Forecasting Energy Minimum Variance Index

Qianlin Song, Shiow-Kang Hu, Tina Behroozi, Xinyuan Liu
STA 2202H
December 2, 2019

Table of Contents

Executive Summary

The realized minimum variance index includes flat currencies, cryptocurrencies and energy assets. Their optimized weights were assigned decreasingly in the same order. This process, which involves variance and covariance of all assets, were then recalibrated every quarter. In order to forecast the realized minimum variance index, three models were considered.

ARIMA model was first tested. To validate its results, time-series cross validations were performed for every quarter. The parameters of this model includes the number of lag observations, the number of times that the raw observations are differenced and the order of the MA model.

Next, GARCH and APARCH models was tested. This model allows volatility changes in time and depend on previous ones. Usually garch and arma are used together. The garch model uses lag of residuals squared and also variance. The APARCH model add some flexibility to GARCH to show leverage effect.

Lastly, Gaussian copula model was able to produce a reasonable prediction of the index. Unfortunately, due to time and resource constraints, we were not able to compare the fitting results with other models. However, it is a promising candidate model to explore in the future.

Models are built specifically on the minimum variance index with key data elements being fully cleaned and reconciled to data sources. This suite of models fully complies with the requirements of this final project.

Based on our models and validations, we propose APARCH model since it has low MSE and provided best fit for the data. GARCH is our close second choice, as it provided almost the exact same fit. Lastly, we ranked ARIMA third as it has materially higher MSE.

Background

An index is a measurement of the selected assets, which are normally computed from the weighted average prices. For this project, our team is tasked to create a minimum variance index for Energy (label 12), which includes the following assets:

Class	Asset
Fiat Currencies	USD, Euro (EUR), Japanaese Yen (JPY), Chinese RenMinBi (RMB)
Crypto Currencies	Bitcoin (BTC), Ethereum (ETH), Ripple (RIP), Monero (MON)
Energy	Brent Crude Oil (OILI, Gasoline (GAS), Natural Gas (NAT), Uranium (URA)

Unlike typical indexes, minimum variance index is constructed from rules-based methodologies to minimize the portfolio volatility. Appetite for this type of index comes from today's low interest rate environment, which increases reliance on the equity market. The bearish performance in the Great Recession also motivated investors as well.

As part of the instructions, all assets were converted from USD to Special Drawing Right currency (XDR). It is an international reserve asset created by the IMF to supplement its member countries' official reserves. This means the final index would include both asset and currency risks. We observed that the asset risk still dominates the resulting index.

There were two significant noteworthy events during the study period. The first event is the 2018 cryptocurrency crash, where their prices collapse from more than 80% of the peak value. The second event is the fall of gas prices due to lower cheaper raw materials. Keeping these events in mind may help us get a better understanding of the results.

Data & Assumptions

Data

The raw data with asset daily prices given to you in an Excel spreadsheet were obtained for the period from September 30, 2016 to September 30, 2019 from the source below:

- https://www.quandl.com/
- http://www.globalfinancialdata.com/

For Flat Currencies and Crypto Currencies, daily data is utilized for model development. Energy Assets have different data frequency: for Brent Crude Oil, daily data is utilized. For Gasoline, weekly data is employed. For both Natural Gas and Uranium, the data frequency is monthly.

Data Quality

We reconciled Excel data provided by the instructor with Quandl and Global Financial Data. The reconciliation matches perfectly.

Data Cleaning and Processing

Data were cleaned and processed in the following steps:

1. Identify the missing data and data frequency of each asset.
2. For monthly and weekly data, use the latest value to impute missing value.
   Interpolation and KNN were also considered but the latest value was chosen due to simplicity and prediction accuracy.
3. Manual fixes of data:
   a. In the first quarter of Chinese RenMinBi, there were artificial spikes in some weekends. We removed them after comparing it with Quantdl.
   b. We inverted the Chinese RenMinBi data since it was not a conversion to USD.
4. For observations that missing from the beginning, compute the average of asset values in 6 months and impute it into the missing observations. This is because we observed, for such index (Gasoline), the time series in 6 months is relatively stable and developers also utilized 6 months as a rolling window to calculate weights.
5. Convert assets into special drawing rights (XDR).
6. Keep current data as price and calculate log return as asset return.

Assumptions

The following assumptions are fundamental for constructing the minimum variance index:

- Weight (w_i) that minimize the variance of the index at each quarter end i would be the weights that minimize the variance of the index for next quarter i+1.
- Index Price ($I(T_i)$), or value of the index at quarter T. Note that the initial index price is constrained to 100 as per instruction.
- Asset Price ($A_n(t)$) for the n-th asset at time t. Together they form an index.

