WAGNER PIAZZA GAGLIANONE
LUIZ RENATO LIMA

Constructing Density Forecasts from Quantile Regressions

The departure from the traditional concern with the central tendency is in line with the increasing recognition that an assessment of the degree of uncertainty surrounding a point forecast is indispensable (Clements 2004). We propose an econometric model to estimate the conditional density without relying on assumptions about the parametric form of the conditional distribution of the target variable. The methodology is applied to the U.S. unemployment rate and the survey of professional forecasts. Specification tests based on Koenker and Xiao (2002) and Gaglianone et al. (2011) indicate that our approach correctly approximates the true conditional density.

JEL codes: C13, C14, C51, C53
Keywords: density forecast, loss function, quantile regression, surveys.

ALTHOUGH FORECASTERS SOMETIMES ATTACH measures of uncertainty, until recently most forecasts were provided in the form of point forecasts (see Diebold, Gunther, and Tay 1998). Density forecasts, usually stated as probability distributions, are becoming more common. They inform the user of the forecast about the risks involved in using the forecast for decision making. For example, monetary authorities in inflation-targeting countries often focus their attention on the probability of future inflation falling within some predefined target range. Users of growth forecasts may be concerned about the probability of recession rather than a specific point estimation of the GDP growth rate. Moreover, volatility forecasts, as measured by the variance, and other measures of risk and uncertainty, can be extracted from the density forecast.

The authors thank the editor Kenneth D. West, two anonymous referees, Emanuel Kohlscheen, Fabio Araujo, and participants of various conferences and seminars for their helpful comments and suggestions. Lima thanks CNPq for financial support as well as seminar participants of the University of Tennessee. The opinions expressed in this article are those of the authors and do not necessarily reflect those of the Banco Central do Brasil. Any remaining errors are ours.

WAGNER PIAZZA GAGLIANONE *is at the Research Department of the Banco Central do Brasil and Fucape Business School* (*E-mail:* wagner.gaglianone@bcb.gov.br). LUIZ RENATO LIMA *is at the Department of Economics of the University of Tennessee at Knoxville and Federal University of Paraíba, Brazil* (*E-mail:* llima@utk.edu).

Received October 22, 2009; and accepted in revised form April 28, 2011.

Journal of Money, Credit and Banking, Vol. 44, No. 8 (December 2012)

The main concern about the current literature of density forecast is that it places some parametric structure on the shape of the conditional distribution. If this parametric representation is misspecified, then density forecasts will probably be misleading. The usage of nonparametric methods normally requires an enormous amount of data, which are not common in macroeconomic forecasting. Besides, nonparametric methods are highly computer intensive, which may discourage applied users. This paper aims to provide a simple, though accurate, approach to compute a density forecast for various forecasting horizons. Our approach is based on quantile regression techniques and can be seen as a generalization of the idea proposed by Capistrán and Timmermann (2009) used to obtain the best point forecast.

We consider a general econometric model with dynamics in the conditional mean and variance. We discuss the curse-of-dimensionality problem that plagues many econometric models and propose to address it by allowing the consensus forecast to affect both location and scale of the distribution of the target variable (y_{t+h}). By imposing a quite weak assumption on the loss function used by individual forecasters, we identify the optimal individual forecast as a function of a common factor (here represented by the consensus forecast) and an idiosyncratic component that depends solely on the individual loss function. We also show that this optimal individual forecast corresponds to some conditional quantile of y_{t+h}. Finally, given our econometric model, a family of conditional quantiles can be estimated and we use it to construct the density forecasts.

We illustrate the applicability of this method by analyzing forecasts of the U.S. unemployment rate during the time period that includes the subprime economic crisis. Our results indicate that the economic crisis seems to have affected the scale of the distribution strongly, resulting in larger probabilities for the event that the unemployment will exceed the 9% or 10% rate in the near future. Finally, in order to validate the estimates presented in this paper, we conduct robustness tests and compare our results with the ones obtained using a benchmark GARCH (1,1) model. Our robustness analysis indicates that the proposed quantile method is superior to other methods that rely on parametric assumptions of the distribution function.

This paper is organized as follows. Section 1 presents the problem, discusses the econometric model and assumptions, and presents our results on forecast optimality. Section 2 shows how to construct a density forecast from quantile regressions and presents our empirical illustration with U.S. unemployment rate data. Section 3 concludes.

1. THE ECONOMETRIC MODEL

Suppose that an agent is interested in forecasting some stationary univariate time series $\{y_{t+h}\}_{h=1}^{\infty}$, given the information available at time t, $\mathcal{F}_t$. We denote the conditional distribution of y_{t+h} given $\mathcal{F}_t$ as $F_{t+h,t}$, and the conditional density as $f_{t+h,t}$. Note that the parametric form of this conditional distribution (density) is not known.

The current literature has proposed a parametric approach to estimate $f_{t+h,t}$, but in this paper we will not adopt this method. Point forecasts conditional on $\mathcal{F}_t$ are denoted as $\widehat{y}_{t+h,t}$ and a density forecast is an estimate $\widehat{f}_{t+h,t}$ of the future density (or distribution function) of a random variable, conditional on $\mathcal{F}_t$. The data-generating process (DGP) with conditional mean and variance dynamics is defined as

$$
\begin{aligned}
y_{t+h} &= X'_{t+h,t}\alpha + (X'_{t+h,t}\gamma)\eta_{t+h}, \\
\eta_{t+h}|\mathcal{F}_t &\sim iid\ F_{\eta,h}\,(0,1)\,,
\end{aligned}
\tag{1}
$$

where $F_{\eta,h}(0, 1)$ is some distribution with mean zero and unit variance, which depends on h but does not depend on $\mathcal{F}_t$; $X_{t+h,t} \in \mathcal{F}_t$ is a $k \times 1$ vector of covariates that can be predicted using information available at time t; and α and γ are $k \times 1$ vectors of parameters. This class of DGPs is very broad and includes most common volatility processes (e.g., ARCH and stochastic volatility). The important thing to notice is that no parametric structure is placed on $F_{\eta,h}$. In this model, covariates affect the location as well as the scale of the distribution. This model is quite general but may suffer from the curse-of-dimensionality problem. Indeed, let T be the sample size available to estimate (1), then if the size of the information set $\mathcal{F}_t$ is large enough, there will be a large amount of parameters k to be estimated. This leads to loss of degrees of freedom and large out-of-sample forecast errors. Besides, standard estimators (such as ordinary least squares [OLS]) are not feasible when $k > T$.

