Network for Studies on Pensions, Aging and Retirement

Annual VaR from High Frequency Data

Alessandro Pollastri and Peter C. Schotman

NETSPAR ACADEMIC SERIES

DP 08/2016-042

Annual VaR from High Frequency Data

Alessandro Pollastri

Peter C. Schotman

August 28, 2016

Abstract

We study the properties of dynamic models for realized variance on long term VaR analyzing the density of future Integrated Variance. Mixing this density with the conditional density of returns given the volatility we derive the predictive density of returns, which we use to estimate VaR. We find that dynamic specifications characterized by higher persistence lead to more conservative VaR estimates when longer horizons are considered. We compare our long term VaR estimates to the ones obtained using the *square root of time rule*. We show that this scaling rule works approximately well.

[1]Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands.
Email: a.pollastri@maastrichtuniversity.nl

[2]Email: p.schotman@maastrichtuniversity.nl

1 Introduction

Risk management is a fundamental task for financial institutions. A widely known measure of risk is given by VaR, introduced by Jorion and since then it has become a standard indicator for supervisors and risk managers.

Regulators of the banking sector are mainly concerned in next day risk measures of a given position for a certain bank. However, pension funds and insurers, for example, have a longer investment horizons.

In this work we study the issue of time aggregation of VaR and propose a new method on how to compute it for any horizon T using High Frequency data.

According to Andersen, Bollerslev, Diebold, and Labys (2003) we compute long term VaR from High Frequency data using a Mixture model that is to say we assume that returns conditional on the volatility are normal and we specify a time series model for log variance following the findings of Andersen, Bollerslev, Diebold, and Labys (2001): they show that log variance is close to be normally distributed. Through the mixture we obtain the unconditional distribution of cumulative returns and with it we can compute long term VaR.

The usage of High Frequency data for multi period VaR computation has been studied in Beltratti and Morana (1999) who find that using High Frequency data improves the estimation of the risk measure. Our approach differs from the one in the above mentioned study in two fundamental ways: first, we do not rely on GARCH specifications and second, our framework allows to study temporal aggregation of VaR using more flexible assumptions.

Furthermore, we study the effect of different time series specifications on the estimates of VaR. More specifically, we consider three time series models: an autoregressive model of order one, the HAR model introduced by Corsi (2009) which is able to approximate the long memory features of volatility. The last time series model we consider is a fractional which parsimoniously capture long memory features of the data as shown in Baillie (1996). We find that a model characterized by a higher persistence implies a more conservative estimate of VaR depending on the current volatility state. When considering a median level of volatility we notice that a more persistent model generates a greater variance for the distribution of the Integrated Variance. On the other hand, when considering low level of volatility the most persistent model displays the lowest VaR estimate because it is slower to revert to the median level of volatility.

We also investigate the validity of the common practice in long term risk management: the *square root of time rule*. This scaling law states that VaR for the investment horizon T is given by the product of next day VaR and the square root of T. It is widely known that such rule only holds under the assumption of *iid* Normal returns. Christoffersen, Diebold, and Schuermann (1998) ,using a GARCH model, show that the above mentioned rule holds reasonably well at the median level of volatility only. They also point out that drawing results on aggregation is extremely difficult and closed form results can be obtained in very restrictive cases as pointed out by Drost and Nijman (1993).

Danielsson and Zigrand (2006) questions the validity of the above mentioned rule assuming an economy characterized by the Merton model which includes jumps but not stochastic volatility. They find that this scaling law underestimates VaR and that this downward bias gets worse when considering higher risk levels and increasing jump intensity.

Wang, Yeh, and Cheng (2011) also raise doubts on the effectiveness of the *square root of time rule*. They study the validity of such rule when the return process is characterized by serial dependence, volatility clustering and jumps. They find that serial dependence and jumps contribute to an overestimation of multi period Value at risk. On the other hand, in our setting we do not consider jumps and we obtain fat tails via the mixture of the conditional density of returns with the density of Integrated Variance. Using a regression model controlling for the persistence of the different time series we show that the above mentioned scaling rule works approximately well. The paper is organized as follows: section two introduces the methodology to obtain value at risk, section three shows the results and section four concludes the paper.

2 Methodology

2.1 The Model

Let $\mathrm{r_t}$ denote the log-return at time t which conditional on the variance we assume to be normally distributed according to the findings of Andersen, Bollerslev, Diebold, and Labys (2003):

$$\mathrm{r}_t|\sigma_t^2 \sim \mathrm{N}(0, \sigma_t^2) \tag{1}$$

and let $p(\mathrm{r_t}|\sigma_t^2)$ denote the conditional density. We aim to study the issue of risk measures aggregation over an investment horizon T characterized by the cumulative return $\mathrm{R_{t+T}}$:

$$\mathrm{R_{t+T}} = \sum_{\mathrm{i=1}}^{\mathrm{T}} \sigma_{t+i}\epsilon_{t+i} \tag{2}$$

Clearly, the variance of integrated returns is given by:

$$\mathrm{S^2_{t+T}} = \sum_{\mathrm{i=1}}^{\mathrm{T}} \sigma^2_{t+i} \tag{3}$$

It follows from equations (1)-(3) that the conditional distribution of cumulative returns given the integrated variance, $p(\mathrm{R_{t+T}}|\mathrm{S^2_{t+T}})$, is normal with first moment equal to zero and variance equal to $\mathrm{S^2_{t+T}}$.

The predictive density of cumulative returns with investment horizon T given the information set up to time t, $\mathrm{I_t}$, can be obtained as:

$$p(\mathrm{R_{t+T}}|\mathrm{I_t}) = \int p(\mathrm{R_{t+T}}|\mathrm{S^2_{t+T}})p(\mathrm{S^2_{t+T}}|\mathrm{I_t})d\mathrm{S^2_{t+T}} \tag{4}$$

We model aggregation through the conditional distribution $p(\mathrm{R_{t+T}}|\mathrm{S^2_{t+T}})$ which under the above mentioned assumptions is still Normal with mean zero and variance $\mathrm{S^2_{t+T}}$. We use realized variance as a proxy for σ_t^2, this is defined as the sum of the intradaily squared log returns.

Furthermore, in order to obtain an $p(\mathrm{R_{t+T}})$ we need the unconditional density of the integrated variance $p(\mathrm{S^2_{t+T}})$. In what follows we assume a time series model for the logarithm of realized variance and we denote it by h_t. For this choice we are motivated by the empirical findings of Andersen, Bollerslev, Diebold, and Labys (2001) who find that realized variance is close to being log normally distributed.

The first model we consider is the HAR model on log realized variance introduced in Corsi (2009), which captures the long memory property of realized variance in a parsimonious way:

$$h_{t+1} = \mu + a_1(h_t - \mu) + a_2(h_t^w - \mu) + a_3(h_t^m - \mu) + \omega\eta_{t+1} \tag{5}$$

where μ is the unconditional mean, η_{t+1} is standard normally distributed and ω is the VoV parameter. Furthermore, h_t^w is the average of the log realized variance of the previous five days whereas h_t^m is the average of the log realized variance of the previous twenty one days.

We also consider as a toy model a simple autoregressive model of order one:

$$h_{t+1} = \mu + a_1(h_t - \mu) + \omega\eta_{t+1} \tag{6}$$

Having a look at (6) it is straightforward to notice that this model is nested in the

HAR model when a_2 and a_3 are equal to zero.

In order to compare the value at risk at different horizons from the HAR model with the ones obtained previously we need to obtain the Integrated Variance, $\mathrm{S}^2_{\mathrm{t+T}}$. For the AR model this can simply obtained by recursion. Exploiting this structure, we recognize that the autoregressive parameter has a strong impact on the variance of the volatility distribution, which results in a more fat-tailed return distribution when the more persistent the dynamics is.

