İstatistik Araştırma Dergisi
Journal of Statistical Research

ISSN: 2791-7616

(2023), 13 (2)
Başvuru / Received :29.05.2023
Kabul / Accepted :19.12.2023

https://dergipark.org.tr/tr/pub/jsstr

ARAŞTIRMA MAKALESİ **RESEARCH ARTICLE**

Forecasting Realized Volatility: Evidence from Nordic Stock Markets

Burak KORKUSUZ
University of Stirling, Division of Accounting and Finance / Dr.
Accounting and Finance Division
burak.korkusuz@stir.ac.uk
Orcid No: 0000-0001-9374-2350

Abstract

This study aims to determine the most effective model for forecasting volatility within the Nordic stock markets. In this regard, the forecasting power of HAR-RV, RSV, and PS models is compared to the ARFIMA-RV model using high frequency data for 7 Nordic stock market indices spanning from 2010 to 2019. One-day-ahead out-of-sample realized volatility forecasts are produced using a recursive window mechanism. The out-of-sample forecast losses are measured by the MSE and QLIKE criteria. The results indicate several noteworthy points. Firstly, the HAR-RV (PS and RSV) models are suggested to be best performing realized volatility models over the ARFIMA-RV model. Secondly, the separation of realized variance into positive and negative realized semivariances, which is known as good and bad volatilities, might offer valuable financial insights in certain situations, aiding the prediction of future realized volatility. Lastly, the results and findings are specific to market, data frequency, time horizon, and some characteristics of data, emphasizing the importance of these factors in interpreting the findings.

Keywords: Volatility, Forecasting, HAR-RV and ARFIMA-RV, Nordic stock markets

Corresponding Author / Sorumlu Yazar: 1- Burak KORKUSUZ, University of Stirling, Accounting and Finance Division.

Citation / Atıf: KORKUSUZ B. (2023). Forecasting Realized Volatility: Evidence from Nordic Stock Markets. *İstatistik Araştırma Dergisi*, 13 (2), 1-12.

Bu çalışma yazarın doktora tezinden türetilmiştir.

Citation / Atıf: KORKUSUZ B. (2023). Forecasting Realized Volatility: Evidence from Nordic Stock Markets. *İstatistik Araştırma Dergisi*, 13 (2), 1-12.

Volatilite Tahmini: İskandinav Hisse Senedi Piyasalarından Bulgular

Özet

Bu çalışma, İskandinav borsaları için en etkin volatilite tahmin modelini belirlemeyi amaçlamaktadır. Bu bağlamda, HAR-(RV, RSV ve PS) modellerinin tahmin gücü, 2010-2019 yılları arasında 7 İskandinav borsa endeksi için yüksek frekanslı veriler kullanılarak ARFIMA-RV modeli ile karşılaştırılmıştır. Özyinelemeli pencere mekanizması kullanılarak bir gün sonra gerçekleşen örneklem dışı volatilite tahminleri üretilmektedir. Örneklem dışı tahmin kayıpları, MSE ve QLIKE kriterleri ile ölçülür. Sonuçlar birkaç önemli noktaya işaret etmektedir. İlk olarak, HAR-RV (PS ve RSV) modellerinin, ARFIMA-RV modeline göre daha iyi performans gösteren model grubu olduğu öne sürülmektedir. İkincisi, varyansın pozitif ve negatif yarı varyanslara veya diğer bir deyişle iyi ve kötü varyanslara ayrıştırılması, bazı durumlarda, gelecekteki varyansın tahminine yardım eden faydalı finansal bilgiler sunabilir. Son olarak, sonuçlar ve bulgular pazara, veri sıklığına, zaman ufkuna ve verilerin bazı karakteristik özelliklerine özgüdür ve bulguların yorumlanmasında bu faktörlerin önemi vurgulanmaktadır.

Anahtar sözcükler: Volatilite, Tahmin, HAR-RV ve ARFIMA-RV, İskandinav hisse senedi piyasaları

1. Introduction

Although numerous studies have explored volatility predictability, a consensus on the optimal forecasting model remains elusive in the literature. Stocks, exchange rates, and crude oil constitute the most examined assets, predominantly using GARCH family models since the 1980s. GARCH models with daily data have become prevalent. However, the availability of minute-wise data prompted the adoption of intraday data for volatility forecasting. Consequently, realized measures, derived from tick-by-tick data using diverse statistical formulas, emerged as a result of these significant advancements. Corsi introduced the paradigm-shifting Heterogeneous Autoregressive model of the Realized Variance (HAR-RV) in 2009, which has since become a prominent model in recent volatility research.