As mentioned by Stock and Watson (2006), the challenge of many predictor forecasting is to "turn dimensionality from a curse into a blessing." There are many methods for forecasting economic time series variables using many predictors. In the dynamic factor model with principal component analysis (PCA), h-step-ahead forecasts are produced by regressing y_{t+h} against the estimated factor $\widehat{C}_t$ and, possibly, lags of $\widehat{C}_t$ and y_t. To obtain iterated h-step-ahead forecasts, one needs to specify a subsidiary model of the dynamic process followed by the common factor C_t. One approach, proposed by Bernanke, Boivin, and Eliasz (2005) models (y_t, C_t) jointly as a vector autoregression, which is estimated by using $\widehat{C}_t$ (PCA estimates) in place of C_t. References on applications of dynamic factor models and PCA to economic forecasting can be found in Stock and Watson and, more recently, in Diebold, Aruoba, and Scotti (2009).

Another method used to address the curse of dimensionality problem is based on the combination of many individual forecasts. Forecast combination is the pooling of two or more individual forecasts from a panel of forecasts to produce a single, pooled forecast. This method was originally developed by Bates and Granger (1969) for pooling forecasts from separate forecasters, whose forecasts may or may not be based on statistical models. Data on forecasters' opinions can be easily obtained from surveys of professional forecasters, such as the Blue Chip Survey published in the Blue Chip Economic Indicators, the Survey of Professional Forecasters (SPF) and the Livingston survey of forecasts, both published by the Federal Reserve Bank of Philadelphia. These surveys summarize the forecasts of experts from industry, government, banking, and academia, and the results of the surveys are often released in terms of mean and median forecasts of all the respondents (consensus forecast), as well as the individual responses from each analyst.

In a recent paper, Capistrán and Timmermann (2009) study the combination of forecasts from survey data by using various combination methods in common use, such as the equal-weighted forecast, previous best forecast, least squares estimation of combination weights, shrinkage, and the odds matrix approach. Since combining forecasts from survey data is complicated by the frequent entry and exit of individual forecasters, they also considered a new method that projects the target variable on the consensus forecast, defined in their paper as being the average forecast from the survey. In other words, they consider the following forecasting equation:

$$y_{t+h} = \alpha_0 + \alpha_1 C_{t+h,t} + \varepsilon_{t+h}, \tag{2}$$

where $C_{t+h,t}$ is the average forecast of y_{t+h} based on the information available at time t, and ε_{t+h} is the forecasting error.

If the goal is to minimize the mean squared error (MSE) loss function, then the optimal point forecasts of y_{t+h} obtained from model (2) will simply be $\widehat{y}_{t+h,t} = \widehat{\alpha}_0 + \widehat{\alpha}_1 C_{t+h,t}$, where $\widehat{\alpha}_0$ and $\widehat{\alpha}_1$ are OLS estimates of the intercept and slope parameters. Although model (2) can be used to obtain optimal point forecasts under MSE loss, it says little about optimality under loss functions other than the MSE. To motivate, assume that there exists a continuum of forecasters who make their best predictions by using loss functions L_i that differs from MSE. In this case, as we will show, the forecast $\widehat{y}_{t+h,t} = \widehat{\alpha}_0 + \widehat{\alpha}_1 C_{t+h,t}$ is no longer optimal and disagreement among forecasters would result from this heterogeneity of loss functions. In what follows, we generalize the model (2) by allowing a heterogeneous population of forecasters with different loss functions. In this new approach, the parameters α_0 and α_1 will depend on L_i and optimal density forecasts can naturally be constructed from quantile regressions.

In order to develop this new approach, we rewrite DGP(1) as follows

$$\begin{aligned} y_{t+h} &= X'_{t+h,t}\alpha + (X'_{t+h,t}\gamma)\eta_{t+h} \\ X'_{t+h,t} &= (1, C_{t+h,t}) \\ \alpha &= (\alpha_0, \alpha_1)' \\ \gamma &= (\gamma_0, \gamma_1)' \\ \eta_{t+h}|\mathcal{F}_t &\sim iid\ F_{\eta,h}(0,1), \end{aligned} \tag{3}$$

where $C_{t+h,t}$ is the consensus forecast, defined as the average or median forecast from some survey of forecasters, and α and γ are 2×1 vectors of parameters. In this model, the consensus forecast affects both the location and scale of the conditional distribution of y_{t+h}. This is also a parsimonious model and, therefore, does not suffer from the curse-of-dimensionality problem.

Our next step is to derive the optimal individual forecast under DGP(3) and some class of unknown loss functions. Following the literature (i.e., Granger 1969, Granger and Newbold 1986, Christoffersen and Diebold 1997, Patton and Timmermann 2007), each individual chooses an optimal forecast $\widehat{y}^i_{t+h,t}$ by minimizing an expected loss function L^i. In this paper we assume that the loss functions are defined according to Assumption 1 below.

ASSUMPTION 1 (LOSS FUNCTION). *The loss function L^i is a homogeneous function solely of the forecast error $e^i_{t+h,t} \equiv y_{t+h} - \widehat{y}^i_{t+h,t}$, that is, $L^i = L(e^i_{t+h,t})$, and $L(ae) = g(a)L(e)$ for some positive function g.*[1]

Proposition 1 presents our result on forecast optimality.

PROPOSITION 1. *Under DGP(3) and a homogeneous loss function (Assumption 1), the optimal forecast will be*

$$\widehat{y}^i_{t+h,t} = \alpha_0(\tau_i) + \alpha_1(\tau_i)C_{t+h,t}, \tag{4}$$

where $\alpha_0(\tau_i) = (\alpha_0 + \gamma_0\gamma^i_h)$, $\alpha_1(\tau_i) = (\alpha_1 + \gamma_1\gamma^i_h)$, and γ^i_h is a constant that depends only on the distribution $F_{\eta,h}(0, 1)$ and the loss function L^i, $i \in (0, 1)$.

PROOF: See the Appendix. □

From equation (4) we see that the individual optimal forecast is an affine function of the common factor $C_{t+h,t}$, whereas the functions $\alpha_0(\tau_i)$ and $\alpha_1(\tau_i)$ are the idiosyncratic components. Thus, Proposition 1 shows that optimal individual forecasts are based on the same information set $\mathcal{F}_t$ and that differences in opinion among professional forecasters are explained by the presence of different loss functions L^i.[2]

Given this optimality result, it is now important to show the relationship between an optimal individual forecast and the conditional quantiles of y_{t+h}. Recall that $F_{t+h,t}$ is the conditional distribution of y_{t+h} and therefore

$$\begin{aligned} F_{t+h,t}\left(\widehat{y}^i_{t+h,t}\right) &= \Pr\left(y_{t+h} < \widehat{y}^i_{t+h,t} | \mathcal{F}_t\right) \\ &= \Pr\begin{pmatrix} y_{t+h} = \alpha_0 + \alpha_1 C_{t+h,t} + (\gamma_0 + \gamma_1 C_{t+h,t})\eta_{t+h} \\ \quad < \alpha_0 + \alpha_1 C_{t+h,t} + (\gamma_0 + \gamma_1 C_{t+h,t})\gamma^i_h | \mathcal{F}_t \end{pmatrix} \\ &= \Pr\left(\eta_{t+h} < \gamma^i_h \middle| \mathcal{F}_t\right) = F_{\eta,h}\left(\gamma^i_h\right), \end{aligned} \tag{5}$$

where $F_{\eta,h}(\gamma^i_h) = \tau_i$ is a fixed value of $\tau \in (0, 1)$. Thus, it follows that, by definition, the optimal forecast $\widehat{y}^i_{t+h,t}$ must coincide with the conditional quantile function of

1. This is exactly the same Assumption L2 of Patton and Timmermann (2007). Although it rules out certain loss functions (e.g., those that also depend on the level of the predicted variable), many common loss functions are of this form, such as MSE, mean absolute error (MAE), lin-lin, and asymmetric quadratic loss.