Methodologies

Index Calculation

Realized Index is calculated at the daily level based on the following steps:

1. The initial index value on March 31, 2017 is calibrated to 100 XDR.
2. Check the quarter end price and weights of each asset as shown below:

$$I(t) = \sum_{n=1}^{N} \frac{w^i{}_n \times I(T_i)}{A_n(T_i)} \times A_n(t)$$

  - i is the number of a calendar quarter
  - $I(T_i)$ is the index value at the end of quarter i
  - $A_n(T_i)$ is the price of the asset n of the index at the end of quarter i
  - $A_n(t)$ the market price of the asset n of the index during calendar quarter $i+1$

   The optimization process will be explained in the next section
3. Check the next quarter end index value and repeat the above step for 10 quarters.

Minimum Variance Weight Optimization

One of the main objectives is to achieve minimum variance for index stabilization. Although asset returns are unknown ahead of time, we can choose to invest in those that have lower historical variances. As for the recalibration frequency, it is done every quarter as instructed. Mathematically, this becomes an optimization problem, and we started by experimenting with a linear programming equation below:

$$Minimize \ \sum_{i=1}^{n} w_i \sigma_i^2$$

$$Such\ that \ \sum_{i=1}^{n} w_i = 1,\ w_i >= 1\%,\ w_i <= 25\%$$

For asset 1 to n, w_i is weight of asset i, while σ_i^2 is its variance. Of course, we want the weight of all assets to add up to 100%. The weight constraints of 1% and 25% are per instruction.

The model above was able to produce reasonable results. For example, it assigned limited exposures to cryptocurrencies, which is the most volatile. However, it is not able to detect correlation between assets, and does not aim for positive returns like real-life investors. For this reason, we based our final optimizer on a research by Laplante, Desrochers, and Prefontaine. The improved quadratic programming equation is shown below:

$$Minimize \ \sum_{i=1}^{n} w^2{}_i \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1,\ i \neq j}^{n} w_i w_j \sigma_{ij}$$

$$Such\ that \ \sum_{i=1}^{n} w_i = 1,\ w_i >= 1\%,\ w_i <= 25\%, \sum_{i=1}^{n} w_i E[r_i] = E[r_p]$$

In addition to the notations above, we took covariance σ_{ij} into consideration. Moreover, a constraint was added such that weighted portfolio return to perform as expected. In our case, we set expected portfolio return, $E[r_p]$, to be actual portfolio return from prior quarter.

Unfortunately, it produced a computational error during the bearish market quarter. We suspect that with more diversified pool of assets, this might be mitigated. In the end, we relaxed this constraint.

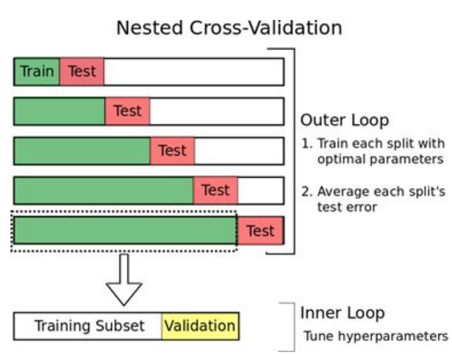
Our final optimizer produces similar weights as the first one, where it limited exposure in high volatility assets. The difference lies in its ability to detect correlations between given assets, resulting in a more diversified portfolio. Given this, the final index returns from here on will be based on weights from this model.

Time-Series Cross Validation

The performance of time-series models is highly sensitive to data. To avoid unrealistic dependencies into the model, developers did not utilize traditional bootstrapping or cross-validation. The general design developers build the model is called cross validation for time series. It provides an almost unbiased estimate of test error, accounting for variability across timeline. This idea is implemented in this model with the following steps:

1. Determine the date where each quarter ends
2. Train the model using all data available before this date
3. Forecast the next quarter given the trained model
4. Keep the prediction and model as a result.
5. Rollforward to the next quarter while adding the new prediction and model into result
6. Analyze the models based on the average test error and in the end, the parameters and model are chosen based on minimized average test error

This process is visualized below:



Analysis of Results

Analysis of Assets

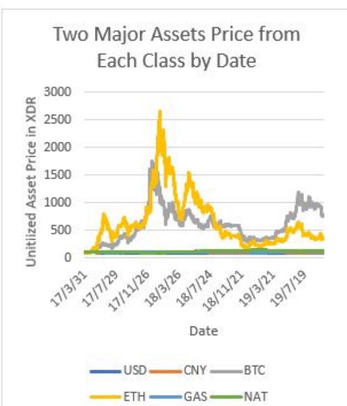
Individual Assets

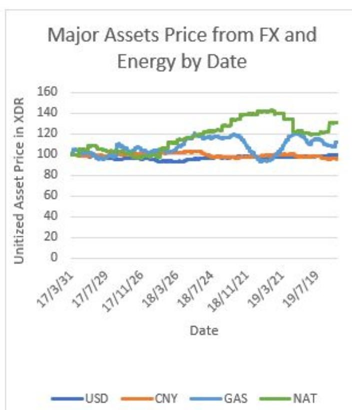
This section contains observations and analysis of each individual asset. We started by experimenting with USD data, and found that it is far from normally distributed. We also found that our assets are time-dependent, and differencing would be needed. The table below summarizes results from an R-function called ndiff, which uses KPSS test to return number of differencing needed in order for the time-series data to be stationary:

