

Unconditional Correlation and Conditional Correlation in Commodity Futures and Equity Markets

Alexander Collins, Ruoning Guan, and Fan Lu

April 30, 2019

Intro:

As we know, global macroeconomic environment is one of many essential factors to affect both equity markets and commodity markets. Due to that, we can assume there is some correlation between these two markets, which is especially obvious during volatile market conditions. For example, the correlation coefficient between BCOM index and SPX index reaches 0.66 in Nov 2008.

However this kind of high correlation coefficient is not common. Before 2008, the correlation between BCOM index and SPX index stayed at very low levels for most years and even approaches zero. For example, in 2001 the correlation coefficient was only -0.0023. This paper will attempt to find the correlation between equity and different commodities in various periods and try to analyze the reasons behind these phenomenon.

In particular, the article studies the unconditional correlation and conditional correlation between 3 commodity futures, 1 commodity index and 4 stock indices. It explains how these assets are correlated with each other and how these correlations change over time depending on market conditions.

Data:

The data, obtained from Bloomberg Terminal, comprise returns of 4 commodities and 4 stock indices. The choice of equity indices was selected in such a way that they represent a diverse group of assets. They are the S&P 500 composite index, the Russell 2000 index, the MSCI Europe Index, and the MSCI Latin America index. This choice of equities allows for a comparison in correlation to commodities between different types equity markets.

The dataset also consists of closing prices for 4 commodities, 3 are futures contracts and 1 is a commodity index. For the futures contracts, we consider copper, oil, and gold. The commodity index we consider is the Bloomberg Commodity Index (BCOM). The choice of commodities allows for a diverse comparison among different commodity markets. Future returns are computed as the percentage change in last price.

We have two frequencies for the data, daily and monthly. The dataset spans January 4, 2000 to March 7, 2019 for most series.

Dynamic Conditional Correlation (DCC):

Product price volatility in financial markets usually presents a time-related change trend and this is usually called heteroscedasticity. Thus, people choose the GARCH model to study the benefits and risks of financial markets. The GARCH model was developed by Bollerslev in 1986 and its foundation comes from the ARCH model. The GARCH model introduces the lag term of the fluctuation and reduces the number of parameters that are needed to be estimated.

Based on the GARCH model, other kinds of models such as VEC, BEKK, CCC-GARCH have been developed. These are useful models in time series analysis, however they have some disadvantages. For example, the VEC model has too many parameters which are needed to be estimated and thus affect the veracity of the results. The BEKK model has introduced the definition of conditional positive-definite, but does not have a clear economic significance for those parameters. The CCC-GARCH model has a premise that the observation sequence share a consistent fluctuation, which will not happen in real world because the fluctuation phenomenon usually gather together.

In order to update the CCC model, Engle and Sheppard proposed the DCC-GARCH model which considered the volatility correlation to be a non-fixed constant, which means simulating the heteroscedasticity. The DCC-GARCH model has obvious advantages in estimating parameters, economic significance, and practicality.

Dynamic conditional correlation (DCC) helps show the relationship between commodity and equity returns at various points in time. This method is better than unconditional techniques such as exponential smoothing or rolling correlations since those are very sensitive to volatility changes thus making it difficult to interpret the true nature of the relationship between the variables. The DCC model is a two-step approach to estimating the time-varying correlation between two series. The first step involves using the GARCH model which is used to estimate time-varying variances. Then, a time-varying correlation matrix is estimated by using the standardized residuals from the first-stage estimation.

Empirical Results:

In general, the return correlation between equity indices and commodities is very low from 2000 to the end of 2007. The daily returns show the correlations between the S&P 500 composite index, BCOM, copper futures, oil futures and gold futures. The monthly returns include those assets along with the Russell 2000 index, the MSCI Europe Index, and the MSCI Latin America index in order to show a more diverse range of correlations.

Daily:

	spx_rtn	bcom_rtn	copper_rtn	oil_rtn	gold_rtn
spx_rtn	1.000	0.033	0.174	-0.030	-0.067
bcom_rtn	0.033	1.000	0.420	0.721	0.418
copper_rtn	0.174	0.420	1.000	0.110	0.330
oil_rtn	-0.030	0.721	0.110	1.000	0.197
gold_rtn	-0.067	0.418	0.330	0.197	1.000

Monthly:

	SPXm_rtn	BCOMm_rtn	copperm_rtn	oil_m_rtn	goldm_rtn	Europe_rtn	Latin_rtn	RTYm_rtn
SPXm_rtn	1.000	0.121	0.317	-0.059	0.003	0.843	0.723	0.706
BCOMm_rtn	0.121	1.000	0.404	0.693	0.448	0.055	0.200	0.153
copperm_rtn	0.317	0.404	1.000	0.171	0.272	0.225	0.253	0.199
oil_m_rtn	-0.059	0.693	0.171	1.000	0.300	-0.010	0.114	0.081
goldm_rtn	0.003	0.448	0.272	0.300	1.000	-0.078	0.286	0.183
Europe_rtn	0.843	0.055	0.225	-0.010	-0.078	1.000	0.735	0.719
Latin_rtn	0.723	0.200	0.253	0.114	0.286	0.735	1.000	0.711
RTYm_rtn	0.706	0.153	0.199	0.081	0.183	0.719	0.711	1.000

The average correlation with the S&P 500 Index for the 4 commodities is .028 for daily returns and .096 for monthly returns during this time period.

Looking at the same data but now spanning 2008 to 2019, there is much higher correlation between commodities and stock indices.

