



Department of Economics and Finance
Major in Finance

**A journey in Realized Variance:
modeling, forecasting and Variance Risk Premium**

Candidate

Fabio Nopor

Student ID 712401

Supervisor

Prof. Paolo Santucci De Magistris

Co-Supervisor

Prof. Tommaso Proietti

Academic Year 2019/2020

Contents

1	Introduction	1
2	Measuring Volatility	4
2.1	Logarithmic Price Process and Quadratic Variation	5
2.2	Realized Variance	9
2.2.1	Microstructure noise	11
2.3	Jumps	12
3	Models and forecast evaluation	18
3.1	Forecasting models	18
3.1.1	HAR-RV	18
3.1.2	HAR-RV-J-D and HAR-RV-J-F	20
3.1.3	HAR-RV-CJ and HAR-RV-LCJ	22
3.1.4	C-HAR	23
3.1.5	Q-Family	23
3.1.6	S-HAR	25
3.1.7	HAR-X	25
3.1.8	Forecast combination	26
3.1.9	Benchmark	27
3.2	Forecast evaluation	27
3.2.1	Point Forecast	27
3.2.1.1	Diebold-Mariano Test and Model Confidence Set	28
3.2.2	Density Forecast	31
3.2.2.1	Berkowitz Test	31

4	In-sample and Out-of-sample analysis	33
4.1	S&P 500 Index	33
4.1.1	In-sample	36
4.1.2	Out-of-sample	40
4.2	iShares China Large-Cap ETF	44
4.2.1	In-sample	46
4.2.2	Out-of-sample	50
4.3	Apple Inc	53
4.3.1	In-sample	55
4.3.2	Out-of-sample	59
5	Variance Risk Premium	63
5.1	Return predictability – full period	66
5.2	Return predictability – excluding Covid-19 pandemic crisis	68
6	Conclusion	71
	Bibliography	74

1 Introduction

The uncertainty of asset returns has for a long time captured the attention of investors and academic researchers. Estimation, modeling and forecasting of financial volatility are of crucial importance in many fields of finance; including asset pricing, risk management, hedging, portfolio management and asset allocation. However, the main issue is that volatility is not directly observable. Early research extracted volatility estimates from asset return data before specifying a parametric time-series model for volatility. For example, Officer (1973) used a rolling standard deviation to estimate volatility at each point in time while German and Klass (1980) and Parkinson (1980) made use of the difference between the high and low prices on a given day, implicitly assuming that volatility is constant over some interval of time. Nevertheless, it is inconsistent and statistically inefficient to use volatility measures based on the assumption of constant volatility. The solution to this problem is to use parametric models to estimate volatility from returns, such as Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) proposed respectively by Engle (1982) and Bollerslev (1986). However, returns distribution shows fat tails, tail crossover, scaling and multiscaling¹ all features that the standard GARCH models and, more in general, short-memory processes are not able to reproduce. Motivated by this, a growing interest in long-memory processes has emerged. Long-memory volatility is usually obtained by employing fractional difference operators as in FIGARCH models of returns or ARFIMA models of realized volatility, nonetheless the fractional integration is a convenient mathematical trick without a clear economic interpretation. To overcome this, Corsi (2009) proposed the Heterogenous Autoregressive model of Realized Volatility (HAR-RV), a simple model able to reproduce the main empirical features observed in the data while remaining parsimonious and easy to estimate. The model is inspired by the Heterogenous Market Hypothesis presented by Muller *et al.* (1993) which recognizes the presence of different heterogeneity in behavior across traders. Each agent is responsible for a different type of volatility in the market, and that leads to the so-called *volatility cascade*.

The main aim of the thesis is to model and forecast the realized variance of the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc on four different forecast horizons: one day, one week, two weeks and one month. Furthermore, to assess whether one (or more) of the models considered in the thesis manages to outperform the other, or if it is possible to

¹ Different scaling exponents for different powers

find a *file rouge* that connects the forecasting performance of the models for the three times series. The starting model which is none other than the basis of all the other models included in the thesis, is the HAR-RV proposed by Corsi (2009). The HAR-RV is estimated with the use of a linear regression, this allows a simple integration with new exogenous variables. The total number of models used is 13. The evaluation of the prediction is carried out by several comparison methodologies starting from the *point forecast* comparison, both from an individual and simultaneous point of view, and *density forecast* comparison represented by the Berkowitz (2001) test. Once we find the best forecasts, on the different time horizons, we use the one month-ahead prediction in order to compute the Variance Risk Premium which is the difference between option implied volatility and realized variance. Therefore, motivated by the work of Bollerslev, Tauchen and Zhou (2009), we use the VRP to predict the returns of the three financial instruments considered.

From the results of the empirical Chapters, it is evident that utilizing more complex models, with respect to the parsimonious HAR-RV, leads to an increase in forecasting performance especially on longer time horizons. In particular, we will see that the decomposition of the realized variance into continuous and jumps components will be of fundamental importance, especially in one month-ahead forecast. For comparison purpose, all the results, both in-sample and out-of-sample, are compared to the commonly used random walk model. Here, we will see that the more complex model always outperforms the benchmark. The second finding in this thesis is relative to the returns prediction, deriving from the use of the Variance Risk Premium. We will see that the VRP which originated from the finest one month-ahead forecasting models will get a superior returns prediction respect to the one deriving from the random walk model.

In particular, the second Chapter is dedicated to the theoretical background of realized variance with significant focus on its decomposition in continuous and jumps component using the threshold bipower variation proposed by Corsi, Pirino and Renò (2010). The third Chapter, is focused on the theoretical analysis of the 13 forecasting models, starting from the basic HAR-RV to more complex models such as HAR-RV-LCJ, HAR-CJ and HAR-Q-F, while the last part of the Chapter is dedicated to the forecasting comparison. However, the most significant Chapters of the thesis are the fourth and fifth. In Chapter 4 we will see the empirical forecast of realized variance for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc stock. All data is downloaded from barchart.com and includes more than half a million observations. The fifth Chapter, is dedicated to the Variance Risk Premium analysis and to returns prediction. This last analysis is conducted both including and

excluding the recent economic and health pandemic crisis related to the spread of Covid-19 at the end of February 2020. The last Chapter deals with the conclusions drawn from the empirical analysis.

2 Measuring Volatility

Since the beginning of the 21st century, with the increasing availability of data and the recognition that higher frequency intraday data is much more informative compared to longer horizon, many authors contributed to the study of the so called quadratic variation and its estimator, the realized variance; notably Andersen, Bollerslev, Diebold and Labys (2001, 2003), Barndorff-Nielsen and Shephard (2001, 2002a, 2004) and Andersen, Bollerslev and X-Huang (2010). Using high frequency data is possible to embed more information than using the canonical open-to-close returns. Figure 1 depicted the intraday price movements of the S&P 500 Index² on 08/04/2015 and 14/02/2020. For both days the open-to-close return was zero so that simply squaring the daily return cannot measure the volatility that occurred intraday.

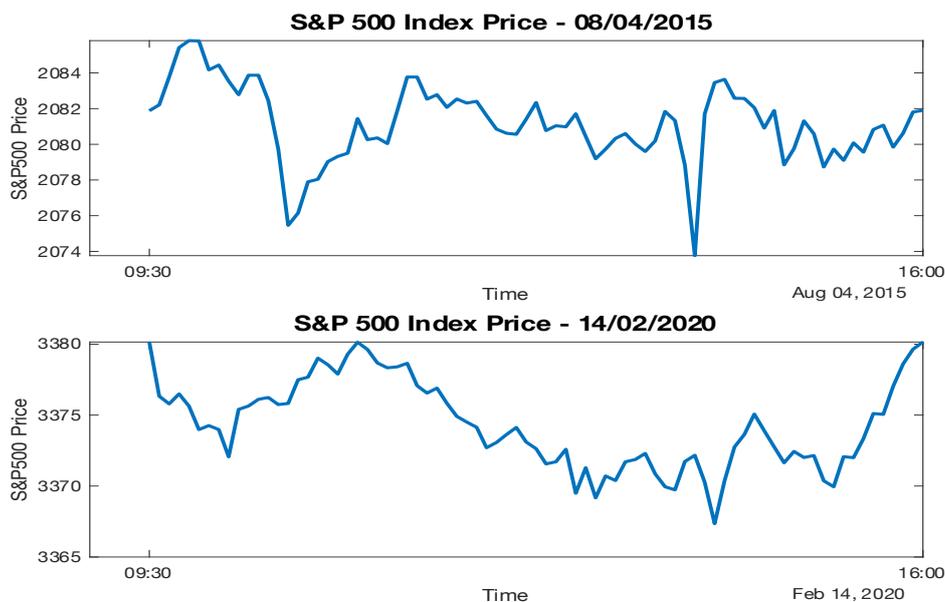


Figure 1: Intraday prices for S&P 500 Index on 08/04/2015 and 14/02/2020 for which open-to-close daily return were both zero.

The Chapter is therefore dedicated to the theoretical and essential background of the non-parametric measure of daily return variability called realized variance which is computed using high-frequency intraday data. After that we are going to break down the realized variance into continuous and jumps component through the method proposed by Corsi, Pirino and Renò (2010) This will be of fundamental importance in the empirical applications

² All data are downloaded from Barchart.com. For more information about S&P 500 Index see paragraph 4.1

in Chapter 4 and 5 where we will see that the models that use it will obtain significant forecast improvement.

2.1 Logarithmic price process and quadratic variation

Let us begin with an univariate logarithmic price process of an asset as p_t ($p_t = \ln P_t$, where P_t denotes the price of process of an asset) on a probability space (Ω, \mathcal{F}, P) evolving over a continuous time interval $[0, T]$ where T denotes a positive integer. We further consider an information filtration denoted as σ -field $(\mathcal{F}_t)_{t \in [0, T]} \subseteq \mathcal{F}$ which satisfies the usual condition of P -completeness and right continuity. We also assume that the information set \mathcal{F}_t contains the asset prices and relevant state variables through time t .

Definition 2.1. The continuous compounded return over the time interval $[t - h, t]$ where $0 \leq h \leq t \leq T$ is the difference between the logarithmic price a time t and the logarithmic price at time $t - h$:

$$r_{t,h} = p_t - p_{t-h} \quad (2.1)$$

For the case of cumulative return up to time t , the return over the $[0, t]$ time interval is:

$$r_t = r_{t,t} = p_t - p_0 \quad (2.2)$$

Using equation (1.1) and equation (1.2) it is possible to define the relationship between the period-by-period return and the cumulative returns:

$$r_{t,h} = r_t - r_{t-h} \quad (2.3)$$

A maintained assumption is that the price process is strictly positive and finite so that p_t and r_t are well defined over $[0, T]$. It follows that the r_t only has countable jump points over $[0, T]$ and the price and return processes are squared integrable. Let us define $r_{t-} = \log_{\tau \rightarrow t, \tau < t} r_t$ and $r_{t+} = \log_{\tau \rightarrow t, \tau > t} r_t$. In particular, $r_t = r_{t+}$ determines the right-continuous left-limit (càdlàg) of the process, while $r_t = r_{t-}$ determines the left-continuous right-limit (càglàd) for all t in $[0, T]$. In the following, we assume to work with a càdlàg version of the return process. We set the jumps in the return process as: $\Delta_{r_t} = r_t - r_{t-}$, $0 \leq t \leq T$. We have at continuity points $\Delta_{r_t} = 0$ and $P[\Delta_{r_t} \neq 0] = 0$ with $t \in [0, T]$. This means that there is a

countable number of jumps in the price process without a set frequency. Jump processes that do not explode are called *regular*. Following Back (1991) and applying the standard assumption of no arbitrage and finite-expected return we have that the log-price process must constitute a semimartingale, a process that is composed by a local martingale and a càdlàg adapted locally bounded process. This affords the following unique canonical return decomposition (Protter, 1992).

Proposition 2.1. *Any arbitrage-free logarithmic price process subject to the regularity conditions outlined above may be uniquely represented as:*

$$r_t \equiv p_t - p_0 = \mu_t + M_t = \mu_t + M_t^c + M_t^J \quad (2.4)$$

Where μ_t is a predictable and finite-variation process, M_t is a local martingale that is further decomposed into a continuous sample path M_t^c and a compensated jump martingale M_t^J . We may normalized the initial conditions such that $\mu_0 \equiv M_0 \equiv M_0^c \equiv M_0^J \equiv 0$

It is convenient to decompose the expected return process, μ_t , into a purely continuous predictable finite-variation part, μ_t^c , and a purely predictable jump part, μ_t^J . There can not be a perfectly anticipated jump in terms of time and size in the mean unless it is accompanied by large jump innovation risk as well, so that $P[\Delta M_t \neq 0] > 0$ in order to overturn the possible gain from the predictable jump. According to Andersen *et al.* (2010) most of the continuous-time asset pricing literature ignores predictable jumps, rather than modifying the standard setup to allow for the presence of predictable (but empirically negligible) jumps, we assume away such jumps.

We also express the price changes and associated return in proposition 2.1 over a discrete time interval and not in continuous-time integral representation. This is due to the fact that the real-time price data are not available at every instant and to the presence of market microstructures noise³. Consequently we focus on measures that represent the (average) volatility over a discrete time interval. This suggests the natural and general notion of volatility based on the quadratic variation process for the local martingale component in the unique semi-martingale return decomposition.

³ For more details on the market microstructure noise and the approach used in the thesis see paragraph 2.2.1

Definition 2.2. Let r_t be a semi-martingale process. The unique *quadratic variation process*, $[r, r]_t$, $t \in [0, T]$, associated with r_t is formally defined as:

$$[r, r]_t = r_t^2 - 2 \int_0^t r_{s-} ds \quad (2.5)$$

where the stochastic integral of the adapted càglàd process, r_{s-} , with respect to the càdlàg semi-martingale, r_s , is well defined.

Following Andersen *et al.* (2003), in the theory of quadratic variation we can identify two key points:

- i. **Proposition 2.2.** For an increasing sequence of random partitions of $[0, T]$, $0 = \tau_{m,0} \leq \tau_{m,1} \leq \tau_{m,2} \leq \dots \leq \tau_{m,m} = T$ such that $\sup_{j \geq 1} (\tau_{m,j+1} - \tau_{m,j}) \rightarrow 0$ and $\sup_{j \geq 1} \tau_{m,j} \rightarrow T$ for $m \rightarrow \infty$ with probability one, we have that

$$\lim_{m \rightarrow \infty} \left\{ \sum_{j \geq 1} \left(r(t \wedge \tau_{m,j}) - r(t \wedge \tau_{m,j-1}) \right)^2 \right\} \rightarrow [r, r]_t \quad (2.6)$$

where $t \wedge \tau \equiv \min(t, \tau)$, $t \in [0, T]$, and the convergence is uniform on $[0, T]$ in probability.

Intuitively, the proposition 2.2 state that the quadratic variation process represents the (cumulative) realized sample-path variability of r_t over the time interval $[0, T]$.

- ii. Under the maintained assumption of no predictable jumps in the returns process we also have

$$[r, r]_t - [r, r]_{t-h} = [M^c, M^c]_t - [M^c, M^c]_{t-h} + \sum_{t-h < s \leq t} \Delta r^2(s) \quad (2.7)$$

The first property suggests that we may approximate the quadratic variation by cumulating high frequency return. The measures computed in this way are generally known as realized variance. While the second property reflects the fact that the quadratic variation of continuous finite-variation process, μ_t^c , is zero, the mean component becomes irrelevant.

In theoretical asset and derivatives pricing literature it is frequently assumed that the sample paths are continuous with the corresponding diffusion process given in the form of stochastic differential equations rather than through (abstract) integral representations for continuous sample path semi-martingale. However, the previous assumption can be made using the martingale Representation Theorem (Protter (1992)) without loss of generality.

Proposition 2.3. *For any univariate, square-integrable, continuous sample path, logarithmic price process, which is not locally riskless, there exists a representation such that for all $0 \leq t \leq T$:*

$$r_{t,h} = \mu_{t,h} + M_{t,h} = \int_{t-h}^t \mu_s ds + \int_{t-h}^t \sigma_s dW_s \quad (2.8)$$

where μ_s is an integrable, predictable, and finite-variation stochastic process, σ_s is a strictly positive càdlàg stochastic process satisfying

$$P \left[\int_{t-h}^t \sigma_s^2 ds < \infty \right] = 1 \quad (2.9)$$

and W_s is a standard Brownian motion.

Let's now introduce a standard continuous time process which defines a very general class of stochastic volatility model. The model has the following form:

$$dp_t = \mu_t dt + \sigma_t dW_t \quad (2.10)$$

where p_t is the logarithm of instantaneous price, μ_t is a càdlàg finite variation process, W_t is a standard Brownian motion and σ_t is a stochastic process independent of W_t . For this process the quadratic variation over $[t-h, t]$ is

$$QV_{t,h} = \int_{t-h}^t \sigma_s^2 ds \quad (2.11)$$

In this setting the quadratic variation and the *integrated variance* coincide. This is no longer true for more general return processes such as stochastic volatility jump-diffusion model

where the quadratic variation is the sum of the integrated variance and a jump component, we come back later to this model. For the time being we consider:

$$IV_{t,h} = QV_{t,h} = \int_{t-h}^t \sigma_s^2 ds \quad (2.12)$$

2.2 Realized Variance

With the disposal of high frequency data it is possible to estimate quadratic variation using the realized variance. The idea of using return realizations for the measurement of return variation comes from not so recent times. (Summers (1986), French et al (1987) etc.). Following the work of Andersen *et al.* (2003) is possible to define the realized volatility.

Definition 2.3. The realized volatility over $[t - h, t]$, for $0 < h \leq t \leq T$, is defined by

$$RV_{t,h} = \sum_{i=1}^n r_{t-h+(\frac{i}{n})h, \frac{h}{n}}^2 \quad (2.13)$$

The realized volatility is simply the second (uncentered) sample moment of the return process over a fixed interval of length h , scaled by the number of observations n (corresponding to the sampling frequency $1/n$) provides a volatility measure calibrated to the h -period measurement interval. If there are no jumps in prices and no market microstructure noise⁴, the realized variance provides a consistent nonparametric measure of the variance of the return process. The theoretical property of realized volatility has been discussed in a number of studies including Andersen and Bollerslev (1998a), Andersen *et al.* (2001b, 2003a), and Barndorff-Nielsen and Shephard (2001, 2002a,b).

Proposition 2.4. *If the return process is square-integrable and $\mu_t \equiv 0$, then for any value of $n \geq 1$ and $h > 0$,*

$$E[\overline{RV}_{t,h} | \mathcal{F}_{t-h}] = E[M_{t,h}^2 | \mathcal{F}_{t-h}] = E[RV_{t,h} | \mathcal{F}_{t-h}] \quad (2.14)$$

where ex-post realized volatility, $RV_{t,h}$, is an unbiased estimator for the ex-ante expected volatility, $\overline{RV}_{t,h}$.

⁴ For more details on the market microstructure noise and the approach used in the thesis see paragraph 2.2.1

Proposition 2.5. *The realized variance converge uniformly in probability to the variance of the returns process*

$$plim_{n \rightarrow \infty} RV_{t,h} = \overline{RV}_{t,h} \quad (2.15)$$

Intuitively, the last proposition implied that the realized variance is a consistent (nonparametric) estimator of variance over any time interval $[t - h, t]$, $h > 0$. Figure 2 shows the realized variance for S&P 500 Index computed using 78, 5-minutes returns interval, taking into account the trading hour from 9:00:00 to 16:30:00, from 24/04/2009 to 30/10/2020.

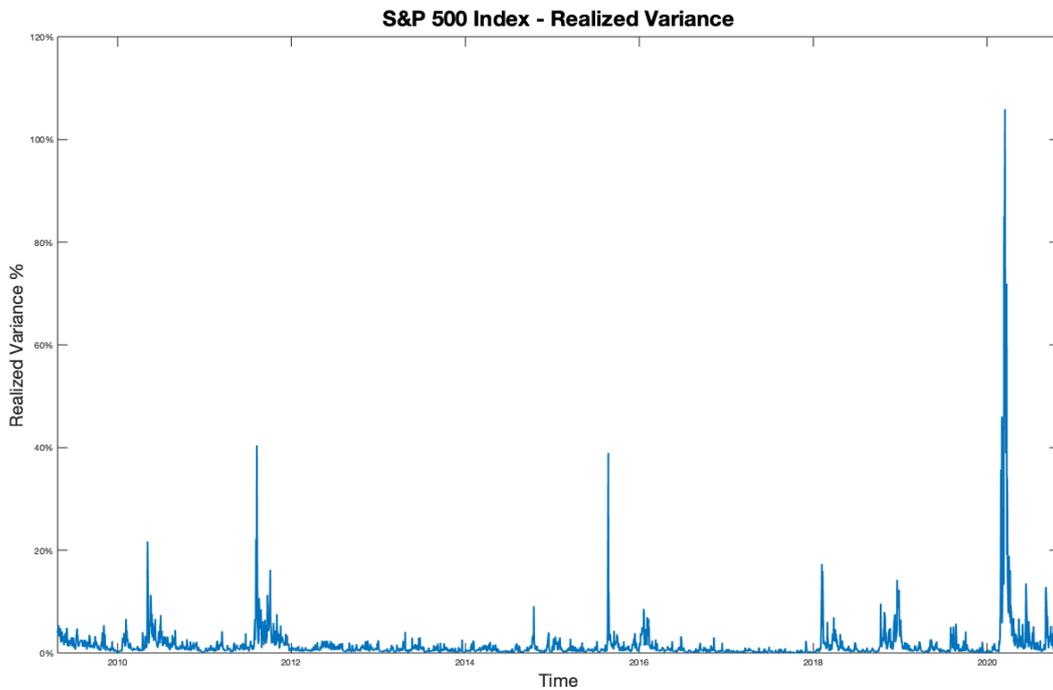


Figure 2: S&P 500 – Realized variance

With no market microstructure noise the asymptotic distribution of realized variance is derived by Barndorff-Nielsen *et. al.* (2002) as:

$$\sqrt{N} \frac{1}{\sqrt{2IQ_t}} (\overline{RV}_{t,h} - RV_{t,h}) \rightarrow N(0,1) \quad (2.16)$$

It implies that:

$$\sqrt{N} (\overline{RV}_{t,h} - RV_{t,h}) \rightarrow N(0, 2IQ_t) \quad (2.17)$$

where IQ_t is the integrated quarticity, defined as:

$$IQ_t = \int_{t-1}^t \sigma_s^4 ds \quad (2.18)$$

The authors further showed that the integrated quarticity (IQ_t) is consistently estimated by the realized quarticity (RQ_t), defined as

$$RQ_t = \frac{n}{3} \sum_{j=1}^n r_{t,j}^4 \quad (2.19)$$

2.2.1 Microstructure Noise

As we have seen in the first lines of the Chapter, the use of high frequency data allows us to embed more information in the computation of the realized variance. In addition, if the data is not affected by market microstructure noise, the average sum of squares of log returns sampled at high frequency would estimate the returns variance. In practice, however, the microstructure noise present in high frequency data leads to non-robust realized variance estimates (Andersen *et al.* 2000). As a result of the presence of noise, realized variance of log-price data has been shown to explode as the sampling interval approaches zero (Zhang *et al.* 2005). Noise is attributed to market imperfections and includes price discretization, bid ask bounce effects, as well as high frequency quoting. A common solution to deal with the bias arising from the market microstructure noise is the choice of an interval between 5 and 30 minutes (e.g., Andersen *et al.* (2001), Barndorff-Nielsen *et al.* (2002), Andersen *et al.* (2003), Bollerslev *et al.* (2015)). This thesis utilizes interval of 5 minutes.

2.3 Jumps

One interesting implementation concerns the addition of jumps component in the quadratic variation. Let's introduce the jump-diffusion model which defines a very general class of stochastic volatility model. The model has the following form:

$$dp_t = \mu_t dt + \sigma_t dW_t + \xi_t dq_t \quad (2.20)$$

where q_t denotes a Poisson process, uncorrelated with the Brownian motion W_t , governed by λ -constant jump intensity. ξ_t is responsible for the magnitude of the jumps, and is normally distributed, $\xi_t \sim N(\bar{\xi}, \sigma_\xi^2)$. We characterized (2.20) as a Brownian semimartingale process with finite jump. The quadratic variation over $[t-h, t]$ is:

$$QV_{t,h} = \int_{t-h}^t \sigma_s^2 ds + \sum_{t-h < s \leq t} J_s^2 \quad (2.21)$$

where the first component on the right is the integrated variance and the second component is the jump variations. In this section, our goal, is to separate the realized variance into continuous and jumps components (for simplicity $RV_{t,h} = RV_t$)

$$RV_t = C_t + J_t \quad (2.22)$$

A solution was proposed by Barndorff-Nielsen *et al.* (2004a; 2006). They used a special case of multipower variation, called bipower variation, constructed from the summation of appropriately scaled cross-products of adjected high-frequency absolute return in order to measure the Integrated Variation.

