

Construction of Minimum Variance Indices with Heteroskedastic Time-varying Asset Covariances

Name: Andrew Muth, Yi (Emily) Wang, Crystal Hai Ying Chen

December 3rd, 2019

1. Goals and Motivation

Time series analysis is useful in terms of identifying trends, seasonal variations, and changes in correlations over time, and thus, is widely used for building market portfolios or indices. The goal of this project is to construct a minimum variance index for Agriculture 12 in XDR for all calendar quarter ends starting March 31, 2017 through September 30, 2019. Note that Agriculture 12 involves the following assets:

- Flat Currency: USD, Euro, JPY, Chinese Yuan/Renminbi
- Cryptocurrency: Bitcoin, Ethereum, Ripple, Monero
- Agri Commodities: Wheat, Soy, Cattle, Lumber

We started by analyzing each individual index component using univariate ARMA models. As indicated by the time series plots (see Figure 1-3), the raw market price data exhibited time-varying mean and volatility, indicating the necessity of introducing GARCH models. The GARCH framework is an extension of the AR process which allows for nonconstant conditional variance among the residuals. Therefore, it is suitable for our analysis in the sense that it reflects time-varying volatility, allowing for the adjustment of index weights based on changes in asset volatilities and correlations. Since time-varying correlations were also present in the time series, we conducted further analysis by building univariate and multivariate models using the dynamic conditional correlation (DCC) method to account for changing covariances and correlations.

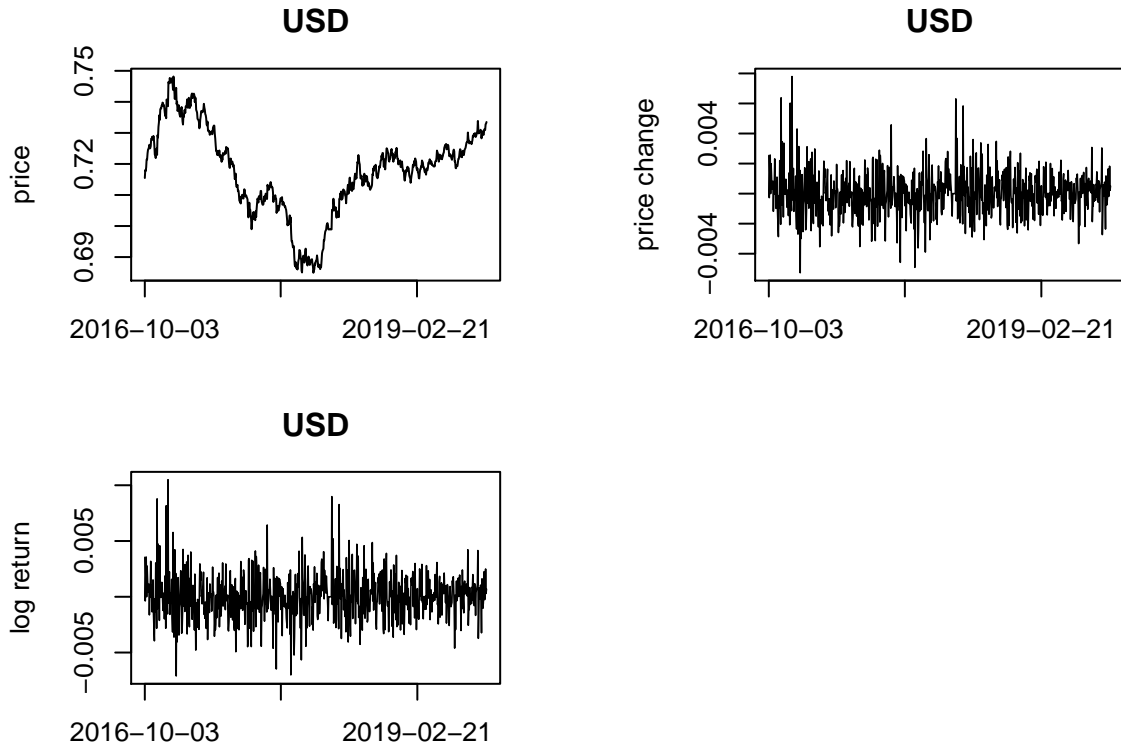


Figure 1: Time series plots of price, change in price and log return of USD from late 2016 to early 2019.

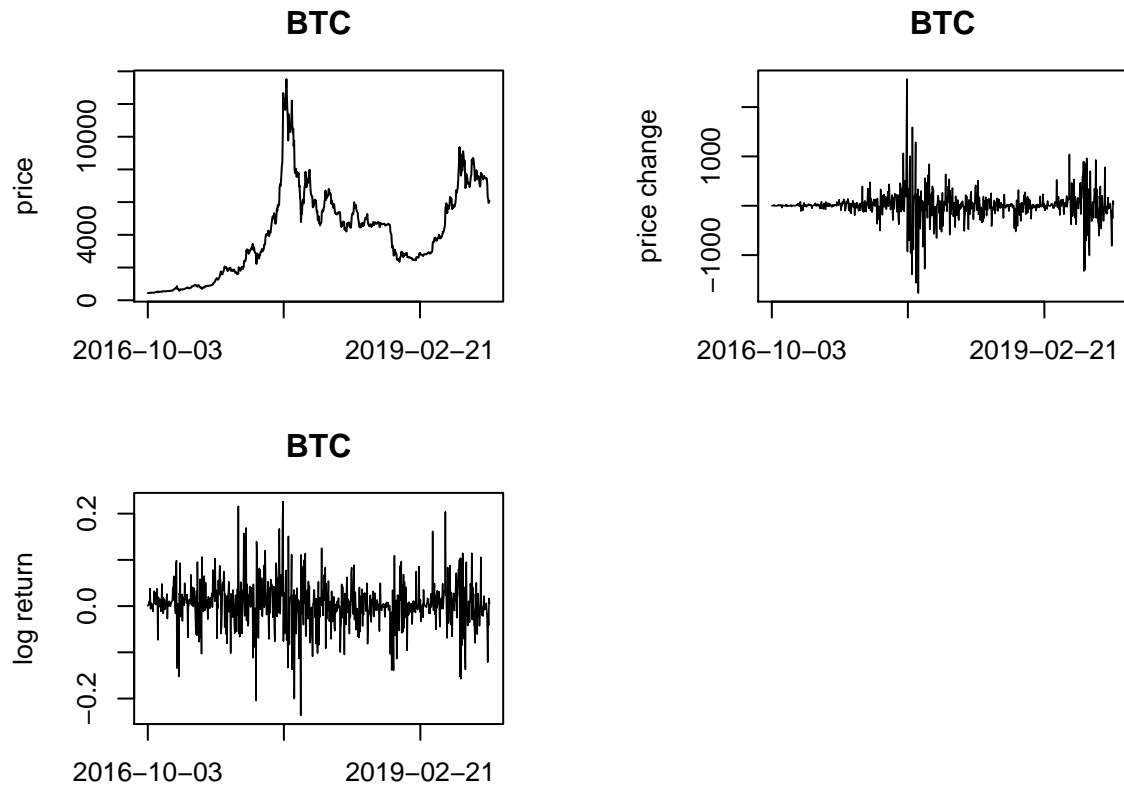


Figure 2: Time series plots of price, change in price and log return of Bitcoin from late 2016 to early 2019.

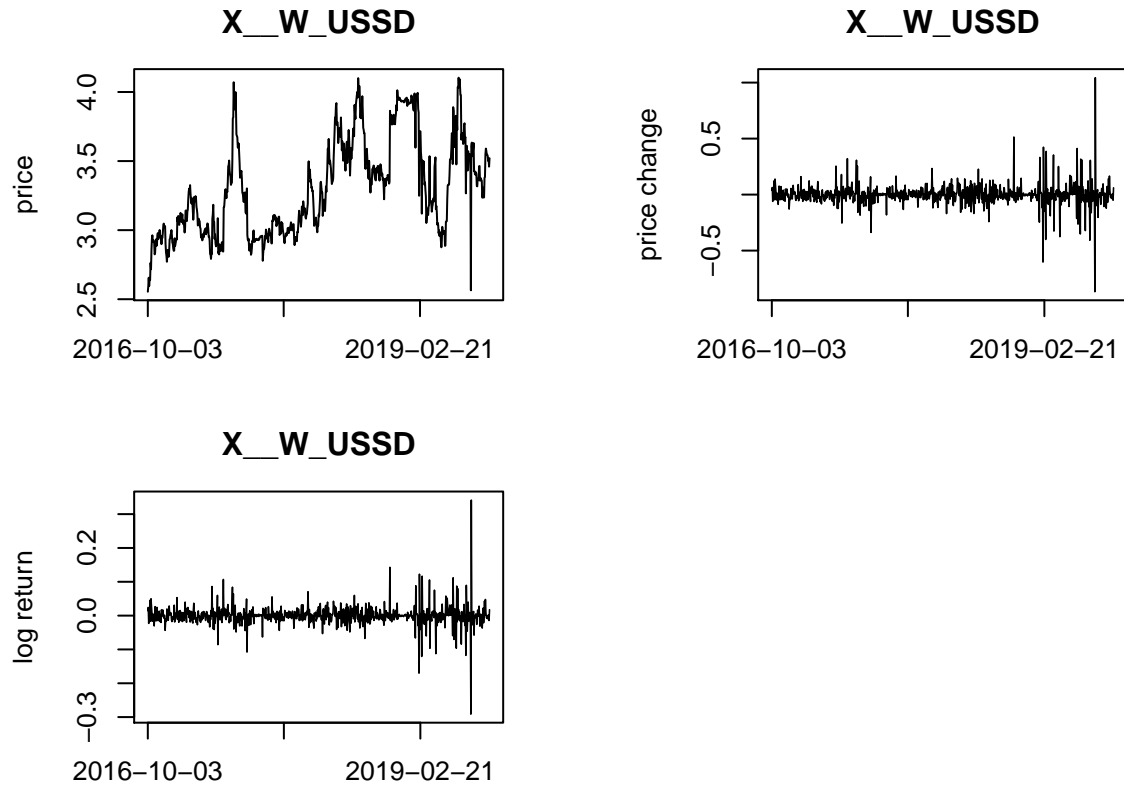


Figure 3: Time series plots of price, change in price and log return of Wheat from late 2016 to early 2019.

2. Practical Issues

Due to the fact that we needed to fit a univariate GARCH model for twelve different assets and each model required different parameter values (i.e. p, q values in ARMA model), we automated the process by utilizing *auto.arima()* function and building custom functions to iterate over each asset and fit a model with the lowest information criterion value. In order to ensure the accuracy and reliability of the automated process, we manually fitted the GARCH models for selective assets. Since both the automated process and the manual process produced the same modeling results, we could conclude that the automated process did not introduce additional errors.

3. Data Wrangling

The data cleaning required prior to constructing the Agriculture 12 index include the following:

- Merging the daily close/settle price of Fiat Currencies assets, Cryptocurrencies assets and Agriculture Commodities assets into one dataframe;
- Setting the data columns to the appropriate data types;
- Removing any weekends or US holiday dates;
- Carrying forward any null values with prices from the last trading day;
- Removing any missing values at the beginning of the series as some data series start from September 30th, 2016 while some starting from October 3rd, 2016;
- Converting the units of measure to Special Drawing Rights (XDR).

In addition to daily prices, we also constructed a daily price difference series and a daily log-return series (for the relevant R script, refer to *process_data.r*).

When building any model, it is important to test its performance on out-of-sample data. In adhering to this principle, the data from the most recent year (i.e. the fourth quarter of 2018 to the third quarter of 2019) were removed from the training set, so that the low volatility indices resulting from the GARCH models could be fairly compared to observed market prices.

