Time Series Modelling and Forecasting

Group 1802 Project Part 3

Yunxiu ZHOU, 1155046976

For
Prof. Chow Ying Foon, Ph.D

FINA 4130
Term 2, AY 17/18

Contents

I

Financial Modeling and Forecasting

CHAPTER 1

Overview

This part of the project develops an empirical thesis about asset price processes. We examine various time series models on multiple asset processes including a stock return as in index, a mutual fund return series comprised of a diversified portfolio, and a mutual fund return made up from Consumer Staples Portfolio.

1.1 Data Collection and Organization

1.1.1 *Asset Process*

Three asset processes are included in this thesis, including FDFAX, POAGX and S&P 500, each of which represents a non-diversified portfolio mutual fund, diversified portfolio mutual fund and a market portfolio stock index.

A summary table is given below in figure 1.1, highlighting data information and their features specifically. We choose these asset processes in the hope of finding different time series patterns in the data and thus assess effectiveness of different models in modelling the asset returns. Profiles of the data are referred from Bloomberg, WRDS CRSP and Money.Us News respectively.

Fund Ticker	Fund Name	Inception Date	Features	Profile
FDFAX	Fidelity Select Consumer Staples Port	1985-07-29	• open-end fund incorporated in the USA; • Objective: capital appreciation • non-diversified	• ≥ 80% of assets in manufacture, sale, or distribution of food and beverage products, agricultural products, and products related to the development of new food technologies
POAGX	Odyssey Aggressive Growth Fund	2004-11-02	• Objective: long-term capital appreciation • 1st in Mid Cap Growth • aggressive • Ranked among the top U.S. Diversified Stock Funds	The fund invests primarily in the common stocks of U.S. companies, emphasizing those companies with prospects for rapid earnings growth. It may invest in stocks across all market sectors and market capitalizations and has historically invested significant portions of its assets in mid- and small-capitalization companies.
S&P	Standard and Poor's 500 Index	1950-01-03 (Investigation Date)	• Market index	Standard and Poor's 500 Index is a capitalization-weighted index of 500 stocks. The index is designed to measure performance of the broad domestic economy through changes in the aggregate market value of 500 stocks representing all major industries. The index was developed with a base level of 10 for the 1941-43 base period.

Figure 1.1: Summary Table for Data Information

1.1.2 *Graphical Visualization and Investigation window*

The first glance of the daily price processes across their entire sampling period reveals that all of the asset closing prices appear to have upward trend, while S&P 500 presents to be the most volatile among the three, with more up and downs. Yet, as their historical length is unequal, such comparison provides little value. Therefore, we select a unified the training period to be from 2005-01-03 to 2017-01-31 for all the three assets. They are visualized in figure 1.3, we see that S&P 500 is of much value in maginitude than the mutual fund prices, Therefore is it hard to compare the asset price processes directly. This point together with the stationary property discussed in the following section motivated us to focus on their log returns, a common subject of interst in financial analyses.