Daily:

	spx_rtn	bcom_rtn	copper_rtn	oil_rtn	gold_rtn
spx_rtn	1.000	0.366	0.333	0.346	-0.010
bcom_rtn	0.366	1.000	0.681	0.751	0.439
copper_rtn	0.333	0.681	1.000	0.448	0.325
oil_rtn	0.346	0.751	0.448	1.000	0.223
gold_rtn	-0.010	0.439	0.325	0.223	1.000

Monthly:

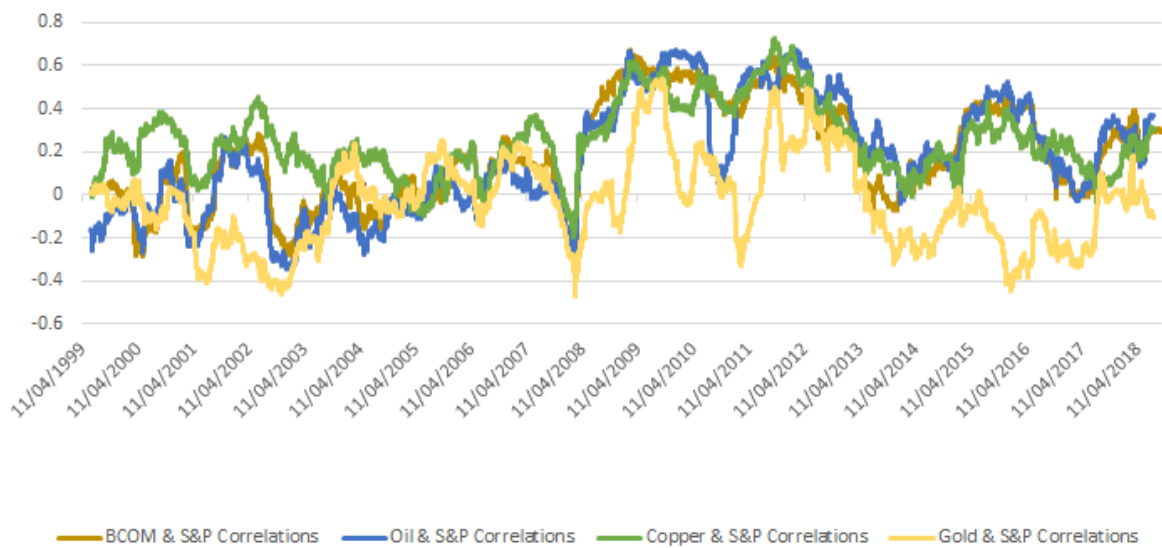
	SPXm_rtn	BCOMm_rtn	copperm_rtn	oilm_rtn	goldm_rtn	Europe_rtn	Latin_rtn	RTYm_rtn
SPXm_rtn	1.000	0.524	0.531	0.472	0.022	0.826	0.660	0.910
BCOMm_rtn	0.524	1.000	0.716	0.754	0.450	0.389	0.662	0.458
copperm_rtn	0.531	0.716	1.000	0.524	0.282	0.401	0.591	0.498
oilm_rtn	0.472	0.754	0.524	1.000	0.169	0.371	0.538	0.441
goldm_rtn	0.022	0.450	0.282	0.169	1.000	-0.076	0.345	-0.002
Europe_rtn	0.826	0.389	0.401	0.371	-0.076	1.000	0.554	0.744
Latin_rtn	0.660	0.662	0.591	0.538	0.345	0.554	1.000	0.553
RTYm_rtn	0.910	0.458	0.498	0.441	-0.002	0.744	0.553	1.000

The average correlation with the S&P 500 Index for the 4 commodities is .218 for daily returns and .387 for monthly returns during this time period.

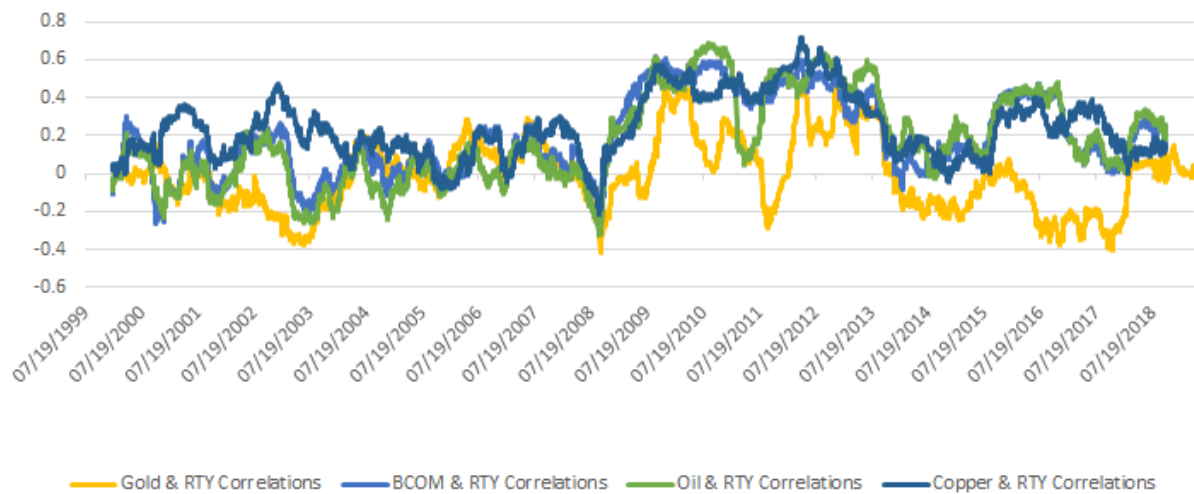
So it is clear that there is a divisive change in correlation between those two periods. If that is so, then how do the correlations change over time in this time frame? This is explained using unconditional and conditional correlations.

Rolling correlations are shown in the following graphs and clearly demonstrate higher correlations beginning in 2008 right after a small period of negative correlation in all 4 commodities. This small period of negative correlation that occurs in mid 2008 is caused by a lag in commodity prices falling. Equity prices fall before the commodity prices did, leading to a small time frame of negative correlations until commodity prices fell as well. These are unconditional correlations and have a rolling period of 120 days. The second figure shows the rolling correlation between the Russell 2000 Index and the Dow Jones Industrial Average index as a benchmark.

