Comparing Predictive Performance of GARCH and Stochastic Volatility Models

Swapnaneel Nath

July 7, 2023

Summary

1. Paper compares predictive performance of two financial models: the Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model, and the Stochastic Volatility model.

2. **Training set** Three 10-year periods (2006-15, 2008-17, and 2010-19) of returns of the S&P-500 Index are used to train the two models.

3. **Test set** Returns for 2016, 2018, and 2020 are predicted.

4. **Estimation method** Hamiltonian Monte Carlo.

5. **Comparison criteria** Sum of squared errors, 95% prediction bands.

6. **Result** Results of the two models are practically indistinguishable by both criteria.

Basic Time Series Model

Observation in a period depends on observation(s) from prior period(s). Basic model:

$$y_t = \alpha + \beta y_{t-1} + z \tag{1}$$
$$z \sim Normal(0, \sigma) \tag{2}$$

Rewritten model:

$$y_t \sim Normal(\alpha + \beta y_{t-1}, \sigma) \tag{3}$$

This is the AR(1) model: first-order autoregressive model.

Extensions - Homoskedastic

AR(p), Autoregressive model of order p:

$$y_t \sim Normal\big(\alpha + \sum_{i=1}^{p} \beta_i y_{t-i}, \sigma\big) \tag{4}$$

MA(q), Moving average model of order q:

$$y_t = \alpha + z_t + \sum_{i=1}^{q} \gamma_i z_{t-i} \tag{5}$$

$$y_t \sim Normal\big(\alpha + \sum_{j=1}^{q} \gamma_j z_{t-j}, \sigma\big) \tag{6}$$

Extensions - Homoskedastic

ARMA(p,q), Autoregressive moving average model:

$$y_t \sim Normal\left(\alpha + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{q} \gamma_j z_{t-j}, \sigma\right) \qquad (7)$$

Heteroskedastic Models

ARCH(q), Autoregressive conditional heteroskedastic model: variance at time t depends on the variances at q prior times.

$$\sigma_t^2 = \delta_0 + \sum_{j=1}^{q} \delta_j z_{t-j}^2 \tag{8}$$

$$y_t \sim Normal(\alpha, \sigma_t) \tag{9}$$

GARCH(p,q), Generalized autoregressive conditional heteroskedasticity model: expands the ARCH model by making the current variance depend on previous errors and their variance.

$$\sigma_t^2 = \delta_0 + \sum_{j=1}^{q} \delta_j z_{t-j}^2 + \sum_{i=1}^{p} \epsilon_i \sigma_{t-i}^2 \tag{10}$$

The observations in GARCH(p,q) are governed by (9).

Stochastic Volatility

Transformation:

$$\lambda_t = \ln(\sigma_t^2) \implies \sigma_t = \exp(\lambda_t/2) \tag{11}$$

In the models described above, the variance in a given time period is perfectly determined by observations in previous periods. There also exist a class of models called stochastic volatility models in which the variance is nondeterministic .

$$y_t = \xi e^{\lambda_t/2} \epsilon_t \tag{12}$$

$$\lambda_t = \eta + \phi(\lambda_{t-1} - \eta) + \zeta_t \tau \tag{13}$$

$$\lambda_1 \sim Normal\left(\eta, \frac{\tau}{\sqrt{1-\phi^2}}\right) \tag{14}$$

Research Questions

Which of the two models, GARCH or Stochastic Volatility, performs better? Do the results hold at different periods and for different time-frames (one month, one quarter, one year)? Critera:

1. Sum of squared errors
2. 95% prediction bands

The performance is also compared for three different time periods: 2016, 2018, and 2020. 2016 had the lowest volatility while 2020 had the highest.

Training Data

The models are trained on the S&P-500 Index (SPX) returns

$$Return_t = \frac{ClosingPrice_t - ClosingPrice_{t-1}}{ClosingPrice_{t-1}}$$

from the following three 10-year periods.

1. **2006-15** containing 2517 observations with minimum -0.0934, maximum 0.1158, mean 0.0002, and standard deviation 0.0130.

2. **2008-17** containing 2518 observations with minimum -0.0903, maximum 0.1158, mean 0.0003, and standard deviation 0.0128.

3. **2010-19** containing 2516 observations with minimum -0.0666, maximum 0.0495, mean 0.0004, and standard deviation 0.0093.

Test Data

The models are tested against data from the year immediately following the training period.

