Tests for Forecast Encompassing." Journal of Business and Economic Statistics, 16 (1998), 254–59.

Harvey, D. and P.Newbold. "Tests for multiple forecast encompassing." Journal of Applied Econometrics, 15 (2000), 471–482.

Hastie, T.; R. Tibshirani, and J.Friedman. "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", Springer, (2009).

Henkel, S.J.; J.S. Martin; and F. Nadari. "Time-varying short-horizon predictability." Journal of Financial Economics, 99 (2011), 560–580.

Hendry, D. F., and M. P. Clements. "Pooling of Forecasts." Econometrics Journal, 7 (2004), 1–31.

Hsiao, C., and S. K. Wan. "Is there an optimal forecast combination?." Journal of Econometrics, (2012), forthcoming.

Jondeau, E., and M. Rockinger. "The Economic Value of Distributional Timing". Swiss Finance Institute Research Paper No. 06-35 (2006).

Judge, G. G.; R. C. Hill; W. E. Griffiths; H. Lutkepohl; and T.-C. Lee. "Introduction to the Theory and Practice of Econometrics." New York, Wiley (1988).

Koenker, R., and G. Bassett. "Regression Quantiles." Econometrica, 46 (1978), 33-50.

Ludvigson, S. C., and S. Ng. "The empirical risk-return relation: a factor analysis approach." Journal of Financial Economics, 83 (2007), 171–222.

Ma, L., and L. Pohlman. "Return forecasts and optimal portfolio construction: a quantile regression approach." European Journal of Finance, 14, 5 (2008), 409–425

Marquering, W., and M. Verbeek. "The Economic Value of Predicting Stock Index Returns and Volatility." Journal of Financial and Quantitative Analysis, 39 (2004), 407–29.

Meligkotsidou, L.; I. D. Vrontos; and S. D. Vrontos. "Quantile Regression Analysis of Hedge Fund Strategies." Journal of Empirical Finance, 16 (2009), 264-279.

Menzly, L.; T. Santos; and P. Veronesi. "Understanding Predictability." Journal of Political Economy, 112, 1 (2004), 1-47.

Neely, C. J.; D. E. Rapach; J. Tu; and G. Zhou. "Forecasting the equity risk premium: the role of technical indicators." Federal Reserve Bank of St. Louis Working Paper 2010-008C, (2011).

Pedersen, T. Q. "Predictable return distributions." Available online at SSRN http://ssrn.com /abstract=1658394, (2010).

Rapach, D. and J. Strauss. "Forecasting U.S. State-Level Employment Growth: An Amalgamation Approach." Journal of International Forecasting, 28 (2012), 315–327.

Rapach, D.; J. Strauss; and G. Zhou. "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy." Review of Financial Studies, 23, 2 (2010), 821-862.

Rapach, D., and G. Zhou. "Forecasting stock returns." In preparation for G. Elliott and A. Timmermann, Eds., Handbook of Economic Forecasting, Vol. 2, (2012).

Spiegel, M. "Forecasting the equity premium: Where we stand today." The Review of Financial Studies, 21 (2008), 1453–1454.

Stock, J. H., and M. W. Watson. "Combination Forecasts of Output Growth in a Seven-Country Data Set." Journal of Forecasting, 23 (2004), 405–30.

Taylor, J. W. "Forecasting daily supermarket sales using exponentially weighted quantile regression." European Journal of Operational Research, 178 (2007), 154-167.

Terasvirta, T. "Forecasting economic variables with nonlinear models." In Handbook of Economic Forecasting, Vol. I, G. Elliott, C. W. J. Granger and A. Timmermann, eds.

Amsterdam, Elsevier (2006).

Tibshirani, R. "Regression Shrinkage and Selection via the Lasso". Journal of the Royal Statistical Society, Series B, 58 (1996), 267-288.

Timmermann, A. "Forecast combinations." In Handbook of Economic Forecasting, Vol. I, G. Elliott, C. W. J. Granger and A. Timmermann, eds. Amsterdam, Elsevier (2006).

