BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

Realized Volatility Modeling of S&P 500 Index Members and the Impact of Temporal Variations in the Mean Levels

Navn:	Fredrik Våland Fusdahl, Fredrik Nordby Johansen

Start:	15.01.2021 09.00
Finish:	01.07.2021 12.00

GRA 1970 - Master Thesis

BI Norwegian Business School

Hand in date: 01.07.2021

Realized Volatility Modeling of S&P 500 Index Members and the Impact of Temporal Variations in the Mean Levels

Supervisor:

Paolo Giordani

By:

Fredrik Nordby Johansen & Fredrik Våland Fusdahl

Abstract

Using a daily panel dataset including almost all the stocks in the S&P 500 dating back to 1985, we document strong similarities in the risk dynamics across stocks. The similarities in risk dynamics are exploited by implementing volatility forecasting models estimated using panel-based methods that aggregate information across stocks and force the coefficients to be the same for each stock. The models that exploit these commonalities in risk characteristics across assets produce highly competitive out-of-sample risk forecasts compared to more conventional individual asset-specific models that implicitly ignore the similarities in risk dynamics. We estimate the models on the daily range of the highest and lowest log intraday stock price, which has been shown to be a good alternative to the high-frequency-based realized volatility (RV) estimator of the integrated volatility. Further, we normalize the RV by the time-varying mean of the RV retrieved from the Kalman Filter and -Smoother as an intermediate step before model estimation. Normalizing the RV before using panel-based estimation methods produces very promising out-of-sample risk forecasts compared to the widely accepted Heterogeneous Autoregressive (HAR) model. Further, it improves the out-of-sample predictive power of the unnormalized models. An important feature of the panel-based models we present is the inclusion of a time-varying mean of each stock. Including this feature mimics introducing an asset-specific intercept for each stock and captures the differences in risk dynamics across assets, as well as the temporal variations in the mean levels.

Acknowledgement

The authors of this thesis would like to thank our supervisor, Paolo Giordani, Department of Finance at BI Norwegian Business School, for his guidance and support throughout this research process.

Contents

1 Introduction

Volatility modeling and -forecasting are of the utmost importance for risk managers, asset managers, derivative traders, regulators, and in general, practitioners affected by outcomes in financial markets. As a result, a large part of the financial econometrics literature has tried to address the challenging task of modeling volatility precisely. For decades the benchmark models of the volatility forecasting literature were ARCH/GARCH type models first proposed by Engel (1982) and Bollerslev (1986). Since then, a long list of competing GARCH- and stochastic volatility models have been proposed in the literature to estimate and forecast volatility in financial markets. These models were estimated on close-to-close squared returns shown to be a poor estimator of daily integrated volatility. In later years, high-frequency data has become more available for many financial assets. With high-frequency returns, one could get much more precise estimates of the integrated volatility than close-to-close returns. However, the latent nature of volatility complicates the implementation of these models and generally does not perform well when estimated directly with intraday data of financial assets. To overcome these shortcomings and effectively exploit the information inherent in high-frequency data, Andersen, Bollerslev, Diebold, and Labys (2003), suggested using reduced form time series forecasting models for the daily volatilities constructed from the intraday sum of squared returns.[1] Set against this background, we present several easy-to-implement volatility forecasting models for S&P 500 index members and draw inspiration from the modeling framework presented by Bollerslev, Hood, Huss, Pedersen (2017). Unlike most of the existing literature, which puts most of its attention on indices, we focus on volatility dynamics at the individual stock level. As an estimator for re-

[1]The use of high-frequency intraday data to precisely measure the true latent integrated volatility was initially presented by Andersen and Bollerslev (1998). This approach has perhaps become the most popular way to model-, measure- and forecasting volatility. See e.g. references in the survey by Andersen, Bollerslev, Christoffersen, and Diebold (2013).

alized volatility (RV), we use the daily range of the intraday highest- and lowest log stock price. However, we will estimate the models on the logarithm of the daily range. Clements, Preve (2019) show that the daily range estimator, in many instances, produces forecasts of similar quality as those of high-frequency-based RV.

Our results are based on a panel dataset including almost all stocks in the S&P 500 index dating back to 1985. The study shows substantial out-of-sample benefits from normalizing the RV by an approximation to the time-varying mean of the RV retrieved from the Kalman Filter and -Smoother before model estimation. Further, we find that including a time-varying "global" mean that captures the strong spillover effects across assets, combined with measures of leverage effects and (correctly) normalizing the RV, greatly enhances the out-of-sample performance compared to the simple benchmark Heterogeneous Autoregressive (HAR) model of Corsi (2009). A large part of the study relates to comparing panel-based regression methods[2] to individual asset-specific estimation techniques. We observe that the class of models that utilize panel-based regression techniques that exploit the strong commonalities in risk dynamics across stocks perform superior out-of-sample compared to more conventional individual asset-specific models for longer forecasting horizons. These results, however, are not as promising for shorter horizons, where the out-of-sample performance of the panel regressions and the individual asset-specific models are more similar.

Further, we investigate the impact of data transformations of the RV to get all the stocks on the same scale before model estimation. The idea behind these data transformations is that we obtain more efficient estimates of the coefficients if we prevent the most volatile stocks from dominating the estimation procedure of the coefficients. The importance of this intermediate step is evident from

[2] The use of panel-based regression techniques to enhance the efficiency of the individual forecast can be compared to Bayesian estimation procedures more generally, as exemplified by Karolyi (1993), who rely on Bayesian shrinkage for improving the forecasts of individual stock return volatilities based on the cross-sectional dispersion.

the out-of-sample performance, which manifests that the (correctly) normalized models do perform better than the unnormalized models.

We start our analysis by calculating the RV each day for every asset. When investigating the distributional characteristics of the stocks RV, we observe strong commonalities in the risk characteristics. For example, ranking stocks by mean levels of volatility and creating quantiles based on this ranking generate unconditional distributions with vast differences in modes. However, when we normalize each asset by its sample mean volatility and look at the same unconditional distributions, very similar patterns emerge, implying that risk dynamics across stocks behave similarly. The distributional similarities across stocks motivate the set of panel-based estimation techniques we employ. Further, we document strong spillover effects across stocks, which are already well-documented in the existing literature (see, for example, Taylor (2005), Andersen, Bollerslev, Christoffersen, and Diebold (2006)). To address the impact of spillovers, we include a ”global” mean that ensures that each stock's forecast moves towards some common time-varying risk factor, in our case a smoothed average of the normalized VIX.

The formulation of our risk models is motivated by the HAR model of Corsi (2009) and can be estimated by standard OLS procedures. In addition, our models draw on important insights from the mixed data sampling (MIDAS) approach of Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2007). First, we implement the most basic HAR model of Corsi (2009). Then we implement more sophisticated models that take leverage effects and spillovers into account. We estimate all models using stock-specific estimation techniques and panel-based regression methods. An important step in the panel regressions is to add a time-varying mean of each stock to account for different risk dynamics across assets. The asset-specific time-varying mean is similar to introducing different intercepts for each stock, which is important due to the differences in mean levels of volatility. Our preferred specification of the models relies on a mixture of ”smooth” Exponential Weighted Moving Average

(EWMA) factors of past realized volatilities. To account for leverage effects and spillovers, we augment the model by including measures of leverage effects and a ”global” risk factor.

In short, our original contributions to the literature are two-folded. First, we take the methods of Corsi (2009) and Bollerslev, Hood, Huss, Pedersen (2017) and apply them to a different and bigger dataset, focusing on individual stocks as opposed to indices and futures. Second, we extend their models by including a better measure of the time-varying mean of the RV of each stock.

The thesis is organized as follows: first, we provide an overview of the related literature in section 3. Then we look at the unconditional distributions and the volatility spillovers in section 4 before moving on to the stylized facts of RV and the models in sections 5 and 6, respectively. Next, in sections 7 and 8 we present the model estimation and forecasting, and a robustness check. Last, you will find the conclusion and references in section 9 and 10, correspondingly.