1. **2016** test set has 252 observations with minimum -0.0359, maximum 0.0247, mean 0.0003, and standard deviation 0.0082.

2. **2018** test set has 251 observations with minimum -0.0409, maximum 0.0495, mean -0.0001, and standard deviation 0.0107.

3. **2020** test set has 253 observations with minimum -0.1198, maximum 0.0938, mean 0.0008, and standard deviation 0.0216.

GARCH

This is a GARCH(1,1) model: it uses the return and volatility from one period prior to compute the volatility for the current period. Let r_t denote the return at time t, which depends on μ, the mean return, and λ_t, the log of volatility (variance) at time t. Furthermore, α_0, α_1, and β_1 are all positive, and $\alpha_1 + \beta_1 < 1$.

$$r_t \sim Normal\big(\mu, \exp(\lambda_t/2)\big) \tag{15}$$

$$\mu \sim Cauchy(0, 10) \tag{16}$$

$$\lambda_1 \sim Cauchy(0, 10) \tag{17}$$

$$\lambda_t = \ln\big[\alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \exp(\lambda_{t-1})\big] \tag{18}$$

$$\alpha_0 \sim LogNormal(0, 1) \tag{19}$$

$$\alpha_1 \sim Uniform(0, 1) \tag{20}$$

$$\beta_1 \sim Uniform(0, 1 - \alpha_1) \tag{21}$$

Stochastic Volatility

The Stochastic Volatility model is specified below. The return, r_t, depends on the mean, μ, and log volatility, $\lambda_t = \ln(\sigma_t^2)$. Persistence of volatility is given by ϕ, which is between -1 to 1, and the white noise scaling factor is given by τ.

$$r_t \sim Normal\big(\mu, \exp(\lambda_t/2)\big) \qquad (22)$$

$$\mu \sim Cauchy(0, 10) \qquad (23)$$

$$\lambda_1 \sim Normal(\eta, \tau) \qquad (24)$$

$$\lambda_t \sim Normal\big(\eta + \phi(\lambda_{t-1} - \eta), \tau\big) \qquad (25)$$

$$\eta \sim Cauchy(0, 10) \qquad (26)$$

$$\tau \sim Cauchy(0, 10) \qquad (27)$$

$$\phi \sim Uniform(-1, 1) \qquad (28)$$

Hamiltonian Monte Carlo — Motivation

- Markov Chain Monte Carlo (MCMC) methods are used to derive sampling distibutions, which can then be used to approximate the required integrals.

- RWMS, Gibbs sampling algorithms are simple to implement, however, in high dimesional parameter spaces with correlated parameters, they prove to be inefficient as they make too many poor proposals.

- Proper exploration becomes compuationally prohibitive; Computationally tractable exploration produces biased results.

Hamiltonian Monte Carlo — Motivation

- In high dimensional parameter spaces that are continuous and differentiable (as is the case with the two models specified above), Hamiltonian Monte Carlo (HMC) provides a solution to this exploration issue.

- HMC Markov Chains are very efficient as consecutive points in each chain tend to be both far apart in the parameter space and from regions with high probability mass.

- While the generation of each new point in the sampling distribution involves more computation than in RWMS or Gibbs Sampling, many fewer points need to be generated to sufficiently cover the target space, leading to overall computational savings.

HMC — Method

- For GARCH, $\theta = (\mu, \alpha_0, \alpha_1, \beta_1, \lambda_1, \lambda_2, ..., \lambda_T)$.

- For Stochastic Volatility, $\theta = (\mu, \eta, \tau, \phi, \lambda_1, \lambda_2, ..., \lambda_T)$

T is the number of days in the training set for which we have a return value. The dimension of the parameter space is $b = 4 + T$.

- The HMC algorithm requires us to provide two functions: the natural logarithm of the joint distribution of the parameters, $\ln \pi(\theta)$, and the gradient of this function in all directions, $\frac{\partial \ln \pi(\theta)}{\partial \theta}$. This gradient will guide the trajectory of the exploration.

HMC — Method

Steps:

1. We augment θ with auxiliary variables m. Each parameter in θ has a corresponding auxiliary variable. Thus, $Dim(\theta) = Dim(m)$, say b. The augmented parameter space, (θ, m) is $2b$-dimensional. It is in this extended space that the algorithm will traverse.