Daily Return Correlations (Rolling) Commodites vs. S&P 2000-2019



Daily Return Correlations (Rolling) of Commodities vs. Russell 2000



Unconditional rolling correlations do not control for time-variations in return volatilities and thus could lead to incorrect inferences. As such, we now use Dynamic Conditional Correlation (DCC) to show these changes in correlation.

DCC Model math part:

The most successful volatility forecasting model is the generalized autoregressive conditional heteroskedasticity model, GARCH(1,1). It describes the volatility dynamics of most of the financial return series. The GARCH(1,1) variance, $h_{ii,t}$, is represented by:

$$x_{i,t} = \mu + \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim N(0, h_{ii,t})$$

$$h_{ii,t} = \gamma_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1} \quad i = 1, \dots, N$$

subject to $\gamma_i > 0, \alpha_i, \beta_i \geq 0, \alpha_i + \beta_i < 1$.

In order to estimate the conditional correlation, we use the dynamic conditional correlation model (DCC) of Engle [2002]. Based on estimating the GARCH(1,1) model and employing its resulting standardized residuals, we also estimate a time-varying $H_t = D_t R_t D_t$ correlation matrix via the DCC(1,1). Thus the covariance matrix can be expressed as

$$D_t = \text{diag}(h_{11,t}^{1/2}, \dots, h_{NN,t}^{1/2})$$

is a diagonal matrix of univariate GARCH volatilities and is $R_t = Q_t^{*-1} Q_t Q_t^{*-1}$ the time varying correlation matrix where $Q_t = (q_{ij,t})$.

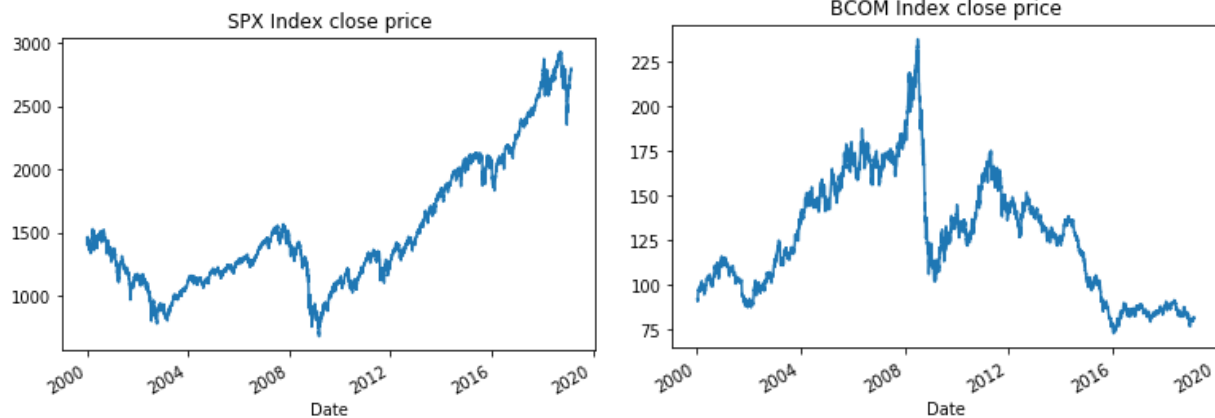
We can use this framework to analyze the conditional returns correlations between commodity futures and traditional assets. We first investigate how they changed over time by regressing them on a constant and a time trend. Second, we use the relation between conditional correlations and conditional volatilities by regressing the former on the latter as:

$$\rho_{TC,t} = \alpha + \beta_T \sqrt{h_{T,t}} + \beta_C \sqrt{h_{C,t}} + \varepsilon_t$$

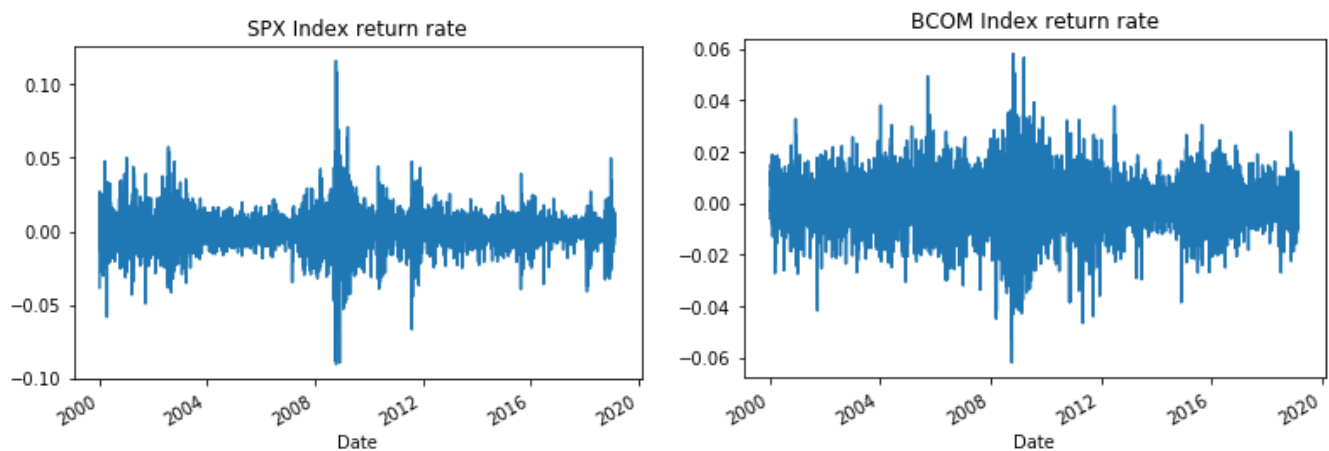
where the subscripts T and C refer to traditional asset class and commodity futures.

