International Journal of Forecasting 32 (2016) 398–410

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Low and high prices can improve volatility forecasts during periods of turmoil

Piotr Fiszeder*, Grzegorz Perczak

Nicolaus Copernicus University, Faculty of Economic Sciences and Management, Department of Econometrics and Statistics, ul. Gagarina 13a, 87-100 Toruń, Poland

ARTICLE INFO

Keywords:
GARCH model
Low and high prices
NIG distribution
Turmoil period
Volatility forecasting

ABSTRACT

In this study, we describe a modification of the GARCH model that we have formulated, where its parameters are estimated based on closing prices as well as on information related to daily minimum and maximum prices. In an empirical application, we show that the use of low and high prices in the derivation of the likelihood function of the GARCH model improved the volatility estimation and increased the accuracy of volatility forecasts based on this model during the period of turmoil, relative to using closing prices only. This analysis was performed for two stock indices from developed markets, i.e., S&P 500 and FTSE 100, and for two stock indices from emerging markets, i.e., the Polish WIG20 index and the Greek Athex Composite Share Price Index. The main result obtained in this study is robust to both the forecast evaluation criterion applied and the proxy used for the daily volatility.

1. Introduction

The modelling and forecasting of the volatility of asset returns is a key issue in many financial and economic applications. One of the most popular volatility models is the GARCH model, and the estimation of its parameters is based solely on the daily closing prices in the majority of cases. However, a single return gives a weak signal for the current level of volatility. This implies that GARCH models are poorly suited to situations where the volatility changes suddenly to a new level. For instance, when the volatility increases sharply on day t and subsequent days, the conditional variance of the GARCH model will not change on day t and will increase only gradually on the subsequent days. Thus, the conditional variance will take many periods to reach a new level of volatility (e.g. Andersen, Bollerslev, Diebold, & Labys, 2003; Hansen, Huang, & Shek, 2012). However, while the commonly available databases do contain the daily closing prices, they also include daily low and high prices, which can be used successfully for volatility estimation. The use of low and high prices is one area in which extensive research, both theoretical and empirical, is currently being conducted (see the review by Chou, Chou, & Liu, 2010). This renewed interest within the scientific community is mainly because the application of such data yields more accurate estimates and forecasts of volatility than those based only on closing prices (e.g., Chou, 2005; Li & Hong, 2011).

The research concerning the use of data on low and high prices can be divided into three main groups (we deliberately omit the use of so-called intraday or high frequency data, and focus on daily data). The first group consists of the so-called price range estimators, which include the best-known estimators of Garman and Klass (1980), Parkinson (1980), Rogers and Satchell (1991) and Yang and Zhang (2000). Financial market practitioners commonly use range estimators to estimate the volatility

* Corresponding author.
E-mail addresses: piotr.fiszeder@umk.pl (P. Fiszeder), grzegorz.perczak@gmail.com (G. Perczak).

http://dx.doi.org/10.1016/j.ijforecast.2015.07.003

because they are significantly more efficient than the estimator calculated as the daily squared return of closing prices. However, these estimators are less popular among scientists because they neglect the temporal dependence of returns (such as conditional heteroscedasticity). The second group are the so-called range-based volatility models, which are widely used for closing prices or their modifications, although they are applied directly to the modelling of the price range (e.g., see Alizadeh, Brandt, & Diebold, 2002; Brandt & Jones, 2006; Chou, 2005; Engle & Gallo, 2006). The third group are models of interval-valued data (e.g., see Arroyo, Espínola, & Maté, 2011; Arroyo, González-Rivera, & Maté, 2010; Maia & de Carvalho, 2011; Maia, de Carvalho, & Ludermir, 2008). High-low intervals are linked naturally to the concept of volatility. This study does not discuss the latter two uses of low and high prices.

Only a few studies have used low and high prices directly for formulating the estimation procedure in existing and known volatility models. These include the GARCH models of Lildholdt (2002) and Venter, De Jongh, and Griebenow (2005), who derived likelihood functions based on low, high and closing prices. The present study makes two main contributions. The first is the presentation of a modification of the GARCH model, where the parameters are estimated based on low, high and closing prices (for details and less complex parameterizations, see Perczak & Fiszeder, 2014). Lildholdt (2002) assumed that, over each day, a new incremental log-price process follows an arithmetic Brownian motion with a constant volatility for that day, and therefore applied the GARCH model with a normal conditional innovation distribution. However, it is well known that the normal distribution is often too light-tailed to be an appropriate distribution for most financial time series. Therefore, similarly to Venter et al. (2005), we assume a normal-inverse Gaussian (NIG) conditional innovation distribution for the GARCH model. Furthermore, our formulation of the model differs in two respects. First, we apply the significantly more efficient range estimator of the variance, instead of the estimator calculated as the daily squared return of closing prices, as is commonly used in the standard GARCH model. Second, we assume slight simplifications where different parameterizations of random variables and stochastic processes are applied.