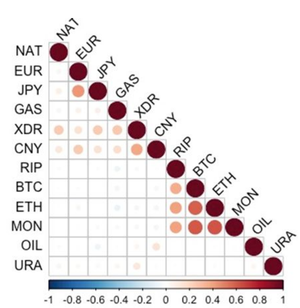
Asset	USD	EUR	JPY	CNY	BTC	ETH	RIP	MON	OIL	GAS	NAT	URA
ndiff	2	1	1	2	1	1	1	1	1	1	1	1

Asset Classes

This section contains observations and analysis of each asset class. We started by plotting the unitized asset prices by date from each asset class to observe the volatility of our data. As shown in the left graph below (left & middle), crypto assets have much higher volatility compared to other asset classes. If we zoom into FX and Energy assets in the right graph, we found that Energy assets have higher volatilities.







The correlations between each asset are summarized above (right). The strongest correlations come from cryptocurrency assets, and the weakest ones were energy. This chart will also help us explain the weights picked by the optimizer in the next section.

Analysis of Indexes

Realized Index

This section only contains actual results (no forecasting models) after weight optimization. The optimized weights are shown below:

Qtr	Currency				Crypto				Energy				Asset Class		
	USD	EUR	JPY	RMB	BIT	ETH	RIP	MON	OIL	GAS	NAT	URA	Currency	Crypto	Energy
Q1	25%	25%	1%	25%	1%	1%	1%	1%	1%	17%	1%	1%	**76%**	**4%**	**20%**
Q2	25%	25%	5%	25%	1%	1%	1%	1%	1%	13%	1%	1%	**80%**	**4%**	**16%**
Q3	25%	25%	17%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q4	25%	17%	25%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q5	25%	17%	25%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q6	25%	23%	19%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q7	25%	23%	17%	25%	1%	1%	1%	1%	1%	3%	1%	1%	**90%**	**4%**	**6%**
Q8	25%	22%	20%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q9	25%	25%	17%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q10	25%	25%	17%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**
Q11	25%	25%	17%	25%	1%	1%	1%	1%	1%	1%	1%	1%	**92%**	**4%**	**4%**

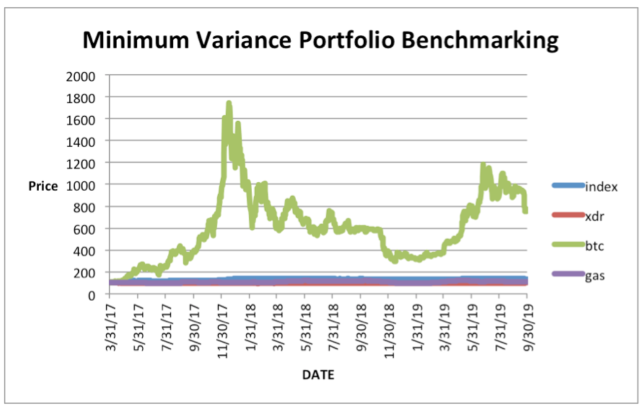
At inception, USD, Bitcoin and Gasoline are major assets. We unitized the initial value of each asset into 100 XDR to satisfy the initial constraint provided by the instructor. The Crypto Currencies is volatile compared to Flat Currencies and Energy and Energy is volatile compared to Flat Currencies, then the ranking order of each asset class percentage within minimum variance index is reasonable. Also, for the change in quarter one and quarter two, it could be explained by changes of volatility in quarter zero (September 2016 to March 2017) and quarter one and quarter zero (January 2017 to June 2017). In general, the variance of Currency from September 2016 to June 2017 is higher than the variance in the following quarters. In contrast, the variance of Energy is relatively lower in these two quarters.