4. Building Dynamic Conditional Correlation (DCC) Models

4.1 Overview of GARCH Models

The autoregressive moving average model (ARMA) assumes that the resulting residuals $\{a_t\}$ are a white noise series with mean zero and constant variance. This assumption is often not adequate in financial time series which are known to exhibit volatility clustering. The generalized autoregressive conditional heteroscedastic (GARCH) model extends the ARMA model by modeling the residuals of the ARMA model for the mean. The innovation at time t , a_t is of the form GARCH(m, s) if

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2$$

where $\{\epsilon_t\}$ is an iid sequence of random variables with mean zero and variance 1 and $\alpha_i \geq 0, \beta_j \geq 0$, and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$.

If we let $\eta_t = \alpha_t^2 - \sigma_t^2$ then $\{\eta_t\}$ is a martingale difference and we may rewrite the above as

$$a_t^2 = \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^s \beta_j \eta_{t-j}$$

Thus the GARCH model may be thought of as an application of an ARMA model to the squared residuals $\{a_t^2\}$ from the ARMA model fit to the mean of the time series.

Multivariate time series may exhibit time-varying correlation in addition to time-varying variance. To capture this behavior the class of dynamic conditional correlation (DCC) models, models the correlation between two time series similar to how a univariate GARCH model models the variance of a time series.

Before fitting the multivariate DCC model, we first needed to construct univariate GARCH models for each individual component of the index. Using the ARMA + GARCH method, we first fitted an ARMA model to the series by setting the maximum number of AR terms and MA terms to 2 respectively. This was followed by fitting a second ARMA model to the squared residuals of the first model. Finally, we used the *ugarchspec()* function from the *rugarch* package to generate the GARCH model by setting the “mean.model” parameter based on the first ARMA model and the “variance.model” parameter based on the second one.

The GARCH model we chose is standard GARCH and we assumed that the conditional density for the innovations follow a normal distribution so that we could limit the number of parameters involved and improve the modeling speed. Note that since the order of the volatility model must be specified for the *ugarchspec* function to work, in the event that the order of the second ARMA model—chosen by *auto.arima*—ends up being $c(0, 0)$, we replaced the result with an ARCH(1) process, which is the simplest GARCH model.

Ljung-Box test conducted on USD daily price indicates the presence of significant serial autocorrelation at $K = 5$ lags, and thus the null hypothesis of white noise process should be rejected:

```
##
## Box-Ljung test
##
## data: price_data[, 2]
## X-squared = 3655.7, df = 5, p-value < 2.2e-16
```

The GARCH model specification for the univariate USD raw price indicates that an ARIMA(0, 0, 2) + GARCH(2, 2) model is the optimal fit.

AIC and BIC for the model fit are -8.4587 and -8.4096, respectively (normalized by dividing by the number of observations). The weighted Ljung-Box tests indicate that while the standardized residuals exhibit significant serial autocorrelation, the standardized squared residuals is a white noise process at $K = 4$ lags. The Adjusted Pearson Goodness-of-Fit test, which compares the distribution of the standardized residuals against the theoretical Gaussian distribution of the standardized residuals, shows that the small p-values strongly reject the null hypothesis that the white noise standardized innovation process follows a Gaussian distribution. Nevertheless, due to technical constraints, when we constructed the multivariate DCC model, we assumed that the distribution of the error term is Gaussian to limit modeling complexity.

```
##
## *-----*
## *          GARCH Model Spec          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model      : sGARCH(2,2)
## Variance Targeting : FALSE
##
## Conditional Mean Dynamics
## -----
## Mean Model       : ARFIMA(0,0,2)
```

```

## Include Mean      : TRUE
## GARCH-in-Mean    : FALSE
##
## Conditional Distribution
## -----
## Distribution : norm
## Includes Skew   : FALSE
## Includes Shape   : FALSE
## Includes Lambda : FALSE
##
## *-----*
## *          GARCH Model Fit          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model   : sGARCH(2,2)
## Mean Model    : ARFIMA(0,0,2)
## Distribution   : norm
##
## Optimal Parameters
## -----
##      Estimate  Std. Error    t value Pr(>|t|)
## mu      0.720084    0.000475  1.5161e+03 0.000000
## ma1     1.335115    0.030871  4.3248e+01 0.000000
## ma2     0.665274    0.026031  2.5557e+01 0.000000
## omega   0.000001    0.000001  8.7124e-01 0.383624
## alpha1  0.000000    0.017454  1.8000e-05 0.999985
## alpha2  0.482768    0.024466  1.9732e+01 0.000000
## beta1   0.472987    0.118307  3.9980e+00 0.000064
## beta2   0.000000    0.076140  2.0000e-06 0.999998
##
## Robust Standard Errors:
##      Estimate  Std. Error    t value Pr(>|t|)
## mu      0.720084    0.001836  392.187636 0.000000
## ma1     1.335115    0.090626  14.732113 0.000000
## ma2     0.665274    0.040789  16.310288 0.000000
## omega   0.000001    0.000004   0.162047 0.871269
## alpha1  0.000000    0.016453   0.000019 0.999984
## alpha2  0.482768    0.386268   1.249824 0.211364
## beta1   0.472987    0.115057   4.110900 0.000039
## beta2   0.000000    0.386510   0.000000 1.000000
##
## LogLikelihood : 3192.718
##
## Information Criteria
## -----
##
## Akaike      -8.4587
## Bayes       -8.4096
## Shibata     -8.4590
## Hannan-Quinn -8.4398
##

```

```

## Weighted Ljung-Box Test on Standardized Residuals
## -----
##               statistic p-value
## Lag[1]                133      0
## Lag[2*(p+q)+(p+q)-1][5]    791      0
## Lag[4*(p+q)+(p+q)-1][9]   1338      0
## d.o.f=2
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -----
##               statistic p-value
## Lag[1]                0.06327 0.8014
## Lag[2*(p+q)+(p+q)-1][11]    8.09239 0.2088
## Lag[4*(p+q)+(p+q)-1][19]   12.54063 0.2244
## d.o.f=4
##
## Weighted ARCH LM Tests
## -----
##           Statistic Shape Scale P-Value
## ARCH Lag[5]      2.710 0.500 2.000 0.09971
## ARCH Lag[7]      4.557 1.473 1.746 0.15113
## ARCH Lag[9]      6.186 2.402 1.619 0.16167
##
## Nyblom stability test
## -----
## Joint Statistic: 16.8538
## Individual Statistics:
## mu      3.4794
## ma1     0.6766
## ma2     0.1926
## omega   3.4896
## alpha1  0.2032
## alpha2  0.5409
## beta1   0.3233
## beta2   0.1683
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:      1.89 2.11 2.59
## Individual Statistic: 0.35 0.47 0.75
##
## Sign Bias Test
## -----
##           t-value   prob sig
## Sign Bias      2.3929 0.01696 **
## Negative Sign Bias 2.0896 0.03699 **
## Positive Sign Bias 0.2976 0.76606
## Joint Effect    7.5758 0.05564 *
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----
##   group statistic p-value(g-1)
## 1    20      215.3    2.995e-35

```

```
## 2    30    242.2    1.662e-35
## 3    40    254.2    2.229e-33
## 4    50    263.7    3.550e-31
##
##
## Elapsed time : 0.1975489
```

4.2 Fitting Univariate USD GARCH Model Residuals to Normal Distribution

4.2.1 Parameters:

```
##          mean          sd
## -0.0008135023    0.0050177060
## ( 0.0001828554) ( 0.0001292983)
```

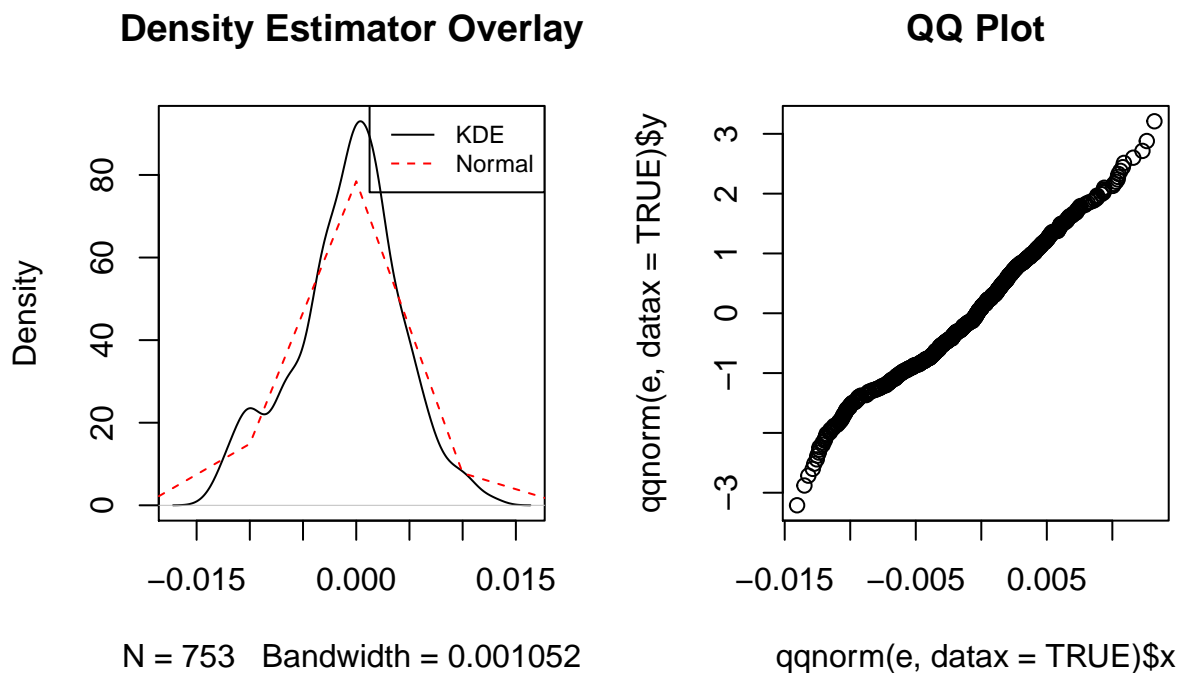


Figure 4: (left) KDE of univariate USD GARCH model residuals versus normal density plot; (right) QQ-plot of univariate USD GARCH model residuals

4.3 Fitting Univariate USD GARCH Model Residuals to t-distribution:

4.3.1 Parameters:

```
## $par
##          mean          sd          nu
## -8.135022e-04  5.017706e-03  8.418954e+06
##
## $objective
## [1] -2918.51
##
## $convergence
## [1] 0
```

```
##
## $iterations
## [1] 68
##
## $evaluations
## function gradient
##      105      337
##
## $message
## [1] "relative convergence (4)"
```

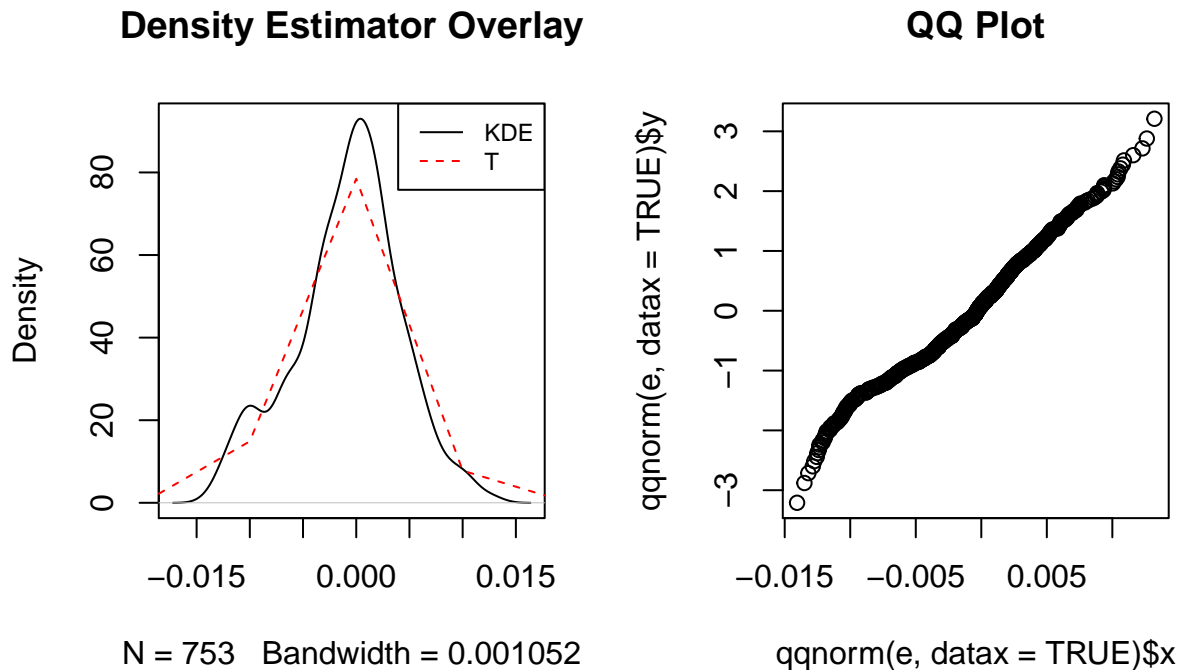


Figure 5: (left) KDE of univariate USD GARCH model residuals versus density plot of t -distribution; (right) QQ-plot of univariate USD GARCH model residuals

Fitting the residuals to the normal distribution and Student's t -distribution resulted in highly similar kernel density estimation plots and QQ plots (see Figure 4-5), which is likely due to the fact that the estimated statistics from both distributions are very close to 0.

