

The Layman's Guide to Volatility Forecasting:

Predicting the Future, One Day at a Time

- Volatility forecasting can work reasonably well—but measuring results is not as easy as it appears.
- Estimation methods have evolved from the 1980s through today as access to more data increased.
- Capturing both intraday and overnight moves is important for proper risk management.
- More sophisticated methods that place more weight on more recent observations tend to outperform.
- Adding high frequency returns can significantly improve forecast accuracy using relatively simple methods.

Maintaining an accurate view on volatility is a critical task in investing. Keeping tabs on market gyrations is not a perfect measure of risk—we care more about volatility on the downside than up. But volatility is a useful gauge for judging how much pain an investor is willing to tolerate or how assets can be combined in ways to diversify a portfolio to maximize risk-adjusted returns.

Although compounding returns over time may be the goal, market volatility can strike quickly in a matter of days with even the most seasoned investors jarred by sudden shocks to their portfolio. Having an effective means of estimating day-to-day risk can help better prepare for volatility when it does arise and act smarter in response.

Measuring, forecasting, and interpreting volatility is another matter. There are very smart people with advanced degrees and training that specialize in the

modeling of volatility and how it gets used in pricing derivatives or managing risk. But there are some basic principles and even advanced techniques that can be used by a wider range of market participants that can help improve their knowledge.

The purpose of this guide is two-fold. The first is to clearly explain some of the foundational principles and different methods of forecasting volatility for practical use. While we attempt to keep the mix of Greek letters and inscrutable squiggles down to a minimum, the target audience is the generally informed, sophisticated investor or market professional.¹ The goal is not to “dumb it down”—only make it more accessible. We take the time to explain key elements and provide additional detail in the footnotes that might be second nature to a quant but could serve as a refresher to those with limited day to day exposure to some of these concepts and methods.²

¹ We count ourselves in this category. While some of us have math degrees and/or training in data science, our approach remains grounded in our practical experience with institutional investors, technology, exchanges, and algorithmic trading. We value the contributions of academics and quants to the process and strive to keep learning from them. But many quantitative studies are unnecessarily dense, poorly written, and riddled with encoded tribal language designed to weed out those without the same training. This is no different than any other discipline only the math and concepts involved much more complicated. Our objective is to share what we have learned wading through the literature and present it to wider, but informed, audience.

² Let's be honest. We have all Googled things we didn't know or forgot that would cause embarrassment to admit to a colleague.

The second is to highlight the efficacy of using higher frequency data to help improve accuracy and responsiveness in forecasting volatility. Although this uses many more data points in the process, even relatively simple methods leveraging high frequency data tend to perform as well or better than more complicated modeling techniques. We aim to demonstrate how powerful forecasting tools can be constructed using this more granular data along with basic math at the undergraduate level.

We begin with an overview of volatility's unusual characteristics followed by a brief survey of different measures and techniques that have been developed over the years. Then we compare how accurately these different measures can predict the volatility in the S&P 500 over a very short period—just one day. We conclude by showing how choices of forecasting tools and benchmarks can impact the ranking of these different models for use in practice.

Characteristics of Volatility

Volatility (and its squared cousin, variance, which we will refer to interchangeably throughout) has some unusual properties.

Volatility is not directly observable, even after the fact.

We can observe prices and changes that are the result of volatility, but there is no single value that equates to a “true” value. As a result, we can only estimate volatility over specific periods of time.³ The inability to precisely measure its true value presents challenges in judging the accuracy of an estimate, even when using multiple benchmarks.

Volatility levels are always changing.

It is apparent to any market watcher that volatility is not constant, adding more complexity to the process of modeling its impact. One of the weaknesses of the well-known Black-Scholes model for pricing options is its assumption of constant or deterministic volatility. Like many economic assumptions, this was done for modeling simplicity rather than any real-world expectations. Many modern techniques assume volatility is a stochastic or time-varying process (a random walk in a time series), which is more realistic but also more difficult to model.

Estimating volatility has both continuous and discrete elements.

Since most assets have set trading sessions with nights, holidays, and weekends, there are both continuous and discrete components that must be modeled somewhat separately and can be tricky to reconcile. Theoretically, volatility is a continuous process even though there may not be prices available to help estimate its level. But that will not be reflected until price forms at the start of the next session, creating a gap or “jump” that must be incorporated into the forecast. Focusing on one component at the expense of the other can lead to mismatches between theory and practice while introducing significant estimation errors.

Volatility tends to be highly autoregressive.

Fortunately for market practitioners, volatility experienced over the recent past can be a useful predictor of volatility in the future. In varying over time, volatility also tends to cluster around events and then takes some time to decay back to some normalized level. These properties lend themselves to some reasonably effective methods of forecasting in comparison to predicting returns, which can be much more difficult.

Due to its autoregressive properties, even a simple 10-day historical standard deviation of returns can be a useful forecast of tomorrow's volatility. If annualized volatility over this 10-day span is rather high at 32%, it is very likely tomorrow's volatility will continue to be high.⁴ But what are the inputs used to calculate this measure? Like many estimates and forecasts, 10-day historical volatility only looks at the daily, close-to-close price returns. But what about the rest of the day?

It is also highly likely that large variations in the close-to-close price indicating higher volatility will be mirrored in the moves intraday as well. But the collective swings in those relatively small daily returns could be an indication of even bigger day to day swings in the days to come.

One of the ways to capture some of this information is by incorporating the daily range (the differences between the open, high, low, and close for the day), which is explored in several of the methods in the next section. But we also delve into using intraday returns themselves, capturing the variation directly in a more continuous process that is closer to the theoretical ideal.

³ This is known as a latent variable, which must be inferred from other observable or measurable variables. With volatility, we use the known values of point to point returns to help estimate variance.

⁴ A 32% annualized volatility equates to an expected daily move of 2% per day. Using the “Rule of 16”, an annualized figure can be converted into an estimated daily price swing and vice versa (1% daily moves would equate to 16% annualized volatility). To directly annualize a standard deviation, the result is multiplied by the square root of time. For daily returns, the square root of the average number of trading days in the US $\sqrt{252} \approx 16$ becomes the scaling factor.

Tools of the Trade

Including all the methods of estimating and forecasting volatility is well beyond the scope of this analysis. But a core range of techniques can be generally organized by their methods, sampling frequencies, and inputs ranging from the very simple to highly sophisticated.

The academic literature often examines estimators more as proxies for the past unobserved latent variance, but many analyze the same tools for use in forecasting. They are highly related but not identical concepts. For our purposes, we focus on how well these measures can predict future short-term volatility by lagging them one day, which obviously needs appropriate proxies to measure results.

| SIMPLE | | WEIGHTED | | HIGH FREQUENCY | |
|-----------------------------|--|--------------|--------------|----------------|-----------------------|
| Historical | Range-Based | Exponential | Model-Based | Realized | "Reduced Form" Model |
| Close Prices | OHLC Prices | Close Prices | Close Prices | Intraday Bars | Close/Intraday Prices |
| 10D HV 22D HV 100D HV | Parkinson Garman-Klass Rogers-Satchell Yang-Zhang | EWMA | GARCH (1,1) | RV | HAR-RV |

Historical

A basic standard deviation of daily returns over the recent past is a logical starting point to begin measuring volatility. Even those with very sophisticated volatility models will find historical volatility a useful datapoint for comparison. Common lookback periods include 10 days (two weeks accounting for weekends), one month (22 trading days or 30 calendar days), or out a few months or longer (100 or even 252 days).⁵

Historical volatility may be simple but there are still some assumptions and modifications to be made. For one, we typically use log returns for many volatility calculations, as they are continuously compounded, instead of simple returns.⁶

$$r_t = \ln\left(\frac{\text{Close Price}_t + \text{Ordinary Dividend}_t}{\text{Close Price}_{t-1}}\right)$$

⁵ We tend to focus on annualized volatility from daily returns even over periods that go well beyond a year. It is certainly possible to use weekly or monthly returns that may be suitable for longer term analysis. But in general, daily volatility computed over the prior 504 days (2 years) will be higher than weekly volatility over the prior 104 weeks (the same two years). At lower sampling frequencies, the distribution gets closer to normal (Gaussian), with the "fat tails" (kurtosis) becoming less pronounced.