2 Related Literature

A large part of the financial econometrics literature has focused on modeling and forecasting volatility. An important class of models proposed in the literature is the GARCH-type models such as GARCH, EGARCH, CJR-GARCH models, and others, which handle the daily unobservable hidden conditional variance (Shin, 2018). The growing availability of high-frequency intra-day asset price data sets has given us a more precise proxy of the latent volatility, called realized volatility (RV). The daily realized variance is the sum of squares of high-frequency intraday returns and may be regarded as a precise proxy of the underlying variance, daily integrated variance. Therefore, a new branch of models was developed, focusing on modeling RV directly using more precise proxies of the unobservable latent variance from high-frequency data. These models were labeled RV models and have been shown to produce more precise forecasts than GARCH-type- and Stochastic Volatility Models. Clements, Preve (2019) show that simple alternatives to the squared intraday RV, such as the logarithmic range (LR), often reproduce similar forecasting quality. He finds that replacing the high-frequency RV estimator with daily LR does surprisingly well in an out-of-sample study. In the remaining part of this section, we will provide an overview of the most important related literature employing various models and variations of the integrated volatility proxy to forecast RV.

In the recent literature for RV forecasts, the heterogeneous autoregressive (HAR) model of Corsi (2009) is a central model for which many modifications have appeared. The model proposed by Corsi (2009) is perhaps the most commonly used benchmark in the RV model literature. In the paper, he presents an additive cascade model of volatility components defined over different periods. This volatility cascade leads to a simple AR-type model in the RV, considering various volatility components realized over different time horizons. Despite the simplicity of its structure and the absence of long-memory properties, simulation results show that the HAR-RV model successfully achieves the purpose of reproducing the main empirical features of RV (long memory, fat tails, and self-similarity) in a tractable and parsimonious way.

Corsi and Reno (2009) try to address the well-known asymmetric response to negative returns by a model called L-HAR. The model is an extension of the model provided in Corsi (2009). It is composed of the same RV components and several factors aiming to capture the leverage effects. In the model, asymmetric leverage effects of returns on RV are captured by truncating the daily, weekly, and monthly returns to be negative, i.e., $x_t I(x_t < 0)$. McAleer and Medeiros (2008b) provided similar leverage consideration by introducing a logistic type transformation, $(1+\exp(-\gamma r_t))^{-1}$. Many studies reported that the LHAR models improve the HAR model in real-world RV forecasting, see Asai, McAleer, and Medeiros (2011), Audrino and Knaus (2016), Byun and Kim (2013), Corsi and Reno (2009), Liu and Maheu (2009), Patton and Sheppard (2015), Scharth and Medeiros (2009), and others.

Bollerslev, Patton, and Quaedvlieg (2015) propose a new family of easy-to-implement RV-based forecasting models (labeled HARQ models) and focus on extending the class of Heterogeneous Autoregressive RV models presented by Corsi (2009). The models exploit the asymptotic theory of high-frequency RV estimation over different frequencies to improve the accuracy of the RV forecasts. They find that allowing the parameters to vary explicitly with the (estimated) degree of measurement errors significantly improves the volatility forecasts compared to some of the most popular existing models, implicitly ignoring the temporal variations in the RV measurement error. They accomplish this by weighting each volatility with the variance of the measurement error. Specifically, the models allow for time-varying parameters, which are high when the variance of the RV error is low and adjusted downward when the variance is high and the signal weak. This feature builds on the implication that when the measurement error variance is small (large), the daily RV is a stronger (weaker) signal for the next day's volatility. By explicitly incorporating the time-varying variance of the measurement errors into the model's parameterization, the estimated HARQ models exhibit more persistence in ”normal times” and quicker mean reversion in ”erratic times” compared to the standard HAR model with constant autoregressive parameters. The models are implemented on the S&P 500 equity index and the individual constituents of the Dow Jones Industrial Average.

Based on high-frequency intraday data for more than fifty commodities, currencies, equity indices, and fixed income instruments Bollerslev, Hood, Huss, Pedersen (2017) document strong volatility dependencies across assets and asset classes. The models exploit the asymptotic theory of high-frequency RV estimation over different frequencies to improve the accuracy of the RV forecasts. Further, they find that normalizing by each asset's average level of volatility results in very similar unconditional distributions across assets and asset classes. The authors also document strong volatility spillover effects across assets and asset classes which are addressed by including a ”global risk factor” in the preferred specification of the models presented. These commonalities in risk characteristics and strong spillover effects are then exploited by estimating risk models by aggregating information across assets using panel-based estimation methods. They find that the new risk models, combined with panel-based techniques, results in statistically significant out-of-sample performance compared to more conventional individually estimated asset-specific risk models. The formulation of their risk models is motivated by the heterogeneous autoregressive model (HAR) proposed by Corsi (2009). In addition, it draws on essential insights from the mixed data sampling (MIDAS) approach of Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2007). First, they show how to simultaneously estimate risk models across many assets using panel regression that adds power by exploiting the similarities in the cross-asset risk characteristics. Second, they introduce new ”smooth” RV models. These ”smoothed” models are built such that the forecasted future volatility depends on past volatilities in a way that is continuous and decreasing in the lag lengths.

Clements, Preve (2019) explored several easily implemented extensions of the HAR model. They evaluate these extensions in an out-of-sample framework. Notably, he argues that the stylized facts about RV (such as spikes/outliers, conditional heteroskedasticity, non-Gaussianity) and well-known properties of OLS make standard OLS procedures far from ideal. They suggest using weighted least squares (WLS) and least absolute deviation (LAD) instead of standard OLS to circumvent these weaknesses. The paper finds that a simple WLS scheme can improve the forecasting performance compared to standard OLS in an out-of-sample study covering three major stock markets, especially over longer time horizons. Further, they show that the log daily range instead of the high-frequency based RV produces highly competitive forecasts coupled with quartic root transformation or WLS.

Unfortunately, all of the models presented above suffer from sudden spikes in RV. Several authors have tried to improve the RV forecast by addressing the sudden jumps in the RV. Most commonly, this is done by decomposing the realized variance into continuous and jump components. Andersen, Bollerslev, and

Diebold (2007) pointed out improved forecasting performance if the continuous and jump diffusive parts were separated. Sudden changes in the log price process are usually characterized by a jump-diffusion process instead of the usual geometric Brownian motion with time-varying volatility. Several extensions of the HAR model have been presented to account for the jump-diffusion part. The HAR-CJ model tries to capture the jump part by including a measure of realized jumps in the model (see, for example, Shin (2018) for a more detailed explanation).

3 Data Sources and Cleaning

Our dataset consists of almost all stocks included in the S&P 500 index dating back to 1985. The data source for price data for these stocks is the CRSP U.S stock database. In our model fitting and out-of-sample calculations, we only want to include S&P 500 index membership stocks; that is, we want our dataset only to include observation of a stock for when it was an S&P 500 index member. CRSP provides a unique dataset consisting of historical S&P 500 index constituents and provides us with a unique identifier of each stock, inclusion date, and exclusion date. We perform several transformations and calculations on this price data. However, these calculations are based on the dataset containing all of the data for each stock. Lastly, we remove stocks with less than 5 years of historical data to have robust estimates of these calculations and transformations. Further, we include price data on the S&P 500 index itself and a measure of its volatility. CRSP provides price data for the S&P 500. As a measure of S&P 500 volatility, we use the time series of the VIX index provided by the Chicago Board Options Exchange. The final panel dataset is 6994489 observations and consists of 1227 unique stocks.

4 Unconditional Distributions and Spillovers

Before starting with risk modeling, we want to gain a deeper understanding of the underlying risk dynamics across stocks. A deeper understanding of the distributional characteristics across assets will subsequently allow us to present appropriate risk models in line with the empirical risk characteristics of the stocks. For that reason, we will investigate the distributional- and risk characteristics across assets in the following section.

There are vast differences in the mean levels of volatility across stocks, as shown in Figure 1. The unconditional distributions of the RV also exhibit apparent dissimilarities. In Figure 1, we plot the unconditional distributions of 5 quantiles ranked by each stock's mean level of volatility. In other words, the first quantile is composed of stocks with the lowest mean levels of volatility, and the fifth quantile is formed by stocks with the highest mean levels of volatility. Despite the apparent similarities in the general patterns, we see that the mode varies across quantiles. However, more interesting traits arise when we normalize each stock by its sample mean, that is, $RV_t/mean(RV)$. The distributional similarities across quantiles are evidenced in Figure 2, which displays the same unconditional distributions of the quantiles as in Figure 1, only normalized by each stock's sample mean. Note that the "width" of these distributions are not similar by construction, which would have been the case if we scaled volatility to unit variance and zero means. These commonalities in unconditional distributions motivate the set of panel regressions we employ in this thesis. By estimating the risk models on the whole panel dataset, instead of individual stock-based estimation, we add power by exploiting the strong commonalities in the risk characteristics of the stocks.