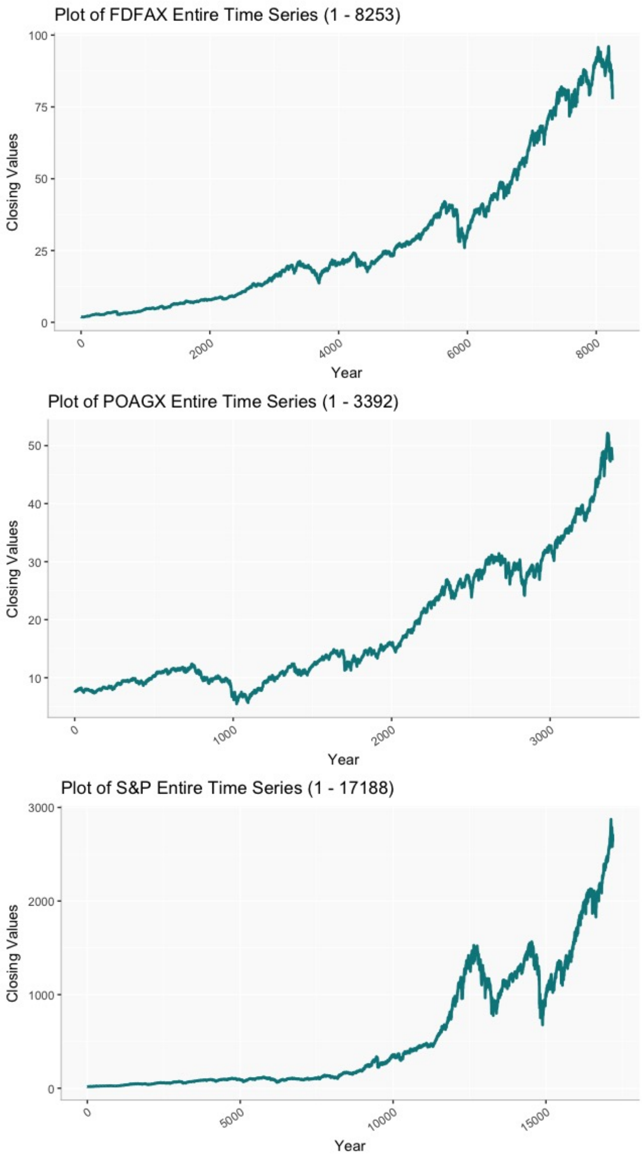


Figure 1.2: Entire Path:: Time Series Plot

Figure 1.3: Aggregated Price Evolution

CHAPTER 2

Statistical Properties of price and return series

2.1 STATIONARITY

2.1.1 *Mathemitical Notions*

Non-stationary data cannot be modeled or forecasted. Results based on non-stationarity can be spurious, and lead to problems such as false serial correlation in stock prices. (Simaan, 2017)

A time series model for $\{X_t\}_t$ is said to be stationary if all its statistics remain unchanged after time shifts, i.e. if they are the same as the statistics of $X_{t_0} + t_t$ for all possible choices of t_0. This is the so-called strong stationarity, which requires strictly that the statistical distribution remain unchanged at any time points.

There is also a weaker notion of stationarity, which requires onlt the first two moments to be stationary while put relaxations on the higher moments. In othe words, a time series model is said to be weakly stationary if its mean function is constant, and its auto-covariance function is a function of the difference of its arguments.

In practical, strong stationarity is rarely observed and hard to prove from both ends, whereas the weak-stationarity is much more useful in practise.

2.1.2 *Checks for Stationarity*

There are many methods to check stationarity for a time series, by conducting investigation on time series including direct observations S_T, series of residuals remaining from the direct observations , or other terms.(Brownlee, 2018)

We investigate the stationarity of our three asset processes in following subsections.

2.1.2.1 Graphical Visualization

A quick review of the time series plot of the data provides a visual check for obvious trends and/or seasonality. From figure 1.2, we have already seen that obvious trend is presented in all of the asset prices. The same is not true after taking the log difference. Figure 2.15 presents the log return of the three assets, no obvious trend is observed, suggesting that their first moment may be constant. But, the variance clustering is very significant even at the first glance at the series: volatility is clearly clustered arounf time of financial crisis around 2008-2009 and 2011-2012.

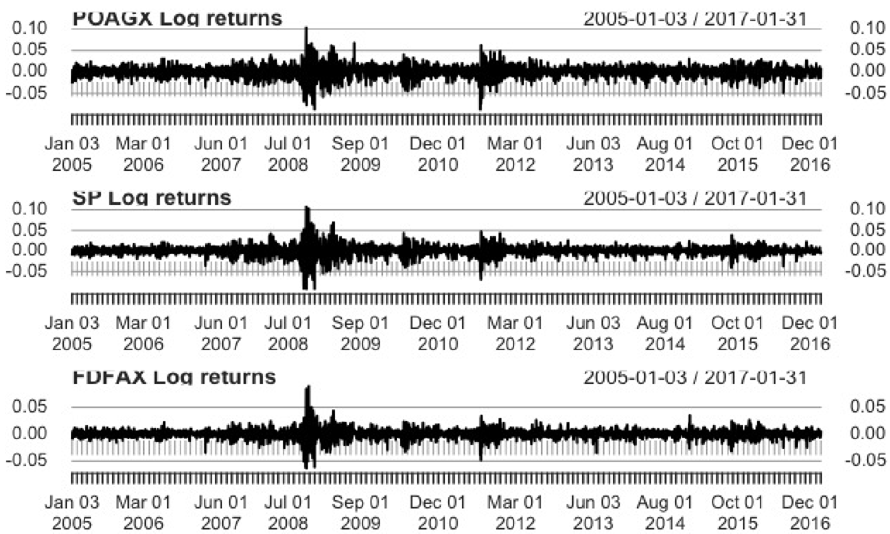


Figure 2.1: Aggregated Price Evolution

2.1.2.2 Summary Statistics

The second tool we can use is the summary statistics. Time series are stationary if they do not have trend or seasonal effects. In other words, common summary statistics calculated on the time series are expected to stay consistent over time, such as mean and variance of the observations.