Andersen, Bollerslev, Diebold, and Labys (2003) use a Fractional model in order to forecast realized variance, this is given by:

$$(1-L)^d(h_{t+1}) = \mu + \omega\eta_{t+1} \tag{7}$$

where d is the fractional parameter and the above equation can be rewritten expanding the lag polynomial:

$$h_{t+1} = \mu + \sum_{s=1}^{\infty} \Delta_s(h_{t+1-s} - \mu) + \omega\eta_{t+1} \tag{8}$$

with the infinite order autoregressive weights Δ_s given by the binomial expansion,

$$(1-L)^d = 1 - dL - (1-d)dL^2/2! - d(1-d)(2-d)L^3/3! + \cdots$$

A recursive formula for the autoregressive weights can be found in the Appendix. This model is stationary for $d \in (0, 0.5)$ and non-stationary but mean reverting for $d \in [0.5, 1)$.

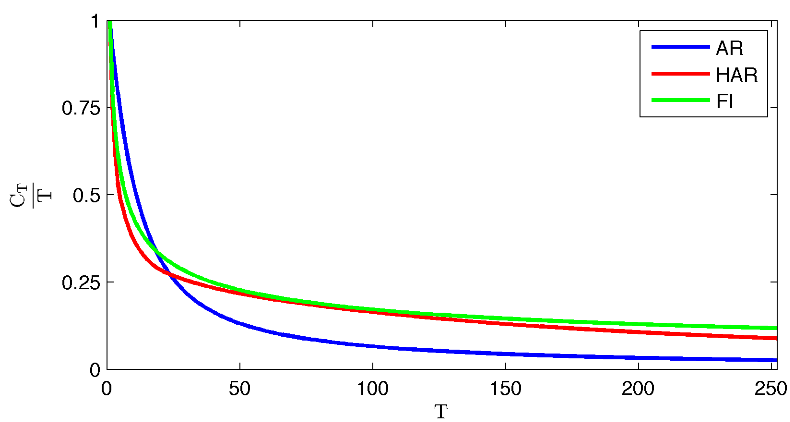
In order to give an idea on how these time series models differ from each other a useful tool is given by the standardized cumulative impulse response functions. These measures how persistent is a shock to volatility. In order to obtain them, we can first rewrite the models in MA(∞) representation:

$$h_{t+1} = \sum_{j=0}^{\infty} \theta_j \eta_{t+1-j} \tag{9}$$

where θ_j differs according to the model. The impulse response of a shock today to volatility at time $t + j$ is given by the coefficients of the above representation. The cumulative impulse response functions are then given by the sum of the θ_j coefficients for the period T.

$$C_T = \sum_{j=0}^{T} \theta_j \tag{10}$$

Figure 1: Standardized Cumulative Impulse Responses Functions



This graph shows the cumulative impulse response functions standardized by T of the three time series models we consider in our analysis with parameters' estimates of the *SPY*.

The standardized cumulative IRFs are then obtained by dividing C_T of equation (10) by the horizon T. Figure 1 shows them for the different models we consider and with the parameter estimates set equal to the one of the SPY. It is clear that for very short horizons there is no big difference among the models and the autoregressive specification displays a stronger cumulative response. However, C_T drops much faster for the autoregressive model compared to the specifications characterized by a longer memory. Finally, from the graph it is also shown that the Fractional model implies a stronger persistence to a shock in volatility.

3 Results

3.1 Data

Our dataset is composed by the SPY an ETF that tracks the SP500 and the intraday prices of current constituents of the Dow Jones starting from November 1995 until November 2015. This leaves us with time series of realized variance of 5036 observations computed using 5-minutes returns.

Table 1 shows the main summary statistics for the returns and of the stocks belonging to our dataset and the last row shows the same statistics for the SPY.

Table 1: Returns Distribution

	$r_{i,t}$				$r_{i,t}/\sigma_{i,t}$			
Stock	Mean	St.dev.	Skew.	Kurt.	Mean	St.dev.	Skew.	Kurt.
Min	-0.069	1.142	-0.202	5.747	-0.016	0.873	0.006	2.696
0.10	-0.034	1.237	-0.004	6.444	-0.007	0.876	0.039	2.767
0.25	-0.003	1.400	0.066	7.221	0.011	0.909	0.046	2.777
0.50	0.010	1.561	0.126	8.728	0.020	0.927	0.060	2.842
0.75	0.026	1.712	0.236	10.533	0.027	0.949	0.096	2.908
0.90	0.040	2.046	0.342	12.476	0.040	0.961	0.114	3.008
Max	0.080	2.367	0.588	16.171	0.084	0.970	0.154	3.060
Mean	0.007	1.638	0.165	9.617	0.023	0.923	0.074	2.865
St. dev	0.049	0.442	0.255	3.731	0.033	0.039	0.051	0.133
SPY	-0.005	1.021	-0.104	10.535	0.053	0.961	0.003	2.789

This table shows the main statistics on the current DJIA constituents. The first four moments are shown for the daily returns and the daily returns standardized by the current volatility.

The first column indicates the percentiles of the relevant summary statistic for the daily returns $r_{i,t}$ of the companies in our dataset. For example, the row that is called *Min* shows the minimum values of the relevant summary statistics the dataset we have whereas *Max* shows the maximum values. The second and third last rows show the mean and the standard deviation of the summary statistics in the different columns. We also show the same summary statistics for $r_{i,t}/\sigma_{i,t}$. It is interesting to notice that returns standardized by the volatility are much close to normality compared to the non standardized ones. This assumption holds reasonably well for the SPY and for the single stocks we have the kurtosis ranging from 2.696 to 3.060 against a kurtosis ranging from 5.747 to 16.171 for $r_{i,t}$.

The second set of summary statistics is related to the realized variances, table 2 shows the same summary statistics for realized variances and log realized variances.

Table 2: Volatility Distribution

	$\sigma^2_{i,t}$				$\ln \sigma^2_{\mathrm{i,t}}$			
Stock	Mean	St.dev.	Skew.	Kurt.	Mean	St.dev.	Skew.	Kurt.
Min	1.585	2.354	6.165	75.501	-0.070	0.832	-0.063	2.250
0.10	1.834	2.860	6.467	79.247	0.126	0.858	0.265	2.523
0.25	2.246	3.406	7.330	108.971	0.334	0.912	0.291	2.837
0.50	2.806	4.102	10.557	215.180	0.477	0.952	0.347	2.975
0.75	3.246	6.267	16.202	449.447	0.668	1.005	0.420	3.232
0.90	4.413	10.491	30.717	1495.202	0.884	1.115	0.476	3.628
Max	7.173	15.194	58.820	3905.738	1.297	1.203	0.514	3.854
Mean	3.329	6.382	19.465	904.183	0.532	0.982	0.322	3.043
St. dev.	1.944	4.782	19.401	1416.553	0.465	0.136	0.193	0.575
SPY	1.148	2.174	9.715	161.728	-0.473	1.046	0.259	3.189

This table shows the main statistics on the current DJIA Realized Variances constituents. The first four moments are shown for the Realized Variances and the log transformation.

The kurtosis of realized variance ranges from a minimum of 75.501 to a maximum of 3905.738 whereas for log realized variances it ranges from a minimum of 2.250 to a maximum of 3.854. For the SPY instead the kurtosis is 161.728 for the realized variance whereas it is 3.189 for log realized variance: this confirms that lognormality of realized variance is a reasonable approximation as stated in Andersen, Bollerslev, Diebold, and Labys (2001).

The AR and HAR model can be estimated using simple OLS whereas for the fractional model we use the Local Whittle Estimator as pointed out by Shimotsu and Phillips (2006)[1]. Table 3 table shows the estimates of the parameters characterizing each model.

Table 3: Parameter estimates

Par.	AR	HAR	FI
d			0.593
			(0.039)
a_1	0.848	0.437	
	(0.01)	(0.02)	
a_2		0.339	
		(0.03)	
a_3		0.182	
		(0.02)	
μ	-0.473	-0.464	
ω	0.555	0.512	0.514

This table shows the estimates of the parameters of the models and the standard errors that we consider for the SPY. For the estimation of d, the number of frequencies used is set equal to $J = T^{\alpha}$ with $\alpha = 0.6$. The asymptotic standard error is then given by $se(d) = 1/\sqrt{4J}$.