The HAR-RV model extends standard realized volatility models, typically based on high-frequency financial data, to incorporate lower-frequency inter-day data. This study employs HAR-RV type models for volatility forecasting across 7 Nordic stock markets between 2010 and 2019, comparing the outcomes with the ARFIMA-RV model. Utilizing 5-minute realized variance series derived from high-frequency data, the study generates one-day-ahead out-of-sample volatility forecasts employing recursive windows forecasting technique. Forecast accuracy is assessed using criteria such as MSE, QLIKE, and the conditional Giacomini-White pairwise test (2006) to determine statistical significance in forecast errors between competing models.

The focus on Nordic stock markets stems from the region's reputation for innovation and technological advancements, attracting successful companies particularly in biotechnology, clean energy, information technology, and gaming. Despite this, there is a lack of empirical studies within the wider body of literature that specifically focused on Nordic stock markets from the perspective of realized volatility. Therefore, this research seeks to identify the most effective model for predicting volatility within these markets. The findings highlight several key points: the superiority of HAR-RV (PS and RSV) models over the ARFIMA-RV model; the potential insights gained from separating realized variance into positive and negative realized semivariances for predicting future volatility; and the specificity of results based on market, data frequency, time horizon, and data characteristics, emphasizing the importance of each contribution in the literature. These findings hold practical implications in financial econometrics, especially in risk management, option pricing, and portfolio management, where accurate realized volatility forecasts are essential.

This paper is organised as follows: Section 2 presents the review of related literature. In Section 3, the data and methods used in this study are explained in more detail. Afterwards, Sections 4 gives the empirical results and their evaluations respectively. Finally, the conclusion is presented in Section 5.

2. Literature Review

During the 2000s, the accessibility of high-frequency data revolutionized financial volatility research. Numerous studies, including works by Andersen and Bollerslev (1997), Andersen, Bollerslev, Diebold, and Labys (2001), Martens and Zen (2004), Koopman, Jungbacker, and Hol (2005), Chortareas, Jiang, and Narkervis (2011), and Sevi (2014), demonstrate that models utilizing intraday or high-frequency data notably enhance the precision of volatility forecasts. This enhancement can be attributed to several reasons. Firstly, owing to the persistence property of volatility, high-frequency data offer a more accurate assessment of current volatility, thus enhancing future volatility predictions. Secondly, these data contribute to better volatility forecast evaluations by minimizing inconsistencies in volatility model rankings. Additionally, high-frequency data aid in comprehending the dynamic nature of financial volatility, a crucial aspect for effective modelling and forecasting.

The availability of high-frequency data prompted the use of intraday data to create more direct proxies for financial volatility, referred to as "realized measures" or "realized variance" in this context. Studies by Barndorff-Nielsen, Kinnebrock, and Shephard (2010), Andersen, Bollerslev, Diebold, and Labys (2003), among others, validate the superiority of realized variance as a proxy for true volatility. As a result, realized variance has become the most commonly used volatility measure among alternative proxies. This measure, derived from the summation of squared intraday returns, necessitates careful consideration of the frequency interval for intraday data to ensure accuracy. Scholarly investigations, such as those by Martens and Zein (2004), suggest that increased intraday observation frequency leads to more precise daily volatility estimations. However, excessively high frequencies, termed ultra-frequency data, may distort data efficiency due to microstructure noise, leading scholars like Hol and Koopman (2002) to propose frequency intervals between 5 and 30 minutes.

Liu, Patton, and Sheppard (2015) conducted a comprehensive analysis comparing various realized measures, concluding that surpassing the accuracy of the five-minute realized variance is challenging. Consequently, the consensus among researchers and practitioners leans towards utilizing the 5-minute realized variance as the target volatility. This study opts for the simple realized variance based on 5-minute squared returns for estimations. Andersen and Bollerslev's (1998) introduction of the realized variance as a more accurate measure of true volatility compared to daily squared returns marked a pivotal moment. Initially employed as an estimator for true volatility assessment in volatility models' forecasting performance, the realized variance gained prominence with the increasing availability of high-frequency data. Andersen et al. (2001, 2003) emphasized that higher data frequencies lead to more accurate volatility forecasts. However, ultra-high frequency data may introduce microstructure noise, destabilizing parameter estimates. Evidence from various studies, including Blair, Poon, and Taylor (2001), Engle (2002), Andersen (2003), Koopman et al. (2005), and Bollerslev (2009), supports the notion that high-frequency returns data outperform daily returns data in measuring true volatility.

