Enhancing the HAR model in an index-specific context

T.R. Janssen (452549)

ERAMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics *

Bachelor Thesis BSc[2] Econometrics/Economics

Supervisor: K.P. de Wit MSc

Second assessor: dr. X. Xiao

July 5, 2020

Abstract

This research extends the work of Corsi (2009) by evaluating an alternative HAR model specification in order to suit the dynamics of indices. In addition, several forecasting and estimation approaches are applied order to evaluate whether they produce more accurate one-day-ahead, one-week-ahead and two-week-ahead out-of-sample forecasts than the original HAR model from Corsi (2009). The following methods are applied: a parametric and a non-parametric weighted least squares scheme, two penalized linear models and a forecast combination technique. We conclude that the weighted least squares schemes yield the largest forecast improvements in comparison with the methodology applied by Corsi (2009). In some contexts, the penalized linear models provide some improvements of the forecasts, whereas applying the forecast combination technique yields forecasts which are inferior to the forecasts of the original HAR model.

*The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Contents

1 Introduction

Volatility forecasting is a widely studied subject to which the paper of Corsi (2009) has made a significant contribution. His paper presents the Heterogeneous Autoregressive model of Realized Volatility (HAR model). He specifies the HAR(1,5,22) model in which one-day-ahead volatility is estimated by means of an ordinary least squares (OLS) regression with three regressors: the realized volatility of the previous day, the average daily realized volatility of the previous five days and the average daily realized volatility of the previous 22 days. This research extends the work of Corsi (2009) by evaluating another HAR specification in order to suit the dynamics of indices. In addition, several alternative forecasting and estimation methods are applied to the HAR model in order to evaluate whether they produce more accurate out-of-sample forecasts than the original HAR model from Corsi (2009).

Being able to forecast volatility is essential for investors with respect to the pricing of financial derivatives and risk management. The ability to model and forecast volatility more precisely gives investors an evident advantage with respect to making investment decisions. More profitable investment decisions yield private investors a larger return on their equity and strengthen the financial position for large investment funds, such that for example a larger indexation of pensions can take place. Additionally, a more prudent risk model may help investors weigh their portfolios in accordance with their personal preference of risk. This could for instance be beneficial for individuals who save their pensions individually. All in all, finding improvements of the original HAR model from Corsi (2009) could help investors to expand their forecasting and risk-management abilities.

Multiple extensions to the HAR model have been presented in academic literature. This thesis extends the literature by considering the HAR model in a solely index-specific context with an HAR(1,2,5,10,22) specification as suggested but not yet evaluated by Corsi (2009). Additionally, this research provides extensions to the original HAR model by combining the original insights of Corsi (2009) about heterogeneous influences of volatility during different time horizons with existing methodologies of constructing weighted least squares schemes (Clements & Preve, 2019), penalized linear models (Audrino, Huang, & Okhrin, 2018) and forecast combinations (Rapach, Strauss, & Zhou, 2010).

This thesis aims to answer the following research question: "Can we enhance the HAR(1,5,22) model from Corsi (2009) given that it improves forecasts in an index-specific context?" In order to answer this research question we explore whether it is more appropriate to apply a HAR(1,2,5,10,22) specification rather than a HAR(1,5,22) specification when forecasting indices. Additionally, we evaluate whether several estimation and forecast procedures as alternatives to OLS can generate more accurate forecasts than original OLS estimated HAR(1,5,22) model. For this purpose we evaluate whether applying several weighted least squares (WLS) schemes, penalized linear models and forecast combination techniques to the HAR(1,2,5,10,22) model yield improvements of the forecasts based on the methodology from Corsi (2009).

We find that with respect to modelling the Standard & Poor's 500 index, the index-specific HAR(1,2,5,10,22) model is the most appropriate, whereas the original HAR(1,5,22) model form Corsi (2009) is more appropriate when forecasting the Financial Times Stock Exchange index and Amsterdam Exchange index. With respect to alternative forecasting approaches applied on the HAR(1,2,5,10,22) model, we conclude that WLS schemes yield the highest forecast improvements while penalized linear methods perform comparably to OLS based HAR estimations and Discounted Mean Squared Forecast Error forecast combinations perform worse than the original HAR(1,5,22) forecasts from Corsi (2009).

In the following section we present a literature review in which we explore existing extensions to the HAR model and we discuss previous applications of weighted least squares, penalized linear models and forecast combinations. Consequently, we explain which data is under consideration in this paper, where we retrieve this data from and how we transform it. Afterwards, the precise methodologies which we apply in this paper are outlined. We present and interpret our findings in the Results section, after which we draw conclusions, reflect on the limitations of this research and make some suggestions for future research in the Conclusion section.

2 Literature Review

As mentioned in the introduction, the HAR(1,5,22) model as proposed by Corsi (2009) has been of major influence. The model entails a regression where volatility forecasts are based on the realized volatility of the preceding day and the average of the daily realized volatility during the preceding week and month. The HAR(1,5,22) model can be considered as an autoregressive model with 22 lags to which several restrictions are imposed in order to establish the heterogeneous influence of different time horizons. The heterogeneous structure of this model is in line with the Heterogeneous Market Hypothesis as presented by Müller et al. (1997). The three time horizons which are distinguished by Corsi (2009) are derived from the empirically relevant daily, weekly & monthly investment horizons.

The realized volatility of these three investment horizons appears to be of predictive relevance with respect to all three financial derivatives under consideration in his paper: S&P 500 index futures, CHF/USD exchange rates and one-month US treasury bills.

In the context of indices, Corsi (2009) suggests that it might be beneficial to investigate an HAR(1,2,5,10,22) model specification as well, where a two-day and biweekly realized volatility term are added to the original HAR(1,5,22) model specification. All in all, the HAR model is relatively simple since it is estimated by a trivial OLS regression and estimates only four parameters. Corsi (2009) compares the in-sample and out-of-sample prediction performance of the HAR(1,5,22) model with AR models of lag order 1 and 3 and finds that the HAR(1,5,22) model provides a superior performance to these autogressive models. In addition, the author makes the same predictions with an ARFIMA model. Corsi (2009) draws the conclusion that the quality of the HAR(1,5,22) predictions is comparable to that of the ARFIMA model or slightly better. However, the ARFIMA model is relatively complex to estimate such that the HAR(1,5,22) model can be considered a welcoming improvement.

In the conclusion of his paper, Corsi (2009) provides some interesting suggestions for further research with respect to the HAR model. One of these extensions entails including different measures of jumps and heterogeneous leverage. Corsi (2009) notes that, although the HAR model is a linear model, the data generating process of volatility may contain some nonlinearities which the original HAR model cannot capture. Corsi and Reno (2009) resolve this issue by taking into account additional determinants of volatility dynamics: lagged negative returns and lagged jumps. Including negative returns in the model enables the authors to incorporate the leverage effect. This is the empirical effect that volatility tends to increase to a larger extent after a negative shock than after a positive shock of a comparable magnitude. Some of the papers which discuss this phenomenon are Christie (1982) & Bollerslev, Litvinova, and Tauchen (2006). Additionally, Corsi and Reno (2009) prove that the lagged negative returns also exhibit features of a daily, weekly & monthly heterogeneous structure. All features which are evaluated by Corsi and Reno (2009) appear to be of significant influence on the future volatility and therefore provide a useful extension to the original HAR model.

In spite of the fact that the aforementioned extensions of the original HAR model do improve the forecast accuracy of the HAR model, they also complicate the model in some sense since they take into account additional features of the nature of the data by adding returns and jumps. It would be interesting to explore whether we can improve the forecasting quality of the HAR model from Corsi (2009) while maintaining its relative simplicity. In academic literature, there are multiple methodologies which may help to reweigh the influence of daily, weekly and monthly HAR components, as well as a two-day and biweekly component.