[1]We have estimated the fractional model also using the Exact Whittle Estimator and the standard GPH algorithm. The Local Whittle Estimator however is robust to autocorrelation. The estimates still lie in the non-stationary region

Table 3 shows that for the Fractional model d lies in the non-stationarity region, to check the stability of this finding we have set $\alpha = 0.5, 0.7$. With these values, the resulting estimates of d are respectively 0.602 and 0.575. This implies that the sample average is not a consistent estimator of the unconditional mean μ. Values for d in the non-stationary region in the context of realized variances have been found by Rossi and Santucci de Magistris (2014) and Luciani and Veredas (2015).

Table 4 displays the parameter estimates for the single stocks in our dataset. For ease of exposition we show three different parameterizations according to the persistence of the stock.

Table 4: Parameter estimates of Single Stocks

Par.	AR			HAR			FI		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
d							0.562	0.640	0.661
a_1	0.756	0.832	0.878	0.334	0.322	0.301			
a_2				0.334	0.372	0.395			
a_3				0.276	0.272	0.282			
μ	0.404	0.681	0.213	0.093	0.464	0.855			
ω	0.559	0.524	0.536	0.511	0.467	0.469	0.517	0.482	0.489

This table shows the estimates of the parameters of the models that we consider for the current constituents of DJIA. For the estimation of d, the number of frequencies used is set equal to $J = T^{\alpha}$ with $\alpha = 0.6$. The asymptotic standard error is then given by $se(d) = 1/\sqrt{4J}$, which in our case is equal to 0.039.

The stocks are chosen according to the degree of persistence defined by autoregressive parameter. The second to fourth columns show the values of the parameters for the autoregressive model and for the different degrees of persistence. Columns five to seven show the parameter values for the HAR model and the last three columns show the parameter values for the Fractional model. The stock characterized by the lowest persistence is *Merck and Co.*, the medium one is *Home Depot* and the high one is *Walmart*. Also for single stocks we find that the fractional parameter lies in the non-stationarity region and that there is not so much heterogeneity among stocks as highlighted in columns 6 to 9. The same applies for the volatility of volatility parameter, ω, which varies from 0.504 to 0.566 for the AR model whereas it is slightly lower for the other models. More interesting is the diversity in the unconditional mean for the AR type of models ranging from the least persistent stock model implied mean annualized volatility of approximately 19% to the most persistent stock one of 17% for the AR. For what concerns the HAR model the values are slightly lower than the ones for the AR model.

3.2 Distribution of Integrated Variance

This section shows the results concerning the distribution of Integrated Variance, $p(\mathrm{S}^2_{\mathrm{t+T}})$, which is a key element in the estimation of VaR in our analysis. Moreover, in order to compare the different models we evaluate all of them at the sample median m given that for the fractional model we can not consistenly estimate the unconditional mean.

Table 5 shows the value of the median for the SPY and the Single Stocks in our dataset. It confirms also that the normality of log realized variances is a reasonable approximation. As shown in table 5 the sample median for the SPY is much lower than the ones for the single stocks of the DJIA. This is not surprising and it is a result of diversification.

Table 5: Sample median m for companies in our dataset

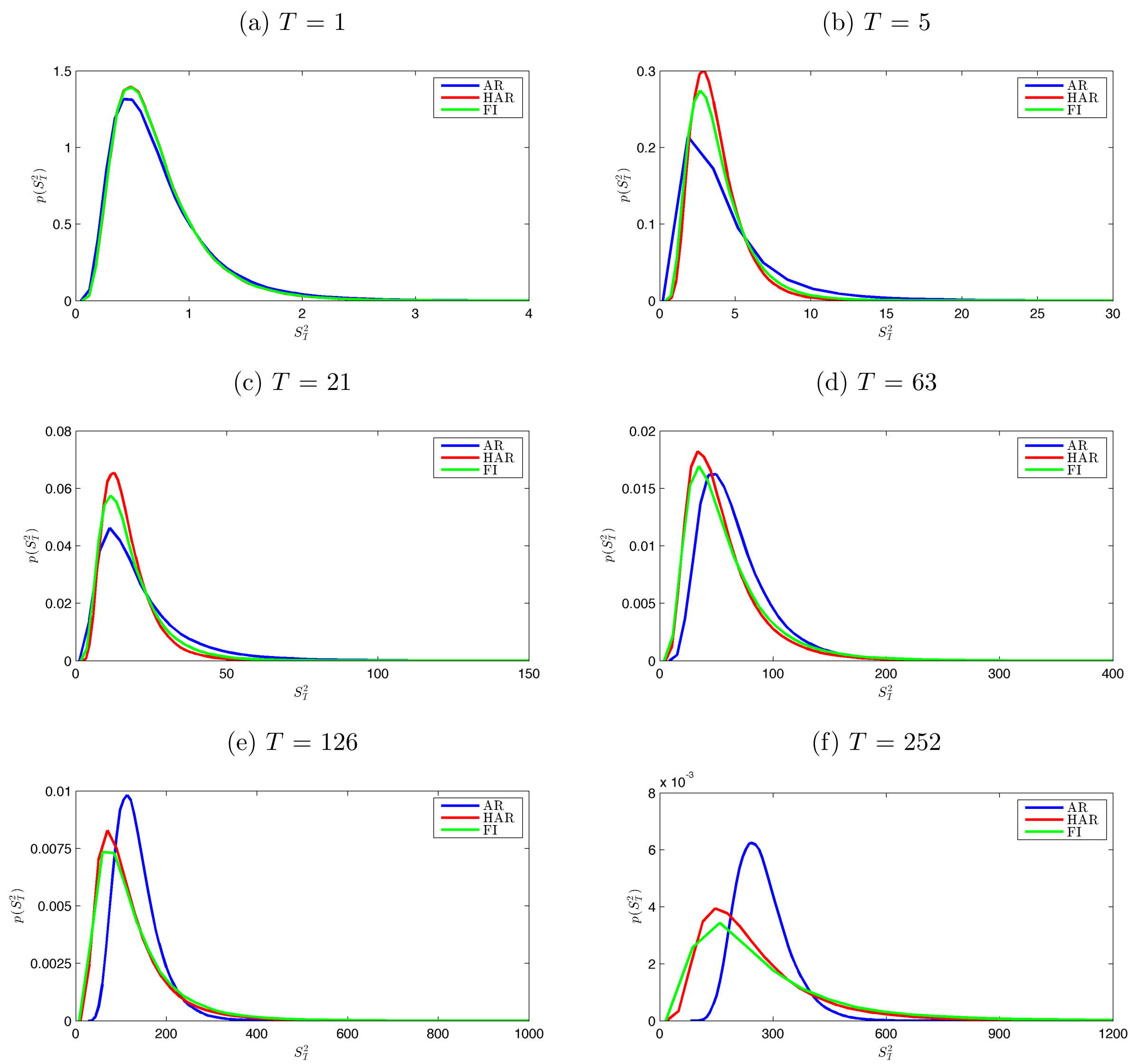
	SPY	Single Stocks		
		Low	Medium	High
m	-0.471	0.404	0.652	0.212

This table shows sample median of log realized variances, h, in our dataset for the SPY and for the single stocks we show the tenth percentile, the median and the ninetieth percentile

Figure 2 shows the kernel densities of the simulated Integrated Variance under the different models in our setting, ranging from the next day, $\mathrm{T} = 1$ until a year, $\mathrm{T} = 252$. It is clear the evolution of the densities through time: for next day integrated variance distribution, this case boils down to the density of σ^2_{t+1} which by construction we know to be a log-normal random variable. As T increases $\mathrm{S}^2_{\mathrm{t+T}}$, defined by 3, is the sum of log-normal random variables and as such has an unknown density for $\mathrm{T} > 1$. However, from figure 2 it is possible to grasp some features among the different models. When $T = 1$ the models are almost not distinguishable, this is due to the fact that models' dynamics hardly play a role for such short aggregation horizon. The difference among the distributions arises from the different values of ω, which affects both the mean and the variance of the distribution. As T increases the difference among the models becomes clear: the AR leads to a wider distribution for the Integrated Variance up to the three months horizon and after the HAR and the Fractional model result in wider densities.

