**ISSN: 2278-4632**
**Vol-13, Issue-06, June 2023**

# REALIZED VOLATILITY PREDICTION IN STOCK MARKET

**M. Rajesh** M.Tech Student, Department of Computer Science & Engineering, Sagi Rama Krishna Raju Engineering College, Bhimavaram, AP, India.
**Dr.V MNSSVKR GUPTA** Assistant Professor, Department of Computer Science& Engineering, Sagi Rama Krishna Raju Engineering College, Bhimavaram, AP, India.

## ABSRACT:

We consider various MIDAS (Mixed Data Sampling) regression models to predict volatility. The models differ in the specification of regressors (squared returns, absolute returns, realized volatility, realized power, and return ranges), in the use of daily or intra-daily (5-minute) data, and in the length of the past history included in the forecasts. The MIDAS framework allows us to compare models across all these dimensions in a very tightly parameterized fashion. Using equity return data, we find that daily realized power (involving 5-minute absolute returns) is the best predictor of future volatility (measured by increments in quadratic variation) and outperforms model based on realized volatility (i.e. past increments in quadratic variation). Surprisingly, the direct use of high-frequency (5-minute) data does not improve volatility predictions. Finally, daily lags of one to two months are sufficient to capture the persistence in volatility. These findings hold both in- and out-of-sample.

## INTRODUCTION:

The conditional volatility literature, starting with Engle's (1982) ARCH-class of models, has been successful at capturing the dynamics of return variance using simple parametric models. A measure of that success is the widespread use of such models in all areas of finance by academics and practitioners alike. And while most researchers would agree that it is important to have a good prediction model of conditional volatility, the question of what model to use is still unsettled**.**

When it comes to forecasting volatility, there are many existing models in addition to the benchmark ARCH/GARCH models of Engle (1982) and Bollerslev (1986) which cast future variance as a polynomial of past squared returns, i.e., σˆ 2 t+1|t ≡ A(L)r 2 t . One alternative is to look for variables, other than squared returns, that relate to future volatility. Ding et al. (1993) and several others show that low-frequency components of volatility might be better captured by absolute returns instead of squared returns. Also, Alizadeh et al. (2002) and Gallant et al. (1999) find daily ranges (high-low price ranges) to be good predictors of volatility. Another rapidly growing research area focuses on data-driven models of realized volatility computed from intra-daily returns sampled at very short intervals such as 5 minutes (Andersen and Bollerslev (1998)).1 All these models suggest a variety of possible ways to forecast volatility. Hence, it seems natural to ask whether some of the suggested predictors are clearly dominated by others and whether there are real benefits from using high-frequency data.2 These questions have proven difficult to answer because the models considered are so different in terms of regressors, frequencies, parameterizations, and

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

return histories, that is it difficult to directly compare them.

We use Mixed Data Sampling (henceforth MIDAS) regression models introduced in Ghysels, Santa-Clara and Valkanov (2002a,b) to provide answers to these questions. MIDAS regressions allow us to run parsimoniously parameterized regressions of data observed at different frequencies. There are several advantages of using mixed data sampling regressions. They allow us to study, in a unified framework, the forecasting performance of a large class of volatility models which involve:

(i) Data sampled at different frequencies;
(ii) Various past data window lengths; and
(iii) Different regressors. The specification of the regressions

combine recent developments regarding estimation of volatility and a not so recent literature on distributed lag models.3 We focus on predicting future conditional variance, measured as increments in quadratic variation (or its log transformation) from one week to one month horizons, because these are the horizons that are most widely used for option pricing, portfolio management, and hedging applications.

First, we use MIDAS regressions to examine whether future volatility is well predicted by past daily squared returns, absolute daily returns, realized daily volatility, realized daily power (sum of intra-daily absolute returns, a measure proposed by Barndorff-Nielsen and Shephard (2003b, 2004)), and daily range. Since all of the regressors are used within a framework with the same number of parameters and the same maximum number of lags, the results from MIDAS regressions are directly comparable. Hence, the MIDAS setup allows us to determine if one of the regressors dominates others. We find that, for the Dow Jones Index and six individual stock return series, the realized power clearly dominates all other daily predictors of volatility at all horizons. Importantly, the predictive content of the realized power is evident not only from in-sample goodness of fit measures, but also from out-ofsample forecasts. The daily range is also a good predictor in the sense that it dominates squared and absolute daily returns. Our method is a significant departure from the usual autoregressive model building approach embedded in the ARCH literature and its recent extensions such as high-frequency data-based approaches. A comparison of the MIDAS regressions with purely autoregressive volatility models reveals that the MIDAS forecasts are better at forecasting future realized volatility in- and out-of-sample.

