Economic Modelling 36 (2014) 332–340

Contents lists available at ScienceDirect

Economic Modelling

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

The impact of trading volume, number of trades and overnight returns on forecasting the daily realized range

Neda Todorova [a,*], Michael Souček [b]

[a] Griffith Business School, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia

[b] Department of Business Administration, European University Viadrina, Große Scharrnstraße 59, D-15230 Frankfurt (Oder), Germany

ARTICLE INFO

Article history:
Accepted 1 October 2013

JEL classification:
C13
C14
C22
C53
G17

Keywords:
Volatility forecasting
HAR model
Realized range
Overnight returns
Trading volume
Number of trades

ABSTRACT

Intraday data of 26 German stocks are used to investigate whether the information contained in trading volume and number of trades as well as in various specifications of overnight returns can improve one-step-ahead volatility forecasts. For this purpose, a HAR model of the realized range adjusted for discrete trading is augmented by each of these variables and compared with the model's default form. The results show that the considered liquidity measures lead to very modest improvements in forecasting performance. The overnight returns exhibit some in-sample forecasting power. However, the accuracy improvement of out-of-sample forecasts is unequivocally non-significant.

1. Introduction

Financial market volatility enjoys a sustained interest in theoretical and empirical literature. Over time, the increasing availability of transaction and quote data has allowed the focus to be set on intraday levels. High-frequency data enable the application of advanced volatility proxies like the realized volatility based on summing squared intraday returns (Andersen and Bollerslev, 1998) or the realized range comprising the sum of scaled intraday price ranges (Christensen and Podolskij, 2007; Martens and van Dijk, 2007). Among others, issues of current research interest are, the development of advanced forecasting approaches (Corsi, 2009; Ghysels et al., 2006), the adjustment of existing estimators in order to deal with empirical drawbacks (Hansen and Lunde, 2006; Zhang et al., 2005) as well as testing and accounting for the existence of jumps (Andersen et al., 2007).

The aim of this study is to examine whether in-sample and out-of-sample forecasting improvements can be achieved by augmenting a heterogeneous autoregressive (HAR) model for the realized range with common liquidity measures and various specifications of overnight returns. Due to its easy application and overall very good forecasting performance, the HAR model is broadly supported in the plethora of modeling and forecasting approaches for financial market volatility.

This article complements the large body of literature in several directions. First, the incremental economic value of trading volume, number of daily transactions and overnight returns for forecasting return volatility is analyzed for a large stock sample. The research is motivated by the pronounced lead–lag correlations of these variables to the daily volatility. Second, the results are sorted by the stocks' liquidity level in order to enhance insights into possible forecast improvements. Third, since the realized range established by Martens and van Dijk (2007) and Christensen and Podolskij (2007) is used instead of realized volatility, as in the original model proposed by Corsi (2009), the study also demonstrates its eligibility for accurate volatility modeling within a (augmented) HAR model.

The notion that liquidity measures like trading volume, number of transactions, bid-ask spreads, or overall market liquidity are related to the return volatility is widespread. A number of studies address the relationship between volume and return volatility. A popular theoretical explanation is based on the mixture of distribution hypothesis which suggests that volatility is positively related to trading volume due to its dependence on a common latent mixing variable, the rate of information arrival (Park, 2010). However, a consensus on this relationship and its economic significance has not been reached yet. Chen et al. (2001) investigate the index volatility of

* Corresponding author. Tel.: +61 7 3735 7219; fax: +61 7 3735 3719.
E-mail addresses: n.todorova@griffith.edu.au (N. Todorova), soucek@europa-uni.de (M. Souček).

0264-9993/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econmod.2013.10.003

nine markets and conclude that trading volume contributes some information to the process of equity index returns. Darrat et al. (2003) find evidence of significant lead–lag relations between trading volume and return volatility in a large number of the DJIA stocks. Considering trading volume as a proxy for the changes in the information set available to market participants, Donaldson and Kamstra (2005) show that trading volume has a switching role between volatility forecasts based on stale information and option-implied estimates. Fuertes et al. (2009) also utilize trading volume for assigning market conditions for volatility forecasting. However, it is not certain whether the volume's informational role can be exploited to enhance performance when volume is explicitly incorporated into forecasting models. Brooks (1998) generates GARCH forecasts based on daily data and shows that augmenting daily stock return volatility models with measures of lagged volume leads only to very modest improvements, in terms of forecasting performance.[1]

Fleming and Kirby (2011) suggest that return volatility and trading volume exhibit similar long-run dynamics and include the logarithms of volatility and volume into a trend-stationary fractionally-integrated process. Their results show only minor gains for improving short-term volatility forecasts. We address this issue by incorporating trading volume information into another economically plausible volatility model with an advanced non-parametric volatility estimator and utilizing an extensive data set from the German market.

Taking standard liquidity measures into account, Jones et al. (1994) argue that the positive volatility–volume relation actually reflects the positive relation between volatility and the number of transactions, concluding that volatility is generated by the frequency of transactions, and not by their size. Thus, trading volume carries no information beyond that contained in the frequency of transactions. Naes and Skjeltorp (2006) present empirical evidence from order data of the Oslo Stock Exchange and also confirm that the number of trades is a more meaningful measure for investigating the relation between volatility and liquidity. Chan and Fong (2000) examine the number and size of trades, as well as order imbalances and provide evidence of their significant role in the volatility–volume relation for a sample of NYSE and Nasdaq stocks. In this context, the incorporation of the number of trades apart from the daily transaction volume appears to be a promising attempt to enhance forecasting accuracy.