This can also be seen from table 6 that shows, with a slight abuse of notation, the mean, $\mu(\mathrm{S}^2_{\mathrm{t+T}})$, where the brackets emphasize that it is related to $\mathrm{S}^2_{\mathrm{t+T}}$. It also displays the volatility, $\omega(\mathrm{S}^2_{\mathrm{t+T}})$, of the Integrated Variance. The first two columns

Figure 2: Kernel Densities of Integrated Variance under different models



This graph shows the kernel density of the Integrated Variance simulated for the models that we consider for the SPY. The different panels display the change in these density as a function of T.

Table 6: Moments of the Integrated Variance Distribution

T	AR		HAR		FI	
	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$
1	0.731	0.439	0.714	0.391	0.715	0.393
5	4.328	3.375	3.719	1.689	3.817	2.007
21	21.269	15.851	16.720	8.124	17.689	10.353
63	66.780	33.741	55.434	33.576	59.149	39.091
126	134.506	49.419	119.446	79.814	127.298	91.681
252	270.347	70.731	253.366	163.454	278.810	219.177

This table shows the moments of the Integrated Variance distributions simulated for the models we consider for the SPY. The different rows display the moments as a function of T.

exhibit the moments for the AR model, the next two for the HAR and the last two for the Fractional model. [2] The rows refer to the aggregation horizon we are considering. As pointed out previously for $T = 1$ the first two moments of the densities are very similar, with the AR presenting higher values compared to the other models. Furthermore, the AR model presents a higher mean of the distribution for all aggregation horizons compared to the HAR model; this does not hold for the Fractional model that presents a higher mean for the one year horizon. Also interesting to analyze is the difference in the second moments of the density for the different models: again the AR exhibits a higher volatility of volatility up to three months compared to the HAR model and up to a month related to the Fractional model. Having a

[2]For the AR model $\mu(\mathrm{S}_{\mathrm{t+T}})$ can be easily computed by the following formula:

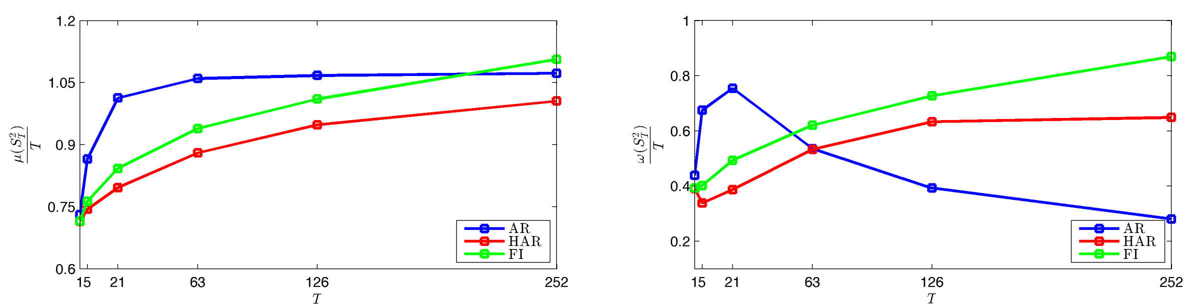
$$\mu(\mathrm{S}_{\mathrm{t+T}}) = \sum_{i=1}^{\mathrm{T}} e^{\mu + \frac{\omega^2}{2}\sum_{s=0}^{i-1} a_1^{2s}}$$

closer look at the table for the models that are characterized by a longer memory we can notice that for any horizon the Fractional model presents higher values for any moment. Moreover, as the horizon grows the difference between the two models becomes more evident.

Figure 3 shows these facts: panel A shows the mean of the distribution of Integrated Variance under the different models whereas panel B shows the standard deviation both scaled by the T. Panel A clearly shows that the AR spikes very quickly com-

Figure 3: First two moments of the Integrated Variance distribution under different models scaled by the horizon T

(a) $\mu(\mathrm{S}^2_{\mathrm{t+T}})$ (b) $\omega(\mathrm{S}^2_{\mathrm{t+T}})$



This graph shows the first two moments of the Integrated Variance distributions simulated for the models for the SPY scaled by the horizon T.

pared to the other models and around the 3 months horizon it starts flattening whereas the other two models have an upward sloped behavior for the mean of the Integrated Variance distribution. Panel B shows that the AR rises very quickly and then it decays sharply whereas the models with longer memory exhibit an increasing standard deviation: this is stronger for the Fractional model.

In what follows we briefly show the results for the different models on the stocks selected according to their persistence as in table 4. The sample median of these stocks is given in table 5 implying a sample annual median volatility of 19.42 % and 17.65 %.

In a similar way to the SPY table 7 shows the moments for the different models and the different horizons for the least persistent stock. The differences in the values of

Table 7: Moments of the Integrated Variance Distribution

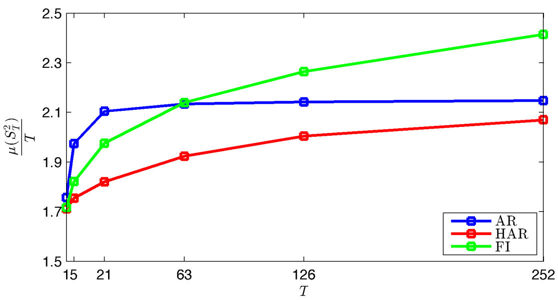
T	AR		HAR		FI	
	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$
1	1.756	1.065	1.711	0.933	1.716	0.950
5	9.869	6.402	8.772	3.494	9.105	4.588
21	44.182	21.484	38.221	14.492	41.483	22.258
63	134.436	40.861	121.152	54.151	134.756	77.927
126	269.903	59.008	252.534	120.790	285.350	173.638
252	541.201	84.670	521.619	237.248	608.426	391.308

This table shows the moments of the Integrated Variance distributions simulated for the models we consider and for the least persistent stock. The different rows display the moments as a function of T.

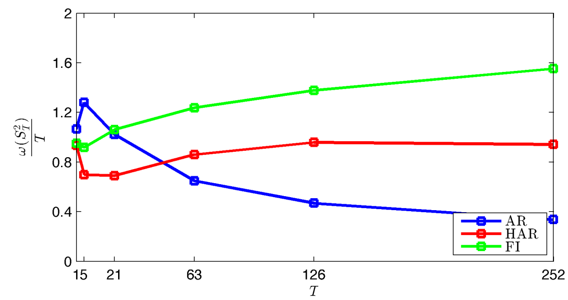
$\mu(\mathrm{S}^2_{\mathrm{t+T}})$ derive from the diversity in the medians, which as previously explained is much lower for the SPY compared to the Single Stocks. For the one day horizon, the difference in the moments stems from the diversity in the VoV parameter for the models we consider. Moreover, the main results found for the SPY hold for this stock: at longer horizons the Fractional model presents a distribution of integrated variance with a higher mean and higher standard deviation compared to the other two models. Moreover, the autoregressive model displays a higher first moment compared to the HAR but a vastly lower standard deviation. In order to make a further comparison to the SPY results, we also show the same graphs displaying the first two moments scaled by the horizon. From figure 4 we can notice in panel A that

Figure 4: First two moments of the Integrated Variance distribution under different models scaled by the horizon T

(a) $\mu(\mathrm{S}^2_{\mathrm{t+T}})$



(b) $\omega(\mathrm{S}^2_{\mathrm{t+T}})$



This graph shows the first two moments of the Integrated Variance distributions simulated for the models for the least persistent stock scaled by the horizon T.

the scaled mean of the distribution for the the autoregressive model flattens at an earlier aggregation horizon compared to the case of the SPY. The same behaviour can be recognized for the volatility of the Integrated variance: it spikes at the weekly horizon and then it flattens down. For the other two models the previous findings are confirmed. The Fractional model shows a scaled mean higher than the autoregressive at the six months horizon whereas for the SPY this holds only at the yearly horizon. The HAR exhibits the same behavior shown in the SPY both for the first and the second moment. However, for the least persistent stock the volatility of Integrated variance flattens out more quickly than for the SPY.