Second, the weights in the MIDAS regressions are parameterized by a flexible function. Obviously, the choice of regressors is as important as is the profile of weights placed on them. In our MIDAS framework, the shape of the weight function is determined by two parameters that are estimated from the data. Hence, the weight profile on the lagged predictors is captured by the shape of the function, whose parameters are estimated from the data with no additional pre-testing or lag-selection procedures. We find that daily lags longer than about 50 days do not help (nor hurt) the forecasts, for any of the regressors.

Third, mixed data regressions allow us to directly project future realized volatility onto high-frequency (say 5-minute) squared and absolute returns without daily pre-filtering and without increasing the number of parameters. Hence, we are able to analyze if there are real benefits from directly using high-frequency data in volatility forecasting. Surprisingly, we find that forecasts using high-frequency data directly do not outperform those that use daily

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

regressors (although the daily regressors are themselves obtained through the aggregation of high-frequency data). It must be noted that none of these results are driven by over-fitting or parameter proliferation. Indeed, all MIDAS specifications – daily and 5-minutes – are directly comparable since they all have the same number of estimated parameters.

In summary, we find that daily realized power is the best predictor of future increments in quadratic variation, followed by the daily range. The prediction equations involve about 50 daily lags and there is no real benefit of using intra-daily data directly. These MIDAS regressions also outperform other linear forecast models involving daily realized volatility. Finally, all of the above results hold in- and out-of-sample. The out-of-sample forecasting precision, which is perhaps the ultimate measure of a model's forecasting potential, indicates that our results are unlikely to be due to sampling error or over-fitting.

The MIDAS regressions can also be used to model asymmetries and the joint forecasting power of the regressors. In fact, Engle and Gallo (2003) use the multiplicative error model (MEM) of Engle (2002) and find improvements in forecasting volatility from the joint use of absolute returns, daily ranges, and realized volatilities using S&P 500 index returns data. Interestingly enough, their results agree with ours, despite the different data set and different method, as they argue that range-based measures in particular provide a very good forecast of future volatility.

The paper is structured as follows. In a first section we introduce and discuss MIDAS volatility models. In section two, we use daily regressors to forecast weekly to monthly volatility in the MIDAS framework. The third section is devoted to MIDAS volatility forecasts involving intra-daily data. Section four concludes.

**MIDAS Models of Conditional Volatility:**

To fix notation, let daily returns be denoted by $r_{t,t-1} = \log(P_t) - \log(P_{t-1})$. Throughout the paper the time index t will refer to daily sampling. When the data is sampled at a higher frequency, say, m-times in a day, we will denote the return over this interval as $r_{t,t-1/m} = \log(P_t) - \log(P_{t-1/m})$. For instance, in our study, returns are sampled every five minutes between the trading hours of 9:30 am and 4:05 pm (corresponding to 80 five-minute intervals within a trading day), and we will write $r_{t,t-1/80} = \log(P_t) - \log(P_{t-1/80})$, which corresponds to the last 5-minute return of day $t - 1$.

Our goal is to predict a measure of volatility over some future horizon H, $V_{t+H,t}$. As a primary measure of volatility for the period t to t + H, we consider the increments in the quadratic variation of the return process $Q_{t+H,t}$. We focus on predicting future realized volatility from one week (H = 5) to one month (H = 20) horizon. These are horizons that matter mostly for option pricing and portfolio management. Focusing on predicting future increments of quadratic variation also allows us to make our analysis directly comparable with a large body of existing literature. The quadratic variation is not observed directly but can be measured with some discretization error. One such measure would be the sum of (future) squared returns, namely PHm j=1[r(t+H)−(j−1)/m,(t+H)−(j−2)/m] 2 , which we will denote by Q˜ (Hm) t+H,t since it involves a discretization based on Hm intra-daily returns. The superscript in parentheses indicates the number of high-frequency data used to compute the variable. Besides

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

increments in quadratic variation, we also consider log Q˜ (Hm) t+H,tas a target variable to forecast. Previous papers, including Andersen et al. (2003), have observed that forecasting the log transformation yields better in- and out-of-sample forecasts of the variance as it puts less weight on extreme realizations of the quadratic variation.