We subset the daily closing price of the asset processes in to sub-samples, one year-long per sample. The descriptive statistics including mean and standard deviation is summarized in the following table 2.2:

	year	mean	sd
1	2005	8.125168	0.4283078
2	2006	9.702427	0.5782656
3	2007	11.347867	0.3778345
4	2008	8.951133	1.3245619
5	2009	8.448094	1.4118173
6	2010	11.568817	0.7402438
7	2011	13.382051	0.8366775
8	2012	14.846768	0.7084887
9	2013	20.369096	2.4415163
10	2014	26.070937	1.2620590
11	2015	29.484594	1.0978291
12	2016	29.890292	2.4399954
13	2017	34.085683	0.2819431

	year	mean	sd
1	2005	0.000370	0.009217
2	2006	0.000778	0.009426
3	2007	-0.000007	0.011072
4	2008	-0.001639	0.024019
5	2009	0.001620	0.018231
6	2010	0.000775	0.012952
7	2011	-0.000019	0.018429
8	2012	0.000770	0.010742
9	2013	0.001736	0.009015
10	2014	0.000606	0.010941
11	2015	0.000182	0.011414
12	2016	0.000439	0.012129
13	2017	0.001573	0.008687

Figure 2.2: Summary Statistics for POAGX: daily closing price v.s.log return

Clearly, the means of prices on the left panel of table 2.2 are significantly different from each other for successive samples, as well as the standard variation. To appreciate the time variability to a larger extent, we also visualized the evolution path of each of the sub sample (shown in figure 2.3) and an aggregated empirical distribution of sub samples in figure 2.4:

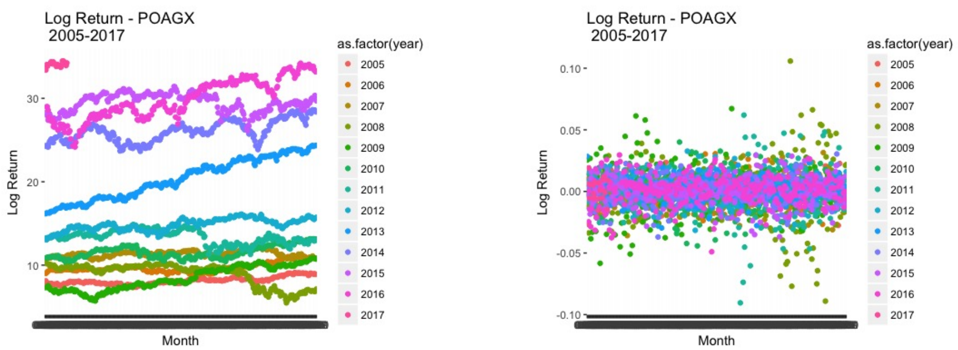


Figure 2.3: Sampling Path Comparison: daily closing price v.s.log return

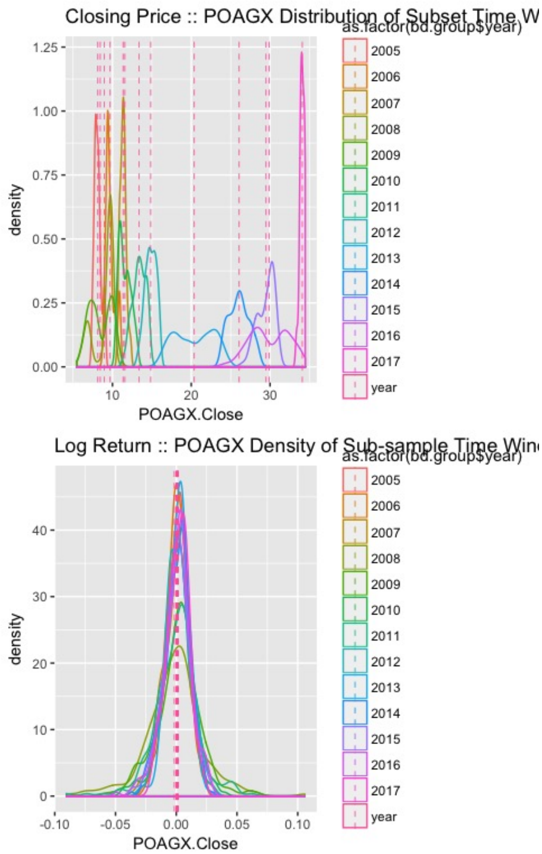


Figure 2.4: Density Comparison: daily closing price v.s.log return

We see that both the position and shape of the sampling paths are different for different subsamples of prices, suggesting different means, trends, volatility and other statistical properties may be fundamentally different for the sub-samples, which further signals that the underlying data structure of asset prices are non-stationary.

Meanwhile, we perform same set of analyses on log prices processes. For POAGX, we see from the summary table that the mean log returns are around for almost all sub-samples across during the entire sampling period, and standard deviation all ranging from 0.8% to 2.4%. The most volatile period to be around 2008, 2009 and 2011, while 2017 appears to be more stable. The time evolution of the return processes shows similar patterns across successsive years, centering around 0 and flutuate in diffeerent times in a year. (Notice that the axis in this plot is month per year, and we draw the sampling points for different years). More overlapping the paths are, more stationary the original process is in terms of different time window. We see that log returns significantly cluster together around zero, while sub samples of daily prices show distinctively different patterns among different years.