2. Recent research by Patton and Timmermann (2010) supports our findings by pointing out that differences in opinion are not primarily driven by differences in information.

y_{t+h} at level τ_i:

$$\widehat{y}^i_{t+h,t} = Q_{y_{t+h}}(\tau_i \mid \mathcal{F}_t), \qquad \text{for some } \tau_i \in (0, 1). \tag{6}$$

This result was first presented by Weiss (1996) and more recently by Patton and Timmermann (2007) who show that in this framework in which the loss function depends solely on the forecast error, the optimal forecast is some summary measure of the true conditional density of the variable y_{t+h} (the mean for quadratic loss, the median for absolute error loss, and the quantile for more general loss). As shown in (6), the consequence of this result is that the forecaster with an unknown loss function that satisfies Assumption 1 will only care about the accuracy of a specific conditional quantile of y_{t+h}.[3]

Thus, the ith forecaster specializes in predicting $Q_{y_{t+h}}(\tau_i \mid \mathcal{F}_t)$ and, if a continuum of individual forecasts were available, we could use it to estimate the conditional density. However, as the cross-section observations in the SPF are very short, no such a continuous amount exist. We notice that it is possible to identify all possible missing individual forecasts by fully exploring the relationship between an individual optimal forecast and some conditional quantile τ_i of y_{t+h}. In other words, from DGP(3) we can identify a family of conditional quantiles $Q_{y_{t+h}}(\tau \mid \mathcal{F}_t)$, $\tau \in (0, 1)$ as follows:

$$\begin{aligned} Q_{y_{t+h}}(\tau \mid \mathcal{F}_t) &= \alpha_0(\tau) + \alpha_1(\tau) C_{t+h,t}; \\ \alpha_0(\tau) &= (\alpha_0 + \gamma_0 \gamma_h) \text{ and } \alpha_1(\tau) = (\alpha_1 + \gamma_1 \gamma_h); \\ \gamma_h &= F^{-1}_{\eta,h}(\tau), \tau \in (0, 1). \end{aligned} \tag{7}$$

Equation (7) states that we can identify the quantiles of $f_{t+h,t}$ through a quantile regression of y_{t+h} on the single covariate $C_{t+h,t}$. This result can be used to estimate the conditional density $f_{t+h,t}$ without imposing any parametric form on $F_{\eta,h}$. To the best of our knowledge, no other work has fully exploited this framework.

Finally, given a family of estimated conditional quantile functions, it is straightforward to estimate the conditional density of the target variable through the formula (see Koenker 2005):

$$\widehat{f}_{t+h,t} = \frac{(\tau_i - \tau_{i-1})}{\widehat{Q}_{y_{t+h}}(\tau_i \mid \mathcal{F}_t) - \widehat{Q}_{y_{t+h}}(\tau_{i-1} \mid \mathcal{F}_t)}.$$

The conditional densities can also be estimated (for instance) by the Epanechnikov kernel, which is a weighting function that determines the shape of the bumps. We prefer the latter because it generates smooth densities, especially when the time series sample size is short, which is going to be the case in our empirical analysis presented in Section 2.

3. Patton and Timmermann (2010) report strong evidence that this type of specialization tends to persist over time. They find that most optimistic (pessimistic) forecasters continue to be optimistic (pessimistic) in the following period.

So far we proposed a quantile method to estimate the conditional density $f_{t+h,t}$ but it is natural now to discuss some forms of evaluating such a method. A conventional procedure for testing distributional assumptions is the Kolmogorov test. Diebold, Gunther, and Tay (1998) were the first ones to propose a framework for evaluating density forecasts based on Kolmogorov tests for conditional distribution in time series. However, they did not consider the effect of parameter estimation on the critical values. Indeed, it is well known that when parameters are estimated, the Kolmogorov test is not asymptotically distribution free (see Durbin 1973). This means that different critical values are needed for different distributions and for different parameter values.

Bai (2003) considered a solution to the above problem by using a martingale transformation approach, proposed by Khmaladze (1981), to derive an asymptotically distribution-free test. Bai established consistency of the test for any location-scale model such as GARCH and IGARCH. Since our quantile method for density forecasting is based on the fact that $F_{\eta,h}$ is unknown, we need a version of the Kolmogorov test that is robust not only against parameter estimation but also against uncertainty caused by the presence of unknown distribution function. Fortunately, Koenker and Xiao (2002) developed a general Kolmogorov test in which the functional form of $F_{\eta,h}$ does not need to be specified under the null hypothesis. It is, therefore, possible to test the null that the true conditional density $f_{t+h,t}$ can be generated by a specific quantile process. In addition to the Kolmogorov test, we can also apply the specification test recently developed by Gaglianone et al. (2011) to evaluate the out-of-sample performance of the proposed quantile method.

2. AN EMPIRICAL ILLUSTRATION: FORECASTING THE DENSITY OF THE U.S. UNEMPLOYMENT RATE

In this section, we illustrate our methodology by using the U.S. unemployment rate y_t (quarterly data), and a set of individual forecasts from the SPF.[4] The sample considered in this paper covers the period 1969.Q1 until 2010.Q3 (167 time series observations). For each sample period, we observe the actual unemployment rate as well as the median ($\widehat{y}^{\text{med}}_{t+h,t}$) and the average forecasts ($\widehat{y}^{\text{aver}}_{t+h,t}$) for the unemployment rate at period $t+h$ that were computed using the information available at period t. Here, h ranges from 1 to 4, meaning that in the third quarter of 2010 (i.e., 2010.Q3) we would have observed the actual unemployment rate for 2010.Q3 as well as the median and average forecasts for 2010.Q4, 2011.Q1, 2011.Q2, and 2011.Q3. We follow the previous literature (i.e., Figlewski 1983, Figlewski and Urich 1983, Capistrán and Timmermann 2009) and use the average forecast as our approximation

4. Forecasts for the quarterly average unemployment rate (seasonally adjusted, percentage points).

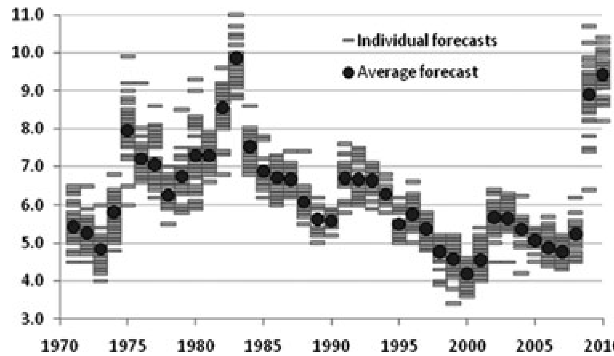

FIG 1. Individual (and Average) Forecasts for the U.S. Unemployment Rate $\{\widehat{y}^{i}_{t+4,t}; \widehat{y}^{\text{aver}}_{t+4,t}\}$.

of the consensus forecast, $C_{t+h,t}$.[5] Figure 1 confirms the existence of disagreement among individual forecasts (gray marks) and shows that the average (consensus) forecast (black dots) is able to track down the changes in the market expectations.