variance in Qi-1 and Qi-2	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q1-Q11
XDR	0.00000480	0.00000380	0.00000254	0.00000295	0.00000323	0.00000378	0.00000341	0.00000290	0.00000209	0.00000146	0.00000129	0.0000029
EUR	0.00001651	0.00001534	0.00001456	0.00001344	0.00001216	0.00001421	0.00001287	0.00001102	0.00001107	0.00000784	0.00000460	0.0000122
JPY	0.00002859	0.00002582	0.00001828	0.00001252	0.00001288	0.00001549	0.00001280	0.00001151	0.00001234	0.00001000	0.00000673	0.000016
CNY	0.00000312	0.00000267	0.00000278	0.00000359	0.00000416	0.00000440	0.00000577	0.00000683	0.00000462	0.00000325	0.00000317	0.0000043
BTC	0.00111274	0.00117145	0.00220584	0.00322739	0.00362989	0.00257024	0.00106960	0.00115844	0.00103226	0.00123781	0.00188052	0.0018145
ETH	0.00370481	0.00427233	0.00554365	0.00422081	0.00374131	0.00364068	0.00256588	0.00276018	0.00245963	0.00190602	0.00202656	0.003355
RIP	0.00215202	0.01138307	0.01694943	0.00880022	0.01100853	0.00570132	0.00370577	0.00348950	0.00168858	0.00158958	0.00232261	0.0062446
MON	0.00361751	0.00439558	0.00704825	0.00682417	0.00635337	0.00505554	0.00288596	0.00287988	0.00218702	0.00175216	0.00222990	0.004082
OIL	0.00032050	0.00029135	0.00021507	0.00016435	0.00013287	0.00018883	0.00021554	0.00029055	0.00032585	0.00025065	0.00026689	0.0000176
GAS	0.00002291	0.00002501	0.00008722	0.00008625	0.00002638	0.00002650	0.00001746	0.00003029	0.00005064	0.00004285	0.00004078	0.0000407
NAT	0.00003348	0.00004206	0.00003504	0.00002190	0.00006967	0.00005936	0.00001971	0.00002705	0.00002053	0.00007088	0.00012711	0.0000453
URA	0.00055376	0.00046270	0.00016777	0.00013786	0.00012787	0.00005001	0.00012872	0.00011448	0.00004946	0.00009494	0.00013699	0.0001823

The realized minimum variance index is plotted above. We observed that the weights of Euro, Japanese Yen, Natural Gas, Gasoline and Uranium fluctuate frequently. The comparison of the volatility of each asset in quarter i and $i-1$ explains the weights change in quarter $i+1$ very well. A high volatility in quarter i and $i-1$ leads to a lower weight in quarter $i+1$ and at the same time. Likewise, a lower volatility in quarter i and $i-1$ leads to a higher weight in quarter $i+1$. Moreover, the correlation between each asset magnifies the influence of volatility. For example, Japanese Yen and Euro currencies has a strong negative correlation. As a result, the Euro has a higher weight than Japanese Yen in quarter six.

Benchmarking

The performance of minimum variance index was compared with levels of asset classes that constitute the index by using one major asset from each asset class at inception from March 31, 2017 through September 30, 2019. The value of the benchmarked assets was set to 100 on March 31, 2017 to compare with minimum variance index. In other words, developers assumed that all money at inception was invested in either minimum variance index or major assets in each class. In this case, the major assets at inception are USD, Bitcoin and Gasoline.



While the initial value of minimum variance index, XDR (USD), BTC (Bitcoin) and GAS (Gasoline) are at 100, the minimum variance index has higher levels of value than GAS and XDR but lower volatility than BTC. The goal of volatility reduction is achieved by minimum variance index and the above average historical returns prove the good performance of minimum variance index.

Analysis of Forecasting Models

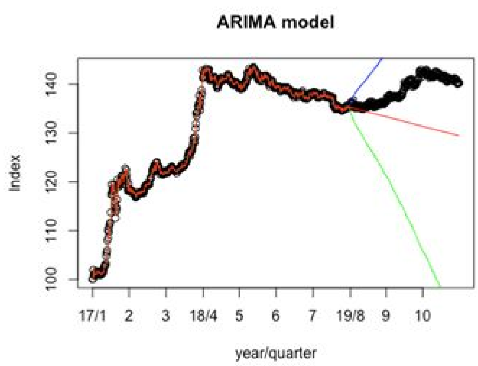
ARIMA Model

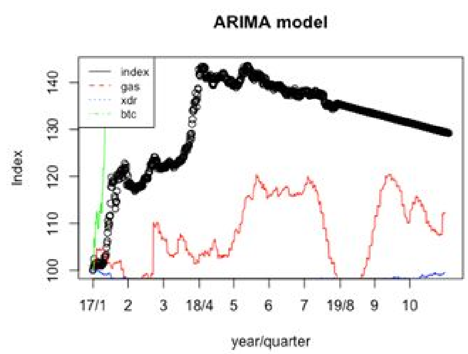
ARIMA Model stands for Auto Regressive Integrated Moving Average. It is a common model for analyzing and forecasting time series data. It is a generalization of the simpler Auto Regressive (AR), Moving Average (MA) and adds the notion of integration.

In order to pick the best parameters, we performed a time series cross-validation. Model and parameters were selected based on relative measures of performance: RMSE, MAPE & MSE. The 7th RMSE and MSE is significantly better than others. However, its MAPE is slightly lower than 6th quarter model; Therefore, the 7th model was selected, which is ARIMA (2,2,3). The summary of model performances can be found below. The best performing model is highlighted in yellow.