Here we Choose SPX Index and BCOM Index as examples to do a detail analysis.

In order to show correlations, we first take a look at their close prices in the most recent two decades. When we look at these two graphs, it's hard to tell what their real correlation is. In fact, one might even deduce that they are negatively correlated since they seem to have opposite trends.



In order to do a more in depth analysis, we graph their return rates. The graphs demonstrate their returns seem to be more correlated, but it is still difficult to tell.

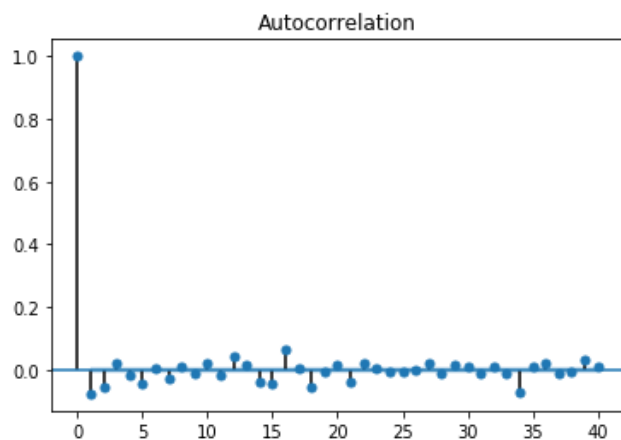


Before we get into a more detail analysis regarding the correlation between SPX Index return rate and BCOM Index return rate, let's take a closer look into the two pictures above. It seems that their return rates could fit a time series model. As you know our topic is using the DCC-GARCH model to find the dynamic correlation between equities and commodities, thus it is important and necessary to prove that their return rates fit the GARCH model.

GARCH model is an expansion of the ARCH model, and the ARCH model has five properties. As an example, we use SPX Index to prove these properties.

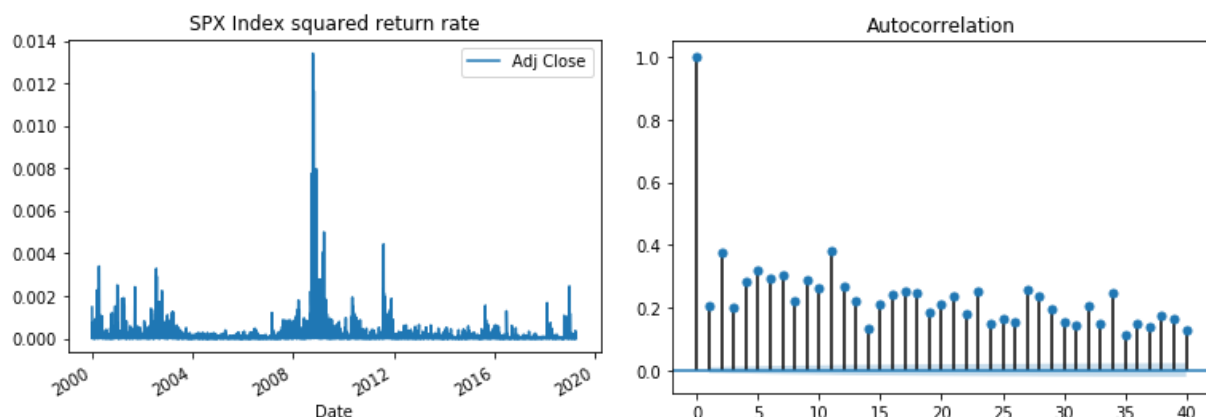
Property 1: No Autocorrelation in Returns

The autocorrelation in the levels of returns demonstrated in following figure shows that predicting the direction of asset returns is not possible.



Property 2: Autocorrelation in Squared Returns

The autocorrelation in the squares of returns demonstrated in following figure shows that while predicting the direction of returns is not possible, predicting their volatility is.



Property 3: Volatility Clustering

The volatility clustering property shows that small movements in returns tend to be followed by small returns in the next period, whereas large movements in returns tend to be followed by large returns in the next period. These movements imply that the autocorrelation of squared returns is positive, a property demonstrated in the last figure.

Property 4: Conditional Normality

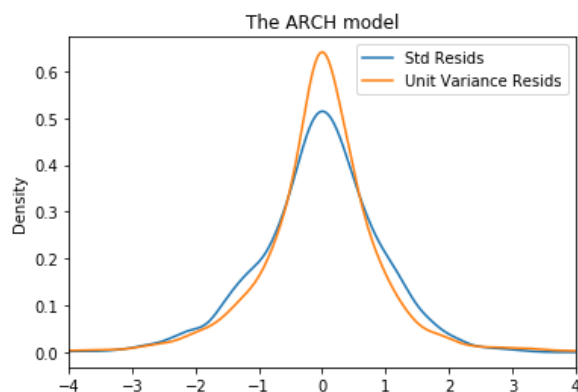
The conditional distribution of returns is normal with conditional mean and conditional variance given by

$$f(y|y_{t-1}) \sim N(0, \alpha_0 + \alpha_1 y_{t-1}^2)$$

For small values of y_{t-1} , the conditional variance is drawn from a relatively compact distribution with mean of zero and approximate variance α_0 . This indicates a high probability of drawing another small value of y in the next period. By contrast, for larger values of y_{t-1} the conditional variance is drawn from a more dispersed distribution with mean zero and variance $\alpha_0 + \alpha_1 y_{t-1}^2$. There is, therefore, a high probability of drawing another large value of y in the next period.

Property 5: Unconditional Leptokurtosis

The unconditional distribution is derived by averaging over all T conditional distributions. Even though the conditional distribution is normal, the unconditional distribution is not. For the relatively low-volatility conditional distributions, the normal distributions are relatively compact with high peaks, whereas, for the relatively high-volatility conditional distributions, the normal distributions are relatively more dispersed with low peaks. Averaging across the conditional distributions yields a nonnormal unconditional distribution, $f(y)$, that has fat-tails and a sharp peak compared to the normal distribution. A distribution with these two properties is said to exhibit leptokurtosis.