Definition 2.4. The multipower (MPV_t) is

$$MPV_t^{[\gamma^1, \dots, \gamma^L]} = \delta^{1-\frac{1}{2}(\gamma^1, \dots, \gamma^L)} \sum_{j=L}^n \prod_{k=1}^L |r_{t,j-k+1}| \quad (2.23)$$

Where δ is the subinterval length on which the M intraday returns $r_{t,j}$ are calculated. As $\delta \rightarrow 0$, MPV_t converges to $\int_0^t \sigma^{1-\frac{1}{2}(\gamma^1, \dots, \gamma^L)}(s) ds$. We can define a special case of multipower variation according to the choices of the vector $\gamma^1, \dots, \gamma^L$. An example is the bipower variation:

$$BPV_t = \mu_1^{-2} MPV_t^{[1,1]} = \mu_1^{-2} \sum_{j=2}^n |r_{t,j}| |r_{t,j-1}| \quad (2.24)$$

where $\mu_1 = \sqrt{2/\pi} \approx 0,7979$. The BPV_t is a consistent estimator for the continuous integrate volatility:

$$BPV_t \rightarrow \int_{t-1}^t \sigma_s^2 ds \text{ as } M \rightarrow \infty \quad (2.25)$$

Following Barndorff-Nielsen *et. al.* (2004a) and combining the equation (2.24) with the estimator of quadratic variation from the previous section we can detect the contribution of the jump variation:

$$\text{plim}_{n \rightarrow \infty} (RV_t - BPV_t) = \sum_{l=1}^{N_t} J_s^2 = J_t \quad (2.26)$$

where J_t denotes the number of non-zero jumps over $[t - h, t]$,

$$J_t = \max [RV_t - BPV_t, 0] \quad (2.27)$$

In empirical applications the result of the previous equation exhibits an unreasonably large number of non-zero small positive values. Therefore we need to distinguish between significant jumps and those that are of no importance. To detect jumps, Huang and Tauchen (2005) apply the following test statistic in which a large positive value implies that a jump occurred:

$$Z_t = \sqrt{n} \frac{(RV_t - BPV_t)RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^2 - 5)\max\{1, TQ_t/BPV_t^2\}}} \quad (2.28)$$

Z_t is standard normally distributed under the null hypothesis of no within-day jumps. The realized tripower quarticity, $TQ_{t,h}$, is defined by:

$$TQ_{t,h} = n\mu_{4/3}^{-3} \left(\frac{n}{n-4}\right) \sum_{i=5}^n |r_{t-h+\frac{(i-4)}{n}h}|^{4/3} |r_{t-h+\frac{(i-2)}{n}h}|^{4/3} |r_{t-h+\frac{i}{n}h}|^{4/3} \quad (2.29)$$

where $\mu_{4/3} = 2^{2/3} \frac{\Gamma(7/6)}{\Gamma(1/2)}$ and $\Gamma(\cdot)$ denotes the Gamma function.

Is now natural to identify the significant jumps by the realization of Z_t in excess of some critical value, Φ_α , that represents the α -quantile of the standard normal distribution function. It follows that:

$$J_t = I(Z_t > \Phi_\alpha)(RV_t - BPV_t) \quad (2.30)$$

where $I(\cdot)$ is the indicator function.

However, BPV_t , has large finite sample bias in presence of jumps. The intuition, according to Corsi *et. al.* (2010) is the following one. We assume that a generic return interval $|r_{t,h}|$ contains a jump. In regards to bipower variation, it multiplies $|r_{t,h-1}|$ and $|r_{t,h+1}|$. Asymptotically the result will vanish and bipower variation will converge into integrated volatility. However for finite samples these returns will not vanish, causing a positive bias which will be larger as $|r_{t,h}|$ increases. The consequence is that the bipower variation will be upward bias, especially in the case of consecutive jumps. This leads to an underestimation of the number of jumps detected. A solution could be to use small intraday returns intervals. Nevertheless, this leads to contamination of market microstructure noise. Following the work of Corsi *et a.l.* (2010) it is possible to introduce a measure that is nearly unbiased for small samples in the presence of jumps, unlike the previous method.

Mancini (2009) provides alternative estimators of squared and fourth power integrated volatility. Threshold realized volatility is defined as follows:

$$TRV_\delta = \sum_{j=1}^M r_{t,j}^2 I_{\{r_{t,j}^2 \leq \Theta(\delta)\}} \quad (2.30)$$

where $\Theta(\delta)$ is the threshold function. In applications, the common practice is to scale the threshold function with respect to the local spot variance:

$$\vartheta_t = c_\delta^2 \hat{V}_t \quad (2.31)$$

Using the threshold function with the multipower variation we obtain the threshold multipower variation:

$$TMPV_t^{[\gamma^1, \dots, \gamma^H]} = \delta^{1-\frac{1}{2}(\gamma^1, \dots, \gamma^H)} \sum_{j=H}^M \prod_{k=1}^H |r_{t,j-k+1}|^{\gamma^k} I_{\{|r_{t,j-k+1}|^2 \leq \vartheta_{t-k+1}\}} \quad (2.32)$$

Recall that the BPV_t is upper biased in case of large jumps in $|r_{t,h}|$. Now, if there is a jump bigger than ϑ_t , it vanishes. The problem is related to the contamination of small jumps that fall outside the exclusion range. To prevent this, Corsi *et al* (2010) proposed a corrected method, called the corrected realized threshold multipower estimator, which includes both measures, BPV_t and $TBPV_t$:

$$C-TMPV_t^{[\gamma^1, \dots, \gamma^H]} = \delta^{1-\frac{1}{2}(\gamma^1, \dots, \gamma^H)} \sum_{j=H}^M \prod_{k=1}^H Z_{\gamma^k}(r_{t,j-k+1}, \vartheta_{j-k+1}) \quad (2.33)$$

where:

$$Z_\gamma(x, y) = \begin{cases} |x|^\gamma & \text{if } x^2 \geq y \\ \frac{1}{2N(-c_\vartheta)\sqrt{\pi}} \left(\frac{2}{c_\vartheta^2} y\right)^{\frac{\gamma}{2}} \Gamma\left(\frac{\gamma+1}{2}, \frac{c_\vartheta^2}{2}\right) & \text{if } x^2 < y \end{cases} \quad (2.34)$$

The corrected threshold bipower variation of (2.24) is defined as:

$$C-TBPV_t = \mu_1^{-2} C-TMPV_t^{[1,1]} = \mu_1^{-2} Z_1(r_{t,j-k+1}, \vartheta_{j-k+1}) \quad (2.35)$$

The test statistic is defined by Corsi *et al.* (2010) as:

$$C-Tz = \delta^{-\frac{1}{2}} \frac{(RV_t - C-TMPV_t)RV_t^{-1}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) \max\left[1, \frac{C-TTriPV_t}{(C-TBPV_t)^2}\right]}} \quad (2.36)$$

where $C-Tz \rightarrow N(0,1)$ as $\delta \rightarrow 0$.

The estimator $TBPV_t$ has better finite sample properties than standard bipower variation and provides a more accurate jumps test. This allows for a corrected separation of continuous and jump components. The authors found that a test based on threshold multipower variation yields a significant advantage with respect to those based on

multipower variation. Setting a confidence level α , and using $C-Tz$ as jump detectors it is possible to estimate the jump component as:

$$J_t = I_{\{C-Tz > \Phi_\alpha\}} \max[RV_t - TBPV_t] \quad (2.37)$$

where Φ_α is the value of the standard Normal distribution corresponding to the confidence level α (we set $\alpha = 0.99$). We can now define the continuous component as:

$$C_t = RV_t - J_t \quad (2.38)$$

In the empirical applications of Chapter 4 and 5 we used the local spot variance as adopted by Corsi *et al.* (2010). \hat{V} is estimated with a non-parametric filter of length $2L + 1$ adapted for the presence of jumps by iterating in Z , while c_v^2 is a scale-free constant. More in detail:

$$\hat{V}_t^Z = \frac{\sum_{i=-L, i \neq -1, 0, 1}^L K\left(\frac{i}{L}\right) r_{t,i}^2 I_{\{(r_{t,i})^2 \leq c_v^2 \cdot \hat{V}_{t+1}^{Z-1}\}}}{\sum_{i=-L, i \neq -1, 0, 1}^L K\left(\frac{i}{L}\right) I_{\{(r_{t,i})^2 \leq c_v^2 \cdot \hat{V}_{t+1}^{Z-1}\}}}, Z = 1, 2, \dots \quad (2.39)$$

With $c_v = 3$ and starting value $\hat{V}_t^0 = +\infty$, it is implied that all observations are used in the first step. At each iteration, large returns are eliminated by the condition $(r_{t,i})^2 > c_v^2 \cdot \hat{V}_{t+1}^{Z-1}$; each estimate of the variance is multiplied by c_v^2 to get the threshold for the following step. We follow the original setting of Corsi *et al.* (2010) setting $L = 25$ and use a Gaussian kernel. Figure 3 shows respectively the continuous and the jumps component both computed using the method proposed by Corsi *et al.* (2010). It is evident that the most important driver of realized volatility is the continuous part.

For reasons of clarity and to avoid confusion in the notation: RV_t is the realized variance, $\sqrt{RV_t}$ is the realized volatility, C_t is the continuous part and J_t is jump process of the quadratic variation. The next Chapter is dedicated to the models that will be used for modeling and forecasting realized variance.

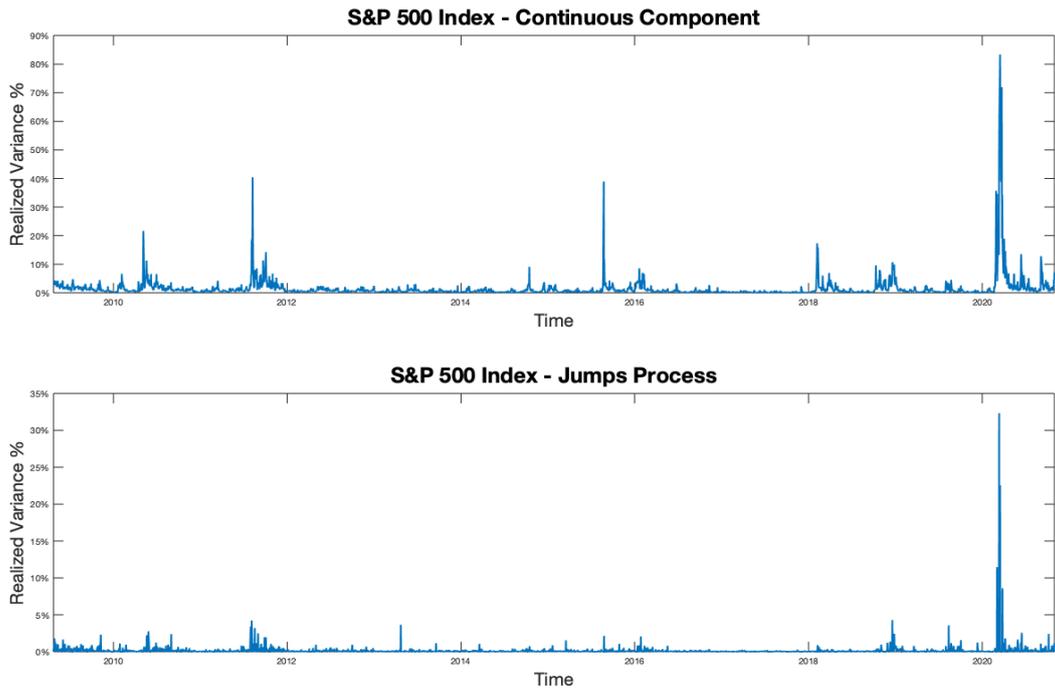


Figure 3 S&P 500 Index - Continuous and Jumps component

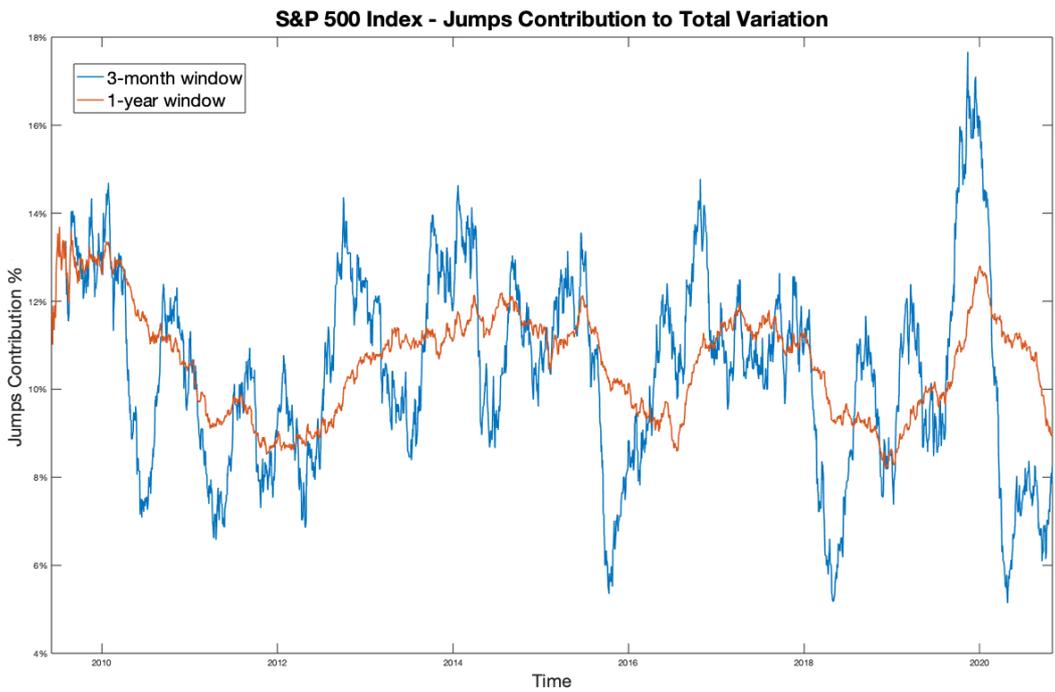


Figure 4 Percentage contribution of daily jump estimated by (2.37) to total quadratic variation measured over a moving window of 3-month and 1-year

3. Models and forecast evaluation

One of the most important reasons to study volatility is for the purpose of prediction. The first part of this Chapter is dedicated to the study of the models used to forecast the realized variance in the empirical application of the next Chapter. The model that is the basis of all other models is the Heterogenous Autoregressive model of Realized Volatility (HAR-RV) proposed by Corsi (2009). The HAR-RV model is based on the Heterogenous Market Hypothesis presented by Muller *et al.* (1993), which recognizes the presence of different heterogeneity in behavior across traders. The heterogeneity may arise from several factors such as geographical locations, risk profiles dissimilarity, institutional constraints etc. In this context we focus on the heterogeneity originates from the difference in the time horizons. Corsi (2009) distinguishes three main types of agents: (i) High intraday frequency trading-market makers, speculators, dealers; (ii) Medium-term investors who typically rebalance their positions weekly; (iii) Long term investors such as insurance companies, pension funds etc. which typically trade less frequently. Each agent is responsible for a different type of volatility in the market. Another important consideration is that volatility over longer time intervals has a stronger influence on volatility over shorter time intervals, and not vice versa. This generates a “volatility cascade” from low frequencies to high frequencies. The economic intuition behind the volatility cascade is that short term traders take into account the level of long-term volatility because it determines the expected future size of trend and risk. On the other side, the long-term traders are not influenced by short term volatility. The last part of the Chapter is dedicated to the methods used to evaluate the forecast, covering both the point and density forecast evaluation.

3.1 Forecasting models

3.1.1 HAR-RV

The parsimonious HAR-RV model proposed by Corsi (2009) assumes that the price process follows the stochastic volatility process, already mentioned above:

$$dp_t = \mu_t dt + \sigma_t dW_t$$

where p_t denotes the logarithmic price process, μ_t is a continuous finite variation process, σ_t is a stochastic process and W_t is a standard Brownian Motion. Following the approach of Corsi (2009) is possible to define the *latent partial volatility* $\tilde{\sigma}_t^{(\cdot)}$ which represents the

volatility generated by a certain market component. We consider three volatility components corresponding to time horizons of one day (1d), one week (1w) and one month (1m) respectively with the following notation: $\tilde{\sigma}_t^{(d)}$, $\tilde{\sigma}_t^{(w)}$, $\tilde{\sigma}_t^{(m)}$. It is also possible to add or consider different time horizons, as shown by Audrino, Huang and Okhrin (2016), however we are going to consider the standard daily, weekly and monthly frequencies. It is now possible to define the previous mentioned volatility cascade from a mathematical point of view:

$$\begin{aligned}\hat{\sigma}_{t+1m}^{(m)} &= c^{(m)} + \phi^{(m)}\sqrt{RV_t^{(m)}} + \tilde{\epsilon}_{t+1m}^{(m)} \\ \hat{\sigma}_{t+1w}^{(w)} &= c^{(w)} + \phi^{(w)}\sqrt{RV_t^{(w)}} + \gamma^{(w)}E_t[\hat{\sigma}_{t+1m}^{(m)}] + \tilde{\epsilon}_{t+1w}^{(w)} \\ \hat{\sigma}_{t+1d}^{(d)} &= c^{(d)} + \phi^{(d)}\sqrt{RV_t^{(d)}} + \gamma^{(d)}E_t[\hat{\sigma}_{t+1w}^{(w)}] + \tilde{\epsilon}_{t+1d}^{(d)}\end{aligned}$$

where $\sqrt{RV_t^{(d)}}$, $\sqrt{RV_t^{(w)}}$ and $\sqrt{RV_t^{(m)}}$ represent respectively the daily, weekly and monthly (ex post) observed realized volatilities. The volatility innovations $\tilde{\epsilon}_{t+1m}^{(m)}$, $\tilde{\epsilon}_{t+1w}^{(w)}$ and $\tilde{\epsilon}_{t+1d}^{(d)}$ are independent serially uncorrelated with zero mean and truncated left tail in order to ensure only positive values of partial volatility. By a recursive substitutions of the upper partial volatility, with the assumption that the daily integrated volatility is determinate by the highest frequency partial volatility, $\sigma_t^{(d)} = \tilde{\sigma}_t^{(d)}$, we get the following model:

$$\sigma_{t+1}^{(d)} = c + \beta^{(d)}\sqrt{RV_t^{(d)}} + \beta^{(w)}\sqrt{RV_t^{(w)}} + \beta^{(m)}\sqrt{RV_t^{(m)}} + \tilde{\epsilon}_{t+1d}^{(d)} \quad (3.1)$$

Knowing that the realized volatility is an estimator of the latent volatility, we have

$$\sigma_{t+1}^{(d)} = \sqrt{RV_{t+1d}^{(d)}} + \tilde{\epsilon}_{t+1d}^{(d)} \quad (3.2)$$

From equation (3.2) we can link the ex post volatility estimate $\sqrt{RV_{t+1d}^{(d)}}$ to the contemporaneous measure of daily latent volatility $\sigma_{t+1}^{(d)}$. Substituting equation (3.2) in equation (3.1) (for simplicity $\sqrt{RV_{t+1d}^{(d)}} = \sqrt{RV_{t+1d}^{(d)}}$) we obtain a very simple time series representation of the cascade model,

$$\sqrt{RV_{t+1d}^{(d)}} = c + \beta^{(d)}\sqrt{RV_t^{(d)}} + \beta^{(w)}\sqrt{RV_t^{(w)}} + \beta^{(m)}\sqrt{RV_t^{(m)}} + \tilde{\epsilon}_{t+1d}^{(d)} \quad (3.3)$$

where the weekly and monthly realized volatility are computed respectively as the average of the last 5 and 22 daily volatilities,

$$\sqrt{RV_t^{(w)}} = \frac{1}{5} \left(\sqrt{RV_t^{(d)}} + \sqrt{RV_{t-1}^{(d)}} + \dots + \sqrt{RV_{t-4}^{(d)}} \right) \quad (3.4)$$

$$\sqrt{RV_t^{(m)}} = \frac{1}{22} \left(\sqrt{RV_t^{(d)}} + \sqrt{RV_{t-1}^{(d)}} + \dots + \sqrt{RV_{t-21}^{(d)}} \right) \quad (3.5)$$

The model is presented taking into account the realized volatility (the square root of the realized variance), similarly it could be written using the realized variance or the logarithmic transformation, respectively:

$$RV_{t+1}^{(d)} = c + \beta^{(d)} RV_t^{(d)} + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \tilde{\epsilon}_{t+1}^{(d)} \quad (3.6)$$

$$\ln(RV_{t+1}^{(d)}) = c + \beta^{(d)} \ln(RV_t^{(d)}) + \beta^{(w)} \ln(RV_t^{(w)}) + \beta^{(m)} \ln(RV_t^{(m)}) + \tilde{\epsilon}_{t+1}^{(d)} \quad (3.7)$$

Since the purpose of this thesis is the forecast of realized variance, we will focus directly on RV_t . One of the peculiarities of the HAR-RV model is its OLS nature, it is therefore easy to add in additional regressors to judge their explanatory power. In the following paragraphs we are going to extend the classic model in equation (3.6) with some exogenous variables.

3.1.2 HAR-RV-J-D and HAR-RV-J-F

To study the HAR-RV-J-D and HAR-RV-F models we need to consider the following jump diffusion model as seen in the theoretical Chapter 2,

$$dp_t = \mu_t dt + \sigma_t dW_t + \xi_t dq_t$$

where the last term is responsible for the jumps. The idea to include jumps in a HAR-RV model was initially proposed by Andersen *et al.* (2007). In particular they proposed to model the jumps and the continuous part separately in order to consider the different dynamics of the components of RV_t . Corsi *et al.* (2010) extended the HAR-RV with the use of C-Tz (2.36) to detect the occurrence of the jumps in a single day, with the use of threshold bipower variation to measure the continuous part of integrated volatility. The weekly and monthly components are computed analogically to the weekly and monthly aggregate realized variance as seen before in (3.4) and (3.5),

$$\begin{aligned}
J_t^{(w)} &= \frac{1}{5} \left(J_t^{(d)} + J_{t-1}^{(d)} + \dots + J_{t-4}^{(d)} \right) \\
J_t^{(m)} &= \frac{1}{22} \left(J_t^{(d)} + J_{t-1}^{(d)} + \dots + J_{t-21}^{(d)} \right) \\
C_t^{(w)} &= \frac{1}{5} \left(C_t^{(d)} + C_{t-1}^{(d)} + \dots + C_{t-4}^{(d)} \right) \\
C_t^{(m)} &= \frac{1}{22} \left(C_t^{(d)} + C_{t-1}^{(d)} + \dots + C_{t-21}^{(d)} \right)
\end{aligned}$$

In the empirical application, we consider the following two extensions to the standard HAR-RV. The first new model, the HAR-RV-J-D, considers only the daily jumps while the second model, the HAR-RV-J-F considers not only the daily, but also the weekly and monthly jumps. Respectively the two models are defined by

$$RV_{t+1}^{(d)} = c + \beta_{RV}^{(d)} RV_t^{(d)} + \beta_{RV}^{(w)} RV_t^{(w)} + \beta_{RV}^{(m)} RV_t^{(m)} + \beta_J^{(d)} J_t^{(d)} + \epsilon_{t+1} \quad (3.8)$$

$$\begin{aligned}
RV_{t+1}^{(d)} &= c + \beta_{RV}^{(d)} RV_t^{(d)} + \beta_{RV}^{(w)} RV_t^{(w)} + \beta_{RV}^{(m)} RV_t^{(m)} + \beta_J^{(d)} J_t^{(d)} \\
&\quad + \beta_J^{(w)} J_t^{(w)} + \beta_J^{(m)} J_t^{(m)} + \epsilon_{t+1}
\end{aligned} \quad (3.9)$$

HAR-RV-J-D and HAR-RV-J-F are based on the realized variance, however, it is possible to construct the models on realized volatility and log transformation. For this last measure we have to consider a correction due to days without jumps where the jumps components are equal to zero. The correction is simply done by the sum of 1 to $J_t^{(d)}$, $J_t^{(w)}$ and $J_t^{(m)}$

$$\begin{aligned}
\ln(RV_{t+1}^{(d)}) &= c + \beta_{RV}^{(d)} \ln(RV_t^{(d)}) + \beta_{RV}^{(w)} \ln(RV_t^{(w)}) + \beta_{RV}^{(m)} \ln(RV_t^{(m)}) \\
&\quad + \beta_J^{(d)} \ln(1 + J_t^{(d)}) + \epsilon_{t+1}
\end{aligned} \quad (3.10)$$

$$\begin{aligned}
\ln(RV_{t+1}^{(d)}) &= c + \beta_{RV}^{(d)} \ln(RV_t^{(d)}) + \beta_{RV}^{(w)} \ln(RV_t^{(w)}) + \beta_{RV}^{(m)} \ln(RV_t^{(m)}) \\
&\quad + \beta_J^{(d)} \ln(1 + J_t^{(d)}) + \beta_J^{(w)} \ln(1 + J_t^{(w)}) + \beta_J^{(m)} \ln(1 + J_t^{(m)}) + \epsilon_{t+1}
\end{aligned} \quad (3.11)$$

3.1.3 HAR-RV-CJ and HAR-RV-LCJ

Andersen *et al.* (2007) suggested another extension to the basic model (3.4). They proposed the HAR-RV-CJ model based on the explicit decomposition of the realized variance into the continuous part and jumps component. The HAR-RV-CJ model is simply defined as

$$RV_{t+1}^{(d)} = c + \beta_c^{(d)} C_t^{(d)} + \beta_c^{(w)} C_t^{(w)} + \beta_c^{(m)} C_t^{(m)} + \beta_j^{(d)} J_t^{(d)} + \beta_j^{(w)} J_t^{(w)} + \beta_j^{(m)} J_t^{(m)} + \epsilon_{t+1} \quad (3.12)$$

This decomposition allows us to understand the specific contribution of the continuous and jumps component that composed the daily realized variance.