Training Quarter	ARIMA (p,d,q)	RMSE	MAPE	MSE
1	(0,1,0)	8.56208	0.05965486	73.30922
2	(0,1,0)	3.250804	0.02153386	10.56772
3	(0,1,0)	8.520436	0.04819091	72.59782
4	(3,1,0)	4.71974	0.02401454	22.27595
5	(3,1,0)	5.962479	0.03574395	35.55115
6	(2,2,3)	1.923359	0.01107697	3.699312
7	**(2,2,3)**	**1.646959**	**0.01056201**	**2.712475**
8	(2,2,3)	3.086849	0.01898104	9.528635
9	(2,2,3)	3.609247	0.02120096	13.02667

The first seven quarters' data was utilized to backtest the index value. We compared the prediction for quarter 8 to quarter 10 with the realized minimum variance index in the left chart below. The best performing model has AIC of 1,188.07 and BIC of 1,214.82.

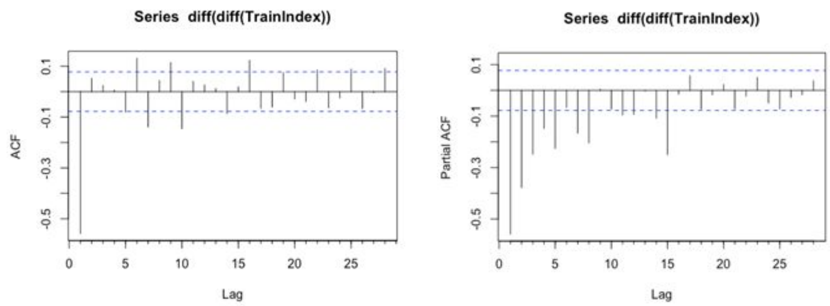




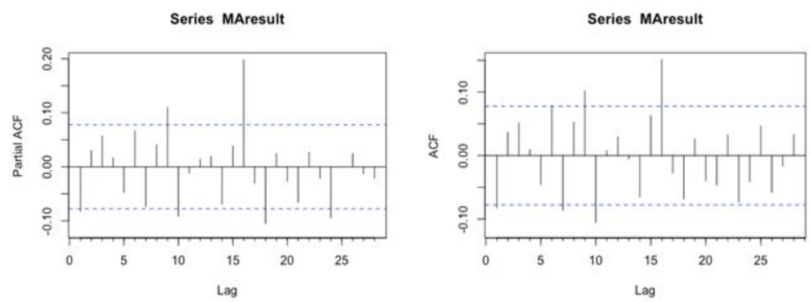
Based on the mean of predicted Index, we may have reduction in index (the weighted return will decrease). If we use the same model for predicting log return of three major assets, the result will be as the right chart above.

ARIMA Model Based on ACF and PACF

For the selected model (7 quarters for training), we also find p and q based on ACF and PACF without considering auto ARIMA function. First, we used standard KPSS test for testing if data is stationary. Based on KPSS results after taking to differentiate of index, the result become stationary (p-value = 0.1).



The PACF variation is less than ACF and also first lag is negative. This shows MA characteristic. Therefore, we applied MA(1) to the data.

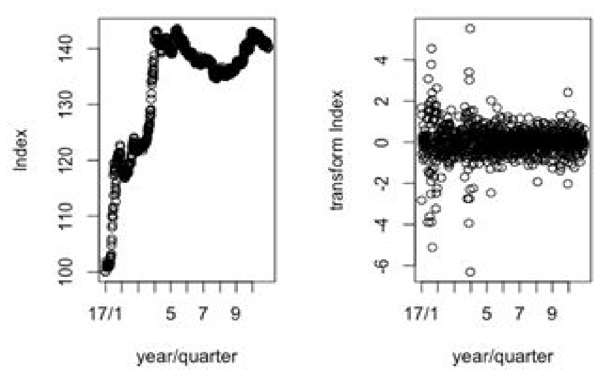


Therefore, if we wanted to choose the model based on ACF and PACF plot, ARIMA(0,2,1) would be chosen. However, auto arima gives better results ,and the two model results are close.