## 1 The MIDAS Specification:

A daily MIDAS volatility model is a regression model

$$V_{t+H,t}^{(Hm)} = \mu_H + \phi_H \sum_{k=0}^{k^{max}} b_H(k,\theta)\tilde{X}_{t-k,t-k-1}^{(m)} + \varepsilon_H$$

where V (Hm) t+H,t is a measure of (future) volatility such as Q˜ (Hm) t+H,t or log Q˜ (Hm) t+H,t. Specification (1.1) has three important features when compared to other models of conditional volatility (discussed below). First, the volatility measure on the left-hand side, V (Hm) t+H,t , and the variables on the right-hand side, X˜ (m) t−k,t−k−1 , might be sampled at different frequencies. Second, the polynomial lag parameters bH are parameterized to be a function of θ, thereby

allowing for a longer history without a proliferation of parameters. Third, MIDAS regressions typically do not exploit an autoregressive scheme, so that X˜ (m) t−k,t−k−1 is not necessarily related to lags of the left hand side variable. Instead, MIDAS regressions are first and foremost regression models and therefore the selection of X˜ (m) t−k,t−k−1 amounts to choosing the best predictor of future quadratic variation from the set of several possible measures of past fluctuations in returns. MIDAS regressions could potentially involve more than one type of regressors, see Ghysels et al. (2004) for further discussion. MIDAS regression models may also be nonlinear and indeed some of the regressors we will consider may provide better results with nonlinear specifications. For instance, Engle and Gallo (2003) provide interesting results along these lines. For simplicity, in this paper we consider only a single regressor linear MIDAS setting.

## Sampling at Different Frequencies :

In equation (1.1), the volatility is measured at weekly, bi-weekly, tri-weekly, and monthly frequency, whereas the forecasting variables X˜ (m) t−k,t−k−1 are available at higher frequencies. For instance, we can use daily data to compute a forecast of next month's volatility (H = 22). In other words, the return volatility over the month of, say, April (from the close of the market during the last day of March to the close of the market during the last day of April) will be forecasted with daily data up to the last day of March. But we could also use hourly or five-minute data to form monthly volatility forecasts. Thus, our model allows us not only to forecast volatility with data sample at different frequencies, but also to compare such forecasts and ultimately to see whether Merton's (1980) well-known continuous asymptotic arguments hold up in practice

In general, the MIDAS framework allows us to investigate whether the use of high-frequency data necessarily leads to better volatility forecasts at various horizons. These issues have motivated much of the recent literature on high-frequency data, see Andersen et al. (2001, 2002, 2003), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2001, 2002a,b, 2003a), among others. In some cases, the right-hand side variables are computed using m high-frequency returns, in which case they are denoted by a superscript (m). For instance, if we want to compute a monthly forecast of volatility using lagged daily

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

volatility estimates obtained from five-minute data, m would be equal to 80, the number of five-minutes in a day.

The MIDAS volatility models allow for a great degree of flexibility. For the sake of systematizing the results, in the next section we consider forecasts at weekly, bi-weekly, tri-weekly, and monthly frequency using daily data. In a subsequent section, we will turn to intra-daily regressors.

**Parsimony of Parameterization:**

Another distinguishing feature of (1.1) is that the lag coefficients bH (k, θ) (weights) are not unrestricted parameters. Rather they are parameterized as a function of θ, where θ is a small-dimensional vector. A flexible parameterization is an important element in the MIDAS specification, as the inclusion of high-frequency data might imply a significant increase in the number of lagged forecasting variables and unrestricted parameters to estimate. It allows us to dramatically reduce the number of parameters to estimate, which is particularly relevant in estimating a persistent process, such as volatility, where distant X˜ (m) t−k,t−k−1 are likely to have an impact on current volatility.

Even with daily forecasting variables, the unrestricted specification of the weights results in a lot of parameters to estimate. The problem only worsens with higher-frequency data. As we will see below, a suitable parameterization bH(k, θ) circumvents the problem of parameter proliferation and of choosing the truncation point k max . Hence, the parameterization bH (k, θ) is one of the most important ingredients in a MIDAS regression.