Corsi (2009) proposed the Heterogeneous Autoregressive model of realized variance (HAR-RV) based on the Heterogeneous Market Hypothesis, showcasing its impressive performance despite its simple structure. Subsequent research by Andersen et al. (2011), Patton and Sheppard (2009), and Bollerslev et al. (2016) corroborated the superior performance of the HAR-RV model. Barndorff-Nielsen, Kinnebrock, and Sheppard (2010) introduced positive and negative realized semivariance measures derived from signed high-frequency intraday returns. Expanding the HAR-RV model, Sevi (2014) decomposed volatility into jump and continuous components, negative and positive realized semivariances, and incorporated the leverage effect. Patton and Sheppard (2015) emphasized the significance of negative realized semivariance in future volatility forecasting, advocating for an asymmetric HAR model that includes both positive and negative realized semivariances. Fang, Jiang, and Luo (2017) highlighted the importance of decomposing only the daily component of the HAR model, suggesting that considering all components alters the influence of explanatory variables.

Research on realized variance has become a highly explored area in forecasting volatility, particularly after significant advancements in this field. The abovementioned studies were among the pioneers in demonstrating that realized variance proves to be a more precise gauge of volatility when compared to squared returns. Using the data of different financial assets many studies aim to find out the best performing volatility forecasting model. However, the literature has still to reach a consensus. Most of the papers concentrate mainly on the stock markets, yet in the context of single (or several) stocks or market indices. Even though stock market indices become one of the most investigated financial assets, there is still a gap in the literature in terms of the most recent developments in the research of tick-by-tick data such as the introduction of new models and applications of those in international markets. Therefore, this study fills this gap in the literature by carrying out a volatility forecasting exercise within 7 Nordic stock market indices between 2010-2019.

3. Data and Methodology

3.1. Data Description

The data used in this study is provided by the Oxford-Man Institute of Quantitative Finance Realized Library. 5-min realized (and semi) variance series are employed for the volatility prediction of 7 Nordic stock market indices which are Amsterdam Exchange Index, Belgium 20 Index, OMX Copenhagen 20 Index, OMX Helsinki All Share Index, OMX Stockholm All Share Index, Oslo Exchange All Share Index, and Swiss Stock Market Index. Majority of them consist of the indices of developed countries. The full list of index names and abbreviations is given at Table 1. The reason of investigating Nordic stock markets is that Nordic region has a reputation for being at the forefront of innovation and technology. The stock markets in this region have seen listings of many successful and innovative companies, especially in sectors such as information technology, biotechnology, clean energy, and gaming. These companies contribute to the growth and development of the Nordic stock markets and attract investor interest.

The dataset of this forecasting exercise is the post-2007/2008 global financial crisis period. Each index covers the period of 9 years, specifically from January 4, 2010 to October 3, 2019.[1] The number of observations in each index is approximately 2400 trading days. However, total trading days in a year can differ between each countries due to different public holidays and nontrading periods. In this forecasting exercise, the initial sample comprises approximately one year period (330 obs. [2010-2011]), whilst the time interval of out-of-sample volatility forecasts is 8 years (2070 obs. [2011-2019]). We arbitrarily choose the in-sample length as 330 observations considering the length at least one year period to let the regression fit normally and obtain a longer out-of-sample period. This is because the main objective of this work is to evaluate the out-of-sample performance of the models.