Since our sample ends in 2010.Q3, we will make density forecasts for the first four quarters out of our sample, that is: 2010.Q4, 2011.Q1, 2011.Q2, and 2011.Q3. According to our approach, we should take the following steps.

Step 1: Estimate the quantile regression (7) by using the one-step-ahead average forecast $\widehat{y}^{\text{aver}}_{t+1,t}$ as covariate, y_{t+1} as dependent variable, and sample t =1969.Q1, ..., 2010.Q3. In other words, we choose $\alpha_0(\tau)$ and $\alpha_1(\tau)$ that solves the following problem:

$$\min_{(\alpha_0,\alpha_1)\in\mathbb{R}^2} \sum_{t=1969.\text{Q1}}^{2010.\text{Q3}} \rho_\tau\left(y_{t+1} - \alpha_0 - \alpha_1 \widehat{y}^{\text{aver}}_{t+1,t}\right), \tag{8}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$. This is the quantile regression optimization problem developed by Koenker and Basset (1978).

The estimated quantile function obtained from solving problem (8) is equal to

$$\widehat{Q}_{y_{t+1}}(\tau \mid \mathcal{F}_t) = \widehat{\alpha}_0(\tau) + \widehat{\alpha}_1(\tau)\, \widehat{y}^{\text{aver}}_{t+1,t}. \tag{9}$$

Step 2: Use (9) to forecast the quantiles of the unemployment rate at 2010.Q4 by just evaluating (9) at the last observation of $\widehat{y}^{\text{aver}}_{t+1,t}$, that is, compute

$$\widehat{Q}_{y_{2010.Q4}}(\tau \mid \mathcal{F}_{2010.\text{Q3}}) = \widehat{\alpha}_0(\tau) + \widehat{\alpha}_1(\tau)\, \widehat{y}^{\text{aver}}_{2010.\text{Q4},2010.\text{Q3}},$$

for various $\tau \in (0, 1)$.

5. Results based on the median forecast are close to the ones obtained using the average forecast and are available upon request. Unlike the mean, the median has the property to be robust against outliers.

Step 3: Construct the density forecast of the unemployment rate at 2010.Q4 by using the quantile interpolating methods suggested in the previous section. Here, we particularly use the Epanechnikov kernel because it generates smooth densities, especially when the time series sample size is short, which is the case in our empirical analysis. The forecast of densities for longer horizons (i.e., $h > 1$) can be generated in a similar way except that we have to replace in (9) the regressand y_{t+1} and the regressor $y^{\text{aver}}_{t+1,t}$ by y_{t+h} and $\widehat{y}^{\text{aver}}_{t+h,t}$, respectively. The quantile functions are estimated for τ ranging from 0.05 to 0.95.

All the results from the proposed quantile method will be compared with the ones from a benchmark model. To motivate such a comparison we chose a model that presents some similarities to the location-scale model (3), but also differs in some relevant aspects. In other words, we keep the conditional mean equal to $\alpha_0 + \alpha_1 C_{t+h,t}$ but we now assume that the distribution function is symmetric and known so that the benchmark model can no longer capture the increased risk of a higher unemployment rate in the U.S. economy. If such an upside risk is an important feature exhibited by the data, then the model that fails to capture it should be rejected empirically.

Given a known parametric form of the distribution function, a specification test can be carried out using the Kolmogorov test proposed by Bai (2003), which is robust against parameter estimation. For the special case of GARCH models with Gaussian innovations, Bai (2003) provides the functional form of the transforming functions used to compute the test statistic. For these reasons, we considered the following benchmark GARCH(1,1) model:

$$\begin{aligned} y_{t+h} &= \alpha_0 + \alpha_1 C_{t+h,t} + \sigma_{t+h}\varepsilon_{t+h}, \\ \varepsilon_{t+h} &\sim i.i.d.\ N(0,1), \end{aligned} \tag{10}$$

where $\sigma^2_{t+h} = \beta_0 + \beta_1\sigma^2_{t+h-1} + \gamma(y_{t+h-1} - \alpha_0 - \alpha_1 C_{t+h-1,t-1})^2$, and $\varepsilon_{t+h} = \sigma^{-1}_{t+h}(y_{t+h} - \alpha_0 - \alpha_1 C_{t+h,t})$.[6] The parameters are assumed to satisfy $\beta_0 > 0$, $\gamma \geq 0$, $\beta_1 \geq 0$, and $(\beta_1 + \gamma) < 1$. Notice that (10) is a location-scale model with location parameter given by $\alpha_0 + \alpha_1 C_{t+h,t}$ and with the consensus forecast affecting the scale of the distribution through the term $\gamma(\sigma_{t+h-1}\varepsilon_{t+h-1})^2$. Moreover, the distribution function is assumed to be known and symmetric about the mean whereas it is unknown in model (3). Thus, we want to investigate whether this forecasting method based on parametric assumptions of the distribution function can be supported by the data. In order to do so, we use (10) to estimate the conditional quantiles $Q_{y_{t+h}}(\tau \mid \mathcal{F}_t)$ for τ ranging from 0.05 to 0.95 and the density is constructed using the same interpolation method of Section 1.

The conditional densities computed using models (3) and (10) appear in Panels 1 and 2 of Figure 2, respectively. Notice that the shape of the conditional densities computed using the location-scale model (3) does not seem to match up with the one from a normal distribution, especially for longer forecast horizons.

