International Journal of Forecasting 29 (2013) 592–604

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Overnight stock returns and realized volatility

Katja Ahoniemi [a,*], Markku Lanne [b]

[a] Imperial College Business School, Imperial College London, United Kingdom
[b] University of Helsinki, Finland

ARTICLE INFO

Keywords:
Non-parametric volatility estimation
Ranking of volatility estimators

ABSTRACT

The information flow in modern financial markets is continuous, but major stock exchanges are open for trading for only a limited number of hours. No consensus has yet emerged on how to deal with overnight returns when calculating and forecasting realized volatility in markets where trading does not take place 24 hours a day. Based on a recently introduced formal testing procedure, we find that for the S&P 500 index, a realized volatility estimator that optimally incorporates overnight information is more accurate in-sample. In contrast, estimators that do not incorporate overnight information are more accurate for individual stocks. We also show that accounting for overnight returns may affect the conclusions drawn in an out-of-sample horserace of forecasting models. Finally, there is considerably less variation in the selection of the best out-of-sample forecasting model when only the most accurate in-sample RV estimators are considered.

1. Introduction

Information that is relevant for investors accumulates around the clock, but stock exchanges tend to be open only for a limited number of hours. The overnight period is becoming more and more important due to the integration of global financial markets, and many news releases are also timed to occur during non-trading hours. During market closures, investors are not able to garner information about asset prices by observing executed trades. Once trading commences at the start of a new day, the information that has arrived during the non-trading hours is reflected in the new day's prices.[1]

The existing literature highlights both the importance of non-trading hours and the fact that prices evolve in different ways during trading and non-trading hours. An early contribution to this literature is that of Oldfield and Rogalski (1980), who assume different data-generating processes for trading and non-trading hours. French and Roll (1986) and Stoll and Whaley (1990a) document that returns over trading hours are more volatile than non-trading hour returns. Hong and Wang (2000) model how market closures affect investors' trading policies and the return-generating process. Focusing on realized volatility, Martens (2002) models the dynamics of returns during trading and non-trading hours differently. Modeling both US and European stocks markets, Tsiakas (2008) documents that the information accumulated overnight contains substantial predictive ability.[2]

The use of high-frequency data to calculate and sum intraday squared returns has become the prevalent method for estimating volatility in recent years. The early literature on realized volatility (RV) dealt with foreign exchange

* Correspondence to: Imperial College Business School, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom.
E-mail address: k.ahoniemi@imperial.ac.uk (K. Ahoniemi).

[1] The emergence of electronic communication networks is gradually shifting equity markets towards around-the-clock trading. However, equity trading outside of established post-close and pre-open sessions is still minimal (Chen, Yu, & Zivot, 2012). The futures market is also not continuous: trading in the S&P 500 futures, for example, is halted during weekends.

[2] Boes, Drost, and Werker (2007) treat overnight returns as a jump, and investigate how these jumps affect option prices.

0169-2070/$ – see front matter
http://dx.doi.org/10.1016/j.ijforecast.2013.03.006

markets, where trading takes place around the clock (see e.g. Andersen & Bollerslev, 1998). However, the same approach of summing intraday squared returns has since been applied to data from other markets that are closed for at least a part of each 24-h period. The best way to incorporate the information that arrives during the times of market closure into a realized volatility estimator is not obvious at the outset.

The existing literature on stock market realized volatility has adopted several approaches to dealing with the time period when the market is closed (in other words, the overnight period). The simplest approach is to ignore the overnight period, in other words, summing only the intraday squared returns (Andersen, Bollerslev, Diebold, & Ebens, 2001; Corsi, Mittnik, Pigorsch, & Pigorsch, 2008; Thomakos & Wang, 2003; Wu, 2011). However, Hansen and Lunde (2006) argue that such an estimator is not a proper proxy of the true volatility because it does not span a full 24-hour period. Another solution in the literature is to calculate the overnight return by subtracting each day's close value from the next day's open, and to add this squared return as one equally-weighted factor in the sum of intraday returns (Becker, Clements, & White, 2007; Blair, Poon, & Taylor, 2001; Bollerslev, Tauchen, & Zhou, 2009; de Pooter, Martens, & van Dijk, 2008; Martens, 2002). A third method is to calculate realized volatility by ignoring the overnight period, but then scaling the resulting value upward so that the volatility estimate covers an entire 24-h day (Angelidis & Degiannakis, 2008; Koopman, Jungbacker, & Hol, 2005; Martens, 2002). Fourth, Hansen and Lunde (2005b) have derived a weighting scheme for the overnight return and the sum of intraday returns. This method is used by e.g. Fleming and Kirby (2011) and Fuertes and Olmo (2013).[3]