One of the papers which proposes several methods for this purpose is Clements and Preve (2019). This paper is centered around evaluating several methodologies which might make an improvement to the HAR(1,5,22) model as introduced by Corsi (2009). Apart from the original HAR model, they estimate the HARQ model as proposed by Bollerslev, Patton, and Quaedvlieg (2016). This model accounts for the errors with which the realized variances are estimated by making use of the realised quadracity. In addition, Clements and Preve (2019) discuss alternatives to OLS with respect to estimating the coefficients of the terms in the original HAR model. Due to the fact that OLS estimates tend to be relatively sensitive to large outliers and volatility is relatively noisy, the authors propose applying the least absolute deviation approach. This consists of minimizing the sum of the absolute errors rather than the squared errors as in OLS. Lastly, Clements and Preve (2019) propose four WLS weighting schemes of which two are parametric and two are non-parametric. The paper displays the predictive accuracy of each of the models and estimation methods in comparison with the original HAR model as presented in Corsi (2009). Clements and Preve (2019) show that for each of the forecast horizons, at least one alternative method outperforms the OLS estimated HAR model. In general, the WLS schemes appear to yield the largest forecast improvements among all evaluated alternatives.

Another methodology which enables us to select relevant autoregressive predictors with respect to index volatility is established by Audrino and Knaus (2016). They apply adaptive Least Absolute Shrinkage and Selection (LASSO) techniques in order to evaluate the appropriateness of the terms in the original the HAR model. Additionally, they evaluate several alternative lag structures and conclude that the structure of the original HAR(1,5,22) model is not always appropriate. The authors of this paper observe two structural breaks of the realized variance dynamics in 2006 and after the financial crisis of 2008. However, despite the the fact that title of the paper by Audrino and Knaus (2016) suggests that they apply adaptive LASSO on the HAR model, they do not apply adapative LASSO on the actual HAR model but on an unrestricted AR(22) model. Audrino et al. (2018) recognise this and again apply adaptive LASSO on an AR(22) model, but now with restrictions such that the autoregressive lags are converted into HAR type lags. All in all, applying adaptive LASSO to the HAR model proves to yield forecasting improvements in some contexts such as the group LASSO where multiple lags are clustered. Apart from the LASSO, there are different loss penalties which might help us construct more accurate predictions. In their paper, Gu, Kelly, and Xiu (2020) apply elastic net and ridge penalties in addition to LASSO when forecasting returns with economic variables as regressors. It might be interesting to apply such penalties in the context of HAR models as well.

A different methodology which might yield some benefits is the econometric technique of forecast combination. Forecast combinations are widely applied in the literature and renowned for their improvement of out-of-sample predictions and reduced volatility among forecasts. Stock and Watson (2004) have empirically applied this technique relatively early. They use numerous indicators to construct forecast combinations to predict GDP growth. A paper which has established the added value of forecast combinations more recently is Rapach et al. (2010). In this paper, 15 economic variables are used to make univariate predictions of the equity premium. Afterwards, forecast combinations of these univariate predictions are constructed. Huang and Lee (2010) analytically prove that under some conditions in finite samples the quality of such forecasts is better than forecasts based on a multiple regression in which all variables are included. The Discounted Mean Squared Forecast Error (DMSFE) forecast combination is one of the most successful forecast combination techniques applied by Rapach et al. (2010). This methodology bases the weights of the individual forecasts on the predictive quality of the univariate forecasts during the period before the out-of-sample period, which is an insight from Diebold and Pauly (1990).

Taking everything into account, there are numerous options to enhance the HAR model in an index-specific context. In this report, we evaluate an index-specific specification of the HAR model. Additionally, we apply several estimation and forecasting methodologies in order to explore whether they succeed in enhancing the original HAR model as introduced by Corsi (2009).

3 Data

The data under consideration is retrieved from the Oxford Realized Library. We retrieve the realized daily variances of three different indices during the time period January 2000 until May 2020. The first index is the Standard & Poor's 500 (SPX) index. This index consists of the 500 companies with the largest market capitalization listed on either the New York Stock Exchange or NASDAQ. Furthermore, we consider the Financial Times Stock Exchange (FTSE) index, which contains the 100 companies with the biggest market capitalization on the London Stock Exchange. Lastly, we evaluate the Amsterdam Exchange (AEX) index, containing the 25 shares with the largest market capitalization listed on Euronext Amsterdam. The dataset of the realized daily variances of the SPX, FTSE & AEX indices contain 5119, 5145 & 5200 daily observations respectively.

The retrieved realized daily variances are calculated on the basis of 5-minute intraday returns. Mathematically notated, that is:

$$RVar_t^{(d)} = \sum_{j \in \mathcal{J}_t} (X_{(j,t)} - X_{(j-1,t)})^2 \tag{1}$$

Here $RVar_t^{(d)}$ equals the realized daily variance at day t. (j,t) is a 5-minute interval in the complete set of 5-minute intervals $\mathcal{J}_t$ on day t, $X_{(j,t)}$ is the value of the index at the end of the 5-minute interval j at day t and $X_{(j-1,t)}$ entails the index value five minutes before the end of the 5-minute interval j at day t. Consequently, we make the following transformation to the realized daily variance in order to obtain the annualized daily realized volatility:

$$RV_t^{(d)} = \sqrt{252 * RVar_t^{(d)}} * 100 \tag{2}$$

Here $RV_t^{(d)}$ is the annualized realized daily volatility in percentages. $RVar_t^{(d)}$ is the five minute realized daily variance as retrieved from the Oxford Realized Library. This is multiplied by the number of trading days in a year of 252 such that we obtain the annualized variance. Lastly, the square root of the annualized realized variance is multiplied by 100 in order to obtain the annualized realized daily volatility in percentages. From now onwards, we refer to annualized realized daily volatility as realized volatility. For the indices SPX, FTSE and AEX we display the mean, median, standard deviation, skewness and kurtosis in Table 1.

Table 1: Descriptive statistics of realized volatility

	SPX	FTSE	AEX
Mean	13.5525	14.6066	14.5054
Median	10.9303	11.9311	11.8904
Standard Deviation	9.9344	9.7726	9.3383
Skewness	3.3444	3.8444	2.8115
Kurtosis	22.2348	31.9658	15.9401

Generally speaking, there are no major differences among the descriptive statistics of the SPX, FTSE and AEX index. For all indices the mean is higher than the median, which is to be expected since volatility has a zero lower bound whereas theoretically there is no upper bound. The positive skewness of each series is also an indication of this feature. Additionally, we observe a kurtosis which is substantially larger than the value of 3 which is expected with a normal distribution, indicating a distribution with relatively fat tails.

In line with Corsi (2009) we construct the realized volatility over multiple time horizons in order to obtain all regressors of the HAR(1,5,22) model. The weekly realized volatility $RV_t^{(w)}$ and the monthly realized volatility $RV_t^{(m)}$, which are the averages of the daily realized volatility over their respective time periods, are constructed as follows:

$$RV_t^{(w)} = \sum_{i=0}^{4} \frac{1}{5} RV_{t-i}^{(d)} \tag{3}$$

$$RV_t^{(m)} = \sum_{i=0}^{21} \frac{1}{22} RV_{t-i}^{(d)} \tag{4}$$

Corsi (2009) mentions that, with respect to modelling indices, a two-day and biweekly lag might also be of some relevance. The two-day realized volatility $RV_t^{(t)}$ and biweekly volatility $RV_t^{(b)}$ are constructed as follows:

$$RV_t^{(t)} = \sum_{i=0}^{1} \frac{1}{2} RV_{t-i}^{(d)} \tag{5}$$

$$RV_t^{(b)} = \sum_{i=0}^{9} \frac{1}{10} RV_{t-i}^{(d)} \tag{6}$$

We now remove the first 21 observations from the sample for each index. We do this because only from observation 22 onwards we can construct all regressors, since the calculation of $RV_t^{(m)}$ requires the 21 observations before time t. Now that we have calculated the transformations and removed the data we need not use anymore, we establish the autocorrelation coefficients of $RV_t^{(d)}$ with $RV_{t-1}^{(d)}$, $RV_{t-1}^{(t)}$, $RV_{t-1}^{(w)}$, $RV_{t-1}^{(b)}$ & $RV_{t-1}^{(m)}$ for each of the three considered indices in Table 2.