The parameterization also allows us to compare MIDAS models at different frequencies as the number of parameters to estimate will be the same even though the weights on the data and the forecasting capabilities might differ across horizons. We don't have to adjust our measures of fit for the number of parameters. In all estimations, we have either one or two parameters determining the pattern of the weights, the former being the case when we restrict our attention to θ1 = 1 and only estimate θ2 > 1. To illustrate the issue of parameter proliferation, consider Figure 1. It displays the estimated unconstrained parameters of equation (1.1) for lags up to 10 days. The figure contains results from various regressors X˜ (m) t−k,t−k−1 , such as Q˜ (m) t−k,t−k−1 , as well as absolute daily returns, daily range, and daily realized power, all of which we discuss at length below. We notice from the results displayed in the figure that the parameter estimates appear to be erratic as the lag increases. Hence, volatility models such as (1.1), whose weights are not tightly parameterized, do only well even with a small number of lags and almost surely will produce poor out of sample forecasts. It must be noted that the robust performance of ARCH/GARCH models can largely be attributed to capturing the dynamics of a large number of past shocks with only a few parameters. This basic idea is also the insight behind the MIDAS regressions.

**Various Regressors:**

In the MIDAS volatility model (1.1), X˜ (m) t−k,t−k−1 can be any variable that has the ability to forecast Q˜ (Hm) t+H,t. To put it differently, MIDAS volatility regressions can involve X˜ (m) t−k,t−k−1

**ISSN: 2278-4632**
**Vol-13, Issue-06, June 2023**

other than past squared returns or past realized volatility, which are the usual regressors considered in the autoregressive conditional volatility literature. The MIDAS approach puts us in the mind set of regression analysis and prompts us to explore various regressors that have the potential to predict future volatility. A number of predictors other than past realized volatility or squared returns have been proposed in various models. Unlike MIDAS, however, these models are typically autoregressive in nature. The MIDAS setup allows us to compare the forecasting ability of different X˜ (m) t−k,t−k−1 's and to choose the model with the best forecasting ability.

In the context of forecasting the quadratic variation Q˜ (Hm) t+H,t, we consider the following regressors.

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^Q_H + \phi^Q_H \sum_{k=0}^{k^{max}} b^Q_H(k,\theta)\tilde{Q}^{(m)}_{t-k,t-k-1} + \varepsilon^Q_{Ht} \quad (1.3)$$

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^G_H + \phi^G_H \sum_{k=0}^{k^{max}} b^G_H(k,\theta)(r_{t-k,t-k-1})^2 + \varepsilon^G_{Ht} \quad (1.4)$$

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^a_H + \phi^a_H \sum_{k=0}^{k^{max}} b^a_H(k,\theta)|r_{t-k,t-k-1}| + \varepsilon^a_{Ht} \quad (1.5)$$

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^r_H + \phi^r_H \sum_{k=0}^{k^{max}} b^r_H(k,\theta)[hi - lo]_{t-k,t-k-1} + \varepsilon^r_{Ht} \quad (1.6)$$

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^p_H + \phi^p_H \sum_{k=0}^{k^{max}} b^p_H(k,\theta)\tilde{P}^{(m)}_{t-k,t-k-1} + \varepsilon^p_{Ht} \quad (1.7)$$

In equation (1.3), past Q˜ (m) t,t−1 are used to predict Q˜ (Hm) t+H,t. Examples of such models have been advocated by Andersen et al. (2001, 2002, 2003) and are discussed at length below. Specification (1.4) is a projection of Q˜ (Hm) t+H,t onto lagged daily returns and corresponds to the ARCH/GARCH class of models (under some parameter restrictions).6

Equations (1.5) and (1.6) involve projecting Q˜ (Hm) t+H,t onto past daily absolute returns and daily ranges, respectively, which are two alternative measures of volatility. Therefore they are natural candidate regressors in the MIDAS specification. It is often argued that in the presence of deviations from normality absolute values could be more robust than squared values for conditional variance estimation (see e.g. Davidian and Caroll (1987)) whereas the virtues of daily range have been explored most recently by Alizadeh et al. (2001) and Gallant et al. (1999). Typically, past absolute returns (ranges) are used to predict future absolute returns (ranges). In particular, when absolute returns (daily ranges) are considered, the autoregressive features of absolute returns (daily ranges) are studied and modeled (see e.g. Ding et al. (1993) for absolute returns and Alizadeh et al. (2001) for daily ranges). Hence, the exploration of alternative measures of volatility has been cast in the context of autoregressive schemes. Here we introduce absolute returns and ranges as alternative predictors and examine their success (relative to the other predictors) at predicting realized volatility. The MIDAS regression format makes this a relatively straightforward exercise.

