THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF STATISTICS

Modeling Stock Return Volatility, a Comparative Approach

ROBERT KRIMETZ
SPRING 2023

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Data Science and Economics
with honors in Data Science

Reviewed and approved* by the following:

Matthew Beckman
Professor of Statistics and Data Science
Thesis Supervisor

Anh Le
Professor of Finance
Honors Adviser

Jeremy Seeman
Professor of Statistics
Faculty Reader

* Electronic approvals are on file.

ABSTRACT

The application of machine learning and probabilistic programming methods on stock return prediction has grown in tandem with the availability of high frequency stock data. With well recorded heteroskedasticity in historical stock returns, modeling attempts have evolved from making general assumptions about the underlying data generating distribution to predicting changes in the underlying distribution of returns. The increase in popularity of 'tradable volatility' through derivative contacts and VIX futures over the past three decades has motivated research efforts to model the variance of daily returns. Along this line of research, three schools of thought have emerged to model return volatility; Time Series Models, Stochastic Models, and Bayesian Models. Given that the preliminary assumptions underlying these models differ, the nature of their results and the varying metrics used to calculate their respective accuracy makes it difficult to directly compare them. Accordingly, the currently available pool of research has diverged along these three separate paths making it unclear the advantages of each. Notably, Bayesian models have largely been neglected in the current pool of research due to their computational intensity. In this paper I derive ten time series and Bayesian models then provide a comprehensive comparative study of the results on real stock data. I found that Bayesian models with intractable posterior distributions significantly outperform time series models at predicting directional change in future volatility, while the GARCH and FIGARCH time series models generate the most accurate point predictions for future volatility. I hope the results outlined in this paper better contextualize different volatility predictions and motivate the creation of more accurate tradeable volatility models.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor and honors advisor, Dr. Matthew Beckman for his guidance throughout my undergraduate career. I would also like to thank my thesis advisor Dr. Anh Le, and Jeremey Seeman for their assistance throughout the research process.

1. Introduction

1.1 Motivation

With respect to daily US stock returns, consensus dictates that normalized returns and detrended returns follow a log-normal distribution [4]. Further analysis across both U.S. and international equity markets have found mixed evidence on the underlying distributions that most accurately characterize daily price action; however, the recurrent appearance of mean reversion and stationarity among other factors implies that the distribution of stock returns changes over time. This understanding has led to continued attempts at modeling short-term stock returns, specifically the variance/volatility of stock returns. From the perspective of fundamental analysis, noticeable changes in the 'typical' volatility of daily stock returns are attributed to different economic regimes. For example, periods of persistently high volatility are attributed to economic uncertainty which then can be rationalized by a variety of qualitative talking points including economic quadrant shifts, changes in federal reserve policy, global supply and demand shocks, etc. The main takeaway here is that there are observable periods of noticeably high volatility during times of economic uncertainty and low volatility during economic booms. This indicates an intuitive understanding that daily returns do not follow a static log-normal distribution as many models assume. A statistical interpretation of this phenomenon is that there is inertia/clustering in the volatility of stock returns [13]. In other words, there is a tendency for periods of high volatility to follow periods of high volatility.

1.2 Problem

Given these observations, there is a clear motivation to model the volatility of stock returns as a function of time, however volatility in equity market is an ambiguous term. There are two ways to measure volatility; realized volatility and implied volatility. Realized volatility is the standard deviation of a stock's price over the past X days. There are many variations of realized volatility calculations based on either closing price or intra-day prices, and the number of observations used to make the calculations. Implied volatility on the other hand is a parameter of the Black Scholes Model which is forward looking in the sense that it indicates the market's expectation of future price volatility based on observed prices in the options market. From an econometrics perspective, implied volatility represents a risk neutral measure of the aggregate market risk premium [18] while realized volatility is deemed a risk adverse measure based on observed historical prices. This explains why implied volatility and realized volatility measures systematically differ. Since there is no clear indication which measure is more useful to predict, there is a divide in pool of literature where some papers predict future realized volatility while other predict implied volatility adding a layer of obscurity to model comparison. Although neither realized volatility nor implied volatility is a perfectly accurate measure for the 'true' volatility of stock returns, in practice both are useful. Most economists focus on predicting future realized volatility given that it can be directly observed from stock prices. There are natural critiques of the accuracy of implied volatility given that Black Scholes model assumes that the implied volatility function of an option should be flat and constant through time [7], which is observably untrue in markets. However, the general adoption of the VIX as a tradeable volatility benchmark and the increased availability of derivative contracts have made accurate predictions of implied volatility a profitable indicator none the less. For example, a market maker who has perfectly hedged all exposure to the greeks will be able to profit from a correct prediction of future implied volatility movement [9]. Analytically there is an unclear correlation between realized volatility and implied volatility that changes overtime, but for the sake of this analysis both are used as independent target variables. Accordingly, this paper derives a variety of novel heteroskedastic Bayesian and time series models to predict both realized and implied volatility. The models are back tested on the current constituents of the S&P 500 and a cross sectional analysis of the results is done by sector volatility classification. This paper is structured as follows, section 2 is a literature review of volatility modeling, section 3 overviews the data sets and methodologies used along with some required preliminaries for each of the models. Section 4 examines the results of the analysis and section 5 ends with a conclusion and a discussion of future works.

2. Literature Review

This section provides an overview of the current literature on volatility modeling. The subsections are broken down based on the three primary schools of thought as outlined in the survey paper Samsudin and Mohamad (2016) [37]. The original volatility models were based purely on observed trading ranges. Parkinson (1980) [35] derived the extreme value method for estimating the variance of stock returns based on the distribution of observed extreme values. Garman and Klass (1980) [19] developed an estimator for historical volatility based on observed daily high, low, and closing prices synthesized by the Average True Range (ATR). Rogers and Satchell (1994) [36] developed a more accurate range estimator for daily volatility based on daily high, low, and closing prices by incorporating a historical look back period. Finally, Yang and

Zhang (2001) [43] modified the Garman and Klass estimator to be independent of the drift and expected return of an underlying asset further improving performance. In a comparative study of the above range estimators by Shu and Zhang (2006) [39] the researchers found that all estimators were effective at capturing past realized volatility if prices followed a geometric Brownian motion with small drift. Both Rogers and Satchell's estimator as well as Yang and Zhang's estimator were shown to be resistant to drift components, but only Yang and Zhang's estimator was resistant to large initial jumps. These range estimators were the starting point for modern volatility modeling; however, their results were by and large snap shots of past price action.

2.1 Time Series Volatility Models

The time series family of volatility models started with the introduction of the ARCH class of models by Engle (1982) [17] which improved variance predictions by allowing the conditional variance of returns to be a function of previous price action. Engle accomplished this by incorporating past squared error terms into future predictions. This model was further improved by Bollerslev (1986) [10] who developed the GARCH model which allowed past conditional variance terms alongside past squared error terms to influence predictions. Nelson (1991) [30] developed the EGARCH model, which extended the GARCH model by allowing there to be asymmetry in the impact of positive and negative shocks with the goal of incorporating the leverage effect into predictions. Bollerslev and Mikkelsen (1996) [11] furthered the development of this family of models by comparing the performance of the ARFIMA (Auto-regressive Fractionally Integrated Moving Average) model, the HYGARCH

(Hyperbolic GARCH) model, and their newly proposed FIGARCH (Fractionally Integrated GARCH) model to incorporate the long memory of stock returns, in other words the observed persistence in volatility tendencies. Outside of expanding the types of regressors in the ARCH family of models, more recent research has looked at changing the underlying assumptions of market structure. Ghysels, Santa-Clara, and Valkanov (2006) [20] looked to optimize time series volatility predictions by implementing a MIDAS (Mixed Data Sampling) framework. Curto (2009) [14] found that the performance of the standard GARCH model could be greatly improved by changing the distributional assumptions from normal/log-normal to the student's t distribution or the stable Pareto distribution. Chistoffersne, Jacobs, and Minmouni (2010) [12] further altered the GARCH model by incorporating realized variance into the Square Root Model (SQR) through a non-affine expansion resulting in the Realized GARCH model. Ardia and Hoogerheide (2010) [5] improved the efficiency of implementing the GARCH models with an assumed Student's t distribution by proposing the use of Bayesian Inference for parameter selection instead of Maximum Likelihood estimation. Ahoniemi (2008) [1] applied the GARCH, EGARCH, and ARIMA-GARCH family of models to tradeable volatility in the form of VIX futures where the latter were found to significantly improve directional prediction accuracy. Rosy, Dong Wan, and Man-Suk (2017) [32] combined many of these previous innovations and prosed a combined ARFIMA and GARCH with Student's T innovations using Bayesian Inference for parameter selection through the use of JAGS. One of the most recent innovations is the space is the development of the HAR (Heterogenous Autoregressive) model by Gong and Lin (2019) [21], which uniquely takes into account realized volatility and uses ensemble empirical mode decomposition to decomposed realized volatility into pure volatility and the leverage effect. Continued research in the time series family of models looks at further altering the distributional assumptions which in many cases is computationally expensive leading to further research in Bayesian methods for parameter approximation. In addition to its use for parameter selection in time series models, an entire family of volatility models is derived using Bayesian inference which is covered next.

2.2 Bayesian Volatility Models

Unlike the other estimators which solve for a point estimate of a parameter, Bayesian models generate a probability distribution of the target variable. For volatility modeling, Bayesian models are directly applicable because they can capture how the distribution of volatility changes over time. Anecdotally, during times of economic uncertainty where stock prices are extremely volatile the resulting distribution of a Bayesian model will 'spread out'. In other words, Bayesian models assume that volatility itself has an underlying data generating distribution, allowing the predictions to reflect changes in both the scale and location parameters of volatility. Karolyi (1993) [27] was the first to use the Bayesian framework to price options contracts using the inverse gamma and normal conjugate pair. The resulting options prices from his implied volatility estimates were more accurate than both implied volatility pricing models and an ex-post estimator (realized volatility). Cunha and Rao (2014) [15] validated these results on a larger set of data where uninformed priors were still able to accurately converge to the observed realized volatility. Yang and Lee (2011) [44] further extended this method using a Bayesian Kernel to generate accurate credible intervals for implied volatilities. Oostdam (2021) [33] applied a Bayesian model to tradeable volatility by generating credible intervals for VIX future contracts, he found that the No-U-Turn sampler significantly improved the efficiency of drawing from the posterior compared Metropolis Hastings and other Hamilton MC methods. Although the literature supports the usefulness of Bayesian inference to predict both implied and realized volatility, the computationally expensive nature of the calculations and the need for advanced Markov Chain Monte Carlo sampling methods made research more difficult. However, with increased computing power, these methods can more readily be employed.