The last set of results of this section is related to the most persistent stock. Table 8 shows the same results as previously explained for the most persistent stock. Although the median for this stock is lower than the least persistent stock, at long

Table 8: Moments of the Integrated Variance Distribution

T	AR		HAR		FI	
	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$	$\mu(\mathrm{S}^2_{\mathrm{t+T}})$	$\omega(\mathrm{S}^2_{\mathrm{t+T}})$
1	1.426	0.824	1.379	0.686	1.392	0.725
5	8.544	6.777	7.039	2.535	7.535	4.097
21	44.465	38.836	30.767	11.307	36.588	25.264
63	141.016	85.994	98.810	47.737	129.311	111.595
126	288.917	131.658	216.020	133.224	303.728	316.369
252	580.739	189.645	473.370	343.449	726.905	918.900

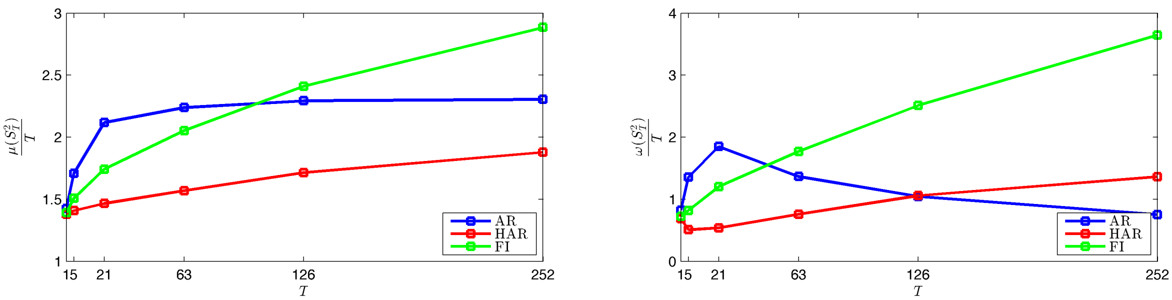
This table shows the moments of the Integrated Variance distributions simulated for the models we consider and for the most persistent stock. The different rows display the moments as a function of T.

horizons the mean of the distribution of Integrated Variance is higher. This means that a higher persistence implies a greater mean and volatility of the Integrated Variance distribution.

Finally, figure 5 shows the first two moments scaled by the horizon T. The results are comparable to the ones obtained for the previous stocks analyzed. For the autoregressive model the mean of the Integrated Variance distribution spikes and then flattens around the same aggregation period as the SPY due to similarity of the autoregressive parameter. The level is of course different and it is due to the value of the median of log realized variance for this stock. For what concerns the models with longer memory, we can notice that the scaled mean is increasing with T and this is particularly evident for the Fractional model. Related to the second moment we can see that the autoregressive model jumps at the one month horizon and then it decays whereas the other models exhibit an increasing scaled volatility. The other two models instead exhibit an increasing scaled volatility of Integrated Variance.

Figure 5: First two moments of the Integrated Variance distribution under different models scaled by the horizon T

(a) $\mu(\mathrm{S}_{\mathrm{T}}^2)$ (b) $\omega(\mathrm{S}_{\mathrm{T}}^2)$



This graph shows the first two moments of the Integrated Variance distributions simulated for the models for the most persistent stock scaled by the horizon T.

The HAR is characterized by the lowest value and then, as T increases, it overtakes the autoregressive model. The Fractional model displays the highest volatility of Integrated Variance compared to the other models as in the previous cases.

3.3 Long Term VaR

In this section we compute VaR under the different time series models and for different horizons T. The results can easily be extended to Expected Shortfall.

We assume that the investor holds a position X_t and that the future (unknown) position is given by:

$$\mathrm{X}_{\mathrm{t+T}} = \mathrm{X}_t e^{\mathrm{R}_{\mathrm{t+T}}} \tag{11}$$

where $\mathrm{R}_{\mathrm{t+T}}$ is defined in equation (2). Furthermore, assuming that $\mathrm{X}_t = 1$, the VaR of the position $\mathrm{X}_{\mathrm{T}} - \mathrm{X}_t$ given the information set at time t at level α is obtained as follows:

$$\text{VaR} = 1 - e^{-q_\alpha} \tag{12}$$

where q_α is the α-quantile of the distribution function of $\text{R}_{\text{t+T}}$ that we obtain solving the following equation;

$$\alpha = \int_{-\infty}^{-q_\alpha} p(\text{R}_{\text{t+T}}) d\text{R}_{\text{t+T}} \tag{13}$$

First we compute VaR according to (12) for the different models and for the maturities that we have considered in the previous section. We start focusing on the *SPY* and then turn to the same stocks considered before. Moreover, the results still consider that the models are evaluated at the median level of realized variance.

Table 9: VaR for the different models we consider and for two confidence levels

T	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	1.381	2.124	1.368	2.080	1.368	2.083
5	3.300	5.219	3.105	4.606	3.134	4.723
21	7.175	11.157	6.462	9.560	6.611	9.982
63	12.475	18.243	11.387	17.030	11.724	17.697
126	17.306	24.375	16.239	24.175	16.680	25.028
252	23.656	32.310	22.737	33.075	23.565	34.970

This table shows the VaR at two confidence levels computed for the different models for the *SPY*. The different rows display the aggregation horizon ranging from one day, first row, to one year, last row.

Table 9 shows the VaR in percentage point computed for the different models and for several horizons. Furthermore, for each model we show the VaR computed at the usual confidence levels $\alpha = 5\%$ and $\alpha = 1\%$. Clearly, VaR is growing due to the monotonically increasing moments of S^2_{T}. When $\text{T} = 1$, there is a very small difference between the models and this is due to the fact that the distribution of the Integrated Variance does not vary too much across the different time series models we have considered. The variety comes from the different values in the VoV parameter, ω. When moving from the next day VaR to longer horizons, we can notice that for $\alpha = 5\%$ the autoregressive model exhibits a higher value. However, for the lower confidence level the autoregressive model is no longer the one presenting the highest VaR at the biannual horizon and at the annual horizon. In fact, at the half year horizon the fractional model overtakes the autoregressive one and at the annual both the HAR and the fractional exhibit a higher VaR. This can be explained remembering the results from the previous section where at longer horizons, a model characterized by a longer memory display either a higher volatility of Integrated Variance, the HAR model, or both a higher mean and volatility of the Integrated Variance distribution, see the fractional model.

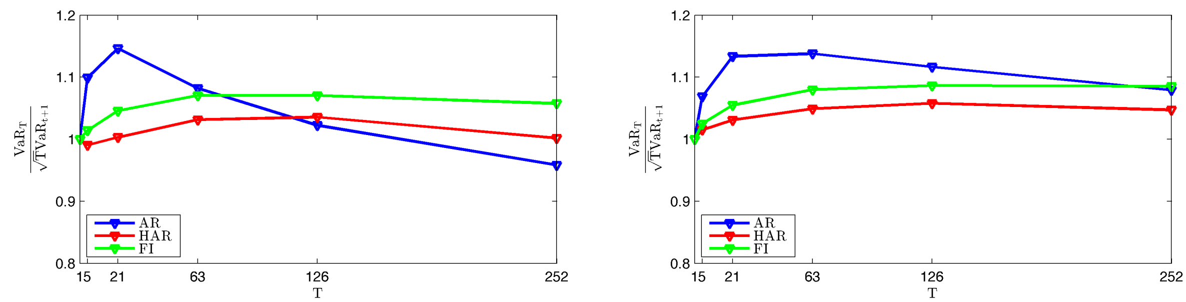
In risk management, the common practice to obtain multi-period VaR is to multiply the one day VaR by the square root of the horizon for which we want to compute it:

$$\mathrm{VaR_T} = \sqrt{\mathrm{T}}\mathrm{VaR_{t+1}} \tag{14}$$

This approach in (14) is called *square root of time rule* and it only holds if the returns are *iid* normal. In what follows we compare our estimates of $\mathrm{VaR_T}$ that is displayed in table 9 to the one obtained using the above mentioned scaling law. Figure 6 shows the ratio between the long term VaR obtained using the methodology that we propose and the one derived adopting the *square root of time rule*. The latter is obtained multiplying the VaR for $\mathrm{T} = 1$ times the square root of the relevant aggregation horizon. Moreover, the graph shows this ratio for the time series models we consider.