4.4 Multivariate model:

Multivariate Ljung-Box tests conducted on daily asset prices, daily price changes and daily log returns all indicate the presence of significant serial autocorrelation in the first $K = 5$ lags, and thus the null hypothesis of white noise process should be rejected:

```
## K Q(K) d.f. p-value
## 1 5 33453.5 720 0

## K Q(K) d.f. p-value
## 1 5 1758.58 720 0

## K Q(K) d.f. p-value
## 1 5 1365.86 720 0
```


Taking the twelfth model as an example where the twelfth model is the DCC GARCH (1, 1) model fit results based on trading days from two consecutive quarters ending September 30, 2019, we found that AIC and BIC for the model fit are -5.5199 and -2.2278, respectively (normalized by dividing by 127).

```
##
## *-----*
## *          DCC GARCH Fit          *
## *-----*
##
## Distribution      : mvnorm
## Model            : DCC(1,1)
## No. Parameters    : 147
## [VAR GARCH DCC UncQ] : [0+79+2+66]
## No. Series        : 12
## No. Obs.          : 127
## Log-Likelihood    : 497.5114
## Av.Log-Likelihood : 3.92
##
## Optimal Parameters
## -----
##              Estimate Std. Error t value Pr(>|t|)
## [USD].mu      7.2328e-01          NA      NA      NA
## [USD].ma1     1.2353e+00          NA      NA      NA
## [USD].ma2     6.9443e-01          NA      NA      NA
## [USD].omega    0.0000e+00          NA      NA      NA
## [USD].alpha1   0.0000e+00          NA      NA      NA
## [USD].alpha2   6.5257e-01          NA      NA      NA
## [USD].beta1    3.0014e-01          NA      NA      NA
## [USD].beta2    0.0000e+00          NA      NA      NA
## [EURUSD].mu    8.1009e-01          NA      NA      NA
## [EURUSD].ma1   7.1117e-01          NA      NA      NA
## [EURUSD].ma2   3.9265e-01          NA      NA      NA
## [EURUSD].omega 0.0000e+00          NA      NA      NA
## [EURUSD].alpha1 9.0312e-02          NA      NA      NA
## [EURUSD].alpha2 4.1836e-02          NA      NA      NA
## [EURUSD].beta1  8.3865e-01          NA      NA      NA
## [JPYUSD].mu    6.6838e-01          NA      NA      NA
## [JPYUSD].ma1   1.1353e+00          NA      NA      NA
## [JPYUSD].ma2   5.7993e-01          NA      NA      NA
## [JPYUSD].omega 4.0000e-06          NA      NA      NA
## [JPYUSD].alpha1 0.0000e+00          NA      NA      NA
## [JPYUSD].alpha2 3.6318e-01          NA      NA      NA
## [JPYUSD].beta1  4.4352e-01          NA      NA      NA
## [JPYUSD].beta2  0.0000e+00          NA      NA      NA
## [CYNUSD].mu    4.9904e+00          NA      NA      NA
## [CYNUSD].ma1   1.1612e+00          NA      NA      NA
## [CYNUSD].ma2   7.4546e-01          NA      NA      NA
## [CYNUSD].omega 7.9000e-05          NA      NA      NA
## [CYNUSD].alpha1 0.0000e+00          NA      NA      NA
## [CYNUSD].alpha2 6.7757e-01          NA      NA      NA
## [CYNUSD].beta1  2.9634e-01          NA      NA      NA
## [BTC].mu       3.2897e+03          NA      NA      NA
## [BTC].ar1      9.9485e-01          NA      NA      NA
## [BTC].omega    6.7672e+04          NA      NA      NA
```

## [BTC].alpha1	6.5670e-01	NA	NA	NA
## [ETH].mu	1.1174e+02	NA	NA	NA
## [ETH].ar1	9.8132e-01	NA	NA	NA
## [ETH].omega	4.8878e+01	NA	NA	NA
## [ETH].alpha1	4.3070e-01	NA	NA	NA
## [XRP].mu	2.1750e-01	NA	NA	NA
## [XRP].ma1	1.1744e+00	NA	NA	NA
## [XRP].ma2	5.4456e-01	NA	NA	NA
## [XRP].omega	5.9000e-05	NA	NA	NA
## [XRP].alpha1	4.3660e-01	NA	NA	NA
## [XRP].alpha2	4.5447e-01	NA	NA	NA
## [XRP].beta1	1.0793e-01	NA	NA	NA
## [XRP].beta2	0.0000e+00	NA	NA	NA
## [XMR].mu	4.5313e+01	NA	NA	NA
## [XMR].ar1	1.0000e+00	NA	NA	NA
## [XMR].omega	9.3212e+00	NA	NA	NA
## [XMR].alpha1	2.7349e-01	NA	NA	NA
## [X__W_USSD].mu	3.2522e+00	NA	NA	NA
## [X__W_USSD].ar1	9.3284e-01	NA	NA	NA
## [X__W_USSD].omega	1.6956e-02	NA	NA	NA
## [X__W_USSD].alpha1	2.4063e-01	NA	NA	NA
## [X__SYB_TD].mu	6.1510e+00	NA	NA	NA
## [X__SYB_TD].ma1	6.5284e-01	NA	NA	NA
## [X__SYB_TD].ma2	5.0392e-01	NA	NA	NA
## [X__SYB_TD].omega	3.0200e-04	NA	NA	NA
## [X__SYB_TD].alpha1	4.9358e-01	NA	NA	NA
## [X__SYB_TD].alpha2	0.0000e+00	NA	NA	NA
## [X__SYB_TD].beta1	2.3555e-01	NA	NA	NA
## [X__SYB_TD].beta2	2.6987e-01	NA	NA	NA
## [X_ICXD].mu	1.0170e+02	NA	NA	NA
## [X_ICXD].ma1	1.3358e+00	NA	NA	NA
## [X_ICXD].ma2	7.4899e-01	NA	NA	NA
## [X_ICXD].omega	1.4624e-01	NA	NA	NA
## [X_ICXD].alpha1	0.0000e+00	NA	NA	NA
## [X_ICXD].alpha2	4.4853e-01	NA	NA	NA
## [X_ICXD].beta1	4.6179e-01	NA	NA	NA
## [X_ICXD].beta2	0.0000e+00	NA	NA	NA
## [CHRIS.CME_LB1].mu	2.6209e+02	NA	NA	NA
## [CHRIS.CME_LB1].ar1	9.4626e-01	NA	NA	NA
## [CHRIS.CME_LB1].ma1	3.7702e-02	NA	NA	NA
## [CHRIS.CME_LB1].ma2	-8.2457e-02	NA	NA	NA
## [CHRIS.CME_LB1].omega	2.7558e+00	NA	NA	NA
## [CHRIS.CME_LB1].alpha1	0.0000e+00	NA	NA	NA
## [CHRIS.CME_LB1].alpha2	3.5734e-02	NA	NA	NA
## [CHRIS.CME_LB1].beta1	9.0444e-01	NA	NA	NA
## [CHRIS.CME_LB1].beta2	1.0000e-06	NA	NA	NA
## [Joint]dcca1	3.1601e-02	NA	NA	NA
## [Joint]dccb1	7.8858e-01	NA	NA	NA
##				
## Information Criteria				
## -----				
##				
## Akaike	-5.5199			
## Bayes	-2.2278			

```
## Shibata          -6.6364
## Hannan-Quinn -4.1823
##
##
## Elapsed time : 1.580407
```

Conducting multivariate Ljung-Box test on the residuals from the above model fit:

```
##   K   Q(K) d.f. p-value
## 1 5 1726.64 720      0
```

The small p-value indicates that the null hypothesis should be rejected. There is significant serial correlation among the residuals in the first $K = 5$ lags, which is an indicator that our model needs further improvement and parameter tuning.

5. Calculating Price Covariance

We obtained index weights by running quarterly rolling forecasts using the DCC model. We fitted a separate DCC model for each pair of consecutive quarters within the training dataset. Then starting from the third quarter (i.e. Q2'2017), we used the model from previous two quarters of data to forecast the daily close/settle price, daily price difference, daily log return for trading days in the following quarter. The *rcov* function extracts the $N \times N \times T$ covariance matrix produced by *dccforecast* with the diagonal of the matrix representing asset variances over the forecast horizon. Note here that N stands for number of assets while T stands for number of forecasting days.

Constructing variance-minimizing portfolio weights given a set of constraints is a quadratic programming problem and we utilized the *solve.QP* function to obtain the vector containing the set of weights that minimize the sum of asset variances. The constraint matrix for index weights include the following two conditions:

1. The weights for risky assets and less risky assets must sum to 1, and
2. Each individual weight must be greater than 1% and less than 25%.

Given that assets with large and volatile variances, such as cryptocurrencies assets, could cause the solver function to crash, we set the maximum variance threshold to be 10^{-6} and manually assigned an index weight of 0.01 to those risky assets. Combining the daily forecasts with quarterly portfolio weights gave us the agriculture index's performance over time.

Note that using *dccforecast* only makes sense for raw daily price predictions, because if we were to use *dccforecast* to obtain daily price differences and log-returns, we would need to run 1-day, 2-day, \dots , n -day ahead predictions for every trading day within a quarter to obtain the end-of-quarter price, which would be computationally inefficient. Therefore, for daily price differences and log-returns, we ran 10,000 iterations of simulations using *dcgsim* to obtain the daily changes and added them back to the preceding price to get the actual daily price.

5.1 Correlation plot associated with the first time period:

Obviously, the historical correlations between USD and Bitcoin, USD and Wheat, Wheat and Bitcoin all vary over time (see Figures 6-8), justifying the need of using a DCC GARCH framework.