Table 2: Autocorrelation coefficients with $RV_t^{(d)}$

	SPX	FTSE	AEX
$RV_{t-1}^{(d)}$	0.8181	0.7358	0.8414
$RV_{t-1}^{(t)}$	0.8417	0.7688	0.7688
$RV_{t-1}^{(w)}$	0.8306	0.7799	0.7799
$RV_{t-1}^{(b)}$	0.8056	0.7642	0.7642
$RV_{t-1}^{(m)}$	0.7454	0.7059	0.7059

The AEX index exhibits decaying autocorrelation, whereas the autocorrelation coefficient of the one-day-ahead realized volatility $RV_t^{(d)}$ is larger with $RV_{t-1}^{(t)}$ than $RV_{t-1}^{(d)}$ for the FTSE and AEX index while further lags show decaying autocorrelation as well. This is an indication that we can expect that the influence of the realized volatility over longer time periods is globally less influential with respect to determining the one-day-ahead forecast than the more recent volatility components.

4 Methodology

For each model specification described in the Methodology section, we make one-day-ahead in-sample forecasts for each of the three indices in order to evaluate the in-sample forecast performance. Additionally, we make one-day-ahead, one-week-ahead, and two-week-ahead out-of-sample forecasts for each of the indices based on the outlined model specifications and forecasting approaches. For the in-sample period we consider a rolling window of length 1000 for each estimation. Such a rolling window contains 979 observations, since the data transformation makes the longest lag $RV_t^{(m)}$ contain information about the 21 preceding days as well. The one-week-ahead and two-week-ahead forecasts are constructed by making five and ten consecutive one-day-ahead forecasts respectively. When forecasting volatility over a longer time horizon, we use the same moving in-sample period of length 1000 as for the one-day-ahead forecasts. Hence we make four less forecasts for the one-week-ahead forecasts and nine less forecasts for the two-week ahead forecasts. Each of the subsections below starts with a brief explanation of the manner in which this section extends existing academic literature.

4.1 HAR models

To start with, we describe the HAR(1,5,22) as introduced by Corsi (2009). In line with the suggestion given in this paper that it might be beneficial to include the two-day realized volatility and the biweekly realized volatility with respect to modelling index volatility, we evaluate the HAR(1,2,5,10,22) model.

The HAR(1,5,22) model from Corsi (2009) is constructed as follows:

$$RV_t^{(d)} = c + \beta^{(d)} RV_{t-1}^{(d)} + \beta^{(w)} RV_{t-1}^{(w)} + \beta^{(m)} RV_{t-1}^{(m)} + \omega_t^{(d)} \tag{7}$$

Here $RV_{t-1}^{(d)}$, $RV_{t-1}^{(w)}$ & $RV_{t-1}^{(m)}$ are the realized daily, weekly and monthly volatility at time $t-1$ respectively. The coefficients c, $\beta^{(d)}$, $\beta^{(w)}$ & $\beta^{(m)}$ are estimated via OLS. These coefficients are used in order to construct the in-sample and out-of-sample forecasts. Furthermore, since we are evaluating data in an index-specific context, we construct the HAR(1,2,5,10,22) where we add the two-day realized volatility and the biweekly volatility as regressors to equation 7. Hence, we construct the HAR(1,2,5,10,22) model as follows:

$$RV_t^{(d)} = c + \beta^{(d)} RV_{t-1}^{(d)} + \beta^{(t)} RV_{t-1}^{(t)} + \beta^{(w)} RV_{t-1}^{(w)} + \beta^{(b)} RV_{t-1}^{(b)} + \beta^{(m)} RV_{t-1}^{(m)} + \omega_t^{(d)} \tag{8}$$

Here $RV_{t-1}^{(t)}$ & $RV_{t-1}^{(b)}$ are the two-day and biweekly realized volatility respectively. The coefficients c, $\beta^{(d)}$, $\beta^{(t)}$, $\beta^{(w)}$, $\beta^{(b)}$ & $\beta^{(m)}$ are estimated via OLS after which the forecasts are constructed.

4.2 Simulation

In line with Corsi (2009), we make a simulation of the returns and volatility based on an HAR(1,5,22) data generating process. Additionally, we simulate an HAR(1,2,5,10,22) model since this research studies volatility in an index-specific context. Afterwards, we compare these series with the actual daily returns and realized volatility of the SPX index. Corsi (2009) generates a simulation based on two-hour returns and volatility, but since we can only retrieve daily return and realized volatility data from the Oxford Realized Library, we choose to simulate daily returns and volatility such that we can make a more appropriate comparison with the realized values of the SPX index.

For the HAR(1,5,22) and HAR(1,2,5,10,22) volatility simulation we use equations 7 & 8 respectively. We retrieve the estimates of c, $\beta^{(d)}$, $\beta^{(t)}$, $\beta^{(w)}$, $\beta^{(b)}$ & $\beta^{(m)}$ from the estimations of equations 7 & 8 for the SPX index over the complete sample period of January 2000 until May 2020. The error term $\omega_t^{(d)}$ is normally distributed with a zero mean and a standard deviation which is equal to the standard deviation of the estimated residuals of the respective HAR estimations of the realized volatility of the SPX index in equations 7 & 8.

We initialize the simulation with 22 observations which are equal to the mean of the realized volatility of the SPX index to which we add the random error term $\omega_t^{(d)}$ with the distribution mentioned above. Afterwards, we construct the one-day-ahead volatility according to the data generating process described in equation 7 & 8 with the normally distributed error $\omega_t^{(d)}$. Since volatility is non-negative by definition, we must ensure that adding the error term $\omega_t^{(d)}$ does not lead to a negative one-day-ahead volatility. If this occurs in the simulation, we generate a new random normally distributed error until the one-day-ahead realized volatility turns out to be positive.

While this is by no means a perfect approach, Corsi (2009) does not provide a clear explanation of his methodology and we believe that for now this is the most appropriate way to incorporate the random errors in order to determine the one-day-ahead volatility. In order to reduce the influence of the initialization values, we generate the realized volatility for twice the sample size of the available SPX data. When we make the comparison between the realized returns and volatility of the SPX series and both simulated HAR series, we only use the last half of the simulated values for either of the HAR data generating processes. When the realized volatility is established, we simulate the returns. In line with Corsi (2009) we generate the simulated returns with the following formula:

$$r_t^{(d)} = RV_t^{(d)} \epsilon_t^{(d)} \tag{9}$$

Here $r_t^{(d)}$ is the simulated daily return on day t, $RV_t^{(d)}$ equals the respective one-day-ahead realized volatility in equations 7 & 8 and error term $\epsilon_t^{(d)}$ which follows that standard normal distribution.

4.3 Autogregressive models

In order to compare the forecasting performance of either of the HAR model specifications, we construct forecasts with autoregressive models of lag order 1 and 3 in line with Corsi (2009). The AR(1) model for realized volatility is specified as follows:

$$RV_t^{(d)} = c + b^{(1)} RV_{t-1}^{(d)} + \omega_t^{(d)} \tag{10}$$

The AR(3) model for realized volatility is specified as follows:

$$RV_t^{(d)} = c + b^{(1)} RV_{t-1}^{(d)} + b^{(2)} RV_{t-2}^{(d)} + b^{(3)} RV_{t-3}^{(d)} + \omega_t^{(d)} \tag{11}$$

As mentioned earlier, the HAR(1,5,22) and HAR(1,2,5,10,22) model can be considered as a restricted autogressive autogressive models of lag order 22. In order to evaluate whether each of the HAR model specifications are a more appropriate specification than the AR(22) model, we perform three goodness-of-fit tests over the complete sample period of each index. The first of these tests is the F-test for which the corresponding test statistic is calculated as follows:

$$F = \frac{\left(\frac{SSR_{HAR} - SSR_{AR22}}{k_{AR22} - k_{HAR}}\right)}{\left(\frac{SSR_{AR22}}{n - k_{AR22}}\right)} \tag{12}$$

Here SRR_{AR22} is the sum of the squared residuals of the AR(22) model, SSR_{HAR} is the sum of the squared residuals of either of the HAR model specifications. k_{AR22} equals 23 for the AR(22) model since one constant and 22 coefficients are estimated. For the HAR(1,5,22) model k_{HAR} equals 4 (3 coefficients plus one constant) and for HAR(1,2,5,10,22) k_{HAR} equals 6 (5 coefficients plus one constant). Lastly, n is the number of observations in the index. Furthermore, we calculate the Akaike (1974) information criterion (AIC) and the Schwarz et al. (1978) information criterion (SIC), which are defined as follows:

$$AIC = 2k - 2ln(\hat{L}) \tag{13}$$

$$SIC = k * ln(n) - 2ln(\hat{L}) \tag{14}$$

Here k is the number of parameters estimated, which equals 23 for the AR(22) model, 4 for the HAR(1,5,22) model and 6 for the HAR(1,2,5,10,22) model. Additionally, n equals the number of observations in the index and $\hat{L}$ is the maximum value of the log-likelihood of each particular model.