The study's second main contribution is to show that the use of additional information related to low and high prices in the derivation of the likelihood function of the GARCH model can improve the volatility estimation and increase the accuracy of volatility forecasts based on a model for periods of turmoil, compared with only applying closing prices. The idea of periods of turmoil refers to periods with large declines in stock prices and very high levels of volatility. To the best of our knowledge, this is the first attempt in the literature to demonstrate the superiority of this approach for forecasting. This issue is important from a practical viewpoint, because low and high prices are almost always commonly available with closing prices for financial series. Therefore, it can be stated that the omission of such data leads to the loss of important information.

It is well known that the extreme values that are associated with turbulent and crisis periods have a significant influence on the estimation results. One of the main weaknesses of the GARCH process where the parameters are estimated based on closing prices is a slow response to abrupt changes in the market (e.g. Andersen et al., 2003; Hansen et al., 2012). The use of low and high prices in the estimation of the parameters should reduce the impact of this negative effect significantly.

The remainder of this paper is organized as follows. Section 2 provides definitions of the distributions and processes employed in this study. Section 3 describes the parameterization of the GARCH model, where the parameters are estimated based on low, high and closing prices. In Section 4, this approach is then used to model the volatility of two well-known stock indices from developed markets, S&P 500 and FTSE 100, as well as the Polish stock index WIG20. Section 5 verifies the forecasting performance both for the usual period and for the period of turmoil due to the financial crisis in the USA. In Section 6, we perform a robustness check for additional proxies for volatility and for a different period of turmoil, namely the Greek debt crisis, using the Athex Composite Share Price Index. Section 7 provides our conclusions.

2. Definitions of the distributions and processes employed in this study

Let $S_{t,\tau}$ be the price of a financial instrument observed on day t ($t \in N, 0 < t$) after time τ ($0 \leq \tau \leq 1$) from the last quotation the day before. Thus, there is the identity $S_{t-1,1} = S_{t,0}$. The daily (24-hour) minimum and maximum prices are defined as $L_t = \min_{0\leq\tau\leq 1} S_{t,\tau}$ and $H_t = \max_{0\leq\tau\leq 1} S_{t,\tau}$, respectively. In addition, we employ the following definitions of daily low, high and closing returns: $A_t = \ln\left(L_t/S_{t,0}\right)$, $C_t = \ln\left(H_t/S_{t,0}\right)$, and $X_t = \ln\left(S_{t,1}/S_{t,0}\right)$.

In practice, only four values of quotations during the day are usually available for each day t (the acquisition of intraday data is usually an added cost, and such data are not available for all assets): today's open price O_t, today's observed low price L_t^*, today's observed high price H_t^*, and today's closing price S_t. If today's open price O_t is different from yesterday's closing price S_{t-1} (the so-called night returns are nonzero), then the variables A_t and C_t can be redefined as: $A_t = \ln\left(\min\left(S_{t-1}, L_t^*\right)/S_{t-1}\right)$, $C_t = \ln\left(\max\left(S_{t-1}, H_t^*\right)/S_{t-1}\right)$ (see Fiszeder & Perczak, 2013).

For a standard Wiener process $\mathcal{B}_\tau$, $\tau \geq 0$, the Brownian motion $X_\tau^{\mathcal{B}} = \mu\tau + \sigma\mathcal{B}_\tau$ is defined. Let $s \in \mathbb{R}_+$ be a fixed value, $A_s^{\mathcal{B}} = \min_{0\leq\tau\leq s} X_\tau^{\mathcal{B}}$ and $C_s^{\mathcal{B}} = \max_{0\leq\tau\leq s} X_\tau^{\mathcal{B}}$. The probability density function of $X_s^{\mathcal{B}}$ with upper and lower absorbing barriers equal to c and a, respectively, is given by the formula (see Cox & Miller, 1965, p. 222, equation 78):

$$f_{X_s^{\mathcal{B}}}\left(a, c, x; \mu s, \sigma^2 s\right) = \frac{1}{dx} P\left(A_s^{\mathcal{B}} > a, C_s^{\mathcal{B}} \leq c, X_s^{\mathcal{B}} \in dx\right)$$
$$= \frac{1}{\sqrt{2\pi}\sigma\sqrt{s}} e^{\frac{2\mu x - \mu^2 s}{2\sigma^2}} \times \sum_{k=-\infty}^{\infty}\left(e^{-\frac{(x-2k(c-a))^2}{2\sigma^2 s}} - e^{-\frac{(x-2c-2k(c-a))^2}{2\sigma^2 s}}\right), \quad (1)$$

where $a \leq 0 \leq c$, $a \leq x \leq c$.