It is natural to extend the HAR-RV-CJ model with the leverage effect to obtain the HAR-RV-LCJ model. Many studies show the relation between negative returns and volatility. For this reason it is straightforward to extend the volatility model with the leverage component. Figure 5 shows the relation between negative returns and realized variance where it is evident that periods with strong negative returns fall back into greater volatility.

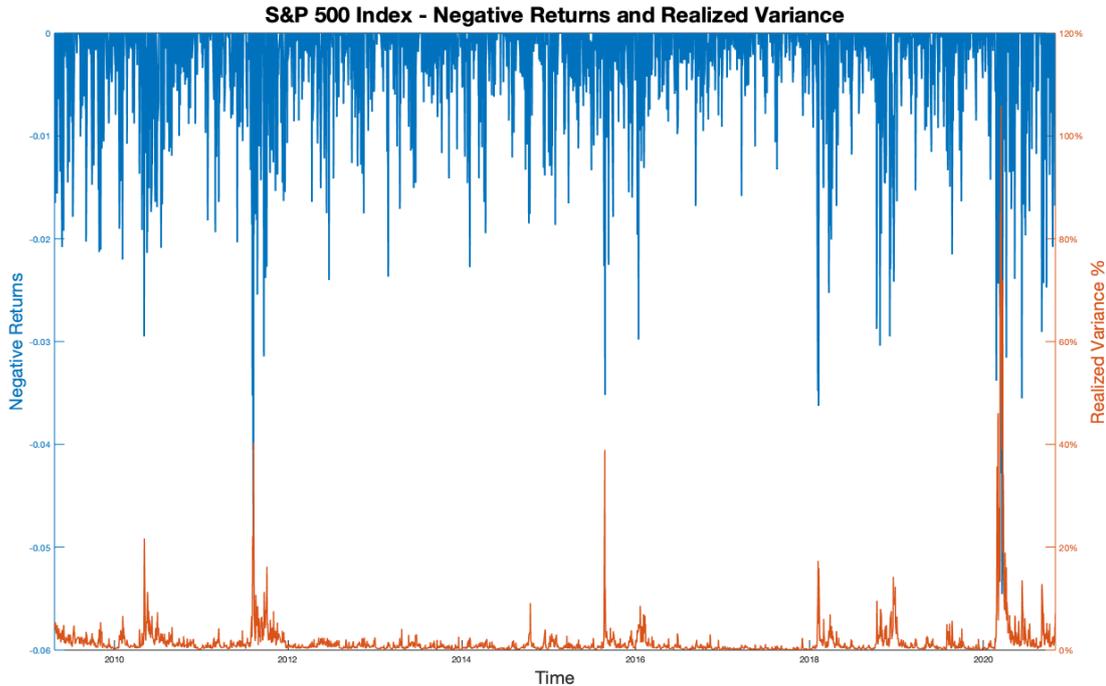


Figure 5: S&P 500 Index - Negative returns and Realized Variance

We model the heterogeneous leverage effects by introducing asymmetric return-volatility, which is dependent on each level of the volatility cascade. By defining daily returns as

$r_t = p_t - p_{t-1}$, where p_t is the logarithmic closing price at time t , we can separate the aggregate positive and negative returns as

$$r_{t+}^{(n)} = \frac{1}{n} (r_t + \dots + r_{t-n+1}) I_{\{(r_t + \dots + r_{t-n+1}) > 0\}} \quad (3.13)$$

$$r_{t-}^{(n)} = \frac{1}{n} (r_t + \dots + r_{t-n+1}) I_{\{(r_t + \dots + r_{t-n+1}) < 0\}} \quad (3.14)$$

where $I_{\{\cdot\}}$ denotes the indicator function and $n = 1, 5, 22$. Following Corsi *et al.* (2012) it is possible to add the aggregate negative return for the three horizons in the HAR-RV-CJ model obtaining the HAR-RV-LCJ model:

$$\begin{aligned} RV_{t+1}^{(d)} = & c + \beta_C^{(d)} C_t^{(d)} + \beta_C^{(w)} C_t^{(w)} + \beta_C^{(m)} C_t^{(m)} + \beta_J^{(d)} J_t^{(d)} + \beta_J^{(w)} J_t^{(w)} \\ & + \beta_J^{(m)} J_t^{(m)} + \gamma^{(d)} r_{t-}^{(d)} + \gamma^{(w)} r_{t-}^{(w)} + \gamma^{(m)} r_{t-}^{(m)} + \epsilon_{t+1} \end{aligned} \quad (3.15)$$

3.1.4 C-HAR

Another simple derivation from the HAR-RV is the C-HAR model. In the theoretical Chapter we saw that the main contribution to realized variance is from continuous part. Motivated by this, the C-HAR model is composed only by the daily, weekly and monthly continuous components

$$RV_{t+1}^{(d)} = c + \beta_C^{(d)} C_t^{(d)} + \beta_C^{(w)} C_t^{(w)} + \beta_C^{(m)} C_t^{(m)} + \epsilon_{t+1} \quad (3.16)$$

3.1.5 Q-Family

The next three models proposed have been developed by Bollerslev, Patton and Quaadvlieg (2015). The models directly exploit the asymptotic theory for high-frequency realized volatility estimation by allowing the dynamic parameters of the models to vary with the degree of estimation error in the realized volatility measure. At the end of the first Chapter we saw the distribution of realized variance proposed by Barndorff-Nielsen and Shephard (2002) and the consistent estimator of the realized quarticity as

$$RQ_t = \frac{n}{3} \sum_{j=1}^n r_{t,j}^4$$

Using RQ_t is possible to define the AR-Q model as

$$RV_{t+1}^{(d)} = c + (\beta_{RV}^{(d)} + \beta_Q^{(d)} RQ_t^{1/2(d)}) RV_t^{(d)} + \epsilon_{t+1} \quad (3.17)$$

where $RQ_t^{1/2}$ is the square root of the realized quarticity. The intuition is that when $\beta_Q^{(d)} < 0$ it follows that uninformative days with large measurement errors, will have smaller impact in the forecast than days where the realized variance is estimated only by $\beta_{RV}^{(d)}$. The next natural step would be to apply the AR-Q intuition to the standard HAR-RV obtaining the HAR-Q-F model:

$$\begin{aligned} RV_{t+1}^{(d)} = c + (\beta_{RV}^{(d)} + \beta_Q^{(d)} RQ_t^{1/2(d)}) RV_t^{(d)} + (\beta_{RV}^{(w)} + \beta_Q^{(w)} RQ_t^{1/2(w)}) RV_t^{(w)} \\ + (\beta_{RV}^{(m)} + \beta_Q^{(m)} RQ_t^{1/2(m)}) RV_t^{(m)} + \epsilon_{t+1} \end{aligned} \quad (3.18)$$

where $RQ_t^{1/2(d)}$, $RQ_t^{1/2(w)}$ and $RQ_t^{1/2(m)}$ are the daily, weekly and monthly aggregation of the square root of realized quarticity

$$RQ_t^{1/2(n)} = \frac{1}{n} (RQ_t^{1/2(d)} + \dots + RQ_{t-n+1}^{1/2(d)})$$

where $n = 1, 5, 22$.

However, the magnitude of the measurement errors in realized variance will generally decrease with the time horizon as the errors are averaged out. The consequence is that the adjustment for the measurement in daily error is likely to be more important than the weekly and monthly adjustment. Hence, we also consider the HAR-Q-D, where we only allow the coefficient of the lagged daily realized variance to vary in function of the square root of realized quarticity

$$RV_{t+1}^{(d)} = c + (\beta_{RV}^{(d)} + \beta_Q^{(d)} RQ_t^{1/2(d)}) RV_t^{(d)} + \beta_{RV}^{(w)} RV_t^{(w)} + \beta_{RV}^{(m)} RV_t^{(m)} + \epsilon_{t+1} \quad (3.19)$$

3.1.6 S-HAR

The S-HAR model, proposed by Patton and Sheppard (2011), decomposes the realized variance into positive and negative semivariance which is derived respectively from positive and negative returns

$$RV_{t-} = \sum_{t=1}^M r_i^2 I_{\{r_i < 0\}} \quad (3.20)$$

$$RV_{t+} = \sum_{t=1}^M r_i^2 I_{\{r_i > 0\}} \quad (3.21)$$

where $RV_t = RV_t^+ + RV_t^-$. With this decomposition we can extend the simple HAR-RV model to obtain the S-HAR model

$$RV_{t+1}^{(d)} = c + \beta_{RV^+}^{(d)} RV_{t+}^{(d)} + \beta_{RV^-}^{(d)} RV_{t-}^{(d)} + \beta_{RV}^{(w)} RV_t^{(w)} + \beta_{RV}^{(m)} RV_t^{(m)} + \epsilon_{t+1} \quad (3.22)$$

The expectation is that the coefficient of negative semivariance will be more significant and larger than that of positive semivariance.

3.1.7 HAR-X

As we have seen in the previous model, is it possible to add exogenous variables to the simple HAR-RV model of Corsi. The next model, besides considering the lagged daily, weekly and monthly value of RV_t , adds implied volatility as an explanatory variable. We can define the HAR-X as

$$RV_{t+1}^{(d)} = c + \beta_{RV}^{(d)} RV_t^{(d)} + \beta_{RV}^{(w)} RV_t^{(w)} + \beta_{RV}^{(m)} RV_t^{(m)} + \beta_X^{(d)} X_t^{(d)} + \epsilon_{t+1} \quad (3.22)$$

where $X_t^{(d)}$ is the daily implied volatility.

The most famous implied volatility index is the VIX, computed by CBOE⁵, also known as the *fear index*. VIX is a financial benchmark designed to be a market estimate of expected volatility of the S&P 500 Index, and is calculated by using the midpoint of real-time S&P

⁵ https://www.cboe.com/tradable_products/vix/

Index (SPX) option bid/ask quotes. The VIX index is intended to provide an instantaneous measure of how much the market thinks the S&P 500 Index will fluctuate in the 30 days ahead. The model-free estimator of the implied volatility that CBOE employs to calculate the VIX Index is

$$\sigma^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) - \frac{1}{T} \left(\frac{F}{K_0} - 1 \right)^2$$

where T is time to expiration, F is the forward index level derived from the index options prices, K_i is the strike price of the i th out-of-the-money option (either a call, if $K_i > F$, or a put, if $K_i < F$), $\Delta K_i = (K_{i+1} - K_{i-1})/2$ is the interval between strike prices minus half the difference between the strike on either side of K_i , K_0 is the first strike below the forward index level, r is the risk-free interest rate to expiration, and $Q(K_i)$ is the mid-quote for the option with strike of K_i . The VIX index is computed as $VIX = 100 * \sigma$. In the empirical application we use the implied volatility index linked to each of the financial instruments considered, as shown in Table 1

Financial Instrument	Implied volatility index
S&P 500 Index	CBOE Volatility Index (VIX)
iShares China Large-Cap ETF	China ETF VIX (VFXI)
Apple Inc	Apple VIX (VXAP)

Table 1: Implied volatility indices

3.1.8 Forecast Combinations

Based on the work of Bates and Granger (1969), forecast combinations have come to be viewed as a simple and effective way to improve the forecasting performance over that offered by individuals' models. The most significant challenge when using this approach, is the choice of weights that must be applied to the single models that will make up the forecast combination. According to Hendry and Clements (2001) a simple weighted average guarantees an "insurance". If two forecasts were differentially biased (one upward, one downward) combining them could result in an improvement. The HAR-Combo model will be the average of the expected realized variance computed by the other models

$$RV_{t+1}^{(Combo)} = \sum_{i=1}^U E [RV_{i,t+1}] \quad (3.23)$$

where U corresponds to the total number of the models considered⁶

3.1.9 Benchmark

One of the simplest and commonly used models in time series forecasting is the random walk (RW) model. The main intuition of the model is that at time t , the h step-ahead forecast is equal to the last observed value,

$$E[RV_{t+h}^{(RW)}] = RV_t \quad (3.24)$$

In the empirical application of Chapter 4 and 5, the random walk model, is used as a benchmark against which to compare the more sophisticated models presented above.

3.2 Forecast evaluation

3.2.1 Point forecast

Once we have a forecast the following natural step is the evaluation of the prediction. In many forecasting problems a recurrent dilemma is that the variable that we are trying to forecast is unobservable, even ex-post; volatility forecast is one such case. A common solution to this problem is to use a conditionally unbiased proxy as the realized variance estimator seen in the previous Chapter. However, according to Andersen *et al.* (1998) and Hansen and Lunde (2006) it is not always true that using a conditionally unbiased proxy leads to the same outcome as if the true, but latent, variable was used. Motivated by this, a solution proposed by Patton (2008) was to employ forecast loss functions less sensitive to large observations than the usual squared forecast error loss function (such as absolute error or proportional error loss function). The use of the aforementioned measures could lead to incorrect inferences and rankings of volatility forecasts.

The two forecast loss functions proposed by Patton (2008) are the Mean Square Error (MSE) and Quasi-Likelihood (QLike) loss functions, respectively

⁶ in the computation of the HAR-Combo model we do not include the benchmark

$$MSE(RV_{t+h}, \widehat{RV}_{t+h}) = (RV_{t+h} - \widehat{RV}_{t+h})^2 \quad (3.25)$$

$$QLike(RV_{t+h}, \widehat{RV}_{t+h}) = \log(\widehat{RV}_{t+h}) + \frac{RV_t}{\widehat{RV}_{t+h}} \quad (3.26)$$

where RV_{t+h} is the observed realization at time h and \widehat{RV}_{t+h} is the forecast of one of the treated models. Another measure commonly used in literature is R^2 , of the Mincer-Zarnowitz regression

$$RV_{t+h} = \alpha + \beta \widehat{RV}_{t+h} + \epsilon_{t+h} \quad (3.27)$$

In this case, as opposed to (3.25) and (3.26) where low relative values indicate better models, high values are preferred.

3.2.1.1 Diebold Mariano test & Model Confidence Set

When we want to make a comparison between two forecasts originated from two different models, a common approach is to use the Diebold-Mariano (DM) test (1995). This test is intended for comparing predictive accuracy of two different forecasts. Precisely, to determine whether the difference in accuracy is statistically significant. The authors proposed a widely applicable test of the null hypothesis of no difference in accuracy between two forecasts. Mathematically, we can indicate the observed values as $\{y_t; t = 1, \dots, T\}$ and the two forecasts as $\{\hat{y}_{1t}; t = 1, \dots, T\}$ and $\{\hat{y}_{2t}; t = 1, \dots, T\}$. The forecast errors are defined as $e_{it} = y_t - \hat{y}_{it}, i = 1, 2$. It is also assumed that the loss associated to the forecast i is a function of the forecast error, e_{it} , denoted as $g(e_{it})$. We can now define the loss differential between the two forecasts by $d_t = g(e_{1t}) - g(e_{2t})$. The hypothesis of the test are

$$H_0: E(d_t) = 0, \forall t \quad (3.28)$$

$$H_1: E(d_t) \neq 0 \quad (3.29)$$

To test this null hypothesis it is possible to use a simple asymptotic z-test. The DM test statistic can be obtained through

$$DM = \frac{\bar{d}_t}{\sqrt{\widehat{Var}(d_t)/T}} \quad (3.30)$$

Where \bar{d}_t is the sample average of the loss differential series and $\widehat{Var}(d_t)$ is a consistent estimator of asymptotic variance of \bar{d}_t obtained with a weighted sum of the available sample autocovariances. Under the null hypothesis the test statistic is asymptotically $N(0,1)$ distributed.

Contrary to the widely used Diebold-Mariano test, that is used to compare only a pair of models, the Model Confidence Set (MCS) proposed by Hansen, Lunde and Nason (2011) allows for comparison of multiple forecast models simultaneously. With MCS it is possible to reduce the set of models to a smaller set known as the *model confidence set* which contains the best model with a given level of confidence. As defined by Hansen *et al.* (2011) the objective of MCS procedure is to determine the set of models, \mathcal{M}^* , that consists of the best models (or model) from a collection of models, \mathcal{M}^0 , where *best* is defined in terms of loss function defined by the user. The procedure is based on an equivalence test, $\delta_{\mathcal{M}}$, and an elimination rule, $e_{\mathcal{M}}$. When $\delta_{\mathcal{M}}$ is rejected there is evidence that the objects in \mathcal{M} are not equally “good” and $e_{\mathcal{M}}$ is used to eliminate an object with poor performance. This procedure is repeated until $\delta_{\mathcal{M}}$ is not rejected and the MCS is now defined by the surviving objects. Following Hansen *et al.* (2011) we consider a set, \mathcal{M}^0 , composed by a finite number of objects, $i = 1, \dots, m_0$. The objects are evaluated according to a loss function $L_{i,t}$ where t is the period. The relative performance variables are defined as $d_{ij,t} \equiv L_{i,t} - L_{j,t}$, for all $i, j \in \mathcal{M}^0$. The set of superior objects is defined by

$$\mathcal{M}^* = \{i \in \mathcal{M}^0: \mu_{i,j} \leq 0 \text{ for all } j \in \mathcal{M}^0\} \quad (3.31)$$

where $\mu_{i,j} = E(d_{ij,t})$ is finite and does not depend on time for all $i, j \in \mathcal{M}^0$. The aim of the method is to determine the set of superior models, \mathcal{M}^* , which can be done via a sequence of significance test where the models that are found to be significantly inferior to the other are eliminated. The null hypothesis is $H_{0,\mathcal{M}}: \mu_{i,j} = 0, \forall i, j \in \mathcal{M}$ with $\mathcal{M} \subseteq \mathcal{M}^0$. The equivalence test, $\delta_{\mathcal{M}}$, is used to test the hypothesis $H_{0,\mathcal{M}}$ for any $\mathcal{M} \subset \mathcal{M}^0$, and the elimination rule, $e_{\mathcal{M}}$, identifies inferior model from \mathcal{M} and remove it when the null hypothesis is rejected. The surviving models end up in $\widehat{\mathcal{M}}_{1-\alpha}^*$. The MCS algorithm is composed by the following steps:

Step 0. Initially set $\mathcal{M} = \mathcal{M}^0$.

Step 1. Test $H_{0,\mathcal{M}}$ using $\delta_{\mathcal{M}}$ at level α .

Step 2. If $H_{0,\mathcal{M}}$ is accepted, define $\widehat{\mathcal{M}}_{1-\alpha}^* = \mathcal{M}$; otherwise, use $e_{\mathcal{M}}$ to eliminate an object from \mathcal{M} and repeat the procedure from Step 1.

In order to test the hypothesis, the two following statistics are constructed

$$t_{i,j} = \frac{\bar{d}_{i,j}}{\sqrt{\widehat{Var}(\bar{d}_{i,j})}} \quad \text{and} \quad t_{i,j} = \frac{\bar{d}_i}{\sqrt{\widehat{Var}(\bar{d}_i)}} \quad \text{for } i, j \in \mathcal{M} \quad (3.31)$$

Where $\bar{d}_{i,j}$ measure the relative sample loss between the i th and j th models, \bar{d}_i is the sample loss of the i th model relative to the average across models in \mathcal{M} , while $\widehat{Var}(\bar{d}_{i,j})$ and $\widehat{Var}(\bar{d}_i)$ are bootstrapped estimates of $Var(\bar{d}_{i,j})$ and $Var(\bar{d}_i)$. The first statistic is the well-known Diebold and Mariano test (1995) seen above. The MCS procedure consists of a sequential testing procedure, which eliminates the worst model at each step, until the hypothesis is accepted for all the models. To eliminate the worst model, we can use an elimination rule coherent with the test statistic

$$e_{max,\mathcal{M}} = \arg \max_{i \in \mathcal{M}} \frac{\bar{d}_t}{\sqrt{\widehat{Var}(d_t)}}, \quad e_{max,\mathcal{M}} = \arg \max_{i \in \mathcal{M}} \left\{ \sup_{j \in \mathcal{M}} \frac{\bar{d}_t}{\sqrt{\widehat{Var}(d_t)}} \right\} \quad (3.32)$$

This procedure yields p -values for all forecast models considered. The MCS p -values are useful to determine which forecast models are included in $\widehat{\mathcal{M}}_{1-\alpha}^*$ at any significance level.

3.2.2 Density forecast

The previous forecast evaluation methodologies were based on the comparison between the prevision and the actual realization. The next step is to study the uncertainty around point forecasts using the so called density forecasts comparison. The advantage of this last approach is that it provides a complete description of the uncertainty around the predicted value. Diebold, Gunther and Tay (1998) evaluate density forecasts with the use of probability integral transform (PIT). Following Rossi (2014) a PIT is the cumulative probability evaluated at the actual realized value of the target variable. According to

Diebold *et al.* (1998) the PIT is uniform, independent and identically distributed if the density forecast is correctly specified. Justified by the difficulty to test for uniformity, especially in small data samples and deriving parametric tests, Berkowitz (2001) proposed a simple transformation to normality. Under the normality it is straightforward to calculate the Gaussian likelihood and construct tests around it.

3.2.2.1 Berkowitz test

Let y_t be a stochastic process that is forecasted at time $t - 1$. Let the probability density of y_t be $f(y_t)$ and the distribution function be $F(y_t) = \int_{-\infty}^{y_t} f(u)du$. It is possible to transform all the realizations into a series of independent and identically distributed random variables using the Rosenblatt (1952) transformation:

$$x_t = \int_{-\infty}^{y_t} \hat{f}(u)du = \hat{F}(y_t) \quad (3.33)$$

Where y_t is the true realization and $\hat{f}(\cdot)$ is the ex-ante forecasted loss density. Under correct model specification x_t should be distributed according to the standard uniform distribution, i.e. $x_t \sim U(0,1)$. A variety of tests would then be available both for independence and for uniformity, for example, the Kolmogorov-Smirnoff (KS) test. However, Berkowitz (2001) proposed a simple transformation to normality.

Let $\Phi^{-1}(\cdot)$ be the inverse of the standard normal distribution function, then the transformation is defined as

$$z_t = \Phi^{-1}[\hat{F}(y_t)], \quad (3.34)$$

Where z_t should be standard normal (i.e. $N(0,1)$). Using a simple LR test it is possible to assess the correct coverage of the full distribution where the null hypothesis is

$$H_0: z_t \sim N(0,1) \quad (3.35)$$

The test statistic is

$$LR_{full} = -2(\ell(0,1) - \ell(\hat{\mu}, \hat{\sigma}^2)) \quad (3.36)$$

which under the null is distributed as a $\chi^2(2)$ (the hats denote estimated values).

Moreover, Berkowitz (2001), proposed a test to assess the right coverage of one tail and not the entire distribution, using a censored likelihood. For example, let the desired cutoff point be 2.32, the 99% quantile for the standard Gaussian distribution, then the new variable of interest is

$$z_t^* = \begin{cases} 2.32 & \text{if } z_t \leq 2.32 \\ z_t & \text{if } z_t > 2.32 \end{cases}$$

As before, is possible to evaluate the correct fit using the LR test, where the null hypothesis again requires that $\mu = 0, \sigma^2 = 1$.

However, in the density forecast comparison in the empirical applications we are going to assess the normality assumption on the entire Gaussian distribution without focusing on a specific quantile.

4 In-sample and Out-of-sample analysis

This Chapter is dedicated to the empirical estimation and forecasting of realized variance of three different time series: S&P500 Index, iShares China Large-Cap ETF and Apple Inc. All the data was downloaded from the data provider barchart.com. For each financial instrument we compute the realized variance, the continuous and the jump component using the method proposed by Corsi *et al.* (2012). After that, we use the twelve models seen in Chapter 3 to carry out a full in-sample analysis in order to assess whether a model can be considered superior in fitting the data. However, the main point of interest is the forecast analysis in order to determine the out-of-sample prediction of each models. The direct forecasts are performed on $h=1, 5, 10,$ and 22 days-ahead corresponding to one day, one week, two weeks and one month utilizing a 1000 day rolling window. As suggested by Swanson and White (1995, 1997), “ignorance” is better than “insanity”, for this reason each forecast is filtered by an “insanity filter” in order to avoid crazy forecasts. The evaluation of the forecasts is made considering both the point and density forecast comparison. The point forecast analysis is computed individually on the basis of R^2 from the Mincer-Zarnowitz regressions, MSE and QLike. To make it easier to read the results, the outcome of the two loss functions are standardized by the loss of the RW model. These results are accompanied by an average score (from 1, best, to 13 worst). The simultaneous comparison is carried out by the Model Confident Set both using the MSE and QLike. The density forecast comparisons are based on the Berkowitz test as seen at the end of Chapter 3.