4.4 Weighted least squares

In this subsection, we apply two WLS schemes. Clements and Preve (2019) successfully apply this methodology in order to improve the HAR(1,5,22) model. The approach outlined below extends the literature by applying one non-parametric and one parametric WLS scheme from Clements and Preve (2019) to the index-specific HAR(1,2,5,10,22) model specification rather than the original HAR(1,5,22) model specification from Corsi (2009).

WLS estimation is equivalent to OLS estimation with the difference being that at each time t we multiply each term in the HAR regression with a weight w_t. The first approach we introduce is the non-paramtric approach. Initially, we estimate the HAR(1,2,5,10,22) coefficients with a moving in-sample period in the same manner as in the original model by means of OLS. Based on these coefficients we construct one-day-ahead forecasts such that we obtain the fitted values of the realized daily volatility at time t. In this methodology, the WLS Fitted weights w_t^{Fitted} are equal to $1/\hat{RV_t}^{(d)}$ where $\hat{RV_t}^{(d)}$ are the fitted values of the one-day-ahead volatility. The parametric WLS scheme we apply is slightly more trivial to estimate. Here the WLS Parameter weights $w_t^{Parameter}$ are equal to $1/RV_t^{(d)}$ where $RV_t^{(d)}$ is the realized daily volatility of at day t.

4.5 Penalized linear models

This section is an extension to the adaptive LASSO approach to the HAR model as presented by Audrino et al. (2018). We apply the adaptive LASSO procedure as introduced by Zou (2006) in order to perform variable selection among the five time horizons included in the HAR(1,2,5,10,22) model. In addition, we introduce adaptive elastic net penalties, where we apply the algorithm from Zou (2006) with an elastic net penalty as performed by Gu et al. (2020).

In line with Zou (2006), we apply the LARS algorithm from Efron, Hastie, Johnstone, Tibshirani, et al. (2004) in order to estimate the adaptive Lasso and adaptive elastic net coefficients. In this algorithm we follow three steps. In the first step we define $x_j^{**} = x_j/\hat{w}_j$ for every HAR(1,2,5,10,22) regressor j. Here x_j is the realized volatility of regressor j and $\hat{w}_j$ is the OLS estimate of regressor j. The second step involves the following LASSO problem for all λ_n:

$$\hat{\theta}^{**} = argmin_{\theta}||RV_t - \sum_{j=1}^{p} x_j^{**}\theta j||^2 + \lambda(1-\rho)\sum_{j=1}^{P}|\theta_j| + \frac{1}{2}\lambda\rho\sum_{j=1}^{P}\theta_j^2 \tag{15}$$

As a last step we construct the eventual coefficients for adaptive LASSO and adaptive elastic net. These are $\hat{\theta} = \hat{\theta}^{**}/\hat{w}_j$ for each regressor j. For the LASSO we have $\rho = 0$. An important characteristic of the LASSO penalty is that it forces less influential HAR coefficients to zero. For the elastic net we have $\rho = 0.5$ and this penalty captures features from both the LASSO operator and ridge regression. Furthermore, we perform 5-fold cross validation, which subdivides the in-sample period in five training and testing samples in order to determine the optimal λ_n. Our code uses the ADMM algorithm from Boyd et al. (2011) & Parikh and Boyd (2014) in order to minimize the least squares loss function. Eventually, for each forecast in the out-of-sample period we use the set of estimators $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$, $\hat{\theta}_4$ & $\hat{\theta}_5$ in order to construct the penalized linear method forecasts $\hat{RV}_t^{(PLM)}$ forecasts for the out-of-sample period:

$$\hat{RV}_t^{(PLM)} = \hat{\theta}_0 + \hat{\theta}_1 RV_{t-1}^{(d)} + \hat{\theta}_2 RV_{t-1}^{(t)} + \hat{\theta}_3 RV_{t-1}^{(w)} + \hat{\theta}_4 RV_{t-1}^{(b)} + \hat{\theta}_5 RV_{t-1}^{(w)} \tag{16}$$

4.6 DMSFE forecast combination

The forecast combination methodology we apply is the Discounted Mean Squared Forecast Error (DMSFE) forecast combination as applied by Stock and Watson (2004) and Rapach et al. (2010). In Rapach et al. (2010), this approach yields significantly better forecasts than an OLS regression which includes all the variables. Huang and Lee (2010) analytically prove that under some conditions in finite samples, combinations of univariate forecasts can deliver forecasts that are superior to a larger model which contains all individual variables. This paper extends the existing literature by applying the DMSFE forecast combination among terms of the an HAR model, in this case the HAR(1,2,5,10,22) model specification.

For this approach we divide the moving in-sample period in two parts: a training sample and a testing sample. The testing sample entails the last 100 observations of the moving in-sample period. The training sample contains all preceding observations in the in-sample period. We make an OLS estimation of the following five univariate equations 17, 18, 19, 20 & 21 during the training sample period:

$$RV_{(t,d)} = a^{(d)} + b^{(d)} RV_{t-1}^{(d)} + v_{(t,d)} \tag{17}$$

$$RV_{(t,t)} = a^{(t)} + b^{(t)} RV_{t-1}^{(t)} + v_{(t,t)} \tag{18}$$

$$RV_{(t,w)} = a^{(w)} + b^{(w)} RV_{t-1}^{(w)} + v_{(t,w)} \tag{19}$$

$$RV_{(t,b)} = a^{(b)} + b^{(b)} RV_{t-1}^{(b)} + v_{(t,b)} \tag{20}$$

$$RV_{(t,m)} = a^{(m)} + b^{(m)} RV_{t-1}^{(m)} + v_{(t,m)} \tag{21}$$

For each equation we construct a one-day ahead forecast for the first observation in the testing period. For the following forecast we make a new estimation in which we expand the training period with the observed volatility of the observation which had just been predicted. We use the forecasts during the testing period in order to ascribe weights to each of the five univariate forecasts. The eventual out of sample forecast is a linear combination of univariate forecasts which is of the following form:

$$R\hat{V}_{(t,FC)} = \pi_{(t,d)} R\hat{V}_{(t,d)} + \pi_{(t,t)} R\hat{V}_{(t,t)} + \pi_{(t,w)} R\hat{V}_{(t,w)} + \pi_{(t,b)} R\hat{V}_{(t,b)} + \pi_{(t,m)} R\hat{V}_{(t,m)} \tag{22}$$

Here $R\hat{V}_{(t,FC)}$ is the forecasted daily volatility in the forecast combination. $R\hat{V}_{(t,d)}$, $R\hat{V}_{(t,t)}$,

$\hat{RV}_{(t,w)}$, $\hat{RV}_{(t,b)}$ & $\hat{RV}_{(t,m)}$ are the forecasts of the daily volatility based on the univariate regression of daily, two-day, weekly, biweekly and monthly volatility respectively. The five weights $\pi_{(t,d)}$, $\pi_{(t,t)}$, $\pi_{(t,w)}$, $\pi_{(t,b)}$ & $\pi_{(t,m)}$ are determined during the testing period and are based on the quality of the individual univariate forecasts during this period. This approach uses the insight from Diebold and Pauly (1990) that weights ought to depend on the historical performance of the individual forecasts.

The individual weights of the DMSFE forecast combination are calculated as follows:

$$\pi_{t,i} = \phi_{t,i}^{-1} / \sum_{j \in \mathcal{J}_t} \phi_{t,j}^{-1} \tag{23}$$

Furthermore:

$$\phi_{t,i} = \sum_{s=m}^{M} \delta^{t-1-s} (RV_{s+1}^{(d)} - \hat{RV}_{s,i})^2 \tag{24}$$

Here m is the first observation in the testing period which corresponds to the particular out-of-sample forecast and M is the last observation in the testing period. Furthermore i and j are elements in the set $\mathcal{I}_t = \{\hat{RV}_{t,d}, \hat{RV}_{t,t}, \hat{RV}_{t,w}\hat{RV}_{t,b}, \hat{RV}_{t,m}\}$. Lastly, δ is the discount factor. In line with Stock and Watson (2004) we compute the DMSFE forecast combinations for $\delta = 1$ & $\delta = 0.9$.