2.3 Stochastic Volatility Models

The third family of volatility models are the Stochastic Volatility (SV) models. Recall from the overview of time series models that the theoretical starting point was allowing for heteroscedasticity in the ARCH models derivation. Stochastic volatility models are slightly different in that they specifically address time-varying volatility in financial data by modeling volatility as a stochastic process. It's important to note that there is some overlap in the classification of stochastic volatility models and time series models. For some of the more sophisticated time series models, the function that define future variance predictions contains a random component. This paper only focuses on the comparison of time series and Bayesian models, however; for completeness an overview of stochastic volatility models is included. Taylor, S.J (1986) [40], independent of Engle (1982) [17], derived the ARCH model but emphasized the stochastic properties of asset returns. Heston (1993) [23] introduced a new options pricing model by assuming that the underlying data generating function of stock price volatility is a stochastic process driven by a Weiner process. The Heston model incorporates the constant elasticity assumption of variance by using a mean-reverting square root diffusion process. Unlike the GARCH model, this created an options pricing framework with a closed form solution. Bates (1996) [8] proposed the Stochastic Volatility Jump Diffusion Model which incorporated both a stochastic volatility and jump diffusion process to better characterize asset returns. Dupire (1994) [16] introduced the concept of local volatility to price options with a volatility smile, his closed form solution allowed for accurate implied volatility predictions without making assumptions about the underlying stochastic process. A drawback of the local volatility approach set forth by Dupire is that when the price of the underlying asset falls (rallies), the local volatility model predicts that the smile will shift to a higher (lower) price which is the opposite of observed market behavior. To correct for this Hagan, Kuman, Lesniewski, and Woodward (2002) [22] proposed the SABR model which allows an asset's forward values and its volatility to correlate. The model assumes that returns follow a geometric Brownian motion with a stochastic volatility component following a log normal process. This innovation resulted in an options pricing model that more accurately captures changes in the volatility smile. Other modeling attempts include Ahoniemi (2007) [2] who analyzed the accuracy of a mixture multiplicative error models to predict directional changes in the Nikkei 225's implied volatilities compared to an ARIMA model. This work was extended in Ahoniemi (2008) [3] where call and put implied volatility for the USD/EUR were jointly modeled in a mixture multiplicative error model to accurately capture regime switching behavior. The success of joint volatility modeling was further explored in Ahoniemi and Lanne (2009) [28] where a bivariate mixture multiplicative error model was used improved call and put implied volatility predictions from previous mixture models on the Nikkei 225. Papanicolaou and Sircar (2013) [34] applied this sharp regime shift mixture effect to the Heston model using asymptotic and Fourier methods resulting in a more accurate prediction for the VIX index. Recent research surrounding stochastic volatility models has further extended and improved the efficiency of these adaptations for the leverage effect. Hosszehni and Kastner (2019) [26] implemented Bayesian Estimation for parameter selection of Stochastic Volatility models with leverage to improve sampling efficiency making the implementation of these methods more practical. Finally, Sepp and Rakhmonov (2022) [38] addressed the short comings of SV models on assets with positive return-volatility correlation by adding a quadratic drift component, extending the feasible scope of these models to assets such as inverse Bitcoin options.

2.4 State of the Art (Machine Learning Models)

Although there are continuing research efforts in all three branches of volatility modeling, most recent publications have involved the application of machine learning algorithms to model implied volatility surfaces using attributes for all three previously mentioned modeling methodologies. For completeness I will now overview this newer class of models. Avellaneda, Carelli, and Stella (2000) [6] implemented the Bayesian Framework through a Neural Network to generate implied volatility surfaces which they found accurately captured the general effect of option smiles. Tino, Nikolaev, and Yao (2005) [41] demonstrated that Sparse Bayesian Kernal models could be used to predict directional changes in implied volatility surfaces leading to a successful straddle trading strategy. Hosker, Djurdjevic, Nguyen, and Slater (2018) [25] compared a variety of machine learning methods (Recurrent Neural Networks, PCA, etc.) and ARIMA models in predicting the future price of VIX futures. The researchers found that the Recurrent Neural Networks and Long Short-Term Memory ML models outperformed all other ML and timeseries models. Medvedev (2022) [29] applied two novel neural network architectures for multi-step timeseries forecasts to build implied volatility Surfaces, encoding the

IV term structure as a discrete dimension. Nybo (2021) [31] conduction a cross sectional analysis of volatility prediction between GARCH and Artificial Neural Networks by sector. He observed that the ANN models outperformed on assets from sectors with historically low volatility whereas the GARCH models performed much better on assets from sectors with medium/high volatility. Vrontos, Galakis, and Vrontos (2021) [42] found that Machine Learning Methods outperformed econometric models at predicting directional changes in implied volatility in virtually all settings. Finally, Hirsa, Osterrieder, and Misheva (2021) [24] found that machine learning could be used to accurately price VIX futures using only a subset of the options data used by the CBOE to update contract value.

3. Data and Methodology

3.1 Data

There were two sources of data used in the following analysis. All 503 current constituents of the S&P 500 were considered to be part of the analysis. The volatility data was pulled from the US Equity Historical and Option Implied Volatilities dataset provided by Quantcha. The dataset contains volatility data on over 8,000 U.S. equities publicly traded on the NASDAQ, NYSE, NYSE ARCA, and NYSE American (formerly AMEX). The dataset included historical realized volatilities generated by both the Close-to-Close and Parkinson models, and at-the-money option-implied volatilities for calls/puts, means, and skew steepness indicators from 10 to 180 days from expiration. All available volatility data was pulled for the current S&P500 constituents from January, 1st, 2016 until January, 1st, 2023. Tickers with more the 5.00% of historical volatility metrics missing from the available range of dates unique to each ticker was dropped from the dataset, resulting in 487 remaining tickers. Historical price data was pulled from the Yahoo Finance API and both datasets were stored in a locally hosted PostgreSQL database, data aggregation and cleaning was done using python, and the back tests were run using Jupyter Notebooks. The use of only current S&P500 constituents in the preceding analysis allows for the influence of survivorship bias in the results. Accordingly, the modeling results should only be interpreted under the context of volatility prediction and not as a proxy for daily return prediction.

3.2 Preliminaries

Recall from the introduction that there are two primary proxies for volatility; realized volatility and implied volatility. There are two main ways to calculate realized volatility; Close-to-Close Volatility and Parkinson's Volatility.

$$CCHV = \sqrt{\frac{1}{\mathrm{n}}\sum \mathrm{X}_{\mathrm{i}}^{2}}$$

$$\mathrm{ParkinsonHV} = \sqrt{\frac{1}{4\mathrm{N}\ln(2)}\sum \ln^{2}(\frac{\mathrm{h}_{\mathrm{i}}}{\mathrm{l}_{\mathrm{i}}})}$$

Close-to-Close Volatility is a simple variance calculation applied to the closing price of a stock over the past X days. Despite its simplicity, this metric closely resembles the variance used in the standard VIX calculation and many corresponding derivative pricing strategies. Parkinson's Volatility is similar but takes a more sophisticated approach by considering the range of intraday prices in addition to the closing price. Given that Close-to-Close HV is more prominently used in volatility modeling research, 10 day rolling Close-to-Close realized volatility was the metric used for realized volatility in this analysis. Implied volatility on the other hand is a parameter of the Black Scholes Model. This analysis used the mean of implied volatilities for at-the-money (ATM) calls and puts with an average expiration of ten calendar days from the measurement date. Now that the two-volatility metrics used in the preceding analysis have been outlined, there are two general modeling approaches used in this analysis. The first approach is to directly model the behavior of rolling historical volatility. Put simply, these models look to directly predict the value of realized volatility and implied volatility given the previous values of realized volatility. The second approach is to build a model which predicts future daily returns based on historical normalized daily returns, and uses the residual variance of the predictions as the estimate of volatility. For this kind of model, daily price action is used to predict future daily price action and the variability in these predictions are used as our volatility estimates. Both of these general approaches are implemented for the time series and Bayesian models. The next section will outline the derivation of the various family specific models.

3.3 Bayesian Models

Bayesian models are an ideological extension of Bayes theorem. The core idea behind the Bayesian framework is that prior beliefs are updated after observing data. Unlike a frequentist approach, the Bayesian framework is fed prior assumptions about the target distribution (call these the prior distributions) and then assigned a likelihood function to update these prior beliefs based on the observed data. The result of this multiplication is a probability distribution representing the likely values of the parameter of interest known as the posterior distribution. In some cases, there is an analytical closed form solution when the prior distribution(s) and likelihood distribution are multiplied, but in many cases this result is intractable. Due to the computational intensity required to sample from this intractable distribution, early practitioners were largely unable to explore Bayesian methods. Recent developments in probabilistic programming methods and more efficient sampling algorithms have made simulating intractable posterior distributions possible. The use of Bayesian methods to directly predict daily stock returns has largely failed because the approximations converge to their frequentist counterparts. Alternatively, the Bayesian framework is more directly applicable to modeling the volatility of equity returns since the underlying distribution is allowed to change given new data. Below I derive five Bayesian Models and generate point predictions and credible intervals for the next days' realized volatility and implied volatility. For each model I calculate the directional accuracy of the point prediction, the root mean squared error term of the point prediction, and the percent of the time the credible interval capture the next days' observed volatility. The models are run on all 487 stocks in the sample space and the sector specific cross-sectional results are compared.

3.3.1 Inverse Gamma, Normal Conjugate Pair

The first model looks to capture the posterior distribution of daily returns. The scale parameter (standard deviation) of the resulting posterior distribution is then used to predict the next day's volatility. This model employs the assumption that the mean of historical daily returns can be used as a reasonable estimate for the location parameter, meaning there is only need for one prior distribution. To begin, the Inverse Gamma distribution is chosen as the prior approximation for the scale parameter and the Gaussian distribution as the likelihood function.