GARCH and APARCH Models

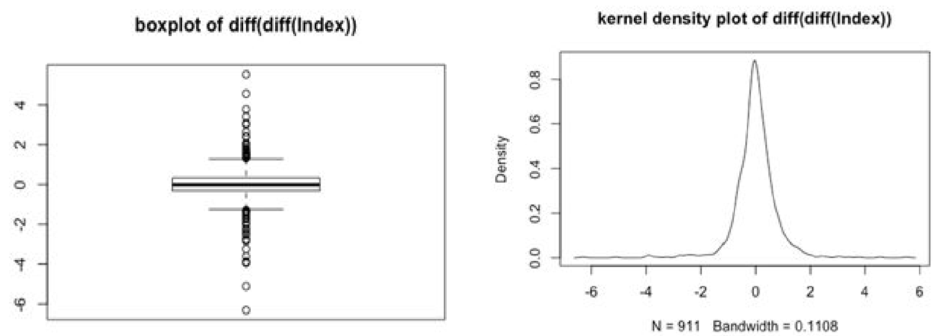
Model Preparations

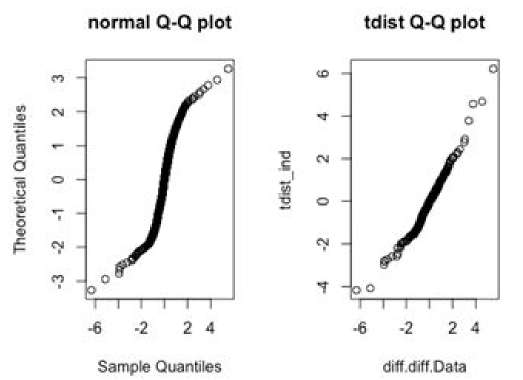
For GARCH and APARCH model, we needed to make index stationary first. Standardized KPSS test was used to check if transformed index is stationary. Square, log, Box-Cox with lamba of 1.95, derivative, and second derivative of index were considered. As mentioned in the ARIMA section, the second derivative of index gave us stationary data (p-value = 0.1). Furthermore, based on index figure and standard ndiffs test index don't have seasonal property.



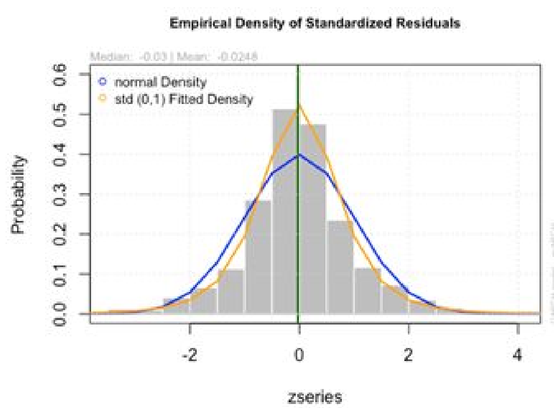
The above figure also shows that when we consider the second derivative of index, the mean and variance is close to constant.

Now we want to find the transformed index distribution. Based on its density, boxplot, and qq-plot, we can see that the data is left skewed and have thicker tails than a normal distribution. Therefore we fit our data in t-distribution.





Based on the plots above, if we don't consider outliner, we can conclude that the second derivative of the data have t-distribution.



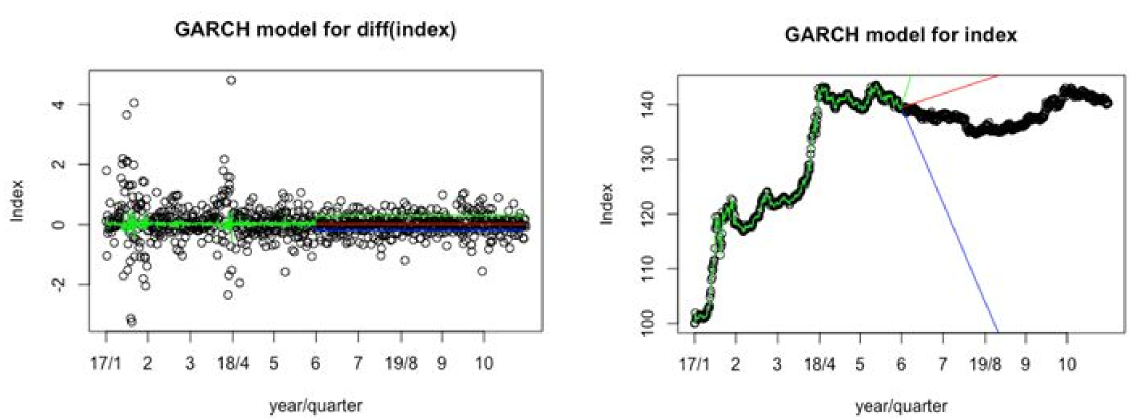
Moreover, the above figure indicates that the standard residuals of the data approximately have t-distribution. Furthermore, if we calculate the skewness (-0.096) and kurtosis (2.288) of the residuals, we see that it is left skewed and have lighter tail than normal distribution. This will also confirm the hypothesis that the standard residuals have t-distribution.

We use this assumption to fit the second derivative of the index into GARCH and APARCH models.