4.1 S&P 500 Index

The Standard & Poor’s 500 Index, commonly known as S&P 500 Index, is a market-capitalization-weighted index of the 505 largest publicly traded companies in the U.S.. The index accounts for 80% of the market value of the U.S.’s equity and it is also considered to be the best representative of the U.S economy. The dataset consists of five-minute observation, considering the canonical trading hours 09:30:00 – 16:00:00, from April 24, 2009 to October 30, 2020 for a total of 2870 days and 226730 observations. The five-minutes log-returns, that are the basis to get the realized variance, were calculated as $r_{t,h} = \ln(P_t/P_{t,h})$, where P_t is the price of the index at time t and h is the five-minute interval.

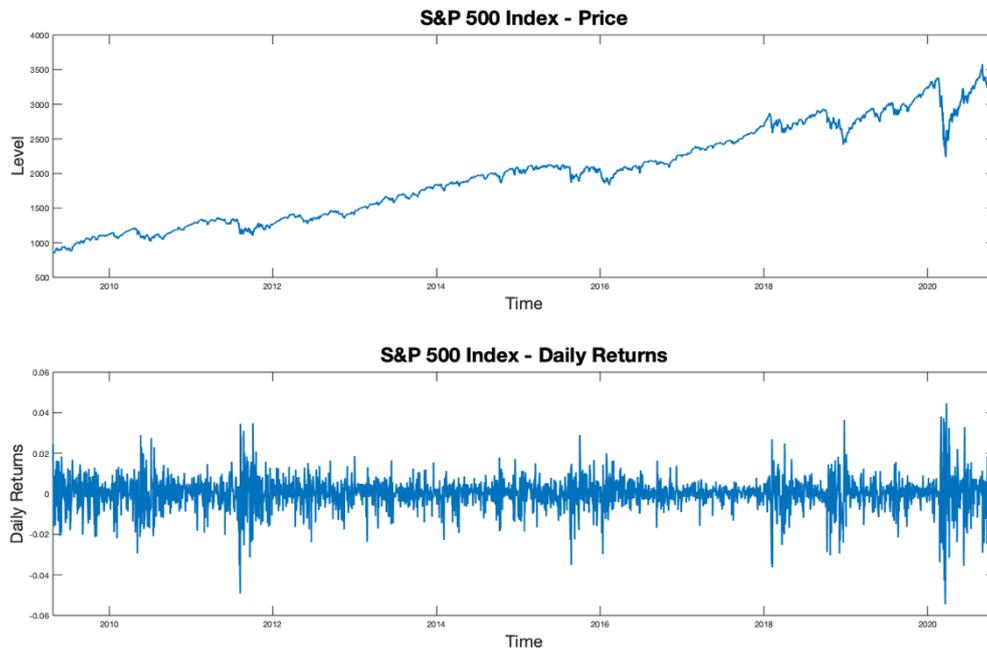


Figure 6: S&P 500 Index - Daily price level and daily returns

Through equation (2.13) is possible to compute the realized variance, while to calculate the continuous and jumps component we use the method proposed by Corsi *et al.* (2012). From Figure 7 is immediately possible to see the dynamics of the three previously mentioned measures during the full period. The realized variance is characterized by four peaks in volatility. The first three are a direct consequence of the Greek crisis (May 2010), sovereign debt crisis (August, 2011) and the so-called China Black Monday (August 2015). The last, but the highest peak in realized variance, was recorded in at the end of February 2020 due to the spread of a new coronavirus called Covid-19. In just 23 days the S&P 500 Index has experienced the greatest market corrections in its history, -33.9%. However this was followed by an impressive rebound from April through August, bringing markets back to their highest recorded level. Nonetheless, the fear of a new wave of infections and new possible lockdown restrictions is reflected in the high levels of volatility in the last part of the dataset. For a better understanding of the data, Table 2 shows the described statistics of RV_t , C_t and J_t based both on the entire time horizon considered and on the Covid-19 pandemic crisis period. The days with jumps are 1518, which makes up approximately 53% of the entire dataset.

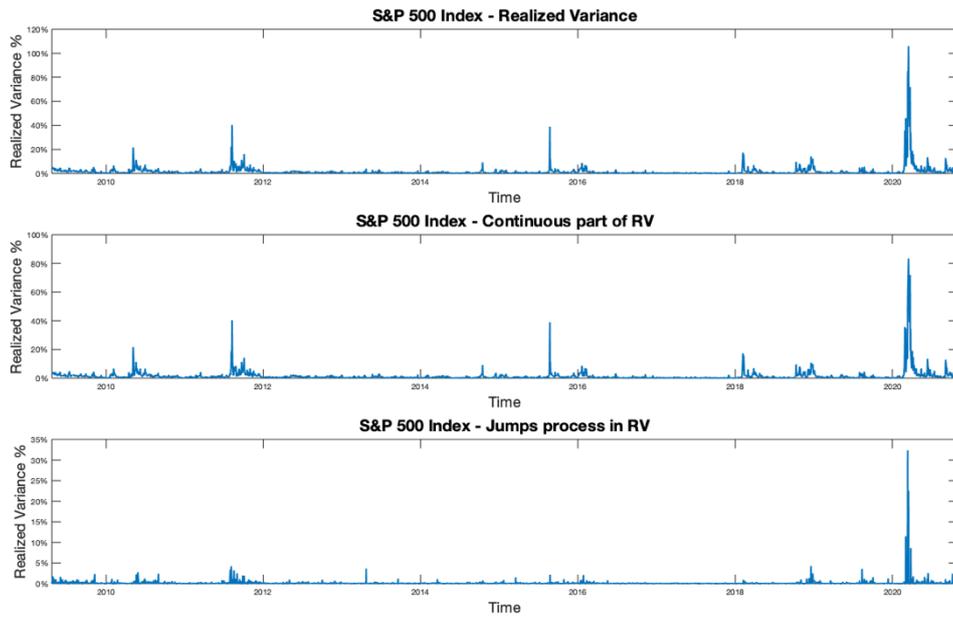


Figure 7: S&P 500 Index - Realized Variance, Continuous component and Jumps process

Full period: 24/04/2009 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	1.4255×10^{-4}	0.016	0.014	0.001
Std. dev.	0.0077	0.044	0.040	0.008
Skewness	-0.466	12.617	11.464	28.108
Kurtosis	8.110	214.600	172.002	961.144
Min	-0.054	4.4220×10^{-4}	4.4220×10^{-4}	0
Max	0.044	1.059	0.837	0.319
Covid-19 period: 24/02/2020 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	-1.2484×10^{-4}	0.075	0.068	0.006
Std. dev.	0.014	0.150	0.133	0.031
Skewness	-0.218	3.890	3.542	8.106
Kurtosis	4.860	19.784	16.136	73.597
Min	-0.054	0.002	0.002	0
Max	0.044	1.059	0.837	0.319

Table 2: S&P 500 Index - Descriptive statistics of r_t , RV_t , C_t and J_t .

4.1.1 In-sample

Tables 3.a and 3.b contain the results of the estimation performed on the S&P 500 Index over the full sample. A very first comment concerns the coefficients of the HAR-RV model. We can see that the daily and weekly realized variance component have the greatest impact on daily realized variance while the impact of the monthly component is negative and less impactful. As shown in Figure 8, the dynamics of the parameters over time are computed with a rolling window of 1000 observations.

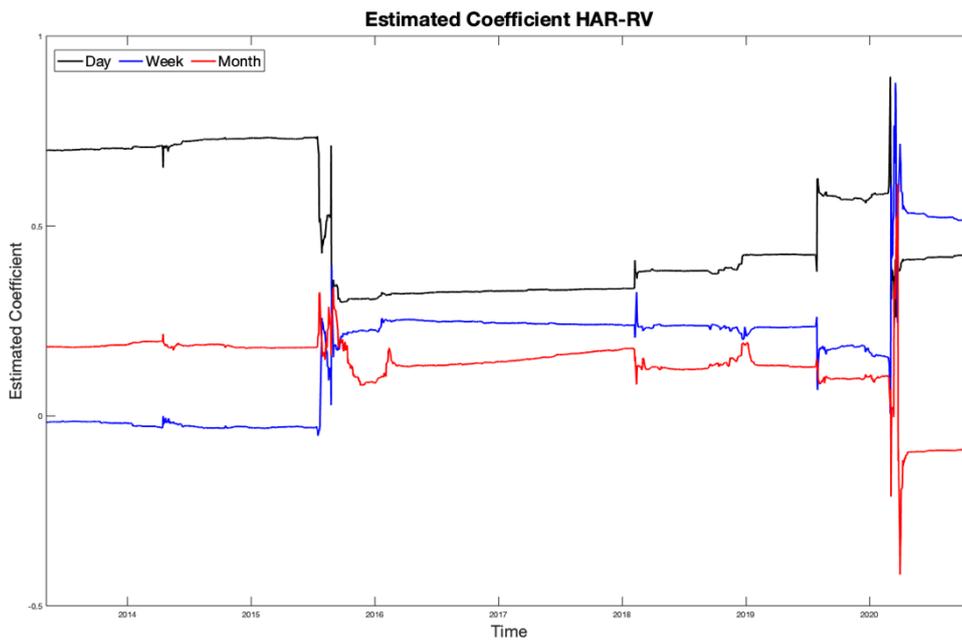


Figure 8: S&P 500 Index – HAR-RV coefficients based on a rolling window of 1000 days.

We can see that after the Chinese Black Monday, the daily parameter has lost about half of its magnitude while the weekly increased from nearly 0 to 0.3. Noteworthy, the weekly parameter increases importance in the period of high volatility after the Covid-19 crisis. The comparison between the HAR-RV and C-HAR models is quite interesting. The results show that the model which considers only the continuous component achieves slightly better in-sample results. Continuing the analysis to other models it is possible to see that the results of the Q-models (AR-Q, HAR-Q-D and HAR-Q-F) are consistent with Bollerslev, Patton and Quaedvlieg (2015). The $\beta_Q^{(d)}$ coefficients are negative and strongly statistically significant. This is in line with the intuition that when the measurement error and the value of RQ increases, the informativeness of the current RV for the future RV decreases. Considering the time-varying nature of the measurement error for all the frequencies in the HAR-Q-F model we get a slightly better result in R^2 and MSE but worse regarding QLike loss function.

Among the Q-family models, the one with the best in-sample fit is the HAR-Q-D. Taking into account the models that include jumps (HAR-RV-J-D, HAR-RV-J-F, HAR-RV-CJ and HAR-RV-LCJ) we can immediately notice how the three jump coefficients are significant for all the models. Importantly, the estimates of the jumps components $\beta_j^{(d)}$ and $\beta_j^{(m)}$ are negative while the weekly coefficient, $\beta_j^{(w)}$, is positive and more significant than the other two. The fact that the realized variance decreases after a jump is in contrast with the results of Corsi *et al.* (2012) who have found a positive contribution by the daily, weekly and monthly jumps. A simple explanation for this is that a jump usually implies a one-off spike upward in realized variance that does not persist on that volatility level. However, a negative coefficient for the daily jumps is consistent with the work of Andersen *et al.* (2007). Comparing HAR-RV-J-D and HAR-RV-J-F we can see that the inclusion of the weekly and monthly jumps component helps to achieve better fitting results. The leverage effect is confirmed by the in-sample results where daily negative returns affect next day volatility. The HAR-Combo is the model with the best in-sample fit.

	HAR-RV	AR-Q	HAR-Q-D	HAR-Q-F	HAR-RV- J-D	HAR-RV- J-F
c	9.291 ⁻² (4.208)	4.69X10 ⁻⁶ (2.080)	4.69X10 ⁻⁶ (2.077)	-2.69X10 ⁻⁶ (-0.960)	9.41X10 ⁻⁶ (3.975)	8.59X10 ⁻⁶ (3.975)
$\beta_{RV}^{(d)}$	0.464 (20.959)	0.962 (50.280)	0.646 (20.523)	0.641 (19.145)	0.563 (20.135)	0.546 (19.795)
$\beta_{RV}^{(w)}$	0.433 (0.433)		0.379 (12.41)	0.263 (5.231)	0.368 (11.502)	0.198 (5.516)
$\beta_{RV}^{(m)}$	-0.046 (-1.978)		-0.072 (-3.064)	0.198 (3.766)	-0.039 (-1.694)	0.157 (2.819)
$\beta_C^{(d)}$						
$\beta_C^{(w)}$						
$\beta_C^{(m)}$						
$\beta_J^{(d)}$					-0.479 (-5.7443)	-0.736 (-8.436)
$\beta_J^{(w)}$						2.197 (10.004)
$\beta_J^{(m)}$						-1.91 (-3.048)
$\gamma^{(d)}$						
$\gamma^{(w)}$						
$\gamma^{(m)}$						
$\beta_{RV+}^{(d)}$						
$\beta_{RV-}^{(d)}$						
$\beta_Q^{(d)}$		-0.611 (-12.432)	-0.412 (-8.033)	-0.429 (-7.569)		
$\beta_Q^{(w)}$				0.379 (3.30)		
$\beta_Q^{(m)}$				-1.262 (-5.604)		
$\beta_X^{(d)}$						
R^2	0.631	0.618	0.639	0.643	0.6358	0.6483
MSE	0.811	0.841	0.793	0.784	0.8021	0.7745
QLike	0.947	0.947	0.936	0.959	0.9439	0.9511
Avg. Score	9	10	5.33	7.33	6.67	5

Table 3.a: S&P 500 Index - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

	C-HAR	S-HAR	HAR-RV- CJ	HAR-RV- LCJ	HAR-X	HAR- Combo	RW
c	9.66x10 ⁻⁶ (4.373)	9.26x10 ⁻⁶ (4.188)	8.59x10 ⁻⁶ (3.975)	-9.60x10 ⁻⁶ (4.188)	-1.70x10 ⁻⁵ (4.188)		
$\beta_{RV}^{(d)}$					0.346 (14.705)		
$\beta_{RV}^{(w)}$		0.433 (14.347)			0.343 (11.324)		
$\beta_{RV}^{(m)}$		-0.046 (-1.973)			-0.2395 (-8.638)		
$\beta_c^{(d)}$	0.625 (24.478)		0.546 (19.795)	0.461 (16.618)			
$\beta_c^{(w)}$	0.328 (9.604)		0.198 (5.516)	0.211 (5.837)			
$\beta_c^{(m)}$	-0.025 (-0.978)		0.157 (2.819)	0.141 (2.229)			
$\beta_J^{(d)}$			-0.189 (-2.527)	-0.281 (-3.847)			
$\beta_J^{(w)}$			2.396 (11.642)	2.579 (12.76)			
$\beta_J^{(m)}$			-1.752 (-3.040)	-1.393 (-2.398)			
$\gamma^{(d)}$				-0.0048 (-10.78)			
$\gamma^{(w)}$				-0.0021 (-1.836)			
$\gamma^{(m)}$				0.00022 (1.252)			
$\beta_{RV+}^{(d)}$		0.442 (7.459)					
$\beta_{RV-}^{(d)}$		0.484 (8.438)					
$\beta_Q^{(d)}$							
$\beta_Q^{(w)}$							
$\beta_Q^{(m)}$							
$\beta_X^{(d)}$					0.341 (12.470)		
R^2	0.630	0.6316	0.648	0.666	0.646	0.6526	0.574
MSE	0.813	0.8113	0.774	0.733	0.780	0.7655	1
QLike	0.943	0.9467	0.951	1.1043	1.1179	0.9269	1
Avg. Score	8.33	7.67	5.33	4.67	7.67	1.67	12.33

Table 3.b: S&P 500 Index - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

4.1.2 Out-of-sample

In this section we evaluate the forecasting performance of the 13 models on the basis of an out-of-sample analysis. The forecasts are based on re-estimating the parameters of the different models, each day with a fixed length rolling windows of 1000 days. The forecast adopted in this thesis is a direct forecast, this way is possible to avoid the forecast of exogenous variables. The forecasts are performed on $h=1, 5, 10$ and 22 days-ahead corresponding to one day, one week, two weeks and one month. In Table 6 we can see the results of the individual forecast analysis based on R^2 , MSE and QLike. An average result of the model based on the three previously motioned measures is also reported. The results obtained from a simultaneous comparison of all the models conducted by the Model Confidence Set, are shown in Table 4. Starting from $h=1$, the more sophisticated models, such as those that include the jumps components, the Q-family (excluding the simplest AR-Q model), and the HAR-X, are the worst. Indeed the benchmark model achieved better results compared to these models.

MCS - MSE							
$h=1$		$h=5$		$h=10$		$h=22$	
Model	P-value	Model	P-value	Model	P-value	Model	P-value
HAR-RV-CJ	0.2913	RW	0.2212	RW	0.5445	RW	0.1361
HAR-RV-J-F	0.3554	HAR-Q-F	0.2723	HAR-Q-D	0.7087	HAR-Q-F	0.1692
HAR-RV-LCJ	0.5035	HAR-RV-J-D	0.3303	HAR-RV	0.7327	AR-Q	0.2462
HAR-RV-J-D	0.6627	C-HAR	0.3353	C-HAR	0.7588	HAR-Q-D	0.2773
S-HAR	0.8008	S-HAR	0.3463	S-HAR	0.7588	HAR-X	0.3463
HAR-Q-D	0.8358	AR-Q	0.4184	HAR-X	0.7588	HAR-RV-J-D	0.3554
HAR-Q-F	0.8599	HAR-RV	0.5165	HAR-Q-F	0.7588	S-HAR	0.3684
RW	0.8889	HAR-Q-D	0.6657	HAR-RV-LCJ	0.7588	HAR-RV	0.4404
HAR-X	0.8889	HAR-RV-LCJ	0.8228	HAR-RV-CJ	0.7588	C-HAR	0.5175
HAR-RV	0.8889	HAR-RV-CJ	0.8228	HAR-RV-J-F	0.7588	HAR-Combo	0.5836
C-HAR	0.8889	HAR-RV-J-F	0.8228	HAR-RV-J-D	0.7588	HAR-RV-J-F	0.5836
AR-Q	0.8889	HAR-X	0.8228	AR-Q	0.7588	HAR-RV-CJ	0.5836
HAR-Combo	1	HAR-Combo	1	HAR-Combo	1	HAR-RV-LCJ	1
MCS - QLike							
$h=1$		$h=5$		$h=10$		$h=22$	
Model	P-value	Model	P-value	Model	P-value	Model	P-value
HAR-RV-LCJ	0	RW	0.0110	RW	0.0060	RW	0.0581
HAR-X	0.0020	HAR-RV-LCJ	0.0490	HAR-Q-F	0.1522	HAR-Q-F	0.2112
HAR-Q-F	0.0280	HAR-Q-F	0.0611	HAR-RV-LCJ	0.1592	S-HAR	0.2302
RW	0.0280	S-HAR	0.0611	S-HAR	0.1832	HAR-Q-D	0.2733
HAR-RV-CJ	0.0320	HAR-RV-J-D	0.0611	HAR-Q-D	0.1832	HAR-X	0.2893
HAR-RV-J-F	0.0320	HAR-RV	0.0611	HAR-RV-CJ	0.1832	HAR-RV	0.2893
S-HAR	0.0320	HAR-RV-CJ	0.0611	HAR-RV-J-D	0.1832	C-HAR	0.3844
HAR-RV-J-D	0.0320	HAR-Q-D	0.0611	AR-Q	0.1832	HAR-RV-J-D	0.4454
HAR-Q-D	0.0320	HAR-RV-J-F	0.0611	HAR-RV	0.1832	AR-Q	0.8068
AR-Q	0.0320	AR-Q	0.0611	HAR-RV-J-F	0.2703	HAR-RV-LCJ	0.9820
C-HAR	0.0320	C-HAR	0.0611	C-HAR	0.2703	HAR-Combo	0.9820
HAR-RV	0.0320	HAR-X	0.1231	HAR-X	0.2703	HAR-RV-CJ	0.9820
HAR-Combo	1	HAR-Combo	1	HAR-Combo	1	HAR-RV-J-F	1

Table 4: S&P 500 Index – Model Confidence Set based on MSE and Q-Like loss function computed with 999 bootstrap replications, average block length of 5 and $\alpha = 0.05$. The included models are all of these with $p\text{-value} > \alpha$

We can affirm that for $h=1$ “simple is better”, in fact, the models that have obtained better results are the simple HAR-RV, C-HAR and AR-Q models. However, the preferred model for one day ahead forecast is the HAR-Combo. These results are also confirmed by the MCS. For one and two weeks-ahead forecast ($h=5, 10$) the results are quite similar. Models that include jumps and VIX get better forecasting performance as forecast time gets longer. The random walk model gets the worst results on both time horizons; results which are also shown by the simultaneous analysis. Also in these two cases the best model is the one that encloses them together, the HAR-Combo. Moving to the last forecast horizon considered, it is possible to deduce that jumps matter. Models that include the jumps component achieved a considerable improvement over shorter forecasting horizons. This conclusion is also confirmed by the results of MCS both using MSE and QLike loss function.

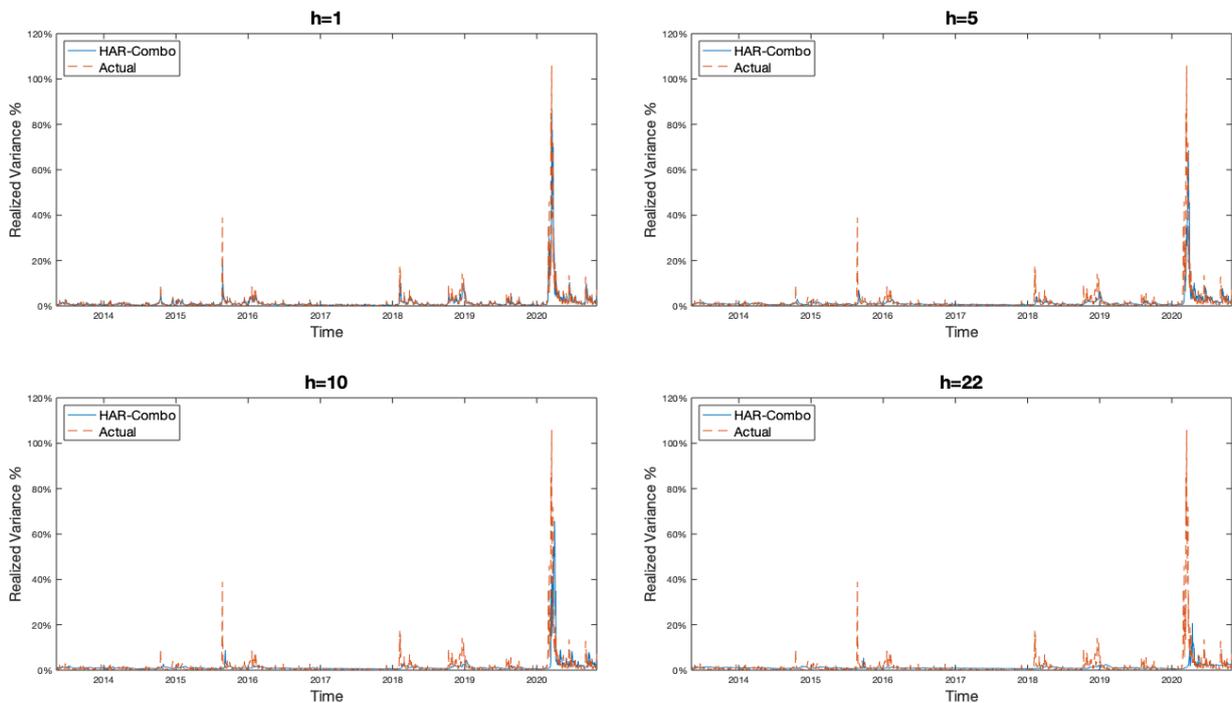


Figure 9: S&P 500 Index: HAR-Combo out-of-sample forecasts

Summarizing the above results we can say that with the increasing of forecast horizons the jumps allow us to obtain better predictions, especially when we consider the daily, weekly and monthly frequency. In a comparison *vis-à-vis* between the HAR-RV and C-HAR, the model that considers only the continuous components beats the classic one. Moving to the Q-family models, taking into account the time-varying nature of the measurement error does not help to get brilliant forecasts compared with other models especially in the one

month ahead-forecast. The model that gets the finest overall performance is the HAR-Combo. Figure 9 shows the out-of-sample forecast of the HAR-Combo for all the forecasting horizons considered. Table 5 reports the p-values of the Berkowitz test relative to the out-of-sample forecast of different models. It is possible to see that the null hypothesis is not rejected, at 5% significance level, for the HAR-RV, HAR-X, HAR-RV-LCJ, S-HAR, C-HAR and HAR-Q-D when $h=10$.