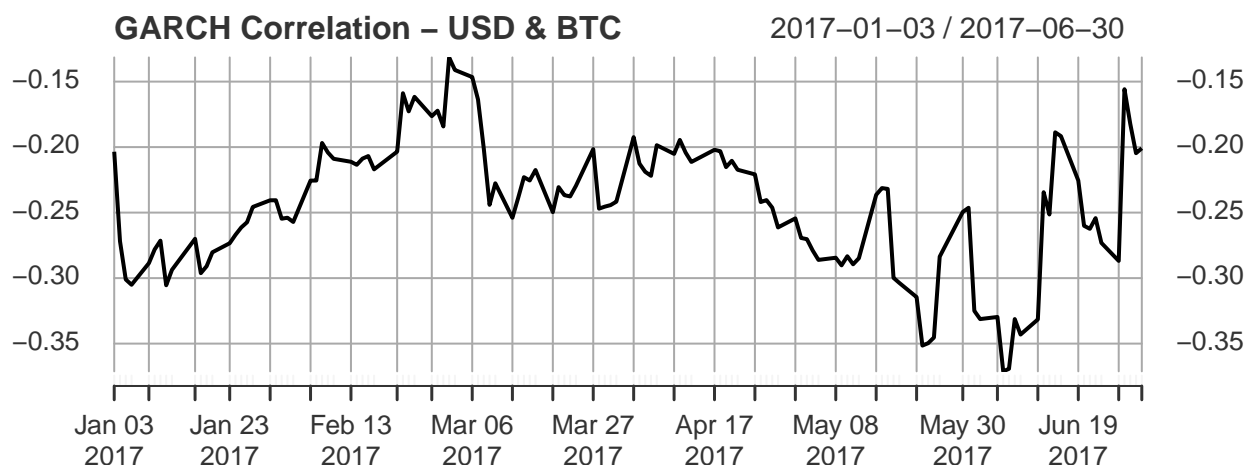


Figure 6: daily correlation estimates between USD and Bitcoin between Jan 3, 2017 to June 30, 2017

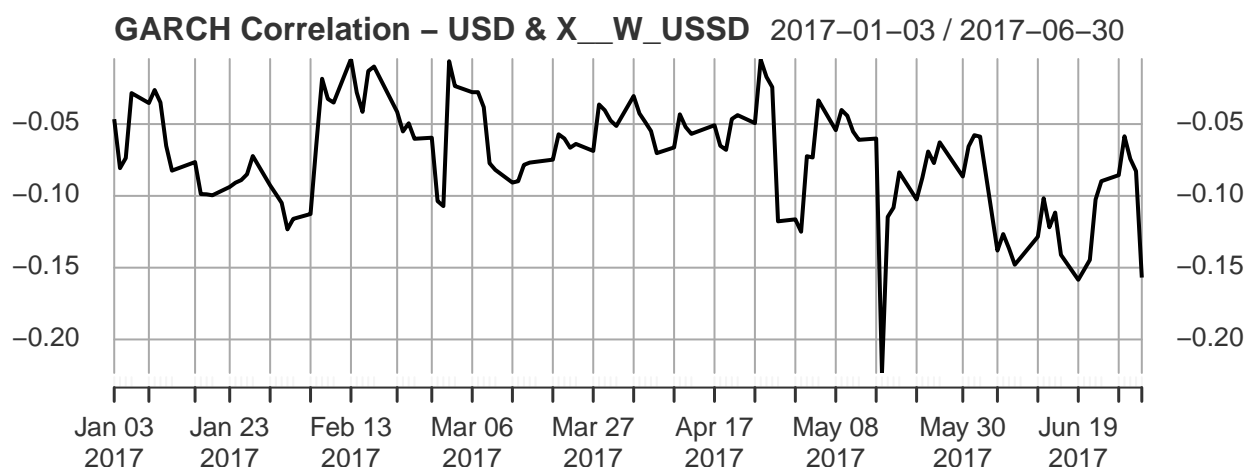


Figure 7: daily correlation estimates between USD and Wheat between Jan 3, 2017 to June 30, 2017

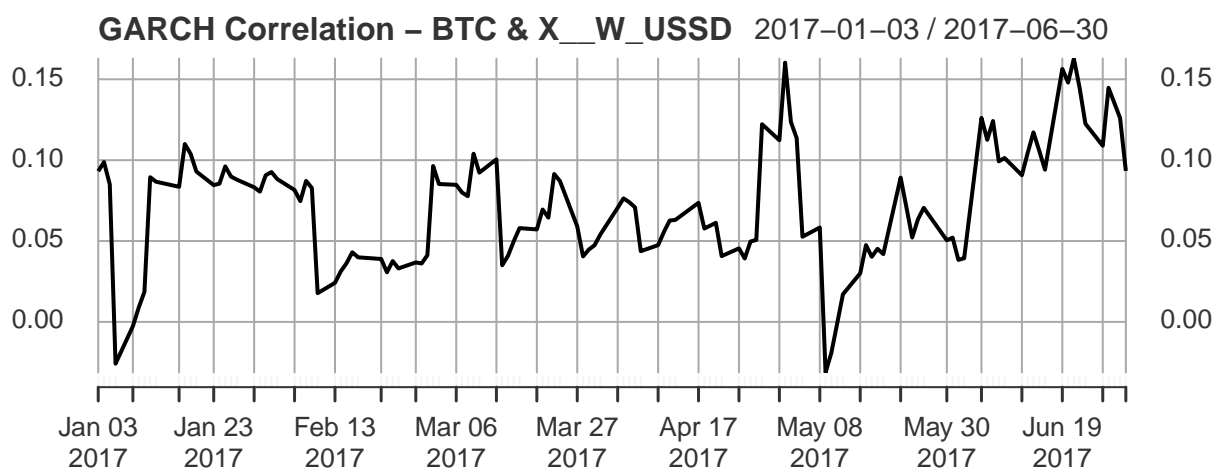


Figure 8: daily correlation estimates between Bitcoin and Wheat between Jan 3, 2017 to June 30, 2017

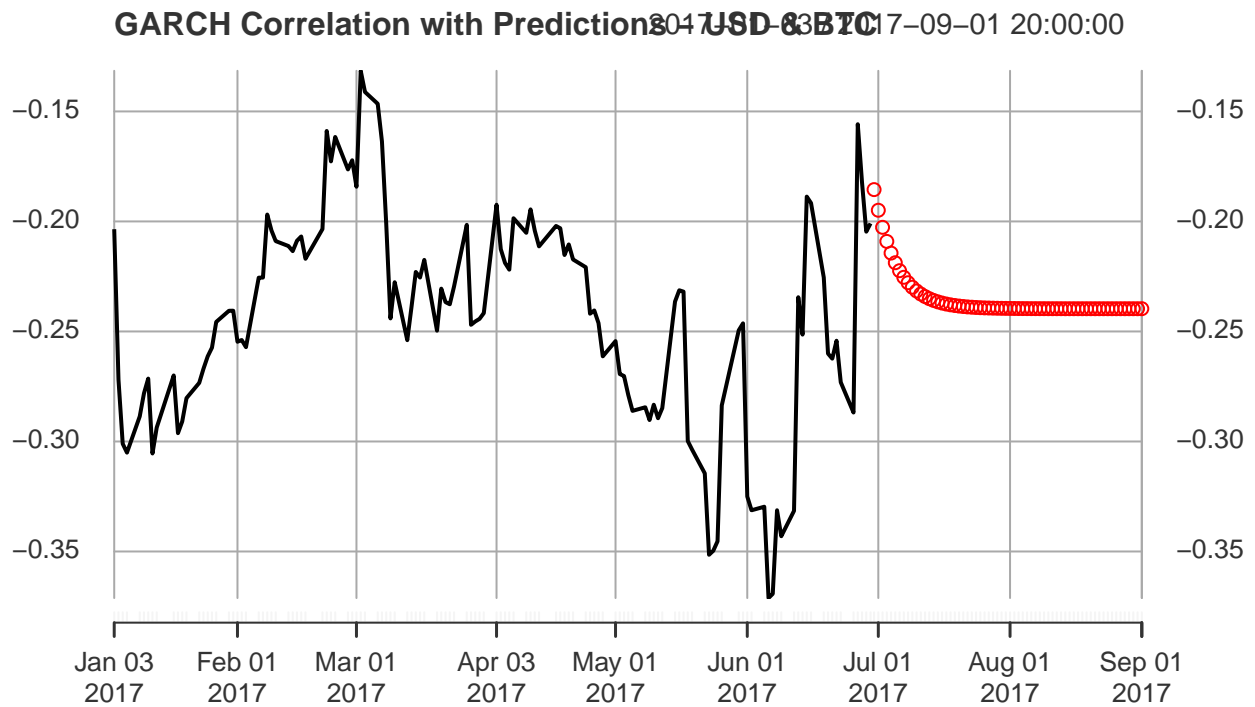


Figure 9: daily correlation estimates between USD and Bitcoin between Jan 3, 2017 to June 30, 2017 with forecasts to Sept 30, 2017

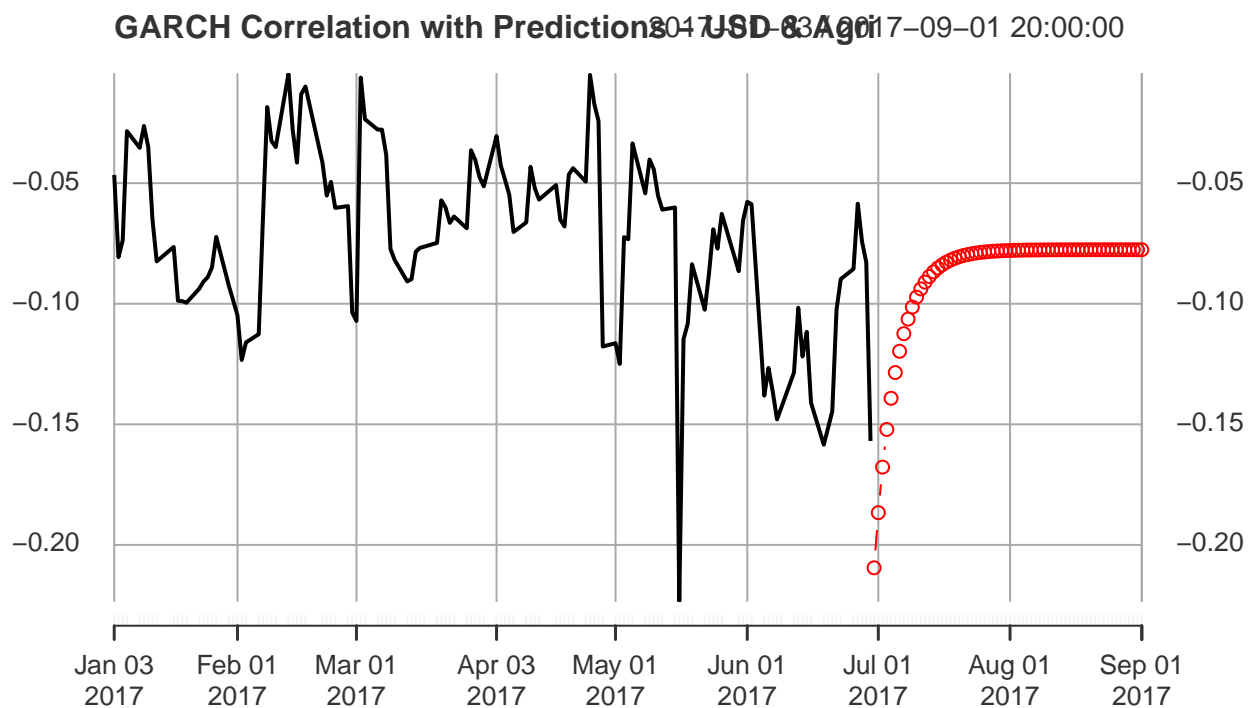


Figure 10: daily correlation estimates between USD and Wheat between Jan 3, 2017 to June 30, 2017 with forecasts to Sept 30, 2017