2. Next, we define the Hamiltonian function, $H(\theta, m)$ as the negative of the natural logarithm of the joint density of the parameters and the auxiliary variables, that is, the augmented parameter space.

$$H(\theta, m) := -\ln \pi(\theta, m) \tag{29}$$

HMC — Method

3. The Hamiltonian function decomposes as follows.

$$
\begin{aligned}
H(\theta, m) &= -\ln \pi(\theta, m) \qquad\qquad (30) \\
&= -\ln[\pi(m|\theta).\pi(\theta)] \\
&= -\ln \pi(m|\theta) - \ln \pi(\theta)
\end{aligned}
$$

Let $K(m, \theta) := -\ln \pi(m|\theta)$ and $V(\theta) := -\ln \pi(\theta)$.

4. The evolution of θ and m is governed by the following Hamilton's equations.

$$
\begin{aligned}
\frac{d\theta}{d\iota} &= \frac{\partial H(\theta, m)}{\partial m} \\
&= \frac{\partial [K(m, \theta) + V(\theta)]}{\partial m} \\
&= \frac{\partial K(m, \theta)}{\partial m}
\end{aligned}
\tag{31}
$$

$$
\begin{aligned}
\frac{dm}{d\iota} &= -\frac{\partial H(\theta, m)}{\partial \theta} \\
&= -\frac{\partial [K(m, \theta) + V(\theta)]}{\partial \theta} \\
&= -\frac{\partial K(m, \theta)}{\partial \theta} - \frac{\partial V(\theta)}{\partial \theta}
\end{aligned}
\tag{32}
$$

Parameter Estimation

- For the GARCH model, 50,000 iterations are run on 4 chains each. 50% of the iterations are used for warmup, resulting in a posterior sample of 100,000 points in total.

- For the Stochastic Volatility model, 1,000,000 iterations are run on 4 chains for 100,000 iterations each. Half of the iterations are used up in the warmup phase, and one out of 20 successive points is selected resulting again in 100,000 points from the posterior distribution. More iterations and thinning are used in this model to compensate for slower convergence and higher autocorrelation in parameters ϕ and τ.

Parameter Updates

- Updating parameters involves the Sequential Monte Carlo step of assigning weights to each of the 100,000 points from the posterior distribution and then resampling (with replacement) as many points according to these weights whenever the Effective Sample Size (ESS) drops under an acceptable threshold (set, here, at 0.9).

- Weights are determined using the Gaussian density function on the predicted return at each point. That is, given the mean and standard deviation (a tansformation of lambda) corresponding to a point, the point will have a higher weight if the Gaussian density for its predicted return is higher, and vice versa. The weights are normalized.

Prediction and Propagation

- Using the parameters found above, each point at time t is used to predict the point at time $t+1$. Lambda at time $t+1$ is estimated first using the parameters at time t, and then λ_{t+1} (along with other parameters) is used to generate return at time $t+1$.

- "Propagation" refers to moving the cloud of 100,000 points one step (one day) forward. Before each propagation, we check the effective sample size and resample if necessary.

Performance Comparison

- Estimates are tested against the returns data for 2016, 2018, and 2020 using the sum of squared errors. Comparisons are done for each month, quarter, and full-year.

- The empirical distributions of predictions are used to compute the fraction of the observed returns data that fall within the 2.5 percentile and the 97.5 percentile of the predicted returns.

Parameter Estimates for the GARCH model

The mean estimates for the GARCH parameters are as follows.

	2006-15	**2008-17**	**2010-19**
μ	0.0006	0.0006	0.0008
α_0	0.0000	0.0000	0.0000
α_1	0.1235	0.1475	0.1871
β_1	0.8502	0.8300	0.7621
λ_T	-9.1817	-10.6268	-10.3061

Table: Mean of GARCH parameter estimates

Parameter Estimates for the Stochastic Volatility Model

The mean estimates for the Stochastic Volatility parameters are as follows.