GARCH Results

It is very common for financial time-series data to have changes in volatility during different periods. The ARMA model consider the conditional variance to be same. Therefore, we are also going to fit our data in Garch model. We defined 9 models like Arima model and then compare them based on RMSE, MAPE, and MSE. The summary of results are shown below:

Training Quarter	ARIMA (p,d,q) +GARCH(1,1)	RMSE	MAPE	MSE
1	(0,1,0)	0.4008973	3.141712	0.1607186
2	(0,1,0)	0.8347413	12.1947	0.6967931
3	(0,1,0)	0.4098489	5.98949	0.1679761
4	(3,1,0)	0.4379748	10.0336	0.1918219
5	**(3,1,0)**	**0.2876346**	**7.82235**	**0.08273366**
6	(2,2,3)	0.4561826	869.4051	0.2081026
7	(2,2,3)	0.4762306	1966.05	0.2267956
8	(2,2,3)	0.6037025	993.2003	0.3644567
9	(2,2,3)	0.4375554	844.6743	0.1914547



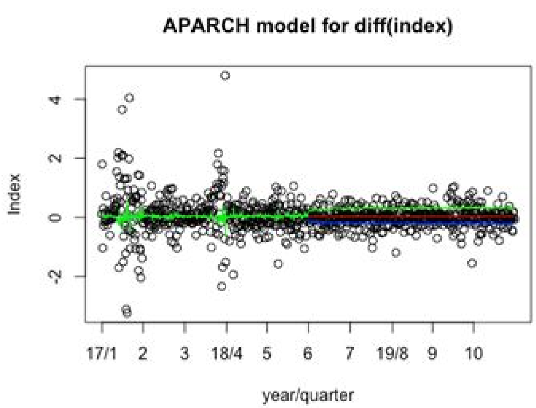
The AIC and BIC is 1.481 and 1.553 respectively.

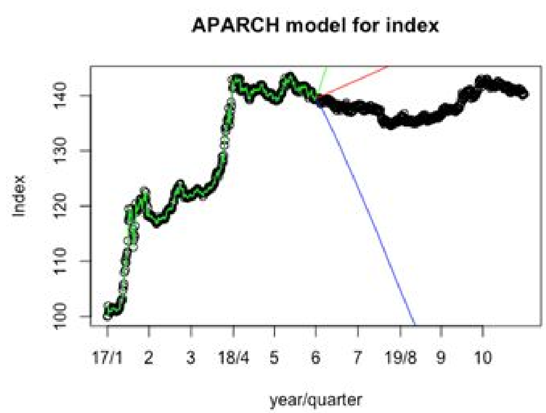
APARCH Results

The garch model cannot consider data with negative value and consider them as positive (does not consider leverage effect). Therefore Aparch model add another constant to the model to be able to show leverage effect. Because Aparch has one more degree of freedom than Garch model, the model can be more flexible than Garch model. However, it will cost more simulation complexity.

Training Quarter	ARIMA (p,d,q) +APARCH(1,1)	gamma	RMSE	MAPE	MSE
1	(0,1,0)	-1.000000	0.4028826	2.807845	0.1623144
2	(0,1,0)	-1.000000	0.8276118	6.21674	0.6849414
3	(0,1,0)	-1.000000	0.4137134	4.298736	0.1711588
4	(3,1,0)	-1.000000	0.4394019	6.865082	0.193074
5	**(3,1,0)**	**-0.985405**	**0.2890811**	**6.29768**	**0.08356791**
6	(2,2,3)	-0.285121	0.456172	531.6295	0.2080928
7	(2,2,3)	-0.214303	0.4762817	573.9417	0.2268443
8	(2,2,3)	-0.255484	0.6036165	590.0321	0.3643528
9	(2,2,3)	-0.295882	0.4383166	367.6671	0.1921215

Based on above table 5th model gives the least error. Although MAPE error of some of the models are higher, however MSE and RMSE of 5th model is better than other methods.

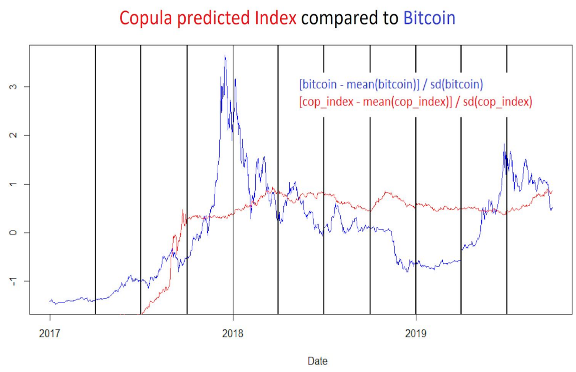


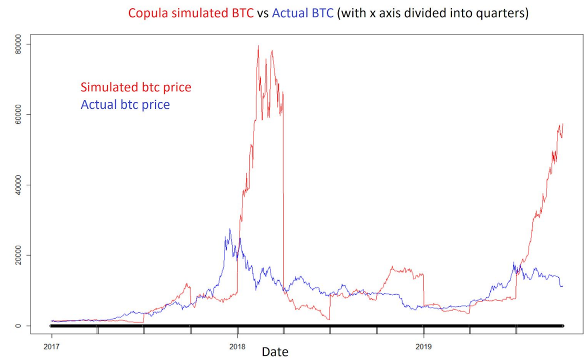


The AIC and BIC is 1.455 and 1.545 respectively.