Figure 6: Ratio of $\mathrm{VaR_T}$ obtained with two different approaches

(a) $\alpha = 1\%$ (b) $\alpha = 5\%$



This graph shows the ratio between $\mathrm{VaR_T}$ obtained using our model and different time series specifications for log realized variance and the long term VaR obtained using the *square root of time rule* for the *SPY*. Moreover, panel a shows the ratio at the confidence level of 1% whereas panel b displays the same ratio at 5%.

From figure 6 it is clear that at 1% confidence level the *square root of time rule* overestimates $\mathrm{VaR_T}$ of approximately 5% for the autoregressive model and underestimates it of almost 6% for the fractional model. The HAR model instead lies in between the two other models: compared to the above mentioned scaling rule the long term VaR would be slightly overestimated. Panel b shows instead that at a higher confidence level this scaling rule underestimates long term VaR for all models. In panel A the ratio for the autoregressive model spikes up to 15% for the monthly horizon and then it decays. In panel B instead, the ratio for this specification grows again quickly and then it flattens around a value slightly above 20%. For the two other time series models, the two panels show the same pattern: the fractional model implies that the *square root of time rule* underestimates the annual VaR of approximately 6% for

$\alpha = 1\%$, whereas at 5% it is slightly above 8%. The HAR, instead, implies that this scaling law underestimates the biannual VaR of almost 4% at $\alpha = 1\%$ whereas for the annual one the *square root of time rule* leads almost to the same value for VaR. At $\alpha = 5\%$ this scaling law underestimates annual VaR of more than 4%.

Another widely used scaling law in finance is related to the family of stable distributions. This class includes inter alia the Normal, the Cauchy and the Levy distribution and they are usually denoted by $\mathbf{S}(\xi, \lambda, \gamma, \delta)$ where the terms in brackets indicated the relevant parameters. The last two for example are respectively the scale and location of the distribution whereas the first two are the index of stability and the skewness parameter. The former parameter is defined in $\xi \in (0, 2]$ whereas the latter is defined in $\lambda \in [-1, 1]$. The Gaussian case is obtained setting $\xi = 2$ and $\lambda = 0$.

The scaling law under this family of distribution is the following:

$$\mathrm{VaR}_{\mathrm{T}} = \mathrm{T}^{\frac{1}{\xi}} \mathrm{VaR}_{\mathrm{t+1}} \tag{15}$$

Clearly, the scaling law in (15) boils down to (14) when $\xi = 2$, which is the Normal distribution case. Usually, the stability index is estimated on returns data using semi-parametric method such as the Hill estimator. However, we are confronted with the opposite problem: given the long term VaR and the aggregation horizons we want to estimate the stability index under the different models.

We obtain ξ through the following regression:

$$\ln \mathrm{VaR}_{\mathrm{T}} = \beta_0 + \beta_1 \ln \mathrm{T} + u_{\mathrm{T}} \tag{16}$$

where the regressand is obtained under the different log variance specifications and $\beta_1 = \frac{1}{\xi}$.

Table 10: Estimates of the parameters of the Stable Regression

Par.	AR	HAR	FI
β_0	0.9949	0.7297	0.7467
	(0.0014)	(0.0005)	(0.0004)
β_1	0.4540	0.5055	0.5108
	(0.0003)	(0.0001)	(0.0001)

This table shows the estimates of the coefficient and standard errors of regression in (16) for the different models we consider.

Table 10 displays the estimates of above mentioned regression. The rows show the values of the parameters and the standard errors. The AR model implies a stability index greater than 2, whereas the more persistent models exhibit a value which is smaller than this value. Moreover, higher persistence leads to decreasing values of such parameter. For the fractional model, this parameter is equal to 1.958.

Table 11 shows the VaR for the least persistent stock in our dataset. Compared to the SPY, VaR estimates are higher because the median of this stock is bigger. Moreover, the table confirms that for T = 1, VaR is quite similar among the different models and the two risk levels. However, when the horizon starts becoming large the difference among the time series specifications is clear and the distance between the fractional and the autoregressive model is of 1% at the highest confidence level and more than 4% at the lower one. Furthermore, this table shows that up to the first quarter the autoregressive model exhibits a higher VaR at both confidence levels whereas when T refers to the six months the fractional model display a higher estimate of VaR at the $\alpha = 1\%$.

Table 11: VaR for the different models we consider and for two confidence levels

T	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	2.131	3.275	2.109	3.199	2.112	3.209
5	4.969	7.633	4.731	6.921	4.802	7.169
21	10.280	15.065	9.632	13.864	9.962	14.754
63	17.325	24.172	16.456	23.558	17.207	25.157
126	23.659	32.131	22.849	32.284	24.003	34.544
252	31.781	42.030	31.166	42.788	32.995	46.337

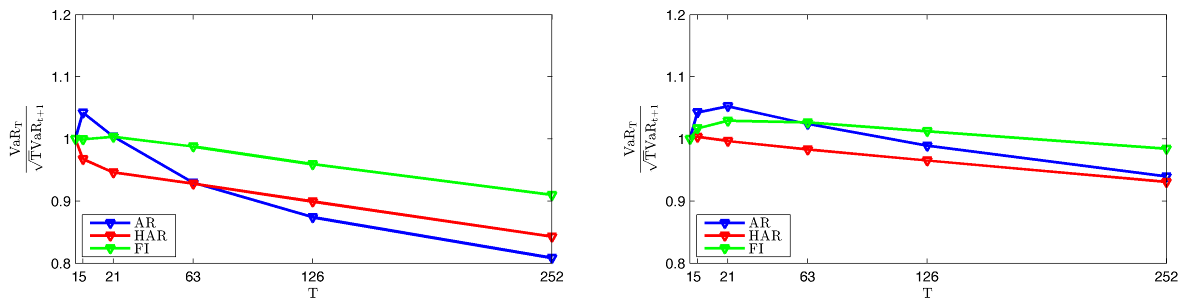
This table shows the VaR at two confidence levels computed for the different models for the least persistent stock. The different rows display the aggregation horizon ranging from one day, first row, to one year, last row.

Figure 7 shows the ratio between VaR obtained using our methodology and the one using the *square root of time rule*: panel A shows the ratio at higher risk level whereas panel B shows the ratio at higher confidence level. Interestingly, for the least persistent stock the *square root of time rule* overestimates VaR at $\alpha = 1\%$. In fact, for the autoregressive model the ratio shows that this scaling law the overestimates VaR of approximately 20% whereas for the fractional the overestimation is of about 10%. For $\alpha = 5\%$, the autoregressive models exhibit that VaR is upward biased using the square root of time of about 5% whereas the fractional model displays a ratio slightly below unity.

Table 12 shows instead the estimates of VaR for the most persistent stock. Compared to the previous case, the next day VaR is smaller because the median realized variance is smaller. From this table it is easy to notice that for annual VaR the fractional model implies a much higher estimate: the difference compared to the other models is of more than 8%. This can easily be explained considering that this specification implies a greater mean of the Integrated Variance distribution and a higher volatility.

Figure 7: Ratio of VaR_T obtained with two different approaches

(a) $\alpha = 1\%$ (b) $\alpha = 5\%$



This graph shows the ratio between VaR_T obtained using our model and different time series specifications for log realized variance and the long term VaR obtained using the *square root of time rule* for the least persistent stock. Moreover, panel a shows the ratio at the confidence level of 1% whereas panel b displays the same ratio at 5%.

From this table we can also notice that the autoregressive exhibits a higher VaR_T than the HAR model for any maturity and both confidence levels. The explanation lies in the values of the first two moments of the Integrated Variance distribution implied by these two models. The HAR model generates a higher volatility than the AR model, however it also generates a much lower mean of the Integrated Variance distribution.