This results in a conjugate pair posterior with a closed form solution that is equivalent to an Inverse Gamma distribution where the alpha and beta terms are calculated as follows, where X represents daily return data.

$$(X|\sigma^2)\pi(\sigma^2|\alpha,\beta) \; = \; (\sigma^2|X_1\ldots X_n) \; = \; IG(\alpha',\beta')$$

$$\pi(\sigma^2|\alpha,\beta) \; = \; \frac{\beta^-}{r(\alpha)}(\sigma^2)^{-(\alpha+1)}e^{\frac{-\beta}{\sigma^2}}$$

$$X|\sigma^2 \sim N(\mu,\sigma^2) \rightarrow \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

$$\alpha' = \alpha \; + \; \frac{n}{2}$$

$$\beta' = \beta \; + \; \frac{1}{2}\sum(X_i-\mu^2)$$

Note that the inverse gamma distribution has two parameters, alpha and beta, that need to be defined for the prior distribution and clearly have an impact on the resulting posterior distribution as evident in the formula above. In this model, the prior distribution represents the prior assumption about the distribution of the variance of daily returns, since this distribution is unknown four uninformed prior specifications were used for alpha and beta (1,1), (2,1), (3,.05), (3,1). The resulting assumed prior assumptions are displayed below.

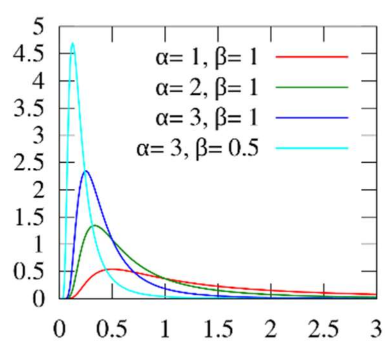

Figure 1: Inverse Gamma Distribution

Recall that the closed-form solution for the posterior represents the expected distribution of the variance of stock returns, where the mean of the resulting distribution can be interpreted as our point estimate for future volatility and the standard deviation of the distribution is a reflection of our confidence in the point estimate. In other words, a wider posterior distribution implies that the volatility of stock volatility is increasing. Accordingly, I used the expected value of the posterior distribution calculated at the close of each trading day plus/minus the variance of the posterior distribution to create the prediction interval for future stock volatility…

$$\mathrm{E[X]} = \frac{\beta'}{(\alpha'\text{-}1)}$$

$$\mathrm{Var[X]} = \frac{(\beta')^2}{(\alpha'\text{-}1)^2(\alpha'\text{-}2)}$$

$$(\mathrm{E[X]}\text{-}\mathrm{Var[X]}, \mathrm{E[X]} + \mathrm{Var[X]})$$

Note that unlike the normal distribution which has a fixed ratio between the area under its cumulative density function and the number of standard deviations an observation is from the mean, the variance of the Inverse Gamma distribution derived above is contingent on the values of alpha and beta, which in turn are contingent on the training sample size and sample values. This gives our interval the unique characteristic that its length dynamically updates as the intra-week variance rises, without the need for arbitrary weightings on more recent data points. Unlike the other four models below, there is a closed form solution for this model making the results computationally more efficient than the other Bayesian models. For each of the four prior specifications, models were trained on the previous 5,7,10,15, and 30 days to predict the next days realized and implied volatility. The results are analyzed in section four.

3.3.2 Bayesian Models Intractable Posterior

A second application of the Bayesian Framework is to directly model the distribution of volatility by feeding in historical realized volatility as the observed data. For this modeling schema we need to figure out a likelihood function and prior distributions that sensibly represent volatility based on our previously understanding of (or lack thereof) the characteristics of volatility. First, volatility (implied and realized volatility) cannot be negative thus the likelihood distribution should not have any probable mass over regions less than zero. Second, generally speaking, periods of extremely high volatility occur less frequently than periods of moderate/low volatility. So, the likelihood function should be concave. From these assumptions, the gamma distribution appears to be a reasonable candidate for the likelihood function.

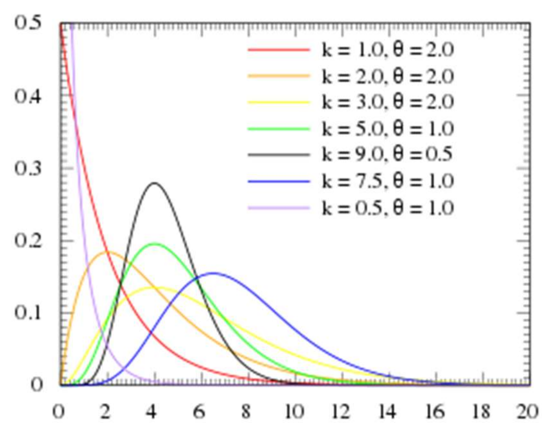

Figure 2 Gamma Distribution

The Gamma distribution has two parameters, a location and scale parameter which both need to be assigned a prior distribution. These hyper parameters have less impact on the model than the likelihood distribution thus uninformed prior assignments are totally acceptable. However, the parameters must be positive and the range of the location parameter is most likely somewhere between zero and one. An implied volatility larger than one implies that the market expects more than a 100% move in the underlying asset, although this is possible, it is reasonable to assume that this is a tail occurrence and the center of the posterior distribution should be less than this. Given these two specifications, the obvious choices for the prior distributions are the Uniform and Exponential distributions. Note a key different between the two is that the Exponential prior attributes a small but decreasing amount of probability mass to all values up to infinity whereas the Uniform prior constrains this range of values. Both of these are reasonable assumptions to make for the location and scale parameters of volatility. Accordingly, I created four models with the following specifications; Gamma likelihood function with a Uniform prior for alpha and an Exponential prior for beta, Gamma likelihood function with a Uniform prior for alpha and a Uniform prior for beta, Gamma likelihood function with an Exponential prior for alpha and an Exponential prior for beta. Finally, a naive model was included with a Log-normal likelihood function with a Uniform prior for alpha and an Exponential prior for beta to be used as a benchmark. Unlike the models in the previous section the resulting posteriors are intractable matrix multiplications only a mathematician could love. Sampling from these distributions is too computationally expensive for most Markov Chain Monte Carlo methods. Accordingly, the NUTS (No-U-Turn) sampler was used. The algorithm uses Hamiltonian dynamics to explore the target distribution by tuning the step size of the MCMC proposal distribution. The python package PyMc3, build on top of optimizing compiler Theano, was used to design the models and sample from the intractable posteriors. A more detailed description of the tuning parameters for the sampling algorithm and an analysis of the results will be discussed in section five.

3.4 Time Series Models

As outlined in the literature review section, time series models are particularly useful in capturing the volatility of financial assets by making the assumption that the conditional variance of a time series changes over time. A defining feature shared by all GARCH type models is that they allow for the conditional variances of predictions and observations to correlate. In this analysis, five time series models were compared; GARCH, FIGARCH, APARCH, HARCH, and EGARCH. Each model is implemented in two ways. The first approach models the series of daily returns given the last 100 days of data to predict the next day's daily return. The variance estimation of the prediction is annualized and then used as the prediction for the next day's volatility. The second approach is to directly predict the next day's volatility by training the model on the historical volatility from the past 100 days. Both of these implantations were used on all five of the models outlined below. All of the models are implemented using Python's ARCH package. The results of the models are discussed in the following results section. The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model makes the assumption that the conditional variance of a time series of is a function of its past values and variances with two parts; a moving average and autoregressive component.

$$\mathrm{X_t} = \mu \ + \ \gamma\sigma_{\mathrm{t}} \ + \alpha_{\mathrm{t}}$$

$$\sigma_{\mathrm{t}}^2 = \omega + \sum_{\mathrm{i=1}}^{\mathrm{p}} \alpha_{\mathrm{i}}\varepsilon_{\mathrm{t\text{-}i}}^2 + \sum_{\mathrm{j=1}}^{\mathrm{q}} \beta_{\mathrm{j}}\sigma_{\mathrm{j\text{-}i}}^2$$

The GARCH model assumes that the conditional variance is stationarity, leading many to characterize it as a 'short memory' model. The FIGARCH (Fractionally Integrated GARCH) model is an extension of the GARCH model where the integration order of the conditional variance can be any real number between 0 and 1.

$$\mathrm{X_t} = \varepsilon_{\mathrm{t}}\sqrt{\sigma_{\mathrm{t}}}$$

$$\sigma_{\mathrm{t}} = \frac{\gamma}{\beta(1)\omega} + \{1\text{-}\frac{\alpha(\mathrm{B})}{\beta(\mathrm{B})}(1\text{-B})^{\mathrm{d}}\}\mathrm{y_t^2}$$

A value of 0 would imply that past variances have no effect on future volatility prediction whereas a value of 1 implied that the process has a unit root and is nonstationary. This feature gives the model 'Long Memory' because the impact of volatility spikes slowly decays over time. Theoretically, this should better capture the volatility clustering effects observed in historical stock data. The APARCH (Asymmetric Power ARCH) model is another extension of the GARCH model, however unlike the GARCH model the impacts of positive and negative shocks do not have symmetric impacts on volatility.

$$\varepsilon_{\mathrm{t}} = \sigma_{\mathrm{t}}\eta_{\mathrm{t}}$$

$$\sigma_{\mathrm{t}}^{\delta} = \omega + \sum_{\mathrm{i}=1}^{\infty} \alpha_{\mathrm{i}}^{+}|\mathrm{e_{t\text{-}i}}|^{\delta}\, 1_{\mathrm{e_{t\text{-}i}}\geq 0} + \alpha_{\mathrm{i}}^{-}|\mathrm{e_{t\text{-}i}}|^{\delta} 1_{\mathrm{e_{t\text{-}i}}<0}$$

The model accomplishes this by incorporating different exponents for positive and negative shocks in the volatility equation. Theoretically, this should capture the disproportionate effect on stock volatility during negative shocks. The HARCH (Hyperbolic ARCH) model is another extension of the GARCH model that attempts to address the long memory of stock returns similar to the FIGARCH model. It accomplishes this by allowing the conditional variance to be affected by the past squared error terms over multiple lags.

$$\mathrm{X_t} = \varepsilon_{\mathrm{t}}\sqrt{\sigma_{\mathrm{t}}}$$

$$\sigma_{\mathrm{t}} = \frac{\gamma}{\beta(1)} + \{1\text{-}\frac{\delta_{\mathrm{h}}(\mathrm{B})}{\beta(\mathrm{B})}[1\text{-}\phi + \phi(1\text{-B})^{\mathrm{d}}]\}\mathrm{y}_{\mathrm{t}}^{2}$$

The contribution of each lagged period decreases by a hyperbolic weighting function. This is another approach which theoretically should capture volatility clustering however the increased sophistication to the residual effect of past volatility spikes could allow for increased responsiveness. The fifth and final model used in this analysis is the EGARCH (Exponential GARCH), another extension of the GARCH model. The EGARCH model uses an exponential function of past error terms and the absolute value of past error terms to predict the conditional variance of future returns, this ensures that the conditional variance is always positive.

$$\log(\sigma_{\mathrm{t}}^{2}) = \omega \; + \sum_{\mathrm{k}=1}^{\mathrm{q}} \beta_{\mathrm{k}}\mathrm{g}(\mathrm{Z_{t\text{-}k}}) + \sum_{\mathrm{k}=1}^{\mathrm{p}} \alpha_{\mathrm{k}}\log\sigma_{\mathrm{t\text{-}k}}^{2}$$

Unlike other GARCH type models, the EGARCH model can also include other variables in the conditional variance equation, such as lagged returns or other economic indicators. Similar to the APACH model, the goal of the EGARCH model is to capture the asymmetric effect of volatility spikes.