	$h=1$	$h=5$	$h=10$	$h=22$
HAR-RV	0.0000	0.0000	0.0819	0.0000
AR-Q	0.0000	0.0000	0.0000	0.0000
HAR-Q-D	0.0000	0.0000	0.3151	0.0000
HAR-Q-F	0.0000	0.0000	0.0000	0.0000
HAR-RV-J-D	0.0000	0.0068	0.0000	0.0000
HAR-RV-J-F	0.0000	0.0000	0.0000	0.0000
C-HAR	0.0000	0.0000	0.3830	0.0000
S-HAR	0.0000	0.0000	0.6062	0.0000
HAR-RV-CJ	0.0000	0.0000	0.0120	0.0000
HAR-RV-LCJ	0.0000	0.0000	0.5424	0.0000
HAR-X	0.0000	0.0000	0.3636	0.0000
RW	0.0000	0.0387	0.0000	0.0000

Table 5: S&P 500 Index – P-value of the Berkowitz test on the entire standard Gaussian distribution, relative to the out-of-sample analysis.

	$h=1$					$h=5$					$h=10$					$h=22$								
	R^2	MSE	QLike	Avg.S.		R^2	MSE	QLike	Avg.S.		R^2	MSE	QLike	Avg.S.		R^2	MSE	QLike	Avg.S.		R^2	MSE	QLike	Avg.S.
HAR-RV	0.594	0.989	0.941	3		0.344	0.951	0.86	7.67		0.204	0.885	0.722	8		0.014	0.545	0.541	5.33		0.014	0.545	0.541	5.33
AR-Q	0.561	0.977	0.956	4.33		0.227	0.962	0.866	9.67		0.171	0.727	0.728	6.33		0.008	0.550	0.533	8.67		0.008	0.550	0.533	8.67
HAR-Q-D	0.548	1.090	0.961	7		0.346	0.815	0.869	7.33		0.162	0.877	0.732	10.33		0.009	0.551	0.552	10		0.009	0.551	0.552	10
HAR-Q-F	0.508	1.099	1.046	10		0.160	1.048	0.939	12.33		0.059	0.830	0.876	11		0.002	0.5641	0.651	11		0.002	0.5641	0.651	11
HAR-RV-J-D	0.537	1.158	0.967	8.67		0.315	1.003	0.859	8.67		0.210	0.754	0.724	5		0.013	0.547	0.535	6.67		0.013	0.547	0.535	6.67
HAR-RV-J-F	0.4972	1.313	0.9703	10.33		0.409	0.733	0.856	3.67		0.211	0.756	0.722	4.33		0.022	0.538	0.538	3.67		0.022	0.538	0.538	3.67
C-HAR	0.592	0.988	0.9431	3.33		0.337	0.950	0.843	6.33		0.207	0.844	0.713	6.67		0.015	0.543	0.539	4.33		0.015	0.543	0.539	4.33
S-HAR	0.564	0.497	0.9656	6.33		0.252	1.057	0.873	11.33		0.215	0.823	0.742	6.67		0.012	0.548	0.550	8.67		0.012	0.548	0.550	8.67
HAR-RV-CJ	0.497	1.314	0.9704	10.67		0.408	0.734	0.859	5.00		0.211	0.757	0.731	6		0.023	0.538	0.521	4		0.023	0.538	0.521	4
HAR-RV-LCJ	0.451	1.399	1.1548	13		0.427	0.753	0.886	6.33		0.256	0.771	0.745	6		0.026	0.536	0.526	4.33		0.026	0.536	0.526	4.33
HAR-X	0.603	1.004	1.134	6.33		0.461	0.725	0.837	1.67		0.195	0.832	0.711	6.67		0.012	0.550	0.542	6.67		0.012	0.550	0.542	6.67
HAR-Combo	0.592	0.940	0.9165	2.33		0.440	0.684	0.817	1.33		0.251	0.659	0.701	1.33		0.017	0.541	0.524	3		0.017	0.541	0.524	3
RW	0.60	1	1	5.67		0.347	1	1	9.67		0.149	1	1	12.67		0.007	1	1	12.67		0.007	1	1	12.67

Table 6: S&P 500 Index – Out-of-sample results for $h=1,5,10,22$ days-ahead forecast. The table shows R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

4.2 iShares China Large-Cap ETF (FXI)

The iShares China Large-Cap ETF, also known as FXI, seeks to track the investment results of an index composed of large-capitalization Chinese equities that trade on the Hong Kong Stock Exchange. FXI is characterized by an exposure to 50 of the largest companies in China. The dataset contains five-minute observation, considering the trading hours 09:30:00 – 15:55:00, from March 11, 2011 to October 30, 2020 for a total of 2400 days and 187200 observations. The five-minutes log-returns, that are the basis to get the realized variance, were calculated as $r_{t,h} = \ln(P_t/P_{t,h})$, where P_t is the price of the index at time t and h is the five-minute interval. The HAR-X model, in this case, use the CBOE China ETF Volatility Index (VXFXI) that reflects the implied volatility for options of the FXI ETF.

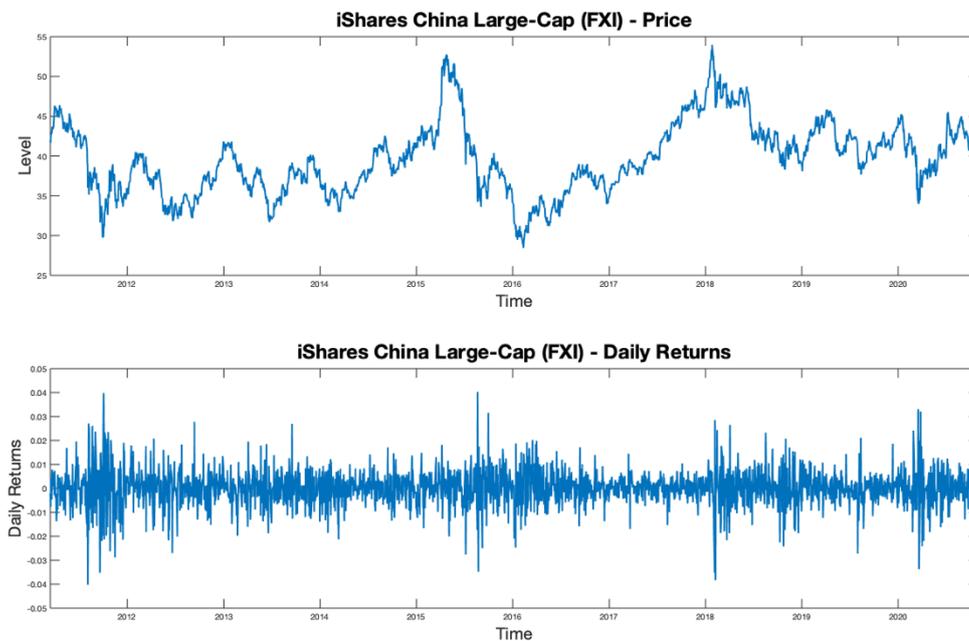


Figure 10: iShares China Large Cap ETF - Daily price level and daily returns

As previously mentioned, Figure 11 shows the RV_t and its decomposition in continuous, C_t , and jumps components, J_t . The highest peak in variance occurred during the China Black Monday in August 2015. The strong uncertainty resulting from this shock is strongly reflected by a bearish market. Moreover, this period is characterized by a highly persistent variance, higher than the one that follows the Covid-19 pandemic crisis. As we can see in the last part of the dataset, the realized variance returned to almost normal pre-crisis level. This countertrend is mainly due to China's ability to stem the pandemic spread. However,

as regards the jumps process, the pike has been reached during the Covid-19 crisis period and jumpy days are 1233 representing 51.37% of the total numbers of days.

Table 7 shows descriptive statistics of r_t , RV_t , C_t and J_t for the entire sample and for the

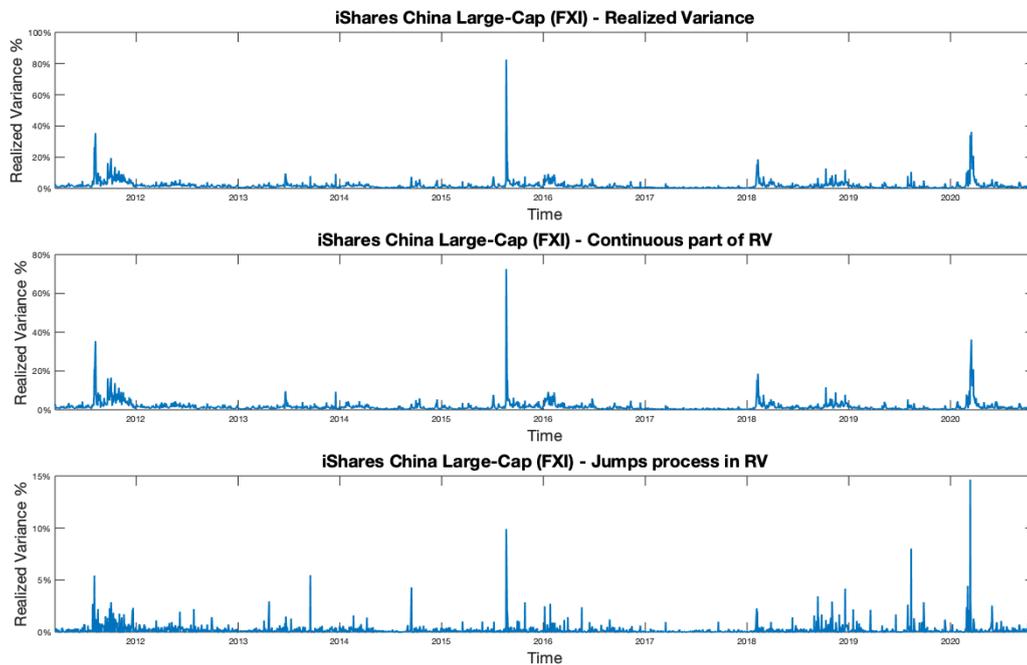


Figure 11: iShares China Large Cap ETF - Realized Variance, Continuous component and Jumps process

Covid-19 period. From the data we can see that the realized variance in the pandemic crisis period, is slightly above the mean but far below the U.S one (2.77% versus 7.51%).

Full period 16/03/2011 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	0.0003	0.018	0.016	0.001
Std. dev.	0.007	0.029	0.026	0.005
Skewness	-0.116	11.974	11.610	12.423
Kurtosis	6.282	259.117	241.448	245.639
Min	-0.040	0.0009	0.0009	0
Max	0.040	0.825	0.726	0.146
Covid-19 period 24/02/2020 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	0.0003	0.027	0.024	0.003
Std. dev.	0.008	0.050	0.045	0.012
Skewness	0.071	4.349	4.462	9.895
Kurtosis	5.817	24.069	26.104	114.101
Min	-0.033	0.002	0.002	0
Max	0.033	0.362	0.362	0.146

Table 7: iShares China Large-Cap ETF –Descriptive statistics of r_t , RV_t , C_t and J_t .

4.2.1 In-sample

A first and interesting step is to assess the dynamics of the parameters of the HAR-RV model. From Figure 12 we can easily see the shock in the parameters occurred after the Chinese Black Monday. The daily coefficient declined from roughly 0.6 to 0.1, while the weekly coefficient increased in magnitude. It is noteworthy how the daily coefficient has acquired new importance after the Covid-19 crisis in February 2020. After taking this into consideration it is necessary to evaluate the fit for each model. Tables 8.a and 8.b contain results of the estimations performed throughout the entire period. The table reports the R^2 , MSE, QLike and as usual, the average score of the three previous mentioned measures. Looking at the coefficients of the HAR-RV and C-HAR model we can easily notice the significance of the daily, weekly and monthly parameters with a slightly higher incidence of the weekly parameter on the daily one in the first model. The results from the S-HAR model where the parameters are in line with the theory proposed by Patton and Sheppard (2015) are very interesting. Respectively the parameter of the negative and positive semivariance are positive and negative. The intuition is that there is a positive relation between the total realized variance and the one caused by negative returns.

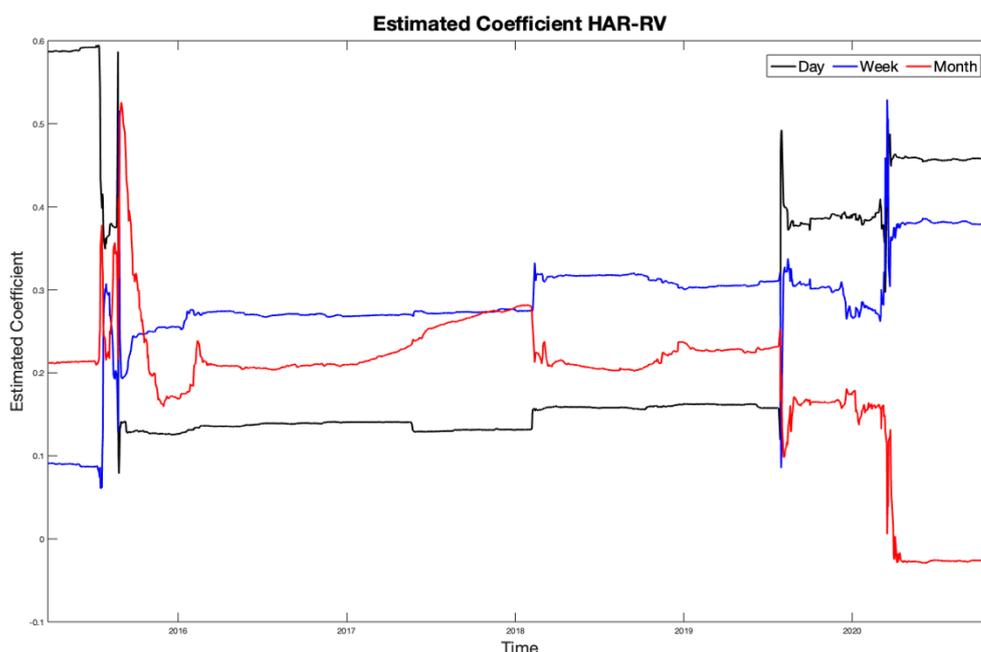


Figure 12 iShares China Large-Cap ETF – HAR-RV coefficients based on a rolling window of 1000 days.

Continuing the analysis we can see that the models that include jumps processes do not get good in-sample results and between the different horizons, the weekly jumps are always significant. Moving on to the results of the Q models (AR-Q, HAR-Q-D and HAR-Q-F) we

can notice that the parameters are consistent with the results of Bollerslev *et al.* (2015). The daily adjustment, $\beta_Q^{(d)}$, is negative and significant for all three models. As expected, the implied volatility parameter is positive and significant in HAR-X. The model with the best in-sample fit is the HAR-Q-F followed by the HAR-Combo while RW is the least accurate.

	HAR-RV	AR-Q	HAR-RV- Q-D	HAR-RV- Q-F	HAR-RV- J-D	HAR-RV- J-F
c	1.46x10 ⁻⁵ (5.280)	1.22x10 ⁻⁵ (5.120)	8.30x10 ⁻⁶ (3.072)	3.38x10 ⁻⁶ (1.000)	1.49x10 ⁻⁵ (5.363)	1.31x10 ⁻⁵ (4.580)
$\beta_{RV}^{(d)}$	0.310 (12.79)		0.711 (19.002)	0.666 (16.513)	0.330 (11.203)	0.341 (11.522)
$\beta_{RV}^{(w)}$	0.353 (9.083)		0.154 (3.831)	0.280 (4.074)	0.345 (8.701)	0.247 (5.047)
$\beta_{RV}^{(m)}$	0.125 (3.105)		0.046 (1.194)	0.061 (0.876)	0.124 (3.070)	0.139 (2.222)
$\beta_C^{(d)}$						
$\beta_C^{(w)}$						
$\beta_C^{(m)}$						
$\beta_J^{(d)}$					-0.133 (-1.201)	-0.310 (-2.548)
$\beta_J^{(w)}$						0.903 (3.066)
$\beta_J^{(m)}$						0.211 (0.353)
$\gamma^{(d)}$						
$\gamma^{(w)}$						
$\gamma^{(m)}$						
$\beta_{RV+}^{(d)}$						
$\beta_{RV-}^{(d)}$						
$\beta_Q^{(d)}$		-1.558 (-19.028)	-1.303 (-13.707)	-1.164 (-10.966)		
$\beta_Q^{(w)}$				-0.808 (-2.039)		
$\beta_Q^{(m)}$				-1.272 (-1.362)		
$\beta_X^{(d)}$						
R^2	0.346	0.387	0.394	0.396	0.3468	0.350
MSE	0.714	0.669	0.661	0.658	0.713	0.709
QLike	0.964	0.965	0.957	0.956	0.964	0.966
Avg. Score	9.33	6	3.33	2.33	8.33	9

Table 8.a: iShares China Large-Cap ETF - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

	C-HAR	S-HAR	HAR-RV- CJ	HAR-RV- LCJ	HAR-X	HAR- Combo	RW
c	1.67x10 ⁻⁵ (6.101)	1.22x10 ⁻⁵ (4.586)	1.31x10 ⁻⁵ (4.580)	-1.69x10 ⁻⁶ (4.580)	-6.42x10 ⁻⁶ (-1.734)		
$\beta_{RV}^{(d)}$					0.293 (12.244)		
$\beta_{RV}^{(w)}$		0.263 (6.944)			0.210 (4.994)		
$\beta_{RV}^{(m)}$		0.102 (2.646)			-0.085 (-1.814)		
$\beta_C^{(d)}$	0.371 (13.550)		0.341 (11.522)	0.296 (10.193)			
$\beta_C^{(w)}$	0.346 (7.989)		0.247 (5.047)	0.213 (4.273)			
$\beta_C^{(m)}$	0.129 (2.920)		0.139 (2.222)	0.026 (0.359)			
$\beta_J^{(d)}$			0.030 (0.284)	0.016 (0.157)			
$\beta_J^{(w)}$			1.150 (4.263)	1.111 (4.209)			
$\beta_J^{(m)}$			0.350 (0.636)	0.774 (1.428)			
$\gamma^{(d)}$				-0.004 (-8.798)			
$\gamma^{(w)}$				-0.002 (-2.094)			
$\gamma^{(m)}$				-0.002 (-1.114)			
$\beta_{RV+}^{(d)}$		-0.389 (-7.269)					
$\beta_{RV-}^{(d)}$		1.326 (17.969)					
$\beta_Q^{(d)}$							
$\beta_Q^{(w)}$							
$\beta_Q^{(m)}$							
$\beta_X^{(d)}$					0.170 (8.410)		
R^2	0.341	0.399	0.350	0.387	0.365	0.392	0.2941
MSE	0.719	0.655	0.709	0.669	0.693	0.664	1
QLike	0.964	0.970	0.966	0.977	0.956	0.956	1
Avg. Score	10.33	4.33	8.66	8	5.33	3	13

Table 8.b: iShares China Large-Cap ETF - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

4.2.2 Out-of-sample

As seen previously, the forecasts are performed on $h=1, 5, 10, 22$ days-ahead corresponding to one day, one week, two weeks and one month out-of-sample analysis. In Table 11 reports the results of the individual forecast analysis based on R^2 , MSE and Qlike. Furthermore it reports, average results of the model based on the three previously motioned measures. The results obtained from a simultaneous comparison conducted by the Model Confidence Set are shown in Table 9. Starting from $h=1$ we can see that the model that allows for a parameter correction achieved better results. HAR-Q-D and AR-Q are the models with the best fit.

MCS - MSE		$h=5$		$h=10$		$h=22$	
$h=1$							
Model	P-value	Model	P-value	Model	P-value	Model	P-value
HAR-RV-J-F	0.0130	RW	0.0270	RW	0.0210	RW	0.0521
HAR-RV-LCJ	0.0130	HAR-RV-CJ	0.1091	AR-Q	0.1181	HAR-RV-CJ	0.3694
HAR-RV-CJ	0.0180	AR-Q	0.1261	HAR-RV-LCJ	0.1491	HAR-Q-F	0.4695
RW	0.0250	HAR-RV-J-F	0.1542	HAR-Q-F	0.1542	HAR-RV-J-F	0.4945
HAR-RV-J-D	0.0350	HAR-RV	0.1542	HAR-RV-CJ	0.1632	HAR-RV-LCJ	0.5786
S-HAR	0.0521	HAR-RV-J-D	0.2022	HAR-RV	0.1932	AR-Q	0.6406
HAR-RV	0.0551	HAR-Q-F	0.3043	HAR-RV-J-F	0.2062	HAR-RV-J-D	0.7347
C-HAR	0.2302	C-HAR	0.3153	HAR-RV-J-D	0.2062	S-HAR	0.7347
HAR-Q-F	0.8088	HAR-RV-LCJ	0.4024	C-HAR	0.2533	HAR-RV	0.7347
AR-Q	0.8719	S-HAR	0.4795	HAR-Q-D	0.2533	HAR-Q-D	0.7818
HAR-Combo	0.9429	HAR-Combo	0.4795	S-HAR	0.2533	C-HAR	0.7958
HAR-X	0.9429	HAR-Q-D	0.4795	HAR-Combo	0.2843	HAR-Combo	0.8058
HAR-Q-D	1	HAR-X	1	HAR-X	1	HAR-X	1

MCS - QLike		$h=5$		$h=10$		$h=22$	
$h=1$							
Model	P-value	Model	P-value	Model	P-value	Model	P-value
HAR-RV-J-F	0.0130	RW	0.0270	RW	0.0210	RW	0.0521
HAR-RV-LCJ	0.0130	HAR-RV-CJ	0.1091	AR-Q	0.1181	HAR-RV-CJ	0.3694
HAR-RV-CJ	0.0180	AR-Q	0.1261	HAR-RV-LCJ	0.1491	HAR-Q-F	0.4695
RW	0.0250	HAR-RV-J-F	0.1542	HAR-Q-F	0.1542	HAR-RV-J-F	0.4945
HAR-RV-J-D	0.0350	HAR-RV	0.1542	HAR-RV-CJ	0.1632	HAR-RV-LCJ	0.5786
S-HAR	0.0521	HAR-RV-J-D	0.2022	HAR-RV	0.1932	AR-Q	0.6406
HAR-RV	0.0551	HAR-Q-F	0.3043	HAR-RV-J-F	0.2062	HAR-RV-J-D	0.7347
C-HAR	0.2302	C-HAR	0.3153	HAR-RV-J-D	0.2062	HAR-Q-D	0.7347
HAR-Q-F	0.8088	HAR-RV-LCJ	0.4024	C-HAR	0.2533	HAR-RV	0.7347
AR-Q	0.8719	S-HAR	0.4795	HAR-Q-D	0.2533	S-HAR	0.7818
HAR-Combo	0.9429	HAR-Combo	0.4795	S-HAR	0.2533	C-HAR	0.7958
HAR-VIX	0.9429	HAR-Q-D	0.4795	HAR-Combo	0.2843	HAR-Combo	0.8058
HAR-Q-D	1	HAR-X	1	HAR-X	1	HAR-X	1

Table 9: iShares China Large Cap ETF– Model Confidence Set based on MSE and Q-Like loss function computed with 999 bootstrap replications, average block length of 5 and $\alpha = 0.05$. The included models are all of these with $p\text{-value} > \alpha$

In contrast, the models that include the jumps components, do not get good performance in the one day-ahead forecast. Excluding the models from the Q family, the best results is achieved by the HAR-X. The previous outcomes are also confirmed by the MCS in Table 9. When continuing the analysis for longer forecast time horizons, is evident that models with jumps acquire predictive power. However, also in the case of $h=5, 10$, the models with the

best out-of-sample performance are the HAR-X, HAR-Q-D and HAR-Combo. These results are also confirmed for the longer forecast period. In the one month-ahead forecast, the best model according to the individual and simultaneous comparison, is the HAR-X followed by the HAR-Combo. This last model, in the S&P 500 Index, was the preferred in almost all of the forecasted times considered, while for the FXI ETF it has never been the first choice. Figure 13 shows the comparison between the best forecast models and the realized value on the different time horizons considered. Where $h=1$ is represented by HAR-Q-D and $h=5, 10, 22$ represented by the HAR-X. Table 10 reports the p-values of the Berkowitz test relative to the out-of-sample forecast of different models. It is possible to see that the null hypothesis is not rejected, at 5% significance level, for the RW ($h=1,5$) and AR-Q ($h=3$).