	2006-15	**2008-17**	**2010-19**
μ	0.0008	0.0008	0.0009
η	-9.4503	-9.3779	-9.8949
ϕ	0.9798	0.9785	0.9469
τ	0.2190	0.2560	0.3366
λ_τ	-9.2065	-11.0425	-10.8228

Table: Mean of Stochastic Volatility parameter estimates

Sum of Squared Errors

	GARCH	SV	Result
Month 1	0.0047	0.0047	Equal
Month 2	0.0027	0.0027	Equal
Month 3	0.0006	0.0006	Equal
Month 4	0.0009	0.0009	Equal
Month 5	0.0008	0.0008	Equal
Month 6	0.0028	0.0028	Equal
Month 7	0.0004	0.0004	Equal
Month 8	0.0002	0.0002	Equal
Month 9	0.0016	0.0016	Equal
Month 10	0.0003	0.0003	Equal
Month 11	0.0009	0.0009	Equal
Month 12	0.0005	0.0005	Equal

Table: Sum of Squared Errors for 2016, by month

Sum of Squared Errors

	GARCH	**SV**	**Result**
Quarter 1	0.0082	0.0082	Equal
Quarter 2	0.0046	0.0046	Equal
Quarter 3	0.0024	0.0024	Equal
Quarter 4	0.0018	0.0018	Equal
Full Year	0.0171	0.0171	Equal

Table: Sum of Squared Errors for 2016, by quarter and full-year

Sum of Squared Errors

	GARCH	SV	Result
Month 1	0.0007	0.0007	Equal
Month 2	0.0056	0.0056	Equal
Month 3	0.0037	0.0037	Equal
Month 4	0.0017	0.0017	Equal
Month 5	0.0009	0.0009	Equal
Month 6	0.0005	0.0005	Equal
Month 7	0.0006	0.0006	Equal
Month 8	0.0004	0.0004	Equal
Month 9	0.0002	0.0002	Equal
Month 10	0.0045	0.0045	Equal
Month 11	0.0028	0.0028	Equal
Month 12	0.0069	0.0069	Equal

Table: Sum of Squared Errors for 2018, by month

Sum of Squared Errors

	GARCH	SV	Result
Quarter 1	0.0100	0.0100	Equal
Quarter 2	0.0032	0.0032	Equal
Quarter 3	0.0012	0.0012	Equal
Quarter 4	0.0144	0.0144	Equal
Full Year	0.0290	0.0290	Equal

Table: Sum of Squared Errors for 2018, by quarter and full-year

Sum of Squared Errors

	GARCH	SV	Result
Month 1	0.0011	0.0011	Equal
Month 2	0.0078	0.0078	Equal
Month 3	0.0724	0.0724	Equal
Month 4	0.0129	0.0129	Equal
Month 5	0.0032	0.0032	Equal
Month 6	0.0071	0.0071	Equal
Month 7	0.0015	0.0015	Equal
Month 8	0.0006	0.0006	Equal
Month 9	0.0049	0.0049	Equal
Month 10	0.0035	0.0035	Equal
Month 11	0.0025	0.0025	Equal
Month 12	0.0006	0.0005	SV is lower

Table: Sum of Squared Errors for 2020, by month

Sum of Squared Errors

	GARCH	**SV**	**Result**
Quarter 1	0.0814	0.0813	SV is lower
Quarter 2	0.0233	0.0233	Equal
Quarter 3	0.0071	0.0071	Equal
Quarter 4	0.0066	0.0066	Equal
Full Year	0.1185	0.1185	Equal

Table: Sum of Squared Errors for 2020, by quarter and full-year

Capture in 95% Prediction Band

- The empirical distributions of the predicted returns for 2016, 2018, and 2020 are used to find the predictions at the 2.5 percentile and the 97.5 percentile for both models in each day of the three years.

- For the GARCH model, the 95 percent prediction band captured 238 out of 252 observations in 2016; 243 out of 251 observations in 2018; and 206 out of 253 observations in 2020.

- For the Stochastic Volatility model, the 95 percent prediction band captured 237 out of 252 observations in 2016; 242 out of 251 observations in 2018; and 203 out of 253 observations in 2020.

Discussion

Results are almost the same. Small differences that occur usually after the ten-thousandths place can be attributed to monte-carlo errors.

1. Out of the twelve month long time-frames assessed for 2016, the GARCH and the Stochastic Volatility results are indistinguishable in every month. These SSEs range from a low of 0.0002 to a high of 0.0047 for both estimates. The highs and the lows occur on the same months for both the models.

2. Out of the twelve month long time-frames assessed for 2018, the GARCH and the Stochastic Volatility results are indistinguishable in every month. These SSEs range from a low of 0.0002 to a high of 0.0069 for the both estimates. The highs and the lows occur on the same months for both the models.