The same conclusion can be reached from the empirical density plots. These plots were made to be the empirical distribution of sample points (i.e. returns or prices) for each year, with the mean value of each sub-sample highlighted by a vertical dashed line in the plot. We see that for closing price, both the centering and shape parameters are significantly different for each year, and largely differ from a barbell shape. As a contrast, log returns generally behaves much similar across the sampling period: centering around mean, with a bell shape and pocess symmetry.

From these three comparison, we believe that while daily close prices do not behave stationarily across the sampling period, stationarity (at least weak-form stationarity) may exists in log returns for POAGX asset series. The same expereinces can be done on the other two asset series.

2.1.2.3 Statistical Tests – Unit Root tests

The third tool for stationarity assessment is statistical tests. There are several statistical tests available to check if the expectations of stationarity are met or have been violated.

Dickey and Fuller developed a test of the null hypothesis that $\alpha = 1$ against an alternative hypothesis that $\alpha < 1$ for the model $x_t = x_{t-1} + \mu_t$ in which μ_t is white noise. A more general test, which is known as the augmented Dickey-Fuller test (Said and Dickey, 1984), allows the differenced series μ_t to be any stationary process, rather than white noise, and approximates the stationary process with an AR model (see in chapter 3).The ADF tests results shown in figure 2.5 revealed that, the null hypotheses of a unit root **cannot** be rejected for all of the three asset prices, while the same set of null hypotheses are rejected for all of the return processes as shown in the following plot, suggesting they have roots smaller than the unit circle, thus are stationary.

```
> adf.test(PO_training)

        Augmented Dickey-Fuller Test

data:  PO_training
Dickey-Fuller = -1.295, Lag order = 14, p-value = 0.8768
alternative hypothesis: stationary

> adf.test(SP_training)

        Augmented Dickey-Fuller Test

data:  SP_training
Dickey-Fuller = -1.2573, Lag order = 14, p-value = 0.8928
alternative hypothesis: stationary

> adf.test(FD_training)

        Augmented Dickey-Fuller Test

data:  FD_training
Dickey-Fuller = -2.0349, Lag order = 14, p-value = 0.5635
alternative hypothesis: stationary
```

```
> adf.test(na.omit(po_logrt))

        Augmented Dickey-Fuller Test

data:  na.omit(po_logrt)
Dickey-Fuller = -15.268, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(po_logrt)) : p-value smaller than printed p-value
> adf.test(na.omit(sp_logrt))

        Augmented Dickey-Fuller Test

data:  na.omit(sp_logrt)
Dickey-Fuller = -15.387, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(sp_logrt)) : p-value smaller than printed p-value
> adf.test(na.omit(fd_logrt))

        Augmented Dickey-Fuller Test

data:  na.omit(fd_logrt)
Dickey-Fuller = -13.925, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(fd_logrt)) : p-value smaller than printed p-value
```

Figure 2.5: ADF Test results: daily closing price v.s. log return

An alternative to the ADF test is known as the Phillips- Perron test (Perron, 1988). PP tests show the same conclusion as the ADF tests, therefore their results are omitted for simplicity reasons.

2.1.3 *ACF and PACF*

Above conclusion can be easily and conveniently examined by ACF and PACF plots. A correlation of a variable with itself at different times is known as autocorrelation or serial correlation. If a time series model is second-order stationary, we can define an autocovariance function (ACVF), γ_k, as a function of its lag k:

$$\gamma_k = E[(x_t - \mu)(x_{t+k} - \mu)].$$

Correspondingly, the lag k auto-correlation function(ACF), γ_k, is defined as

$$\rho_k = \frac{\gamma_k}{\sigma^2}$$

.

We examine the auto-correlation function of both the asset prices and returns. The correlogram for three asset processes are shown in figure 2.6, figure 2.7 and figure 2.8.