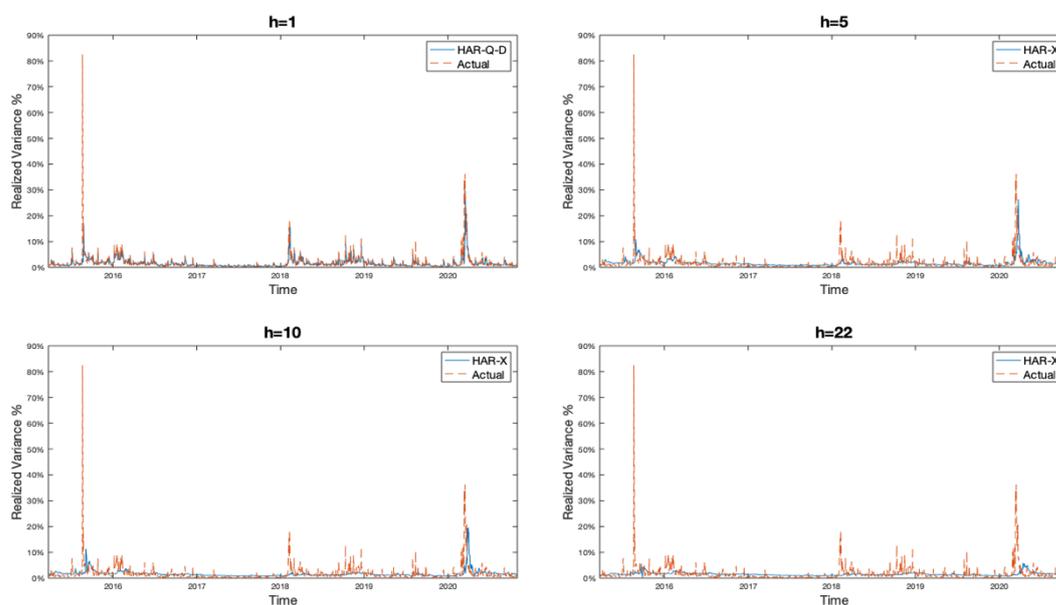


Figure 13: iShares China Large Cap ETF – Out-of-sample forecasts . $h=1$: HAR-Q-D, $h=5,10,22$ HAR-X

	$h=1$	$h=5$	$h=10$	$h=22$
HAR-RV	0.0000	0.0000	0.0000	0.0000
AR-Q	0.0000	0.0000	0.3492	0.0000
HAR-Q-D	0.0000	0.0000	0.0000	0.0000
HAR-Q-F	0.0000	0.0000	0.0000	0.0008
HAR-RV-J-D	0.0000	0.0000	0.0000	0.0000
HAR-RV-J-F	0.0000	0.0000	0.0000	0.0000
C-HAR	0.0000	0.0000	0.0000	0.0000
S-HAR	0.0000	0.0000	0.0000	0.0000
HAR-RV-CJ	0.0000	0.0000	0.0000	0.0000
HAR-RV-LCJ	0.0000	0.0000	0.0000	0.0000
HAR-X	0.0000	0.0000	0.0000	0.0000
RW	0.4051	0.0802	0.0000	0.0000

Table 10: iShares China Large Cap ETF – P-value of the Berkowitz test on the entire standard Gaussian distribution, relative to the out-of-sample analysis.

	$h=1$			$h=5$			$h=10$			$h=22$						
	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.
HAR-RV	0.219	0.772	0.981	8.33	0.089	0.629	0.917	8.67	0.029	0.590	0.839	6.33	0.003	0.527	0.739	6
AR-Q	0.282	0.670	0.970	2	0.083	0.622	0.929	7	0.004	0.799	0.853	11.33	0.002	0.526	0.743	7.33
HAR-Q-D	0.275	0.681	0.966	1.66	0.093	0.622	0.910	4.33	0.033	0.583	0.838	3	0.003	0.526	0.739	4
HAR-Q-F	0.223	0.738	0.975	5.66	0.077	0.654	0.927	12	0.017	0.604	0.869	11.67	0.0009	0.762	0.752	11.67
HAR-RV-J-D	0.217	0.775	0.988	9.67	0.089	0.628	0.917	8	0.029	0.590	0.839	6	0.003	0.527	0.739	7.33
HAR-RV-J-F	0.215	0.780	1.004	12	0.100	0.618	0.933	3.67	0.033	0.586	0.845	4	0.003	0.528	0.750	7.33
C-HAR	0.220	0.772	0.978	7.33	0.087	0.628	0.915	8.33	0.025	0.596	0.837	7	0.003	0.529	0.529	7.33
S-HAR	0.272	0.700	0.983	4.67	0.091	0.627	0.910	6	0.031	0.593	0.837	5.33	0.003	0.527	0.739	6.67
HAR-RV-CJ	0.213	0.787	1.003	12.33	0.096	0.623	0.939	7.33	0.028	0.596	0.854	9	0.002	0.533	0.755	11
HAR-RV-LCJ	0.253	0.729	1.000	6.33	0.097	0.628	0.918	6.67	0.028	0.600	0.855	10.33	0.007	0.534	0.752	7.33
HAR-X	0.247	0.739	0.969	4.33	0.105	0.626	0.897	2.67	0.041	0.588	0.820	1.67	0.006	0.525	0.734	1.33
HAR-Combo	0.239	0.742	0.969	5.33	0.099	0.620	0.9082	2.33	0.032	0.588	0.832	3.33	0.005	0.526	0.736	2.33
RW	0.215	1	1	11.33	0.069	1	1	13	0.023	1	1	12.33	0.002	1	1	12.33

Table 11: iShares China Large Cap ETF – Out-of-sample results for $h=1,5,10,22$ days-ahead forecast. The table shows R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

4.3 Apple Inc (AAPL)

Founded in 1976 by Steve Jobs and Steve Wozniak, Apple Inc is an American multinational corporation specialized in computer, software, consumer electronics and online services. With more than 506 stores and more than 147000 employees. It is the 3th largest company in the world and also the one with the greatest weight within the S&P 500 Index at the end of December 2020. After the 2018 correction, the Apple Inc stock price increased steadily until the Covid-19 crisis. However, a strong rebound allowed the title to reach historic highs at the beginning of September 2020. The dataset contains five-minute observation, considering the trading hours 09:30:00 – 15:55:00, from February 24, 2011 to October 30, 2020 for a total of 2412 days and 188136 observations. The HAR-X model, in this case, uses the CBOE Apple VIX Volatility Index (VXAPL) that estimates the expected 30-day volatility of Apple Inc stock returns.

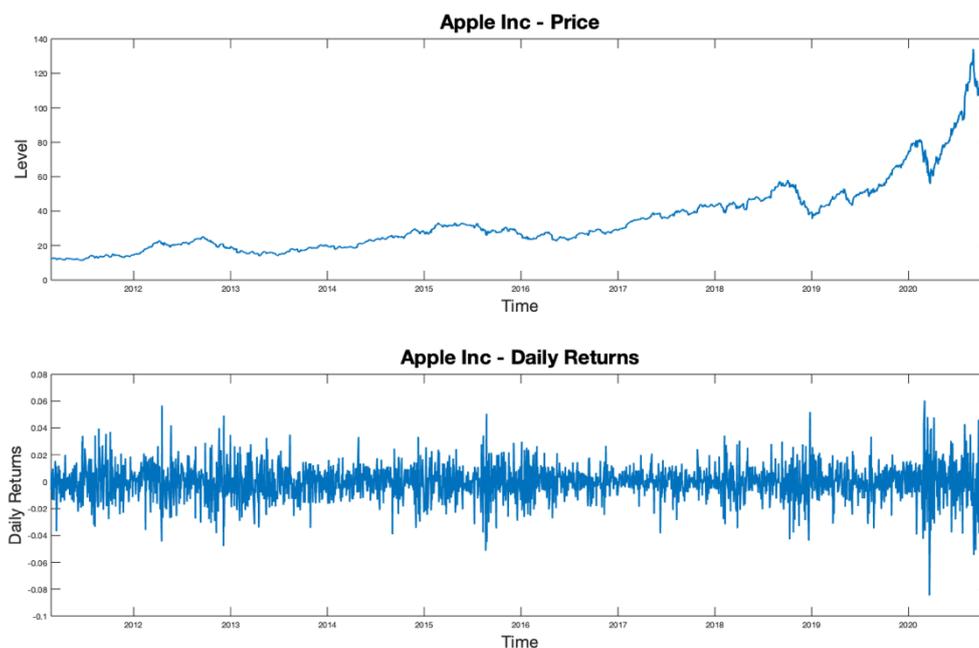


Figure 14: Apple Inc - Daily price level and daily returns

As usual the five-minutes log-returns (that are the basis to get the realized variance) were calculated as $r_{t,h} = \ln(P_t/P_{t,h})$, where P_t is the price of the index at time t and h is the five-minute interval. Figure 15 depicts the realized variance and its decomposition in continuous and jumps components. From the trend of realized variance we can extract two key moments: China Black Monday, August 2015 and Covid-19 crisis, February 2020. In

particular, the highest level of volatility was achieved during the pandemic crisis, and the following period is characterized by a level of volatility above average with another significant peak in September 2020.

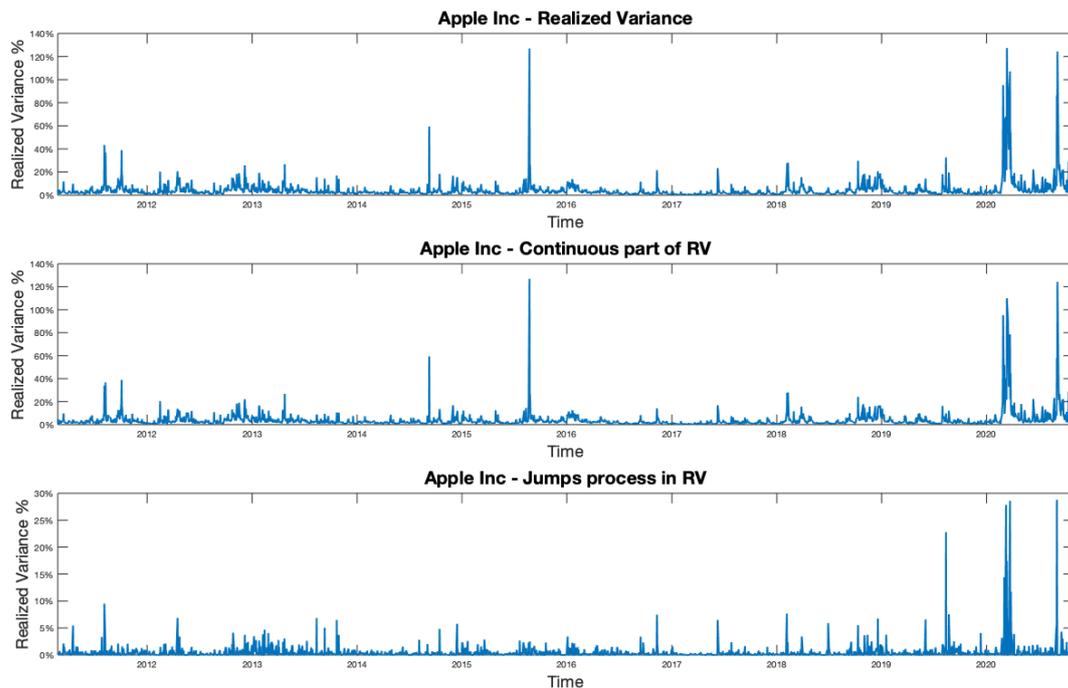


Figure 15: Apple Inc- Realized Variance, Continuous component and Jumps process

Looking at Table 12 we can see that the mean value of the realized variance recorded during the Covid-19 crisis, is much higher than that of the whole period, 14.69% vs 4.40%. In addition, with a more general outlook to the other financial instruments considered in this thesis, the mean realized variance of Apple Inc is the highest evidence that the title is impacted not only by the systematic risk, but also by the idiosyncratic one. It should be noted that the jumpy days are about 47% of the entire period, slightly less compared to the S&P 500 and iShares China Large-Cap ETF.

Full period 24/02/2011 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	-0.0003	0.044	0.039	0.004
Std. dev.	0.012	0.080	0.073	0.014
Skewness	-0.127	9.063	9.525	12.396
Kurtosis	5.542	112.158	125.915	207.611
Min	-0.084	0.002	0.0021	0
Max	0.060	1.274	1.269	0.287

Covid-19 period 14/02/2020 – 30/10/2020				
	r_t	RV_t	C_t	J_t
Mean	4.46×10^{-5}	0.146	0.133	0.013
Std. dev.	0.020	0.220	0.198	0.042
Skewness	-0.276	3.212	3.374	5.140
Kurtosis	4.226	13.731	15.405	31.181
Min	-0.084	0.011	0.006	0
Max	0.060	1.274	1.243	0.287

Table 12: Apple Inc –Descriptive statistics of r_t , RV_t , C_t and J_t .

4.3.1 In-sample

Looking at Figure 16 we can see that until the Covid-19 crisis the coefficient with the biggest impact on realized variance are the daily and the monthly parameters. After the pandemic shock, the weekly coefficient increased in importance. This phenomenon recurs in each financial instrument seen previously. Tables 13.a and 13.b show the results of the full sample estimation of the different models, also taking into consideration R^2 , MSE, Qlike and the average score. The daily, weekly parameters of the HAR-RV model are significant while the monthly is not. Interestingly, if we only consider the continuous component of realized variance, in C-HAR model, the monthly parameter becomes significant. However, the full sample fit of these two models is very similar.

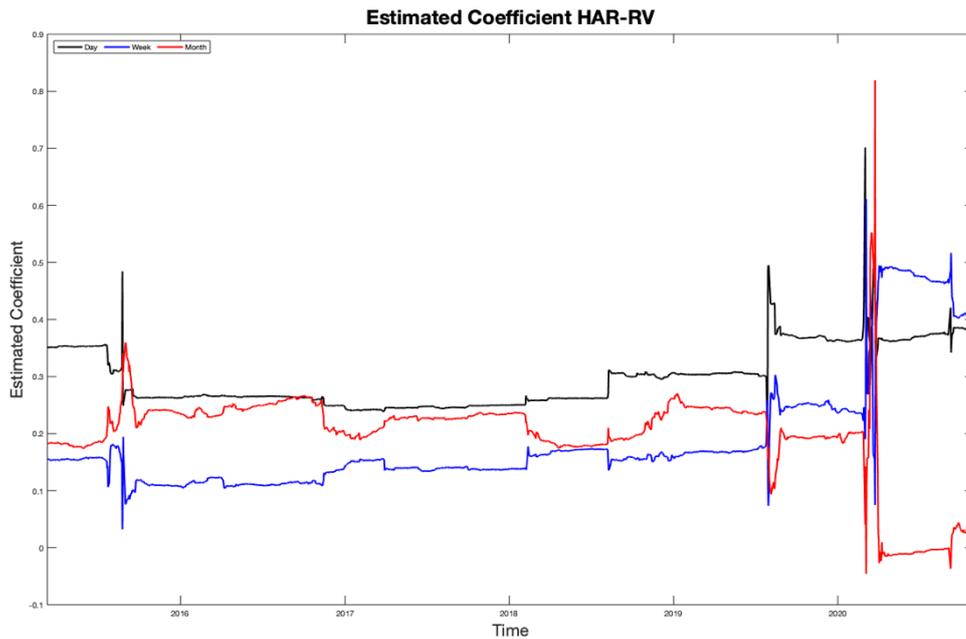


Figure 16: Apple Inc – HAR-RV coefficients based on a rolling window of 1000 days.

The parameters of the S-HAR confirm that the variance, deriving from negative returns, has a positive and direct impact on the total realized variance, confirming the result of Patton and Sheppard (2015). The results of the AR-Q, HAR-Q-D and HAR-Q-F confirm the negativity and significance of the $\beta_Q^{(d)}$. Taking into account the time-varying nature of the measurement error, allows us to obtain better results than those of the parsimonious HAR-RV model. In regards to the models that include the jumps components we can see that the daily jumps are not significant. However, the weekly and monthly jumps are significant for all the models we have considered. In addition the weekly jumps contribute positively to the daily realized variance while the monthly jumps contribute negatively. A good fitting is achieved by the HAR-RV-LCJ, the model that includes the continuous components of realized variance, jumps and leverage effect. A very strong fit is also obtained by the model that includes the implied volatility. From the results is evident that the $\beta_x^{(d)}$ is strongly significative and positive. As for the S&P 500 Index, the model with the best in the sample fit is the one that collects all the others, the HAR-Combo.

	HAR-RV	AR-Q	HAR-RV- Q-Day	HAR-RV- Q-F	HAR-RV- J-D	HAR-RV- J-F
c	3.41×10^{-5} (5.218)	3.21×10^{-5} (4.742)	1.93×10^{-5} (2.740)	3.57×10^{-6} (0.351)	3.40×10^{-5} (5.214)	3.50×10^{-5} (5.384)
$\beta_{RV}^{(d)}$	0.360 (14.919)	0.861 (27.458)	0.544 (12.864)	0.540 (11.878)	0.364 (13.356)	0.364 (13.401)
$\beta_{RV}^{(w)}$	0.360 (9.537)		0.314 (8.146)	0.235 (3.319)	0.358 (9.383)	0.256 (5.899)
$\beta_{RV}^{(m)}$	0.07 (2.182)		0.057 (1.568)	0.25304 (3.072)	0.080 (2.199)	0.220 (3.080)
$\beta_C^{(d)}$						
$\beta_C^{(w)}$						
$\beta_C^{(m)}$						
$\beta_J^{(d)}$					-0.030 (-0.287)	-0.232 (-2.017)
$\beta_J^{(w)}$						1.357 (4.830)
$\beta_J^{(m)}$						-1.591 (-2.60)
$\gamma^{(d)}$						
$\gamma^{(w)}$						
$\gamma^{(m)}$						
$\beta_{RV+}^{(d)}$						
$\beta_{RV-}^{(d)}$						
$\beta_Q^{(d)}$		-0.665 (-9.651)	-0.382 (-5.279)	-0.381 (-4.745)		
$\beta_Q^{(w)}$				0.236 (1.376)		
$\beta_Q^{(m)}$				-0.8 (-2.644)		
$\beta_X^{(d)}$						
R^2	0.407	0.385	0.414	0.416	0.407	0.413
MSE	0.743	0.770	0.734	0.732	0.743	0.735
QLike	0.968	0.970	0.965	0.966	0.967	0.971
Avg. Score	8.66	10.66	4.67	4.33	7.67	8.33

Table 13.a: Apple Inc - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model In bold the models with the best result.

	C-HAR	S-HAR	HAR-RV- CJ	HAR-RV- LCJ	HAR-X	HAR- Combo	RW
c	3.42×10^{-5} (5.215)	3.15×10^{-5} (4.890)	3.50×10^{-5} (5.384)	1.12×10^{-5} (0.979)	-5.52×10^{-5} (-6.446)		
$\beta_{RV}^{(d)}$					0.264 (11.047)		
$\beta_{RV}^{(w)}$		0.365 (9.790)			0.264 (4.814)		
$\beta_{RV}^{(m)}$		0.076 (2.120)			-0.246 (-6.001)		
$\beta_C^{(d)}$	0.410 (15.603)		0.364 (13.401)	0.312 (11.775)			
$\beta_C^{(w)}$	0.336 (8.240)		0.256 (5.899)	0.283 (6.483)			
$\beta_C^{(m)}$	0.138 (3.430)		0.220 (3.080)	0.190 (2.442)			
$\beta_J^{(d)}$			0.131 (1.239)	-0.038 (-0.374)			
$\beta_J^{(w)}$			1.614 (6.133)	1.802 (7.011)			
$\beta_J^{(m)}$			-1.371 (-2.488)	-1.440 (-2.699)			
$\gamma^{(d)}$				-0.008 (-12.324)			
$\gamma^{(w)}$				0.003 (1.697)			
$\gamma^{(m)}$				-0.001 (-0.425)			
$\beta_{RV+}^{(d)}$		-0.103 (-1.623)					
$\beta_{RV-}^{(d)}$		0.831 (12.919)					
$\beta_Q^{(d)}$							
$\beta_Q^{(w)}$							
$\beta_Q^{(m)}$							
$\beta_X^{(d)}$					0.528 (15.204)		
R^2	0.399	0.422	0.413	0.453	0.458	0.434	0.361
MSE	0.735	0.724	0.735	0.686	0.679	0.710	1
QLike	0.967	0.968	0.971	0.986	0.997	0.962	1
Avg. Score	8.67	5	8	5	4.67	2.33	13

Table 13.b: Apple Inc - In-sample analysis. The table shows the OLS estimation of parameters reported with t-statistics in brackets, R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

4.3.2 Out-of-sample

As previously, the forecasts are performed on $h=1, 5, 10, 22$ days-ahead corresponding to one day, one week, two weeks and one month out of sample analysis. Table 16 reports results of the individual forecast analysis based on R^2 , MSE and QLike while the outcome of the simultaneous comparison performed by the Model Confident Set, both with MSE and QLike loss function are shown in Table 14. Starting from the classic HAR-RV model we can see that especially for one and five days-ahead the model is able to achieve a better forecast than more complex models. In the shortest forecast period this model is one of the finest. The HAR-RV is also generally preferred to the C-HAR model for $h=1,5$ and 10. The C-HAR model achieved a better result in one month-ahead forecast. Moving to the Q-family models the simple AR-Q is preferred to the HAR-Q-D and HAR-Q-F particularly for two weeks and one month ahead forecast. In addition, for $h=10$, the AR-Q is the model with the greatest predictive power. For one days-ahead the HAR-Q-D, HAR-Q-F and AR-Q are not the best choice.

MCS - MSE		$h=5$		$h=10$		$h=22$	
$h=1$							
Model	P-value	Model	P-value	Model	P-value	Model	P-value
RW	0.2763	RW	0.1211	RW	0.2503	RW	0.0230
HAR-RV-CJ	0.3994	HAR-Q-D	0.3153	HAR-RV-J-D	0.4945	HAR-Q-F	0.0360
HAR-RV-J-F	0.4915	HAR-RV-LCJ	0.3153	C-HAR	0.4945	HAR-RV-J-D	0.0450
HAR-Q-D	0.5906	HAR-RV-J-D	0.3153	HAR-RV-CJ	0.4945	S-HAR	0.0561
HAR-RV-J-D	0.5906	HAR-RV-CJ	0.3153	S-HAR	0.5235	C-HAR	0.0891
C-HAR	0.6286	HAR-Q-F	0.3153	HAR-RV-J-F	0.5235	HAR-RV	0.1311
HAR-Q-F	0.7097	C-HAR	0.3153	HAR-RV	0.5516	HAR-Q-D	0.3323
S-HAR	0.7097	HAR-RV-J-F	0.3153	HAR-Q-F	0.5516	HAR-Combo	0.4354
HAR-RV-LCJ	0.7097	S-HAR	0.3153	HAR-Combo	0.5516	HAR-X	0.5155
HAR-RV	0.7097	HAR-RV	0.3153	HAR-X	0.5846	AR-Q	0.5155
AR-Q	0.7097	AR-Q	0.3153	HAR-RV-LCJ	0.7227	HAR-RV-CJ	0.5155
HAR-Combo	0.7097	HAR-Combo	0.3153	HAR-Q-D	0.8428	HAR-RV-J-F	0.5155
HAR-X	1	HAR-X	1	AR-Q	1	HAR-RV-LCJ	1

MCS - QLike		$h=5$		$h=10$		$h=22$	
$h=1$							
Model	P-value	Model	P-value	Model	P-value	Model	P-value
HAR-RV-LCJ	0.0090	RW	0.005	RW	0	RW	0.01
RW	0.0130	AR-Q	0.0741	HAR-Q-D	0.0340	HAR-Q-F	0.0501
S-HAR	0.0130	HAR-RV	0.0741	HAR-Q-F	0.0360	S-HAR	0.0511
HAR-X	0.0160	HAR-RV-CJ	0.0741	S-HAR	0.0360	HAR-RV-J-D	0.0651
HAR-Q-F	0.0160	HAR-RV-J-D	0.0741	C-HAR	0.0360	HAR-RV	0.0761
C-HAR	0.0160	HAR-Q-D	0.0741	HAR-RV	0.0360	C-HAR	0.1321
AR-Q	0.0160	C-HAR	0.0741	HAR-RV-J-D	0.0360	HAR-Q-D	0.1912
HAR-RV-CJ	0.0160	HAR-RV-CJ	0.0741	HAR-RV-LCJ	0.0360	AR-Q	0.3383
HAR-RV	0.0160	S-HAR	0.0741	AR-Q	0.0360	HAR-Combo	0.7187
HAR-Q-F	0.0160	HAR-RV-J-F	0.0741	HAR-RV-CJ	0.0360	HAR-RV-LCJ	0.7187
HAR-Q-D	0.0160	HAR-Q-F	0.0741	HAR-RV-J-F	0.0360	HAR-RV-CJ	0.7187
HAR-RV-J-D	0.0160	HAR-Combo	0.0741	HAR-Combo	0.0360	HAR-RV-J-F	0.7187
HAR-Combo	1	HAR-X	1	HAR-X	1	HAR-X	1

Table 14: Apple Inc– Model Confidence Set based on MSE and Q-Like loss function computed with 999 bootstrap replications, average block length of 5 and $\alpha = 0.05$. The included models are all of these with $p\text{-value} > \alpha$

Looking at the results of the models that include jumps components we note that the prediction power increases with the forecast time horizon. All the models get unsatisfactory results for one day ahead forecast. When comparing the HAR-RV-J-D and HAR-RV-J-F models, we see the one which includes, not only the daily but also the weekly and monthly jumps achieved better results. It is clearly visible that the jumps matter especially in the one month-ahead forecast where the best models are the HAR-RV-J-F, HAR-RV-CJ and HAR-RV-LCJ. Unlike the results of the S&P500 Index, but in line with those seen for the iShares China Large-Cap, the HAR-X model achieved solid outcomes on every forecast period considered, especially for $h=10$.