Figure 8 shows the ratio between the VaR obtained using our methodology and the one given by *square root of time rule.* Interestingly, panel A shows that for the most persistent stock at very long term model uncertainty is a relevant issue. The fractional model implies that the above mentioned scaling law underestimates annual VaR at $\alpha = 1\%$ of more than 10% whereas for the other two models such rule overestimates it. Moreover, from panel A it is interesting to see the behaviour of this ratio for the autoregressive model: it spikes reaching almost 20% and then it decays very quickly getting to the HAR model. Panel B shows approximately the same

Table 12: VaR for the different models we consider and for two confidence levels

T	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	1.925	2.941	1.899	2.852	1.906	2.879
5	4.598	7.263	4.254	6.187	4.371	6.585
21	10.137	15.972	8.688	12.506	9.322	14.247
63	17.541	25.758	14.975	21.664	16.674	25.653
126	24.232	33.974	21.232	30.813	24.147	37.108
252	32.633	43.806	29.624	42.582	34.423	51.904

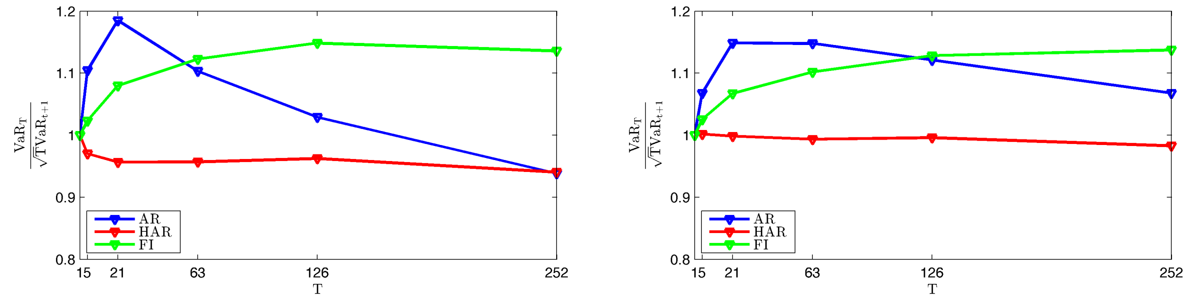
This table shows the VaR at two confidence levels computed for the different models for the most persistent stock. The different rows display the aggregation horizon ranging from one day, first row, to one year, last row.

behavior, the autoregressive model spikes at the monthly horizon touching a ratio of 15% and from there on it starts decaying and showing a monotonically decreasing behavior from there on. The fractional model, instead, shows a monotonically increasing behavior reaching a ratio of approximately 12%.

Table 13 shows the estimates of VaR for the *SPY* when volatility at time t is high. More precisely we set this high value of volatility equal to the seventy-fifth percentile of the unconditional volatility distribution. It is clear from this table that a model with higher persistence implies a greater estimate of VaR. This is due to the effect that the dynamics of these models have on the mean of $p(\mathrm{S}_{\mathrm{t+T}})$. The autoregressive model reverts to the median level much more quickly compared to the other models: there is a small difference when the aggregation horizon is near. On the other hand,

Figure 8: Ratio of VaR_T obtained with two different approaches

(a) $\alpha = 1\%$ (b) $\alpha = 5\%$



This graph shows the ratio between VaR_T obtained using our model and different time series specifications for log realized variance and the long term VaR obtained using the *square root of time rule* for the most persistent stock. Moreover, panel a shows the ratio at the confidence level of 1% whereas panel b displays the same ratio at 5%.

the contrast among models becomes evident when the horizon is above the three months.

Table 13: VaR for the different models we consider and for two confidence levels

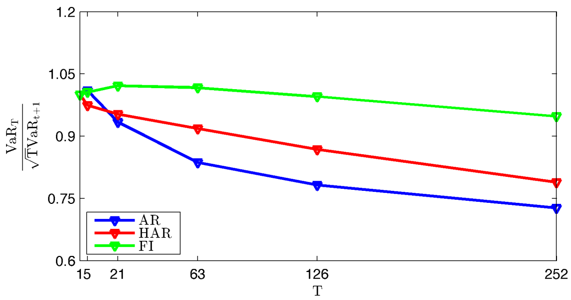
T	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	1.833	2.814	1.883	2.860	1.906	2.896
5	4.048	6.356	4.212	6.232	4.342	6.516
21	7.798	12.045	8.480	12.486	9.042	13.558
63	12.828	18.688	14.102	20.841	15.699	23.376
126	17.555	24.716	18.942	27.856	21.942	32.361
252	23.839	32.471	24.925	35.782	30.086	43.570

This table shows the VaR at two confidence levels computed for the different models for the *SPY* in a high volatility period. The different rows display the aggregation horizon ranging from one day, first row, to one year, last row.

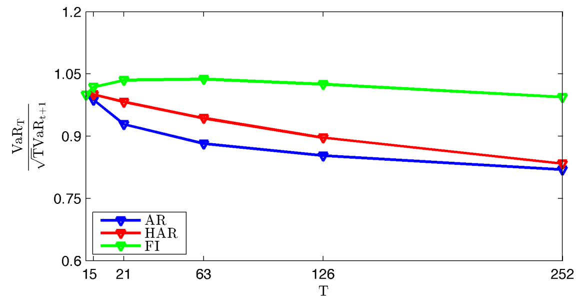
Inspecting the last row of this table also shows and comparing it to table 9 we can notice that the autoregressive model has reached approximately the same estimate of VaR when considering the initial volatility to being at the median level whereas the other two models have not.

Figure 9: Ratio of $\mathrm{VaR_T}$ obtained with two different approaches

(a) $\alpha = 1\%$



(b) $\alpha = 5\%$



This graph shows the ratio between $\mathrm{VaR_T}$ obtained using our model and different time series specifications for log realized variance and the long term VaR obtained using the *square root of time rule* for the SPY when volatility is high. Moreover, panel a shows the ratio at the confidence level of 1% whereas panel b displays the same ratio at 5%.

Figure 9 shows the ratio between VaR obtained using the methodology that we propose and the *square root of time rule*. Both panels show qualitatively the same result: if time t volatility is high this scaling law overestimates VaR. Of course the degree of overestimation is way stronger for a model characterized by short memory such as the autoregressive model.

Table 14 shows the estimates of VaR for the SPY when volatility is low, more precisely we set it equal to the twenty-fifth percentile of the unconditional distribution of volatility. Compared to the previous case we can notice that the reverse happens: the autoregressive model presents the highest estimate of VaR at every horizon. As explained previously this is due to the characteristics of the different models. In fact, when volatility is low the autoregressive model reaches the median level more quickly compared to the other models.

Table 14: VaR for the different models we consider and for two confidence levels

T	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	0.991	1.526	0.941	1.431	0.929	1.414
5	2.610	4.166	2.169	3.232	2.144	3.232
21	6.600	10.341	4.694	6.994	4.574	6.952
63	12.142	17.775	8.912	13.467	8.261	12.600
126	17.119	24.118	13.674	20.696	12.032	18.390
252	23.565	32.147	20.675	30.420	17.578	26.613

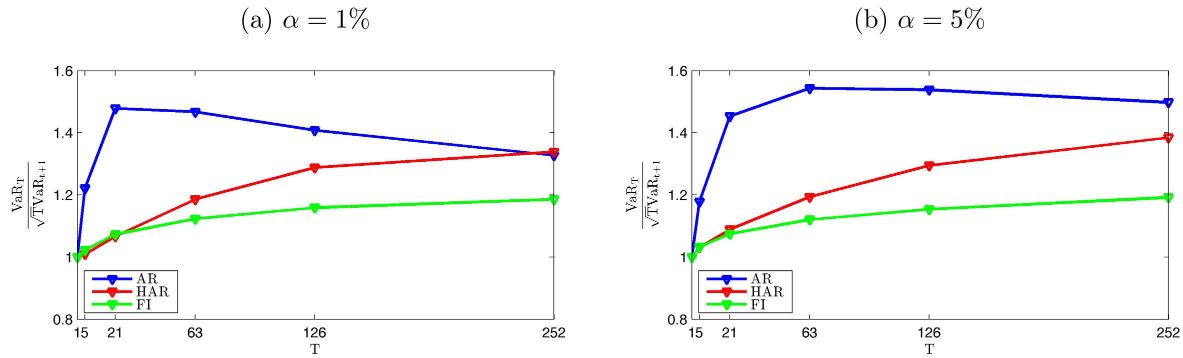
This table shows the VaR at two confidence levels computed for the different models for the SPY in a low volatility period. The different rows display the aggregation horizon ranging from one day, first row, to one year, last row.