Volatility spillover effects in international markets are well known and documented in the related literature (see, for example, Bollerslev, Hood, Huss, Pedersen, (2017)). We also document prominent spillover effects between the indi-

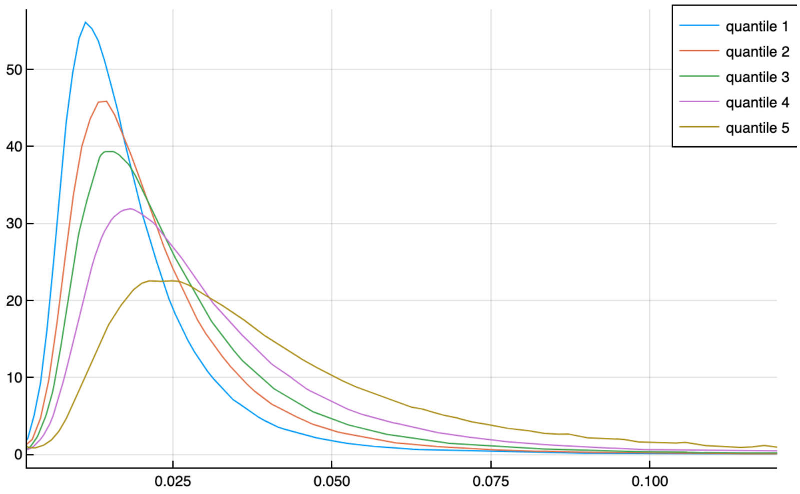


Figure 1: Kernel distributions of realized volatility of quantiles ranked by stocks mean level of volatility

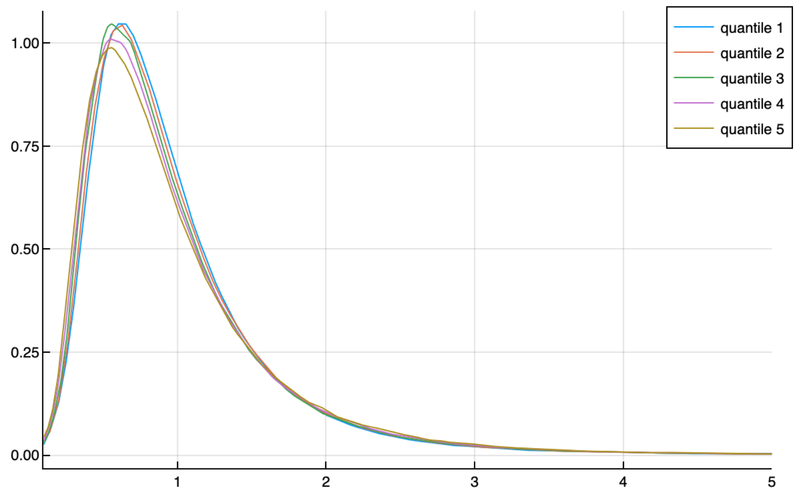


Figure 2: Kernel distribution of the normalized realized volatility of quantiles ranked by stocks mean level of volatility

vidual constituents in the S&P 500. The presence of spillovers is, for example, evident by looking at the improvement of the out-of-sample risk forecasts of the models that include a ”global” mean, which ensures that the forecasts of each stock’s RV move toward a common risk factor. In our case, this is captured by including several smoothed averages of the normalized VIX that has different decay rates to both capture the long- and short-run volatility spillover effects across stocks.

5 Stylized Facts of Realized Volatility

In the academic literature on RV, we find some well-documented and known features of RV.

5.1 Volatility Persistence

RV is highly persistent and displays significant autocorrelations even at very long lags. This property is often ascribed to a long memory data generating process. It is well documented that there exists a strong presence of long-memory in the RV of the S&P 500 (Corsi, 2012). Therefore, a natural question to ask is if the same holds for individual S&P 500 index members. In figure 3, we have plotted the average autocorrelation for each lag across all stocks and the 95% empirical confidence band. It is important to note that we are only interested in the autocorrelation of a given stock when it is included in the S&P 500 index. Hence, the autocorrelation function is applied to a dataset containing stock data for when a given stock is included in the index. Moreover, to protect ourselves against spurious correlations, we exclude stocks with less than ten years of data, i.e., the stock has to be at least a ten-year member of the S&P 500 to be included in the autocorrelation calculation. Figure 3 suggests that we, on average, will see the same strong dependency we see in the S&P 500. Notably, while the confidence bands could indicate that the dependencies in RV for S&P 500 index members

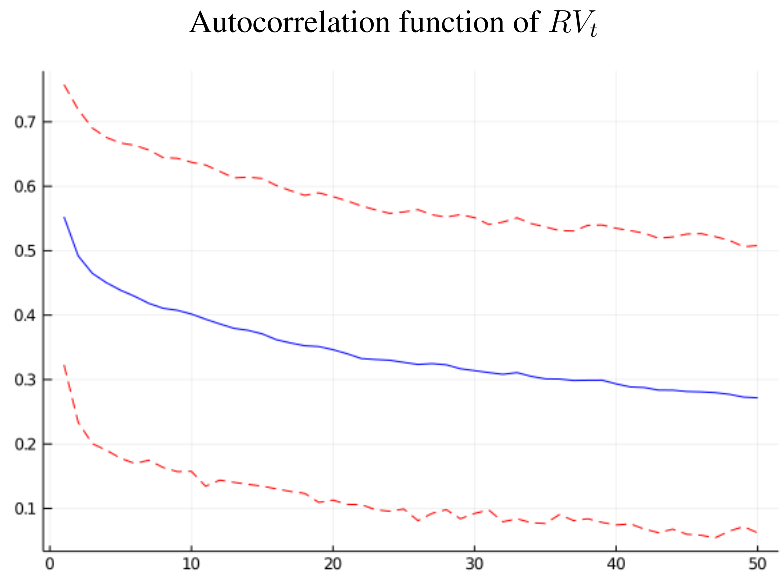


Figure 3: Average autocorrelation for S&P 500 members (BLUE) for each lag. 95% empirical confidence band for each lag (RED)

vary significantly, they should be interpreted with care. Empirical confidence band is simply empirical quantiles, and we cannot say to what extent the true autocorrelations really differ or differ due to sampling error.

5.2 The Leverage Effect

For equity indices, it is a well-established fact that there is an asymmetry in the relationship between returns and volatility. This relationship is also referred to as the leverage effect, which refers to the tendency of an asset's volatility to be negatively correlated with the asset's returns. Typically, rising asset prices are accompanied by declining volatility and vice versa. The term "leverage" refers to one possible economic interpretation of this phenomenon and was first coined by Black (1976): as asset prices decline, companies become mechanically more leveraged since the relative value of their debt rises relative to that of their equity. As a result, it is natural to expect that their stock becomes riskier, hence more volatile (Sahalia, Fang, Li, 2013). In figure 4, 5, 6 we have plotted the function corr(r_t^-, RV_{t+h}), corr(r_t^{5-}, RV_{t+h}) and corr(r_t^{20-}, RV_{t+h}). See our Model

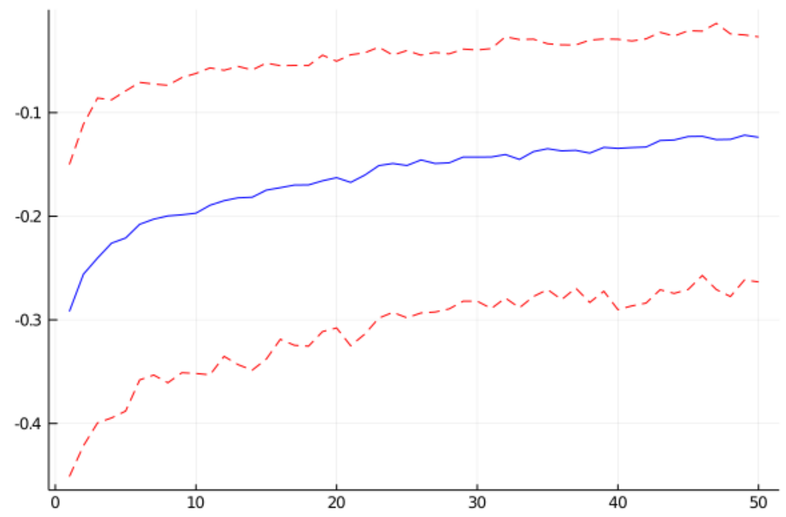


Figure 4: Average Corr(r_t^-, RV_{t+h}) for S&P 500 members (BLUE) for each lag. 95% empirical confidence band for each lag (RED)

section for definitions of r_t^-, r_t^{5-} and r_t^{20-}. The three figures show the leverage effect. The figures indicate that volatility for index members is correlated with lagged negative returns on average. However, as with figure 1, we see a large dispersion in how present the leverage effect is from index member to index member. Moreover, while there seems to be evidence of the leverage effect for S&P 500 members, the effect is not as strong as seen in the S&P 500 itself (Corsi, 2012).