4.7 Forecast evaluation

For the forecast evaluation, we broadly follow the same procedure as Corsi (2009). For each forecast we display the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). These two forecast performance measurements are calculated as follows:

$$RMSE = \sqrt{\frac{1}{P} \sum_{i=m+1}^{m+P} (\hat{RV}_i^{(d)} - RV_i^{(d)})^2} \tag{25}$$

$$MAE = \frac{1}{P} \sum_{i=m+1}^{m+P} |\hat{RV}_i^{(d)} - RV_i^{(d)}| \tag{26}$$

Here $m + 1$ is the time of the first forecast, P is the number of forecasts, $\hat{RV}_i^{(d)}$ is the forecast at time i and $RV_i^{(d)}$ is the observed value at time period i. Additionally, we calculate the R^2 value of the following Mincer and Zarnowitz (1969) regression for each of the forecast methods.

$$RV_t^{(d)} = b_0 + b_1 E_{t-1}[(\hat{RV}_t^{(d)}] + error \tag{27}$$

In this regression, the observed realized volatility $RV_t^{(d)}$ is regressed on a constant and the forecast for each forecast $\hat{RV}_t^{(d)}$. Applying this performance measurement in this context is in line with analysis from Andersen and Bollerslev (1998), Andersen, Bollerslev, and Diebold (2007) & Aït-Sahalia and Mancini (2008). When we perform the one-week-ahead and one-month-ahead forecasts the regression becomes:

$$\sum_{j=0}^{h} RV_{t+j}^{(d)} = b_0 + b_1 E_{t-1}[\sum_{j=0}^{h} \hat{RV}_{t+j}^{(d)}] + error, \tag{28}$$

with $h = 4$ for the one-week-ahead forecast and $h = 9$ for the two-week-ahead forecast.

5 Results

5.1 Estimated HAR coefficients

Table 3: Estimated coefficients of HAR(1,5,22)

	SPX		FTSE		AEX	
$\hat{c}$	0.8342	(3.2734)	1.0834	(3.8309)	0.8176	(4.5479)
$\hat{\beta}^{(d)}$	0.3825	(10.1028)	0.2539	(6.4208)	0.4229	(10.6935)
$\hat{\beta}^{(w)}$	0.4559	(7.0000)	0.5262	(7.4820)	0.4397	(9.2129)
$\hat{\beta}^{(m)}$	0.1000	(2.2916)	0.1458	(3.1452)	0.0808	(2.1115)

Notes. This table shows the results of the in-sample estimation of the ordinary least squares regression of the HAR(1,5,22) model for daily observations of the following three indices: S&P500 (5119 observations), Financial Times Stock Exchange (5145 observations) and Amsterdam Stock Exchange (5200 observations). The time horizon under consideration is January 2000 - May 2020. Between parentheses we report t-values which are based on heteroskedasticity-consistent standard errors.

In Table 3 we display the estimates of the coefficients of the HAR(1,5,22) model as well as the corresponding t-values. It is noteworthy that every estimated parameter for each of the three indices under consideration is significantly different from zero. For the SPX, FTSE & AEX index we observe that the parameter $\hat{\beta}^{(m)}$ is the lowest, implying that the realized monthly volatility is less influential than the realized daily and weekly volatility with respect to making one-day-ahead forecasts of the daily volatility. For the SPX and AEX indices $\hat{\beta}^{(d)}$ and $\hat{\beta}^{(w)}$ are of a comparable size, whereas the difference between $\hat{\beta}^{(d)}$ and $\hat{\beta}^{(w)}$ is substantially larger for the FTSE index where it holds that $\hat{\beta}^{(d)} < \hat{\beta}^{(w)}$.

In Table 4 we depict the estimates of the coefficients of the HAR(1,2,5,10,22) model as well as the

Table 4: Estimated coefficients of HAR(1,2,5,10,22)

	SPX		FTSE		AEX	
$\hat{c}$	0.8315	(3.4265)	1.1180	(3.9184)	0.8452	(4.6393)
$\hat{\beta}^{(d)}$	0.2048	(3.7325)	0.1768	(3.0066)	0.3101	(4.9630)
$\hat{\beta}^{(t)}$	0.3731	(4.6270)	0.1829	(1.9912)	0.2301	(2.7615)
$\hat{\beta}^{(w)}$	0.1476	(1.6419)	0.2587	(2.4109)	0.1870	(2.3321)
$\hat{\beta}^{(b)}$	0.1685	(1.5023)	0.2737	(2.2090)	0.2114	(2.2789)
$\hat{\beta}^{(m)}$	0.0446	(0.6926)	0.0313	(0.4629)	0.0030	(0.0562)

Notes. This table shows the results of the in-sample estimation of the ordinary least squares regression of the HAR(1,2,5,10,22) model for daily observations of the following three indices: S&P500 (5119 observations), Financial Times Stock Exchange (5145 observations) and Amsterdam Stock Exchange (5200 observations). The time horizon under consideration is January 2000 - May 2020. Between parentheses we report t-values which are based on heteroskedasticity-consistent standard errors.

corresponding t-values. Unlike for the estimation of the HAR(1,5,22) model, for the HAR(1,2,5,10,22) model not all coefficients for each index are significant. For the SPX index, we get estimates which are significantly different than zero for $\hat{c}$, $\hat{\beta}^{(d)}$ & $\hat{\beta}^{(t)}$, whereas for the FSTE and AEX index we obtain significant nonzero estimates for $\hat{\beta}^{(w)}$ and $\hat{\beta}^{(b)}$ as well, while for no index we obtain a significant nonzero estimation of $\hat{\beta}^{(m)}$. Furthermore, it is noteworthy that for each index the estimated value of constant $\hat{c}$ is is approximately equal for the the HAR(1,5,22) and HAR(1,2,5,10,22) model specification.

5.2 Simulation

Here we present the result of our simulation of an HAR(1,5,22) and HAR(1,2,5,10,22) model. As mentioned in the methodology, these simulations use the estimated coefficients which are displayed in Tables 3 & 4. The actual returns of the SPX index can be found in Figure 1, the simulated returns based on the HAR(1,5,22) model can be found in Figure 2 and the simulated returns constructed on the basis of the HAR(1,2,5,10,22) model are shown in Figure 3. Visually, there are no noteworthy differences between the HAR(1,5,22) and HAR(1,2,5,10,22) return simulations. There are however quite some differences between the realized daily returns of the SPX index and the simulated returns. The average absolute SPX returns appear to be lower than the absolute simulated returns, although the maximum and minimum of the daily realized SPX returns are respectively higher and lower than the maximum and minimum of either of the HAR simulations. We may conclude that unfortunately the simulation does not appear to yield an near-perfect replication of the data generating process

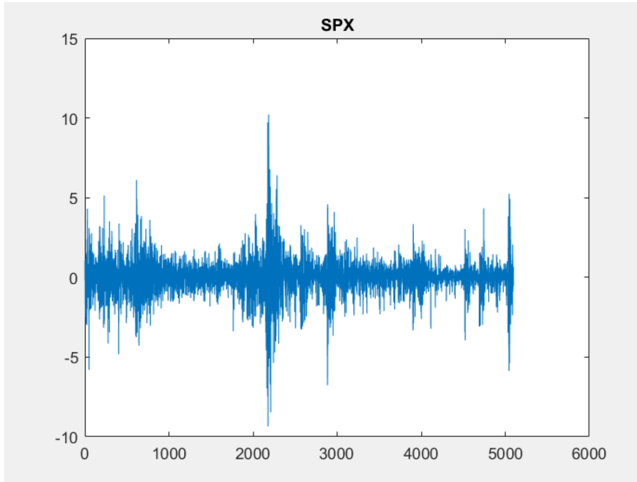

Figure 1: Realized daily returns of the SPX index

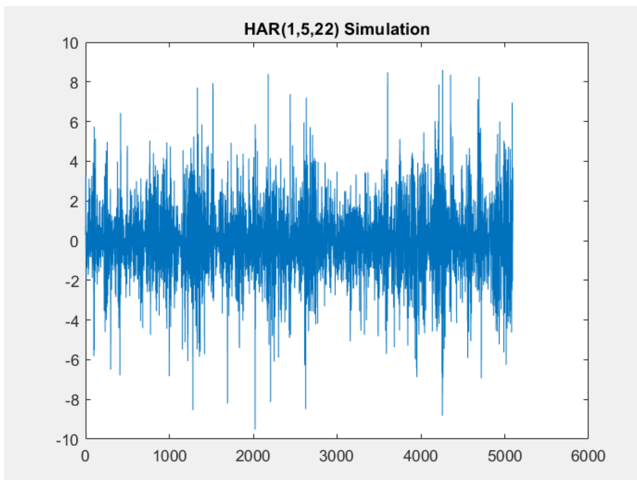

Figure 2: Simulated daily returns based on an HAR(1,5,22) data generating process

