Journal of Banking & Finance 40 (2014) 321–329

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

The empirical similarity approach for volatility prediction

Vasyl Golosnoy [a], Alain Hamid [b], Yarema Okhrin [b,*]

[a] Department of Statistics and Econometrics, Ruhr-Universität Bochum, Universitaetstr. 150, D-44801 Bochum, Germany
[b] Department of Statistics, University of Augsburg, Universitaetstr. 16, D-86159 Augsburg, Germany

ARTICLE INFO

Article history:
Received 15 July 2013
Accepted 5 December 2013
Available online 14 December 2013

JEL classification:
G17
C53

Keywords:
Case based decisions
Empirical similarity
Forecasting combinations
Volatility forecasts

ABSTRACT

In this paper we adapt the empirical similarity (ES) concept for the purpose of combining volatility forecasts originating from different models. Our ES approach is suitable for situations where a decision maker refrains from evaluating success probabilities of forecasting models but prefers to think by analogy. It allows to determine weights of the forecasting combination by quantifying distances between model predictions and corresponding realizations of the process of interest as they are perceived by decision makers. The proposed ES approach is applied for combining models in order to forecast daily volatility of the major stock market indices.

1. Introduction

Consider the task to forecast a process of interest. There are often several competing models available for this purpose with their strong and weak sides. How should forecasts from these different models be combined? Starting from the seminal contribution of Bates and Granger (1969) there has been suggested a large number of approaches to determine weights for model combination. These weights are usually related to the model prediction success probabilities (cf. Elliott and Timmermann, 2004) which can be interpreted as occurrence probabilities for the states in the coming period. The ability to evaluate probabilities is crucial for the classical decision theory in spirit of von Neumann–Morgenstern. However, such probabilistic approach is not always possible or desired.

By making decisions in situations under uncertainty or ignorance a decision maker can be unable or unwilling to evaluate probabilities but prefers to rely on *thinking by analogy* for learning from the past about the future. The analogical (case based) reasoning is widely applied for decision making in medicine, law, business, politics, or artificial intelligence (cf. Gilboa and Schmeidler, 2001). The case based decision theory presumes analogous thinking of human beings in cases where the current situation is evaluated by considering its similarity to previously experienced (past) situations (cf. Gilboa and Schmeidler, 2001). Cases which are more similar to the current situations obtain larger weights compared to those which are less similar. The concept of *empirical similarity* (Gilboa et al., 2006; Gilboa et al., 2011) provides the econometric framework for estimation of the similarity function from the data (Gilboa and Schmeidler, 2012). It allows to measure distances between cases (problems, situations) as they are perceived by decision makers.

In this paper we suggest and apply a methodology how to use the empirical similarity (ES) concept in order to combine forecasts from different models in a non-probabilistic manner. In our setting alternative forecasts originating from competing models could be evaluated as cases, which are to some extent similar to the currently observed state or realization. A model which recently provides more precise point forecasts should obtain a larger current weight compared to alternatives. The core idea of our approach is to measure the empirical similarity distance between the current observation and the last one-period-ahead forecasts from different models. This similarity distance determines model weights for the next period forecasts. Thus, our approach exploits the information about the recent performance of different models in order to determine the weights of the forecasting model combination. The advantages of such ES combination approach compared to the probabilistic alternatives are that (i) it does not require knowledge of model success probabilities; (ii) it relates the weights of the forecasting models to the preferences of economic agents; and (iii) it reveals from the data how decision makers evaluate the similarity between forecasts and realizations.

* Corresponding author. Tel.: +49 821 598 4152; fax: +49 821 598 4227.
E-mail addresses: vasyl.golosnoy@rub.de (V. Golosnoy), alain.hamid@wiwi.uni-augsburg.de (A. Hamid), yarema.okhrin@wiwi.uni-augsburg.de (Y. Okhrin).

0378-4266/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2013.12.009

We illustrate the application of the proposed ES forecast combination approach by modeling the daily process of realized volatilities. For this purpose we evaluate empirical similarities for combining volatility models which can be treated as approaches reflecting different investment horizons (cf. Ghysels et al., 2006; Corsi, 2009). In particular, we model series of daily realized volatilities of the leading world financial indices for about 13 recent years characterized by both high and low volatility periods. The parameters of the ES approach are estimated with the maximum likelihood methodology (cf. Lieberman, 2010) for both full sample and mowing windows of 250 daily observations. We compare forecasting performance of the ES approach with a set of popular volatility models by conducting both in-sample and out-of-sample predictions. The obtained estimation results reveal how forecasts from various volatility models are aggregated via the empirical similarities in the perception of decision makers. A special attention is drawn to the analysis of volatility patterns during and immediately after the recent subprime crisis with a highly complex volatility dynamics. The proposed empirical similarity model appears to provide the most suitable description of the volatility process during that period.

The rest of the paper is organized as follows. In Section 2 we propose a novel empirical similarity approach which allows to combine forecasts from different models. The ES methodology for combining volatility forecasts or components is presented in Section 3. The empirical study in Section 4 is devoted to the estimation and forecast comparison of competing volatility models. Moreover, we draw a special attention to the recent subprime crisis period which is characterized by a highly nonlinear volatility dynamics. Section 5 concludes the paper.