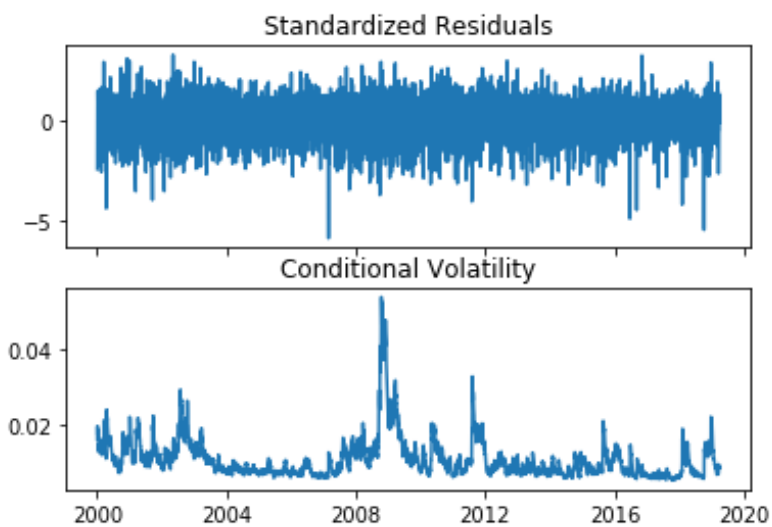


By using Python GARCH package called arch, we can fit SPX Index return rate into a GARCH model. And following is the fit result.

```
<bound method ARCHModelResult.summary of
ults
=====
Dep. Variable:          Adj Close    R-squared:              -0.001
Mean Model:            Constant Mean  Adj. R-squared:         -0.001
Vol Model:             GARCH         Log-Likelihood:         15664.7
Distribution:          Normal        AIC:                   -31321.5
Method:               Maximum Likelihood  BIC:                   -31295.5
                                           No. Observations:      4841
Date:                 Sun, Apr 28 2019  Df Residuals:          4837
Time:                 17:08:27          Df Model:              4
                               Mean Model
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu          5.7069e-04  1.052e-04      5.424  5.829e-08  [3.645e-04,7.769e-04]
                               Volatility Model
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega       2.8917e-06  2.059e-11  1.404e+05      0.000  [2.892e-06,2.892e-06]
alpha[1]    0.1000  8.637e-03     11.578  5.332e-31  [8.307e-02, 0.117]
beta[1]     0.8800  5.543e-03     158.766  0.000    [ 0.869, 0.891]
=====

Covariance estimator: robust

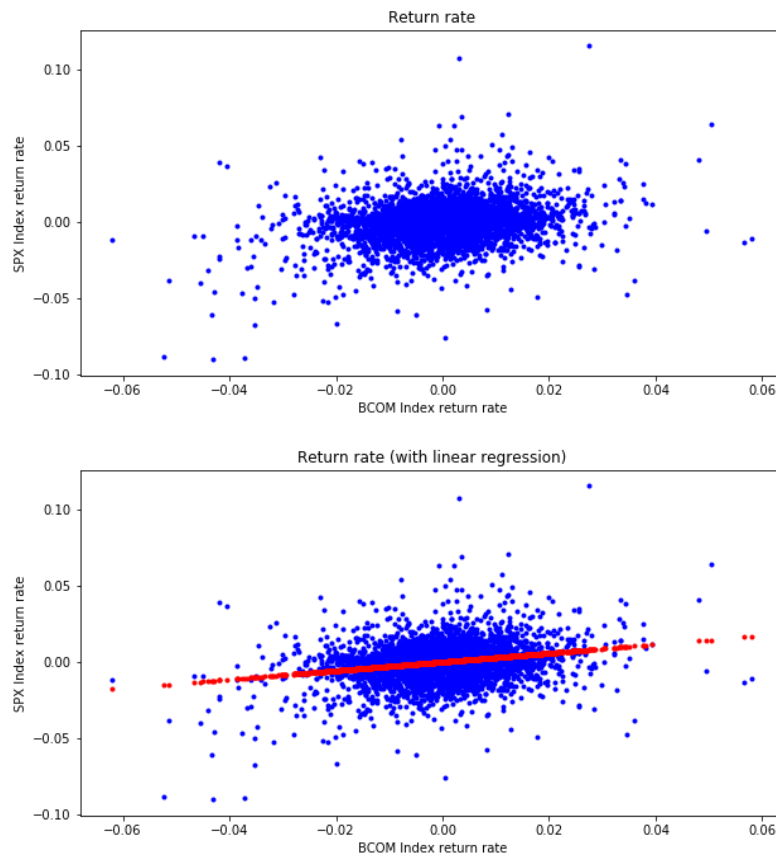
WARNING: The optimizer did not indicate successful convergence. The message was
Positive directional derivative for linesearch. See convergence_flag.
ARCHModelResult, id: 0x1c2b6f98d0>
```



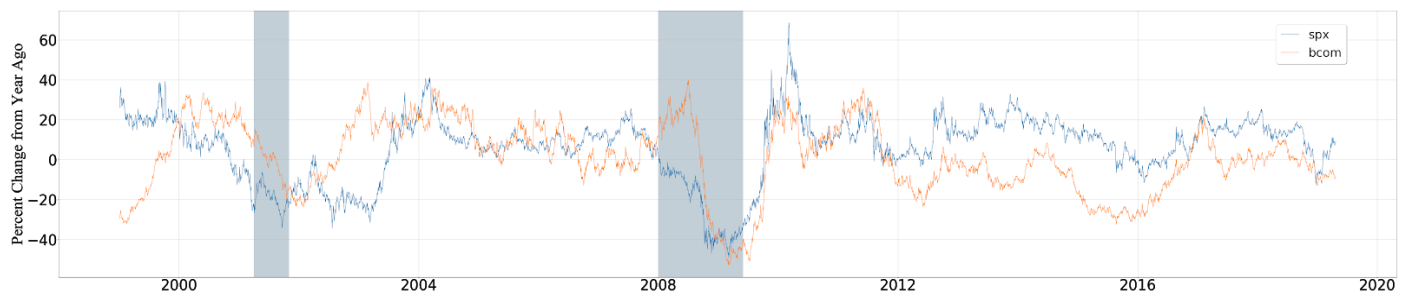
Before using DCC-GARCH model to find the dynamic correlation, let me introduce some commonly used ways to find correlation between two assets, and later compare their results with our DCC-GARCH dynamic correlation result.