⁶ At higher frequencies of sampling (daily), there is not much difference between simple and logarithmic returns. But log returns are computationally easier to deal with as you can simply add them to get a compounded end of period return. There are also some quirky mathematical modeling benefits that come with using log returns that are explained in detail (but plain English) here: <https://tinyurl.com/2mu2a3zp>.

⁷ While we use variance and volatility interchangeable, we present all formulas in the paper in their squared form as a daily variance. To convert to an annualized daily volatility, simply take the square root and multiply by $\sqrt{252}$. Some will point out that multiplying by the square root of time is an imperfect method and tends to either under- or over-estimate volatility—and they are right—but everyone does it anyway. For additional details, see "Converting 1-Day Volatility to h-Day Volatility: Scaling by [Square Root of Time] is Worse than You Think".

We also make some small changes to the standard variance formula⁷:

$$\sigma_t^2 = \frac{1}{n-1} \sum_{t=1}^n (r_t)^2$$

A standard sample variance is calculated by the sum of the squares of the differences between the observations and the sample mean divided by the number of observations minus 1. However, there is one modification we make that differs from standard variance in statistics: we assume the average daily log return (the "drift") will be zero. This reduces it to the sum of the squares of the daily log returns divided by the number of observations minus one (as shown in the formula above). When looking at relatively short lookback periods such as 10 days, it

is possible for average returns to be so extraordinarily high or low, they become unrealistic to annualize and significantly distort the volatility calculation. As a result, it is often simpler to assume zero drift when computing historical volatility.

Easy to calculate and understand, historical volatility does have its limitations. For one, it is entirely backward-looking and makes no attempt to make any adjustments to forecast future volatility. It treats every single day in the time series equally, giving more distant data points the same ability to influence the outcome even though conditions may have since changed. It also forces the selection of a lookback period, which can be an arbitrary decision with outsized impact. **Lastly, the use of just close-to-close price information ignores intraday movements that may provide a more complete picture.**

Range-Based

Another class of estimators attempts to address the limitation of only using closing prices by incorporating the daily range into the calculation. Introduced over time from about 1980 to 2000, the methods below build on each other, adding elements that were missing from prior versions to capture more real-world dynamics and improve accuracy.

Parkinson

The earliest range-based estimator, Parkinson uses the range defined by the difference between the low and high of the day instead of closing prices.⁸

$$\sigma_{PARK}^2 = \frac{1}{4\ln(2)} \sum_{t=1}^n \ln\left(\frac{High_t}{Low_t}\right)^2$$

One key limitation is the assumption of continuous trading, as no information from the overnight gap (the prior close to next day's open or "jump" in academic parlance) is included, which tends to underestimate volatility. Approximately one-fifth of the total daily price variation occurs overnight—a significant portion to be omitted in the estimate.⁹

Garman-Klass

Extending Parkinson, the Garman-Klass (GK) estimator adds two more range-based measures, the opening and

closing prices. If the opening price is unavailable, the model allows for the prior day's close to be used in place.

$$\sigma_{GK}^2 = \sum_{t=1}^n \frac{1}{2} \ln\left(\frac{High_t}{Low_t}\right)^2 - (2\ln(2) - 1) \ln\left(\frac{Close_t}{Open_t}\right)^2$$

But absent that condition it normally ignores the overnight return so suffers from the same systematic underestimation error as Parkinson.

Rogers-Satchell

Unlike the previous estimators, Rogers-Satchell assumes a non-zero mean return with the same OHLC data points to construct a more sophisticated measure.

$$\sigma_{RS}^2 = \frac{1}{n} \sum_{t=1}^n \ln\left(\frac{High_t}{Close_t}\right) \ln\left(\frac{High_t}{Open_t}\right) + \ln\left(\frac{Low_t}{Close_t}\right) \ln\left(\frac{Low_t}{Open_t}\right)$$

Despite the addition of drift to the model, it still ignores any of the information in the overnight jump, leading this measure to also under-estimate full day volatility.

Yang-Zhang

Finally, Yang-Zhang brings together the elements of the prior range-based models and attempts to address all their shortcomings.¹⁰ It combines overnight volatility with a weighted average of open to close and Rogers-Satchell volatility.

$$\sigma_{YZ}^2 = \sigma_{ON}^2 + k\sigma_{OC}^2 + (1 - k)\sigma_{RS}^2$$

Where:

$$k = \frac{0.34}{1.34 + \frac{n+1}{n-1}}$$

Overnight Volatility

$$\sigma_{ON}^2 = \frac{1}{n-1} \sum_{t=1}^n \left(\ln\left(\frac{Open_t}{Close_{t-1}}\right) - \overline{\ln\left(\frac{Open_t}{Close_{t-1}}\right)} \right)^2$$

Open to Close Volatility

$$\sigma_{OC}^2 = \frac{1}{n-1} \sum_{t=1}^n \left(\ln\left(\frac{Close_t}{Open_t}\right) - \overline{\ln\left(\frac{Close_t}{Open_t}\right)} \right)^2$$

The measure incorporates the overnight jumps while capturing elements of the trend (assumes non-zero mean return), creating a more robust measure than the other

⁸ The squared range is divided by a constant ($4\ln(2) \approx 2.77$) to adjust it to be similar in magnitude to a statistical variance sampled over a number of days.

⁹ Hansen and Lunde (2005).

¹⁰ There is another variant called Garman-Klass-Yang-Zhang that we omit for brevity that just adds an overnight return to Garman-Klass.

range-based estimators.¹¹ We address additional aspects of weighting overnight versus intraday volatility in a subsequent section on measures using high frequency, intraday returns instead of daily OHLC.

Exponential Weighting

While the Yang-Zhang model manages to address many of the weaknesses in the earlier range-based models, it still has a significant limitation: all the data points in the historical lookback are weighted equally. Given the tendency for volatility to spike and then take some time to fade, **placing more weight on more recent observations** yet accommodating the residual effects of an event that is still lingering in the market's memory **can help better model real-world conditions**. Without some smoothing of the data, a volatility event will abruptly drop off from a moving window of historical data.

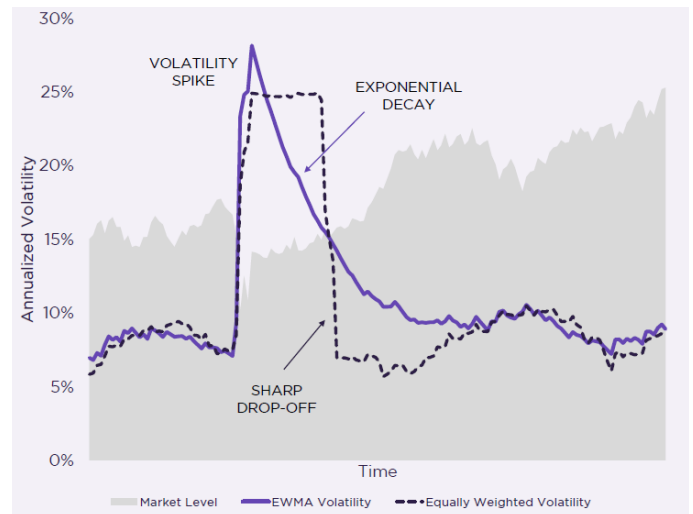
One approach is to use exponentially weighted moving average (EWMA) volatility, which applies a decay factor to prior observations to smooth out the series.¹²

$$\sigma_{EWMA}^2 = (1 - \lambda) \sum_{t=1}^n \lambda^{t-1} r_t^2$$

The decay (represented by lambda, λ) must be a value less than one, typically set to 0.94 for daily data but can be adjusted a little higher (for slower decay) or lower (for faster decay). Using the default value of 0.94, starting at index 1 in the formula above results in a weight of $(1-0.94) * (0.94)^0 = 6.00\%$ for the most recent day followed by $(1-0.94) * (0.94)^1 = 5.64\%$ for day prior, $(1-0.94) * (0.94)^2 = 5.30\%$ for the third prior day and so on.¹³

The ease of computation and performance of the EWMA estimator makes it a popular choice among practitioners. It is also used for index products such as structured notes and fixed index annuities that seek to target volatility levels with daily rebalancing.¹⁴

Figure 1 - Exponential Smoothing Following a Spike



Despite its utility and simplicity as workhorse, EWMA also has its limitations. For one, it only uses daily returns, leaving it susceptible to shocks potentially concealed in large intraday swings that may be a precursor to a larger volatility event. Secondly, it forces you to somewhat arbitrarily select a lambda that can significantly impact the outcome (too fast or too slow). However, using a combination of two or more lambdas can help mitigate this, especially with simple rules that use the higher of the two values (a fast and a slow) to be on the conservative side.