6. The Gaussian maximum likelihood method is used to estimate the parameters of this model.

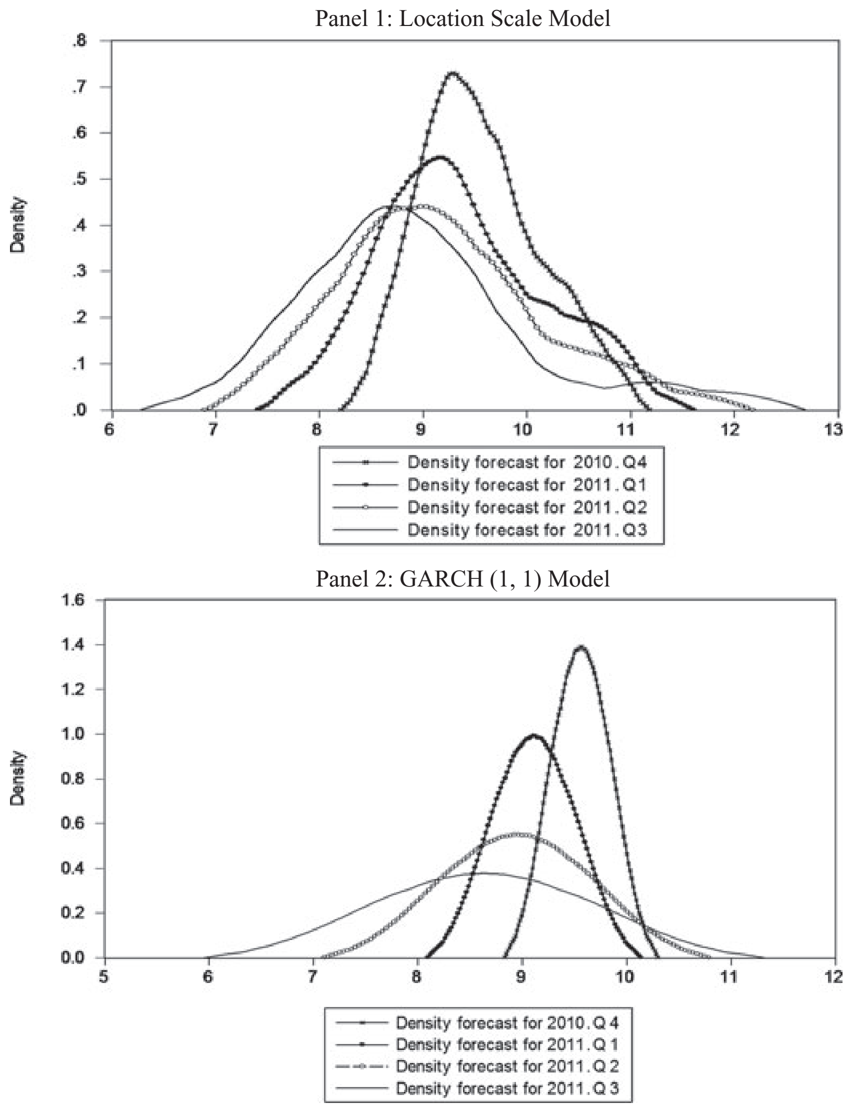

FIG 2. Density Forecast of the U.S. Unemployment Rate Conditional at SPF Average Forecasts.

Indeed, unlike the standard parametric approach currently adopted in practice, the quantile regression setup allows us to forecast the density without relying on the normality assumption. In this sense, the resulting density forecast is far from normal and is therefore able to reflect the current increased risk of a higher unemployment rate in the U.S. economy provoked by the recent subprime crisis. Panel 2 shows the conditional density computed using the Gaussian GARCH(1,1) model. Compared to the density in Panel 1, it is clear that the GARCH(1,1) model is unable to capture the

TABLE 1

SPF Average Forecast, the Bias-Adjusted Mean Forecast, the Quantile Regression-Based Median and Probability Forecasts, and the GARCH(1,1)-Based Median and Probability Forecasts

Period	Observed unemployment rate (%)	SPF average forecast	Bias-adjusted mean forecast	Median conditional quantile (%)	Prob. (unp > 9%)	Prob. (unp > 10%)	GARCH(1,1) median forecast	Prob. (unp > 9%)	Prob. (unp > 10%)
2010.Q2	9.70	–	–	–	–	–	–	–	–
2010.Q3	9.57	–	–	–	–	–	–	–	–
2010.Q4	–	9.55	9.53	9.49	84%	25%	9.56	98%	6%
2011.Q1	–	9.39	9.37	9.23	66%	23%	9.11	61%	1%
2011.Q2	–	9.21	9.16	9.14	56%	19%	8.95	47%	7%
2011.Q3	–	9.02	8.92	8.77	44%	15%	8.63	36%	9%

Note: unp = unemployment.

increased risk of a higher unemployment rate for the 2011.Q1, 2011.Q2, and 2011.Q3 periods.

Given the conditional density $\widehat{f}_{t+h,t}$, we can compute the probability that the unemployment rate will exceed a given level, which might be a useful tool for making decisions on relevant macroeconomic policies. For example, the current economic crisis severely affected the unemployment rate in the U.S., raising it to more than 10% in the last quarter of 2009 and to more than 9% during the first three quarters of 2010. Thus, a policymaker might be interested in forecasting the probability that the unemployment rate will surpass say 10% or 9% in the near future. The consensus forecast says nothing at all about the conditional density of y_{t+h}, which is the relevant information needed to answer such questions. Table 1 reports the average forecast from SPF, the bias-adjusted average forecast proposed by Capistrán and Timmermann (2009), the median conditional quantile obtained by evaluating (7) at $\tau = 0.5$, as well as the forecast of an unemployment rate greater than 10% and 9%, respectively. All the point and probability forecasts are for the periods 2010.Q4, 2011.Q1, 2011.Q2, and 2011.Q3. The same median and probability forecasts were also computed using the GARCH(1,1) model and are displayed in the last three columns of Table 1.

There are many interesting results we can derive from Table 1. First, the forecasts obtained from the proposed location-scale model (3) indicate that the U.S. economy will recover very slowly from the current economic crisis. Indeed, columns 6 and 7 of Table 1 show that the probability that the unemployment rate will be above 9% and 10% in the near future is diminishing over time although the probability of a higher than 9% rate is still quite large (44%) at 2011.Q3. Second, the conditional mean of $\widehat{f}_{t+h,t}$ (fourth column) is greater than its conditional median (fifth column), confirming the right skewness of $\widehat{f}_{t+h,t}$ reported in Figure 2, and consequently the presence of an increased risk of a higher unemployment rate in the United States. If we assume that the 10% rate is an indicator of economic recession, then the risk of a double-dip recession in the near future is not negligible, although it is going down to 15% in 2011.Q3 (seventh column of Table 1). Third, since all point forecasts (columns 3–5) are lower than 10% for all out-of-sample periods, one could be misled to conclude that the probability that the unemployment rate will exceed 10% in the near future is close to zero. Finally, columns 8–10 show that the GARCH(1,1) fails to predict the high (although decreasing) probability that the unemployment rate will remain above 9% by the third quarter of 2011, and it also fails to predict the existence of a nonnegligible probability of an above 10% rate. Thus, the benchmark model seems to fail to predict the increased risk of a higher unemployment rate in the near future.

In sum, the point forecasts are just estimates of the location parameter of $\widehat{f}_{t+h,t}$ and therefore cannot say much about either the increased risk of high unemployment rates or the low recovery pace of the U.S. economy. The methodology developed in this paper shows, however, that the consensus (average) forecast, when used as a common factor of all individual forecasts, can be very useful for computing the conditional density of y_{t+h} without imposing any assumption on the parametric form of the conditional distribution. This is useful to recover an important piece of missing information from the data. We suggest that this piece of information can be relevant for the design of macroeconomic policies that aim to avoid the negative effects of an extreme economic crisis.