Table 1. The full list of index names and abbreviations

Symbol	Name	Mean	St. Dev.	Skew.	Ex. Kur.
AEX	Amsterdam Exchange Index	7.23E-05	0.00011	9.1184	135.49
BFX	Belgium 20 Index	6.66E-05	8.51E-05	8.1125	109.33
OMXC20	OMX Copenhagen 20 Index	8.74E-05	0.00031	29.286	1039.0
OMXHPI	OMX Helsinki All Share Index	7.33E-05	0.00046	46.596	2253.9
OMXSPI	OMX Stockholm All Share Index	6.31E-05	0.00024	34.113	1410.1
OSEAX	Oslo Exchange All Share Index	8.74E-05	0.00015	8.8676	129.06
SSMI	Swiss Stock Market Index	5.51E-05	0.00012	21.799	680.43

Source: (Author's calculation)

Table 1 presents the first four statistical moments of 5-minute realized variance series for different indices respectively, namely; mean, standard deviation, skewness, and excess kurtosis. The values of the moments are as commonly seen in the literature. The means of realized variance series are close to zero for each index that is consistent with the literature. The series also have a high positive skew. Lastly, the values of the fourth moment indicate the leptokurtic distribution for all the dataset. Therefore, it can be pointed out that the series have non-Gaussian distribution.

Liu, Patton, and Sheppard (2015) conducted a comparison of more than 400 realized measures, noting that the challenge in outperforming the five-minute realized variance significantly has failed. Hence, we opt for the 5-minute realized variance as a stand-in for true volatility. Generally, higher data frequency tends to enhance the precision of volatility estimation. However, increased frequency may introduce errors in measurement and price discreteness due to microstructure noise, potentially affecting the efficiency of the data in higher frequencies.

[1] The Oxford-Man Institute's Realized Library is no longer available after 2020 and they have no future plans to replace this.

3.2. Methods

3.2.1. Realized Volatility and Realized Volatility Models

Volatility, being latent, requires a substitute to represent the true volatility. Initially, researchers commonly utilized daily squared returns until Andersen and Bollerslev's work in 1998 revealed their inadequacy compared to cumulative intraday squared returns. Subsequently, Andersen, Bollerslev, Diebold, and Labys (ABDL; 2003) introduced the concept of realized variance, which sums the squared intraday returns. Both realized variance and daily squared returns serve as unbiased estimates of volatility, but realized variance is renowned for its high efficiency as a measure of volatility.

$$RV_t = \sum_{i=1}^{m} r_{t,i}^2 \tag{1}$$

As expressed in Equation (1), realized variance is computed by summing the squared intraday returns, where 'm' signifies the number of intraday observations on day 't'. In theory, a higher 'm' value leads to a more precise estimation of daily volatility. However, excessively high 'm' numbers can distort the efficiency of high-frequency data due to microstructure noise effects. ABDL (2003), Martens (2001), and Hol and Koopman (2002) recommend a frequency interval between 5 and 30 minutes. In a recent study by Liu, Patton, and Sheppard (2015), comparing over 400 realized measures, it was noted that surpassing the accuracy of the five-minute realized variance is challenging. Consequently, we opt for the 5-minute realized variance in estimating the HAR and ARFIMA models. Barndorff-Nielson et al. (2010) separate realized variance into positive and negative realized semivariances, depicting good and bad volatilities.

3.2.2. HAR-RV Models

The HAR-RV model is based on the heterogeneous market hypothesis of Muller, Dacorogna, Dave, Olsen, Pictet and von Weizsacker (1997). According to this hypothesis, there are three types of investors that have different risk preferences and different reactions to the same new market information. In addition to the hypothesis, the same researchers develop the Heterogenous Autoregressive Conditional Heteroskedasticity (HARCH) model. Inspired by the HARCH model and its background hypothesis, Corsi (2009) proposes the HAR-RV model that is an additive cascade model of different volatility components. The model is specified as:

$$RV_{t+h}^d = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+h} \tag{2}$$

where RV_t^d is daily realized volatility; RV_t^w refers to weekly realized volatility, and then RV_t^m indicates monthly realized volatility. RV_t^w and RV_t^m can easily be calculated as follows:

$$RV_t^w = \frac{1}{5}(RV_{t-5}^d + RV_{t-4}^d + \cdots + RV_{t-1}^d)$$

$$RV_t^m = \frac{1}{22}(RV_{t-22}^d + RV_{t-21}^d + \cdots + RV_{t-1}^d$$

The main point of the HAR-RV model is to predict future volatility using three different volatility components; a daily (RV_t^d), a weekly (RV_t^w), and a monthly (RV_t^m) components. The HAR-RV model can simply be estimated by the ordinary least square (OLS) method. The model is such a good alternative to the ARFIMA model. The HAR-RV model can also capture long memory characteristics of volatility even though it is not in the class of long memory models. In practice, the HAR-RV model is found to be such a promising model as the model performance is remarkably good in spite of its simple structure.