4. Results

4.1 Bayesian Conjugate Pair Results

Recall from the preliminaries section that the first Bayesian model generated predictions using the Inverse Gamma and Normal conjugate pair. The models were run on all of the tickers in the dataset to serve as a bench mark for the other Bayesian models specified below. There were two variables altered in this modeling schema. The first was the number of days feed in as observations and the second was the alpha and beta values assigned to the gamma prior distribution. Each table in the following three images represents a different alpha and beta pairing, and each column represents the number of observations feed into the iteration of the model predicting implied volatility. Starting with the 95% credible intervals generated by the models, The prior assignments of (1,1) for alpha and beta with a five-day training period was the only model that accurately captured the next days implied volatility more than 95% of the time. This is likely due to the length of the credible interval which explains why the confidence interval accuracy decreases as a function of the number of training days.

Alpha = 1, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.981812	0.900645	0.627890	0.171137	0.000063
std	0.033469	0.131700	0.236264	0.171695	0.000626
min	0.553254	0.000000	0.000000	0.000000	0.000000
25%	0.982964	0.886396	0.498296	0.018171	0.000000
50%	0.991482	0.953436	0.680295	0.119672	0.000000
75%	0.994889	0.973886	0.806644	0.281658	0.000000
max	1.000000	1.000000	0.964225	0.714367	0.011357

Alpha = 2, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.900373	0.733971	0.415301	0.077723	0.000030
std	0.131988	0.214393	0.237067	0.109808	0.000392
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.886112	0.625674	0.237081	0.000568	0.000000
50%	0.953436	0.801817	0.423623	0.022714	0.000000
75%	0.973878	0.896934	0.591993	0.123225	0.000000
max	1.000000	0.976136	0.913685	0.582624	0.006814

Alpha = 3, Beta = 0.5

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.110561	0.018281	0.000710	0.000011	0.000011
std	0.135417	0.043525	0.005182	0.000186	0.000233
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.004543	0.000000	0.000000	0.000000	0.000000
50%	0.048268	0.000000	0.000000	0.000000	0.000000
75%	0.173197	0.011925	0.000000	0.000000	0.000000
max	0.630892	0.370812	0.080636	0.003975	0.005150

Alpha = 3, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.733619	0.518881	0.239269	0.029985	0.000022
std	0.214545	0.242367	0.200116	0.059785	0.000313
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.625674	0.364566	0.057638	0.000000	0.000000
50%	0.801249	0.547416	0.196479	0.001136	0.000000
75%	0.896934	0.704797	0.383873	0.030664	0.000000
max	0.976136	0.955139	0.784781	0.433277	0.005150

Figure 3 Conjugate Pair Credible Interval Results for Implied Volatility

Moving on to the directional predictions of implied volatility, there is no model specification that consistently outperforms the others. The highest directional accuracy achieved is 53.08% which is outperformed by the models outlined in the literature review but sets a good benchmark for the Intractable Bayesian directional prediction results outlined in the next section.

Alpha = 1, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.506615	0.524456	0.530035	0.516474	0.512830
std	0.020397	0.018221	0.018162	0.017810	0.015834
min	0.461670	0.471323	0.470187	0.465645	0.465645
25%	0.491766	0.511641	0.519023	0.504827	0.502555
50%	0.505963	0.523566	0.529813	0.516752	0.513345
75%	0.519171	0.536852	0.541761	0.529529	0.524039
max	0.647815	0.574074	0.586031	0.561045	0.561045

Alpha = 2, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.524458	0.530822	0.524655	0.514029	0.512826
std	0.018170	0.017434	0.019464	0.016570	0.015833
min	0.471323	0.477570	0.465645	0.465645	0.465645
25%	0.511641	0.519591	0.511364	0.503123	0.502555
50%	0.523566	0.530380	0.524702	0.514480	0.513345
75%	0.536627	0.542306	0.538898	0.525838	0.524039
max	0.572970	0.578648	0.578409	0.561045	0.561045

Alpha = 3, Beta = 0.5

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.513528	0.512900	0.512842	0.512832	0.512823
std	0.016312	0.015878	0.015823	0.015820	0.015833
min	0.465645	0.465645	0.465645	0.465645	0.465645
25%	0.502555	0.502555	0.502555	0.502555	0.502555
50%	0.513913	0.513345	0.513345	0.513345	0.513345
75%	0.525270	0.524039	0.524039	0.524039	0.524039
max	0.561045	0.561045	0.561045	0.561045	0.561045

Alpha = 3, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.530835	0.528000	0.518962	0.513211	0.512825
std	0.017446	0.019123	0.018903	0.016102	0.015833
min	0.477570	0.465645	0.465645	0.465645	0.465645
25%	0.519591	0.515048	0.505922	0.502555	0.502555
50%	0.530380	0.528977	0.519591	0.513913	0.513345
75%	0.542306	0.541170	0.533147	0.524702	0.524039
max	0.578648	0.592277	0.566156	0.561045	0.561045

Figure 4 Conjugate Pair Directional Predictor Results for Implied Volatility

Lastly, the Root Mean Squared Errors (RSME) of the point predictions are calculate for implied volatility. The seven-day models with the prior specification (1,1) and (2,1) and the five-day model with prior specification (1,2) are the most accurate, and also have notably lower standard deviations than the predictions from the other models.

Alpha = 1, Beta = 1

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.179845	0.141836	0.164318	0.208844	0.264400
std	0.039360	0.060798	0.079980	0.086002	0.087363
min	0.068896	0.050044	0.065394	0.076904	0.124223
25%	0.155605	0.107820	0.111044	0.150312	0.204577
50%	0.173826	0.120794	0.139820	0.188728	0.244465
75%	0.199454	0.151925	0.198795	0.249291	0.304333
max	0.419364	0.520106	0.599005	0.661965	0.725771

Alpha = 1, Beta = 2

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.141995	0.154112	0.184645	0.221330	0.268094
std	0.060944	0.076223	0.083955	0.086604	0.087400
min	0.050011	0.066493	0.065606	0.085773	0.127765
25%	0.107873	0.102218	0.128257	0.162005	0.208285
50%	0.120934	0.128496	0.162739	0.201076	0.248385
75%	0.152076	0.184404	0.223430	0.261731	0.308027
max	0.520928	0.578864	0.630544	0.677007	0.729825

Alpha = 3, Beta = 0.5

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.226862	0.243650	0.260358	0.276862	0.297674
std	0.086908	0.087244	0.087397	0.087440	0.087407
min	0.090081	0.104662	0.120272	0.136282	0.156923
25%	0.167324	0.183972	0.200498	0.217160	0.238361
50%	0.206586	0.223542	0.240416	0.256988	0.277702
75%	0.267481	0.284481	0.300374	0.317116	0.337944
max	0.683634	0.702508	0.720820	0.738704	0.760953

Alpha = 3, Beta = 1

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.154265	0.174978	0.201756	0.231516	0.271366
std	0.076325	0.082519	0.085642	0.086945	0.087429
min	0.066529	0.064550	0.072724	0.094045	0.130922
25%	0.102324	0.120213	0.143642	0.171923	0.211571
50%	0.128577	0.151646	0.181216	0.211330	0.251549
75%	0.184534	0.211582	0.241969	0.272368	0.311354
max	0.579502	0.616796	0.653204	0.688910	0.733404

Figure 5 Conjugate Pair RMSE for Implied Volatility

Next, these same model specifications were used to predict point estimates and credible intervals for realized volatility. The performance of the credible intervals for all model specifications were notably worse when attempting to capture realized volatility, however, the prior specification (1,1) trained on the previous five days of data was still the best performing interval.

Alpha = 1, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.941113	0.793818	0.606962	0.335605	0.038944
std	0.033019	0.078350	0.130354	0.140258	0.032881
min	0.610454	0.221893	0.038462	0.000000	0.000000
25%	0.932425	0.766610	0.562748	0.235662	0.014764
50%	0.947826	0.813742	0.646792	0.342987	0.030664
75%	0.959682	0.839580	0.695739	0.441794	0.055082
max	1.000000	0.920308	0.782946	0.603634	0.188529

Alpha = 2, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.793583	0.668922	0.487934	0.255981	0.029513
std	0.078519	0.116233	0.145586	0.125971	0.026035
min	0.221893	0.054945	0.014793	0.000000	0.000000
25%	0.766326	0.639125	0.409710	0.157297	0.009654
50%	0.813174	0.704017	0.521863	0.254401	0.022147
75%	0.839296	0.745104	0.590006	0.353208	0.040886
max	0.920308	0.844961	0.724588	0.572970	0.147076

Alpha = 3, Beta = 0.5

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.312256	0.181326	0.078037	0.020667	0.000998
std	0.151976	0.110238	0.058908	0.019851	0.002392
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.197615	0.094832	0.033759	0.006246	0.000000
50%	0.307780	0.168086	0.064736	0.015900	0.000000
75%	0.431005	0.260080	0.111584	0.029529	0.001136
max	0.683135	0.508234	0.304940	0.123225	0.041454

Alpha = 3, Beta = 1

	5_Day_CI	7_Day_CI	10_Day_CI	15_Day_CI	30_Day_CI
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.668687	0.546036	0.381723	0.193131	0.022597
std	0.116349	0.140255	0.144528	0.107883	0.020639
min	0.054945	0.019724	0.004931	0.000000	0.000000
25%	0.639125	0.478137	0.281454	0.110417	0.007098
50%	0.704713	0.581488	0.398295	0.185122	0.017036
75%	0.744820	0.644703	0.489639	0.271754	0.032084
max	0.844961	0.760931	0.651334	0.491766	0.116411

Figure 6 Conjugate Pair Credible Interval Results for Realized Volatility

The directional predictions of realized volatility were slightly better with five of the models achieving 54% accuracy including two from the prior specification (3,1), a substantial improvement from the implied volatility predictions. Finally, the RMSE values were worse for all models when used to predict realized volatility compared to implied volatility.