The last regression (1.7) involves similar arguments using techniques and developments of more recent date. The preference for absolute returns is a subject that has received much attention recently, see in particular Barndorff-Nielsen and Shephard (2003b, 2004) and Woerner (2002). Recall that Q˜ (m) t,t+1 is defined as the sum of m intra-daily squared returns. Instead of taking squared returns, say every five minutes, Barndorff-Nielsen and Shephard suggest to consider the sum of high-frequency absolute returns, or the so-called "realized power" variation P˜ (m) t+1,t, which is defined as Pm j=1 |rt−(j−1)/m,t−(j−2)/m|.

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

Regression (1.7) projects future realized volatility on past daily realized power.

Finally, to forecast the log of the quadratic variation, we consider log transformations of the five regressors in equations (1.3-1.7). In this fashion, our results would be directly comparable with those in the previous literature. To summarize, the MIDAS framework offers the ability to mix data sampled at different frequency, combined with a tightly parameterized model that allows different regressors to forecast volatility.

**Comparison of MIDAS with Other Volatility Models :**

TO further understand the flexibility of the MIDAS volatility models, it useful to compare them with other widely used models of conditional volatility, which for the purpose of presentation will be written as:

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^{Q}_{H} + \phi^{Q}_{H} \sum_{k=0}^{k^{max}} b^{Q}_{H}(k) \tilde{Q}^{(m)}_{t-k,t-k-1} + \varepsilon^{Q}_{Ht} \qquad (1.8)$$

$$\tilde{Q}^{(Hm)}_{t+H,t} = \mu^{G}_{H} + \phi^{G}_{H} \sum_{k=0}^{k^{max}} b^{G}_{H}(k;\theta) (r_{t-k,t-k-1})^{2} + \varepsilon^{G}_{Ht} \qquad (1.9)$$

In equation (1.8), past Q˜ (m) t,t−1 are used to predict Q˜ (Hm) t+H,t and the weights b Q H(k) are not parameterized. When H = 1 such models for so called realized volatility, analyzed by Andersen et al. (2001, 2003), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2001, 2002a,b, 2003a), and Taylor and Xu (1997), often rely on Merton's (1980) arguments that arbitrarily accurate estimates of volatility can be obtained by increasingly finer sampling of returns. The above papers show that the use of high-frequency data is beneficial in predicting volatility. Again, when H = 1, the difference between (1.8) and (1.3) is the specification of the weights b Q H . In this regard, it is important to note that Andersen et al. (2003) advocate the use of long memory models to parsimoniously parameterize the weights. In particular, they consider models of the following type:

$$(1 - \sum_{k=1}^{5} b_{A}(k) L^{k})(1 - L)^{d} \log \tilde{Q}^{(m)}_{t+1,t} = \mu + \varepsilon_{t} \qquad (1.10)$$

Hence, using a fractional differencing approach one can capture with a single parameter d slowly decaying response patterns associated with long memory. In addition to the fractional differencing parameter d in equation (1.10) Andersen et al. (2003) advocate the use of an AR(5) autoregressive expansion appearing on the left hand side of the equation. Hence, a total of 6 parameters (not including the constant) are used to model the autoregressive dynamics of realized volatility. Model (1.10) is expressed in terms of log volatility as Andersen et al. argue that the log transformation induces normality and therefore justifies the use of linear autoregressive models. This model will be our benchmark for all in-sample and outof-sample forecast comparisons and is henceforth referred to as the "ABDL" model.

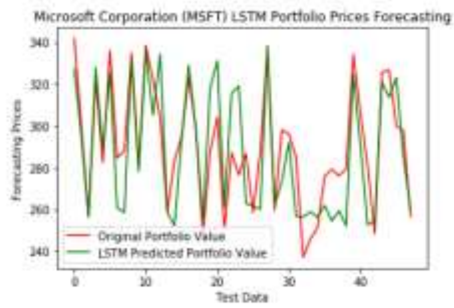
It is important to stress the differences between MIDAS regression models and the benchmark ABDL ARFI(5,d) model specification appearing in equation (1.10). None of the MIDAS regressions operate through autoregression whereas the ABDL specification follows much closer the tradition of ARCH-type models since Q˜ (m) t+1,t is projected onto its lagged values.