5.3 Jumps in Prices and Spikes in Volatility

An important feature of financial data is the presence of discontinuous jumps. For the S&P 500, Corsi (2012) argues that jumps are not very frequent and, for any practicality, unpredictable, but that they have a strong impact on future volatility. We do not investigate discontinuous jumps and their implication on future volatility for S&P 500 members in our thesis. However, we acknowledge that discontinuous jumps occur much more frequently for S&P 500 members than the S&P 500 index itself. We suspect that understanding the impact of

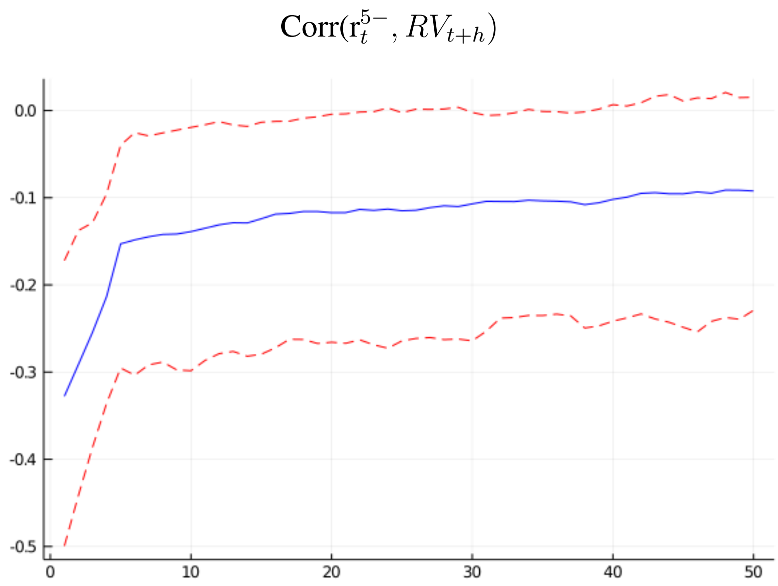


Figure 5: Average Corr(r_t^{5-}, RV_{t+h}) for S&P 500 members (BLUE) for each lag. 95% empirical confidence band for each lag (RED)

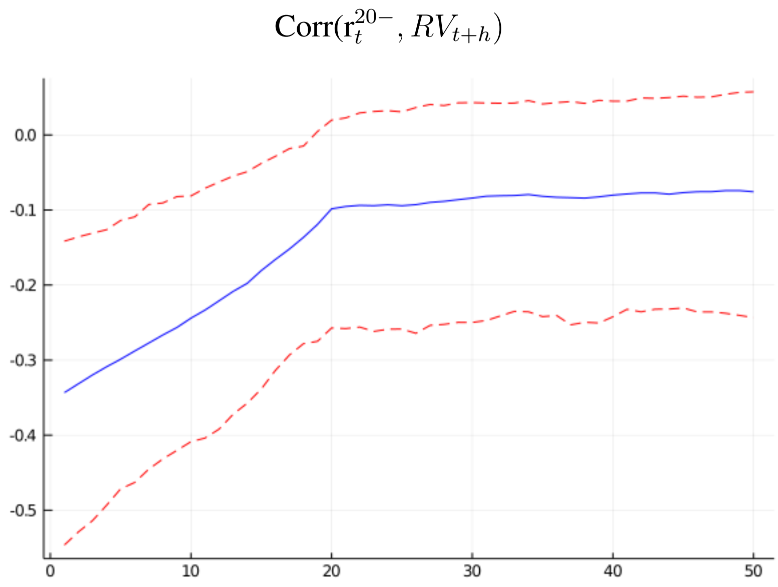


Figure 6: Average Corr(r_t^{20-}, RV_{t+h}) for S&P 500 members (BLUE) for each lag. 95% empirical confidence band for each lag (RED)

these discontinuous jumps on future stock volatility would significantly increase forecasting performance.

5.4 Volatility Cascade and Scaling

Scaling and volatility cascade refers to the stylized fact that volatility over longer periods strongly influences those at shorter periods than conversely (Muller et al. 1997, Arneodo et al. 1998, Lynch 2000, Gencay et al. 2002). Corsi (2009) proposed an economic explanation for the volatility cascade. By noticing that for short-term traders, the level of long-term volatility could matter because it might influence the expected future size of positions and risk. Thus, short-term traders in the financial market react and revise their trading to changes in long-term volatility. However, short-term volatility does not affect long-term traders and strategies, as short-term fluctuation is not paid attention to by these market participants. To capture this, people started modeling volatility using sums of autoregressive processes. Interestingly, simple additive models defined as the sums of autoregressive processes display a decaying memory pattern that can be mistaken for a hyperbolic one. Ganger (1980) shows that the sum of an infinite number of short-term processes can produce long memory. However, an approximated long memory process can be obtained by aggregating only a few heterogeneous time scales. (Corsi, (2012))

6 Models

In this section, we define our assumed data-generating process and our chosen measure of volatility before we specify features of interest and the models investigated in the out-of-sample forecast.

We assume the standard continuous time process

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

where $X(t)$ is the $\log(\text{price}_t)$, $\mu(t)$ is the finite variation process, $W(t)$ is a standard Brownian motion, and $\sigma(t)$ is a stochastic process independent of $W(t)$. For this diffusion process, the integrated volatility associated with day t is the integral of the instantaneous variance over the one-day interval $[t-1d, t]$, where a full-trading day is represented by the time interval 1d (Corsi, 2009),

$$IV_t^{(d)} = \int_{t-id}^{t} \sigma^2(w)dw$$

The integrated volatility is not directly observable and needs to be estimated. There is a large literature on how we can approximate the integrated volatility, and state of the art is to approximate it by the RV defined as the sum of intraday squared returns. The RV over a time interval of one day is defined by

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{N-1} r_{t-j\Delta}^2}$$

where $\Delta = 1d/N$ is the frequency of our return data. $r_{t-j\Delta} = \log(\text{price}_{t-j\Delta}) - \log(\text{price}_{t-j+1\Delta})$. Note that t define the day, while j define the time within day t.

Unfortunately, we do not have access to intraday data for the S&P 500 members, and therefore need to use other estimators for volatility. Range-based volatility estimators were first suggested by Parkinson (1980) and later popularized by Alizadeh, Brandt, and Diebold (2002) for the purpose of estimating stochastic volatility models. In the older literature on GARCH and SV models for volatility forecasting, models were estimated on close-to-close prices. For example, for a simple GARCH(1,1) model, one usually assumed for the data-generated process

$$r_t = c + \sigma_t z_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Where r_{t-1}^2 is the squared close-to-close return. Alizadeh, Brandt, and Diebold showed that log range estimators measure volatility much more efficiently than close-to-close squared returns. Models estimated on range-based volatility estimators resulted in better-performing models. For our analysis, we use the daily range estimator as a proxy for the integrated volatility, one of the simplest range-based estimators. It is defined as

$$RV_t^{(d)} = \log(\text{High}_t) - \log(\text{Low}_t)$$

where High_t and Low_t is the intraday highest and lowest stock price observed, respectively. We will estimate the models on $\log(RV_t^{(d)})$.