	$h=1$	$h=5$	$h=10$	$h=22$
HAR-RV	0.0640	0.0000	0.1492	0.0000
AR-Q	0.0000	0.0000	0.0000	0.0000
HAR-Q-D	0.0000	0.0000	0.0000	0.0000
HAR-Q-F	0.0000	0.1076	0.0000	0.0000
HAR-RV-J-D	0.0000	0.0000	0.3645	0.0000
HAR-RV-J-F	0.0903	0.0000	0.3170	0.0000
C-HAR	0.0024	0.0000	0.3697	0.0000
S-HAR	0.0302	0.0000	0.2650	0.0000
HAR-RV-CJ	0.0000	0.0000	0.3245	0.0000
HAR-RV-LCJ	0.0000	0.0169	0.0000	0.0000
HAR-X	0.0047	0.0000	0.7148	0.0000
RW	0.0000	0.0000	0.0000	0.0000

Table 15: Apple Inc – P-value of the Berkowitz test on the entire standard Gaussian distribution, relative to the out-of-sample analysis.

The results obtained from density forecast analysis in Table 15 are noteworthy. The null hypothesis is not rejected at the 5% significance level for the following models:

- $h=1$: HAR-RV, HAR-RV-Q-F
- $h=5$: HAR-Q-F
- $h=10$: HAR-RV, HAR-RV-J-D, HAR-RV-J-D, HAR-RV-J-D, HAR-RV-CJ, HAR-RV-LCJ and HAR-X.

It is important to note that on a 2 weeks forecasting horizon the inclusion of the jumps components allows us to obtain good density forecast results as proof that for long forecast times the jumps contribute to improve the prediction performance. Figure 17 shows the comparison between the best forecast model and the realized value on the different time horizons considered. For different forecast time horizons the models with the finest results are respectively: HAR-Combo, HAR-X, HAR-Combo and HAR-RV-J-F

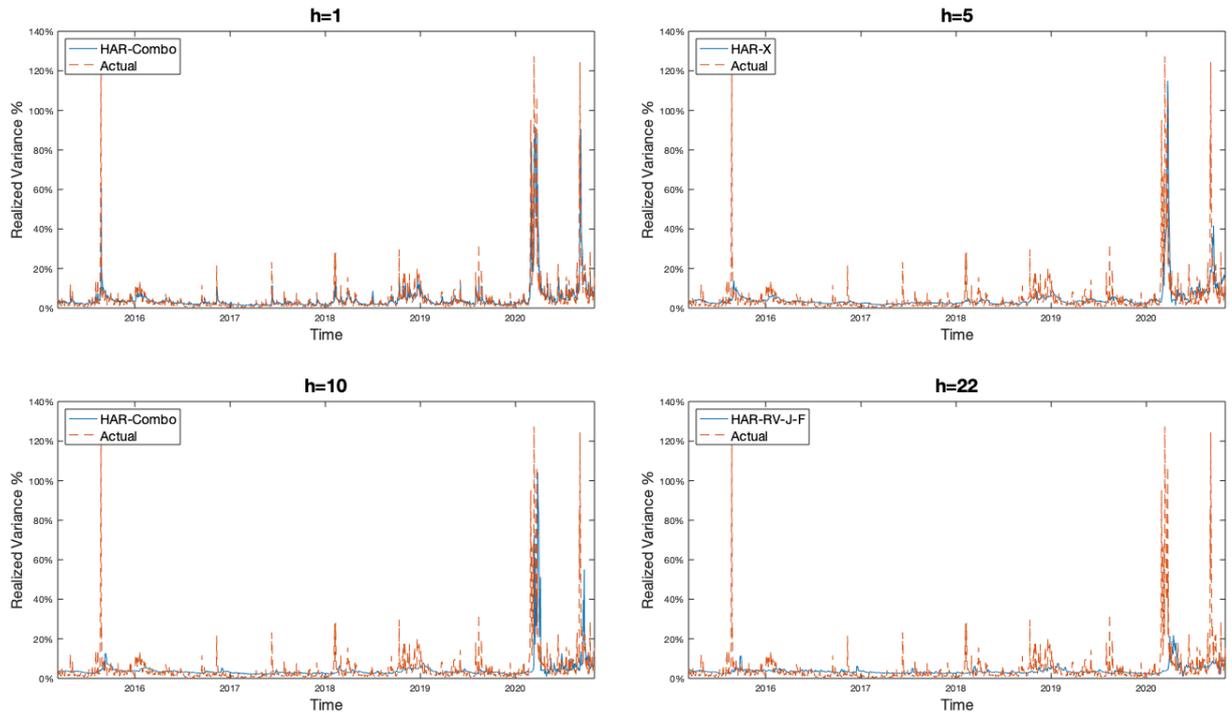


Figure 17: Apple Inc - HAR-Combo Out-of-sample forecasts. $h=1$, HAR-Combo; $h=2$, HAR-X; $h=10$ HAR-Combo; $h=22$, HAR-RV-J-F

	$h=1$				$h=5$				$h=10$				$h=22$							
	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.	R^2	MSE	QLike	Avg.S.
HAR-RV	0.395	0.824	0.983	4	0.217	0.661	0.937	4.67	0.143	0.717	0.905	6.67	0.013	0.540	0.857	9.33	0.013	0.540	0.857	9.33
AR-Q	0.362	0.842	0.986	8.67	0.158	0.682	0.941	11.67	0.166	0.656	0.901	2.67	0.023	0.533	0.855	5.67	0.023	0.533	0.855	5.67
HAR-Q-D	0.360	0.863	0.984	10.33	0.206	0.675	0.938	9.67	0.173	0.659	0.908	4.67	0.018	0.536	0.855	6.67	0.018	0.536	0.855	6.67
HAR-Q-F	0.371	0.862	0.988	10.33	0.252	0.690	0.937	7.33	0.089	0.735	0.926	12	0.005	0.568	0.892	12	0.005	0.568	0.892	12
HAR-RV-J-D	0.388	0.848	0.982	5.33	0.206	0.673	0.937	8	0.129	0.737	0.901	9.67	0.013	0.541	0.857	9.67	0.013	0.541	0.857	9.67
HAR-RV-J-F	0.384	0.864	0.983	7	0.209	0.671	0.934	6	0.133	0.732	0.896	7	0.028	0.530	0.846	2	0.028	0.530	0.846	2
C-HAR	0.387	0.844	0.985	7	0.212	0.669	0.939	7.33	0.136	0.731	0.903	7.67	0.013	0.540	0.856	8	0.013	0.540	0.856	8
S-HAR	0.398	0.826	0.991	6.33	0.212	0.667	0.933	4.33	0.139	0.722	0.906	7.67	0.012	0.541	0.857	11	0.012	0.541	0.857	11
HAR-RV-CJ	0.384	0.861	0.984	8.33	0.208	0.672	0.935	7	0.132	0.733	0.897	8	0.028	0.530	0.846	3	0.028	0.530	0.846	3
HAR-RV-LCJ	0.413	0.827	1.002	6.67	0.215	0.681	0.940	8.67	0.153	0.682	0.904	4.67	0.034	0.527	0.848	2.33	0.034	0.527	0.848	2.33
HAR-X	0.438	0.766	0.990	4	0.284	0.604	0.911	1	0.148	0.723	0.884	4.	0.023	0.535	0.840	3.67	0.023	0.535	0.840	3.67
HAR-Combo	0.408	0.811	0.976	2	0.225	0.655	0.923	2.33	0.145	0.708	0.893	3.67	0.0227	0.533	0.847	4.67	0.0227	0.533	0.847	4.67
RW	0.384	1	1	11	0.144	1	1	13	0.123	1	1	12.67	0.0051	1	1	13	0.0051	1	1	13

Table 16: Apple Inc – Out-of-sample results for $h=1,5,10,22$ days-ahead forecast. The table shows R^2 , MSE, and QLike losses standardized by the loss of the random walk model. In bold the models with the best result.

5 Variance risk premium and returns prediction

When investing in a security, investors face at least two sources of uncertainty, namely the uncertainty of returns, represented by the returns variance, and the uncertainty of the returns variance itself. This gives rise to the so called variance risk premium, introduced by Carr and Wu (2009), which is measured as the difference between the risk neutral and physical expectations of an asset's total return variation. The variance risk premium not only captures aggregate market risk aversion but also has been found to carry predictability for future equity returns (Bollerslev, Tauchen and Zhou, 2008). The VRP is therefore computed as the difference between a squared measure of implied volatility deriving from option prices and the expected realized variance computed with high frequency historical returns as shown in the theoretical Chapter 2. Table 1 shows the financial instruments used in the following analysis with the relative implied volatility measures. Following the notation of Rombouts, Stentoft and Violante (2017), a generic VRP is defined as

$$\Pi_{t,t+\tau}^i = E_t^Q[QV_{t,t+\tau}] - E_t^P[QV_{t,t+\tau}] \quad (5.1)$$

Where $\tau=1$ month, $QV_{t,t+\tau}$ is the (latent) quadratic variation of the underlying price process and the conditional expectations are under the risk neutral (Q) and physical (P) measures, respectively. According to Carr and Wu (2009) the VRP is the expected profit to the long position of a variance swap contract and is expected to be positive. In practice, the risk neutral measure, $E_t^Q[QV_{t,t+\tau}]$, is directly computed from the square of option market implied volatility index, while the physical measure, $E_t^P[QV_{t,t+\tau}]$, is the forecast of realized variance over the next month using five minutes returns computed with the models that achieved the best out-of-sample performance in the previous Chapter. In particular the models used in the computation of the VRP are shown in Table 16. Figure 18 depicts the variance risk premium for the different financial instruments considered. As expected VRP is generally positive for all the three financial instruments. However, it becomes slightly negative for S&P 500 Index and Apple Inc during the crisis period, particularly during the Covid-19 crisis. Regarding the difference in VRP which results from the use of different models, it is possible to note that during less volatile period the spread is limited, while during more volatile periods the difference is more marked. This is notable in the variance risk premium deriving from the S&P500 Index during the pandemic crisis where the second negative spike in VRP is not present if we consider the HAR-Combo. For a better understanding of the data, Table 17 shows the descriptive statistics of the different VRP's.

	Model
S&P 500 Index	HAR-Combo
	HAR-RV-J-F
iShares China Large-Cap ETF	HAR-Combo
	HAR-X
Apple Inc	HAR-RV-LCJ
	HAR-RV-J-F

Table 16: One month-ahead forecast winning models

	S&P 500 Index		iShares China Large Cap ETF		Apple Inc	
	$\Pi_{t,t+1}^{HAR-Combo}$	$\Pi_{t,t+1}^{HAR-RV-J-F}$	$\Pi_{t,t+1}^{HAR-X}$	$\Pi_{t,t+1}^{HAR-Combo}$	$\Pi_{t,t+1}^{HAR-RV-J-F}$	$\Pi_{t,t+1}^{HAR-RV-LCJ}$
Mean	2.1669	2.1899	5.374	5.348	5.654	5.722
Std. dev.	4.415	4.387	4.317	4.379	7.616	7.600
Skewness	8.076	8.026	3.544	3.448	5.278	5.313
Kurtosis	88.001	87.566	22.813	21.808	45.799	46.282
Min	-6.489	-7.402	1.104	0.990	-2.885	-3.226
Max	67.115	66.836	46.638	46.630	99.532	99.703

Table 17: Descriptive statistics for different VRP computed with the winning models

The VRP can also be estimated with a random walk model for realized variance, however, from the results in the previous paragraph we can assess that the predictions of random walk are always inferior to the other models, particularly in one month-ahead forecast. Figure 18 depicts the VRP's based on random walk model. These time series present high negative values mainly due to the asynchronicity in time between the risk neutral and physical measures.

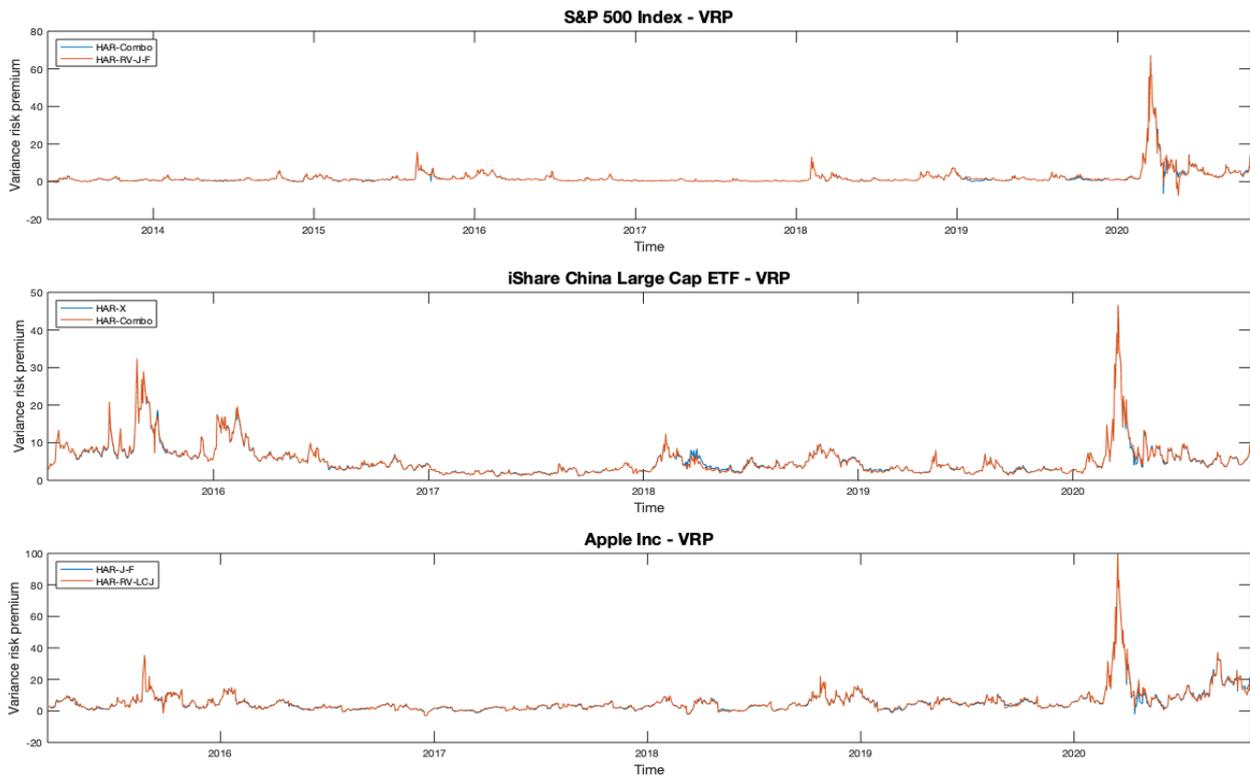


Figure 18: Variance Risk Premium of the winning models

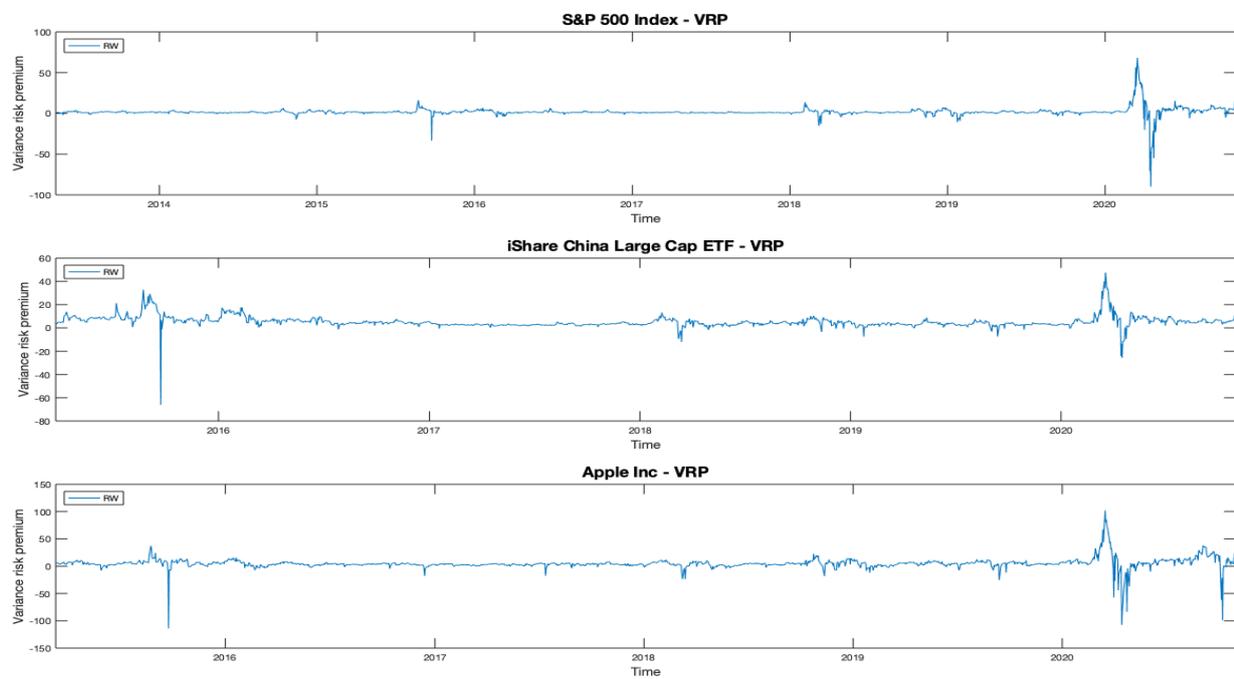


Figure 19: Variance Risk Premium of the random walk model

5.1 Returns predictability – full period

Motivated by the fact that the variance risk premium harbors information about aggregate risk aversion, Bollerslev *et al.* (2009) demonstrated that it has predictive power for future monthly returns. The predictive regression is defined as

$$\frac{1}{h} \sum_{j=1}^h r_{t+j} = \alpha + \beta \Pi_{t,t+1}^i + \varepsilon_{t+h,t} \quad (5.2)$$

where h is the forecast horizon, r_t is the monthly return for month t and $\Pi_{t,t+1}^i$ is one of the VRP's estimated with the winning model exhibit in Table 16. Predictability is measured by the adjusted R^2 , and the possible serial correlation generated by the overlapping of the averaging returns is corrected by the Newey-West Standard Errors. To compare the returns predictably obtained with the winning model we also include the results obtained with the random walk, $\Pi_{t,t+1}^i$, computed by

$$\Pi_{t,t+1}^{RW} = E_t^Q [QV_{t,t+\tau}] - RV_{t-1,t} \quad (5.3)$$

Table 18 show the results of the above regression respectively for S&P 500 Index, iShares China-Large-Cap ETF and Apple Inc stock. Starting from the first asset we find that the peak in R^2 is reached in $h=1$ for all the models considered. Particularly for the one month-ahead forecast the results are in favor of the $\Pi_{t,t+1}^{RW}$ while, for longer prediction time horizons, the results of the two more sophisticated model are superior and quite similar. For a better understand of the data, Figure 20 depicts the pattern in the R^2 for all the models. Moving to the Chinese ETF we can observe puzzling results. The R^2 are mixed until $h=5$, after that the clear superior performance of $\Pi_{t,t+1}^{HAR-X}$ and $\Pi_{t,t+1}^{HAR-Combo}$ is evident. The peak of 3.51% is reached by the $\Pi_{t,t+1}^{HAR-X}$ for the 12 month ahead forecasting. For Apple Inc the results are consistent with those of Bollerslev *et al.* (2009). At the 6th months ahead the pattern in R^2 follows an inverse U-shaped curve, with a peak of 6.68% shown by the $\Pi_{t,t+1}^{HAR-RV-LCJ}$. In this case, the difference in using the VRP computed with the winning model compared to the random walk is particularly evident.

S&P 500 Index

		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^i$									
$\Pi_{t,t+1}^{HAR-Combo}$	α	0.546 (7.666)	0.421 (8.681)	0.334 (9.110)	0.283 (9.690)	0.277 (11.108)	0.292 (13.675)	0.299 (18.532)	0.270 (18.533)
	β	-0.167 (-7.514)	-0.099 (-8.090)	-0.049 (-6.114)	-0.022 (-4.425)	-0.012 (-3.057)	-0.017 (-4.981)	-0.027 (-7.060)	-0.016 (-5.847)
	R^2	6.262	4.908	2.097	0.673	0.291	0.745	3.286	1.640
$\Pi_{t,t+1}^{HAR-RV-J-F}$	α	0.531 (7.619)	0.430 (8.742)	0.339 (9.194)	0.287 (9.840)	0.283 (11.288)	0.297 (13.840)	0.301 (18.537)	0.274 (18.761)
	β	-0.159 (-7.501)	-0.102 (-8.017)	-0.051 (-6.128)	-0.024 (-4.521)	-0.014 (-3.346)	-0.019 (-5.106)	-0.027 (-7.044)	-0.018 (-5.816)
	R^2	5.571	5.119	2.221	0.773	0.405	0.926	3.377	2.015
$\Pi_{t,t+1}^{RW}$	α	0.503 (7.244)	0.295 (6.141)	0.259 (7.321)	0.252 (8.849)	0.267 (10.963)	0.278 (13.391)	0.258 (16.669)	0.243 (17.745)
	β	-0.180 (-10.316)	-0.049 (-5.723)	-0.018 (-4.390)	-0.009 (-3.188)	-0.009 (-3.252)	-0.012 (-5.255)	-0.010 (-4.871)	-0.005 (-4.646)
	R^2	12.697	2.156	0.505	0.226	0.300	0.741	0.787	0.318

iShares China Large-Cap ETF

		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^i$									
$\Pi_{t,t+1}^{HAR-X}$	α	1.088 (10.080)	1.046 (14.738)	0.808 (12.922)	0.643 (12.733)	0.586 (14.126)	0.604 (16.470)	0.838 (33.832)	0.830 (39.531)
	β	-0.060 (-3.095)	-0.053 (-5.052)	-0.011 (-1.209)	0.016 (2.276)	0.025 (5.396)	0.020 (5.512)	-0.023 (-8.490)	-0.024 (-10.486)
	R^2	0.730	1.242	0.095	0.303	1.049	0.980	2.338	3.517
$\Pi_{t,t+1}^{HAR-Combo}$	α	1.060 (9.941)	1.003 (14.096)	0.774 (12.443)	0.615 (12.047)	0.566 (13.455)	0.590 (15.907)	0.829 (33.278)	0.824 (39.488)
	β	-0.055 (-2.892)	-0.045 (-4.466)	-0.005 (-0.555)	0.021 (2.977)	0.029 (6.040)	0.023 (5.982)	-0.022 (-7.869)	-0.023 (-10.203)
	R^2	0.634	0.931	0.019	0.550	1.439	1.276	2.085	3.296
$\Pi_{t,t+1}^{RW}$	α	1.054 (9.638)	0.902 (13.399)	0.684 (12.149)	0.624 (13.459)	0.610 (15.531)	0.635 (17.363)	0.756 (28.001)	0.754 (31.869)
	β	-0.055 (-3.069)	-0.026 (-3.069)	0.011 (1.531)	0.020 (3.290)	0.021 (5.086)	0.015 (3.717)	-0.008 (-2.370)	-0.011 (-3.508)
	R^2	0.824	0.425	0.138	0.632	1.005	0.710	0.403	0.928

Apple Inc

		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^i$									
$\Pi_{t,t+1}^{HAR-RV-J-F}$	α	0.359 (2.209)	0.231 (2.222)	-0.128 (-1.746)	-0.246 (-4.291)	-0.327 (-6.973)	-0.330 (-7.966)	-0.235 (-8.004)	-0.295 (-11.750)
	β	-0.0779 (-2.814)	0.047 (-3.180)	0.020 (2.085)	0.037 (5.399)	0.052 (9.934)	0.048 (10.500)	0.019 (4.753)	0.021 (6.059)
	R^2	1.641	1.413	0.405	2.038	5.608	6.043	1.657	3.189
$\Pi_{t,t+1}^{HAR-RV-LCJ}$	α	0.385 (2.314)	0.192 (1.889)	-0.146 (-1.990)	-0.269 (-4.735)	-0.348 (-7.435)	-0.348 (-8.430)	-0.235 (-7.919)	-0.296 (-11.71)
	β	-0.081 (-2.900)	-0.040 (-2.916)	0.022 (2.472)	0.040 (6.236)	0.055 (10.653)	0.051 (11.004)	0.019 (4.825)	0.021 (6.154)
	R^2	1.788	1.004	0.525	2.438	6.228	6.682	1.627	3.163
$\Pi_{t,t+1}^{RW}$	α	0.159 (1.191)	0.013 (0.158)	-0.099 (-1.476)	-0.096 (-1.537)	-0.087 (-1.523)	-0.094 (-1.801)	-0.156 (-4.890)	(0.202) (-7.104)
	β	-0.050 (-3.016)	-0.010 (-1.483)	0.017 (2.590)	0.012 (1.650)	0.011 (1.489)	0.008 (1.116)	0.006 (1.869)	0.006 (1.827)
	R^2	1.378	0.145	0.623	0.485	0.559	0.336	0.379	0.531

Table 18: Return predictability results – Full sample. The Table shows the results of the return predictability regression for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc. t-statistics in brackets is computed using Newey-West standard errors. R^2 is in percentage.