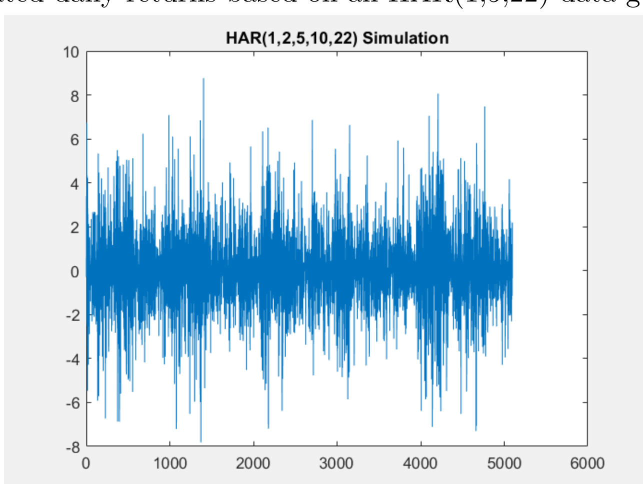

Figure 3: Simulated daily returns based on an HAR(1,2,5,10,22) data generating process

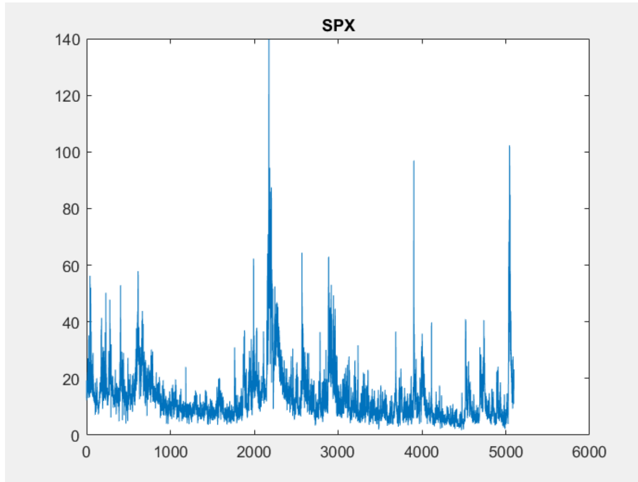

Figure 4: Realized daily volatility of the SPX index

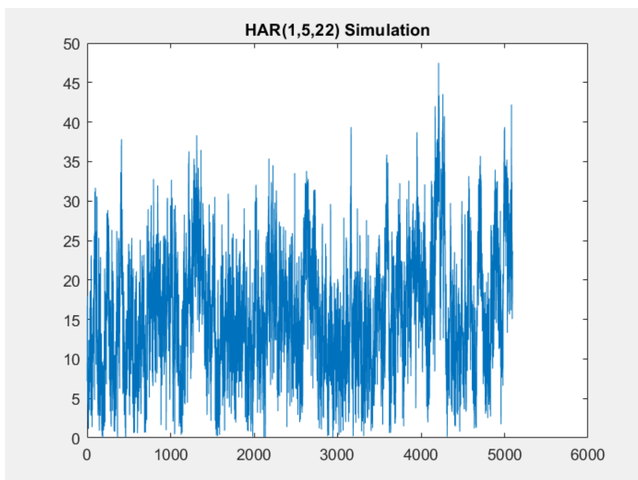

Figure 5: Simulated daily volatility based on an HAR(1,5,22) data generating process

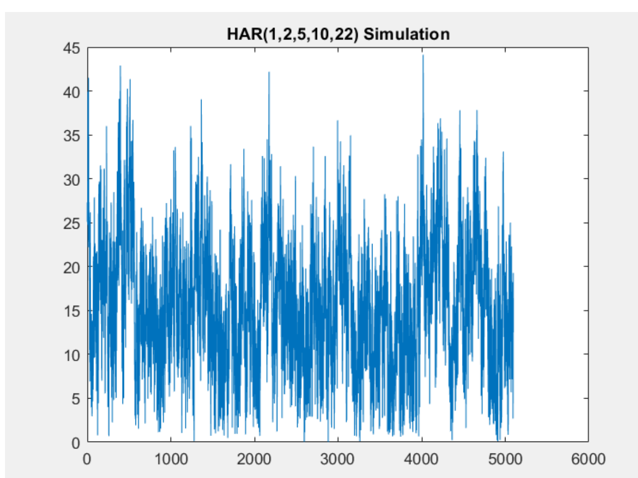

Figure 6: Simulated daily volatility based on an HAR(1,2,5,10,22) data generating process

Table 5: Realized volatility autocorrelation coefficients of SPX, HAR(1,5,22) and HAR(1,2,5,10,22)

Lag	SPX	HAR(1,5,22)	HAR(1,2,5,10,22)
1	0.8182	0.7327	0.6927
2	0.3559	0.2289	0.3092
3	0.1298	0.1702	0.1275
4	0.1234	0.1163	0.0891
5	0.0874	0.1098	0.0901
6	0.0348	0.0247	0.0383
7	0.0500	0.0148	0.0233
8	0.0570	0.0257	0.0676
9	0.0841	0.0061	0.0493
10	-0.0238	0.0313	0.0149
11	0.0185	0.0162	0.0227
12	0.0062	0.0259	-0.0272
13	-0.0016	0.0205	-0.0059
14	0.0117	0.0247	0.0223
15	0.0100	0.0225	0.0044
16	-0.0004	0.0109	0.0077
17	0.0045	0.0278	0.0031
18	0.0101	0.0058	-0.0033
19	0.0158	-0.0107	0.0191
20	0.0069	0.0088	0.0165

of the daily SPX returns.

The realized volatility of the SPX index can be found in Figure 4, the simulated volatility of the HAR(1,5,22) model in Figure 5 and that of the simulated HAR(1,2,5,10,22) data generating process in Figure 6. We observe no major visual differences between the simulated volatility of either of the HAR models. However, we note that the SPX index shows substantially more positive outliers than the simulated HAR models. The error distribution used in the simulation is not close to the distribution of the error in the data generating process of the volatility of the SPX index. Probably a distribution with fatter tails than a normal distribution ought to be used in order to obtain an improved simulation.

Lastly, we present the autocorrelation coefficients until lag 20 of the daily returns of the SPX index, the simulated HAR(1,5,22) & HAR(1,2,5,10,22) models in Table 5. These results appear to be broadly in line with findings in the simulation by Corsi (2009). We appear to have reproduced some long memory of realized volatility, although we formally use a short-memory model. For both the realized SPX values and the simulated volatility we obtain comparable autocorrelation coefficients values for lags 2 to 20. For lag 1, the autocorrelation coefficient of the SPX realized volatility appears somewhat larger than the autocorrelation coefficient of the first lag in both simulated series.

5.3 Goodness-of-fit tests

In Table 6 we present results of the goodness-of-fit tests. When we consider the results of the F-tests, we note that for both the HAR(1,5,22) and HAR(1,2,5,10,22) model we can reject the null hypothesis of equal goodness-of-fit in comparison with the AR(22) model for each of the indices under consideration. We obtain conflicting results among the indices when we consider the AIC. With respect to the SPX index, the HAR(1,2,5,10,22) model minimizes the AIC, whereas AR(22) yields the highest AIC. This is an indication of HAR(1,2,5,10,22) being the most appropriate specification for the SPX index. However, for the FTSE and AEX index this is the other way around. Nevertheless, when we take into consideration the SIC we obtain less conflicting results. For this information statistic it holds for each index that $SIC_{HAR(1,2,5,10,22)} > SIC_{HAR(1,5,22)} > SIC_{AR(22)}$ which could be an indication that indeed the index-specific HAR(1,2,5,10,22) model is the most appropriate specification overall.