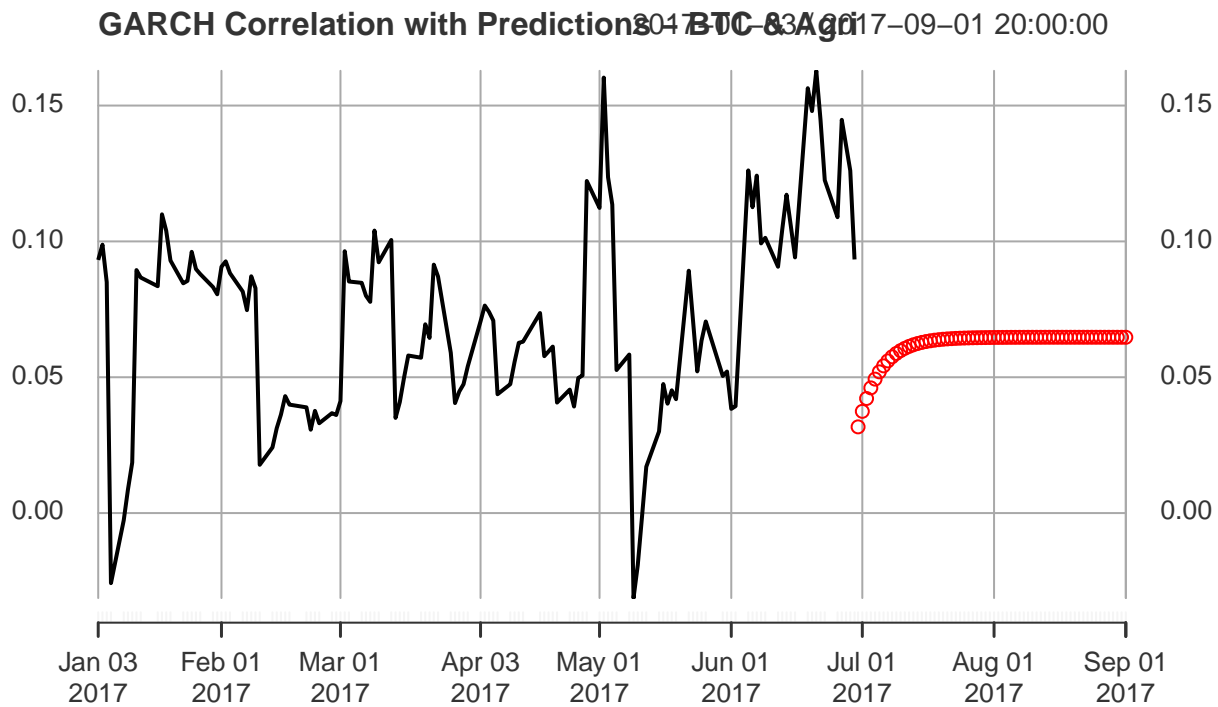


Figure 11: daily correlation estimates between Bitcoin and Wheat between Jan 3, 2017 to June 30, 2017 with forecasts to Sept 30, 2017

Looking at the ACF and CCF plots between USD, BTC, and X__W__USDD in Figures 12-14, we can see some interesting autocorrelation decay and correlation decay patterns:

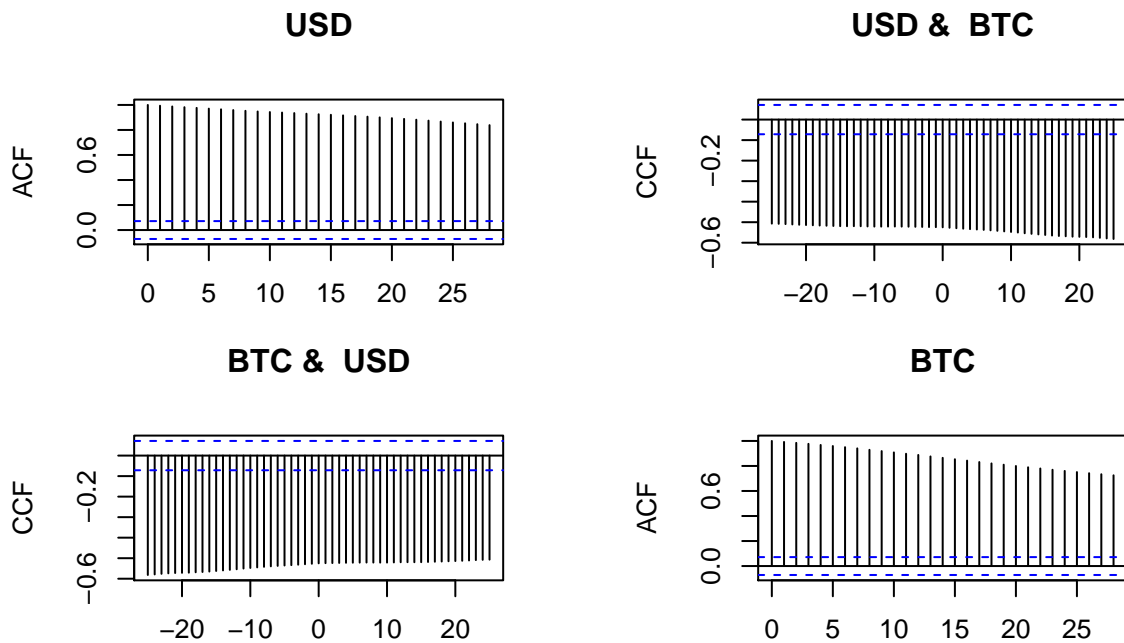


Figure 12: (top left) ACF of USD; (top right) CCF of USD vs. Bitcoin; (bottom left) CCF of Bitcoin vs. USD; (bottom right) ACF of Bitcoin

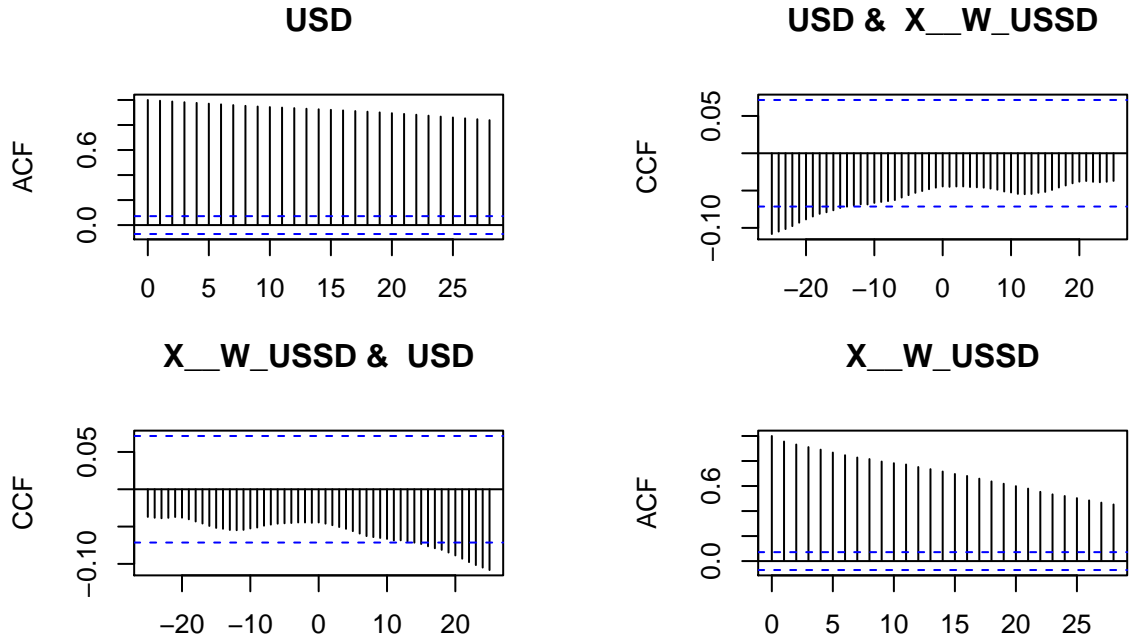


Figure 13: (top left) ACF of USD; (top right) CCF of USD vs. Wheat; (bottom left) CCF of Wheat vs. USD; (bottom right) ACF of Wheat

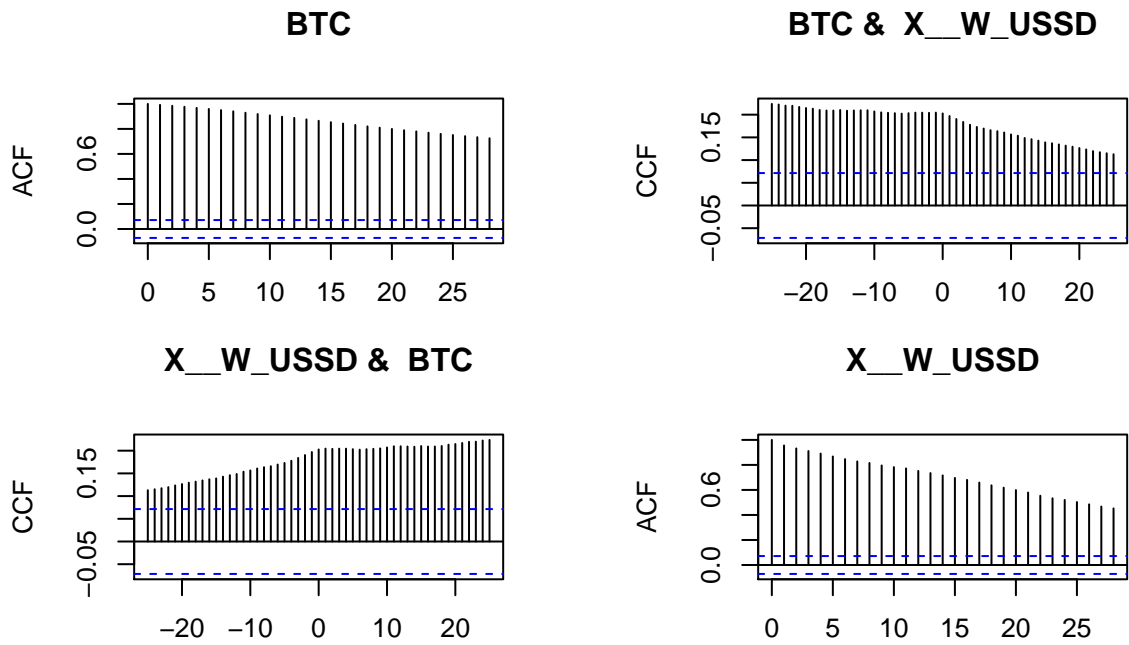


Figure 14: (top left) ACF of BTC; (top right) CCF of BTC vs. Wheat; (bottom left) CCF of Wheat vs. BTC; (bottom right) ACF of Wheat

Examining the residuals for the twelfth models by looking at ACF, PACF, and CCF plots in Figures 15-17:

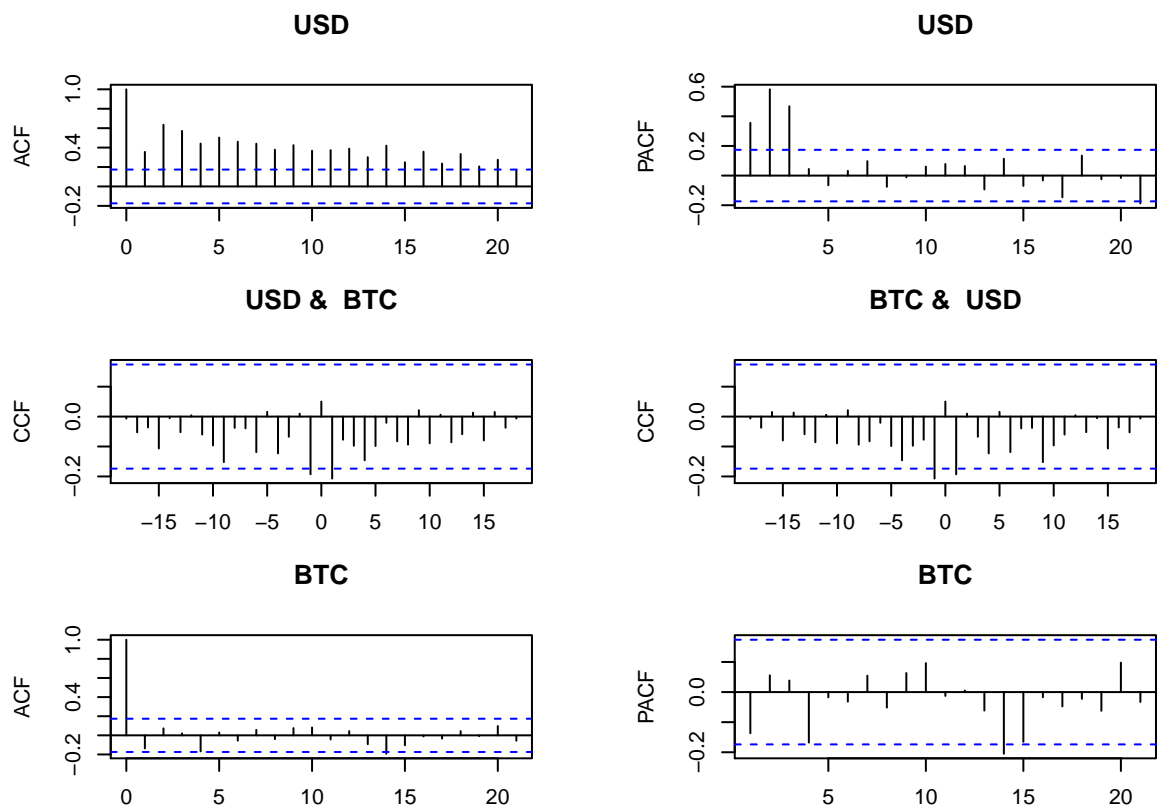


Figure 15: (top left) ACF of USD; (top right) PACF of USD; (middle left) CCF of USD vs. Bitcoin; (middle right) CCF of Bitcoin vs. USD; (bottom left) ACF of Bitcoin; (bottom right) PACF of Bitcoin