6.1 HAR

For a long time, the academia of dynamic modeling of volatility paid most attention to models such as Generalised Autoregressive Conditional Heteroskedasticity models (GARCH) and the Stochastic Volatility models (SV) estimated on close-to-close absolute returns. Both GARCH and SV models could capture stylized facts such as fat-tails and fat-tail crossover [3] while also being straightforward to estimate and have multivariable extensions. However, both GARCH and SV had drawbacks by not capturing and reproducing all of the stylized facts observed in the data. First, they did not reproduce long-term memory. For example, in both GARCH and SV models, volatility shocks decay exponentially, while we see a hyperbolic decay rate in the data. Second, they were not able to capture scaling and volatility cascade.

Recently, high-frequency data has become more effortlessly available for many financial assets. With high-frequency returns, one could get much more precise estimates of volatility than close-to-close returns. Unfortunately, GARCH and SV models did not perform well on high-frequency data. However, by moving to a more precise and higher quality estimate of volatility using high-frequency returns, more complex models could be specified, and sharper results

[3] Fat-tail and fat-tail cross-over refer to the excess of kurtosis of the daily returns, and the empirical cross-over from fat tail too thin tail distributions as the aggregation interval increases, respectively.

could be obtained. Hence, high-frequency data led to a whole new range of rich volatility models. Corsi (2009) proposed the heterogeneous autoregressive (HAR) model that modeled volatility directly by using additive processes with heterogeneous components. The model has proven to be very successful and is today perhaps the most popular benchmark model for RV forecasting (Clements, Preve (2019)). However, it is essential to note that several heterogeneous processes for GARCH and SV had been suggested before Corsi (e.g., fractional GARCH), but the results are clearer when using RV. The HAR model presented by Corsi (2009) is defined as

$$RV_{t+h}^{h} = \alpha + \beta_1 RV_t + \beta_2 RV_t^5 + \beta_3 RV_t^{20} + \epsilon_t$$

Where RV_t^5 and RV_t^{20} is the daily (1 day), weekly (5 days), and monthly (20 days) RV, respectively, and defined as

$$RV_t^h = \frac{1}{h} \sum_{i=1}^{h} RV_{t-i+1}$$

Corsi shows that despite the simplicity of the HAR model, it can capture rich dynamics of both returns and volatility observed in the data, and compared to SV and GARCH, it can reproduce the long-term memory of volatility, and by construction, capture the volatility cascade effect.

For all our models in our thesis, we include a stock-specific time-varying mean μ_t. We further elaborate on this in our estimation section. Thus, the HAR model we focus on in our thesis is given by

$$RV_{t+h}^{h} = \alpha + \beta_1 RV_t + \beta_2 RV_t^5 + \beta_3 RV_t^{20} + \rho\mu_t + \epsilon_t$$

6.2 LHAR

Given the stylized facts presented earlier, several extensions of the original HAR model have been presented. A natural extension is to capture the asymmetric response RV has to positive and negative returns, i.e., the leverage effect. We follow the methodology presented by Corsi and Rose (2012) to extend the heterogeneous structure to leverage effects. Corsi and Rose assume that RV has asymmetric responses to previous daily returns, past 5-day, and 20-day returns. Let r_t be the daily log return. We define daily, weekly, and monthly truncated returns to model the leverage effect at different frequencies.

$$r_t^{h-} = \min(0, r_t^h)$$

where r_t^h is the mean return of h days, i.e $r_t^h = \frac{1}{h}\sum_{i=1}^{h} r_{t-i+1}$. The proposed extended model is called the HAR model with leverage effects (HARL) and given by

$$RV_{t+h}^h = \alpha + \beta_1 RV_t + \beta_2 RV_t^5 + \beta_3 RV_t^{20} + \gamma_1 r_t^- + \gamma_2 r_t^{5-} + \gamma_3 r_t^{20-} + \rho\mu_t + \epsilon_t$$

Note that there are other ways to take into account the leverage effect. The most straightforward and standard way would be to include the simple return r_t^h. In this setting, one would most likely see γ_1, γ_2, and γ_3 being negative; that is, negative returns lead to upward adjustments in the volatility forecast and a downward adjustment for positive returns. However, by using simple returns, you assume symmetric responses in volatility from positive and negative returns. It is well known that volatility tends to increase more after a negative shock than after a positive shock of the same magnitude (see Campbell and L. Hentschel (1992) and Christie (1982)). Therefore, using the truncated return r_t^{h-}, we assume a non-linear behavior in the leverage effect, i.e., an asymmetric behavior to positive and negative returns of the same magnitude.

6.3 Hexp-LG

Motivated by section 6.2, we extend our LHAR model even further by including measures to capture spillover effects. Inspired by Bollerslev, Hood, Huss, Pedersen (2017), we introduce a common global risk factor for S&P 500 index members. A natural choice for a global risk factor is the VIX, which serves as a good approximation for the volatility of the S&P 500.

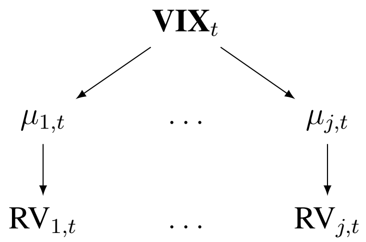
Moreover, from a statistical perspective, it is good practice to include such a variable. First, since we have included a time-varying mean for all the stocks, we should include something in our model with longer memory of the mean volatility across longer cycles. The idea behind this is that it allows us to remedy the loss of information coming from using only the last few years of data to compute the time-varying mean for each stock. By including the VIX and its global average, we can correct this. If the VIX, and implicitly the volatility of S&P 500 has been lower on average the past five years, the five-year average of the S&P 500 index members has also been below the global average, and the volatility should therefore be adjusted upwards, which the estimates will do. Second, it is a simple and cheap way to introduce shrinkage of the estimates by allowing stocks to converge to some global mean. We are doing it in the simplest way possible by a linear model where we will have a single coefficient for all the stocks, implying all the stocks converge at the same rate. To capture differences in convergence speed for stocks, we would need to move to non-linear models, which is outside the scope of this thesis. We could also take a Bayesian view on this. A reasonable assumption in our opinion is that the expected volatility of S&P 500 index members should be regarded as related or connected in some way by the construction of the index. Bayesian hierarchical models are a way to capture such dependencies. If we assume that RV of stock j is a function of the time-varying mean μ_j,

$$RV_{t,j} = \mu_{t,j} + \epsilon_t$$

to force dependency between the expected volatilities of S&P 500 index members, we could give $\mu_{t,j}$ a probabilistic specification in terms of further parameters, known as hyperparameters. For example, in the above model, we could make a further assumption that the time-varying mean of stock j is a function of the global risk factor (VIX),

$$\mu_{t,j} = VIX_t + \epsilon_t$$

The following graph illustrates the hierarchical structure and dependence in the example above.



It is a simplified model, but thinking in hierarchical terms provides us with an understanding of the multiparameter problem and what the inclusion of a global risk factor is trying to capture.

Following the Hexp model proposed by Bollerslev, Hood, Huss, Pedersen (2017), we implement a model that relies on a mixture of ”smooth” Exponential Weighted Moving Averages (EWMA) of the past realized volatilities and the global risk factor. By introducing these smoothed averages, we avoid the step-wise changes inherent in the forecast from the HAR models. The reason for this is that we construct the factors of past realized volatilities in a way that is continuous and decreasing in the lag lengths. Hence, we avoid abrupt changes in the RV estimates when large observations drop out of the sum. We explicitly estimate the relative importance of different EMWA factors constructed from the past daily RV’s,

$$\text{ExpRV}_t^{\text{CoM}(\lambda)} = \sum_{i=1}^{500} \frac{e^{-i\lambda}}{e^{-\lambda} + e^{-2\lambda} + \cdots + e^{-500\lambda}}$$

Where λ is the decay rate of the weights, $\text{CoM}(\lambda)$ is the center-of-mass $\text{CoM}(\lambda) = \frac{\exp(-\lambda)}{1-\exp(-\lambda)}$. The center-of-mass effectively summarizes the effective sample size of lagged realized volatilities that it uses. To calculate the rate of decay corresponding to the effective sample size we compute the inverse of $\text{CoM}(\lambda)$, and is expressed as $\lambda = \log(1 + 1/CoM)$. Consequently, $\lambda = \log(1 + 1/125) = 0.08$

can be thought of as a decay rate corresponding to a center-of-mass equivalent to 125 trading days (see for example Bollerslev, Hood, Huss, Pedersen, (2017)). Similarly to the other models, we focus on spanning the universe of past realized volatilities. To span the universe of past realized volatilities we choose λ's that correspond to a center-of-mass of 1, 5, 25, 125 trading days.