Different types of investors have different objectives in financial markets. For instance, some investors are completely hedgers whilst some others are completely speculators. Hence the HAR-RV model is based on capturing different reactions of different investors through the simple autoregressive process. Financial interpretation of the model is that the investors are divided into three different categories. In the model, RV_t^d, RV_t^w, and RV_t^m components represent short-term, middle-term, and long-term investors respectively and indicate the degree of different investors' impact on current realized volatility. In other words, the model coefficients provide an understanding of how these different market participants react and perceive to volatility. Moreover, the HAR-RV model can successfully capture the persistence feature of realized volatility.

$$RSV_t^+ = \sum_{i=1}^{m} r_{t,i}^2\, I\,\{r_{t,i} > 0\} \tag{3}$$

$$RSV_t^- = \sum_{i=1}^{m} r_{t,i}^2\, I\,\{r_{t,i} < 0\} \tag{4}$$

where $I\{\cdot\}$ is an indicator function. We should also note that $RV_t = RSV_t^+ + RSV_t^-$. Barndorff-Nielsen, Kinnebrock, and Sheppard (2010) first introduce positive and negative realized semivariance measures, which are obtained from the signed high frequency intraday returns. Patton and Sheppard (2015) decompose only the daily explanatory HAR model component into negative and positive realized semivariances. In this study, we call Patton and Sheppard (2015)'s model as the HAR-PS model. The HAR-PS specification is presented in Equation (5):

$$RV_{t+h}^d = \beta_0 + \beta_d^- RSV_t^- + \beta_d^+ RSV_t^+ + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+h} \tag{5}$$

Following that it is added one more realized semivariance specification to the model comparison that decomposes not only the daily component, but also separates weekly and monthly components. The model of Patton and Sheppard (2011) is called here as the HAR-RSV and the model is given as follows:

$$RV_{t+h}^d = \beta_0 + \beta_d^- RSV_t^- + \beta_d^+ RSV_t^+ + \beta_w^- RSV_t^- + \beta_w^+ RSV_t^+ + \beta_m^- RSV_t^- + \beta_m^+ RSV_t^+ + \varepsilon_{t+h} \tag{6}$$

According to the seminal research of Patton and Sheppard (2011) and Barndorff-Nielsen, Kinnebrock, and Sheppard (2010), the decomposition of realized variance into positive and negative realized semivariances (or good and bad volatilities) adds more information for the prediction of future volatility.

3.2.3. ARFIMA-RV model

The long memory autoregressive fractionally integrated moving average (ARFIMA) model is in the class of long memory models and therefore can successfully capture the persistency feature of volatility. Andersen et al. (2003) suggest the univariate ARFIMA model in order to model the realized volatility. An ARFIMA (p, d, q) model is presented by:

$$\varphi(L)(1-L)^d(RVt - \mu) = \theta(L)\varepsilon_t \tag{7}$$

where $\varphi(L)$ and $\theta(L)$ are the lag polynomials of the autoregressive *(AR)* and moving average *(MA)* components. ε_t is the error term which is distributed approximately as a Gaussian white noise [$N(0,\sigma_u^2)$]. The fractional differencing parameter is represented by d in equation (7). The *AR* and *MA* components explain the short memory properties of volatility and as for the d, it accounts for the long memory properties of volatility. The value of d is expected between 0 and 0.5 in order to capture long memory property. Andersen et al. (2003) found $d=0.401$. In this context, a general empirical conclusion with ARFIMA model is that this framework outperforms traditional GARCH models which are based on daily returns (Hansen and Lunde, 2010).

3.2.4. Recursive window forecasting method and loss functions

The recursive window method is used for obtaining the volatility forecasts. The loss functions; the mean squared error (MSE) and the quasi-Gaussian log-likelihood (QLIKE) are considered in order to compare the models. Lastly, the Giacomini and White (2006) pairwise test is employed to evaluate the forecasting performance of two models. Initially, the whole sample needs to be divided into two subgroups such as the initial sample and out-of-sample windows. In the literature, there is no consensus on how to select an appropriate forecasting window. Since the main objective of this work is to evaluate the out-of-sample performance of the models, we arbitrarily choose the initial and out-of-sample windows considering a length that allows the regression fit normally and obtain longer out-of-sample period. The recursive window's working principle does work the way that the estimation sample is then rolled forward by adding one new observation and not dropping the most distant observation. In this way, the size of initial sample window used to estimate the models grows in each step.