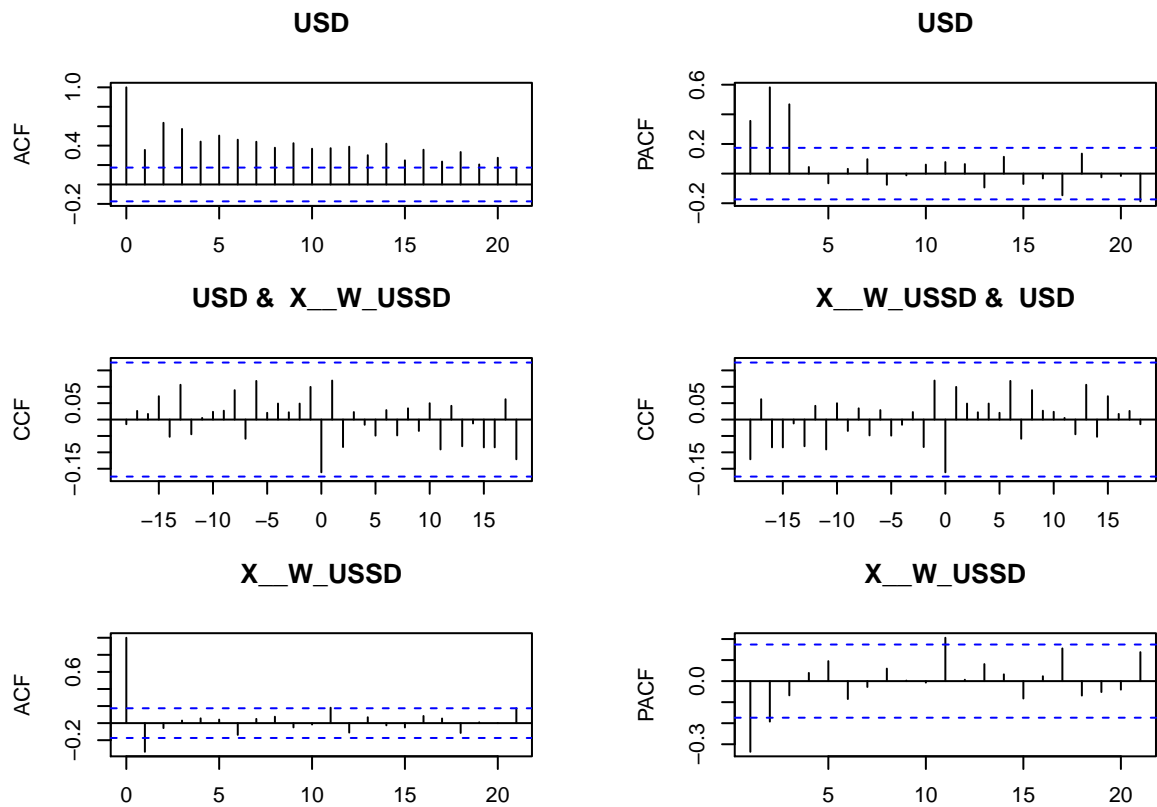


Figure 16: (top left) ACF of USD; (top right) PACF of USD; (middle left) CCF of USD vs. Wheat; (middle

right) CCF of Wheat vs. USD; (bottom left) ACF of Wheat; (bottom right) PACF of Wheat

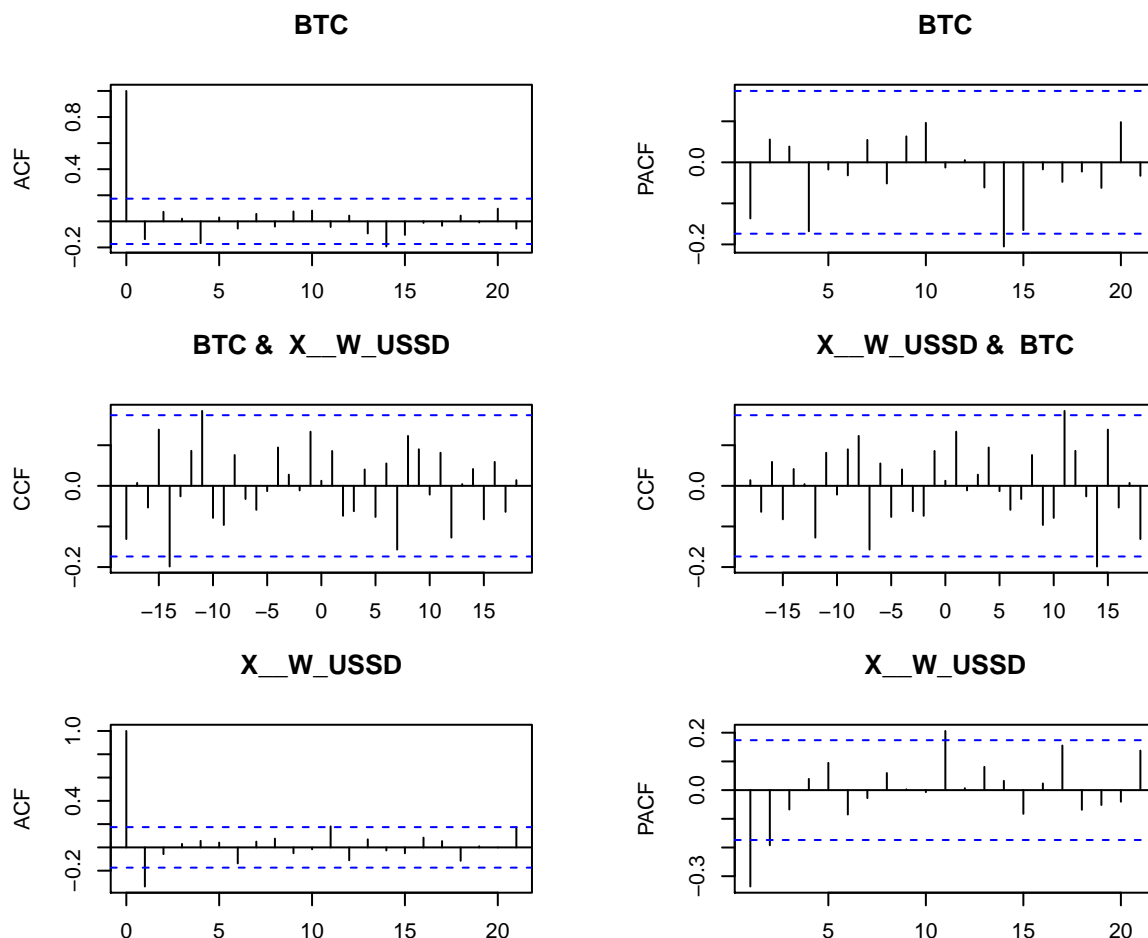


Figure 17: (top left) ACF of Bitcoin; (top right) PACF of Bitcoin; (middle left) CCF of Bitcoin vs. Wheat; (middle right) CCF of Wheat vs. Bitcoin; (bottom left) ACF of Wheat; (bottom right) PACF of Wheat

6. Portfolio Weights

The stacked portfolio weights plots indicate that the currency assets make up most of the entire portfolio, regardless whether the modeling process was based on daily raw prices, daily price changes or daily log-returns (see Figures 18-20).

The first and third portfolios appear similar to each other. The portfolios only consist of a very small proportion of the Cryptocurrency assets, which is in line with our expectations since their price and return are highly volatile. Among the Agricultural Commodities, Wheat and Soy seem to take up a significant portion of the portfolio weights, although towards the end of the analysis time frame, the variance-minimizing policy decided to sell the agricultural commodity and instead allocate the majority weights to Fiat Currencies assets.

For the portfolio based on log-returns, it still retained a significant proportion of Agricultural Commodities towards the end. Interestingly, the re-weighting scheme seems to be balancing the weight between Wheat and Soy: for each four-quarter period, the portfolio started with a majority of Wheat (among the Agricultural Commodities asset class) and then gradually decrease the weights assigned to Wheat while buying more Soy assets instead.

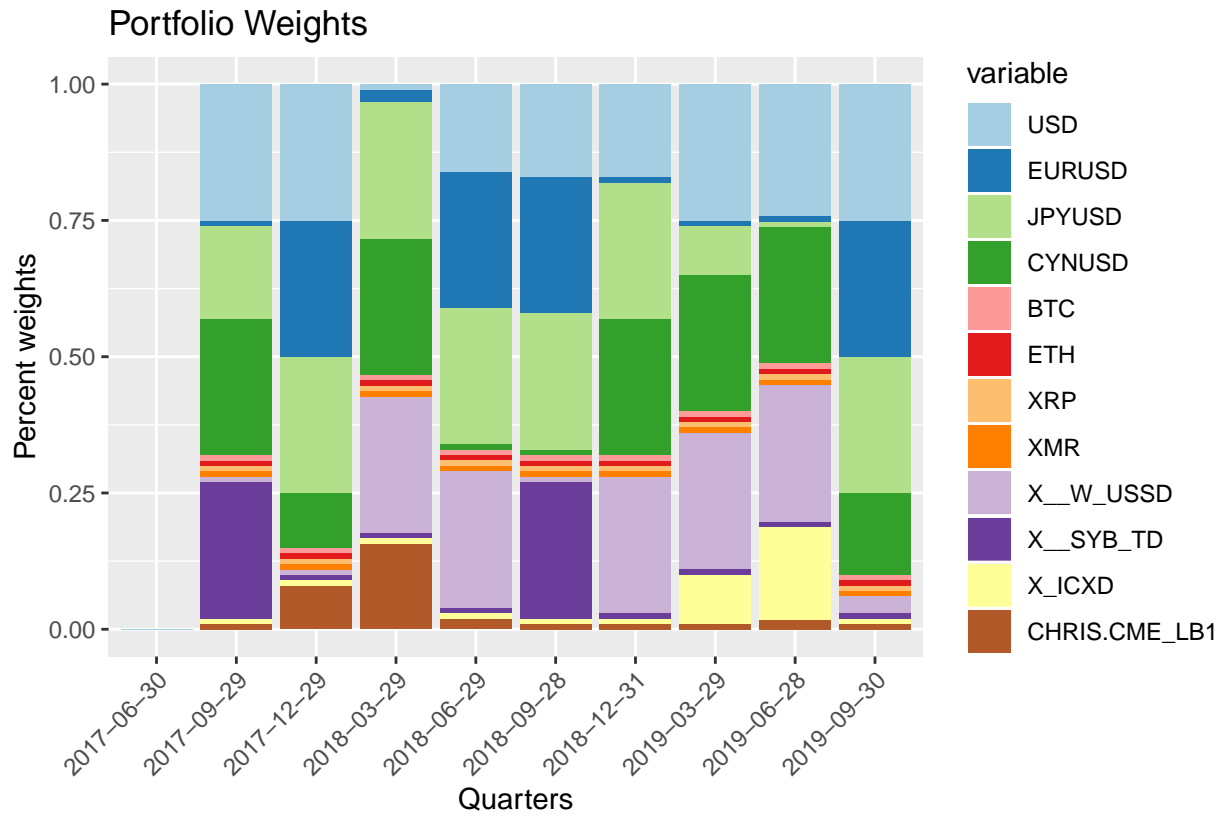


Figure 18: Portfolio reweights for raw prices over June 30, 2017 to Sept 30, 2019