GARCH

GARCH stands for Generalized Autoregressive Conditional Heteroskedasticity, and is a parameterized econometric modeling technique that can be used to estimate volatility. The key innovation with this model begins with the last part of the name—heteroskedasticity. In standard linear regression, the assumption is that the error term—the noise that will cause the dependent variable to deviate from its relationship with the independent one—will have a constant variance (homoscedasticity). With GARCH, the assumption is the variance of the error term will be time-varying but is also conditional on the overall level of variance. Cutting through a lot of the jargon, a GARCH process is well-matched to the clustering behavior and persistence of volatility following certain events in the

¹¹ The 1.34 in the formula is a constant α (the numerator is $\alpha - 1$) that is intended to minimize variation in the estimator itself. This value is the default in the original Yang-Zhang paper.

¹² Exponential smoothing was popularized in the financial industry by JP Morgan's RiskMetrics in the mid-1990s as a part of the widely adopted Value at Risk (VaR) framework. The formula shown above is a little easier to grasp conceptually but EWMA is also written in a reduced, recursive form that is easier for computation: $\sigma_{EWMA}^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_t^2$.

¹³ The lookback period for an exponentially weighted moving average is theoretically infinity as the weight will decay asymptotically into smaller numbers that become meaningless to the output. For the default value of 0.94, the cumulative weight remaining drops below 1% after about 75 days.

¹⁴ Known as "risk-controlled" strategies, these products use a combination of assets and a risk-free asset such as T-Bills to dynamically target a specific volatility level. For example, the index may consist of the S&P 500 and cash and seek to target a 5% volatility level. If the estimator predicts the S&P 500 will have 20% volatility tomorrow, the strategy holds 25% in stocks and the rest in cash ($25\% = 5\% \text{ target} / 20\% \text{ estimate}$), rebalancing daily as volatility changes.

market. But with these more realistic assumptions come more complicated techniques to manage those conditions.

GARCH was developed in 1986 by Tim Bollerslev following the work of Robert Engle who introduced the “ARCH” part in 1982.¹⁵ The most commonly used GARCH (1,1) model can be specified as follows¹⁶:

$$\sigma_{GARCH}^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

There is a lot to unpack in this relatively simple equation, but we begin with a plain English description of what the model is doing. Paraphrasing Engle, GARCH is a weighted average of the long run variance, the predicted variance, and the variance of the most recent squared residual, which correspond to the terms in the equation above in that order.¹⁷

To fit the model, the constants ω , α and β must be estimated and have some constraints.¹⁸ Getting these parameters involves using maximum likelihood estimation, a statistical technique designed to find the values with the highest probability that the process in the model produced the observed data.¹⁹ Once fit, a variety of tests for significance are performed to confirm the model is correctly specified.²⁰ Once specified, GARCH produces an entire time series of variance forecasts, easily updating with just the prior prediction and residual.

While somewhat complex, GARCH is a very powerful tool and tends to perform well in forecasting. While not immediately apparent from the description of the process, the GARCH formula places more weight on the more recent data, similar to the previously described exponentially weighted volatility. In fact, EWMA is

a special case of GARCH (1,1)—and GARCH (1,1) is a generalized form of EWMA.²¹ What EWMA lacks is the “long memory” in GARCH that assumes reversion to some long-term mean level of variance over time. GARCH is not computationally difficult and open-source tools in Python and other platforms make it easier to manage with access to price data and some programming skill. It is widely used for tasks such as managing portfolio risk or pricing derivatives. However, it is not without drawbacks. One is that the basic GARCH (1,1) assumes that positive or negative shocks have an equal impact on volatility when in reality there is a strong negative relationship between returns and volatility, commonly referred to as the leverage effect.²² There have been a whole family of GARCH-based variants (EGARCH, FIGARCH, GJR-GARCH, and others) developed to address asymmetry and add other features to the process, but they come at a cost of higher complexity and difficulty in fitting the model. More importantly, **most GARCH models still generally use only daily closing prices, leaving them vulnerable to more sudden volatility shocks** that could potentially be detected with higher-frequency intraday returns.

Realized Volatility

The use of high-frequency returns in modeling volatility has increased with the wider availability of intraday data by both practitioners and academics alike over the last 20 years. Known as realized volatility (“RV”)²³, it traces its intellectual roots back to a paper by Robert Merton from 1980, who concluded that volatility could be more precisely estimated from realized returns that improve

¹⁵ The “generalized” extension of the model is the third term in the equation that helps the model adapt by using its past prediction errors. Bollerslev was Engle’s research assistant at the University of California San Diego and wrote the GARCH paper as a doctoral student. He currently teaches at Duke University and continues his work on volatility forecasting, lately more focused on realized volatility techniques using intraday returns.

¹⁶ The “1,1” follows the general specification of the model (p,q) which defines p lags in the prior variances and q lags in the residual errors. In addition to being a representation of the most common model, the GARCH (1,1) notation is cleaner to show for our purposes.

¹⁷ Googling “GARCH” and reading through many of the explanations yields some dreadful results. There are some very basic descriptions with no real details and then a lot of very dense academic references but not much in the middle. However, Engle’s GARCH 101 intro (available at <https://www.stern.nyu.edu/rengle/GARCH101.PDF>) is an accessible read that is clearly targeted towards his students (he is currently a professor at NYU’s Stern School of Business).

¹⁸ The parameters must sum to one but the ω term is actually the long run variance multiplied by another weight (gamma). All parameters must also be greater than zero and alpha and beta must sum to a value less than 1.0.

¹⁹ This is commonly done using statistical packages or Python or R libraries. For additional information, check out this very clear and concise explainer available at [Probability concepts explained: Maximum likelihood estimation | by Jonny Brooks-Bartlett | Towards Data Science](#).

²⁰ As with maximum likelihood, GARCH can be a lot simpler to specify and use with software packages and Python/R libraries that can include a number of these tests for significance in the package.

²¹ To get EWMA from GARCH, the long run variance term is set to 0 and the rest reduces to two weights that equal 1.0 (lambda in EWMA, alpha and beta in GARCH), each multiplied by the identical variance and squared return terms in the recursive form of the EWMA equation and GARCH above (see footnote 12 for more details).

²² Black, F. (1976) Studies in Stock Price Volatility Changes of the Nominal Excess Return on Stocks. Proceedings of the American Statistical Association, Business and Economics Statistics Section, 177-181.

²³ This is no different than the commonly accepted realized volatility, in which we are referring to the variability of returns that has been historically realized at a daily or lower frequency. Intraday realized volatility is simply a compressed version of this, measuring the variability within a single day via smaller (potentially tick-by-tick) intervals.

with higher sampling frequencies.²⁴ The realized variance for a single day is simply the sum of squared intraday returns:

$$\sigma_{RV}^2 = \sum_{i=1}^M r_{t,i}^2$$

The sampling frequency used generally ranges from 5 to 30 minutes, with the “M” in the formula above corresponding to the number of bars that constitute a full day (78 for 5-minute bars in a 6.5-hour US trading day from 9:30AM-4PM).²⁵ As with some of the earlier historical estimators, there is no “drift” or mean return subtracted from each squared return as the assumed return from intraday returns is even closer to zero than daily.

As a forecasting tool, RV suffers from some of the same limitations as some of the earlier range-based estimators. Without the overnight return, an RV-only estimator will tend to underestimate the total variation that is important to a practitioner. The overnight gap plus the variability during the day are both metrics that matter for a position or hedge and would be measured by a real-time risk management system.

In the next section, we focus on RV as a building block of more sophisticated forecasting techniques. But for completeness, we also evaluate its efficacy as a forecast tool itself alongside the other estimators.

HAR-RV

Introduced by Fulvio Corsi, the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model has a relatively simple structure featuring three main drivers.²⁶ The first is the use of daily RV—the same measure defined

above that sums the squares of intraday returns for the day at a fixed interval such as 5 minutes. This adds the rich information available from higher frequency data, sampled at a low enough frequency to avoid distortion from market microstructure noise. Secondly, it uses a cascading series of lagged terms that are designed to capture the actions of market participants with differing time horizons in their investment process.²⁷ Lastly, it uses simple linear regression (OLS) to estimate its three coefficients, capturing some longer-term historical patterns across multiple market regimes. This component mimics, but does not completely replicate, the “long memory” feature of GARCH.