Alpha = 1, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.518370	0.533329	0.541067	0.532549	0.503463
std	0.018401	0.016548	0.014396	0.016766	0.010622
min	0.478123	0.486087	0.489495	0.474162	0.469555
25%	0.505395	0.522661	0.532652	0.521295	0.496309
50%	0.516184	0.533220	0.541488	0.533220	0.503123
75%	0.529813	0.543441	0.549688	0.544009	0.509938
max	0.609865	0.623457	0.651163	0.578081	0.540034

Alpha = 2, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.533350	0.540268	0.539268	0.526883	0.502122
std	0.016547	0.014456	0.014987	0.016795	0.010197
min	0.486087	0.495741	0.480977	0.475866	0.470187
25%	0.522661	0.531516	0.530664	0.514617	0.495173
50%	0.533220	0.540034	0.540034	0.527541	0.502555
75%	0.543312	0.548552	0.549532	0.539194	0.508234
max	0.623457	0.651163	0.582056	0.576945	0.537037

Alpha = 3, Beta = 0.5

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.524025	0.514166	0.505387	0.499414	0.496835
std	0.016396	0.014129	0.011339	0.009484	0.009217
min	0.475298	0.471891	0.470187	0.466042	0.462529
25%	0.511641	0.504259	0.497601	0.492902	0.490062
50%	0.524134	0.514480	0.504827	0.499148	0.496877
75%	0.536470	0.522863	0.512493	0.505395	0.502555
max	0.566156	0.551391	0.544009	0.537037	0.537037

Alpha = 3, Beta = 1

	5_Day_EV	7_Day_EV	10_Day_EV	15_Day_EV	30_Day_EV
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.540304	0.540793	0.535069	0.521403	0.500949
std	0.014455	0.014418	0.016416	0.016008	0.009887
min	0.495741	0.484384	0.477002	0.475298	0.469555
25%	0.531516	0.531516	0.524702	0.508802	0.494605
50%	0.540602	0.542306	0.535491	0.521863	0.500852
75%	0.548552	0.549688	0.547416	0.532652	0.506599
max	0.651163	0.596899	0.582624	0.560477	0.537037

Figure 7 Conjugate Pair Directional Predictor Results for Realized Volatility

For the first set of Bayesian Models there are three main takeaways, the conjugate pair generally made more accurate predictions for implied volatility than realized volatility. The conjugate pair made better credible intervals for realized volatility than it did for implied volatility. Lastly, the shorter training sets overall lead to much more accurate predictions with 30-day training windows having the worst results. This first class of Bayesian models will be referenced as a bench mark for the remaining models.

Alpha = 1, Beta = 1

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.245389	0.207364	0.215732	0.245324	0.289094
std	0.047340	0.069205	0.086021	0.093329	0.095898
min	0.158069	0.115256	0.100192	0.100002	0.137312
25%	0.220016	0.165345	0.155742	0.180522	0.223529
50%	0.236108	0.183268	0.189639	0.218487	0.262191
75%	0.254717	0.223825	0.248366	0.286027	0.333662
max	0.648451	0.638826	0.644949	0.665374	0.724537

Alpha = 2, Beta = 1

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.207523	0.210272	0.228421	0.254661	0.292178
std	0.069336	0.082392	0.090450	0.094303	0.095974
min	0.115236	0.109667	0.094399	0.106239	0.140376
25%	0.165410	0.154005	0.165546	0.188680	0.226624
50%	0.183399	0.183887	0.201193	0.227783	0.265379
75%	0.223940	0.238386	0.265720	0.295972	0.336941
max	0.638892	0.642320	0.650414	0.679235	0.728321

Alpha = 3, Beta = 0.5

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.258886	0.272085	0.285690	0.299509	0.317385
std	0.094706	0.095425	0.095853	0.096078	0.096165
min	0.109429	0.120893	0.133916	0.147833	0.166362
25%	0.192750	0.206279	0.220108	0.233998	0.251685
50%	0.231969	0.245071	0.258690	0.272954	0.291437
75%	0.300786	0.315420	0.330097	0.344718	0.363124
max	0.685215	0.702696	0.719735	0.736495	0.757426

Alpha = 3, Beta = 1

	5_Day_RMSE	7_Day_RMSE	10_Day_RMSE	15_Day_RMSE	30_Day_RMSE
count	487.000000	487.000000	487.000000	487.000000	487.000000
mean	0.210422	0.222229	0.240232	0.262523	0.294925
std	0.082489	0.088724	0.092715	0.094904	0.096034
min	0.109671	0.095664	0.097305	0.112475	0.143126
25%	0.154082	0.160576	0.176144	0.196374	0.229366
50%	0.184036	0.195518	0.213260	0.235582	0.268281
75%	0.238623	0.257569	0.280829	0.304848	0.339851
max	0.642380	0.647790	0.657293	0.690241	0.731664

Figure 8 Conjugate Pair RMSE for Realized Volatility

4.2 Bayesian Intractable Posterior Results

For the Bayesian models with intractable posterior distributions, the process of generating results was notably different from the conjugate pair models. Unlike the other models, generating enough samples to accurately depict the posterior distribution was too computationally expensive to be done for each day in the back test. Additionally, since 100 previous datapoints were used as the observations for each prediction, changing only a fraction of the observations and then sampling from the posterior would not have significantly changed the resulting posterior samples. Accordingly, every ticker was labeled with a list of dates approximately 100 trading days apart. Two of the dates were chosen at random for each of the tickers and the models were built on the previous 100 days. During the modeling phase, 500 iterators of prior parameter tuning with a NUTS sampler was used to create a trace vector. The trace vector was then used to generate 100,000 samples from the posterior distribution for each prediction. A 10,000-sample burn-in period was implemented and the resulting 90,000 samples were used to represent the intractable posterior distribution. From these remaining samples two intervals and two point estimators were created. An 85% credible interval was used by looking at the ordered percentiles of the samples and a novel 'skew interval' was created by taking three times the distance between the mean and the median then centering it over the posterior distribution. Secondly, both the mean and the median were used as point estimates and the directional accuracy and RMSE was calculated for each. The analysis was run on all of the tickers in the data set then replicated on three cross sectional slices of the data based on historical sector volatility. Information Technology, Consumer Discretionary, Health Care, Energy, and Communication Services were classified as 'High Volatility Sectors', Materials, Industrials, and Financials as 'Medium Volatility Sectors', and Utilities, Real Estate, and Consumer Staples as 'Low Volatility Sectors'. Finally, the models were used to predict both realized and implied volatility. The results are displayed below.

All Tickers

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.938080	0.841073	0.862745	0.668731	0.293413	0.321763
1	Gamma(Uniform_Uniform)	0.983488	0.963880	0.758514	0.829721	0.471844	0.315584
2	Gamma(Exponential_Exponential)	0.912281	0.514964	0.863777	0.834881	0.293336	0.309542
3	LogNormal(Uniform_Exponential)	0.865841	0.934985	0.725490	0.756450	7.770775	0.800073

High Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.948718	0.839744	0.833333	0.675214	0.257950	0.281823
1	Gamma(Uniform_Uniform)	0.989316	0.974359	0.745726	0.807692	0.467212	0.281718
2	Gamma(Exponential_Exponential)	0.910256	0.487179	0.833333	0.818376	0.258006	0.268151
3	LogNormal(Uniform_Exponential)	0.835470	0.901709	0.735043	0.745726	6.687688	0.780439

Medium Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.929878	0.838415	0.902439	0.661585	0.334487	0.368636
1	Gamma(Uniform_Uniform)	0.978659	0.948171	0.783537	0.871951	0.484183	0.354126
2	Gamma(Exponential_Exponential)	0.914634	0.545732	0.905488	0.856707	0.334384	0.354913
3	LogNormal(Uniform_Exponential)	0.884146	0.957317	0.722561	0.777439	6.676078	0.804130

Low Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.924855	0.849711	0.867052	0.664740	0.300148	0.327710
1	Gamma(Uniform_Uniform)	0.976879	0.965318	0.745665	0.809249	0.460475	0.324614
2	Gamma(Exponential_Exponential)	0.913295	0.531792	0.867052	0.838150	0.299816	0.321470
3	LogNormal(Uniform_Exponential)	0.913295	0.982659	0.705202	0.745665	11.520943	0.843639

Figure 9 Intractable Bayesian Results for Realized Volatility

Starting with the results for realized volatility, the Gamma distribution with Uniform priors generated the best credible intervals of all the Bayesian models capturing the next day's realized volatility over 98.00% of the time. This result hold across all cross-sectional volatility breakdowns as well, notably outperforming all conjugate pair models. All three models which use the gamma distribution as the likelihood function have at least one point estimate that correctly predicts the directional change of realized volatility over 80.00% of the time. This is arguably the most significant result from the analysis. The mean and median estimators of the model with a Gamma likelihood function and exponential priors, are the most accurate at predicting directional changes in realized volatility for all cross-sectional breakdowns. The mean and median estimators for this model capture directional changes for all stocks

86% and 83% of the time respectively, 83% and 81% of the time for high volatility sectors, 91% and 86% of the time for medium volatility sectors, and 87% and 84% of the time for low volatility sectors. As outlined in the literature review section, some of the best preforming published directional indicators were only able to capture the correct change in volatility ~75% of the time. Accordingly, expanding on the use of these posterior estimators as directional indicators will be a topic of future work. Finally, the RMSE of the estimators are calculated for each model. Note that models with the Log Normal likelihood functions are intentionally inaccurate (it allows volatility predictions to be negative) to contextualize what bad results look like. Accordingly, the root mean squared error values for the models with a Gamma likelihood function are relatively accurate as point estimators compared to the intentionally uninformed models varying by as little as 15 basis points on average. In summation, the use of intractable Bayesian models appears to be exceptionally good at modeling realized volatility. Another notable feature of the results is the success of the Mean Median interval for the model with a Gamma likelihood function and Uniform priors. Given that this range is based only on the distance between the mean and the median, this interval dynamically adapts to changes in the skew of the volatility distribution. Additionally, it will almost always be shorter than the 85% credible interval. Despite its shorter length, the interval captures realized volatility 96.3% of the time when run all tickers, 97.4% for high volatility stocks, 94.8% for medium volatility stocks, and 96.5% for low volatility. Given that this is significantly higher than the models which use an exponential prior, using this interval as a trading range for volatility will also be a topic of future research.