We do a similar transformation of the normalized VIX, where we have normalized by the sample mean. Let ExpVIX_t^h be the moving average computed using the exponential filter above. Bollerslev, Hood, Huss, Pedersen (2017) include only the exponential smoothed global risk factor with $h = 5$. We take it a step further and include the exponential smoothed global risk factor for $h = 25, 125, 500$. The new augmented HAR model is then,

$$RV_{t+h}^h = \alpha + \sum_{h=1,5,125} \beta_h \text{ExpRV}_t^h + \gamma_1 r_t^- + \gamma_2 r_t^{5-} + \gamma_3 r_t^{20-} + \sum_{h=5,125,500} \delta_h \text{ExpVIX}_t^h + \rho\mu_t + \epsilon_t$$

We will refer to this model as Hexp-LG. [4]

7 Model Estimation and Forecasting

7.1 Estimation

In our analysis, RV_t is the realized volatility. The models are estimated on the transformation $log(RV_t)$ to avoid negativity issues and get closer to a Normal distribution. The estimation is straightforward using standard OLS. We present several different implementations. A large part of our study relates to comparing panel-based regression techniques to more conventional individual asset-specific model estimation. The panel-based estimation framework is motivated by the methodology introduced by Bollerslev, Hood, Huss, Pedersen (2017). However,

[4]The pedantic reader will wonder why r_t^{h-} is not smoothed as well. We did estimate the model using a smoothed version of r_t^{h-}. However, this led to worse out-of-sample performance compared to the Hexp-LG model.

we only focus on the individual constituents of the S&P 500, as opposed to cross-asset class estimation. Further, we investigate the impact of different data transformations by normalizing the RV[5] with two variations of a time-varying mean of the RV[6]. We start by estimating the models individually on an asset-specific basis, with stock-specific coefficients. Then we use panel regressions where we force the coefficients to be the same across all assets.

Longer-run forecasts, say over weekly or monthly horizons, may be obtained by iteratively substituting the one-day ahead forecasted RV_t into the right-hand side of the model. Subsequently adding up the one, two, three, etc., one-day ahead forecasts to achieve forecasts for longer horizons. Instead, we use a much simpler approach. To forecast more than one-day ahead we replace RV_{t+1} on the left-hand side of the model with the RV over the forecast horizon h of interest, that is, $RV_t^h = \frac{1}{h}\sum_{i=1}^{h} RV_{t-h+i}$. In the forecasting literature, this is often labeled the direct approach, contrary to iterated forecasts.

We normalize the RV of each stock with an approximation to the time-varying mean of the RV to obtain risk measures that are on the same scale across assets. This will, in turn, prevent the most volatile assets from dominating the estimation procedure. Additionally, normalizing the RV makes sure that the parameters are "scale-free" in the sense that they do not depend on the mean level of volatility. Further, normalizing with a time-varying retrieved from the Kalman Filter and -Smoother, contrary to the sample mean or the expanding mean[7], will most likely lead to more efficient estimates of the coefficients (normalize by a

[5]We say that we normalize the RV with the time-varying mean of the RV. This is just for notational convenience; we normalize the log(RV) by the time-varying mean of the log(RV).

[6]We also perform normalization using the constant sample mean. However, this transformation leads to 25-30% worse performance compared to normalizing with time-varying mean and no normalizing.

[7]Even though an expanding mean is somewhat time-varying, it is not optimal. Our thinking: normalizing by a noisy time-varying expanding mean is similar to introducing measurement errors in the RV. Hence, the coefficients will be estimated less precisely, resulting in worse forecasts.

more precise approximation to the time-varying mean). The underlying assumption for this argument is a suspicion that the mean level of volatility across assets varies over time. Expanding further on this idea, if the mean evolves over time, then normalizing with the simple sample mean or the expanding mean is similar to introducing measurement errors in the RV (normalize with an incorrect- and noisy mean), resulting in worse forecasts. This argument seems reasonable, considering that the out-of-sample performance drops more than 15% when we normalize by the sample mean, even compared to model estimation on the raw RV.

Another important reason we employ the Kalman Filter and -Smoother is the inclusion of a time-varying mean level of stock volatility in our models. Since we include an intercept in our panel-based models, we need something that captures the difference in risk dynamics across stocks. Including an asset-specific time-varying mean will achieve just this. Considering the dispersion in mean levels of volatility across assets, it seems unreasonable to force the intercept to be the same for each stock. For that reason, it is essential to include a time-varying mean in the models that accounts for this feature. The inclusion of a time-varying mean can also be interpreted as introducing a time-varying intercept for each stock that varies in line with the temporal variations of the mean level of volatility. Note that we do not normalize the longer-run time-varying mean.

To derive the time-varying means, consider the univariate time series y_t satisfying,

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_e^2) \tag{1}$$

$$\mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \tag{2}$$

Where $\{\epsilon_t\}$ and $\{\eta_t\}$ are two independent Gaussian white noise series and $t = 1, 2, .., T$. In this case, μ_t is the underlying log volatility of the asset at time t, y_t is $log(RV_t)$, and the ratio $\frac{\sigma_e^2}{\sigma_\eta^2}$ determines the average sample size/decay[8].

[8]As starting value for the state variable, μ_t, we use the sample mean, calculated using data

We want a very persistent or ”slow” moving mean. Consequently, we set the ratio such that the decay rate is equivalent to the effective sample size equal to five years of data (1000 trading days). The true volatility can not be directly observed through time but evolves according to a random walk model. The standard deviation of ϵ_t can be interpreted as the scale used to capture the impact of shocks to y_t. For instance, if $\sigma_\epsilon^2 = 0$, then y_t just follows a random walk, $y_{t+1} = y_t + \eta_t$. The model in Equations 1 and 2 is a special linear Gaussian state-space model. The variable μ_t is called the state of the system at time t and is not directly observable over time. Equation 1 provides the link between the observable data y_t and the state μ_t and is called the observation equation with measurement error ϵ_t. Equation 2 governs the time evolution of the state variable and is the state equation with innovation η_t. The model is also called a local-level model in Durbin and Koopman (2001, Chapter 2), which is a simple case of the structural time series model of Harvey (1993).

The purpose of using the local-level model is to infer the properties of the state μ_t from the data $\{y_t|t = 1, ..., T\}$ and the model. Note, however, that this is an intermediate step. Our main objective is not to infer properties of μ_t in itself. Rather we first compute the time-varying mean μ_t before using it as input in the model for the RV. Most commonly, three methods are used to infer the properties of μ_t. They are filtering, smoothing, and prediction. In our analysis, we only infer the properties of μ_t by utilizing smoothing and filtering. Let $F_t = \{y_1, y_2, ..., y_t\}$ be the available information at time t. In short, filtering means recovering the state variable μ_t given F_t. Smoothing is to estimate μ_t from F_T where $T > t$. Put differently, smoothing exploits the entire dataset, thereby called a two-sided filter, while filtering uses data until time t, and hence called a one-sided filter. To retrieve the time-varying mean given information up until t and T, for the one-sided and two-sided filter, respectively, we make use of the Kalman Filter and -Smoother. Let $\mu_{t|j}$ be $E(\mu_t|F_j)$. Then we want to

until the origin of the forecasting period.

calculate $E(\mu_t|F_T)$ and $E(\mu_t|F_t)$ for the two-sided- and one-sided filter correspondingly. For a full-fledged derivation, see, for example, Tsay (2010, Chapter 11). Our idea is then to estimate the models on the RV that are normalized by an approximation to the time-varying mean that is more precise than the sample mean or the expanding mean. The transformation is an extension of the idea presented by Bollerslev et al. (2017), who ”center” the models by subtracting the expanding sample mean of the past realized volatilities. The simple normalized HAR model then takes the following form:

$$RV_{t+h}^{h} - \mu_{t|j} = \alpha + \beta_1(RV_t - \mu_{t|j}) + \beta_5(RV_t^5 - \mu_{t|j}) + \beta_{20}(RV_t^{20} - \mu_{t|j}) + \rho\mu_{t|j} + \epsilon_t$$