3. Out of the twelve month long time-frames assessed for 2020, the GARCH results are indistinguishable from the Stochastic Volatility results on 11 occasions. The only small noticeable difference occur in month 12, for which the Stochastic Volatility estimate shows a lower sum. These SSEs range from a low of 0.0006 to a high of 0.0724 for the GARCH estimates, while for the Stochastic Volatility estimates, they range from a low of 0.0005 to a high of 0.0724. The highs and the lows occur on the same months for both the models.

4. Out of the four quarters assessed for 2016, GARCH and Stochastic Volatility perform equally well in every quarter. These SSEs range from a low of 0.0018 to a high of 0.0082 for both estimates. The highs and the lows for both models occur in the same quarters.

5. Out of the four quarters assessed for 2018, GARCH and Stochastic Volatility perform equally well in every quarter. These SSEs range from a low of 0.0012 to a high of 0.0144 for both estimates. The highs and the lows for both models occur in the same quarters.

6. Out of the four quarters assessed for 2020, GARCH and Stochastic Volatility are indistinguishable in every quarter except the first. In quarter 1, the sum for Stochastic Volatility is slightly lower. The SSEs range from a low of 0.0066 to a high of 0.0814 for the GARCH estimates, while for the Stochastic Volatility estimates, they range from a low of 0.0066 to a high of 0.0813. The highs and the lows for both models occur in the same quarters.

7. For 2016, the full-year SSE is 0.0171.

8. For 2018, the full-year SSE is 0.0290.

9. For 2020, despite small differences in monthly or quarterly results, the full-year SSE is 0.1185 for both models.

10. In 2016, the prediction bands for the GARCH model capture 238 out of 252 observations (94.44% accuracy) while those for the Stochastic Volatility model capture 237 out of 252 observations (94.04% accuracy).

11. In 2018, they capture 243 out of 251 observations (96.81% accuracy) for GARCH, and 242 out of 251 observations for Stochastic Volatility (96.41% accuracy).

12. In 2020, they capture 206 out of 253 observations for GARCH (81.42% accuracy), and 203 out of 253 observations for Stochastic Volatility (80.23% accuracy).

13. While the GARCH model is slightly more predictive in each year above, they differ by only one day. Also, in the high-volatility year of 2020, the predictive accuracy of both models reduce significantly. Hence, in practical terms, both models are indistinguishable.

Further Studies

1. Empirical distributions of parameters are un/under-utilized in this paper. Further studies might take into account the full distributions of these parameters for generating predictions.

2. Analyses in this paper compressess the distribution of predicted returns. Computing SSE requires distributions to be reduced to single point estimates. Prediction bands use only two points from each distribution.

3. Further studies could make use of the full empirical distributions of predicted returns. These might include an analysis of distribution-based properties such as the width of the prediction bands and the thickness of the tails.

4. The outputs of the Stochastic Volatility models generally result in greater kurtosis in all three test years.

5. In 2016, the skewness of the distribution of kurtosis of the daily predicted returns for the GARCH model is 0.1964 as opposed to 1.0573 for the Stochastic Volatility model. This implies that the kurtosis of daily returns is more right-skewed for the Stochastic Volatility model, further implying that the distribution of predicted returns from the Stochastic Volatility model is thicker-tailed.

6. This result holds for 2018 with the skewness of the kurtosis at 1.0021 for GARCH and 4.8058 for Stochastic Volatility; and for 2020 with the skewness of kurtosis at 1.9687 and 4.4532 for GARCH and Stochastic Volatility respectively.

7. An exploration of the process resulting in these differences and their implications might be of interest, especially when the analysis is done from a risk-management/ruin-avoidance framework.

8. Probabilistic forecasting is another approach that may be taken to make claims about the returns predictions. In this case, probabilistic forecasting would generate various ranges of returns and the associated probabilities for observing returns in those ranges.

9. Parameters μ and λ_1 for GARCH, and μ, η, and τ for Stochastic Volatility begin with noninformative *Cauchy*$(0, 10)$ priors. Convergence can be vastly sped up by the use of informative priors. Posterior sampling distributions generated in this paper may be used to generate priors in future studies.

10. Methodology used in this paper may also be used on assets other than the S&P-500 Index to see how the parameter estimates differ are and whether the prediction accuracy holds.

11. An aspect of modeling this paper does not cover is computation time. In this study, the parameter estimation for Stochastic Volatility took several times as long as the parameter estimation for GARCH, and yet resulted in similar outputs. Further studies could involve a closer analysis of computation time.

Thank you!