Tukey, J. W. "Explanatory Data Analysis." Addison-Wesley, Reading, MA (1977).

Yu, K.; Z. Lu; and J. Stander. "Quantile regression: applications and current research areas." The Statistician, 52 (2003), 331-350.

Wachter, J., and M. Warusawitharana. "Predictable returns and asset allocation: Should a skeptical investor time the market?." Journal of Econometrics, 148, 2 (2009), 162-178.

White, H. "Approximate nonlinear forecasting methods." In Handbook of Economic Forecasting, Vol. I, G. Elliott, C. W. J. Granger and A. Timmermann, eds. Amsterdam, Elsevier (2006).

West, K.; H. Edison; and D. Cho. "A Utility-based Comparison of Some Models of Exchange Rate Volatility." Journal of International Economics, 35 (1993), 23-46.

Wu, Y. and Y. Liu. "Variable selection in quantile regression". Statistica Sinica, 19 (2009), 801-817.

Table 1. Conditional Quantile Predictive Ability

Panel A: Individual predictive models

Predictor	Q5	Q10	Q15	Q20	Q25	Q30	Q35	Q40	Q45	Q50	Q55	Q60	Q65	Q70	Q75	Q80	Q85	Q90	Q95
D/P		█	█						█	█	█	█	█	█	█	█			
D/Y		█	█						█	█	█	█	█	█	█	█			
E/P	█	█				█			█	█	█	█	█						
D/E										█	█	█			█	█	█	█	█
SVAR	█																█	█	█
B/M										█									
NTIS		█									█	█							█
TBL							█	█	█	█									
LTY																			
LTR								█	█	█				█		█			
TMS																			
DFY		█	█									█	█	█	█	█	█	█	
DFR	█		█	█	█	█	█	█	█		█	█							
INFL		█	█	█		█		█	█	█	█								
I/K		█	█	█	█	█	█	█	█	█						█	█		

Panel B: Combining Methods

	Q5	Q10	Q15	Q20	Q25	Q30	Q35	Q40	Q45	Q50	Q55	Q60	Q65	Q70	Q75	Q80	Q85	Q90	Q95
Mean	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Median	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█		█	█
Trimmed Mean	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
DALFE(1)	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
DALFE(0.9)		█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
AL Cluster 2	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
AL Cluster 3	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
AL Principal Components	█	█	█	█	█	█	█		█			█	█	█	█	█	█	█	█
AL Lasso	█	█		█	█	█	█	█	█	█				█	█	█		█	█
AL Ridge	█	█		█	█	█	█	█	█	█				█	█	█		█	█

Notes: Q5- Q95 denote the 5% to 95% quantiles of the return distribution. Grey cells denote superior predictive ability, i.e. lower out-of-sample values of the expected asymmetric loss, $\sum_t \rho_\tau\left(r_{t+1} - \hat{r}_{t+1}^{(\tau)}\right)$, associated with the quantile forecasts of each model specification or combining method (shown in the first column of the table), than the value associated with the forecasts of the prevailing quantile (PQ) model.

Table 2. Out-of-sample performance of individual predictive mean regression models and Combining Mean Forecasts (CMF) approach

Predictor	MSFE Ratio	λ_A	λ_B	CMF approach	MSFE Ratio	λ_A	λ_B
D/P	0.9928	0.580**	0.420	Mean	0.9703	2.597***	-1.597
D/Y	0.9900	0.590**	0.410	Median	0.9781	3.209***	-2.209
E/P	1.0109	0.287	0.713	Trimmed Mean	0.9715	2.943***	-1.943
D/E	1.0160	0.177	0.824*	DMSFE(1)	0.9704	2.463***	-1.463
SVAR	1.0665	0.110	0.890	DMSFE(0.9)	0.9702	2.444***	-1.444
B/M	1.0180	0.089	0.911**	Cluster 2	0.9766	1.244**	-0.244
NTIS	1.0210	-0.192	1.192**	Cluster 3	0.9878	0.766*	0.234
TBL	1.0243	0.406**	0.594**	Principal Components	1.0169	0.347	0.653*
LTY	1.0259	0.360	0.640**				
LTR	1.0115	0.261	0.739*				
TMS	1.0265	0.373*	0.627***				
DFY	1.0271	-0.398	1.398**				
DFR	0.9909	0.627	0.373				
INFL	1.0076	0.340	0.660				
I/K	0.9768	0.693***	0.307				