So far, we have interpreted the previous results without a proper evaluation of the accuracy of our estimates. Our first concern regards the choice of the location-scale model (3). This choice, in turn, affects the specification of the quantile regression (7) that is used to estimate the density $f_{t+h,t}$. If model (3) is misspecified, then the estimated density $\widehat{f}_{t+h,t}$ would not be a correct approximation of the true unknown density $f_{t+h,t}$. Robust inference on quantile regression models can be carried out by using the test developed by Koenker and Xiao (2002). It is, therefore, possible to test the null hypothesis that the true conditional density $f_{t+h,t}$ can be generated by a specific quantile process as, for example, the location-scale model (3). It is also important to recall that conventional procedures for testing density forecasts as the one proposed by Diebold, Gunther, and Tay (1998) and Bai (2003) are not robust against uncertainty caused by both parameter estimation and unknown distribution function. The robustness of the Koenker and Xiao test is based on the simple fact that what can be done for tests based on the parametric empirical process (as the Kolmogorov tests) can also be adapted for tests based on the parametric empirical *quantile* process, which renders robustness against parameter estimation via the Khmaladze transformation (as in the Bai test), and robustness against unknown distribution functions via quantile regression estimation.

Our results have indicated the presence of an upside risk in the U.S. unemployment rate. Thus, we should expect that a model that fails to capture such a risk should be rejected empirically. Table 2 reports the test statistics of the Koenker and Xiao (2002) and Bai (2003) tests applied to the location-scale and benchmark GARCH(1,1) models, respectively.[7] According to Table 2, the null hypothesis of location-scale model (3) cannot be rejected against a more general quantile process, using the entire

7. The Koenker and Xiao (2002) test is equivalent to Bai (2003) test when the distribution function is known.

TABLE 2

KOENKER AND XIAO (2002) AND BAI (2003) TEST STATISTICS

Horizons	Koenker–Xiao	Bai
$h = 1$	0.15	3.01*
$h = 2$	0.76	7.61*
$h = 3$	0.97	3.20*
$h = 4$	0.67	2.62*

NOTE: * indicates rejection of the null at 5% of significance.

sample size from 1969.Q1 to 2010.Q3. This implies that the probability forecasts in Table 1 obtained from model (3) are accurate approximations of the true probability. It also shows that the null hypothesis of correct specification of the benchmark GARCH(1,1) model with Gaussian innovations is rejected at 5% level of significance for all forecasting horizons h, suggesting that the density forecast obtained from such a model does not correctly approximate the true one.

The above results lead us to conclude that the quantile regression-based approach proposed in this paper performs very well in generating a conditional density that is close to the true one $f_{t+h,t}$. Given the fact that our quantile approach is quite general, in the sense that we do not need to assume a parametric form for the distribution function $F_{\eta,h}$, and it is also easy to implement, we believe that the computation of a density forecast no longer needs to be considered a complex econometric exercise.

Now we turn to the out-of-sample forecast exercise. In this case, one rejects a given forecasting model based on its out-of-sample performance. Under the MSE loss function the optimal forecast is the conditional mean, and a simple test of out-of-sample performance could be based on the following linear regression:

$$\begin{aligned} y_{t+h} &= \alpha_0 + \alpha_1 \widehat{y}^i_{t+h,t} + e_{t+h}, \\ E\left[e_{t+h} | \mathcal{F}_t\right] &= 0. \end{aligned} \tag{11}$$

From equation (11) we have that $E[y_{t+h}|\mathcal{F}_t] = \alpha_0 + \alpha_1 \widehat{y}^i_{t+h,t}$ and therefore the null hypothesis of optimality (under the MSE loss) of the model that generated $\widehat{y}^i_{t+h,t}$ can be tested through the coefficient restrictions $H_0\colon \alpha_0 = 0$ and $\alpha_1 = 1$.

When we turn to investigate the out-of-sample performance of the model used to forecast the conditional density, we need to evaluate the performance of this model in predicting each quantile of the conditional density and not just its mean. The fact that we can use a quantile model to estimate $f_{t+h,t}$ allows us to give a value-at-risk model interpretation to (3). Thus, we can implement specification tests available in the risk-management literature to evaluate the (pseudo) out-of-sample forecast performance of our location-scale model. We use the Gaglianone, Lima, Linton, and Smith (GLLS) test to evaluate the out-of-sample performance of the location-scale model because it has the advantage of exhibiting a good finite sample power against a misspecified

TABLE 3
THE GLLS TEST (p -VALUES)

	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.5$	$\tau = 0.6$	$\tau = 0.7$	$\tau = 0.8$
$h = 1$	0.64	0.34	0.44	0.94	0.81	0.99	0.76
	(0.00)	(0.06)	(0.09)	(0.41)	(0.01)	(0.00)	(0.60)
$h = 2$	0.25	0.46	0.78	0.97	0.65	0.98	0.62
	(0.00)	(0.00)	(0.00)	(0.15)	(0.14)	(0.24)	(0.98)
$h = 3$	0.19	0.48	0.72	0.73	0.59	0.25	0.90
	(0.00)	(0.06)	(0.05)	(0.01)	(0.03)	(0.22)	(0.00)
$h = 4$	0.73	0.45	0.56	0.35	0.17	0.45	0.68
	(0.60)	(0.07)	(0.71)	(0.24)	(0.00)	(0.00)	(0.06)

model. To construct our pseudo out-of-sample forecast, we considered a subsample $\{y_{t+h}; C_{t+h,t}\}_{t=1}^{T^*}$, where $T^* = 90$ is initially used to estimate the quantile regression (7). Then we make h-step-ahead forecasts of the τth conditional quantile based on $\mathcal{F}_{T^*}$ by evaluating (7) at $C_{t+h,t} = C_{T^*+h,T^*}$ (the last observation of $C_{t+h,t}$ in the subsample). For a given τ, the forecast quantile is called $VaR(\tau)_{T^*+h,T^*}$, and we keep updating the sample until $T^* = T = 167$ (full sample). For each new observation added to the estimating sample, the quantile regressions (7) are re-estimated and a new set of h-step ahead forecasts are calculated by using the same above-mentioned procedure. At the end, for a given τ, we end up with four time series denoted by $\{VaR(\tau)_{T^*+h,T^*}\}_{T^*=90}^{167}$ and $h = 1, 2, 3, 4$. The GLLS test is implemented through the following quantile regressions using the pseudo out-of-sample observations.