Since the main goal of this work is to compare the performance of the competing models, we need to measure the ability of the models using some loss functions. Many different forecasting criteria can be used for comparison purpose. Lopez (2001) points out that it is not clear to decide which measure is the most accurate to which model. On the other hand, Patton (2011) documents the robustness of the QLIKE and MSE criteria. The reason is explained as such: in the case of such a noisy volatility proxy, the QLIKE and MSE provide consistent rankings for volatility models. In this regard, these two criteria are selected, namely the mean squared error (MSE) and the quasi-Gaussian log-likelihood (QLIKE). The loss functions are specified as follows:

$$QLIKE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[log\widehat{RV_t^2} + \frac{\widehat{RV_t^2}}{RV_t^2}] \quad (8)$$

$$MSE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[RV_t^2 - \widehat{RV_t^2}]^2 \quad (9)$$

where RV_t^2 is the proxy of the true volatility and $\widehat{RV_t^2}$ is the volatility forecast. The number of observations is represented by τ. The QLIKE and MSE loss functions are frequently used criterion in the literature due to being robust to the noisy volatility proxy. Patton and Sheppard (2009) indicate that the QLIKE is powerful in the Diebold-Mariano test, which is quite similar test to the Giacomini and White (GW) test that we use here. Although the MSE and QLIKE are the most frequently used criteria, there is still a possibility that such a model with the lower error may not be exactly better than the other model. For further robustness, it is necessity to apply the GW test.

4. Empirical Results

Employing one-step-ahead recursive window forecasting method this study assesses the forecast results of four competing models (HAR-RV, HAR-PS, HAR-RSV, and ARFIMA-RV) for seven Nordic stock market indices. The forecasting exercise results of the four models are given in Table 2 and 3. Those forecasts are generated using the recursive window technique and then the forecasts' losses are measured by the QLIKE and MSE criterion. Lower QLIKE and MSE values in these tables indicate better performance and higher accuracy in forecasting the future volatility of stock market indices.

According to the results of QLIKE and MSE loss functions, it is clear that HAR-RV type models (e.g. HAR-RV, HAR-PS, and HAR-RSV) outperform ARFIMA-RV model. This means that HAR-type models are found to be promising models at forecasting realized volatility and exhibits remarkably good performance in spite of its simple structure in comparison with ARFIMA-RV model.

Table 2. QLIKE for recursive window forecast models

Index/Model	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	**-8.8430**	-8.8420	-8.8407	-8.6583
BFX	-8.8527	-8.8528	**-8.8532**	-8.7515
OMXC20	**-8.4475**	-8.4293	-8.3141	-8.1909
OMXHPI	**-8.6986**	-8.6938	-8.6664	-8.6894
OMXSPI	-8.9579	-8.9805	**-9.0016**	-7.9089
OSEAX	-8.6303	**-8.6371**	-8.6183	-8.4479
SSMI	-9.0730	**-9.0746**	-9.0708	-8.8375

Source: (Author's estimation)

Table 3. MSE for recursive windows forecast models

Index/Model	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	**8.29E-09**	8.40E-09	8.42E-09	9.42E-09
BFX	**4.85E-09**	5.06E-09	5.19E-09	5.47E-09
OMXC20	**3.90E-08**	3.93E-08	4.05E-08	4.42E-08
OMXHPI	7.85E-09	7.81E-09	**7.63E-09**	1.15E-08
OMXSPI	8.77E-09	**8.25E-09**	8.29E-09	6.77E-08
OSEAX	**1.76E-08**	1.79E-08	1.81E-08	1.95E-08
SSMI	**1.48E-08**	1.69E-08	1.78E-08	1.52E-08

Source: (Author's estimation)