The model in its simplest form can be specified as follows:

$$\sigma_{HAR-RV}^2 = \beta_0 + \beta_d RV_t^{(d)} + \beta_w RV_t^{(w)} + \beta_m RV_t^{(m)} + \varepsilon,$$

where $RV_t^{(d)}$, $RV_t^{(w)}$, and $RV_t^{(m)}$ represent realized variance (squared terms) for a one-day, one-week (5 day), and one-month (22 day) period on day t .²⁸ To get the multi-day measure, the “building block” of a single day’s RV is simply averaged over the desired period (e.g. sum up the RV for 5 days and divide by 5).

Estimating the coefficients involves initializing the model at a point in the past using as much historical data as possible. However, for a more robust and practical implementation, we use an expanding window—only letting the model use as much information that existed up to the point of the forecast date. This “walk-forward” approach helps avoid look-ahead bias and reduces the potential for overfitting the model.²⁹

HAR-RV tends to perform very well against the more complex GARCH and is easier to fit. It possesses some of the mean-reverting properties of GARCH and places more weight on recent observations while capturing the dynamics of intraday price action for improved response to volatility shocks. Like GARCH, HAR-RV has spawned a number of

²⁴ High-quality intraday historical data was not generally available in the 1980s. But as Merton also pointed out, there are theoretical limits to sampling at very high frequencies as the benefits of the increased precision can be swamped by the effects of market microstructure noise—the bouncing around between the bid/offer and any discrepancies in latency. The original paper can be accessed at https://www.nber.org/system/files/working_papers/w0444/w0444.pdf.

²⁵ The standard in the academic literature is 5-minute returns but similar results are also generated from bars using anywhere from 10-30 minutes. Any more frequent sampling than 5 minutes is generally thought to introduce too much noise in the process to be useful although there are a handful of studies going down to 1 minute, especially on broad large-cap indices that are aggregates of many stocks and somewhat less susceptible to noisy microstructure effects.

²⁶ Corsi (2009).

²⁷ The rationale behind the Heterogeneous Markets Hypothesis (HMH) from Ulrich Muller (1993) is that volatility is a function of trading activity from multiple agents in the market with distinct investing horizons: short term speculators (intraday), medium term traders (days), and long-term investors (weeks or longer). A brief article by Muller himself that succinctly describes the theory is available at <http://m.e-m-h.org/DMOP01.pdf>.

²⁸ While some researchers use overlapping data points in the regression (the one-day RV is a component of the 5-day RV), we find better results in an alternative approach that only spans unique information in each term. Effectively, the lags are computed as the simple averages of the Zero-th, 1st to 4th, and 5th to 21st lags, respectively.

²⁹ This process purposefully weakens the predictive power of the model in the earlier periods as less information is available to use in the forecast. It involves recursively fitting the OLS regression every day, adding information to the model as it becomes available. Since the coefficients tend to be very unstable with an insufficient amount of data in the beginning, we “burn off” a year of historical data to provide time for them to stabilize. As more time passes, the marginal impact of a single day on the coefficients becomes very small, but will respond to true volatility shocks, which is a desired feature.

extensions, focusing on correcting for asymmetry (using semi-variance in SHAR) or dynamically adjusting RV due to inherent measurement error (“integrated quarticity” in Bollerslev’s HARQ). But the biggest item to address is the glaring weakness in an RV-only estimator—the lack of any contribution from the overnight jump.

Without addressing the jump, HAR-based models will tend to underestimate volatility in comparison to methods like EWMA or GARCH that capture this information. A HAR-based model may be very effective in forecasting the RV for a single day, but the overnight return represents risk to practitioners that cannot be ignored even if RV is a “better” representation of the integrated, unobservable variance.

Bollerslev added a jump term to the basic model with his HAR-J, simply folding in the previous day’s overnight return as another term in the regression. Others decompose the jump into separate components matching the 1-, 5-, 22-day lags in the RV terms. More recently, Bollerslev concluded that simply adding the squared overnight return to the sum in calculating RV itself can be a more effective means of adding the information.³⁰ However, we find an alternative approach performs even better: scaling the RV-based forecast up to better approximate a measure using daily data.

As referenced before, Hansen and Lunde (2005) estimate that overnight returns contribute about one-fifth of the daily price variation based on their historical analysis of

Dow Jones Industrial Average stocks. In their paper that has been extensively referenced and replicated, they propose three ways to combine RV with overnight returns to get an equivalent variance for the entire day.³¹ The paper primarily focuses on the problem of getting the best estimator of integrated variance, but we use similar techniques to adjust HAR-RV-based forecasts to daily equivalents.

The first is to simply multiply the RV-based measure by a static scalar representing the average contribution from the overnight return. Using Hansen and Lunde’s estimate of one-fifth, this translates to a scalar of 1.25x (squared for variance).³² This static scalar can perform well, especially for broad market indices, but can be improved with a more dynamic approach. The second method just adds the overnight return to the RV measure as Bollerslev suggests. But Hansen and Lunde argue this method adds a very noisy measure to the estimate, potentially distorting the result.

The last method involves an optimized weighting of the overnight and RV components that helps smooth out some of the noisiness in just adding the overnight return. The weighting method is complex and beyond the scope of this paper to go through in any detail but has been well-supported by subsequent research since its initial publication.³³ We combine this optimization process with the HAR-RV in forming our preferred forecast detailed in the next section.

³⁰ Bollerslev, T, Tauchen, G, and Zhou, H (2009), “Expected Stock Returns and Variance Risk Premia”, *Review of Financial Studies*, 22, 4463–4492.

³¹ [Hansen and Lunde \(2005\)](#).

³² If one-fifth of whole-day variation is overnight that leaves 80% for the open to close period. Scaling up from 80% to 100% is 1.25x.

³³ Hansen and Lunde described differences in the relationship between the overnight and intraday volatility across different types of assets, noting a much larger impact from overnight returns in the limited number of technology stocks in their narrow universe of DJIA stocks. However, Todorova and Soucek and Ahoniemi and Lanne further demonstrate these differing relationships using the broader market (S&P 500) and international markets (using Australian equities) in concluding the superiority of the optimal weighting to other methods.

What about the VIX?

Most investors are familiar with the so-called “Fear Gauge”, the Cboe Volatility Index (VIX). The VIX was designed to measure market expectations of 30-day volatility implied by the pricing of S&P 500 index options. While forward-looking and driven by market prices, we omit the VIX for this analysis of volatility forecasting for several reasons.

For starters, while the VIX is a very useful indicator for the level of market volatility, the index itself is not tradable. Futures, options, and exchange-traded funds on the VIX are available but do not track the current spot index all that well—they align to what the market expects the VIX will be on expiration. They also have a term structure, with differing market-driven expectations of where VIX will be one month out, two months out, and so on.

More importantly, VIX is not well-suited for a one-day forecast as it measures expectations over the next 30 days. In a [prior study](#)*, we did use VIX as an estimate for a simple volatility-targeting strategy with daily rebalancing and it performed relatively poorly, even compared to a naïve 10-day historical volatility estimate.

To be clear, there is predictive value in the volatility implied by options pricing and some volatility models make use of this data. But we can only cover so much material and VIX and implied volatility could span several volumes.

*Barchetto, A and Poirier, R., “Risk Before Return: Targeting Volatility with Higher Frequency Data”, available at <https://tinyurl.com/6nvjhrnc>.

Comparing Forecast Accuracy

Now that we have defined a general set of estimators and forecasting methods ranging from basic to advanced, we compared the predictive power of each along with their pros and cons. For our test, we analyzed the ability of each estimator to forecast volatility one day ahead using a variety of proxies as benchmarks.

Data

As a proxy for US equity market volatility, we used returns from the SPDR S&P 500 ETF (SPY), with data from January 1, 1997 through December 31, 2020. We sourced open, high, low, and closing prices from Bloomberg and intraday 15-minute bars from Cboe’s Livevol Datashop.

Variance Proxies

Since true volatility is unobservable, selecting more than one proxy is advisable. To avoid biasing conclusions towards the more intraday-focused realized measures versus less frequent sampling using daily data, we used a range of both as proxies for all estimators.