All Tickers

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.913313	0.808050	0.466460	0.665635	0.286908	0.326675
1	Gamma(Uniform_Uniform)	0.970072	0.921569	0.299278	0.351909	0.434633	0.300130
2	Gamma(Exponential_Exponential)	0.787410	0.269350	0.464396	0.540764	0.286869	0.306638
3	LogNormal(Uniform_Exponential)	0.538700	0.926729	0.318885	0.308566	7.739948	0.757874

High Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.923077	0.809829	0.487179	0.653846	0.294327	0.332763
1	Gamma(Uniform_Uniform)	0.985043	0.931624	0.324786	0.399573	0.448290	0.307149
2	Gamma(Exponential_Exponential)	0.779915	0.269231	0.489316	0.542735	0.294447	0.310374
3	LogNormal(Uniform_Exponential)	0.534188	0.888889	0.337607	0.329060	6.651224	0.744328

Medium Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.896341	0.798780	0.454268	0.667683	0.307367	0.350636
1	Gamma(Uniform_Uniform)	0.960366	0.902439	0.289634	0.320122	0.437918	0.319673
2	Gamma(Exponential_Exponential)	0.789634	0.265244	0.442073	0.539634	0.307214	0.330943
3	LogNormal(Uniform_Exponential)	0.548780	0.948171	0.307927	0.298780	6.648797	0.758297

Low Volatility Sectors

	Model	Credible_Interval	Mean_Median_Interval	Mean_Direction_Acc	Median_Direction_Acc	Mean_RMSE	Median_RMSE
0	Gamma(Uniform_Exponential)	0.919075	0.820809	0.433526	0.693642	0.218175	0.255122
1	Gamma(Uniform_Uniform)	0.947977	0.930636	0.248555	0.283237	0.388397	0.235759
2	Gamma(Exponential_Exponential)	0.803468	0.277457	0.439306	0.537572	0.217859	0.241681
3	LogNormal(Uniform_Exponential)	0.531792	0.988439	0.289017	0.271676	11.491690	0.792597

Figure 10 Intractable Bayesian Results for Implied Volatility

Looking at the results of the models for predicting implied volatility, there is similar success in creating credible interval across all sector volatility profiles. However, the accuracy of the directional predictions for implied volatility are significantly worse than the results from the realized volatility analysis. This is most likely attributed to fact that the modes are trained on rolling realized volatility, which occasionally diverges from implied volatility. Given that the observed values greatly influence the shape and skew of the posterior distribution this likely is the cause for the poor results. Given the success of the intractable models when predicting realized volatility, a topic of future work will be to re-run the back test by training the model on implied volatility. A notable setback of these models is how computationally expensive they are to generate and how much memory is required to store all the samples for each iteration of the back test. For this reason, future work to expand these models onto implied volatility data will require computational assistance.

4.3 Time Series Results

Recall that there are two general schemas for the five timeseries models. The first set of models were trained on daily returns and the variance of the prediction was used as the point estimate for future realized, these results are contained in the next three tables. The second set of models were trained on the past 100 days of realized volatility, and the value of the next day's prediction was used as the point estimate. These results are contained in the last three tables. As evident from the results of the Bayesian models, there is a notable deprecation in the accuracy of the models trained on realized volatility when predicting implied volatility. Due to this observation and the computational intensity of training the models every day in the back test period for each ticker, the results below only indicate the performance of predicting realized volatility. Due to the computational limitations of the resources available for this analysis, recreating the back test for implied volatility is reserved for future work. Staring with the model trained on daily returns, the FIGARCH model marginally outperforms the other models at creating confidence intervals for realized volatility with ~98% accuracy across all cross sections. Additionally, all of the models generated more accurate credible intervals on low volatility sectors.

All Tickers

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI	EGARCH_CI
count	486.000000	486.000000	486.000000	486.000000	486.000000
mean	0.979896	0.980746	0.979664	0.969126	0.830524
std	0.023839	0.023163	0.023167	0.022526	0.039364
min	0.489209	0.506680	0.500000	0.521069	0.581622
25%	0.974769	0.975904	0.975215	0.963236	0.806994
50%	0.981325	0.982530	0.981280	0.970155	0.832570
75%	0.987296	0.987952	0.986126	0.976506	0.856808
max	1.000000	1.000000	1.000000	1.000000	0.987654

High Volatility

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI	EGARCH_CI
count	233.000000	233.000000	233.000000	233.000000	233.000000
mean	0.978556	0.979476	0.978800	0.967327	0.822766
std	0.008544	0.008603	0.007923	0.009102	0.036858
min	0.945783	0.946956	0.941390	0.940171	0.696581
25%	0.974081	0.973494	0.974034	0.960843	0.801383
50%	0.978916	0.980120	0.978865	0.966847	0.822309
75%	0.984337	0.985542	0.984190	0.973494	0.847436
max	1.000000	1.000000	1.000000	0.991489	0.987654

Medium Volatility Sectors

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI	EGARCH_CI
count	165.000000	165.000000	165.000000	165.000000	165.000000
mean	0.979944	0.980937	0.979186	0.970319	0.837695
std	0.039059	0.037793	0.038103	0.036243	0.043671
min	0.489209	0.506680	0.500000	0.521069	0.581622
25%	0.978313	0.978326	0.978208	0.967431	0.817358
50%	0.983735	0.984337	0.982498	0.972892	0.843852
75%	0.987952	0.988554	0.986683	0.977724	0.864597
max	1.000000	1.000000	1.000000	1.000000	0.924841

Low Volatility Sectors

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI	EGARCH_CI
count	88.000000	88.000000	88.000000	88.000000	88.000000
mean	0.983355	0.983752	0.982847	0.971655	0.837622
std	0.008982	0.009363	0.008667	0.010663	0.033325
min	0.963253	0.962048	0.963120	0.950602	0.740168
25%	0.977108	0.976506	0.975845	0.963705	0.822586
50%	0.983740	0.985542	0.984616	0.973201	0.835229
75%	0.990512	0.990361	0.989377	0.978916	0.857622
max	0.998557	0.999398	0.998555	0.990361	0.917868

Figure 11 Time Series Confidence Intervals Generated from Daily Returns

Looking at the directional accuracy of the model predictions, the GARCH model was marginally better than all other models capturing the change in direction ~55% of the time. The GARCH model also had the lowest standard deviations in prediction results across all cross sections. Although these directional predictions are in line with the Bayesian benchmark models, the results are significantly worse than the previously outlined intractable Bayesian predictors. This further emphasis the utility of the Bayesian framework and the motivation for future research.

All Tickers

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction	EGARCH_Direction
count	486.000000	486.000000	486.000000	486.000000	486.000000
mean	0.549369	0.548939	0.546381	0.546286	0.531236
std	0.012570	0.012882	0.013401	0.012976	0.013669
min	0.515060	0.506329	0.509639	0.495726	0.482639
25%	0.542169	0.540964	0.537940	0.538554	0.522892
50%	0.548795	0.548825	0.545783	0.546386	0.531508
75%	0.556627	0.556024	0.553614	0.554217	0.539157
max	0.622951	0.639344	0.641975	0.606557	0.606557

High Volatility

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction	EGARCH_Direction
count	233.000000	233.000000	233.000000	233.000000	233.000000
mean	0.549245	0.548669	0.546063	0.545615	0.529886
std	0.012363	0.012779	0.013883	0.013616	0.013637
min	0.515060	0.506329	0.515371	0.495726	0.482639
25%	0.540964	0.540361	0.537349	0.536747	0.521687
50%	0.548795	0.548193	0.545181	0.546386	0.530120
75%	0.556627	0.556627	0.554217	0.553614	0.537349
max	0.592593	0.592593	0.641975	0.604938	0.604938

Medium Volatility Sectors

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction	EGARCH_Direction
count	165.000000	165.000000	165.000000	165.000000	165.000000
mean	0.551048	0.550561	0.547921	0.548806	0.534324
std	0.013021	0.013464	0.012890	0.012055	0.014037
min	0.516867	0.513855	0.509639	0.511149	0.494578
25%	0.543976	0.543373	0.540299	0.542169	0.526189
50%	0.550000	0.550000	0.546988	0.548193	0.534805
75%	0.559036	0.556669	0.555422	0.556024	0.542020
max	0.622951	0.639344	0.622951	0.606557	0.606557

Low Volatility Sectors

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction	EGARCH_Direction
count	88.000000	88.000000	88.000000	88.000000	88.000000
mean	0.546551	0.546614	0.544335	0.543335	0.529020
std	0.011840	0.011714	0.012843	0.012178	0.012096
min	0.516867	0.518072	0.513855	0.514311	0.493068
25%	0.539608	0.537801	0.537199	0.535542	0.521687
50%	0.547475	0.547289	0.544880	0.542470	0.529518
75%	0.553163	0.553163	0.552560	0.552410	0.537107
max	0.578644	0.572289	0.587302	0.569880	0.556999

Figure 12 Time Series Directional Predictor Generated from Daily Returns

Finally, the RMSE scores were calculated for each of the models, the results for the EGARCH and APARCH models were nonsensical and indicated that there was a convergence problem in the results. Although the directional accuracies and credible intervals for these two models appears in line with the other models, the RMSE scores indicate that the EGARCH and APARCH models not able to accurately capture the dynamics of the problem so they have been dropped from the results. Looking at the remaining three models, the RMSE scores indicate that the point estimates were notably closer to observed realized volatility than the intractable Bayesian models, with the GARCH and the FIGARCH models having the closest estimates. Based on all three results the GARCH and FIGARCH modes appear to create more precise point estimates than the Bayesian models, but they are less sensitive to changes in the skew of the distribution resulting in notably lower directional predictions for realized volatility.