When these results are examined in more details, the best performing HAR-RV genre models differ from one market index to another index and also between the employed two loss functions. Therefore, this analysis evaluates those results index by index separately and then try to draw a more precise picture in the end. In terms of the AEX and OMXC20 stock market indices, HAR-RV model is superior to the other counterparts such as HAR-PS and HAR-RSV models. We should remember here that while the HAR-PS model decomposes only the daily component, the HAR-RSV model decomposes the daily, weekly and monthly components. When the BFX index is evaluated, it can be seen that the loss functions give opposite results. For instance, the HAR-RSV is the best performing volatility forecasting model for the QLIKE criterion whereas the MSE loss function favours the HAR-RV for this stock market index. Afterwards, when it comes to the OMXHPI index, the findings of the QLIKE and MSE tell vice versa compared to the BFX such that the HAR-RSV is the best performing volatility forecasting model for the MSE criterion, whilst the QLIKE loss function suggests the HAR-RV for this stock market index. The decomposition of positive and negative semi variances does work for the index of the OMXSPI. However, the only difference is that while the QLIKE supports the HAR-RSV, the MSE criteria shows the HAR-PS as a best performing volatility forecasting model for the OMXSPI. Lastly, the results of the OSEAX and SSMI stock market indices are in the same direction even though the loss functions yield opposite results. For example, the QLIKE selects the HAR-PS as a best performing model for both the indices, but the HAR-RV is suggested by the MSE for the same indices. It is important here to note that each one of these loss functions has a specific calculation method that could cause to yield different results, which is unsurprising.

Table 4. Conditional Giacomini-White test results

AEX	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.084 (-)	0.095 (-)	0.050 (-)
HAR-PS	-	0.023 (-)	0.002 (-)
HAR-RSV	-	-	0.005 (-)

BFX	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.181 (-)	0.184 (-)	0.197 (-)
HAR-PS	-	0.086 (-)	0.046 (-)
HAR-RSV	-	-	0.030 (-)

OMXC20	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.079 (-)	0.000 (-)	0.000 (-)
HAR-PS	-	0.000 (-)	0.000 (-)
HAR-RSV	-	-	0.000 (-)

OMXHPI	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.045 (+)	0.018 (+)	0.000 (-)
HAR-PS	-	0.000 (+)	0.000 (-)
HAR-RSV	-	-	0.000 (-)

OMXSPI	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.000 (+)	0.001 (+)	0.000 (-)
HAR-PS	-	0.017 (-)	0.000 (-)
HAR-RSV	-	-	0.000 (-)

OSEAX	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.264 (-)	0.024 (-)	0.000 (-)
HAR-PS	-	0.181 (-)	0.000 (-)
HAR-RSV	-	-	0.000 (-)

SSMI	**HAR-PS**	**HAR-RSV**	**ARFIMA-RV**
HAR-RV	0.451 (-)	0.151 (-)	0.084 (-)
HAR-PS	-	0.089 (-)	0.110 (+)
HAR-RSV	-	-	0.286 (+)

Source: (Author’s estimation)

In order to underpin those results, further robustness tests need to be done, in particular the pairwise GW test to test the equal conditional predictive ability of the forecasts produced by the competing models. For instance, we have two different forecasted series, namely X and Y. Assuming that the values of loss functions of X are lower than Y. Can it be said that the forecast X has a superior performance compared to the forecast Y? Or is it possible that the difference between the forecasts X and Y is inherently insignificant? In order to test conditional predictive ability Giacomini and White (2006) suggest a pairwise test on equal conditional predictive ability, which examines whether two different forecasting models statistically have the same accuracy or not. In short, this test evaluates the forecasting performance of two competing models.

The p-values of the conditional GW test results are reported in Table 4. The null hypothesis is that "the two models (row and column) statistically have the equal predictive accuracy" is tested in terms of squared forecast error. The signs, + and –, in bracket show which model outperforms best and which model is outperformed. A positive sign indicates the superiority of the column model, whilst a negative sign means that the model in row outperforms the column model. In more detail, a positive sign means that the model in row has larger forecast loss in comparison with the model in column, which implies that the column model is significantly superior. In a similar vein, a negative sign does imply that the row model forecast performs significantly better compared to the column model forecast, since the latter produces larger loss. If the test statistics higher than 0.05 critical value (which implies the null hypothesis cannot be rejected), this means that the column and row models perform equally well, so that it is difficult to say that whether the row model or column model is superior. The results in Table 4 further confirm the superiority of the HAR specification over the ARFIMA-RV model as the last column signs are mostly negative and the null hypothesis "row and column models statistically have the equal predictive accuracy" is rejected. Therefore, column model (HAR) is superior to the row model (ARFIMA-RV). However, the evaluation among the HAR type models is unclear, meaning equal forecasting performance between the column and row (HAR-RV, HAR-PS, and HAR-RSV) models. This is because the values of loss functions between the winning model and the second winning model in Tables 2 and 3 are quite close to each other. It is difficult to say which HAR model is best among only HAR models regardless of ARFIMA-RV model. These findings are also in line with the results of both loss functions.