Where $\mu_{t|j}$ is the time-varying mean of the RV at time t, and RV_t^h is log-transformed. For the two-sided filter $j = T$ and $j = t$ for the one-sided filter. Using the approximation to the time-varying mean retrieved from smoothing (two-sided filter), contrary to the expanding mean or the sample mean, we hopefully normalize by a more efficient measure of the time-varying mean of the RV at time t and hence obtain a more precise estimate of the coefficients in our models. The procedure also mimics what a Bayesian model would do in the sense that we exploit all the data available to us[9] and continually update the estimates as time passes and we get access to more information. Again, by normalizing the RV, we ensure that the most volatile assets don't dominate the estimation of the coefficients because we estimate the model on stocks RV that are on the same scale. Hence, we obtain more efficient estimates of the coefficients, which is evident from the out-of-sample forecasting analysis. This manifests that the out-of-sample risk forecasts of the (correctly) normalized models perform superior to the unnormalized models.

There are, of course, weaknesses with our model and estimation procedure. Several shortcomings have been addressed and highlighted in the existing literature. Common issues that are highlighted relate to the known properties of OLS,

[9]The two-sided filter is both forward-looking and backward-looking, and conditions on the entire dataset. In contrast, the one-sided filter only conditions on the data up until time t.

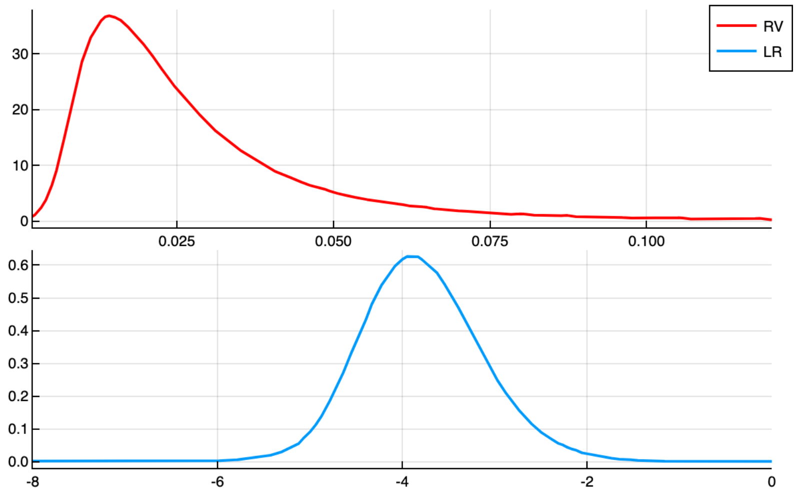


Figure 7: Kernel distributions of realized volatility (red) and log daily range (blue) for all stocks.

such as conditional heteroscedasticity, sensitivity to outliers, and suboptimality in the presence of non-Gaussianity, combined with the stylized facts about RV (spikes/outliers, conditional heteroscedasticity) (see Clements, Preve (2019)). We would, however, point out that violations of these assumptions usually affect the standard errors of the estimated parameters. Hence, these violations only raise problems if we want to make meaningful inferences about the estimated parameters by conducting classical hypothesis tests. Considering that our main objective is to produce precise forecasts of future RV, these problems are not that severe in our case. On the other hand, we acknowledge that OLS is sensitive to outliers and can, in fact, overfit in the presence of outliers. However, by estimating the model on log(RV), we circumvent the issue associated with outliers, as we obtain a distribution with much thinner tails compared to the distribution of the RV. This is clearly evident in Figure 7.

7.2 Out-of-Sample Forecast Results

The purpose of this section is to conduct an empirical analysis of the performance of the models presented in section 2 by doing an out-of-sample forecast analysis. We will focus our discussion on a one-month forecast horizon and perform a 10-year out-of-sample forecast starting from January 2010, with a 1-year expanding window, i.e., and we reestimate our model each year. For our models, estimated using both panel and individual regression, we only predict the volatility of stocks included in the S&P 500 for a given forecasting year. For the individual regression estimation, we use the full sample of each stock to fit the model. The out-of-sample forecast results for S&P 500 index members for a monthly forecast for panel and individual regression are reported in Tables 1 and 2, respectively. The panel regression estimation that restricts the coefficients to be the same across all stocks outperforms the individual regression approach on the monthly forecast horizon. Focusing on Table 1, looking across the table's columns, the results show a clear ranking of the models, with Hexp-LG being the superior model both for the panel-based- and individual model. However, as we show in our robustness section, the gain in performance moving from HAR/HAR-L to Hexp-LG depends on the forecast horizon, with the monthly horizon showing the most substantial increase in performance. Thus, showing that improvements in forecasting power compared to the original HAR model manifest over longer horizons. Looking at the monthly forecast, we obtain close to a 14.5% reduction in RMSE on average by choosing the best performing Hexp-LG (normalized by two-sided mean) model, compared to the benchmark HAR model estimated by individual regression.

When we look across transformations, that is, looking across how we have normalized our data, we find some interesting results. For example, looking at the HAR model, we can reduce the RMSE by normalizing the volatility by the two-sided time-varying mean. As a result, we obtain a reduction of 4.5% for the HAR model by estimating the model on normalized volatility rather than raw volatility. Not presented in the table, there is also significant improvement by including a time-varying mean. Compared to a HAR-model without time-varying mean and estimated on panel data (no normalization), we obtain a reduction in RMSE of 7%. For the Hexp-LG, the drop is similar, with an improvement of 4.6%. Interestingly, comparing one-sided mean and two-sided mean, we see that both approaches present an enhanced performance; however, the two-sided mean shows a slight improvement compared to the one-sided mean.

We want to isolate the effect of including a global mean into the Hexp-LG model. Hence, we run a Hexp-L model without the global risk factors, i.e.

$$RV_{t+h}^{h} = \alpha + \sum_{h=1,5,125} \beta_h \text{ExpRV}_t^h + \gamma_1 r_t^{-} + \gamma_2 r_t^{5-} + \gamma_3 r_t^{20-} + \rho\mu_t + \epsilon_t$$

And compare it to the Hexp-LG. The Hexp-L model produces an RMSE equal to 0.00439. The model is estimated on two-sided normalized volatility and panel regression. Interestingly, by removing the smoothed global risk factor, the out-of-sample performance drops significantly. Thus, showing the importance of including longer-run volatility factors.

From Tables 1 and 2, we see that the out-of-sample performance of the pooled regression models is somewhat better than the individual asset-specific models for the monthly forecasting horizon. The best performing pooled regression model almost reduces the RMSE by 10% compared to the best individual asset-specific model. One explanation for the drop in RMSE can be attributed to two competing properties of statistical learning methods: bias and variance. For a given value x_0, the MSE can always be decomposed into the sum of three fundamental quantities: the variance of $\hat{f}(x_0)$, the squared error bias of $\hat{f}(x_0)$ and the variance of the error term ϵ. We introduce some bias by forcing the coefficients to be the same across all assets. However, by forcing the coefficients to be the same, the resulting forecasting variance will be lower. This is in line with the findings of Bollerslev, Hood, Huss, Pedersen (2017), who show that the average forecast variance is reduced significantly when going from individual asset-specific estimation to panel-based regression methods that exploit the strong commonalities in the risk characteristics across assets.

h	**HAR**	**HAR-L**	**Hexp-LG**
		Unnormalized	
20	0.004 71	0.004 61	0.004 20
		Two-sided mean	
20	0.004 52	0.004 46	0.004 03
		One-sided mean	
20	0.004 59	0.004 51	0.004 09

Table 1: Panel Out-of-Sample Forecast Results

h	**HAR**	**HAR-L**	**Hexp-LG**
		Unnormalized	
20	0.005 05	0.004 94	0.004 66
		Two-sided mean	
20	0.004 79	0.004 71	0.004 35
		One-sided mean	
20	0.004 87	0.004 80	0.004 41

Table 2: Individual Out-of-Sample Forecast Results

8 Robustness

In our thesis, we have decided to focus on future volatility instead of future variance. There is no specific reason for this other than that most investment decisions and risk measurements often depend on future volatility and not future variance. However, we perform a robustness check where we also estimate our models on realized variance. In doing so, the RV and volatility risk factors on the right-hand side in the models are replaced by the variance estimator. The VIX should be transformed to match realized variance. When estimating on variance, using VIX, which is a standard deviation, may not work as well. Hence, we transform the VIX from standard deviation to variance $\text{VIX}^2 = (\text{VIX}/100)^2$. The results obtained by using realized variance instead of volatility confirm our general conclusion. First, the models estimated using panel-based methods outperform the individually estimated risk models in general. Second, our HEXP-LG model normalized with a two-sided filter outperforms all other models.