Notes: The table reports the MSFE ratios of the individual predictive mean regression models and of the Combining Mean Forecasts (CMF) approach with respect to the prevailing mean (PM) benchmark model for the out-of-sample period 1965:1-2010:4. The MSFE of the PM model is equal to 0.0071. Values of the MSFE ratio below unity indicate superior forecasting performance of the predictive models with respect to the historical average forecast. Statistical significance of the out-of-sample forecasts is assessed by pairs of encompassing tests: (i) one for testing if the PM model forecasts encompass the forecasts of the individual predictive models or the CMF approach (associated with the parameter λ_A), and (ii) a second one for testing if the individual predictive models' or the CMF approach's forecasts encompass the PM model forecasts (associated with the parameter λ_B). *, **, *** indicate significance at the 10%, 5% and 1% confidence levels, respectively.

Table 3. Out-of-sample performance of robust point forecasts and Combining Robust Forecasts (CRF) approach
Panel A: Individual Predictive Models – Fixed weighting (FW) schemes

	FW1			FW2			FW3			FW4		
	MSFE Ratio	λ_A	λ_B	MSFE Ratio	λ_A	λ_B	MSFE Ratio	λ_A	λ_B	MSFE Ratio	λ_A	λ_B
D/P	0.9841	0.901*	0.099	0.9839	0.940*	0.060	0.9842	0.841**	0.159	0.9899	0.656**	0.344
D/Y	0.9786	0.939**	0.061	0.9753	1.045**	-0.045	0.9789	0.860**	0.140	0.9837	0.703**	0.297
E/P	1.0165	0.070	0.930*	1.0190	-0.007	1.007*	1.0143	0.114	0.886	1.0156	0.127	0.873
D/E	1.0013	0.477	0.523	1.0048	0.408	0.592	1.0062	0.385	0.615	1.0131	0.245	0.755*
SVAR	1.0996	0.061	0.940*	1.1044	0.042	0.958*	1.0927	0.077	0.923*	1.0904	0.051	0.949*
B/M	1.0370	-0.021	1.021**	1.0411	-0.063	1.063***	1.0310	0.030	0.970**	1.0231	0.059	0.941**
NTIS	1.0476	-0.085	1.085**	1.0543	-0.063	1.063***	1.0403	-0.111	1.111***	1.0317	-0.156	1.156***
TBL	1.0131	0.444**	0.556**	1.0204	0.414*	0.586**	1.0131	0.446**	0.555**	1.0239	0.407**	0.593**
LTY	1.0138	0.409	0.591**	1.0208	0.344	0.656**	1.0156	0.400	0.600**	1.0243	0.359	0.641**
LTR	1.0523	0.145	0.855***	1.0558	0.145	0.855***	1.0426	0.166	0.834***	1.0251	0.203	0.797**
TMS	1.0642	0.278*	0.722***	1.0638	0.293*	0.707***	1.0576	0.293*	0.708***	1.0433	0.327*	0.674***
DFY	1.0717	-0.389	1.389***	1.0655	-0.236	1.236***	1.0580	-0.454	1.454**	1.0344	-0.462	1.462**
DFR	1.0057	0.404	0.596*	1.0026	0.444	0.556	1.0055	0.405	0.595	0.9958	0.578	0.422
INFL	1.0067	0.301	0.699	1.0103	0.268	0.732	1.0071	0.298	0.702	1.0082	0.298	0.702
I/K	0.9765	0.685***	0.315	0.9763	0.647**	0.353*	0.9775	0.685	0.315	0.9742	0.724***	0.276

Panel B: Combining Robust Forecasts (CRF) approach – Fixed weighting (FW) schemes