All Tickers

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	486.000000	486.000000	486.000000
mean	0.127327	0.125714	0.163786
std	0.055799	0.055301	0.058951
min	0.056059	0.056119	0.082503
25%	0.093238	0.091531	0.126805
50%	0.114415	0.113809	0.148763
75%	0.144073	0.143979	0.186336
max	0.638602	0.638358	0.648526

High Volatility

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	233.000000	233.000000	233.000000
mean	0.144294	0.141998	0.179288
std	0.055010	0.054321	0.058214
min	0.067771	0.068888	0.085415
25%	0.106431	0.103580	0.138276
50%	0.130766	0.128424	0.165936
75%	0.165771	0.164611	0.202361
max	0.394875	0.396042	0.431125

Medium Volatility Sectors

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	165.000000	165.000000	165.000000
mean	0.118176	0.117214	0.155512
std	0.056601	0.056579	0.056985
min	0.060010	0.060547	0.082503
25%	0.091069	0.091197	0.128128
50%	0.109144	0.108391	0.144740
75%	0.123906	0.124727	0.169312
max	0.638602	0.638358	0.648526

Low Volatility Sectors

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	88.000000	88.000000	88.000000
mean	0.099561	0.098536	0.138254
std	0.039227	0.039387	0.052744
min	0.056059	0.056119	0.086575
25%	0.075589	0.074611	0.109186
50%	0.091349	0.089203	0.122418
75%	0.113172	0.112505	0.151724
max	0.346087	0.347084	0.514268

Figure 13 Time Series RMSE Generated from Daily Returns

The five time series models were run again but alternatively trained on the realized volatility from the past 100 days. Surprisingly, the resulting confidence intervals were significantly less accurate than the confidence intervals from the previous modeling schema. Notably, when trained on past realized volatility the APARCH model generated the most accurate interval estimator. The EGARCH model performed even worse than in the previous results and was dropped from the results for conciseness.

All Tickers

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI
count	486.000000	486.000000	486.000000	486.000000
mean	0.889414	0.892147	0.893038	0.889291
std	0.025873	0.024842	0.024649	0.025948
min	0.806024	0.812048	0.816867	0.812048
25%	0.873645	0.877711	0.877108	0.872289
50%	0.890328	0.892771	0.894214	0.889759
75%	0.905422	0.906362	0.907831	0.906024
max	1.000000	1.000000	1.000000	1.000000

High Volatility Sectors

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI
count	233.000000	233.000000	233.000000	233.000000
mean	0.883451	0.886900	0.887838	0.883717
std	0.026546	0.025265	0.024832	0.026456
min	0.806024	0.812048	0.816867	0.812048
25%	0.865663	0.871084	0.872289	0.866265
50%	0.881254	0.884268	0.885542	0.881928
75%	0.900000	0.901205	0.903012	0.900000
max	0.962963	0.962963	0.950617	0.962963

Medium Volatility Sectors

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI
count	165.000000	165.000000	165.000000	165.000000
mean	0.896301	0.898278	0.899030	0.895725
std	0.025224	0.024811	0.024533	0.025784
min	0.842169	0.840964	0.839662	0.822182
25%	0.881928	0.884337	0.884198	0.882530
50%	0.894578	0.896988	0.898673	0.894578
75%	0.906627	0.907831	0.909910	0.908434
max	1.000000	1.000000	1.000000	1.000000

Low Volatility Sectors

	GARCH_CI	FIGARCH_CI	APARCH_CI	HARCH_CI
count	88.000000	88.000000	88.000000	88.000000
mean	0.892290	0.894542	0.895570	0.891985
std	0.021442	0.020662	0.021476	0.021560
min	0.836747	0.838554	0.837952	0.839759
25%	0.878765	0.882530	0.882078	0.879367
50%	0.894880	0.897590	0.898223	0.895783
75%	0.909173	0.908936	0.911596	0.908133
max	0.939157	0.938554	0.940325	0.939759

Figure 14 Time Series Confidence Intervals Generated from Historical Volatility

The directional predictions from all five models were similar to the previous results with the GARCH and FIGARCH models notably outperforming the APARH and HARCH models. However, the directional results were still significantly worse than the Bayesian predictions.

All Tickers

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction
count	486.000000	486.000000	486.000000	486.000000
mean	0.547093	0.547387	0.546672	0.546896
std	0.014938	0.015482	0.014330	0.014038
min	0.513855	0.516265	0.513889	0.510417
25%	0.538554	0.538554	0.537952	0.537952
50%	0.546386	0.546386	0.546084	0.545800
75%	0.554819	0.555422	0.554217	0.554217
max	0.737705	0.737705	0.721311	0.688525

High Volatility Sectors

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction
count	233.000000	233.000000	233.000000	233.000000
mean	0.546394	0.546544	0.545647	0.546554
std	0.012316	0.013322	0.012225	0.012680
min	0.517361	0.517361	0.513889	0.510417
25%	0.537952	0.537349	0.537349	0.537952
50%	0.545181	0.545783	0.545181	0.545783
75%	0.554819	0.554819	0.553614	0.553614
max	0.592593	0.629630	0.592593	0.592593

Medium Volatility Sectors

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction
count	165.000000	165.000000	165.000000	165.000000
mean	0.548579	0.548985	0.548293	0.547943
std	0.018826	0.018954	0.017665	0.016279
min	0.523494	0.520482	0.519880	0.518675
25%	0.539157	0.539759	0.539157	0.538554
50%	0.547590	0.548193	0.547261	0.546988
75%	0.555422	0.556627	0.556024	0.555422
max	0.737705	0.737705	0.721311	0.688525

Low Volatility Sectors

	GARCH_Direction	FIGARCH_Direction	APARCH_Direction	HARCH_Direction
count	88.000000	88.000000	88.000000	88.000000
mean	0.546158	0.546624	0.546344	0.545837
std	0.012842	0.013355	0.012253	0.012914
min	0.513855	0.516265	0.523494	0.514458
25%	0.538554	0.537952	0.537349	0.537349
50%	0.545181	0.545181	0.546084	0.545482
75%	0.554015	0.555422	0.553614	0.554289
max	0.590164	0.598361	0.598361	0.590164

Figure 15 Time Series Directional Predictor Generated from Historical Volatility

Once again, the EGARCH and APARCH models had nonsensical RMSE terms and were removed from analysis. The remaining three models all have higher RMSE than the previous set of models, implying that the timeseries models trained on daily returns were more accurate than those trained on past realized volatility. In the next section these finding are condensed and future works are discussed.

All Tickers

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	486.000000	486.000000	486.000000
mean	0.199517	0.202369	0.199930
std	0.061676	0.092929	0.062132
min	0.093401	0.092204	0.093401
25%	0.160396	0.159838	0.160613
50%	0.184755	0.184637	0.185305
75%	0.223148	0.223164	0.223440
max	0.667321	1.732980	0.674106

High Volatility Sectors

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	233.000000	233.000000	233.000000
mean	0.211949	0.218251	0.212189
std	0.063378	0.118063	0.063779
min	0.104494	0.104363	0.104602
25%	0.165573	0.165385	0.165474
50%	0.198416	0.198321	0.198975
75%	0.235143	0.234600	0.235440
max	0.441308	1.732980	0.441668

Medium Volatility Sectors

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	165.000000	165.000000	165.000000
mean	0.194518	0.194157	0.195105
std	0.058699	0.058693	0.059388
min	0.093401	0.092204	0.093401
25%	0.160542	0.160331	0.162164
50%	0.185014	0.184326	0.184490
75%	0.218270	0.218083	0.218497
max	0.667321	0.667723	0.674106

Low Volatility Sectors

	GARCH_RMSE	FIGARCH_RMSE	HARCH_RMSE
count	88.000000	88.000000	88.000000
mean	0.175971	0.175715	0.176521
std	0.054565	0.054372	0.054989
min	0.118580	0.118220	0.118842
25%	0.144648	0.145149	0.144412
50%	0.164699	0.164657	0.164513
75%	0.182849	0.183843	0.185555
max	0.529762	0.528036	0.531368

Figure 16 Time Series RMSE Generated from Historical Volatility

5. Conclusions

The goal of this analysis was to compare a variety of volatility models in the three most applicable settings to for asset managers and market makers. Each model was used to creating confidence interval for future movement, predict daily directional changes in volatility, and generate point estimates of future volatility for both realized and implied volatility. Unique to this analysis, probabilistic programming methods were used to developed Bayesian models with intractable posterior distributions, and in many of the evaluation criteria these had the best performing predictions. There are a few key takeaways. First, intractable Bayesian models assigned a Gamma likelihood function are able to predict directional changes of realized volatility between 80% and 90% of the time. This is significantly better than all other models in this analysis and all published directional predictors included in the literature review. Additionally, the use of the mean and median from the posterior distribution samples with a Gamma Likelihood function and Uniform prior assumptions were able to accurately capture the changing skew of the distribution of realized volatility. Secondly, timeseries models trained on daily returns data generate significantly better volatility predictions than those trained on past measures of realized volatility. Furthermore, the FIGARCH and the GARCH models make the most accurate point predictions of all the models analyzed. Finally, both Bayesian and timeseries models trained on historical realized volatility were much worse at predicting future implied volatility than predicting future realized volatility, indicating that realized volatility is not a strong predictor for implied volatility. A major limitation of this analysis was the computation and memory requirements for building and storing the modeling results. In future work this analysis will be extended to train all of the models on implied volatility. Specifically, intractable Bayesian model will be trained on implied volatility to see if the directional predictions are just as accurate as the realized volatility directional predictions. The magnitude of the directional estimator will also be incorporated to optimize the performance of the indicator. The results above demonstrate the ability of the Bayesian framework to outperform some of the most popular timeseries modes, and I hope this paper motivates further applications of these models in practice.