The most common way is to do a linear regression on the return rates. The red points on the following graph are original points projected on the linear regression line. This method can also be easily achieved on Bloomberg Terminal, and lets one obtain the slope of the linear regression line.

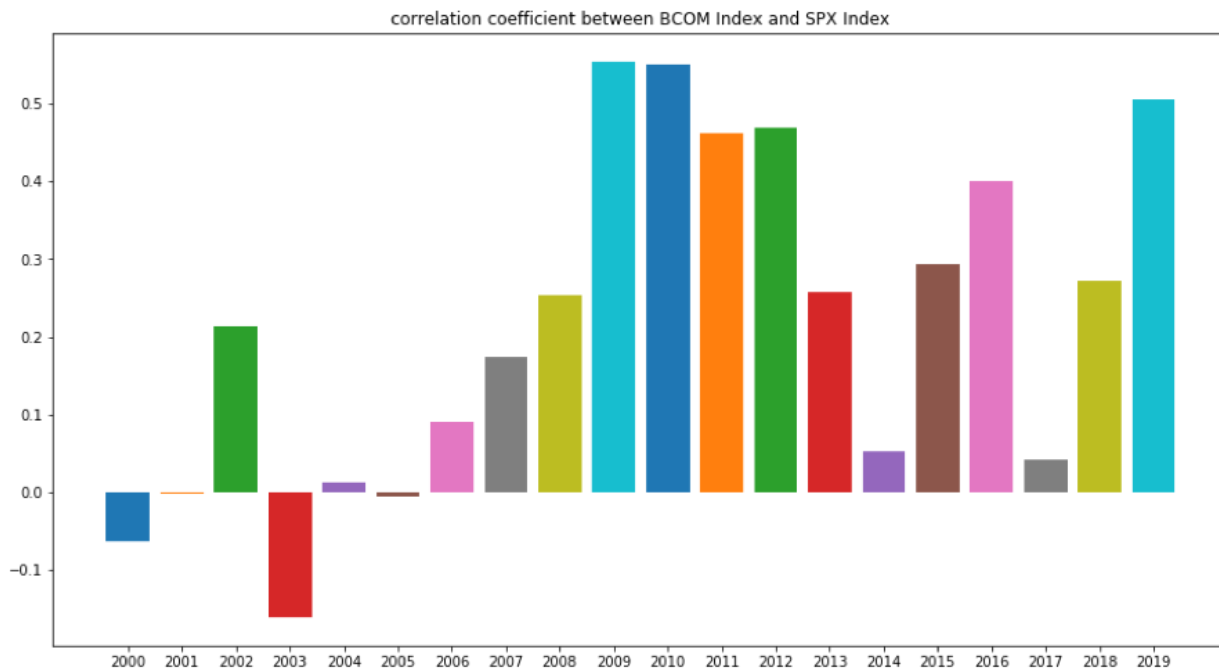
However, despite its convenience, it can only give us a single number to describe the correlation between two assets in a specific period. In the other word, we cannot find how the correlation changes through this period.



In order to find out the change of correlation during some time period, let us introduce another easy way to visually find the change in correlation. Technically, this way cannot give us an exact numerical result, but will help us to visualize the change in correlation between assets. We call it the percent change from year ago graph.