	CRF-FW1			CRF-FW2			CRF-FW3			CRF-FW4		
Mean	0.9761	2.050**	-1.050	0.9768	1.861**	-0.861	0.9741	2.456***	-1.456	0.9720	3.144***	-2.144
Median	0.9865	1.458*	-0.458	0.9893	1.257*	-0.257	0.9848	1.850**	-0.850	0.9794	3.600***	-2.600
Trimmed Mean	0.9778	2.255**	-1.255	0.9786	2.057**	-1.057	0.9761	2.751***	-1.751	0.9737	3.757***	-2.757
DMSFE(1)	0.9755	2.081**	-1.081	0.9763	1.878**	-0.878	0.9737	2.441***	-1.441	0.9719	2.986***	-1.986
DMSFE(0.9)	0.9747	2.022**	-1.022	0.9760	1.814*	-0.814	0.9731	2.343**	-1.343	0.9716	2.842***	-1.842
Cluster 2	0.9726	1.446**	-0.446	0.9778	1.280*	-0.280	0.9744	1.394**	-0.394	0.9769	1.317**	-0.317
Cluster 3	1.0059	0.393	0.608	0.9992	0.517	0.484	1.0017	0.466	0.534	0.9861	0.791*	0.209
Principal Components	1.0289	0.317	0.683**	1.0256	0.330	0.668**	1.0284	0.318	0.682**	1.0287	0.295	0.705**

(Overlapping/misaligned values appear in the CRF-FW2 column: 2.144, 2.600, 2.757, 1.986, 1.842, 1.101, 0.317, 0.209, 0.705)

Table 3. (continued)
Panel C: Individual Predictive Models – Time-varying weighting (TVW) schemes

	TVW1			TVW2			TVW3		
	MSFE Ratio	λ_A	λ_B	**MSFE Ratio**	λ_A	λ_B	**MSFE Ratio**	λ_A	λ_B
D/P	0.9879	0.742*	0.258	0.9853	0.884*	0.116	0.9905	0.680*	0.320
D/Y	0.9876	0.729*	0.271	0.9786	0.938**	0.063	0.9882	0.694*	0.306
E/P	1.0062	0.341	0.659	1.0075	0.237	0.764	1.0111	0.274	0.726
D/E	1.0035	0.462	0.539	1.0027	0.467	0.533	1.0046	0.461	0.539*
SVAR	1.0399	0.251*	0.749	1.0643	0.170*	0.830	1.0146	0.412**	0.588
B/M	1.0065	0.357	0.643	1.0145	0.131	0.869*	1.0020	0.459	0.541
NTIS	1.0044	0.396	0.605*	1.0168	0.211	0.789**	0.9927	0.660*	0.340
TBL	1.0100	0.449*	0.551**	1.0158	0.423*	0.578**	1.0198	0.396*	0.604**
LTY	1.0140	0.396	0.604*	1.0222	0.313	0.687**	1.0168	0.375	0.626*
LTR	1.0230	0.338	0.663**	1.0256	0.315	0.685**	1.0140	0.396*	0.604**
TMS	1.0198	0.420**	0.580***	1.0314	0.382**	0.618***	1.0085	0.464**	0.536**
DFY	1.0261	-0.227	1.227**	1.0252	-0.099	1.099*	1.0148	0.169	0.831*
DFR	0.9820	0.783**	0.218	0.9761	1.080*	-0.080	0.9865	0.650*	0.350
INFL	0.9960	0.571	0.429	1.0083	0.393	0.607	0.9997	0.503	0.497
I/K	0.9821	0.616**	0.384*	0.9827	0.592***	0.408**	0.9865	0.585**	0.415*

Panel D: Combining Robust Forecasts (CRF) approach – Time-varying schemes

	CRF-TVW1			CRF-TVW2			CRF-TVW3		
Mean	0.9635	2.829***	-1.829	0.9654	2.907***	-1.907	0.9633	1.817***	-0.817
Median	0.9718	3.756***	-2.756	0.9760	5.199***	-4.199	0.9669	1.660***	-0.660
Trimmed Mean	0.9650	3.037***	-2.037	0.9677	3.314***	-2.314	0.9667	1.730***	-0.730