In a nutshell, each one of the loss functions has a specific calculation method, causing to produce different results. The QLIKE and MSE are the most popular and frequently used ones in the literature due to being robust to the noisy volatility proxies. In this work, the QLIKE and MSE criteria indicate that the HAR specification is the winner against the ARFIMA-RV model in all the indices, whereas the superiority among only the HAR models differs in the stock market indices.

5. Conclusion

This exercise compares the forecasting power of HAR-RV, RSV, and PS models to the ARFIMA-RV models which are derived from high frequency data. In this regard, 7 different Nordic stock market indices in the region between 2010-2019 are included. One-day-ahead out-of-sample realized volatility forecasts are produced using the recursive window mechanism. The out-of-sample forecast losses are measured by the MSE and QLIKE loss functions. Afterwards, the conditional Giacomini-White pairwise test is used to test the forecasting accuracy of the competing models. In this volatility forecasting exercise, the HAR-type models are found to be promising models against the ARFIMA-RV model. Moreover, the decomposition of realized variance into positive and negative realized semivariances (or good and bad volatilities), in certain cases, could add more information for the prediction of future volatility.

In the end, the aim of this empirical exercise is an attempt to find which model best fits in the data of Nordic stock markets. Nordic region has an enviable reputation for being at the forefront of innovation and technology and therefore the Nordic stock markets have seen listings of many successful and innovative companies. These companies contribute to the growth and development of the Nordic markets and attract institution and investor attention. Therefore, the findings of this study could be utilized in many practices in the Nordic stock markets where the realized volatility forecasts are required to generate accurate volatility forecasts, especially for the applications of risk and portfolio managements.

References

Andersen, T. G. & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885-905.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4, 115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association*, 96, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modelling and Forecasting Realized Volatility. *Econometrica*, 71(2), 579-625.

Barndorff-Nielsen, O., S. Kinnebrock, & N. Shephard. (2010). Measuring Downside Risk: Realized Semi-variance. In Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, T. Bollerslev, J. Russell, and M. Watson, eds. Oxford; New York: Oxford University Press, 117–136.

Blair, B. J., Poon, S.-H., & Taylor, S. J. (2001). Forecasting S&P100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, 105(1), 5-26.

Bollerslev, T.; S. Z. Li; & V. Todorov. (2016). Roughing up Beta: Continuous vs. Discontinuous Betas and the Cross-Section of Expected Stock Returns. *Journal of Financial Economics*, 120, 464–490.

Chortareas, G., Jiang, Y., & Nankervis, J. (2011). Forecasting exchange rate volatility using high-frequency data: Is the euro different? *International Journal of Forecasting*, 27, 1089-1107.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174-196.

Engle, R. (2002). New frontiers for arch models. *Journal of Applied Econometrics*, 17(5), 425-446.

Fang, N., Jiang, W., & Luo, R. (2017). Realized Semivariances and the Variation of Signed Jumps in China's Stock Market. *Emerging Markets Finance and Trade*, 53(3), 563-586.

Hansen, P. R., & Lunde, A., (2010). Forecasting volatility using high frequency data. A systematic review, 1-37.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements. *Journal of Empirical Finance*, 12, 445–475.

Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. Journal of Econometrics, 187(1), 293-311.

Martens, M., & Zein, J. (2004). Predicting financial volatility: high-frequency time-series forecasts vis-`a-vis implied volatility. Journal of Futures Markets, 24, 1005–1028.

Martens, M., & Zein, J. (2004). Predicting financial volatility: high-frequency time-series forecasts vis-`a-vis implied volatility. Journal of Futures Markets, 24, 1005–1028.

Merton, Robert C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation, Journal of Financial Economics, 8, 1-39.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V. Von Weizsäcker, J. E. (1997). Volatilities of different time resolutions, Analysing the dynamics of market components, Journal of Empirical Finance, 4 (2-3) (1997), 213-239.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1), 246 - 256.

Patton, A. J. & Sheppard, K. (2009). Optimal combinations of realized volatility estimators. International Journal of Forecasting, 25(2), 218-238.

Patton, A. J. & Sheppard, K. (2015). Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility. The Review of Economics and Statistics, 97(3), 683-697.

Sevi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. European Journal of Operational Research, 235, 643–659.