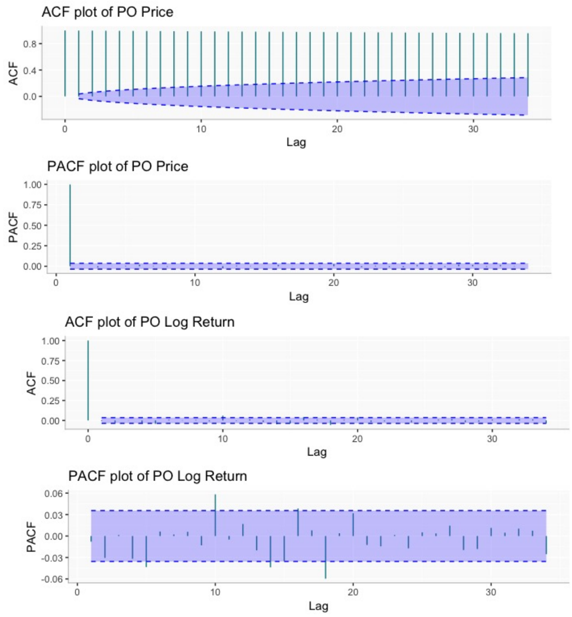


Figure 2.6: Correlogram for POAGX: log return v.s.daily closing price

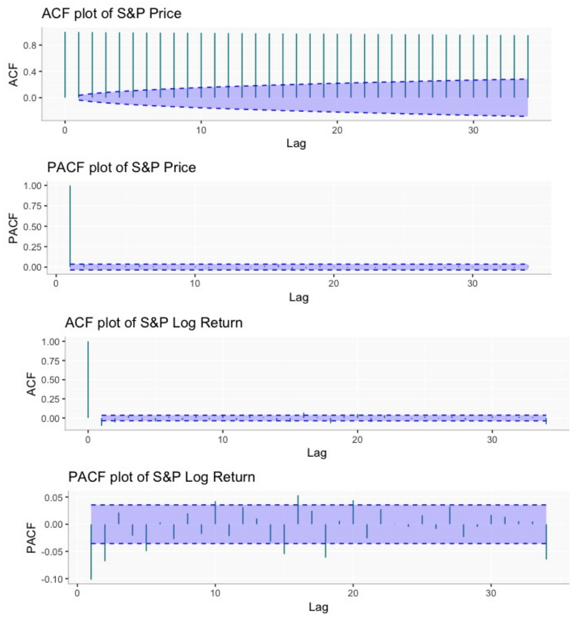


Figure 2.7: Correlogram for S&P: log return v.s.daily closing price

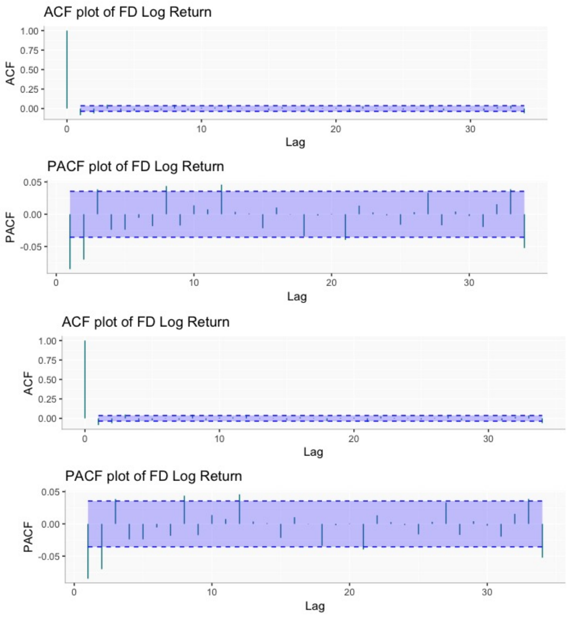


Figure 2.8: Correlogram for FDFAX: log return v.s.daily closing price

We see that for all of the three asset prices, their ACFs do not want to disappear after a specific time points. According to Metcalfe and Cowpertwait, 2009, a gradual decay from a high serial correlation is a notable feature of a random walk series. In addition, if a series follows a random walk, the differenced series will be white noise. These two arguments strengthen our believe that the close price are random works while the log returns are stationary, looking similar to a while noise process(with some special pattern in variance).

2.2 INDEPENDENCE

In this section, we introduce two statistical tests for Independence checking.

Before formally introduce two statistical tests for Independence checking, we first examine weather it is necessary to conduct the Independence test. In other words, we are to address the question that should one care about the Independence of time series as we have already have the powerful auto-correlation tests.

2.2.0.1 Importance of Independence check

Time series are regarded as independent if there is absolutely no relationship between the current variable and past variables. Yet, zero correlation does not imply independence. and one do want to check for the Independence of time series even little correlation has been tested to possess in the data sample.

Here are two examples illustrate the ideas above. Consider the following two simulated data structures. The first vector is constructed as i.i.d while the second is dependent. Both of the auto-corregrams shown in figure 2.9 reveals zero correlation among successive lags, suggesting no auto-correlation exsit for two vectors. but BDF tests show different results on these two: for series x, the p value is high, therefore we do not reject of the null hypothesis that the data is i.i.d at 1% significance level, while for series y, we see that p-value are not different from zero, therefore we reject the null hypothesis and suspect dependence among lags.