$$Q_{y_{t+h}}(\tau|\mathcal{F}_t) = \beta_0(\tau) + \beta_1(\tau)\, VaR(\tau)_{t+h,t}, \tag{12}$$

$$t = T^* = 90, \ldots, 167 \quad \text{and} \quad h = 1, 2, 3, 4.$$

The null hypothesis of correct specification of the location-scale model at quantile level τ is given by $H_0: \beta_0(\tau) = 0$ and $\beta_1(\tau) = 1$, which implies that $Q_{y_{t+h}}(\tau|\mathcal{F}_t) = VaR(\tau)_{t+h,t}$. Gaglianone et al. (2011) conducts some Monte Carlo simulations and concludes that for a sample size as small as 250 observations, the GLLS test can be oversized for τ as low as 0.01 and 0.05. This implies that our null hypothesis of correct specification of the location-scale model could be overrejected at extreme values of τ and small sample sizes. Since we only have 77 out-of-sample observations, we can avoid this size-distortion problem on the GLLS test by moving away from extreme values of τ. For this reason, we considered $\tau = 0.20, 0.30, 0.40, 0.50, 0.60, 0.70$, and 0.80. Table 3 shows p-values of the GLLS test applied to both models for forecasting horizons $h = 1, 2, 3$, and 4, with the p-values of the GARCH(1,1) model appearing within the parentheses.

The results in Table 3 indicate that for a 5% test, the quantile regression specification (7), which is derived from the location-scale model (3), yields forecasts of conditional quantiles of y_{t+h}, $VaR(\tau)_{t+h,t}$, that are statistically close to the conditional

TABLE 4

COMPARISON BETWEEN QUANTILE REGRESSION PROBABILITY FORECASTS (UNRESTRICTED AND RESTRICTED ESTIMATION)

Period	Median conditional quantile (%)	Prob. (unp > 9%)	Prob. (unp > 10%)	He's (1997) estimation prob. (unp > 9%)	He's (1997) estimation prob. (unp > 10%)
2010.Q4	9.49	84%	25%	83%	24%
2011.Q1	9.23	66%	23%	67%	23%
2011.Q2	9.14	56%	19%	56%	20%
2011.Q3	8.77	44%	15%	44%	17%

NOTE: unp = unemployment.

quantiles of y_{t+h}, $Q_{y_{t+h}}(\tau|\mathcal{F}_t)$ at the out-of-sample period. This guarantees out-of-sample optimality of the location-scale model at the chosen τ's and complements the results from Table 2 regarding the use of the quantile regression-based method proposed in this paper to forecast densities. On the other hand, the GARCH(1,1) model fails to predict almost 50% of the suggested conditional quantiles.

Our last concern regarding the use of quantile regression to produce predictive densities is about the presence of "crossings" of the conditional quantile functions (see Koenker 2005, p. 56). He (1997) points out that the crossing problem occurs more frequently in multiple-variable regressions. In this sense, we should not expect crossing to be an issue in our empirical results because they are based on estimates from single-variable regressions. Nonetheless, we considered a robustness check and reestimated the quantile regressions (7) by using the robust method proposed by He.[8] We forecast the probabilities of the unemployment rate to surpass 9% and 10% by using the He's method and the results are presented in Table 4 along with the probability forecasts displayed in Table 1.

Table 4 shows that the robust restricted quantile regression method proposed by He (1997) produces probability forecasts that are very similar to the ones computed by using the unrestricted quantile regression method of Koenker and Basset (1978). This confirms our analysis in which crossing is not supposed to be a problem in single-variable quantile regressions, which comes out as an additional advantage of the density forecasting method proposed in this paper.

3. CONCLUSION

This paper proposed a semiparametric approach for forecasting the conditional density of a time series process y_{t+h}. It was showed that individual optimal forecasts have a common factor which is represented by the consensus forecast. From a

8. For a location-scale model, Chernozhukov, Fernandez-Val, and Galichon (2010) recognize that the method proposed by He (1997) avoids crossings of the estimated quantile functions.

statistical point of view, this common factor framework allows us to define any conditional quantile of y_{t+h} as an affine function of the average forecast.

The methodology developed in this paper provides a simple and efficient way to estimate the uncertainty behind an economic forecast, and therefore can be useful in identifying the correct economic policy under different circumstances. Perhaps most importantly, our approach is applicable under a wide variety of structures, since it does not depend on the underlying model used to generate the consensus forecast and does not require knowledge of the parametric form of the conditional distribution function. Given this semiparametric approach, we were able to make h-step-ahead forecasts of any quantile of y_{t+h} and, therefore, forecast the entire density.

We illustrated the applicability of this method by analyzing forecasts on the U.S. unemployment rate during the time period that includes the subprime economic crisis. Our results indicate that the U.S. economy is expected to have a slow recovery from the current economic crisis and that the risk of double-dip recession in the near future is not negligible, although it has been going down over time. In order to validate the estimates presented in this paper, we used the test for quantile regression inference developed by Koenker and Xiao (2002). The results suggested that the density forecast obtained by using the proposed quantile approach is statistically close to the true density $f_{t+h,t}$.

Although the proposed methodology has several appealing properties, it should be viewed as complementary to existing approaches rather than competing with them. An empirical comparison in which the quantile approach proposed in this paper is implemented with survey factors as well as estimated factors from observable economic variables would be very interesting. In the latter, we would have to take a stand on which particular variables agents use to compute their forecasts, a decision that, in practice, would lead us to consider a dynamic factor model with PCA plus a subsidiary model to compute iterated forecasts. Although more complicated than the approach based on survey factors, the structural approach has the advantage of being general enough to allow counterfactual exercises. The first step toward this type of analysis has recently been taken by De Nicolò and Lucchetta (2010), who propose new measures of systemic risk by estimating quantile models with estimates of factors as conditioning variables. Furthermore, several important topics remain open for future research, such as: (i) the linear representation of the DGP is something that could be relaxed in future research, (ii) nonlinear dynamics across quantiles should also be explored, and (iii) forecasts of counterfactual distributions using data from experimental economics (lab or field) in which the effect of economics incentives from forecast users on the behavior of economic forecasters can be investigated.