Table 1 - Variance Proxy Calculations

| Proxy | Formula | Description |
|----------------------------|--|--|
| Squared Return | $r_t^2 = \left(\frac{Close_t}{Close_{t-1}} - 1 \right)^2$ | Daily return, squared |
| Realized Variance | $\sigma_{RV}^2 = \sum_{i=1}^M r_{t,i}^2$ | The sum of squared intraday 15-minute returns (where M=26) |
| Range | $RG_t^2 = (\ln(High) - \ln(Low))^2$ | The difference between the high and low of the day, squared ³⁴ |
| Full Day Realized Variance | $RV_t^{*2} = RV_t^2 + \left(\frac{Open_t}{Close_{t-1}} - 1 \right)^2$ | Adds the squared overnight return to the intraday realized variance to capture the whole day ³⁵ |

Loss Functions

In addition to benchmarks to measure against, the tools used to measure the errors from the predictions are equally as important. Like with the proxies, we used a range of loss functions³⁶ to provide more context. In the formulas in the following table, $\hat{\sigma}^2$ represents the model-estimated value and the h represents the (true) variance proxy.

³⁴ The high-low range and squared range are traditional variance proxies in the academic literature. However, the range as a proxy is subject to some unrealistic assumptions in the process that generated the data to be properly evaluated—zero mean return, no jumps, and constant conditional volatility during the day. Patton (2008) proposes an adjustment to the range (dividing it by a constant, $2\sqrt{\ln(2)} \approx 1.665$), which when squared becomes an unbiased proxy for conditional variance. However, this does not change the ranking of our estimators in any way and adds some complexity, so we use the more traditional range for simplicity. For additional information on adjusted range and evaluating variance proxies, please see: http://public.econ.duke.edu/~ap172/Patton_robust_forecast_eval_11dec08.pdf.

³⁵ This adjustment to realized variance is our preferred measure as it addresses both the intraday fluctuations as well as the opening gap. For practitioners, both are very relevant to a live position or hedge that is being monitored day to day. A miscalculation that leads to a mark that produces a significant hit to the P&L at the opening of trading or later in the day could lead to a tap on the shoulder—or worse.

³⁶ A “loss function” is a way of measuring how different a group of predicted values are from the true/known values and then expressing this as a single number. Lower values are preferred, indicating “less loss/error” from the true/known values.

Table 2 - Loss Function Calculations

| Loss Function | Formula | Interpretation |
|--------------------------------|--|--|
| Mean Squared Error (MSE) | $L(\hat{\sigma}^2, h) = (\hat{\sigma}^2 - h)^2$ | Penalizes more extreme errors |
| Root Mean Squared Error (RMSE) | $L(\hat{\sigma}^2, h) = \sqrt{MSE}$ | Denominated in the same units as the underlying data, making it easier to compare across models (like standard deviation compared to variance) |
| Mean Absolute Error (MAE) | $L(\hat{\sigma}^2, h) = \hat{\sigma}^2 - h $ | More robust to outliers, which may be an undesirable feature |
| Quasi-Likelihood (QLIKE) | $L(\hat{\sigma}^2, h) = \frac{h}{\hat{\sigma}^2} - \ln\left(\frac{h}{\hat{\sigma}^2}\right) - 1$ | An asymmetric measure that penalizes under-estimation more harshly |

Use of the “wrong” loss function can lead to some significant errors in ranking the efficacy of the forecasts. In prior work on volatility forecasting, MSE and QLIKE are shown to be more robust to noise in the proxy, making them generally preferred in terms of ranking.³⁷ While we focus on these two robust measures in our analysis, we include the other loss functions to highlight strengths and weaknesses of the estimators and show how their rankings can be distorted.

Forecast Methods

We use all the estimators presented in this paper for a comprehensive comparison. For measures that require a historical lookback, we specified short (10 day), medium (22 day), and long (100 day) periods and calculated a set of results for each.

Table 3 - Estimators and Specifications

| Estimator | Short Name | Additional Information |
|---------------------|------------|--|
| Historical Variance | HV | 10, 22, 100 days |
| Parkinson | Park | 10, 22, 100 days |
| Garman-Klass | GK | 10, 22, 100 days |
| Rogers-Satchell | RS | 10, 22, 100 days |
| Yang-Zhang | YZ | 10, 22, 100 days |
| EWMA | EWMA | 0.94 lambda |
| GARCH (1,1) | GARCH | Fit using maximum likelihood via arch v4.19 for Python |
| Realized Variance | RV | Prior day, uses 15-minute intraday bars |
| HAR-RV | HAR-RV | 1, 5, 22 lags, unique spans only ³⁸ |
| Scaled HAR-RV | HAR-RV+ | Multiplies HAR-RV by 1.25 ² |
| Optimized HAR-RV | HAR-RV* | A weighted sum of HAR-RV and the average squared overnight return over the prior month ³⁹ |

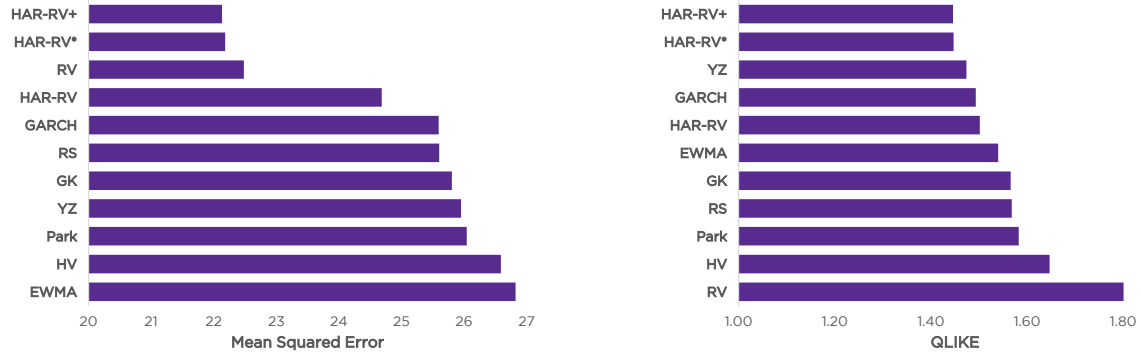
³⁷ See the previously referenced paper by Andrew Patton for more details on the robustness of forecasts, available at: http://public.econ.duke.edu/~ap172/Patton_robust_forecast_eval_11dec08.pdf.

³⁸ See footnote 28 on modification to the standard Corsi HAR process. Also uses the same 15-minute bars as RV.

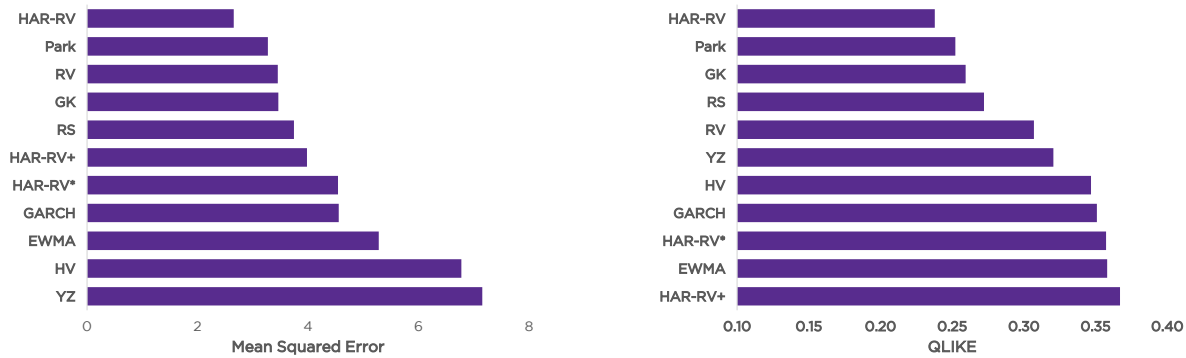
³⁹ This process uses the same weighting technique from Hansen and Lunde only with different estimators for the intraday and overnight components. We substitute HAR-RV for RV in the intraday and the mean squared close to open return over the prior 22 trading days for the one day overnight. Both are intended to improve accuracy and smooth out some of the noise in the overnight return from day to day.