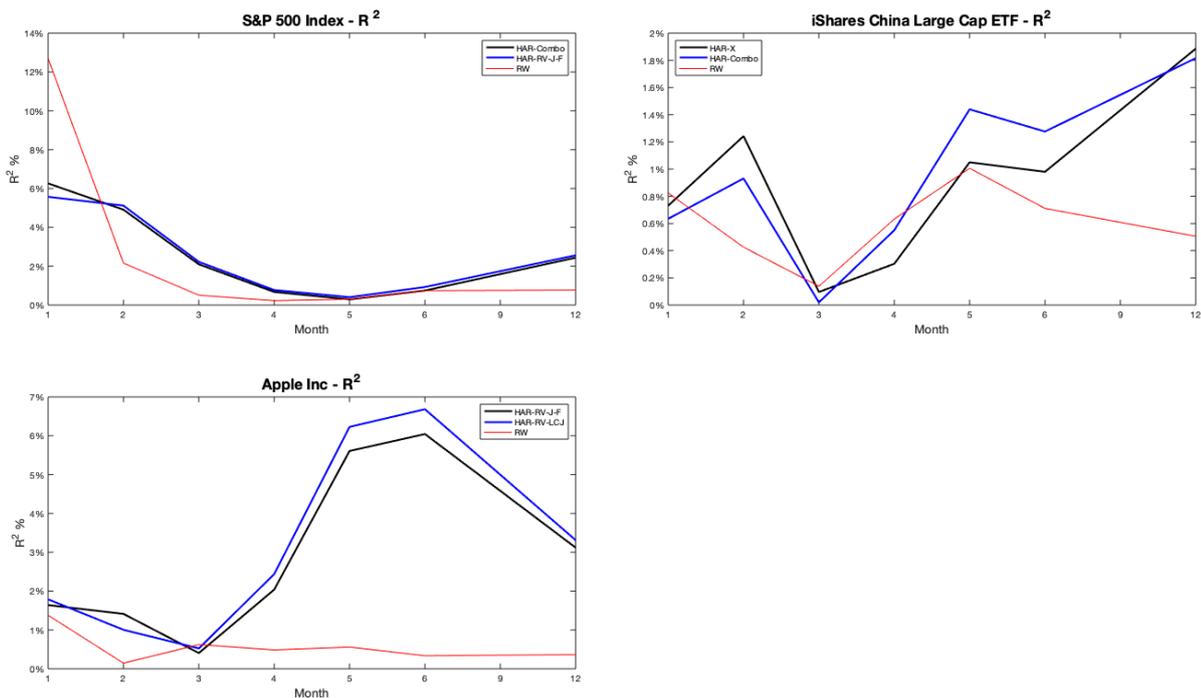


Figure 20:Returns predictability R^2 pattern – Full sample. The figure shows the pattern of the R^2 deriving from the returns prediction regression for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc.

5.2 Returns predictability – excluding Covid-19 pandemic crisis

The previous analysis is carried out considering the Covid-19 pandemic crisis with its extreme impact on both the realized variance side and returns side. The following results, presented in Table 19, show the outcome of return prediction regression, without considering the pandemic shock. The first result of excluding this crisis is an important and general increase in R^2 except for the iShares China Large-Cap. For this ETF the outcomes are puzzling considering that the results of using the random walk are in line with the other more sophisticated model. Nevertheless, this may be due to fact that the pandemic crisis is not the only relevant shock. We can observe from Figure 18 that the VRP of the ETF is affected by a period of high instability in 2015, much more that the other two financial instruments. The greatest improvement is for the S&P 500 Index, especially if we compare the results of the most sophisticated models with the random walk. The increase in R^2 is also confirmed by the results of Apple Inc stock. Even in this case the forecast of the VRP generated by the winning model is superior to the random walk. These results lead us to the conclusion that the VRP has predictive power for returns especially when we exclude

the Covid-19 pandemic crisis or more in general, during a low volatility period. It is also evident that using the VRP deriving from the winning model, summarized in Table 16, it is possible to obtain better prediction results than when using the random walk model.

S&P 500 Index – Excluding Covid-19 crisis

$\Pi_{t,t+1}^i$		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^{HAR-Combo}$	α	1.631 (13.165)	1.075 (14.626)	0.804 (13.713)	0.574 (13.272)	0.507 (13.851)	0.517 (15.530)	0.524 (19.022)	0.451 (17.519)
	β	-1.071 (-7.514)	-0.667 (-8.090)	-0.468 (-6.114)	-0.289 (-4.425)	-0.224 (-3.057)	-0.222 (-4.981)	-0.226 (-7.060)	-0.180 (-5.847)
	R^2	26.660	21.196	17.340	10.325	8.730	11.722	20.866	17.358
$\Pi_{t,t+1}^{HAR-RV-J-F}$	α	1.582 (13.424)	1.079 (14.884)	0.839 (13.990)	0.618 (13.981)	0.546 (14.625)	0.551 (16.082)	0.532 (19.190)	0.484 (18.153)
	β	-1.000 (-7.501)	-0.647 (-8.017)	-0.476 (-6.128)	-0.310 (-4.521)	-0.244 (-3.346)	-0.238 (-5.106)	-0.223 (-7.044)	-0.197 (-5.816)
	R^2	23.577	20.233	18.184	12.025	10.502	13.640	20.726	21.072
$\Pi_{t,t+1}^{RW}$	α	1.051 (4.773)	0.412 (4.502)	0.274 (4.140)	0.199 (4.474)	0.229 (5.758)	0.2613 (6.895)	0.274 (8.888)	0.231 (10.079)
	β	-0.180 (-4.516)	-0.049 (-3.423)	-0.018 (-2.122)	-0.009 (-0.686)	-0.009 (-1.024)	-0.012 (-1.687)	-0.010 (-2.320)	-0.005 (-1.708)
	R^2	27.749	21.530	16.226	8.298	6.749	9.506	19.186	13.796

iShares China Large-Cap ETF - Excluding Covid-19 crisis

$\Pi_{t,t+1}^i$		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^{HAR-X}$	α	1.225 (10.713)	1.122 (14.616)	0.909 (13.113)	0.704 (11.027)	0.622 (12.197)	0.649 (15.121)	0.862 (31.273)	0.885 (37.269)
	β	-0.091 (-3.095)	-0.070 (-5.052)	-0.028 (-1.209)	0.012 (2.276)	0.031 (5.396)	0.030 (5.512)	-0.009 (-8.490)	-0.014 (-10.486)
	R^2	1.200	1.448	0.386	0.119	1.046	1.363	0.271	1.051
$\Pi_{t,t+1}^{HAR-Combo}$	α	1.185 (10.494)	1.052 (13.26)	0.852 (11.914)	0.660 (10.103)	0.591 (11.457)	0.628 (14.489)	0.843 (30.268)	0.868 (36.234)
	β	-0.084 (-3.695)	-0.056 (-4.210)	-0.016 (-1.396)	0.021 (1.905)	0.037 (5.098)	0.034 (6.320)	-0.005 (-1.591)	-0.011 (-3.708)
	R^2	1.059	0.983	0.142	0.369	1.597	1.875	0.100	0.653
$\Pi_{t,t+1}^{RW}$	α	1.104 (7.756)	0.909 (11.010)	0.700 (9.883)	0.609 (9.891)	0.608 (11.475)	0.663 (12.578)	0.805 (31.139)	0.822 (35.238)
	β	-0.067 (-2.587)	-0.027 (-2.101)	0.013 (1.180)	0.031 (3.041)	0.0343 (4.413)	0.027 (3.367)	0.002 (0.738)	-0.002 (-0.673)
	R^2	0.891	0.306	0.127	1.046	1.709	1.532	0.022	0.030

Apple Inc - Excluding Covid-19 crisis

$\Pi_{t,t+1}^i$		Horizon (h)							
		1	2	3	4	5	6	9	12
$\Pi_{t,t+1}^{HAR-RV-J-F}$	α	1.608 (8.394)	1.036 (8.209)	0.424 (4.926)	0.245 (3.497)	0.052 (0.993)	0.001 (0.010)	0.063 (1.700)	-0.026 (-0.784)
	β	-0.496 (-10.595)	-0.352 (-11.861)	-0.214 (-10.800)	-0.179 (-11.896)	-0.129 (-12.941)	-0.114 (-11.170)	-0.123 (-15.353)	-0.097 (-13.553)
	R^2	18.005	18.799	12.519	14.397	11.494	12.085	26.871	21.149
$\Pi_{t,t+1}^{HAR-RV-LCJ}$	α	1.685 (8.584)	0.967 (7.728)	0.384 (4.442)	0.225 (3.159)	0.038 (0.715)	-0.010 (-0.213)	0.055 (1.489)	-0.033 (-0.998)
	β	-0.507 (-10.735)	-0.329 (-11.710)	-0.201 (-10.453)	-0.172 (-11.606)	-0.124 (-12.524)	-0.109 (-10.866)	-0.119 (-15.211)	-0.094 (-13.421)
	R^2	18.628	16.2721	10.885	13.060	10.407	11.003	24.993	19.487
$\Pi_{t,t+1}^{RW}$	α	0.570 (1.466)	0.008 (0.046)	-0.297 (-3.172)	-0.330 (-3.786)	-0.341 (-4.225)	-0.354 (-4.824)	-0.320 (-4.420)	-0.343 (-6.819)
	β	-0.229 (-2.606)	-0.089 (-2.406)	-0.030 (-1.740)	-0.033 (-1.880)	-0.029 (-1.723)	-0.023 (-1.507)	-0.025 (-1.586)	-0.017 (-1.572)
	R^2	9.533	3.029	0.637	1.217	1.446	1.286	2.915	1.611

Table 19: Return predictability results – Excluding Covid-19 crisis. The Table shows the results of the return predictability regression for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc. t-statistics in brackets is computed using Newey-West standard errors. R^2 is in percentage.

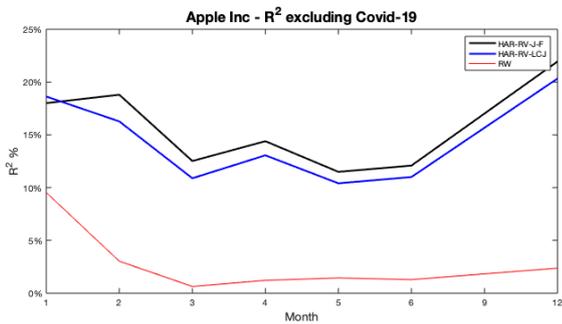
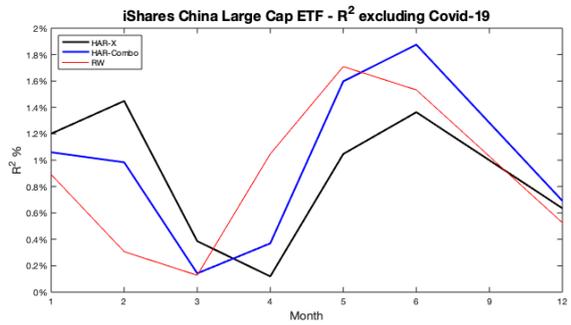
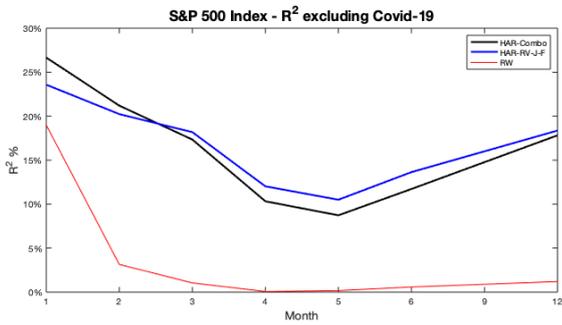


Table 20: Returns predictability R^2 pattern – Excluding Covid-19 crisis. The figure shows the pattern of the R^2 deriving from the returns prediction regression for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc.

6 Conclusion

This thesis deals with the measurement as well as the forecast of realized variance, computed by using high frequency data of three time series: S&P 500 Index, iShares China Large-Cap ETF and Apple Inc. The main problem with asset volatility is that it is not directly observable and must be estimated. A common measure is to use the daily square return, but as we have seen in the second Chapter, we can embed more information in the computation of volatility using high frequency data.

Specifically, in the first part of the thesis we introduced the theory behind realized volatility by defining the concept of the logarithmic price process, quadratic variation and realized variance computed with the average sum of squares of log returns samples at high frequency. To deal with the market microstructure noise we use 5 minute interval. The last part of Chapter 2 is dedicated to the decomposition of realized variance into continuous and jump components using the method proposed by Corsi *et al.* (2010). This method allowed us to overcome the problem of underestimating jumps. Then, in the 3rd Chapter, we introduced the 12 models used to forecast the realized variance, starting from the parsimonious HAR-RV model proposed by Corsi (2009) and addressed more complex models such as HAR-RV-LCJ, HAR-RV-CJ and HAR-Q-F. Thus, we question whether adding or modifying the starting model leads to an improvement in predictive power. The last part of the Chapter is dedicated to the evaluation of the forecast, introducing both the point forecast and density forecast evaluation methodology.

Moving to the empirical application, we can immediately see many points of interest. The first one is the dynamics of the HAR-RV components for the three financial instruments. In particular, we see that during the less volatile periods, for the S&P 500 Index and Apple Inc stock, the volatility is driven by the daily component while after the Covid-19 pandemic crisis, the weekly components take over. However, for iShares China Large-Cap we observed a different outcome. During low volatility periods, the weekly component drives the volatility while the daily parameter increases in importance only after the turmoil period at the end of 2019 and after the pandemic crisis. Another key point of interest is the in-sample fitting results. The most complex models get the best results with respect to the simplest such as the HAR-RV and C-HAR model. In particular, we can observe that the results deriving from Apple Inc and iShares China Large-Cap are quite similar. For both instruments the best models are the HAR-Combo, HAR-Q-D and HAR-Q-F models. If we

consider the three instruments together the model with the best in-sample fitting is the HAR-Combo.

The most interesting part of the thesis is the out-of-sample forecast comparison based on four different forecast horizons: one day, one week, two weeks and one month. Looking at the results for each financial instrument, we can see that there is not a single model that constantly outperform the others. However, it is possible to see a *fil rouge* that connects the three assets:

- i. In one day-ahead forecast *simple is better* in fact the parsimonious HAR-RV model generally outperformed the most complex models.
- ii. The performance of the Q-family models generally decreases with an increase in forecast time horizons. This is especially evident for the S&P 500 Index and iShares China Large-Cap ETF. Excluding the random walk, the HAR-Q-F model is the least performing of all the three instruments and for all the forecast periods.
- iii. Jumps matter. The models that included the jumps component get good out-of-sample forecasting results especially on longer horizons. For example, the two best models for the one month ahead forecast with regards to the theiShares China Large-Cap ETF are the HAR-RV-LCJ and HAR-RV-J-F.
- iv. HAR-Combo is a solid model that achieves very good results not only in-sample but also in the out-of-sample analysis.

It is also possible to predict returns with volatility. In Chapter 5, we used the one month ahead volatility forecast in order to compute the Variance Risk Premium. In particular, we saw that using the VRP which derives from the best one month-ahead realized variance forecasting models (Table 12) we can obtain a better returns prediction with respect to the use of the random walk model. Moreover, we saw that when excluding the Covid-19 pandemic shock we obtain an important increment in the predictive power of the VRP especially for the S&P 500 Index and Apple Inc stock while for the iShares China Large-Cap ETF the results are quite puzzling.

In conclusion, we can affirm that for the three different financial instruments considered, no forecast model definitively prevails in comparison to the others. This leads us to two possible improvements, the first one directly concerns the forecasting models, and the latter the typology of assets which object of our analysis. Related to the models, we saw the good results of the HAR-Combo, however, this model is simply computed with an average of the

other 11 models. A simple improvement may be to exclude, or underweight, models that have had unsatisfactory predictions. We can also move to models that are not based on the HAR-RV such as the MEM (Engle and Gallo (2006), Mixture-MEM (Lanne (2006)), HEAVY (Shephard and Sheppard (2010)) and Realized GARCH (Hansen et al. (2012)). Concerning the second point of improvement, we can continue the analysis to other Indices, ETFs and stocks to assess whether the results obtained for the S&P 500 Index, iShares China Large-Cap ETF and Apple Inc, are confirmed. However, the biggest disadvantage of the use of high frequency data remains the difficulty in dealing with a considerable amount of data as well as it's availability.

Bibliography

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

Andersen, T., T. Bollerslev, F. Diebold, & P. Labys (2001): "The distribution of realized exchange rate volatility." *Journal of the American Statistical Association* (96): pp. 42-55.

Andersen, T., L. Benzoni, J. Lund (2002): "An empirical investigation of continuous-time equity return models." *Journal of Finance* (57): pp.1239-1284.

Andersen, T., T. Bollerslev, F. Diebold, & P. Labys (2003): "Modeling and Forecasting Realized Volatility." *Econometrica* 71 (2): pp. 579-625.

Andersen, T., L. Benzoni (2007): "Realized Volatility. In T. Andersen, R. Davis, J. Kreiss, and T. Mikosch (Eds.), "Volatility. In J. Birge and V. Linetsky (Eds.), *Handbook of Financial time Series*. Springer Verlag.

Andersen, T., T. Bollerslev & F. Diebold (2007): "Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility." *NBER Working Paper Series 11775*, National Bureau of Economic Research, Inc.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2010). "Parametric and nonparametric volatility measurement." . *Handbook of Financial Econometrics: Tools and Techniques*, pp. 67-137 North-Holland.

Andersen, T., T. Bollerslev & X. Huang (2010): "A reduced form framework for modeling volatility of speculative prices based on realized variation measures." *Journal of Econometrics* 160(1): pp. 176-189.

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

Audrino, Francesco, and Yujia Hu (2018). "Volatility forecasting: Downside risk, jumps and leverage effect." *Econometrics* 4, 8.

- Back, K. (1991): "Asset prices for general processes." *Journal of Mathematical Economics* (20): pp. 317-395.
- Bandi, F. & J. Russell (2006a): "Separating microstructure noise from volatility." *Journal of Financial Economics*. (79): pp. 655-692.
- Bandi, F. & J. Russell (2006b): "Volatility. In J. Birge and V. Linetsky (Eds.)", *Handbook of Financial Engineering*. Elsevier.
- Barndorff-Nielsen, O. & N. Shephard (2001): "Non-gaussian Orstein-Uhlenbeck-based models and some of their uses in financial economics." *Journal of the Royal Statistical Society, Series B* (63): pp. 167-241.
- Barndorff-Nielsen, O. & N. Shephard (2002a): "Econometric analysis of realised volatility and its use in estimating stochastic volatility models." *Journal of the Royal Statistical Society, Series B* (64): pp. 253-280.
- Barndorff-Nielsen, O. & N. Shephard (2002b): "Estimating quadratic variation using realized variance." *Journal of Applied Econometrics* (17): pp. 457-477.
- Barndorff-Nielsen, O. & N. Shephard (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps." *Journal of Financial Econometrics* 2(1): pp. 1-37.
- Barndorff-Nielsen, O. & N. Shephard (2006): "Econometrics of testing for jumps in financial economics using bipower variation." *Journal of Financial Econometrics* (4): pp. 1-30.
- Barndorff-Nielsen, O., P. Hansen, A. Lunde & N. Shephard (2008): "Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise." *Econometrica* 76(6): pp. 1481-1536.
- Bates, D. (2000): "Post-'87 Crash Fears in S&P 500 Futures Options." *Journal of Econometrics* (94): pp. 181-238.
- Bates, J. M., & Granger, C. W. (1969). "The combination of forecasts." *Journal of the Operational Research Society*, 20(4): pp. 451-468.

Bekaert, Geert, and Marie Hoerova. "The VIX, the variance premium and stock market volatility." *Journal of econometrics* 183.2 (2014): pp.181-192.

Berkowitz, J. (2001). "Testing density forecasts, with applications to risk management". *Journal of Business & Economic Statistics*, 19(4): pp. 465-474.

Black, F. & M. Scholes (1973): "The pricing of options and corporate liabilities." *Journal of Political Economy* (81): pp. 637-654.

Bollerslev, T., T. H. Law & G. Tauchen (2008): "Risk, jumps, and diversification." *Journal of Econometrics* 144(1): pp. 234-256.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). "Expected stock returns and variance risk premia." *The Review of Financial Studies*, 22(11): pp. 4463-4492.

Bollerslev, Tim, Andrew J. Patton, and Rogier Quaedvlieg (2016). "Exploiting the errors: A simple approach for improved volatility forecasting." *Journal of Econometrics* 192.1 pp. 1-18.

Caporin, M., Rossi, E., & de Magistris, P. S. (2017). "Chasing volatility: A persistent multiplicative error model with jumps." *Journal of Econometrics*, 198(1), pp. 122-145.

Carr, P., & Wu, L. (2009). "Variance risk premiums." *The Review of Financial Studies*, 22(3), pp. 1311-1341.

Clements, M. P., & Hendry, D. F. (Eds.). (2002). *A companion to economic forecasting*. Oxford: Blackwell.

Christensen, B. J. & M. _ . Nielsen (2005): "The Implied-Realized Volatility Relation with Jumps in Underlying Asset Prices." *Working Papers 1186*, Queen's University, Department of Economics

Corsi, F. (2009). "A simple approximate long-memory model of realized volatility." *Journal of Financial Econometrics*: vol.7, no. 2, pp.174-196.

Corsi, Fulvio, Davide Pirino, and Roberto Reno (2010). "Threshold bipower variation and the impact of jumps on volatility forecasting." *Journal of Econometrics* 159.2: pp. 276-288.

Corsi, F., Renò, R. (2012). "Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling." *Journal of Business & Economic Statistics*: vol. 3, pp. 368-380.

Corsi, F., Audrino, F. and Reno, R. (2012). "HAR Modeling for Realized Volatility Forecasting". In: *Handbook of Volatility Models and Their Applications*. pp. 363-382

Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998). "Evaluating density forecasts with applications to financial risk management." *International economic review*, pp. 863-883.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, 33(1), 1-1.

Engle R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation." *Econometrica*, 50, 987-1007

Eraker, B., M. Johannes & N. Polson (2003): "The Impact of Jumps in Volatility." *Journal of Finance* (58): pp. 1269-1300.

Fassas, A. P., & Papadamou, S. (2018). "Variance risk premium and equity returns." *Research in International Business and Finance*, 46: pp. 462-470.

Fleming, J. & B. S. Paye (2010): "High-frequency returns, jumps and the mixture of normals hypothesis." *Journal of Econometrics* 160(1): pp. 119-128.

French, K., G. Schwert & R. Stambaugh (1987): "Expected stock returns and volatility." *Journal of Financial Economics* (19): pp. 3-29.

Hansen, P. & A. Lunde (2006): "Realized variance and market microstructure noise." *Journal of Business and Economic Statistics* 24(2): pp. 127-161.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). "The model confidence set." *Econometrica*, 79(2): pp 453-497.

Huang, X., Tauchen, G., (2005). "The relative contribution of jumps to total price variance." *Journal of Financial Econometrics* (3): pp. 456-499.

Mancini, C. (2009). "Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps." *Scandinavian Journal of Statistics*, 36(2): pp. 270-296.

McAleer, M. & M. Medeiros (2008): "Realized Volatility: A Review." *Econometric Reviews* 27(1-3): pp. 10-45.

Merton, R. (1976): "Option pricing when underlying stock returns are discontinuous." *Journal of Financial Economics* (3): pp. 125-144.

Mincer, J. & V. Zarnowitz (1969): "The evaluation of economic forecasts." New York: National Bureau of Economic Research.

Muller, U., M. Dacorogna, R. Dav, R. Olsen, O. Picket & J. von-Weizsacker (1997): "Volatilities of different time resolutions – analysing the dynamics of market components." *Journal of Empirical Finance* (4): pp. 213-239.

Pan, J. (2002): "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time Series Study." *Journal of Financial Economics* (63): pp. 3-50.

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

Patton, Andrew J., and Kevin Sheppard (2015). "Good volatility, bad volatility: Signed jumps and the persistence of volatility." *Review of Economics and Statistics* 97.3: pp. 683-697.

Protter, P. (1992): "Stochastic integration and differential equations: A new approach." New York: Springer-Verlag.

Rombouts, J. V., Stentoft, L., & Violante, F. (2020). "Dynamics of variance risk premia: A new model for disentangling the price of risk." *Journal of Econometrics*, 217(2): pp. 312-334.

Rossi, B. (2014). "Density forecasts in economics, forecasting and policymaking."

Swanson, N. R., & White, H. (1997). A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics*, 79(4): pp. 540-550.

Theil, H. (1966): *Applied Economic Forecasting* Amsterdam: North-Holland. Zhang, L. (2006): "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach." *Bernoulli* (12): pp. 1019-1043.

Zhang, L., P. Mykland & Y. Ait-Sahalia (2005): "A tale of two time scales: Determining integrated volatility with noisy high frequency data." *Journal of the American Statistical Association* (100): pp. 1394-1411.