Notes: The table reports the MSFE ratios of the robust point forecasts and of the Combining Robust Forecasts (CRF) approach, under fixed weighting (FW) and time-varying weighting (TVW) schemes, with respect to the prevailing mean (PM) benchmark model for the out-of-sample period 1965:1-2010:4. Values of the MSFE ratio below unity indicate superior forecasting performance of the predictive models with respect to the historical average forecast. Statistical significance of the out-of-sample forecasts is assessed by pairs of encompassing tests: (i) one for testing if the PM model forecasts encompass the robust point forecasts or the forecasts of the CRF approach (associated with the parameter λ_A), and (ii) a second one for testing if the robust point forecasts or the CRF approach's forecasts encompass the PM model forecasts (associated with the parameter λ_B). *, **, *** indicate significance at the 10%, 5% and 1% confidence levels, respectively.

Table 4. Out-of-sample performance of Quantile Forecasts Synthesis (QFS) approach
Panel A: Fixed weighting (FW) schemes

	QFS -FW1			QFS -FW2			QFS -FW3			QFS -FW4		
	MSFE Ratio	λ_A	λ_B	**MSFE Ratio**	λ_A	λ_B	**MSFE Ratio**	λ_A	λ_B	**MSFE Ratio**	λ_A	λ_B
Mean	0.9761	2.050^{**}	-1.050	0.9768	1.861^{**}	-0.861	0.9741	2.456^{***}	-1.456	0.9720	3.144^{***}	-2.144
Median	0.9886	1.354^{*}	-0.354	0.9903	1.208	-0.208	0.9866	1.768^{**}	-0.768	0.9830	3.498^{***}	-2.498
Trimmed Mean	0.9785	2.214^{**}	-1.214	0.9791	2.044^{**}	-1.044	0.9768	2.727^{***}	-1.727	0.9746	3.808^{***}	-2.808
DALFE(1)	0.9758	2.047^{**}	-1.047	0.9766	1.846^{**}	-0.846	0.9738	2.430^{***}	-1.430	0.9719	3.043^{***}	-2.043
DALFE(0.9)	0.9752	2.011^{**}	-1.011	0.9760	1.825^{**}	-0.825	0.9731	2.419^{***}	-1.419	0.9711	3.092^{***}	-2.092
AL Cluster 2	0.9768	1.331^{**}	-0.331	0.9809	1.115^{**}	-0.115	0.9754	1.489^{**}	-0.489	0.9733	1.864^{**}	-0.864
AL Cluster 3	0.9798	0.965^{*}	0.035	0.9753	1.084^{**}	-0.084	0.9787	1.051^{**}	-0.051	0.9785	1.237^{**}	-0.237
AL Principal Components	1.0079	0.448^{*}	0.552^{*}	1.0160	0.417^{*}	0.583^{**}	1.0062	0.456^{*}	0.544^{*}	1.0181	0.383^{**}	0.617^{**}
AL Lasso	0.9777	0.747^{**}	0.253	0.9899	0.592^{**}	0.408	0.9782	0.755^{**}	0.245	0.9866	0.657^{**}	0.343
AL Ridge	0.9696	1.157^{**}	-0.157	0.9719	1.022^{**}	-0.022	0.9680	1.234^{**}	-0.234	0.9705	1.215^{**}	-0.215

Panel B: Time-varying weighting (TVW) schemes

	QFS -TVW1			QFS -TVW2			QFS -TVW3		
Mean	0.9594	2.138^{***}	-1.138	0.9619	2.553^{***}	-1.553	0.9677	1.387^{**}	-0.387
Median	0.9669	2.748^{***}	-1.748	0.9717	3.736^{***}	-2.736	0.9746	1.495^{**}	-0.495
Trimmed Mean	0.9619	2.216^{***}	-1.216	0.9648	2.799^{***}	-1.799	0.9702	1.382^{**}	-0.382

Notes: The table reports the MSFE ratios of the Quantile Forecasts Synthesis (QFS) approach, under fixed weighting (FW) and time-varying weighting (TVW) schemes, with respect to the prevailing mean (PM) benchmark model for the out-of-sample period 1965:1-2010:4. Values of the MSFE ratio below unity indicate superior forecasting performance of the predictive models with respect to the historical average forecast. Statistical significance of the out-of-sample forecasts is assessed by pairs of encompassing tests: (i) one for testing if the PM model forecasts encompass the QFS forecasts (associated with the parameter λ_A), and (ii) a second one for testing if the QFS forecasts encompass the PM model forecasts (associated with the parameter λ_B). *, **, *** indicate significance at the 10%, 5% and 1% confidence levels, respectively.