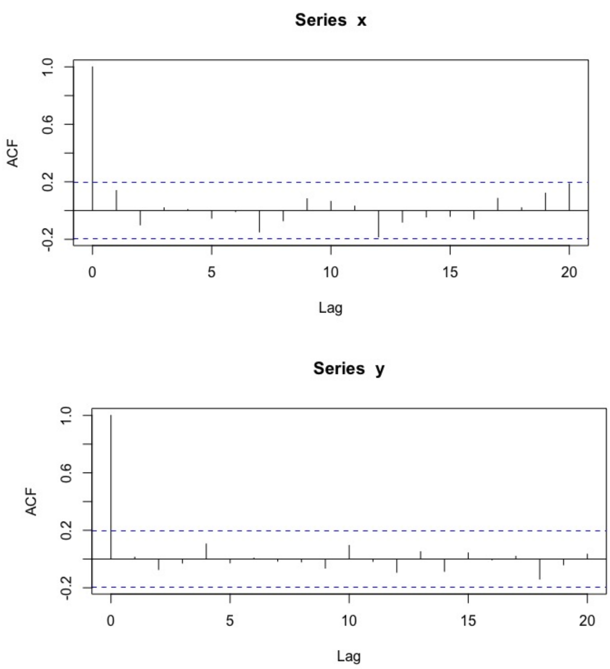


Figure 2.9: ACF ILLUSTRATION Example

```
> bds.test(x,m=3)

         BDS Test

data:  x

Embedding dimension =  2 3

Epsilon for close points =  0.5258 1.0516 1.5774 2.1033

Standard Normal =
      [ 0.5258 ] [ 1.0516 ] [ 1.5774 ] [ 2.1033 ]
[ 2 ]     1.4525     1.1702     2.0147     1.6471
[ 3 ]     0.1622     1.0470     1.6528     1.2591

p-value =
      [ 0.5258 ] [ 1.0516 ] [ 1.5774 ] [ 2.1033 ]
[ 2 ]     0.1464     0.2419     0.0439     0.0995
[ 3 ]     0.8712     0.2951     0.0984     0.2080
```

```
> bds.test(y)  # not independent

         BDS Test

data:  y

Embedding dimension =  2 3

Epsilon for close points =  0.1786 0.3573 0.5359 0.7145

Standard Normal =
      [ 0.1786 ] [ 0.3573 ] [ 0.5359 ] [ 0.7145 ]
[ 2 ]   547.6341   123.2873     2.0759    -9.3362
[ 3 ]   743.4724   115.3152    -2.1764    -7.9260

p-value =
      [ 0.1786 ] [ 0.3573 ] [ 0.5359 ] [ 0.7145 ]
[ 2 ]          0          0     0.0379          0
[ 3 ]          0          0     0.0295          0
```

Figure 2.10: BDS Test Illustration Example

Once the importance of Independence is recognized, we now introduce some of the most widely used tests for Independence check and their limitations.

2.2.1 *BDF Test*

The Brock- Dechert-Scheinkman test (a.k.a. BDS test) of serial independence (Brock, Dechert Scheinkman, 1996) checks whether a sequence of random variables are i.i.d.

Let $Y_t^n = (y_t, y_{t+1}, \dots, y_{t+n1})$. we have

$$
\begin{aligned}
Y_1^n &= (y_1, y_2, \dots, y_n), \\
Y_2^n &= (y_2, y_3, \dots, y_{n+1}), \\
&\dots \\
Y_{T-n}^n &= (y_{T-n}, y_{T-n+1}, \dots, y_T),
\end{aligned}
$$

Define a correlation integral with dimension n and distance ϵ as:

$$C_T(n,\epsilon) = \frac{2}{T-n+1} \sum_{t<m} I_\epsilon(y_t^n - y_s^n).$$

where $I_\epsilon(y_t^n, y_s^n) = 1$ if the maximal norm $y_t^n - y_s^n < \epsilon$ and 0 otherwise. If y_t are indeed i.i.d., then Y_t^n should exhibit no pattern in the n-dimensional space, so that $C(n,\epsilon) = C(1,\epsilon)^n$. The asymptotic null distribution of the BDS test is $N(0,1)$. For detailed description of BDS test, one may refer to (Kuan, 2018) for more information.

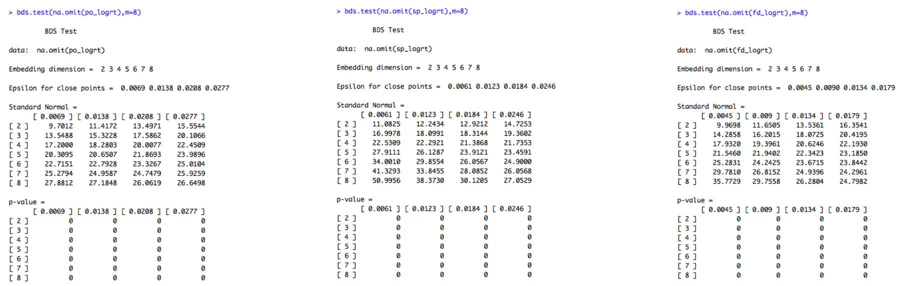
To summarize, the structure of BDS test is:

H_0: the test statistics follow $N(0,1)$, in other words, the underlying y_ts are i.i.d.