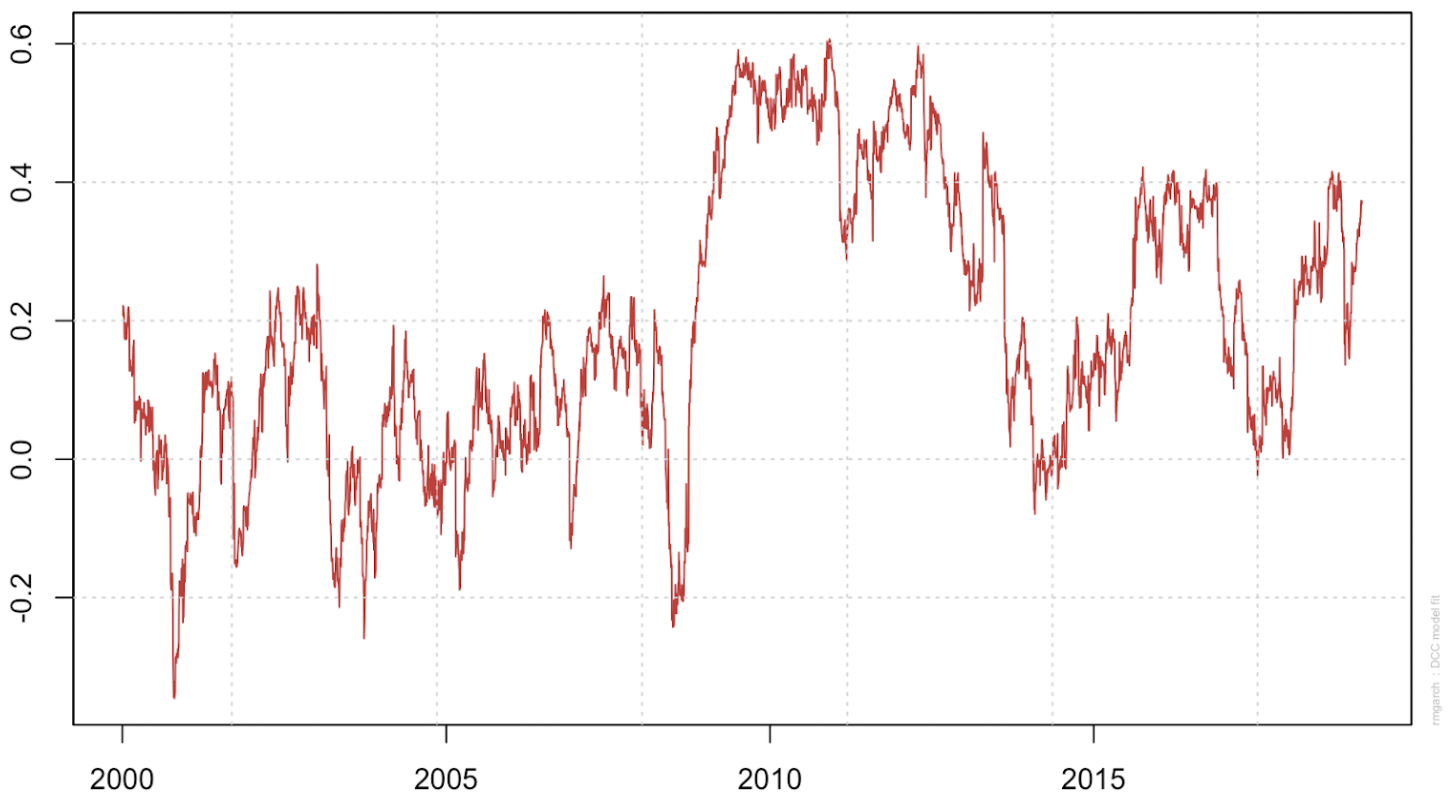
But still, as you can see, this way still cannot satisfy all of our requirements. We need to find a way that can tell us both the change of correlation and the numerical measurement of it. Therefore, we find graphing the yearly correlation coefficient in a bar graph is a good way to solve this problem.



Now let's compare this yearly correlation coefficient bar graph with the graph of return rates' linear regression. We find after year 2006, there are many correlation coefficient that are greater than 0.2, however, on the linear regression graph the regression line is much flatter than our expectation. The reason for that is due to the fact that the data before year 2006 has correlation coefficient close to zero or even negative, as demonstrated earlier in the paper. These parts of the data will influence the overall correlation for long time periods. It is also possible to use a rolling window to get the rolling monthly correlation line chart as we showed before, which uses the same idea with the yearly correlation coefficient bar graph.

Finally, let us introduce DCC-GARCH model, the most advanced way to find the dynamic correlation between two assets. By using the R package rmgarch and rugarch, we can find the DCC-GARCH fit result between SPX Index and BCOM Index in the most recent two decades. The result of this dynamic correlation is similar with the rolling correlation line chart graph. However, the difference between these two is that the dynamic correlation less influenced by historical data containing periods of high volatility.

**DCC Conditional Correlation
spx_close-bcom_close**



```

*-----*
*      DCC GARCH Fit      *
*-----*

```

```

Distribution      : mvnrm
Model            : DCC(1,1)
No. Parameters    : 11
[VAR GARCH DCC UncQ] : [0+8+2+1]
No. Series        : 2
No. Obs.          : 4801
Log-Likelihood    : 31327.21
Av.Log-Likelihood : 6.53

```

Optimal Parameters

```

-----
                Estimate Std. Error  t value Pr(>|t|)
[bcom_close].mu   -0.000035   0.000124  -0.28333 0.776921
[bcom_close].omega 0.000000   0.000000   2.04510 0.040845
[bcom_close].alpha1 0.038938   0.002273  17.12749 0.000000
[bcom_close].beta1 0.956507   0.001613 592.84045 0.000000
[spx_close].mu     0.000552   0.000116   4.73750 0.000002
[spx_close].omega 0.000002   0.000001   1.60162 0.109240
[spx_close].alpha1 0.107606   0.016534   6.50798 0.000000
[spx_close].beta1 0.877259   0.017101  51.29839 0.000000
[Joint]dcca1       0.018662   0.003943   4.73276 0.000002
[Joint]dccb1       0.978193   0.005044 193.94392 0.000000

```

Information Criteria

```

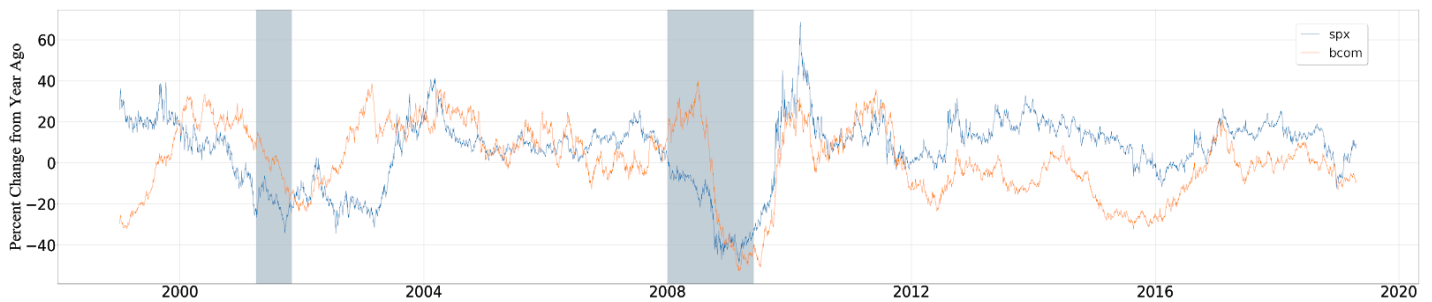
-----
Akaike      -13.046
Bayes       -13.031
Shibata     -13.046
Hannan-Quinn -13.040