Aiming to explain the persistence of volatility established by using GARCH-type models for 10 actively traded US stocks, Gallo and Pacini (2000) conclude that the information about the trading activity of the previous day is inferior to the information which arrives during the time when the market is closed. Taking overnight jumps into account when estimating multiday volatility is an important issue especially for practical purposes. There is a number of studies proposing different ways to include the impact of information arrivals during the closed-market time in the context of realized variance (e.g. Fleming et al., 2003; Hansen and Lunde, 2005; Koopman et al., 2004; Martens, 2002). Hansen and Lunde (2005) argue that, despite being very noisy, overnight returns do contain useful information beyond that included in realized volatility gained by returns observed during open market time. Using the functional coefficient model of Cai et al. (2000), Gallo (2001) provides some significant in-sample evidence of the impact of overnight surprises on intraday returns. However, the overnight surprises are found not to contribute to a significant improvement of out-of-sample volatility forecasts. On the other hand, Tseng et al. (2012) investigate three stock indices on the Taiwan Stock Exchange incorporating the previous nights' absolute returns into a HAR model. Their findings indicate that taking information arrivals from the market's nontrading time into account improves the performance of volatility forecasts in both in-sample and out-of-sample analyses.

Table 1
Descriptive statistics of the annualized daily realized ranges.

	Mean	SD	Skew	Kurt	Min	Max
Merck	0.2712	0.1128	1.3738	2.6463	0.0753	0.8231
Fresenius	0.2458	0.1245	1.8079	3.9195	0.0632	0.9451
Henkel	0.2087	0.0847	1.3673	2.5912	0.0514	0.6809
Linde	0.2381	0.1182	1.9317	4.3816	0.0727	0.8718
Lufthansa	0.2805	0.1159	1.8309	4.0017	0.1237	0.9256
Metro	0.2567	0.1322	1.9507	4.1621	0.0806	0.9850
Adidas	0.2176	0.0774	1.1716	1.7647	0.0661	0.5907
MAN	0.2957	0.1331	1.5352	2.9206	0.0975	0.9982
Dt Post	0.2550	0.1096	1.6788	3.3237	0.0993	0.8660
Dt Börse	0.2406	0.0915	1.3007	3.2270	0.0646	0.7828
Continental	0.2455	0.0933	1.8808	7.7579	0.0823	1.0882
ThyssenKrupp	0.2895	0.1280	1.4191	2.3029	0.1008	0.9818
Infineon	0.3268	0.1533	2.1779	5.8454	0.1237	1.3922
Commerzbank	0.2981	0.1492	1.8672	3.9204	0.1035	1.1405
BMW	0.2436	0.1054	2.0664	5.9258	0.0920	0.9736
Bayer	0.2730	0.1474	2.9616	14.7257	0.0983	1.7662
RWE	0.2452	0.1137	2.4037	7.7157	0.0905	0.9936
Volkswagen	0.2530	0.1123	2.0276	5.8363	0.1039	1.0209
BASF	0.2239	0.1028	2.0238	4.9590	0.0878	0.7961
MRück	0.2474	0.1444	2.2727	5.8581	0.0791	1.0328
Dt Telekom	0.2228	0.1085	2.6613	8.2224	0.1115	0.8928
SAP	0.2371	0.1015	1.8407	4.8580	0.0878	0.8843
Daimler	0.2632	0.1170	1.7609	3.4299	0.1068	0.8015
Dt Bank	0.2300	0.1078	2.0729	5.3355	0.0854	0.8434
Siemens	0.2356	0.1087	2.0010	4.6473	0.0974	0.7713
Allianz	0.2490	0.1450	2.2905	6.3322	0.0757	1.1309

Overall, these findings leave the question of the statistical significance of the impact of overnight returns and liquidity measures for volatility forecasting not conclusively clarified. The current study adopts a similar approach like Tseng et al. (2012) but runs an empirical analysis for a substantially larger data base, a different sample period and another asset class, investigating liquidity measures as well as various specifications of the overnight returns.

The article is arranged as follows. The next section presents the methodology of the study. Subsequent sections describe the data, in-sample and out-of-sample results. The final section concludes.

2. Methodology

2.1. Volatility proxy

The sum of squared intraday returns which is known in the recent literature as realized variance (often also referred to as realized volatility) firstly proposed by Andersen and Bollerslev (1998) is currently the most adopted estimator of integrated variance of one day in discrete settings. Andersen et al. (2003) show that the realized variance which uses all available data is a consistent estimator of the integrated variance when there is no microstructure noise. The increasing availability of high-frequency data and the theoretical properties of realized volatility make this estimator very appealing. However, very high-frequency prices are heavily contaminated by market microstructure effects, such as the bid-ask bounce, which distort the realized volatility to an extent dependent on the properties of the noise. The bias caused by microstructure noise is subject of active research interest (e.g., Bandi and Russell, 2006; Hansen and Lunde, 2006, among others). Various ways of dealing with the distortion due to microstructure noise are possible. A simple, often adopted solution of this problem is to sample at lower frequencies, for example every 5 min. Alternatively, several bias correction procedures have been proposed, such as the subsampling and the kernel-based approaches.[2]

[1] See also Andersen (1996) and Bollerslev and Jubinski (1999) for a literature review of older studies.

[2] A summary of basic assumptions about microstructure noise, their implications for realized volatility as well as various adjustment approaches can be found in McAleer and Medeiros (2008).