2. Empirical similarity for model combination

Assume that there are p models (forecasts, recommendations) which could be combined in order to forecast the variable of interest y_{t+1}. Define a finite set of distinct forecasts from different models as $\{x_{1,t},\dots x_{p,t}\}$ and consider the task of combining them in a parsimonious manner. A family of linear forecast combinations remains popular starting from the seminal paper of Bates and Granger (1969). A linear forecast combination is given as

$$\hat{y}_{t+1} = \sum_{i=1}^{p} a_{i,t} x_{i,t}, \quad (1)$$

where non-negative $a_{i,t}$s are the proportions of the ith model with $\sum_{i=1}^{p} a_{i,t} \equiv 1$. There is a straightforward probabilistic interpretation for the weights $a_{i,t}$, which are in general related to model success probabilities (cf. Elliott and Timmermann, 2004). Weighting models as in (1) presumes *the ability* to choose the weights $a_{i,t}$ appropriately by considering some given objective functions. Various probabilistic approaches are proposed for the choice of the proportions $a_{i,t}$, however, there is no dominating methodology up to now.

Now let us consider situations under uncertainty or ignorance where economic agents do not have specific (probabilistic) beliefs about model weights in future but simply prefer models which performed well in similar cases in the past. In these situations the agents should form their decisions relying on analogical case based reasoning (Gilboa and Schmeidler, 2001). The case based decision theory (cf. Gilboa and Schmeidler, 2001) is developed for situations where decision makers refrain from evaluating probabilities but relies on their experience in order to evaluate distances (similarities) between past cases (situations) and the current state of nature.

The empirical similarity (ES) approach of Gilboa et al. (2006) provides the econometric framework for estimation of the similarity functions from the data. In order to describe their concept assume that there is a vector of variables $\mathbf{z}_t$ characterizing the current situation, which is followed by the realization y_{t+1} in the next period. The ES postulates that the model combination weights $a_{i,t}$ should be replaced by non-negative similarity-based frequencies $\phi[\mathbf{z}_s, \mathbf{z}_t]$, which sum up to unity and serve as weights for the experienced realizations y_{s+1}. In this setting the DGP is driven directly by its historical observations weighted by $\phi[\mathbf{z}_s, \mathbf{z}_t]$'s. Then the corresponding ES model equation is given as

$$y_{t+1} = \sum_{s<t} \phi[\mathbf{z}_s, \mathbf{z}_t] y_{s+1} + \varepsilon_{t+1}, \quad \varepsilon_t \sim (0, \sigma^2), \quad (2)$$

where $\mathbf{z}_s$ is a vector characterizing the situation at time s, y_{s+1} is the realization of the process of interest experienced in the next period. Thus, the similarity function measures the distance between the vectors $\mathbf{z}_t$ and $\mathbf{z}_s$ as it is assessed by a decision maker.

Relying on the ES concept of Gilboa et al. (2006), we suggest an ES approach for combining forecasting models. For this purpose we unite the ideas behind the forecasting Eq. (1) and the ES model in (2). The resulting ES forecast combination is given as

$$y_{t+1} = \sum_{i=1}^{p} \phi[y_t, x_{i,t-1}] x_{i,t} + \varepsilon_{t+1}, \quad \varepsilon_t \sim (0, \sigma^2). \quad (3)$$

The essential difference to Eq. (2) is that we replace the vector of characteristics $\mathbf{z}_s$ by the forecast from the ith model $x_{i,t-1}$, so that we now measure the distance between the previous forecast $x_{i,t-1}$ and the corresponding realization y_t in order to obtain the weights $\phi[y_t, x_{i,t-1}]$. Then the forecast combination which is a weighted sum of the forecasts $\{x_{1,t},\dots,x_{p,t}\}$ is given as

$$\hat{y}_{t+1} = \sum_{i=1}^{p} \phi[y_t, x_{i,t-1}] x_{i,t}.$$

In our ES combination setting the process of interest y_{t+1} is driven directly by the alternative forecasts $x_{i,t}$s, allowing to interpret (3) as a proxy for the true DGP as it is perceived by decision makers.

The model in (3) incorporates nonlinear *autoregressive* features due to the fact that y_t enters the similarity function $\phi[\cdot,\cdot]$ which determines the DGP of y_{t+1}. Moreover, it has a *spatial* property by measuring distances between the forecasts and the realization, which are used for weighting $x_{i,t}$s in order to assess y_{t+1}. This point corresponds to the suggestion of Gilboa et al. (2006, pp. 437–438) that for time series the current observation could be compared not with a history but with a profile (cross-section) of components.

The weights $\phi[\cdot,\cdot]$ depend on the previous experience of decision makers. The distance between the proxy of the current realization and the ith model forecast is measured in our case as

$$\phi[y_t, x_{i,t-1}] = \frac{\theta[y_t, x_{i,t-1}]}{\sum_{j=1}^{p} \theta[y_t, x_{j,t-1}]}. \quad (4)$$

The weights $\phi[y_t, x_{i,t-1}] \in [0,1]$ can be interpreted as normalized relative empirical similarities with the property $\sum_{i=1}^{p} \phi[y_t, x_{i,t-1}] \equiv 1$, whereas $\theta[y_t, x_{i,t-1}]$ is the similarity (distance) function parameterized below. The interpretation of the similarity measures $\theta[y_t, x_{i,t-1}]$ is straightforward, namely a small distance between y_t and $x_{i,t-1}$ implies a high similarity value of $\theta[y_t, x_{i,t-1}]$, while a large distance indicates on low similarity.

There are several possibilities to specify the similarity function $\theta[y_t, x_{i,t-1}] \geqslant 0$ (cf. Golosnoy and Okhrin, 2008; Guerdjikova, 2008; Lieberman, 2010). In this paper we exploit a flexible specification of the exponential similarity function of Billot et al. (2008), which is given as

$$\theta[y_t, x_{i,t-1}] = \exp\left(-\omega_i (y_t - x_{i,t-1})^2\right), \quad \text{with} \quad \omega_i \in \mathbb{R}. \quad (5)$$