H_1: series y_ts are not i.i.d.

Trapletti's tseries package *R Programming* 2018 for R provides a BDS test function for financial time series. The command bds.test(x, m) computes the BDS statistic on a time series object x for m dimensions, and $\epsilon = 0.5\sigma, 1.0\sigma, 1.5\sigma$ and 2.0σ by default. The function outputs two m by 4 matrices. The first matrix, labelled as 'Standard Normal', contains the normalized BDS statistics for each embedding dimension and each epsilon value, whereas the second matrix, labelled 'p-value', contains the two sided p-values.

The BDS testing results for the three log return series reveals that all of them, although zero correlation with lags, are not independent with their history at different lags, as all the p values in the second matrix are around 0.



```
> bds.test(na.omit(po_logrt),m=8)

         BDS Test

data:  na.omit(po_logrt)

Embedding dimension =  2 3 4 5 6 7 8

Epsilon for close points =  0.0069 0.0138 0.0208 0.0277

Standard Normal =
      [ 0.0069 ] [ 0.0138 ] [ 0.0208 ] [ 0.0277 ]
[ 2 ]     9.7012    11.4172    13.4971    15.5544
[ 3 ]    13.5488    15.3228    17.5862    20.1066
[ 4 ]    17.2000    18.2803    20.0077    22.4509
[ 5 ]    20.3095    20.6507    21.8693    23.9896
[ 6 ]    22.7151    22.7928    23.3267    25.0104
[ 7 ]    25.2794    24.9587    24.7479    25.9259
[ 8 ]    27.8812    27.1848    26.0619    26.6498

p-value =
      [ 0.0069 ] [ 0.0138 ] [ 0.0208 ] [ 0.0277 ]
[ 2 ]          0          0          0          0
[ 3 ]          0          0          0          0
[ 4 ]          0          0          0          0
[ 5 ]          0          0          0          0
[ 6 ]          0          0          0          0
[ 7 ]          0          0          0          0
[ 8 ]          0          0          0          0
```

```
> bds.test(na.omit(sp_logrt),m=8)

         BDS Test

data:  na.omit(sp_logrt)

Embedding dimension =  2 3 4 5 6 7 8

Epsilon for close points =  0.0061 0.0123 0.0184 0.0246

Standard Normal =
      [ 0.0061 ] [ 0.0123 ] [ 0.0184 ] [ 0.0246 ]
[ 2 ]    11.0825    12.2434    12.9212    14.7253
[ 3 ]    16.9978    18.0991    18.3144    19.3602
[ 4 ]    22.5309    22.2921    21.3868    21.7353
[ 5 ]    27.9111    26.1287    23.9121    23.4591
[ 6 ]    34.0010    29.8554    26.0567    24.9000
[ 7 ]    41.3293    33.8455    28.0852    26.0568
[ 8 ]    50.9956    38.3730    30.1205    27.0529

p-value =
      [ 0.0061 ] [ 0.0123 ] [ 0.0184 ] [ 0.0246 ]
[ 2 ]          0          0          0          0
[ 3 ]          0          0          0          0
[ 4 ]          0          0          0          0
[ 5 ]          0          0          0          0
[ 6 ]          0          0          0          0
[ 7 ]          0          0          0          0
[ 8 ]          0          0          0          0
```

```
> bds.test(na.omit(fd_logrt),m=8)

         BDS Test

data:  na.omit(fd_logrt)

Embedding dimension =  2 3 4 5 6 7 8

Epsilon for close points =  0.0045 0.0090 0.0134 0.0179

Standard Normal =
      [ 0.0045 ] [ 0.009 ] [ 0.0134 ] [ 0.0179 ]
[ 2 ]     9.9698   11.6505    13.5361    16.3541
[ 3 ]    14.2858   16.2015    18.0725    20.4195
[ 4 ]    17.9320   19.3961    20.6246    22.1930
[ 5 ]    21.5460   21.9402    22.3423    23.1850
[ 6 ]    25.2831   24.2425    23.6715    23.8442
[ 7 ]    29.7810   26.8152    24.9396    24.2961
[ 8 ]    35.7729   29.7558    26.2804    24.7982

p-value =
      [ 0.0045 ] [ 0.009 ] [ 0.0134 ] [ 0.0179 ]
[ 2 ]          0         0          0          0
[ 3 ]          0         0          0          0
[ 4 ]          0         0          0          0
[ 5 ]          0         0          0          0
[ 6 ]          0         0          0          0
[ 7 ]          0         0          0          0
[ 8 ]          0         0          0          0
```