```

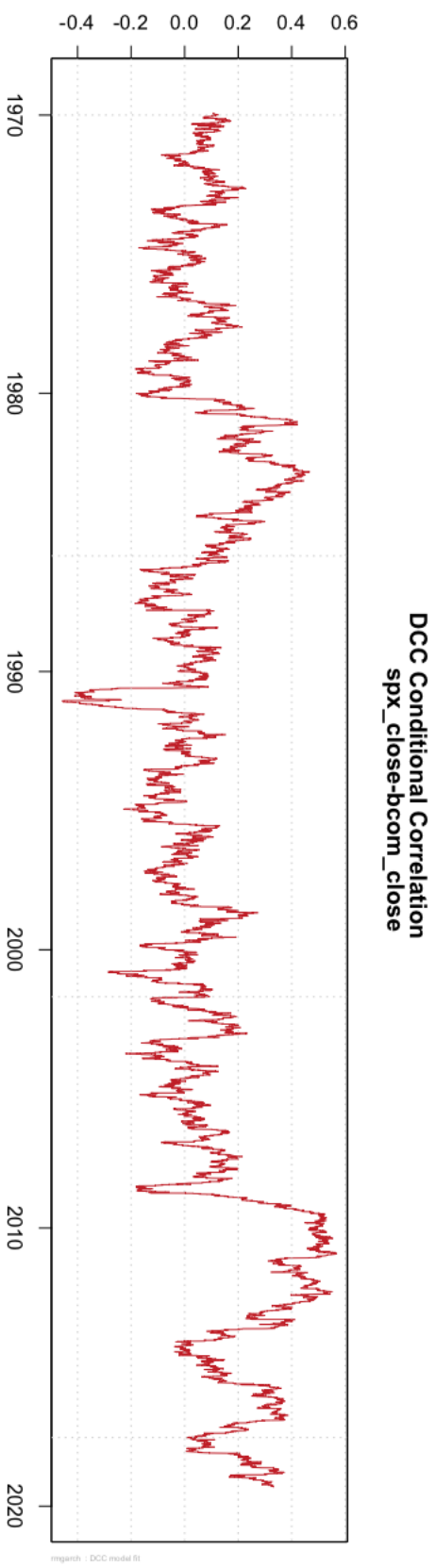
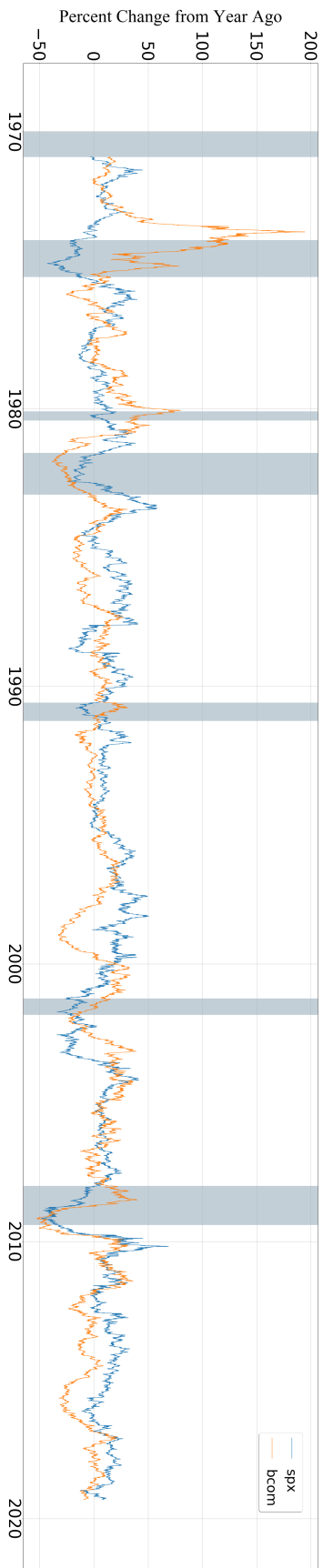
Elapsed time : 2.836058

Conclusion:

Using the result of dynamic correlation graph, we can find a positive correlation between equity market and commodity market. Now, let's combine the dynamic correlation graph with business cycle. In the following graph, dash areas are economy recession periods. We find when there is a large recession, the correlation between stock price and commodity price will suddenly increase, especially during the global finance crisis of 2007 to 2008. Therefore we can conclude that the correlation between equity market and commodity market will increase when the economic environment is unstable.



In next page, we attached DCC conditional correlation graph and business cycle graph both from year 1970 until year 2019. It's helpful to combine these two graph together to analysis how economy recession influence the correlation between commodity market and equity market.



References

V. L. Martin, A. S. Hurn and D. Harris 2011. Econometric Modelling with Time Series Specification, Estimation and Testing (ARCH properties in Chapter 20 Nonlinearities in Variance)

James Chong* and Joëlle Miffre**. March 2009. Conditional Correlation and Volatility in Commodity Futures and Traditional Asset Markets (DCC-GARCH methodology part)

The Following Content is not included in our report

(putting here just for helping people's further study about analysis which asset is in leading position)

Although we can conclude how the correlation change between Commodity market and Equity market overtime, we still cannot clearly identify which one is in the leading position when there is a financial turbulence. In order to observe this we show some pictures of return rates in high correlation months. Notice, in order to get the effective and systematized analysis, the months with the highest correlation coefficient during each year were chosen. We start with year 2008. The reason we begin with year 2008, is because before year 2008 all month's correlation coefficient is smaller than 0.5.

Just by looking at the graphs of returns, it seems as though the S&P 500 returns slightly lead the BCOM returns, however it is hard to tell.