Table 6: Goodness-of-fit tests

		SPX	FTSE	AEX
F-Test	HAR(1,5,22)	7.6671	5.6480	6.2486
		(1.9084)	(1.9084)	(1.9083)
	HAR(1,2,5,10,22)	2.2356	3.4244	3.7256
		(1.9688)	(1.9688)	(1.9687)
AIC	AR(22)	32802	31308	30602
	HAR(1,5,22)	31414	32871	30682
	HAR(1,2,5,10,22)	31312	32826	30632
SIC	AR(22)	31458	32953	30753
	HAR(1,5,22)	31440	32897	30709
	HAR(1,2,5,10,22)	31351	32865	30671

Notes. This table shows various goodness of fit tests. First of all it depicts the reuslts of the F-test for multiple hypothesis testing between the HAR(1,5,22) model from Corsi (2009) and the AR(22) model. Besides, the F-test for multiple hypothesis testing between the HAR(1,2,5,10,22) and the AR(22) model is performed. Between parentheses the 1% critical values are established. Lastly, we show the Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC) for the AR(22), HAR(1,5,22) and HAR(1,2,5,10,22) model.

5.4 In-sample forecast evaluation

In Table 7 we present the results of our one-day-ahead in-sample forecast evaluation. In line with Corsi (2009) we find that the AR(1) and AR(3) model appear to have less predictive in-sample quality than the HAR(1,5,22) model. For each of the indices under consideration, it holds that the

RSME and MAE of the AR(1) model is higher than the AR(3) model and the RSME and MAE of the AR(3) model is higher than the RSME and MAE of the the HAR(1,5,22) model. Interestingly, the HAR(1,2,5,10,22) model has a lower RMSE and MAE than the HAR(1,5,22) model for each index and a higher Mincer and Zarnowitz (1969) R^2 value as well. This shows that with respect to in-sample forecasts, the HAR(1,2,5,10,22) model appears to superior to the HAR(1,5,22) model.

Table 7: In-sample Forecast Evaluation

	SPX			FTSE			AEX		
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
AR(1)	5.7172	3.5221	0.6688	6.6186	3.8511	0.5412	5.0525	3.1678	0.7073
AR(3)	5.2975	3.2514	0.7156	6.1112	3.5270	0.6090	4.7508	2.9339	0.7410
HAR(1,5,22)	5.2694	3.2077	0.7193	5.9800	3.4017	0.6267	4.6789	2.8643	0.7494
HAR(1,2,5,10,22)	5.2148	3.1883	0.7251	5.9518	3.3868	0.6302	4.6542	2.8561	0.7520

Notes. This table shows a comparison of the in-sample performance of one-day ahead forecasts of the autoregressive models of lag order 1 and 3, as well as the HAR(1,5,22) model from Corsi (2009) and the HAR(1,2,5,10,22) model. These forecasts are made for the SPX, FTSE and AEX index during the time period January 2000 - May 2020. The three performance measures are the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE) and the R^2 of the Mincer-Zarnowitz regression.

5.5 Out-of-sample forecast evaluation

In Table 8 we present the results of our one-day-ahead, one-week-ahead and two-week ahead out-of-sample forecast evaluation. When we consider Table 8 we note that overall the WLS Fitted and WLS Parameters methods appear to deliver the strongest improvements in the forecast performance measures. Apart from the RMSE of the one-day-ahead FTSE index forecasts, at least one of the WLS schemes is included in the two most beneficial values of the RMSE, MAE and R^2 for each of the indices and forecast horizons. Furthermore, it is interesting to note that the longer the time horizon the higher the relative quality of the WLS approaches. Apart from that, we observe that WLS Fitted appears to outperform WLS Parameter for one-day-ahead forecasts while this advantage disappears when we consider the two-week-ahead forecasts. All in all, the WLS schemes appear to provide the largest improvement of forecast accuracy despite them being one of the less complex enhancements of the HAR model which we consider in this research.

It ought to be noted that there are also a few instances in which adaptive elastic net appears to be the best performing method for the SPX index. However, the performance of adaptive elastic net is not consistent among all indices under consideration since adaptive elastic net appears to deliver the worst performance among all applied models and methods for the FTSE index. Nevertheless,

Table 8: Out-of-sample forecast evaluation

One-day-ahead		SPX			FTSE			AEX	
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
AR(1)	15.228	3.4990	0.6804	9.1608	3.8604	0.5415	10.685	2.9948	0.6853
AR(3)	8.0155	3.2749	0.7079	**4.2910**	3.5186	0.5978	5.6203	2.7691	0.7147
HAR(1,5,22)	5.3746	3.1616	0.7274	**5.9914**	3.3465	0.6313	**4.4596**	2.6885	0.7294
HAR(1,2,5,10,22)	5.3443	3.1465	0.7303	5.9966	3.3373	0.6307	4.4676	2.6894	0.7285
WLS Fitted	**5.2954**	**3.1305**	**0.7349**	5.9447	3.3196	**0.6369**	**4.4250**	**2.6800**	**0.7331**
WLS Parameter	5.4659	**3.0674**	0.7331	6.1427	**3.2589**	**0.6331**	4.5670	**2.6591**	**0.7302**
Adaptive LASSO	5.3298	3.1385	0.7319	6.0350	**3.2968**	0.6285	4.4894	2.6974	0.7258
Adaptive ENET	**5.3069**	3.1434	**0.7336**	7.9166	3.3666	0.4340	4.5074	2.7109	0.7229
DMSFE FC $\delta = 1$	5.3934	3.2154	0.7254	6.0449	3.3952	0.6262	4.5385	2.7589	0.7201
DMSFE FC $\delta = 0.9$	5.4146	3.2304	0.7233	6.0406	3.3948	0.6266	4.5482	2.7623	0.7191
One-week-ahead		SPX			FTSE			AEX	
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
AR(1)	8.2743	5.2861	0.4702	9.0816	5.6914	0.3220	7.1183	4.6315	0.4487
AR(3)	7.2328	4.3238	0.5287	7.6059	4.4438	0.4441	6.1770	3.7603	0.5182
HAR(1,5,22)	6.9774	3.9948	0.5583	**7.1265**	3.9444	0.4906	5.9198	3.4025	0.5483
HAR(1,2,5,10,22)	6.9777	3.9658	0.5549	7.1423	3.9514	0.4871	5.8773	3.3988	0.5463
WLS Fitted	**6.8848**	3.9516	**0.5625**	**7.0476**	3.9179	**0.4974**	**5.7604**	**3.3644**	**0.5572**
WLS Parameter	7.0646	**3.8377**	0.5601	7.2021	**3.8025**	**0.4968**	**5.8584**	**3.2877**	**0.5605**
Adaptive LASSO	6.9989	3.9547	0.5542	7.2353	**3.9146**	0.4796	5.9253	3.4095	0.5404
Adaptive ENET	**6.9086**	**3.9421**	**0.5614**	10.420	3.9907	0.2507	5.9129	3.4155	0.5384
DMSFE FC $\delta = 1$	7.7736	4.6313	0.5139	7.9996	4.7107	0.4343	7.3258	4.1051	0.4690
DMSFE FC $\delta = 0.9$	7.8285	4.6371	0.5107	8.0054	4.7064	0.4334	7.3832	4.1074	0.4665
Two-week-ahead		SPX			FTSE			AEX	
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
AR(1)	8.7091	5.4567	0.3683	9.2088	5.7432	0.2690	7.5394	4.8120	0.3271
AR(3)	8.0413	4.7150	0.4173	8.0375	4.6212	0.3649	7.0140	4.1458	0.3748
HAR(1,5,22)	**7.9084**	4.4737	0.4468	**7.8099**	4.2969	0.3967	6.9251	3.8834	0.4035
HAR(1,2,5,10,22)	7.9428	4.4843	0.4437	7.8513	4.3139	0.3937	7.0527	3.9379	0.3931
WLS Fitted	**7.8122**	**4.4647**	**0.4488**	**7.7486**	**4.2749**	**0.4020**	**6.7542**	**3.8791**	**0.4112**
WLS Parameter	8.0397	**4.2638**	**0.4511**	8.0020	**4.1648**	**0.4022**	**6.8151**	**3.7319**	**0.4192**
Adaptive LASSO	7.9544	4.4701	0.4457	8.1265	4.3060	0.3761	7.1418	3.9670	0.3847
Adaptive ENET	7.9339	4.4693	0.4464	13.995	4.4592	0.1446	7.1320	3.9784	0.3825
DMSFE FC $\delta = 1$	8.5588	4.9436	0.4058	8.5751	4.9350	0.3409	8.2619	4.4340	0.3253
DMSFE FC $\delta = 1$	8.6176	4.9572	0.4031	8.5801	4.9308	0.3402	8.3268	4.4398	0.3230