Table 5. Encompassing tests for pairs of forecasts from the CMF, CRF and QFS approaches

	Mean Combination Method		Median Combination Method		Trimmed Mean Combination Method	
	λ_A	λ_B	λ_A	λ_B	λ_A	λ_B
CMF, CRF-TVW1	6.262**	-5.262	3.069	-2.069	5.492*	-4.492
CMF, CRF-TVW2	2.373	-1.373	1.291	-0.291	1.964	-0.964
CMF, CRF-TVW3	1.704	-0.704	1.194*	-0.194	1.174	-0.174
CMF, QFS-TVW1	3.625**	-2.625	3.647**	-2.647	3.179*	-2.179
CMF, QFS-TVW2	4.552**	-3.552	3.280*	-2.280	3.520*	-2.520
CMF, QFS-TVW3	0.723	0.277	0.771	0.229	0.606	0.394
CRF-TVW1, QFS-TVW1	1.827	-0.827	2.509	-1.509	1.672	-0.672
CRF-TVW2, QFS-TVW2	2.250	-1.250	2.611	-1.611	2.102	-1.102
CRF-TVW3, QFS-TVW3	-3.394	4.394*	-1.573	2.573*	-3.023	4.023

Notes: The table reports results on the encompassing tests for all pairs of forecasts obtained by the Combining Mean Forecasts (CMF) approach, the Combining Robust Forecasts (CRF) approach and the Quantile Forecasts Synthesis (QFS) approach, under the three time-varying weighting schemes (TVW1-TVW3). For each pair of approaches, shown in the first column of the table, statistical significance of the out-of-sample forecasts is assessed by pairs of encompassing tests: (i) one for testing if the forecasts produced by the first approach encompass the forecasts produced by the second (associated with parameter λ_A), and (ii) a second one for testing if the forecasts produced by the second approach encompass the forecasts produced by the first (associated with parameter λ_B). *, **, *** indicate significance at the 10%, 5% and 1% confidence levels, respectively.

Table 6. Multiple encompassing tests and amalgam forecasts

	MS^* CMF	MS^* CRF-TVW	MS^* QFS-TVW	MSFE Ratio	λ_A	λ_B
Panel A: Mean Combination Methods						
CMF, CRF-TVW1, QFS-TVW1	1.981	1.644	1.310	0.9640	2.577**	-1.577
CMF, CRF-TVW2, QFS-TVW2	1.418	1.174	0.875	0.9655	2.807**	-1.807
CMF, CRF-TVW3, QFS-TVW3	3.549**	3.381**	3.669**	0.9661	1.907**	-0.907
Panel B: Median Combination Methods						
CMF, CRF-TVW1, QFS-TVW1	1.642	1.293	0.880	0.9676	3.176**	-2.176
CMF, CRF-TVW2, QFS-TVW2	0.999	0.744	0.489	0.9683	1.940**	-0.940
CMF, CRF-TVW3, QFS-TVW3	1.486	0.522	1.329	0.9827	2.529**	-1.529
Panel C: Trimmed Mean Combination Methods						
CMF, CRF-TVW1, QFS-TVW1	1.553	1.194	0.944	0.9746	3.172**	-2.172
CMF, CRF-TVW2, QFS-TVW2	0.955	0.729	0.495	0.9737	3.508**	-2.508
CMF, CRF-TVW3, QFS-TVW3	1.209	0.920	1.229	0.9728	3.760**	-2.760