Exhibit 1 - Ranked Forecast Performance by Estimator and Variance Proxy

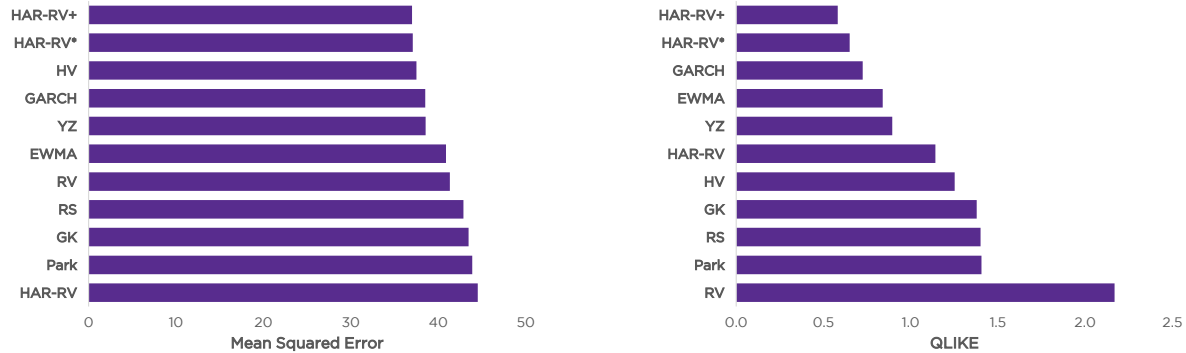
Panel A: Squared Return



Panel B: Realized Volatility



Panel C: Range



Panel D: Full Day Realized Variance

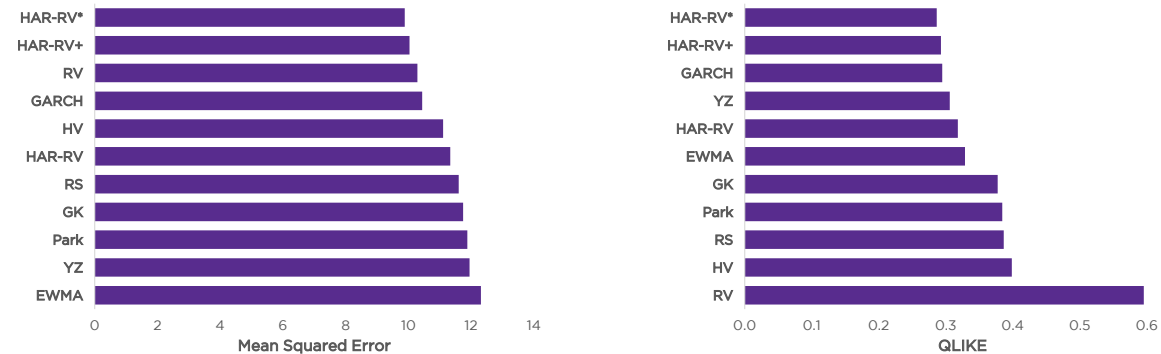


Exhibit 2 - All Forecast Results and Performance Relative to EWMA

| | <i>All Estimators and Loss Functions</i> | | | | <i>Relative to EWMA</i> | | | |
|--------------------------|--|------|-------|-------|-------------------------|--------|--------|--------|
| | MSE | RMSE | MAE | QLIKE | MSE | RMSE | MAE | QLIKE |
| Squared Return | | | | | | | | |
| Historical Variance | 26.59 | 5.16 | 1.664 | 1.646 | 99.1% | 99.6% | 99.4% | 106.9% |
| Parkinson | 26.04 | 5.10 | 1.458 | 1.583 | 97.1% | 98.5% | 87.1% | 102.8% |
| Garman-Klass | 25.80 | 5.08 | 1.457 | 1.566 | 96.2% | 98.1% | 87.0% | 101.7% |
| Rogers-Satchell | 25.60 | 5.06 | 1.465 | 1.568 | 95.5% | 97.7% | 87.5% | 101.8% |
| Yang-Zhang | 25.95 | 5.09 | 1.639 | 1.474 | 96.8% | 98.4% | 97.9% | 95.7% |
| EWMA | 26.82 | 5.18 | 1.674 | 1.540 | 100.0% | 100.0% | 100.0% | 100.0% |
| GARCH | 25.59 | 5.06 | 1.633 | 1.493 | 95.4% | 97.7% | 97.6% | 97.0% |
| Realized Variance | 22.48 | 4.74 | 1.384 | 1.800 | 83.8% | 91.6% | 82.7% | 116.9% |
| HAR-RV | 24.68 | 4.97 | 1.403 | 1.501 | 92.0% | 95.9% | 83.8% | 97.5% |
| Scaled HAR-RV | 22.13 | 4.70 | 1.561 | 1.446 | 82.5% | 90.8% | 93.2% | 93.9% |
| Optimized HAR-RV | 22.18 | 4.71 | 1.550 | 1.447 | 82.7% | 90.9% | 92.6% | 94.0% |
| Realized Variance | | | | | | | | |
| Historical Variance | 6.77 | 2.60 | 0.936 | 0.347 | 128.3% | 113.3% | 98.7% | 96.9% |
| Parkinson | 3.27 | 1.81 | 0.611 | 0.252 | 62.0% | 78.7% | 64.4% | 70.4% |
| Garman-Klass | 3.46 | 1.86 | 0.626 | 0.259 | 65.6% | 81.0% | 66.0% | 72.4% |
| Rogers-Satchell | 3.74 | 1.93 | 0.650 | 0.272 | 70.9% | 84.2% | 68.5% | 76.0% |
| Yang-Zhang | 7.15 | 2.67 | 0.939 | 0.321 | 135.4% | 116.4% | 99.0% | 89.5% |
| EWMA | 5.28 | 2.30 | 0.949 | 0.358 | 100.0% | 100.0% | 100.0% | 100.0% |
| GARCH | 4.56 | 2.13 | 0.878 | 0.351 | 86.3% | 92.9% | 92.5% | 98.0% |
| Realized Variance | 3.45 | 1.86 | 0.561 | 0.307 | 65.4% | 80.9% | 59.2% | 85.7% |
| HAR-RV | 2.65 | 1.63 | 0.524 | 0.238 | 50.3% | 70.9% | 55.2% | 66.4% |
| Scaled HAR-RV | 3.98 | 1.99 | 0.841 | 0.367 | 75.4% | 86.8% | 88.6% | 102.5% |
| Optimized HAR-RV | 4.54 | 2.13 | 0.838 | 0.357 | 86.0% | 92.7% | 88.4% | 99.8% |
| Range | | | | | | | | |
| Historical Variance | 37.49 | 6.12 | 2.004 | 1.251 | 91.7% | 95.8% | 99.5% | 149.1% |
| Parkinson | 43.87 | 6.62 | 2.127 | 1.405 | 107.3% | 103.6% | 105.5% | 167.4% |
| Garman-Klass | 43.46 | 6.59 | 2.121 | 1.378 | 106.3% | 103.1% | 105.2% | 164.2% |
| Rogers-Satchell | 42.87 | 6.55 | 2.122 | 1.400 | 104.9% | 102.4% | 105.3% | 166.8% |
| Yang-Zhang | 38.53 | 6.21 | 2.002 | 0.894 | 94.3% | 97.1% | 99.3% | 106.5% |
| EWMA | 40.87 | 6.39 | 2.016 | 0.839 | 100.0% | 100.0% | 100.0% | 100.0% |
| GARCH | 38.50 | 6.21 | 1.955 | 0.724 | 94.2% | 97.1% | 97.0% | 86.3% |
| Realized Variance | 41.32 | 6.43 | 2.203 | 2.167 | 101.1% | 100.6% | 109.3% | 258.2% |
| HAR-RV | 44.51 | 6.67 | 2.163 | 1.141 | 108.9% | 104.4% | 107.3% | 135.9% |
| Scaled HAR-RV | 36.99 | 6.08 | 1.886 | 0.581 | 90.5% | 95.1% | 93.6% | 69.2% |
| Optimized HAR-RV | 37.08 | 6.09 | 1.927 | 0.649 | 90.7% | 95.3% | 95.6% | 77.4% |
| Full Day RV | | | | | | | | |
| Historical Variance | 11.10 | 3.33 | 1.000 | 0.398 | 90.2% | 95.0% | 97.8% | 121.2% |
| Parkinson | 11.88 | 3.45 | 0.836 | 0.384 | 96.5% | 98.2% | 81.8% | 116.9% |
| Garman-Klass | 11.74 | 3.43 | 0.839 | 0.377 | 95.4% | 97.7% | 82.0% | 114.9% |
| Rogers-Satchell | 11.60 | 3.41 | 0.853 | 0.386 | 94.2% | 97.1% | 83.4% | 117.6% |
| Yang-Zhang | 11.94 | 3.46 | 0.990 | 0.306 | 97.0% | 98.5% | 96.8% | 93.1% |
| EWMA | 12.31 | 3.51 | 1.023 | 0.329 | 100.0% | 100.0% | 100.0% | 100.0% |
| GARCH | 10.43 | 3.23 | 0.940 | 0.295 | 84.8% | 92.1% | 91.9% | 89.7% |
| Realized Variance | 10.28 | 3.21 | 0.838 | 0.595 | 83.5% | 91.4% | 81.9% | 181.2% |
| HAR-RV | 11.33 | 3.37 | 0.790 | 0.318 | 92.1% | 95.9% | 77.3% | 96.8% |
| Scaled HAR-RV | 10.03 | 3.17 | 0.894 | 0.293 | 81.5% | 90.3% | 87.4% | 89.1% |
| Optimized HAR-RV | 9.88 | 3.14 | 0.886 | 0.286 | 80.3% | 89.6% | 86.7% | 87.2% |