The other two models have a slower reversion to the median level and according to the speed to which they revert the estimates of VaR is higher or lower. In fact, the Fractional model is the one that presents the lower estimate of VaR. This is also evident from figure 10, that shows the above mentioned ratio. The panels show two different confidence levels. We can notice that the ratio spikes for the autoregressive model and then it starts decaying whereas for the other models it has a monotonically increasing behaviour. In case of low volatility, using the *square root of time rule* implies that we are underestimating VaR.

We also aim to derive a general scaling law using the framework we suggest to compute long term VaR. We propose to do it using the regression given in equation (17).

This regression model extends the previous one in two ways: first consider the full

Figure 10: Ratio of $\mathrm{VaR_T}$ obtained with two different approaches



This graph shows the ratio between $\mathrm{VaR_T}$ obtained using our model and different time series specifications for log realized variance and the long term VaR obtained using the *square root of time rule* for the *SPY* when volatility is low. Moreover, panel a shows the ratio at the confidence level of 1% whereas panel b displays the same ratio at 5%.

cross-section of stocks in our dataset and it also consider the effects deriving from a different persistence of the time series in our dataset measured by λ_{i}. For the AR model λ_{i} is the autoregressive coefficient, a_1. For the HAR model this is set equal to the largest eigenvalue of the state space matrix F whereas for the fractional model this is set equal to the fractional parameter d.

$$\mathrm{x_{i,T}} = \beta_0 + \beta_1 \ln \mathrm{T} + \beta_2 \lambda_i + \beta_3 \lambda_i \ln \mathrm{T} + u_{\mathrm{i,T}} \tag{17}$$

The regressand is instead given by:

$$\mathrm{x_{i,T}} = \ln \mathrm{VaR_{i,T}} - \frac{1}{2} \ln \mathrm{T} - \ln \mathrm{VaR_{i,t+1}} \tag{18}$$

This is done in order to avoid the issue arising from imposing a common constant for the cross-section. Table 15 shows the results from the above mentioned regression. We are interested in how VaR scales under the different models when T grows. In order to see it we need to consider the total effect measured by the partial derivative of $x_{i,T}$ with respect to the horizon.

Table 15: Estimates of the parameters of Extended Stable Regression

Par.	AR		HAR		FI	
	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
β_0	0.044	-0.329	0.585	1.105	0.344	0.396
	(0.145)	(0.208)	(0.319)	(0.558)	(0.057)	(0.094)
β_1	-0.149	-0.131	-0.222	-0.458	-0.203	-0.296
	(0.043)	(0.063)	(0.128)	(0.216)	(0.025)	(0.038)
β_2	0.080	0.653	-0.576	-1.131	-0.526	-0.589
	(0.177)	(0.256)	(0.332)	(0.581)	(0.095)	(0.156)
β_3	0.159	0.085	0.218	0.453	0.338	0.475
	(0.053)	(0.078)	(0.134)	(0.222)	(0.041)	(0.064)

This table shows the estimates of the coefficient and standard errors of regression in (17) for the different models we consider at two confidence levels.

Let us consider first the autoregressive model for the three stocks given in table 4: at 5% confidence level the implied scaling law for all stocks is slightly below the *square root of time rule*. At 1% the implied scaling law is still lower than the above mentioned rule but the difference is way smaller compared to the previous analized confidence level. Let us now analyze the same derivative for the second mode il our table for the same stocks. At both confidence levels the HAR model implies a VaR scaling which is quite similar to the *square root of time rule*. For example, the total above mentioned derivative for the most persistent stock is -0.009 at 5% and -0.015 at 1%. Finally, for the Fractional model for the least persistent stock the implied scaling law is slightly lower than the above mentioned one, in fact the total effect for this stock is -0.013 and 0.021 for the most persistent stock meaning a faster scaling than square root. At 1%, for the most persistent stock the total effect is 0.018 and for the least persistent the total effect is -0.029.

4 Conclusions

We study the properties of different time series models on Integrated Variance and point out that dynamic specification strongly affects long-term VaR. We point out that the fractional model implies a higher mean of the Integrated Variance for very long maturity and also it generates a higher volatility of volatility.

Moreover, regulatory recommendations suggest the usage of the *square root of time rule*. However, it is well known that this only holds if returns are *iid* normal. In our setting, we show that this rule works approximately well for long term VaR.

Model uncertainty is one of the issues we deal with in this paper: however, the current results can be extended to consider more time series models for realized variance. Furthermore, we also plan to investigate the effects of jumps on long term VaR. These last two issues are high on our agenda.

References

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2001, The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association* 96, 42–55.

Andersen, Torben G, Tim Bollerslev, Francis X Diebold, and Paul Labys, 2003, Modeling and Forecasting Realized Volatility, *Econometrica* 71, 579–625.

Baillie, Richard T, 1996, Long memory processes and Fractionally integrated in econometrics, *Journal of Econometrics* 73, 5–59.

Beltratti, Andrea, and Claudio Morana, 1999, Computing value at risk with high frequency data, *Journal of Empirical Finance* 6, 431–455.

Christoffersen, P., F. Diebold, and T Schuermann, 1998, Horizon problems and extreme events in financial risk management, *Reserve Bank NY Econ.Policy Rev* Policy Rev, 109–118.

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

Danielsson, Jon, and Jean Pierre Zigrand, 2006, On time-scaling of risk and the square-root-of-time rule, *Journal of Banking and Finance* 30, 2701–2713.

Drost, C., and T. F. Nijman, 1993, Temporal Aggregation of Garch Processes, *Econometrica* 61, 909–927.

Luciani, Matteo, and David Veredas, 2015, Estimating and forecasting large panels of volatilities with approximate dynamic factor models, *Journal of Forecasting* 34, 163–176.

Rossi, Eduardo, and Paolo Santucci de Magistris, 2014, Estimation of long memory in integrated variance, *Econometric Reviews* 33, 785–814.

Shimotsu, Katsumi, and Peter C B Phillips, 2006, Local Whittle estimation of fractional integration and some of its variants, *Journal of Econometrics* 130, 209–233.

Wang, Jying N., Jin Huei Yeh, and N. Y P Cheng, 2011, How accurate is the square-root-of-time rule in scaling tail risk: A global study, *Journal of Banking and Finance* 35, 1158–1169.

A Appendix

A.1 Monte Carlo Simulation

A.1.1 AR models

The AR(1) and HAR model can be written in state space form as follows:

$$Y_{t+1} = FY_t + GE_{t+1}$$

where Y_{t+1} is a vector containing h_{t+1} in the first position and its lags until h_{t-p+2} where $p = 22$ is the order of the restricted autoregressive process. Moreover, F is the matrix containing in the first row the autoregressive coefficients: for the AR(1) it is a row containing a non zero element in the first position and zero elsewhere whereas for the HAR model it contains the coefficient estimates of the transformed AR(22) process. The lower $p-1 \times p-1$ block contains an identity matrix and the last column is composed by $p-1$ zeros. Finally G, is a column vector containing a one in the first position and zeros elsewhere. By simple recursion,

$$\begin{aligned}
\mathbb{E}[Y_{t+T}] &= F^T Y_t \\
\mathbb{V}[Y_{t+T}] &= \omega^2 \sum_{j=0}^{T-1} F^j G G' (F^j)'
\end{aligned}$$

A.1.2 FI Model

Recalling the FI model as in equation :

$$(1 - L)^d (h_t - \mu) = \omega \eta_{t+1}$$

Define $y_t = h_t - \mu$. Simulations of y_{t+T} is performed using the recursive structure of the coefficients of the Fractional model:

$$y_{t+T} = \sum_{s=1}^{T} \Delta_s y_{t+T-s} + \omega \eta_{t+T}$$

where

$$\Delta_{s+1} = \Delta_s \frac{s - d}{s + 1}$$

with $\Delta_0 = 1$.