Notes: The table reports results on multiple encompassing tests employed to compare the forecasts obtained by the Combining Mean Forecasts (CMF) approach, the Combining Robust Forecasts (CRF) approach and the Quantile Forecasts Synthesis (QFS) approach, under the three time-varying weighting schemes (TVW1-TVW3). Columns (2) - (4) report the MS^* statistics to test the null of multiple forecast encompassing. The test is conducted three times for every triad by employing the model in the first row as the reference model. Columns (5) - (7) report the MSFE ratios of an amalgam forecast constructed by averaging the forecasts of the three approaches, shown in the first column. Statistical significance of the out-of-sample forecasts is assessed by pairs of encompassing tests: (i) one for testing if the amalgam forecasts encompass the PM forecasts (associated with the parameter λ_A), and (ii) a second one for testing if the PM forecasts encompass the amalgam forecasts (associated with the parameter λ_B). *, **, *** indicate significance at the 10%, 5% and 1% confidence levels, respectively.

Table 7. Economic evaluation of Combining Mean Forecasts (CMF) approach and Combining Robust Forecasts (CRF) approach

Panel A: Combining Mean Forecasts		
	Φ	Θ
Mean	297.41	371.86
Median	145.30	194.80
Trimmed Mean	270.86	343.78
DMSFE(1)	304.75	379.79
DMSFE(0.9)	320.79	396.46
Cluster 2	248.87	286.90
Cluster 3	242.84	269.37
Principal Components	235.71	246.97

Panel B: Combining Robust Forecasts - Fixed weighting schemes								
	CRF-FW1		CRF-FW2		CRF-FW3		CRF-FW4	
	Φ	Θ	Φ	Θ	Φ	Θ	Φ	Θ
Mean	186.79	220.43	190.6	221.98	207.32	250.62	236.59	294.95
Median	59.01	67.84	18.42	8.10	76.89	92.33	108.40	137.98
Trimmed Mean	159.17	191.64	162.38	195.73	178.16	219.10	216.80	274.95
DMSFE(1)	204.59	241.21	206.26	239.85	223.17	268.92	249.46	309.62
DMSFE(0.9)	245.70	284.84	242.50	278.06	258.98	307.33	275.16	337.48
Cluster 2	252.01	266.45	229.06	245.53	236.64	252.93	217.11	240.74
Cluster 3	175.07	196.52	224.49	247.22	216.39	239.84	263.24	290.35
Principal Components	250.88	257.87	255.38	261.98	252.54	259.64	238.68	243.91

Panel C: Combining Robust Forecasts – Time-varying weighting schemes						
	CRF-TVW1		CRF-TVW2		CRF-TVW3	
	Φ	Θ	Φ	Θ	Φ	Θ
Mean	371.49	448.74	338.90	407.92	394.85	480.13
Median	231.73	296.58	158.60	212.87	327.55	412.53
Trimmed Mean	352.71	429.44	306.40	375.15	373.76	459.03

Notes: The performance fee, Φ, is the difference between the realized utilities of competing models, $\Phi = \Delta U = \overline{U}^i - \overline{U}^{PM}$, where $\overline{U}^i, \overline{U}^{PM}$ denote the average mean-variance utility of an investor with risk aversion coefficient of three over the forecast evaluation period from using the ith model/specification and the historical average benchmark model (PM), respectively. The weight on stocks in the investor's portfolio is restricted to lie between zero and 1.5. The mean-variance utility for the ith model/specification is given by:

$$\overline{U}^i = \frac{W_o}{P - P_0}\left[\sum_{t=0}^{P-P_0-1}(R_{p,t+1}^i) - \frac{\gamma}{2}\sum_{t=0}^{P-P_0-1}(R_{p,t+1}^i - \overline{R_p^i})^2 \right],$$

where $P - P_0$ is the number of out-of-sample forecasts, W_o is the initial wealth of the investor and γ denotes the coefficient of relative risk aversion. Θ is the difference between the manipulation-proof performance measure of the two competing models, $\Theta = M(R_p)^i - M(R_p)^{PM}$, where

$$M(R_p) = \frac{1}{1-\gamma}\ln\left\{ \frac{1}{P-P_0}\sum_{t=0}^{P-P_0-1}\left(\frac{R_{p,t+1}^{PM}}{R_{f,t}} \right)^{1-\gamma} \right\}.$$ Φ and Θ are reported in annualized basis points.