Results Overview

In **Exhibit 1**, we ranked all estimators by their forecasting performance against each proxy, focusing on the two robust loss functions, MSE and QLIKE. We limited analysis to the 10-day lookback period for all estimators that require one as they clearly outperformed the longer term 22-day and 100-day lookbacks across all metrics by an overwhelming margin.

Despite being considered a noisy proxy, Squared Return in **Panel A** in **Exhibit 1** provides some initial insights into the strengths and weaknesses of each estimator. The realized variance measures (RV, HAR-RV*, HAR-RV+) performed best against MSE, but RV itself drops to the bottom using QLIKE. Since Squared Return includes the overnight jump, the intraday-only RV is penalized more for its under-estimation. EWMA performed relatively poorly against MSE but was much better using QLIKE, helped by its inclusion of the overnight return in its calculation and higher weighting of more recent observations. GARCH, and the three HAR variants are all consistently strong against both loss functions.

Since Realized Variance removes the overnight component, it should be no surprise that RV, HAR-RV, and the range-based estimators without a jump component performed relatively better in **Panel B**. The relatively poor performance of standard industry methods GARCH and EWMA against QLIKE suggests that Realized Variance is not well-aligned with practice as a benchmark, as both are widely used in real-world applications. And the addition of an overnight component to HAR-RV+ and HAR-RV* is designed to make them equivalent to measures like GARCH and EWMA.

Panel C shows that despite their name, the range-based estimators (Parkinson, Garman-Klass, Rogers-Satchell) except for Yang-Zhang were relatively poor predictors of the Range. The more sophisticated models performed well, especially against QLIKE, with the weighting of more recent results helping EWMA and GARCH guard against under-estimation. HAR-RV* and HAR-RV+ also place more weight on recent results, but also capture intraday variation which apparently leads to some value in forecasting the daily range.

Lastly, using the full range of intraday and overnight variation as a proxy in **Panel D**, the more sophisticated estimators that either directly capture or model both components perform well whereas the range-based (Parkinson, Garman-Klass, Rogers-Satchell) and more naïve ones (RV, HV) performed poorly, especially RV against QLIKE.

Exhibit 2 shows the full range of values by variance proxy for each estimator against all four loss functions. The top two most accurate estimators in each column for each proxy are colored in green and the least accurate two in red. To the right of the performance against each

proxy, we also report performance of all estimators benchmarked relative to EWMA, a standard estimator with wide application throughout the financial industry. All values below one (more accurate than EWMA) are colored green whereas values greater than one (less accurate than EWMA) are colored red.

Focusing on the results relative to EWMA highlights some broad differences in how these estimators work. The HAR models and GARCH are clustered fairly closely to EWMA across the range of variance proxies, owing to their inclusion of jump returns and weighting of more recent values.

The output for QLIKE and MSE (and indirectly RMSE as it is just the square root of MSE) aligns with the rankings in **Exhibit 1**. But looking at the other metrics such as MAE is quite instructive as it demonstrates the importance of using the correct loss function for the task. Estimators like RV or HAR-RV that use intraday data and exclude the jump component look strong compared to EWMA, GARCH, HAR-RV*, or HAR-RV+ when measured against a Squared Return proxy and MAE. But when factoring in the risk of under-estimation, HAR-RV slips and RV falls to the worst ranked estimator in using QLIKE to evaluate.

Across the loss functions and multiple proxies, the realized volatility estimators that adjust for overnight returns—Scaled HAR-RV (HAR-RV+) and Optimized HAR-RV (HAR-RV*)—are consistently top performers. While measuring against the combination of intraday squared returns and the squared overnight gap might seem tailor-made for these estimators, their strong performance against all the proxies further supports their case as very effective forecasting tools. And while not showing up in the top two for any column, EWMA and especially GARCH also performed very well, reflective of their traditional and continued use by practitioners. The more naïve early estimators based on historical standard deviation or daily range are simply less effective, only scoring well against a limited combination of proxy and loss functions.

Conclusion

The wider availability of intraday data over the last twenty years has encouraged more research into its use for volatility forecasting. But while realized volatility is used in practice among quantitatively-oriented asset managers, investment banks, and other financial institutions, anecdotally EWMA and GARCH are still more prevalent in practice. For specialized volatility-targeted products such as risk control indices used in structured products or fixed index annuities, EWMA-based forecasts for targeting are clearly dominant (although different models are used to help hedge their exposure). For portfolio risk management and derivatives pricing, more sophisticated variants of GARCH are widely in use, which may perform significantly better than the standard GARCH (1,1) presented here. There are also some GARCH models that use realized volatility components, adding

the increased responsiveness of intraday data to improve their forecasts. But they come at the cost of increased complexity in fitting and maintaining these models. For accuracy, responsiveness, and ease of calculation, it is difficult to beat some of the HAR-based estimators that are properly tuned for the full day's volatility.

While the basic HAR-RV has been in use for over a decade, there are still ongoing efforts to improve its forecasting accuracy using the same basic structure. Bollerslev, who initially developed GARCH, has been primarily focused on research using realized volatility and HAR-based models in recent years. The fine-tuning of the continuous intraday and overnight components into an estimate that best captures the desired risk characteristics for a given day is of critical importance. While the realized variance may be an unbiased proxy for the latent integrated variance, it fails to account for all the practical risks that a professional with capital at risk faces. **A sharp loss at the opening of trading in a volatile market that is exacerbated by an inferior risk model is a threat to one's P&L and potentially their job.** In our opinion, the impact of the overnight return plus the variation experienced throughout the trading session is most aligned with the practitioner and the best proxy for volatility. As a result, a risk forecast that explicitly accounts for these components with solid performance across a range of variance proxies such as HAR-RV+ or

HAR-RV* is very well suited for practical application in day-to-day risk management applications.

As stated previously, covering all forms of volatility forecasting in a relatively brief overview would be impossible. In addition to many variants in the GARCH family of models, there are other methods we do not even mention (the Heston model, Markov chains, Monte Carlo simulations, and other stochastic volatility models). Market-derived inputs such as option implied volatility or VIX are another form of forward-looking measure that can be used to help predict future volatility.

Regardless of the tools and methods used, we believe that higher frequency data will continue to be an evolving part of volatility forecasting and risk management for years to come. Intuitively, having more frequent sampling of data should lead to more accurate results. If asked to forecast tomorrow's temperature at 3pm, would you rather use the seasonal average, the average over the last two weeks, or the hourly temperature over the last 48 hours? With the low cost of computing power and data storage compared to even ten years ago, we expect the future of forecasting will harness this data for even more sophisticated techniques, especially where they tend to be most effective in the short-term.

About the Authors

Anthony Barchetto, CFA

tony@saltfinancial.com

Tony is the Founder and Chief Executive Officer of Salt Financial, a leading provider of index solutions and risk analytics for investment advisors, fund sponsors, investment banks, and insurance carriers. Salt develops risk-focused index strategies built for higher precision and responsiveness to changing markets using its truVol[®] and truBeta[®] analytics.

Prior to founding Salt, Tony led corporate development for Bats Global Markets, helping guide the firm from its merger with rival Direct Edge to initial public offering and subsequent acquisition by Cboe Global Markets. Before Bats, Tony served in a similar capacity for global institutional trading platform Liquidnet. Earlier in his career, he served in leadership roles in trading, sales, and product management with Citigroup and Knight Capital Group.

Tony earned a BS in International Economics from Georgetown University and is a CFA Charterholder.