APPENDIX: PROOF OF PROPOSITION 1

The proof is similar to the one shown by Granger (1969), Christoffersen and Diebold (1997), and Patton and Timmermann (2007) in the first part of their

Proposition 2. Thus, by homogeneity of the loss function and DGP (3) we have that:

$$
\begin{aligned}
\widehat{y}^{i}_{t+h,t} &= \arg\min_{\widehat{y}} \int L^{i}(y-\widehat{y})dF_{t+h,t}(y) \\
&\arg\min_{\widehat{y}} \int \left[g\left(\frac{1}{X'_{t+h,t}\gamma}\right)\right]^{-1} L^{i}\left(\frac{1}{X'_{t+h,t}\gamma}(y-\widehat{y})\right) dF_{t+h,t}(y) \\
&= \arg\min_{\widehat{y}} \int \left[g\left(\frac{1}{(\gamma_0+\gamma_1 C_{t+h,t})}\right)\right]^{-1} \\
&\quad \times L^{i}\left(\frac{1}{(\gamma_0+\gamma_1 C_{t+h,t})}(y-\widehat{y})\right) dF_{t+h,t}(y) \\
&= \arg\min_{\widehat{y}} \int L^{i}\left(\frac{1}{(\gamma_0+\gamma_1 C_{t+h,t})}(y-\widehat{y})\right) dF_{t+h,t}(y) \\
&= \arg\min_{\widehat{y}} \int L^{i}\Bigg(\frac{1}{(\gamma_0+\gamma_1 C_{t+h,t})}(\alpha_0+\alpha_1 C_{t+h,t}+\gamma_0\eta_{t+h} \\
&\qquad + \gamma_1 C_{t+h,t}\eta_{t+h} - \widehat{y})\Bigg) dF_{\eta}, h(\eta).
\end{aligned}
$$

Let us represent a forecast by $\alpha_0+\alpha_1 C_{t+h,t}+(\gamma_0+\gamma_1 C_{t+h,t})\widehat{\gamma}$. This way, it follows that:

$$
\begin{aligned}
\widehat{y}^{i}_{t+h,t} &= \alpha_0+\alpha_1 C_{t+h,t}+(\gamma_0+\gamma_1 C_{t+h,t}) \\
&\quad \times \arg\min_{\widehat{\gamma}} \int L^{i}\Bigg(\frac{1}{(\gamma_0+\gamma_1 C_{t+h,t})}(\alpha_0+\alpha_1 C_{t+h,t} \\
&\quad + (\gamma_0+\gamma_1 C_{t+h,t})\eta_{t+h} - \alpha_0 - \alpha_1 C_{t+h,t} \\
&\quad - (\gamma_0+\gamma_1 C_{t+h,t})\widehat{\gamma})\Bigg) dF_{\eta,h}(\eta) \\
&= \alpha_0+\alpha_1 C_{t+h,t}+(\gamma_0+\gamma_1 C_{t+h,t}) \\
&\quad \times \arg\min_{\widehat{\gamma}} \int L^{i}(\eta_{t+h}-\widehat{\gamma}) dF_{\eta,h}(\eta) \\
&= \alpha_0+\gamma_0\gamma^{i}_{h}+\alpha_1 C_{t+h,t}+\gamma_1 C_{t+h,t}\gamma^{i}_{h} \\
&= \alpha_0(\tau_i)+\alpha_1(\tau_i) C_{t+h,t}, \\
&\quad \text{where } \alpha_0(\tau_i)=(\alpha_0+\gamma_0\gamma^{i}_{h}) \text{ and } \alpha_1(\tau_i)=(\alpha_1+\gamma_1\gamma^{i}_{h}),
\end{aligned}
$$

in which we have used the fact that $F_{\eta,h}(\eta)$ is time invariant by definition. □

LITERATURE CITED

Bai, Jushan. (2003) “Testing Parametric Conditional Distributions of Dynamic Models.” *Review of Economics and Statistics*, 85, 531–49.

Bates, John M., and Clive W.J. Granger. (1969) “The Combination of Forecasts.” *Operational Research Quarterly*, 20, 451–68.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz. (2005) “Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach.” *Quarterly Journal of Economics*, 120, 387–422.

Capistrán, Carlos, and Allan Timmermann. (2009) “Forecast Combination with Entry and Exit of Experts.” *Journal of Business and Economic Statistics*, 27, 428–40.

Chernozhukov, Victor, Ivan Fernandez-Val, and Alfred Galichon. (2010) “Quantile and Probability Curves without Crossing.” *Econometrica*, 78, 1093–1125.

Christoffersen, Peter F., and Francis X. Diebold. (1997) “Optimal Prediction under Asymmetric Loss.” *Econometric Theory*, 13, 808–17.

Clements, Michael P. (2004) “Evaluating the Bank of England Density Forecasts of Inflation.” *Economic Journal*, 114, 844–66.

De Nicolò, Gianni, and Marcella Lucchetta. (2010) “Systemic Real and Financial Risks: Measurement, Forecasting, and Stress Testing.” SSRN Working Paper No. 1763852.

Diebold, Francis X., S. Borağan Aruoba, and Chiara Scotti. (2009) “Real-Time Measurement of Business Conditions.” *Journal of Business and Economic Statistics*, 27, 417–27.

Diebold, Francis X., Todd A. Gunther, and Anthony S. Tay. (1998) “Evaluating Density Forecasts, with Applications to Financial Risk Management.” *International Economic Review*, 39, 863–83.

Durbin, James. (1973) “Weak Convergence of Sample Distribution Functions When Parameters Are Estimated.” *Annals of Statistics*, 1, 279–90.

Figlewski, Stephen. (1983) “Optimal Price Forecasting Using Survey Data.” *Review of Economics and Statistics*, 65, 813–36.

Figlewski, Stephen, and Thomas Urich. (1983) “Optimal Aggregation of Money Supply Forecasts: Accuracy, Profitability and Market Efficiency.” *Journal of Finance*, 28, 695–710.

Gaglianone, Wagner P., Luiz Renato Lima, Oliver Linton, and Daniel R. Smith. (2011) “Evaluating Value-at-Risk Models via Quantile Regressions.”*Journal of Business and Economic Statistics*, 29, 150–60.

Granger, Clive W.J. (1969) “Prediction with a Generalized Cost of Error Function.” *Operational Research Quarterly*, 20, 199–207.

Granger, Clive W.J., and Paul Newbold. (1986) *Forecasting Economic Time Series*, 2nd ed. San Diego, CA: Academic Press.

He, Xuming. (1997) “Quantile Curves without Crossing.” *American Statistician*, 51, 186–92.

Khmaladze, Estate V. (1981) “Martingale Approach in the Theory of Goodness-of-Fit Tests.” *Theory of Probability and Its Applications*, 26, 240–57.

Koenker, Roger. (2005) *Quantile Regression*, 1st ed. Cambridge, UK: Cambridge University Press.

Koenker, Roger, and Gilbert Bassett, Jr. (1978) “Regression Quantiles.” *Econometrica*, 46, 33–50.

Koenker, Roger, and Zhijie Xiao. (2002) “Inference on the Quantile Regression Process.” *Econometrica*, 70, 1583–612.

Patton, Andrew J., and Allan Timmermann. (2007) “Testing Forecast Optimality under Unknown Loss.” *Journal of the American Statistical Association*, 102, 1172–84.

Patton, Andrew J., and Allan Timmermann. (2010) “Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion.” *Journal of Monetary Economics*, 57, 803–20.

Stock, James H., and Mark W. Watson. (2006) “Forecasting with Many Predictors.” In *Handbook of Economic Forecasting*, Vol. 1, edited by G. Elliott, C. Granger, and A. Timmermann, pp. 515–50. Amsterdam: Elsevier.

Weiss, Andrew A. (1996) “Estimating Time Series Models Using the Relevant Cost Function.” *Journal of Applied Econometrics*, 11, 539–60.