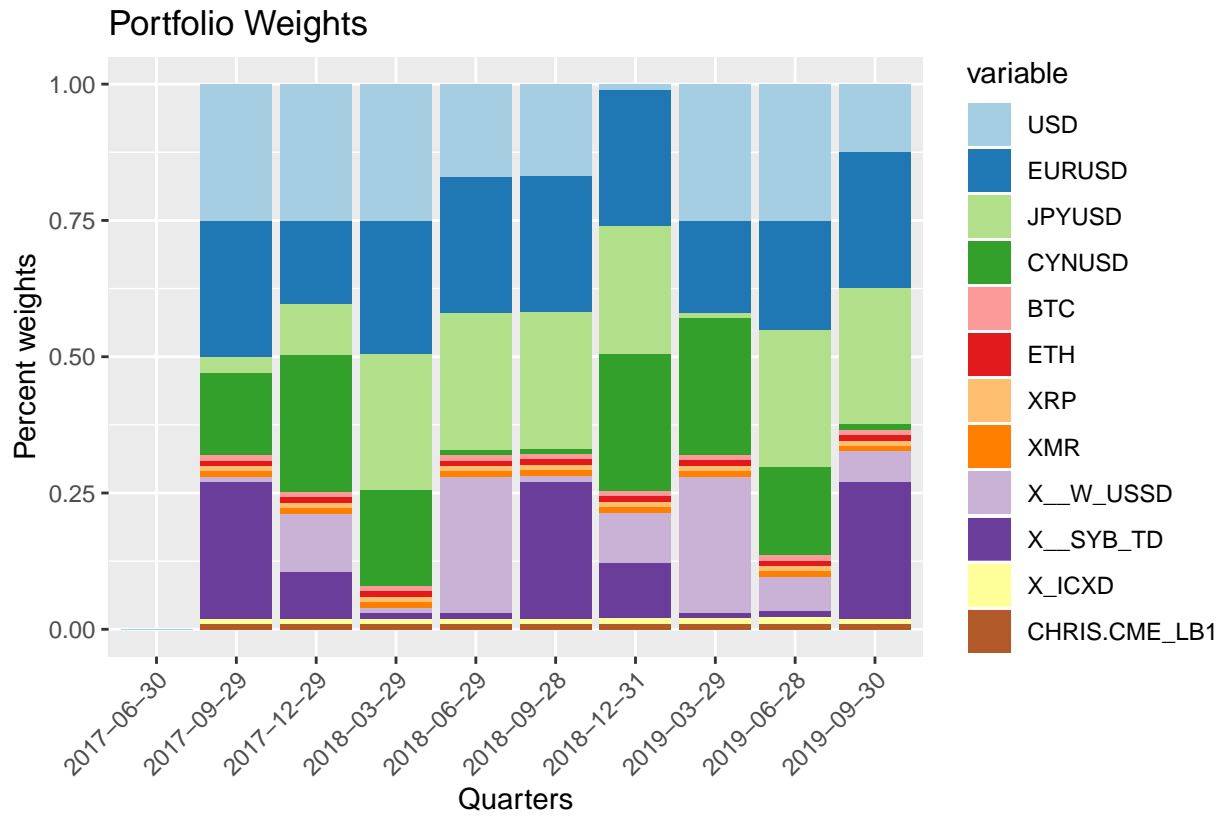


Figure 19: Portfolio reweights for prices difference over June 30, 2017 to Sept 30, 2019

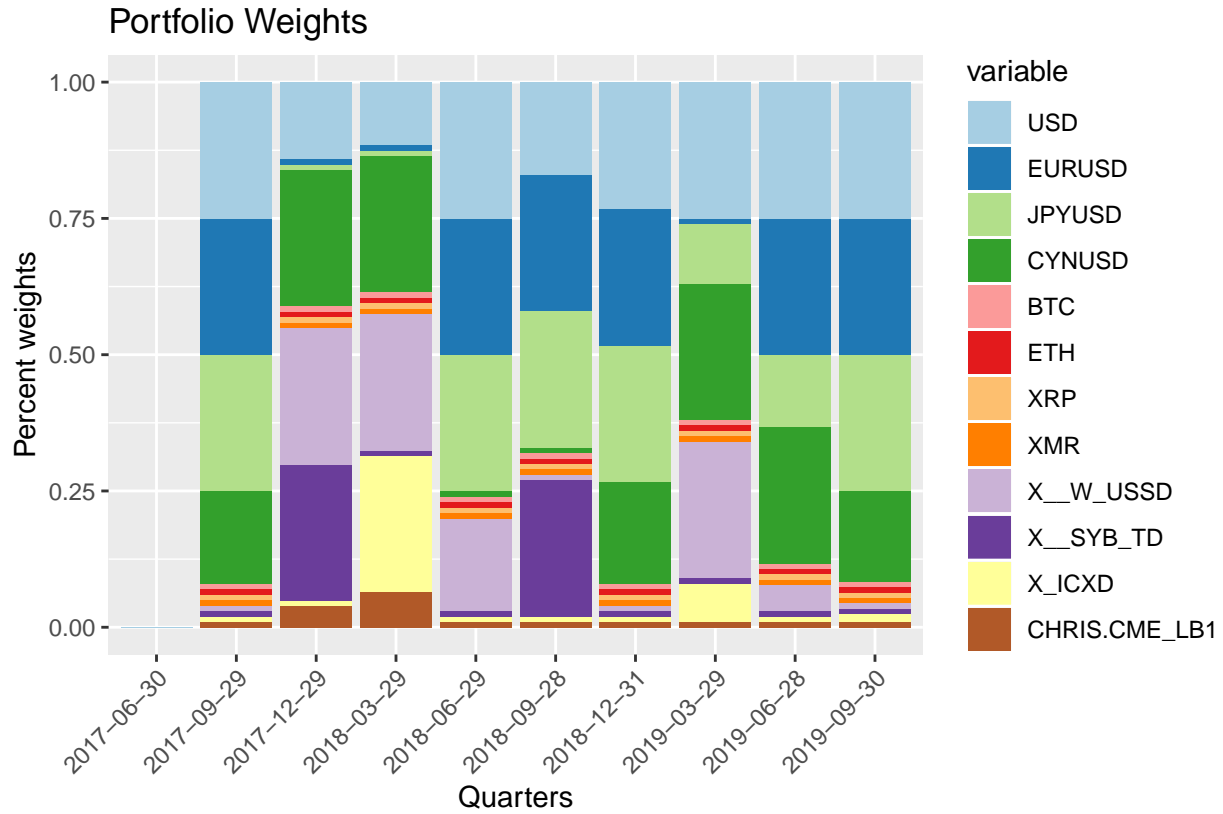


Figure 20: Portfolio reweights for log-returns over June 30, 2017 to Sept 30, 2019

7. Index Performance

From June 30, 2017 to September 30, 2019, the S&P 500 Index achieved a gross return of 22.743% (based on a closing price of 2425.18 on July 3, 2017 and a closing price of 2976.74 on September 30, 2019). Our Agricultural index (modeling and reweighting based on raw daily prices) outperformed the market. Evaluating our portfolio construction mechanism on the out-of-sample data, it achieved a sizable return from October 1, 2018 to September 30, 2019, though the return seems to be reverting back around second half of 2019 (see Figures 21-25).

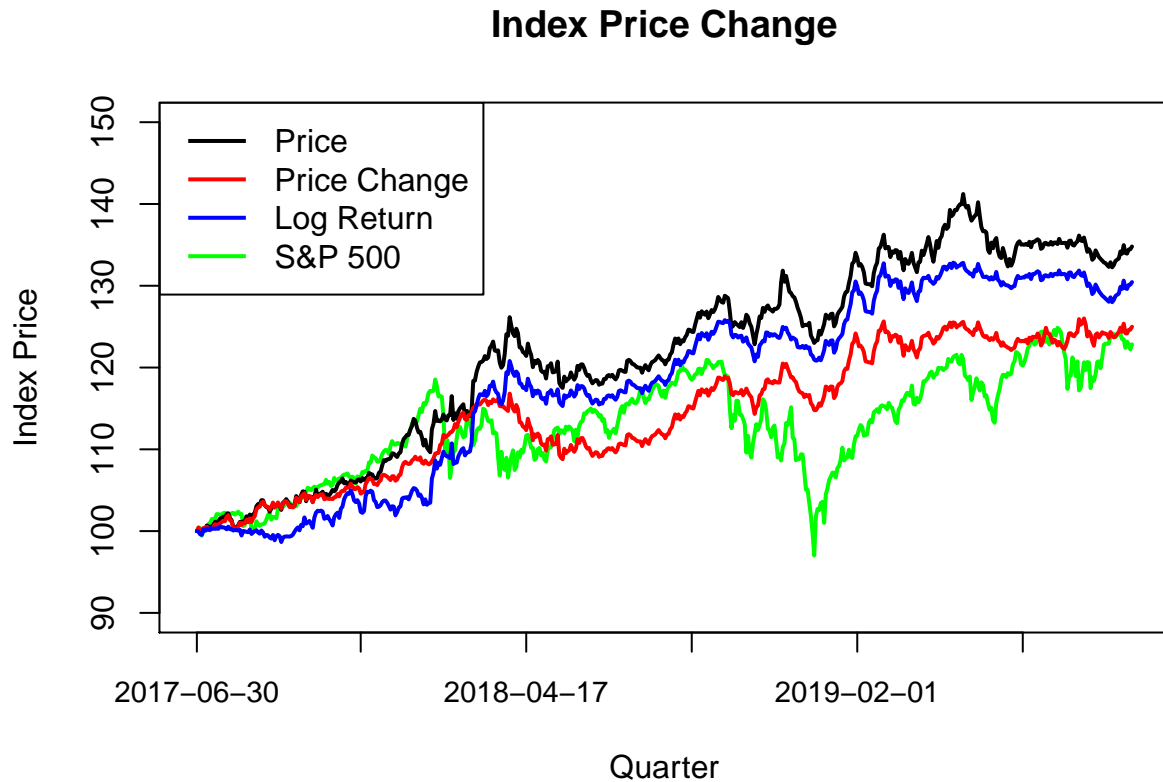


Figure 21: Index price change of raw prices, price difference, log-returns benchmarked against SMP 500 from June 30, 2017 to Sept 30, 2019. (Source: Yahoo Finance SMP 500 Index Adj.close)