Ryan Poirier, ASA, CFA, FRM

ryan@saltfinancial.com

Ryan is Director, Index Products & Research for Salt Financial. In his role, he is responsible for the conceptualization and design of index-based strategies leveraging Salt Financials' full suite of truVol[®] and truBeta[®] analytics. He also leads the research efforts, focusing primarily on risk forecasting and portfolio construction.

Prior to joining Salt Financial, Ryan spent three years at S&P Dow Jones Indices as part of the Global Research & Design team, where he grew the quantitative index offerings for banks, asset managers, and insurance companies. He is a CFA Charterholder, FRM Charterholder, and an Associate of the Society of Actuaries. Ryan earned a MS in Financial Engineering from New York University and a BS in Mathematics and Finance from SUNY Plattsburgh.

Junseung (Eddy) Bae

eddy@saltfinancial.com

Eddy is Manager, Products and Technology for Salt Financial. Eddy is responsible for driving the product development and technology processes at Salt. Eddy has held several technical and analytical roles with e-commerce and fintech startups.

He is a former army intelligence analyst with the Republic of Korea and holds a BS in Business with a concentration in Computing and Data Science from New York University.

About Salt Financial

Salt Financial LLC is a leading provider of index solutions and risk analytics, powered by the patent pending truVol[®] Risk Control Engine (RCE) and proprietary truBeta[®] portfolio construction tools. We leverage the rich information contained in intraday prices to better estimate volatility to develop index-based investment products for insurance carriers, investment banks, asset managers, and ETF issuers. Salt is committed to collaborating with industry leaders to empower the pursuit of financial outperformance for investors worldwide.

For more information, please visit www.saltfinancial.com.

Academic References

- Ahoniemi, K. & Lanne, M. (2010). Realized volatility and overnight returns. Bank of Finland Research Discussion Papers, 19.
- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885-905.
- Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up Including Jump Components In the Measurement Modeling and Forecasting of Return Volatility. *The Review of Economic and Statistics*, 89(4), 701-720.
- Audrino, F., & Knaus, S. (2016). Lassoing the HAR Model: A Model Selection Perspective on Realized Volatility Dynamics. *Econometric Reviews*, 35(8-10), 1485-1521.
- Audrino, F., Huang, C., & Okhrin, O. (2017). Flexible HAR Model for Realized Volatility. Working Paper.
- Bandi, F. M., & Russell, J. R. (2008). Microstructure Noise, Realized Variance, and Optimal Sampling. *The Review of Economic Studies*, 75(2), 339-369.
- Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society*, 64(2), 253-280.
- Bekierman, J., & Manner, H. (2016). Improved Forecasting of Realized Variance Measures. Working Paper.
- Biais, B., Glosten, L., & Spatt, C. (2005). Market Microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications. *Journal of Financial Markets*, 8(2), 217-264.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192, 1-18.
- Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected Stocks Returns and Variance Risk Premia. *Review of Financial Studies*, 22, 4463-4492.
- Buccheri, G., & Corsi, F. (2019). Hark the Shark: Realized Volatility Modelling with Measurement Errors and Nonlinear Dependencies. Working Paper.
- Corsi, F. (2009). A Simple Long Memory Model of Realized Volatility. *The Journal of Financial Econometrics*, 7, 174-196.
- Corsi, F., & Reno, R. (2009). HAR volatility modelling with heterogeneous leverage and jumps. Working Paper.
- Fleming, J., Kirby, C., & Ostdiek, B. (2003). The Economic Value of Volatility Timing Using 'Realized' Volatility. *Journal of Financial Economics*, 63(3), 473-509.
- French, K. R., & Roll, R. (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, 17, 5-26.
- Hansen, P. R., & Lunde, A. (2005). A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data. *Journal of Financial Econometrics*, 3(4), 525-554.
- Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realized and Implied Volatility Measurements. *Journal of Empirical Finance*, 12, 445-475.
- Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5 minute RV A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187, 293-311.
- Luong, C., & Dokuchaev, N. (2018). Forecasting of Realised Volatility with the Random Forests Algorithm. *Journal of Risk and Financial Management*, 11(4).
- Madhavan, A. (2000). Market Microstructure: A Survey. *Journal of Financial Markets*, 3(3), 205-258.
- Martens, M. (2002). Measuring and Forecasting S&P 500 Index Futures Volatility Using High-Frequency Data. *Journal of Futures Markets*, 22, 497-518.
- McAleer, M., & Medeiros, M. (2008). A Multiple Regime Smooth Transition Heterogeneous Autoregressive Model for Long Memory and Asymmetries. *Journal of Econometrics*, 147(1), 104-119.
- McAleer, M., & Medeiros, M. (2011). Forecasting Realized Volatility with Linear and Nonlinear Models. *Journal of Economic Surveys*, 25(1), 6-18.
- Merton, R. (1980). On Estimating the Expected Return on the Market. *Journal of Financial Economics*, 8, 323-361.
- Patton, A. J., & Sheppard, K. (2015). Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility. *The Review of Economic and Statistics*, 97(3), 683-697.
- Todorova, N. & Soucek, M. (2014). Overnight information flow and realized volatility forecasting. *Finance Research Letters*, 11(4), 420-428.

Disclosures

Copyright® 2021 Salt Financial Indices LLC. Salt Financial Indices LLC is a division of Salt Financial LLC. “Salt Financial” and “TRUBETA”, and “TRUVOL” are registered trademarks of Salt Financial Indices LLC. These trademarks together with others have been licensed to Salt Financial LLC and/or designated third parties. The redistribution, reproduction and/or photocopying of these materials in whole or in part are prohibited without written permission. This document does not constitute an offer of services in jurisdictions where Salt Financial Indices LLC, Salt Financial LLC or their respective affiliates (collectively “Salt Financial Indices”) do not have the necessary licenses. All information provided by Salt Financial Indices is impersonal and not tailored to the needs of any person, entity or group of persons. Salt Financial Indices receives compensation in connection with licensing its indices to third parties. Past performance of an index is not a guarantee of future results.

It is not possible to invest directly in an index. Exposure to an asset class represented by an index is available through investable instruments based on that index. Salt Financial Indices does not sponsor, endorse, sell, promote or manage any investment fund or other investment vehicle that is offered by third parties and that seeks to provide an investment return based on the performance of any index. Salt Financial Indices makes no assurance that investment products based on the indices will accurately track index performance or provide positive investment returns. Salt Financial Indices is not an investment advisor and makes no representation regarding the advisability of investing in any such investment fund or other investment vehicle. A decision to invest in any such investment fund or other investment vehicle should not be made in reliance on any of the statements set forth in this document. Prospective investors are advised to make an investment in any such fund or other vehicle only after carefully considering the risks associated with investing in such funds, as detailed in an offering memorandum or other similar document that is prepared by or on behalf of the issuer of the investment fund or other investment product or vehicle. Salt Financial Indices is not a tax advisor. A tax advisor should be consulted to evaluate the impact of any tax-exempt securities on portfolios and the tax consequences of making any particular investment decision. Inclusion of a security within an index is not a recommendation by Salt Financial Indices to buy, sell, or hold such security, nor is it intended to be investment advice and should not be construed as such.

These materials have been prepared solely for informational purposes based upon information generally available to the public and from sources believed to be reliable. No content contained in these materials (including index data, ratings, credit-related analyses and data, research, valuations, model, software or other application or output therefrom) or any part thereof (“Content”) may be modified, reverse engineered, reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of Salt Financial Indices. The Content shall not be used for any unlawful or unauthorized purposes. Salt Financial Indices and its third-party data providers and licensors (collectively, the “Salt Financial Indices LLC Parties”) do not guarantee the accuracy, completeness, timeliness or availability of the Content. The Salt Financial Indices LLC Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON AN “AS IS” BASIS. THE SALT FINANCIAL INDICES LLC PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT’S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall the Salt Financial Indices LLC Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs) in connection with any use of the Content even if advised of the possibility of such damages.

In addition, Salt Financial Indices provides services to, or relating to, many organizations, including but not limited to issuers of securities, investment advisers, broker-dealers, investment banks, other financial institutions and financial intermediaries, and accordingly may receive fees or other economic benefits from those organizations, including organizations whose securities or services they may recommend, rate, include in model portfolios, evaluate or otherwise address. Salt Financial Indices has established policies and procedures to maintain the confidentiality of certain nonpublic information received in connection with each analytical process it performs in connection with the services it provides.