BIBLIOGRAPHY

[1] Katja Ahoniemi. Modeling and forecasting the vix index. *Available at SSRN 1033812*, 2008.

[2] Katja Ahoniemi. Multiplicative models for implied volatility. *Available at SSRN 967254*, 2008.

[3] Katja Ahoniemi and Markku Lanne. Joint modeling of call and put implied volatility. *International Journal of Forecasting*, 25(2):239–258, 2009.

[4] I Antoniou, Vi V Ivanov, Va V Ivanov, and PV Zrelov. On the log-normal distribution of stock market data. *Physica A: Statistical Mechanics and its Applications*, 331(3-4):617–638, 2004.

[5] David Ardia and Lennart F Hoogerheide. Bayesian estimation of the garch (1, 1) model with student-t innovations. *The R Journal*, 2(2):41–47, 2010.

[6] M Avellaneda, A Carelli, and F Stella. A bayesian approach for constructing implied volatility surfaces through neural networks. *Journal of Computational Finance*, 4(1):83–107, 2000.

[7] Gurdip Bakshi, Nikunj Kapadia, and Dilip Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1):101– 143, 2003.

[8] David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.

[9] Nicolas PB Bollen and Robert E Whaley. Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance*, 59(2):711–753, 2004.

[10] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[11] Tim Bollerslev and Hans Ole Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of econometrics*, 73(1):151–184, 1996.

[12] Peter Christoffersen, Kris Jacobs, and Karim Mimouni. Volatility dynamics for the s&p500: Evidence from realized volatility, daily returns, and option prices. *The Review of Financial Studies*, 23(8):3141–3189, 2010.

[13] Rama Cont. Volatility clustering in financial markets: empirical facts and agent-based models. *Long memory in economics*, pages 289–309, 2007.

[14] Jos´e Curto, Jos´e Castro Pinto, and Gonc¸alo Nuno Tavares. Modeling stock markets' volatility using garch models with normal, student's t and stable paretian distributions. *Modeling stock markets' volatility using GARCH models with normal, Student's t and stable Paretian distributions*, (2):311–321, 2009.

[15] Juliet G D'Cunha and K Aruna Rao. Bayesian inference for volatility of stock prices. *Journal of Modern Applied Statistical Methods*, 13(2):29, 2014.

[16] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[17] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.

[18] Stephen Figlewski. What goes into risk-neutral volatility? empirical estimates of risk and subjective risk preferences. *The Journal of Portfolio Management*, 43(1):29–42, 2016.

[19] Mark B Garman and Michael J Klass. On the estimation of security price volatilities from historical data. *Journal of business*, pages 67–78, 1980.

[20] Eric Ghysels, Pedro Santa-Clara, and Rossen Valkanov. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95, 2006.

[21] Xu Gong and Boqiang Lin. Modeling stock market volatility using new har-type models. *Physica A: Statistical Mechanics and its Applications*, 516:194–211, 2019.

[22] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *The Best of Wilmott*, 1:249–296, 2002.

[23] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[24] Ali Hirsa, Joerg Osterrieder, Branka Hadji Misheva, Wenxin Cao, Yiwen Fu, Hanze Sun, and Kin Wai Wong. The vix index under scrutiny of machine learning techniques and neural networks. *arXiv preprint arXiv:2102.02119*, 2021.

[25] James Hosker, Slobodan Djurdjevic, Hieu Nguyen, and Robert Slater. Improving vix futures forecasts using machine learning methods. *SMU Data Science Review*, 1(4):6, 2018.

[26] Darjus Hosszejni and Gregor Kastner. Approaches toward the bayesian estimation of the stochastic volatility model with leverage. In *Bayesian Statistics and New Generations: BAYSM 2018, Warwick, UK, July 2-3 Selected Contributions*, pages 75–83. Springer, 2019.

[27] G Andrew Karolyi. A bayesian approach to modeling stock return volatility for option valuation. *Journal of Financial and Quantitative Analysis*, 28(4):579–594, 1993.

[28] Markku Lanne and Katja Ahoniemi. Implied volatility with time-varying regime probabilities. 2008.

[29] Nikita Medvedev and Zhiguang Wang. Multistep forecast of the implied volatility surface using deep learning. *Journal of Futures Markets*, 42(4):645–667, 2022.

[30] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the econometric society*, pages 347–370, 1991.

[31] Curtis Nybo. Sector volatility prediction performance using garch models and artificial neural networks. *arXiv preprint arXiv:2110.09489*, 2021.

[32] Rosy Oh, Dong Wan Shin, and Man-Suk Oh. Bayesian analysis of financial volatilities addressing long-memory, conditional heteroscedasticity and skewed error distribution. *Communications for Statistical Applications and Methods*, 24(5):507–518, 2017.

[33] Oskar Oostdam Oostdam. Investment strategy for vix futures based on a bayesian approach to term structure modelling. 2021.

[34] Andrew Papanicolaou and Ronnie Sircar. A regime-switching heston model for vix and s&p 500 implied volatilities. *Quantitative Finance*, 14(10):1811–1827, 2014.

[35] Michael Parkinson. The extreme value method for estimating the variance of the rate of return. *Journal of business*, pages 61–65, 1980.

[36] Leonard CG Rogers, Stephen E Satchell, and Y Yoon. Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics*, 4(3):241–247, 1994.

[37] Najmi Ismail Murad Samsudin and Azhar Mohamad. Implied volatility forecasting in the options market: A survey. *Sains Humanika*, 8(2), 2016.

[38] Artur Sepp and Parviz Rakhmonov. Log-normal stochastic volatility model with quadratic drift: Applications to assets with positive return-volatility correlation and to inverse martingale measures. *Available at SSRN 2522425*, 2022.

[39] Jinghong Shu and Jin E Zhang. Testing range estimators of historical volatility. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 26(3):297–313, 2006.

[40] Stephen J Taylor. *Modelling financial time series*. world scientific, 2008.

[41] Peter Tino, Nikolay Nikolaev, and Xin Yao. Volatility forecasting with sparse bayesian kernel models. In *4th International conference on computational intelligence in economics and finance*, pages 1052–1058, 2005.

[42] Spyridon D Vrontos, John Galakis, and Ioannis D Vrontos. Implied volatility directional forecasting: a machine learning approach. *Quantitative Finance*, 21(10):1687–1706, 2021.

[43] Dennis Yang and Qiang Zhang. Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business*, 73(3):477–492, 2000.

[44] Seung-Ho Yang and Jaewook Lee. Predicting a distribution of implied volatilities for option pricing. *Expert Systems with Applications*, 38(3):1702–1708, 2011.

ACADEMIC VITA
Robert J. Krimetz

EDUCATION

The Pennsylvania State University | Schreyer Honors College — **University Park, PA**
Eberly College of Science | B.S. in Data Science focus: Statistical Modeling — Class of 2023
College of the Liberal Arts | B.S. in Economics

RELEVANT EXPERIENCE

Citigroup Sales and Trading — **New York, NY**
Summer Analyst — *May 2022 – Aug 2022*
- Completed the ten-week S&T program rotating on the Short-Term Credit Trading, Equities Platform Sales, and FX & Currencies Structuring desks providing daily market commentary and product related pitch decks
- Wrote python scripts to automate a Value At Risk calculation using Monte Carlo simulation for digital & vanilla option structures and visualized the results for the FX Structuring desk to use in client presentations
- Developed equity momentum indicators based on order flow volume data from the bank's electronic execution algos using PyCaret to compare the accuracy of different machine learning methods given novel feature spaces

Nittany Lion Fund — **University Park, PA**
Director of Pitch Quality | Healthcare Lead Analyst | Director of Compliance — *Dec 2019 – Present*
- Developed the overarching investment thesis for the NLF's ~ $1.70 MM healthcare portfolio by rebalancing subsector allocations to optimize market cap, beta, and active exposure through fundamental stock selection
- Built and updated discounted cash flow models for all healthcare holdings and oversaw the completion of the sector's deliverables including weekly, monthly, quarterly, annual, and other healthcare specific reports
- Directed the training for new/prospective analysts and provided weekly lessons for PSIA's healthcare members

PSU Quant — **University Park, PA**
Founder — *Jun 2019 – Aug 2021*
- Created analytical tools for the Nittany Lion Fund in Python to perform event studies on historical trades to analyze managers' behavioral biases and evaluate the effectiveness of a compliance mandated stop-limit order
- Researched and back tested multi-factor models using Python through the Quantopian IDE and Alphalens
- Built a supervised machine learning model to predict the monthly change in the 10-year treasury yield

Nittany AI Challenge | Trace — **University Park, PA**
Design Lead — *Feb 2021 – Aug 2021*
- Collaborated with teammates to develop a web application that can monitor and predict urban activity volumes using Distributed Acoustic Sensing (DAS) data produced by the Penn State Fiber Optic Array (FORESEE)
- Created a Django web application to store and display FORESEE's recorded vibrations in real time
- Assisted in creating visualization methods to highlight urban activity volume concerns to third party candidates

Hintz Capital Management — **Morristown, NJ**
Summer Intern — *May 2019 – Aug 2019*
Summer Analyst — *May 2021 – Aug 2021*
- Drafted research reports relevant to proposed investment positions spanning equities, commodities, and options trading; streamlining the firm's research process by providing analysts with condensed and relevant research
- Created pitch decks and valuations (DCF and CCA) for prospective holdings and presented to portfolio managers leading to the initiation of new positions and revisions in the firm's targeted subsector allocations
- Calculated and monitored net and gross exposure across multiple accounts, tracking which margins were furthest from the firm's compliance target and accordingly advised analysts through which accounts to make trades

ASA DataFest 2022 — **University Park, PA**
Winner – Most Insightful Analysis — *Mar 2022*
- Cleaned and analyzed data from Yale Medical School's Play2Prevent adolescent drug resistance program using R
- Awarded Most Insightful project from a panel of judges for being the only team to demonstrate and visualize a statistically significant correlation between gameplay data and Drug Use Resistance Survey Results

LEADERSHIP

Eagle Scout — **Mar 11th, 2017**
- Eagle project involved the renovation of a local church's landscaping and construction of a retainment wall
- Received a bronze, gold, and silver palm in addition to holding positions as SPL and ASPL

INTERESTS

- Phillies Baseball, 76ers basketball, skiing, golf, camping/hiking National Parks, espresso drinks