Furthermore, MIDAS regressions involve at most two parameters for the Beta polynomial, a scaling parameter and an intercept, i.e. less than the typical ABDL setting. The challenge is to outperform the ABDL specification while choosing: (1) the type of regressors; and (2) the decay patterns through judicious choice of parameterizations of the polynomial weighting schemes. The success of this challenge is the main argument of this paper.7

Finally, it should be noted that equation (1.9) is a version of the most widely used specification of conditional volatility, namely, the ARCH-type models of Engle (1982) (see also Bollerslev (1986)). In (1.9), future volatility is projected onto lagged daily squared returns (and a constant), where the weights are tightly parameterized via an autoregression such as in the popular GARCH(1,1) specification. Andersen et al. (2003) show that models appearing in equation (1.10) outperform ARCH-type models, which is why we use the former as a benchmark.

Results:


In above screen in blue colour text we can see LSTM MSE (mean square error) is 66 and in graph X-axis represents number of days and y-axis represents asset values and red line represents original asset value and green line represents Predicted value. In below screen we can see output for MDT dataset




In above screen we can see MSE value as 328 on MSFT dataset and below is the TJX dataset output

## CONCLUSION:

We study the predictability of return volatility with MIDAS regressions. Our approach allows us to compare forecasting models with different measures of volatility, frequencies, and lag lengths. While the main focus of this paper is volatility forecasting, it is clear that the MIDAS framework is general in nature and can find a good use in any empirical investigation that involves data sampled at different frequencies. Simplicity, robustness, and parsimony are three of its main attributes.

We report several intriguing findings regarding the predictability of weekly to monthly realized volatility in equity markets. First, we find that daily realized power outperforms daily realized volatility and that daily and intra-daily absolute returns outperform respectively daily and intra-daily squared returns. This set of results suggests that absolute returns are very successful at capturing fluctuations in future return volatility, despite the predominant emphasis in the literature on squared returns. Also, we find that daily ranges are extremely good forecasters of future volatility and are only second to realized power. This last finding is consistent with results in Gallant et al. (1999), Alizadeh et al. (2002) and Engle and Gallo (2003), among others, who use different methods and different data. Finally, we show

ISSN: 2278-4632
Vol-13, Issue-06, June 2023

that the direct use of high-frequency data does not necessarily lead to better volatility forecasts.

## REFERENCES :

[1]. Heterogeneous Information Arrivals and Return Volatility: survey by Andersen, T. and T. Bollerslev [2016]

[2]. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts: survey by Andersen, T. and T. Bollerslev [2012]

[3]. The Distribution of Exchange Rate Volatility: survey by Andersen, T., T. Bollerslev, F. X. Diebold and P. Labys[2007]

[4]. Modeling and Forecasting Realized Volatility: survey by Andersen, T., T. Bollerslev, F. X. Diebold and P. Labys[2018]

[5]. Some Like it Smooth, and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modeling, and Forecasting Asset Return Volatility: survey by Andersen, T.G., Bollerslev, T. and F.X. Diebold[2020]

[6]. Rolling sample volatility estimators: some new theoretical, simulation and empirical results: survey by Andreou, E. and E. Ghysels[2002]

[7] Andersen, T., T. Bollerslev, F. X. Diebold and P. Labys (2003), "Modeling and Forecasting Realized Volatility", Econometrica, 71, 529-626.

[8] Andersen, T.G., Bollerslev, T. and F.X. Diebold, (2003), "Some Like it Smooth, and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modeling, and Forecasting Asset Return Volatility," Manuscript, Duke, Northwestern and University of Pennsylvania.

[9] Andreou, E. and E. Ghysels (2002), "Rolling sample volatility estimators: some new theoretical, simulation and empirical results" Journal of Business and Economic Statistics, 20, 3, 363-376.

[10] Baillie, R.T., T. Bollerslev and H. O. Mikkelsen (1996), "Fractionally integrated generalized autoregressive conditional heteroskedasticity" Journal of Econometrics, 74, 3-30.

[11] Bandi, F.M. and J.R. Russell,(2003), "Microstructure noise, realized volatility, and optimal sampling", Working paper, Graduate School of Business, The University of Chicago.

[12] Barndorff-Nielsen, O. and N. Shephard (2001), "Non-Gaussian Ornstein–Uhlenbeckbased models and some of their uses in financial economics (with discussion), Journal of the Royal Statistical Society, Series B, 63, 167-241.