Gaussian Copula Simulation of Index

For each quarter except the first one, our Gaussian copula model attempts to simulate the quarters prices for all of the assets and fit the model using Kendall's tau coefficient, as well as the empirical CDFs of each of the previous months assets. Combined with the weights that we obtained through optimization, this allows us to construct a simulated index for that month. Of course, any approximation of the inverse CDF of the assets can be used, which would produce different results.





The index the Gaussian copula has predicted is shown on the left. The copula's simulation for Bitcoin price is compared to the actual price on the right. Note that the jumps at the end of each quarter for the graph on the right are because the actual price of bitcoin at the end of the previous month is set to the initial value of the copula at that month.

The copula model appears to overestimate the variance of bitcoin during the Bitcoin spike. The copula is created by first making the underlying assets stationary, in this case by taking the difference of the log of consecutive prices of each asset. Thus the changes in the assets is no longer autocorrelated in the copula simulations. This is not a very realistic assumption.

Conclusions

In conclusion, our minimum variance index performed as expected. It invested heavily in low-volatility assets like currencies, and avoided high-volatility assets like cryptocurrencies. As for the forecasting models that we have analyzed, their results are summarized below:

Model	Parameters	MSE
ARIMA	(2,2,3)	2.712
GARCH	**(3,1,0)**	**0.083**
APARCH	(3,1,0)	0.084

The best ARIMA, GARCH, and APARCH models returned low mean-squared errors and can be used for forecasting results. However, based on the mean-squared error and AIC, APARCH is best fit for the data. APARCH result is better than GARCH one because it has more flexibility. However, APARCH and GARCH results are very close because our index don't have negative value.

Moreover, the GARCH's mean-squared error is less than Arima because the index have non-constant volatility. This can also be seen in numerous number of lags in ACF of data. Furthermore, AIC and BIC of GARCH model are much smaller than ARIMA model. This means GARCH model has less leakage in information than ARIMA model. Since GARCH allows conditional variance, and the market data exhibits non-constant volatility, especially during the crypto market crash of 2018. Thus, we believe that GARCH is a reasonable model for our forecast. It was reasonable that GARCH gives better results, because it exhibits heavy tail which is very common for stock returns.

Our conclusion is that APARCH provided the best fit for our data with GARCH as a close second, and ARIMA as our third place.

Due to time and resource constraints, we were not able to compare model fit of Gaussian copula with other models. However, it showed reasonable predictions and is a promising candidate model to further explore in the future.

Appendix A: Performance Measures

To measure the performance of models, RMSE (square root of the MSE), MAPE (mean absolute percentage error) and MSE (mean squared error) are utilized. They indicate the distance between realized data points and prediction and thus measure the performance of each model. They are very common criterion for model performance. In the Garch(1,1) model as a risk predictor for international portfolios, ME (margin of error), RMSE, MAPE and MSE are employed for statistical error measuring. The lower the measure is, the better the model performance is.

- RMSE is an absolute measure of fit and it can be interpreted as the standard deviation of the unexplained variance. It is the square root of MSE.
- MAPE expresses accuracy as a percentage and it is a relative error.

$$M = \frac{100\%}{n\%} = \sum_{t=1}^{n} = \left|\frac{A_t - F_t}{A_t}\right|$$

  where A_t is the actual value and F_t is the forecast value. The difference between A_t and F_t is divided by the actual value A_t again. The absolute value in this calculation is summed for every predicted value and divided by the number of fitted points n. Multiplying by 100% makes it a percentage error.

- MSE measures the average of the squares of the errors. It is the average squared difference between estimated value and the realized value.

References

1. *Volatility reduction: How minimum variance indexes work*. (ftserussell.com).

2. Badr, W. (n.d.). 6 Different Ways to Compensate for Missing Values In a Dataset (Data Imputation with examples). *Towards Data Science*. Retrieved from https://towardsdatascience.com/6-different-ways-to-compensate-for-missing-values-data-imputation-with-examples-6022d9ca0779

3. Laplante, J.-F., Desrochers, J., & Préfontaine, J. (2011). The GARCH (1, 1) Model As A Risk Predictor For International Portfolios. *International Business & Economics Research Journal (IBER)*, *7*(11). doi: 10.19030/iber.v7i11.3305

4. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. Heathmont, Vic.: OTexts.

5. Cochrane, C. (2018, May 19). Time Series Nested Cross-Validation. Retrieved from https://towardsdatascience.com/time-series-nested-cross-validation-76adba623eb.

6. RUPPERT, D. (2016). *Statistics and data analysis for financial engineering*. Place of publication not identified: SPRINGER-VERLAG NEW YORK.