Table 8. Economic evaluation of Quantile Forecasts Synthesis (QFS) approach and amalgam forecasts

Panel A: Quantile Forecasts Synthesis - Fixed weighting schemes							
	QFS - FW1		QFS – FW2		QFS – FW3		QFS – FW4	
	Φ	Θ	Φ	Θ	Φ	Θ	Φ	Θ
Mean	186.79	220.43	190.60	221.98	207.32	250.62	236.59	294.95
Median	28.75	23.30	14.76	-1.80	45.79	48.79	83.49	107.68
Trimmed Mean	157.37	190.92	161.32	194.91	176.63	218.84	211.95	270.90
DALFE(1)	195.92	231.15	197.97	230.42	215.21	259.83	242.23	301.36
DALFE(0.9)	220.51	258.57	223.47	258.96	240.17	287.89	263.13	324.45
AL Cluster 2	250.28	261.19	243.45	256.26	248.28	261.21	261.00	293.72
AL Cluster 3	276.53	292.82	338.32	360.18	270.11	289.06	257.20	279.87
AL Principal	262.00	269.61	289.17	297.40	217.64	220.09	173.78	195.27
AL Lasso	328.87	328.74	317.67	339.1	292.53	285.12	226.94	224.99
AL Ridge	316.86	336.58	310.97	321.94	314.81	343.52	295.4	337.64

Panel B: Quantile Forecasts Synthesis – Time-varying weighting schemes						
	QFS - TVW1		QFS - TVW2		QFS - TVW3	
	Φ	Θ	Φ	Θ	Φ	Θ
Mean	424.97	511.01	383.53	463.91	358.33	444.58
Median	314.28	394.97	242.92	317.86	294.50	378.65
Trimmed Mean	405.73	490.10	356.23	436.45	342.67	427.31

Panel C: Amalgam Forecasts						
	Mean Combination Method		Median Combination Method		Trimmed Mean Combination Method	
	Φ	Θ	Φ	Θ	Φ	Θ
CMF, CRF-TVW1, QFS-TVW1	372.38	452.66	316.12	391.60	198.58	248.85
CMF, CRF-TVW2, QFS-TVW2	346.79	422.88	352.95	435.21	214.42	270.29
CMF, CRF-TVW3, QFS-TVW3	374.93	457.80	79.06	98.46	235.60	299.62

Notes: The performance fee, Φ, is the difference between the realized utilities of competing models, $\Phi = \Delta U = \overline{U}^i - \overline{U}^{PM}$, where $\overline{U}^i, \overline{U}^{PM}$ denote the average mean-variance utility of an investor with risk aversion coefficient of three over the forecast evaluation period from using the ith model/specification and the historical average benchmark model (PM), respectively. The weight on stocks in the investor's portfolio is restricted to lie between zero and 1.5. The mean-variance utility for the ith model/specification is given by:

$$\overline{U}^i = \frac{W_o}{P - P_0}\left[\sum_{t=0}^{P-P_0-1}(R_{p,t+1}^i) - \frac{\gamma}{2}\sum_{t=0}^{P-P_0-1}(R_{p,t+1}^i - \overline{R_p^i})^2 \right],$$

where $P - P_0$ is the number of out-of-sample forecasts, W_o is the initial wealth of the investor and γ denotes the coefficient of relative risk aversion. Θ is the difference between the manipulation-proof performance measure of the two competing models, $\Theta = M(R_p)^i - M(R_p)^{PM}$, where

$$M(R_p) = \frac{1}{1-\gamma}\ln\left\{ \frac{1}{P-P_0}\sum_{t=0}^{P-P_0-1}\left(\frac{R_{p,t+1}^{PM}}{R_{f,t}}\right)^{1-\gamma} \right\}.$$ Φ and Θ are reported in annualized basis points.