In the absence of a consensus approach, this paper comprehensively compares the existing solutions to dealing with returns from the overnight period. In our empirical analysis, we consider the returns on the S&P 500 index as well as on the thirty equities included in the Dow Jones Industrial Average. As the true volatility is unobservable, we cannot compare the various volatility measures directly. However, such a comparison is facilitated by the procedure recently introduced by Patton (2011b), which allows for the ranking of volatility estimators even without observing the true volatility. For S&P 500 returns, this test procedure selects the Hansen and Lunde (2005b, henceforth HL) realized volatility estimator over the other alternatives when using a mean squared error type loss function. For individual stocks, the result is different: estimators that utilize only intraday data are most accurate.

In addition to in-sample comparisons, we also investigate the role of overnight information in volatility forecasting. In particular, our results highlight the significance of choosing how to treat overnight returns when determining the RV estimator to which different out-of-sample volatility forecasts are compared. Using two in-sample periods of different lengths, and two out-of-sample periods with and without the recent financial crisis, we find that the selection of the best forecasting model can depend on the treatment of overnight returns. Therefore, in typical volatility forecasting exercises where one RV estimator is used as the proxy for the latent volatility in loss functions, the way in which overnight information is included in that proxy can affect the outcome of the horserace between the forecasting models. This outcome is much more stable, however, if overnight information is included or excluded in accordance with the recommendations of the Patton (2011b) testing procedure. In fact, when choosing the forecasting model based on the most accurate estimators only, the number of return series for which model selection is unaffected by the treatment of overnight returns rises from 10 to 22 out of 31 time series under analysis. Therefore, we recommend that researchers employ a two-step approach of first evaluating in-sample accuracy then basing the forecast model selection on only the most accurate RV estimators.

Defining the overnight return on an index, such as the S&P 500 index we consider, is particularly complicated, and thus warrants additional attention. The published opening quote of the index tends to be, if not equal, very close to the previous day's close due to the fact that trading in all 500 constituents does not commence immediately at 9:30 AM. This poses an additional challenge in determining the overnight returns for the index, as a stale index quote will not fully reflect the information that accumulated during the non-trading hours. To that end, we propose two alternative solutions: using the return from the previous close up until 9:35 AM (a five-minute return), and using the difference between the previous close and the so-called special opening quote (SOQ), which is calculated from the opening prices of each of the 500 constituents. Empirically, both overnight return proxies fare equally well.

The rest of the paper is organized as follows. Section 2 describes all of the RV measures that are compared in the paper. Section 3 describes the data sets, and details the two proposed overnight return proxies for the S&P 500 index. Section 4 provides details on the Patton (2011b) testing framework and shows which RV estimators it deems most accurate in-sample. Section 5 contains the empirical forecast analysis. Section 6 concludes.

2. Realized volatility estimators

This section outlines the various competing realized volatility estimators that are compared in the later sections. In what follows, the term realized volatility refers to both the realized variance and its square root. The basic measure of RV is defined in the standard way:

$$RV_t = \sum_{i=1}^{m} (p_{t,i} - p_{t,i-1})^2, \quad (1)$$

where p is the log price of an asset and m denotes the number of intraday returns to be summed. For example, in the US stock market, there are 78 five-minute returns in one trading day. The squared overnight return ON_t can be added to this simple intraday measure as a 79th factor in

[3] There are also numerous studies with realized volatility applications that do not mention how the overnight return is treated, if accounted for at all. These include those of Engle and Gallo (2006) and Giot and Laurent (2007).