Notes. This displays the one-day-ahead, one-week-ahead and two-week-ahead out-of-sample performance of the AR(1), AR(3), HAR(1,5,22) & HAR(1,2,5,10,22) models for the indices SPX, FTSE and AEX. Furthermore, we apply two weighted least squares schemes on a model with a HAR(1,2,5,22) lag structure. WLS Fitted multiplies each observation with the inverse of the fitted daily volatility. WLS Parameter multiplies each observation with the inverse of the realized daily volatility of the forecasted observation. We apply two penalized linear methods. Adaptive LASSO applies a LASSO penalty function with the algorithm from Zou (2006). Adaptive elastic net applies the same procedure in which we take an average of the LASSO and ridge regression. Lastly, we perform a forecast combination among univariate forcasts of the five time horizons included in the HAR(1,2,5,10,22) model. We calculate the Discounted Mean Forecasts Error forecast combination from Rapach et al. (2010) with discount factors $\delta = 1$ & $\delta = 0.9$. The three performance measures are the Root Mean Squared Error, the Mean Absolute Error (MAE) and the R^2 of the Mincer-Zarnowitz regression. The two lowest RMSE and MAE values and the two highest R^2 values among the models for each index and forecast horizon are indicated in bold.

we must note that the penalized linear methods require substantial computational time due to the necessary cross-validations and computationally demanding ADMM algorithm which we use to minimize the loss function and penalty function. Lastly, it is noteworthy that the quality of the DMSFE forecast combination forecasts is relatively low, indicating that this methodology is not as beneficial in the context of HAR models as in the context of forecasting returns with independent economic variables as regressors as performed by Rapach et al. (2010).

6 Conclusion

In this paper we answer the following research question: "Can we enhance the HAR(1,5,22) model from Corsi (2009) given that it improves forecasts in an index-specific context?" In order to answer this question we have considered the HAR(1,2,5,10,22) model, since Corsi (2009) suggests that in an index-specific context this might be a more appropriate specification than the HAR(1,5,22) model. The HAR(1,2,5,10,22) model appears to be superior to the HAR(1,5,22) model when forecasting the Standard & Poor's 500 index, but inferior to the HAR(1,5,22) model when forecasting the Financial Times Stock Exchange and Amsterdam Stock Exchange index during January 2000 until May 2020.

Furthermore we have studied three types of forecasting and estimation procedures among the components of the HAR(1,2,5,10,22) model as an alternative to the OLS performed by Corsi (2009). These are WLS, penalized linear models and forecast combinations. While at different time horizons, different methodologies appear to improve both the HAR(1,5,22) and HAR(1,2,5,10,22) model which are estimated via OLS, we can conclude that generally speaking the non-parametric and parametric WLS schemes deliver the largest improvements of the original HAR(1,5,22) model from Corsi (2009) among all indices and forecast time horizons under consideration. Penalized linear methods applied on the HAR(1,2,5,10,22) model generally yield forecasts of a quality which is comparable to the OLS estimated HAR models, but require substantially more computational power which has to be traded of against its the minor forecast improvements. Lastly, the DMSFE forecast combination technique applied in this research does not provide a beneficial increase in forecast accuracy.

There are several limitations to this research which we would like to outline before concluding this paper. Unlike Corsi (2009), we could not derive the realized volatility from high-frequency data due to a lack of availability. Instead we retrieved realized variances from the Oxford Realized Library which are based on 5-minute returns rather than high-frequency data. Furthermore, we cannot retrieve the raw data from which these realized variances are calculated, such that we cannot verify the correctness of the applied transformations of the raw data by the Oxford Realized Library.

Additionally, we sum up some potential improvements of this research. To start with, we have only considered recent data and only applied our methodologies in one particular sample within the time span under consideration of January 2000 until May 2020. Subdividing our dataset into multiple sample periods might enable us to evaluate whether we can draw the same conclusions for each of the individual subsamples. Furthermore, it might be interesting to consider older data in order to evaluate whether volatility dynamics have changed over a longer period of time. The latter might however be slightly complicated due to the lack of available high-frequency data which are necessary in order to calculate the volatility which we apply in this research.

Another improvement of this paper could be evaluating more specifications of penalized linear models such as a ridge regression or different specifications of the penalty function. In addition to this, it might be appropriate to perform a larger amount of cross-validations. We were not able to make a more elaborate study with respect to the penalized linear models since our programming software MATLAB required substantial time when estimating the coefficients of penalized linear methods. This limited our ability to assess multiple penalty specifications and perform substantially more cross-validations. Perhaps when we use different statistical software such as R or another programming language such as Python we might be able to expand the scope of the penalized linear model study in a convenient manner.

Lastly, we mention several suggestions for future research. This paper tries to enhance the HAR model in an index-specific context by using a HAR(1,2,5,10,22) specification which is suggested by Corsi (2009) for the SPX index. The other financial derivatives studied by Corsi (2009), which are exchange rates and treasury bill rates, might as well have their own dynamics and require different HAR specifications in order to make more accurate predictions. Hence it would be interesting to apply the WLS schemes, penalized linear models and the forecast combination technique applied in this research to other types of financial derivatives with an appropriate derivative-specific lag structure.

Another suggestion for future research would be combining the most successful approaches in this thesis with the extension that is performed by Corsi and Reno (2009). In this paper, substantial gains with respect to predictive accuracy are gained by including negative returns, heterogeneous leverage and jumps to the HAR model. When the successful methodologies in this paper are combined with the approaches in the paper Corsi and Reno (2009) we might be able to make even more precise predictions.

References

Aït-Sahalia, Y., & Mancini, L. (2008). Out of sample forecasts of quadratic variation. *Journal of Econometrics*, *147*(1), 17–33.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 885–905.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics*, *89*(4), 701–720.

Audrino, F., Huang, C., & Okhrin, O. (2018). Flexible har model for realized volatility. *Studies in Nonlinear Dynamics & Econometrics*, *23*(3).

Audrino, F., & Knaus, S. D. (2016). Lassoing the har model: A model selection perspective on realized volatility dynamics. *Econometric Reviews*, *35*(8-10), 1485–1521.

Bollerslev, T., Litvinova, J., & Tauchen, G. (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, *4*(3), 353–384.

Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, *192*(1), 1–18.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, *3*(1), 1–122.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics*, *10*(4), 407–432.

Clements, A., & Preve, D. (2019). A practical guide to harnessing the har volatility model. *Available at SSRN 3369484*.

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

Corsi, F., & Reno, R. (2009). Har volatility modelling with heterogeneous leverage and jumps. *Available at SSRN 1316953*.

Diebold, F. X., & Pauly, P. (1990). The use of prior information in forecast combination. *International Journal of Forecasting*, *6*(4), 503–508.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al. (2004). Least angle regression. *The Annals of statistics*, *32*(2), 407–499.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Huang, H., & Lee, T.-H. (2010). To combine forecasts or to combine information? *Econometric Reviews*, *29*(5-6), 534–570.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). NBER.

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—analyzing the dynamics of market components. *Journal of Empirical Finance*, *4*(2-3), 213–239.

Parikh, N., & Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in optimization*, *1*(3), 127–239.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, *23*(2), 821–862.

Schwarz, G., et al. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461–464.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, *23*(6), 405–430.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.