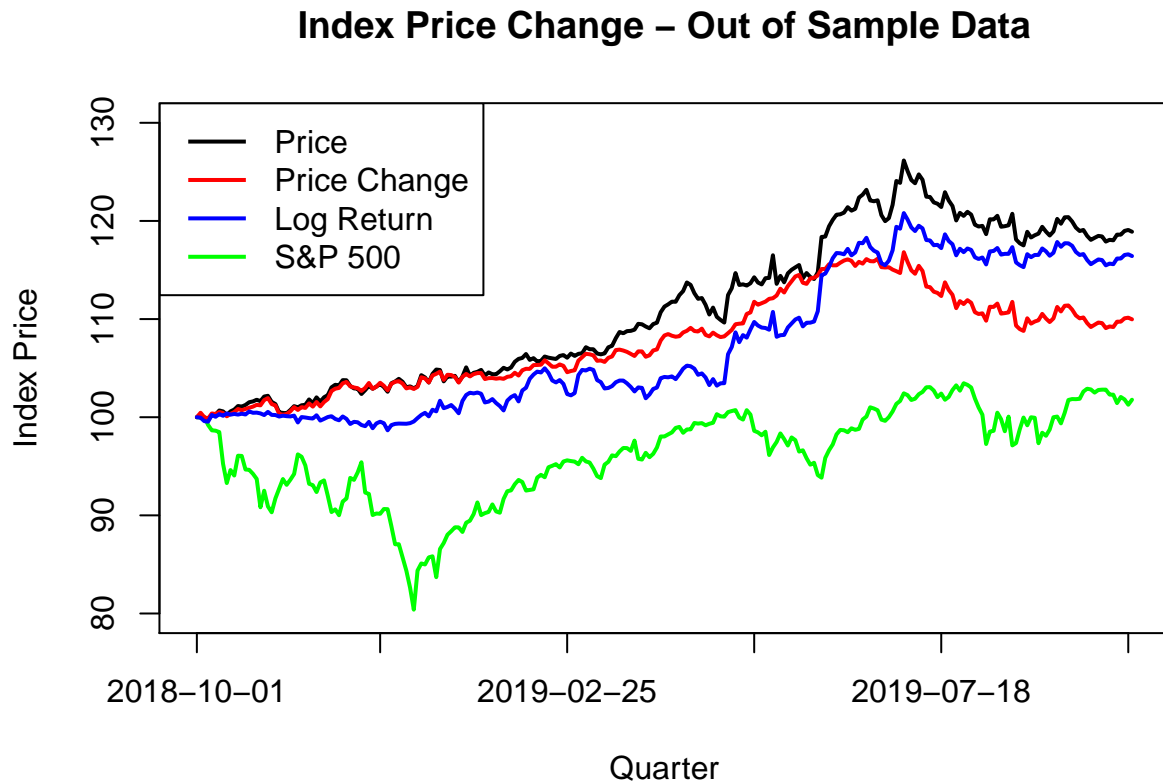


Figure 22: Index price change of raw prices, price difference, log-returns benchmarked against SMP 500 from Oct 1, 2018 to Sept 30, 2019. (Source: Yahoo Finance SMP 500 Index Adj.close)

Index price based on raw price

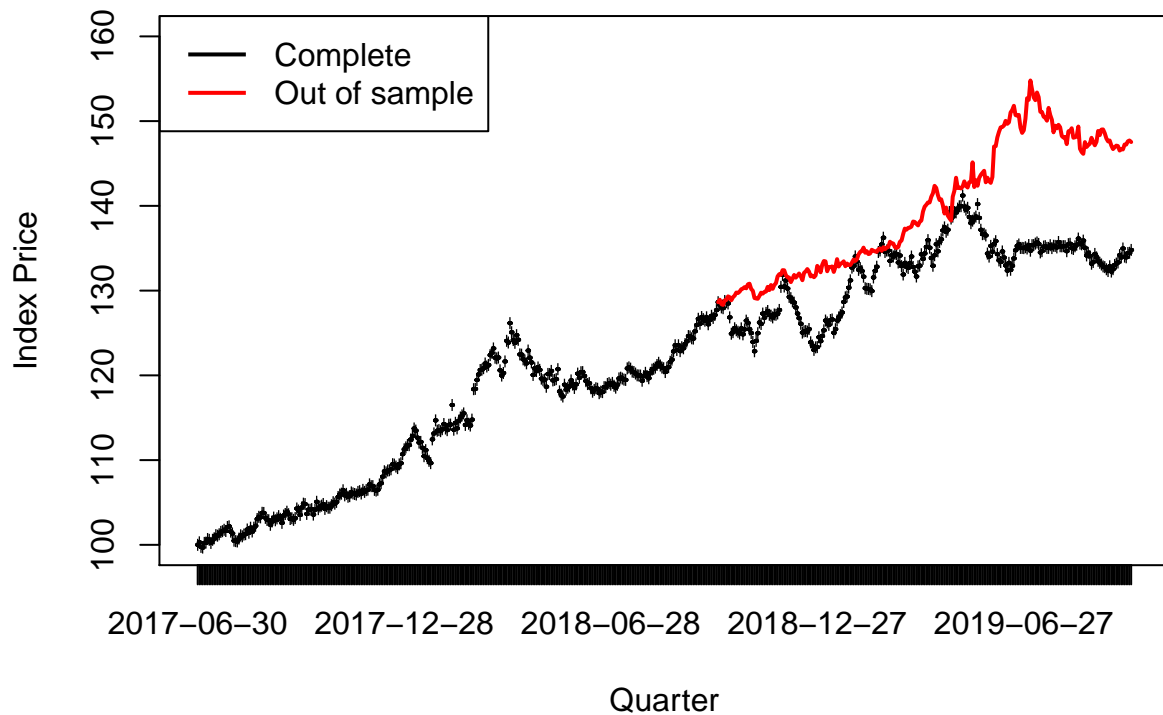


Figure 23: Raw prices index of complete series vs. out-of-sample data

Index price based on price change

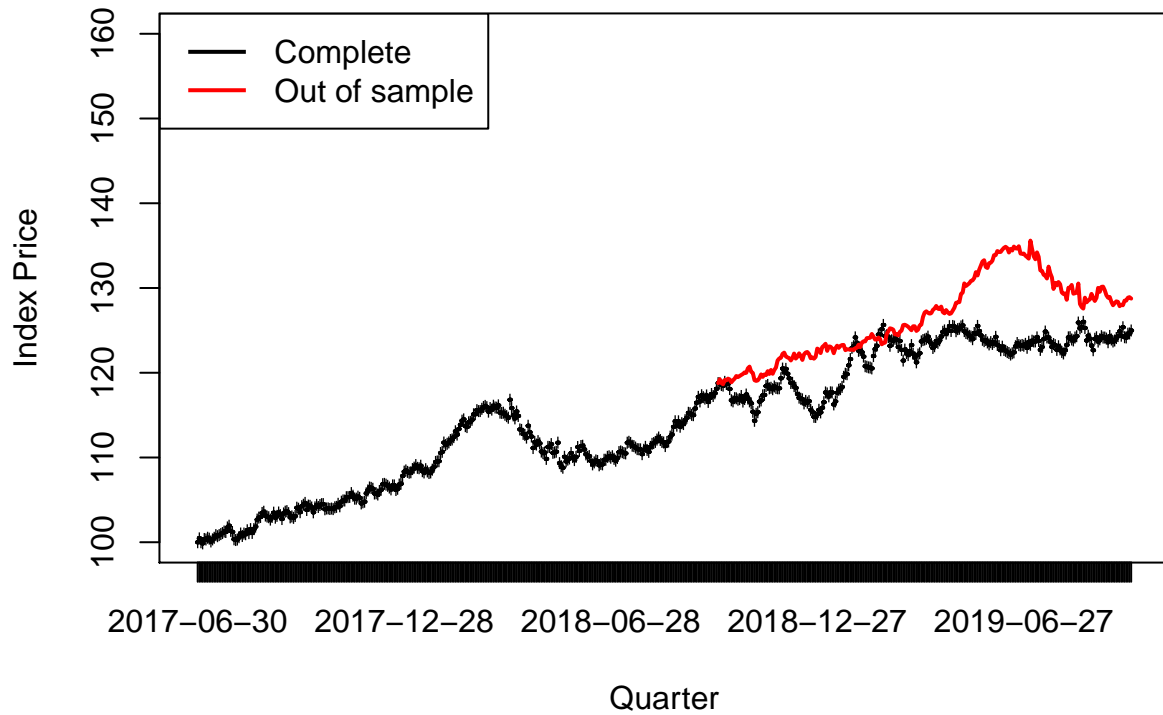


Figure 24: Price differen index of complete series vs. out-of-sample data

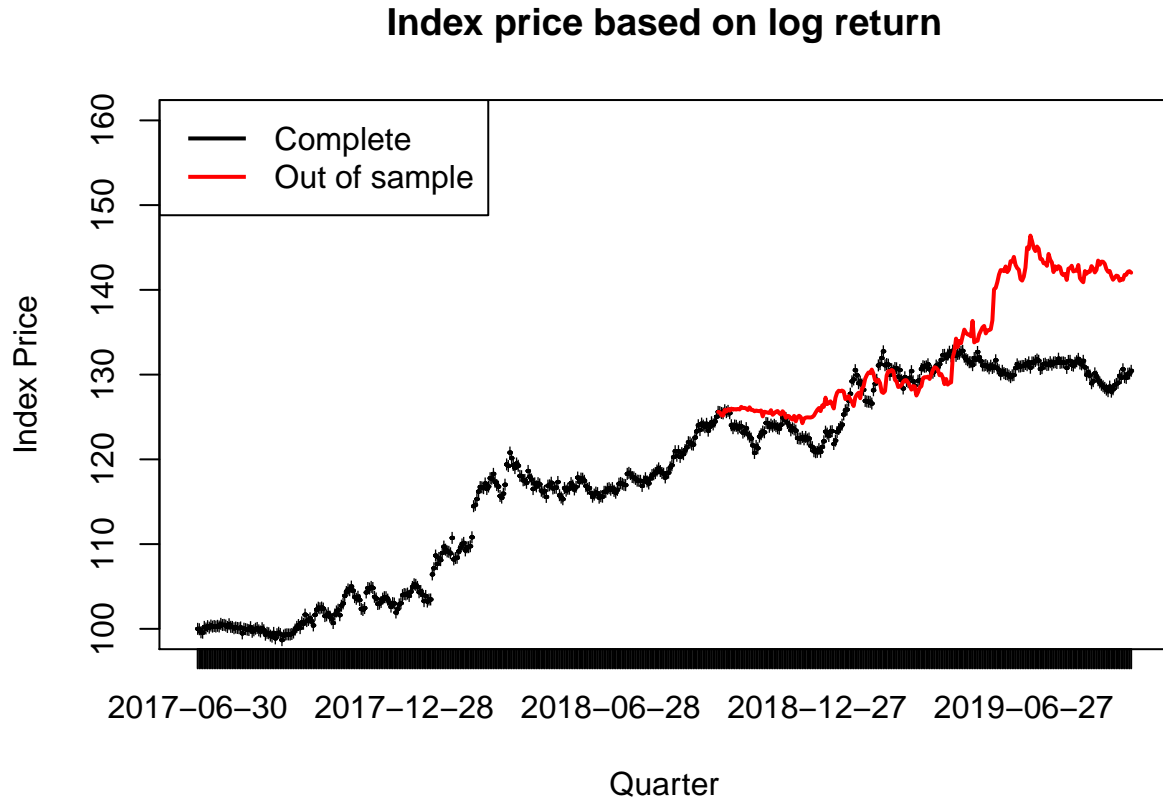


Figure 25: Log-return index of complete series vs. out-of-sample data for log-returns

8. Extension/Future Application

We believe that an immediate improvement to our work could be had by fitting each model on more than just the past two quarters of data. This would allow the GARCH residuals to be modeled by a more appropriate distribution such as a symmetric or skewed t-distribution. It would also allow for us to test more complex GARCH models which may better represent the ground truth of the return series.

In the construction of the low volatility indices, specifically in the formulation of the quadratic programming problem for the portfolio weights, we did not specify a minimum expected return for the portfolio. Specifying a minimum portfolio return would be a simple extension of our work. As most indices are weighted by market cap, we hypothesize that the quadratic programming problem of choosing the minimum volatility weights would still be well defined if we were to introduce the constraint that the low volatility index must be expected to outperform the market cap weighted index by a fixed percent. This extension of our work could be used to construct smart beta indices designed to outperform a particular market cap weighted index while still being composed of the same assets.