Further, we look at the daily and weekly forecast horizon. The gain in performance moving from HAR/HAR- L to Hexp-LG depends on the forecast horizon, with the monthly horizon showing the most substantial increase in performance. The gain in forecast performance moving from individually estimated risk models to panel-based methods also depends on the forecast horizon. For the weekly horizon, we obtain less strong but similar results compared to the monthly forecast horizon. However, the individually estimated risk models seem to slightly outperform the panel-based risk models on average for the daily horizon.

We also remove the global risk factor and run the HEXP-L model for daily and weekly horizons. Interestingly we do not see the same drop in performance when the global risk factor is removed on the shorter horizons. The VIX is a measure of implied volatility for the next 30 days for the S&P 500 index. Hence, we think an improvement to our model would be to include implied volatility for the S&P 500 on shorter horizons; this can be done using shorter-dated (such as 1-day and 5-days) options for the S&P 500 and back out the implied volatility.

9 Conclusion

In this thesis, we study the risk dynamics of the individual constituents of the S&P 500 and present several easy-to-implement volatility forecasting models, including some novel ones. Based on a panel dataset including all the stocks in the S&P 500 index dating back to 1985, we observe similar patterns in the risk dynamics across stocks. Furthermore, normalizing the asset's RV by each stock's sample mean generates surprisingly similar unconditional distributions of the normalized risk measures. These commonalities in risk characteristics can be exploited by aggregating information across assets using cross-asset panel regressions that force the coefficients to be the same for each stock. The panel-based models produce very competitive out-of-sample risk forecasts compared to more conventional asset-specific models, especially for longer time horizons. The out-of-sample performance of the panel-based models can be further improved by normalizing each stock's RV (except for the longer-run mean) with an approximation to the time-varying mean retrieved from the Kalman Filter and -Smoother. In addition, we demonstrate the importance of including a stock-specific time-varying mean in the panel regressions that captures the impact of differences in the risk dynamics across stocks and show that incorporating this feature improves the out-of-sample predictive power.

Further, we document strong volatility spillover effects and find that a ”global” risk factor contains much information about future volatilities that the individual past realized volatilities fail to capture. Thus, including this common time-varying risk factor in the models enhances the out-of-sample forecasting performance. Lastly, we implement a set of ”smooth” models that are dependent on past realized volatilities in a way that is continuous and decreasing in the lag lengths. We show that these ”smooth” models perform even better in out-of-sample risk forecasting compared to the other models.

In short, our original contributions to the literature are two-folded. First, we take the methods of Corsi, Bollerslev, Hood, Huss, and Pedersen and apply them to a different and bigger dataset, focusing on individual stocks as opposed to indices and futures. Second, we extend their models by including a better measure of the time-varying mean of the RV.

10 References

- Andersen, T. G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Econometric Review, 39, 885–905

- Andersen, T. G., Bollerslev, T., Diebold, F. X. (2007). Roughing it up including jump components in the measurement, modeling, and forecasting. The Review of Economics and Statistics, 89, 701–720

- Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003a. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

- Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2003b. Micro effects of macro announcements: Real-time price discovery in foreign exchange. American Economic Review 93, 38–62.

- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F. X., 2013. Financial risk measurement for financial risk management. In: Handbook of the Economics of Finance (eds. G. Constantinides, M. Harris, and R. Stulz).

- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F. X., 2006. Volatility and correlation forecasting. In: Handbook of Economic Forecasting (eds. G. Elliott, C.W.J. Granger and A. Timmermann).

- Alizadeh, S., Brandt, M.W., and Diebold, F.X. Range-based estimation of stochastic volatility models. The Journal of Finance, 57(3):1047–1091, 2002.

- Asai, M., McAleer, M., Medeiros, M. C. (2011). Asymmetry and long memory in volatility modeling. Journal of Financial Econometrics, 0, 1–18.

- Audrino, F., Knaus, S., 2016. Lassoing the HAR model: A model selection perspective on realized volatility dynamics. Econometric Review, 35, 1485–1521.

- Alizadeh, S., Brandt, M. and Diebold, F.X. 2002, "Range-Based Estimation of Stochastic Volatility Models," Journal of Finance, 57, 1047-1092.

- Bollerslev, T., Hood, B., Huss, J., Pedersen, L., H., 2018. Risk Everywhere: Modeling and Managing Volatility, Review of Financial Studies, Society for Financial Studies, vol. 31(7), pages 2729-2773.

- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31, 307–327.

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

- Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the American Statistical Association. pp. 171–181.

- Byun, S. J., Kim, J. S. 2013. The information content of risk-neutral skewness for volatility forecasting. Journal of Empirical Finance, 23, 142–161

- Corsi, F., Reno, R. 2012. HAR volatility modeling with heterogeneous leverage and jumps. Working Paper, University of Siena.

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

- Corsi, F., and R. Renò. 2012. "Discrete-time Volatility Forecasting with Persistent Leverage Effect and the Link with Continuous-time Volatility Modeling." Journal of Business Economic Statistics 30: 368–380.

- Campbell, J. Y. and Hentschel, L., 1992. No news is good news: an asymmetric model of changing 20 volatility in stock returns. Journal of Financial Economics, 31:281–318.

- Christie, A, 1982. The stochastic behavior of common stock variances: value, leverage and 26 interest rate effects. Journal of Financial Economics, 10:407–432.

- Clements, A., Preve, D., 2019. A practical guide to harnessing the har volatility model. Available at SSRN 3369484.

- Durbin, J. and Koopman, S. J., 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford, UK.

- Engle, Robert F, 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." Econometrica, vol. 50, no. 4, 1982, pp. 987–1007. JSTOR, www.jstor.org/stable/1912773.

- Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: how to get the most out of returns data sampled at different frequencies. Journal of Econometrics, 131, 59–95.

- Harvey, A. C., Ruiz, E., and Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61: 247–264.

- Karolyi, A. G., 1993. A Bayesian approach to modeling stock return volatility for optin valuation. Journal of Financial and Quantitative Analysis 28, 579–593.

- Liu, C., Maheu, J. M., 2009. Forecasting realized volatility: A Bayesian model-averaging approach. Journal of Applied Econometrics, 24, 709–733.

- McAleer, M., Medeiros, M. C., 2008b. A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. Journal of Econometrics, 147, 104–119.

- Muller, U., Dacorogna, M., Dave, R., Olsen, R., Pictet, O. and J. Weizsacker, 1997. Volatilities of different time resolutions – analyzing the dynamics of market components. Journal of Empirical Finance, 4:213–239, 1997.

- Patton, A. J., Sheppard, K., 2015. Good volatility, Bad volatility: signed jumps and the persistence of volatility. The Review of Economics and Statistics, 97, 683–697.

- Parkinson, M., 1980. The Extreme Value Method for Estimating the Variance of the Rate of Return. The Journal of Business 53(1):61-65.

- Sahalia, A., Fan, J. Li, Y. 2013. The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency. Princeton University, USA.

- Scharth, M., Medeiros, M., C., 2009. Asymmetric effects and long memory in the volatility of Dow Jones stocks. International Journal of Forecasting, 25, 304–327.

- Shin, W., S., 2018. Forecasting realized volatility: A review. Journal of the Korean Statistical Society. Volume 47, Issue 4, December 2018, Pages 395-404.

- Tsay, S., R., 2010. Analysis of Financial Time Series, 3d Edition. Wiley; 3rd edition.