Figure 2.11: BDS test results for three log return series

While the BDS test does have an advantage that it is robust to random variables that do not possess high-order moments, it usually needs a large sample to ensure proper performance (Kuan, 2018). As a result, test statistics for dimensions larger than 5 are usually believed to be less robust. (Belaire-Franch and Contreras, 2002) Although we have choosen up to 8 dimenstions in this project, all of them show significant evidence against the null hypothesis. Moreover, it has been found that the BDS test has low power against various forms of nonlinearity.

2.2.2 *Variance Ratio Test*

Lo and MacKinlay (Lo and MacKinlay, 1988) variance ratio tests are based on the property that, if returns are i.i.d., the variance of the k-period return should be approximately k times of the variance of the one-period return. Mathematically, this variance ratio is given by

$$VR(k) = \frac{\frac{1}{Tk}\sum_{t=k} T(y_t + y_{t-1} + y_{t-2} + \cdots + y_{t-k+1} - k\hat{\mu})2}{\frac{1}{T}\sum_1 T(y_t - \hat{\mu})}2$$

where $\hat{\mu} = T{-}1\sum_{t=1}^{T} y_t$. Lo and MacKinlay showed that, if the returns are i.i.d., then the test statistic

$$M_1(k) = (VR(k) - 1)\Phi(k)^{-1/2}$$

follows the standard normal distribution asymptotically, under the null hypothesis that $VR(k) = 1$. A later modified version[1] of LoMac model proposed an alternative test statistic that is robust against the presence of conditional heteroscedasticity, given by

$$M_2(k) = (VR(k) - 1)[\sum_{j=1}^{k-1}[\frac{2(k-j}{k}\theta_j]^{-1/2}$$

.

In other words, the Lo Mac model states that if the sample statistics computed from the sample is not too far away from 0 in terms of standard error, then the series are believed to be i.i.d. with confidence.

We conduct the Lo Mac models on the three return series for various holding period k. While the 1% and 5% two-tail critical Z values are 2.576 and 1.960 respectively, we observe that the test statistics are all negative and significant, which helps us reach the same conclusion of dependence among return series as those from BDS tests.

	POAGX			S&P			FDFAX	
	M1	M2		M1	M2		M1	M2
k=2	-0.4712995	-0.2930923	k=2	-5.6270136	-2.8633696	k=2	-4.6956697	-2.3534105
k=5	-1.5618203	-0.9000867	k=5	-5.3735393	-2.4967356	k=5	-4.5962537	-2.1046475
k=10	-2.1781496	-1.2417899	k=10	-4.8549638	-2.2033788	k=10	-3.8321323	-1.7480214
k=30	-1.7252435	-1.0222173	k=30	-2.8803218	-1.3133448	k=30	-2.1505988	-1.0131322
k=250	-0.5022234	-0.3679080	k=250	-0.6421867	-0.3766211	k=250	-0.8155726	-0.5297342
k=500	-0.5477778	-0.4449060	k=500	-0.5932454	-0.4000879	k=500	-1.0437452	-0.7806342

Figure 2.12: Lo Mac Test for Independence

[1] Interested person may refer to the orginal paper for more details

2.3 NORMALITY CHECK

In addition to the stationary and independent property, we are also interested in the normality in our time series. There are usually two ways for one to check the normality for time series, graphical visualization of density plots and normality statistical tests. One can also use QQ-plot for the same purpose.

2.3.1 *Density Plots*

The density plots of the three log returns series during the entire training periods are shown in figure 2.13, with the red dotted line companioned as the theoretical normal density plot with mean as the sample mean and standard deviation as the sample estimated standard deviation. We see that, in general, the returns appear not to be identical to their standard normal counterparts: although with the same centering location at zero, their shapes are different and with fat tail performance.

2.3.1.1 Normality Statistical Tests

We can corroborate these empirical density distribution plots with one of many statistical test for the null hypothesis that a sample time series $R_1, \dots, R_n$ of returns come from a normally distributed population. One popular such tests is Shapiro-Wilk. If the p-value of obtaining the Shapiro-Wilk test computes is less than a given confidence level then the null hypothesis should be rejected (i.e. the data does not come from a normal distributed population).

The test results shown in figure 2.14 shows that the p-values for the test statistics for three returns are all around zero, we reject the null hypotheses that these returns are drawn from a normal distribution. Such conclusion is supported by another famous test, known as Kolmogorov-Smirnov normality test. The results leads us to the same conclusion that the log returns do not fit the normal distributions exactly.

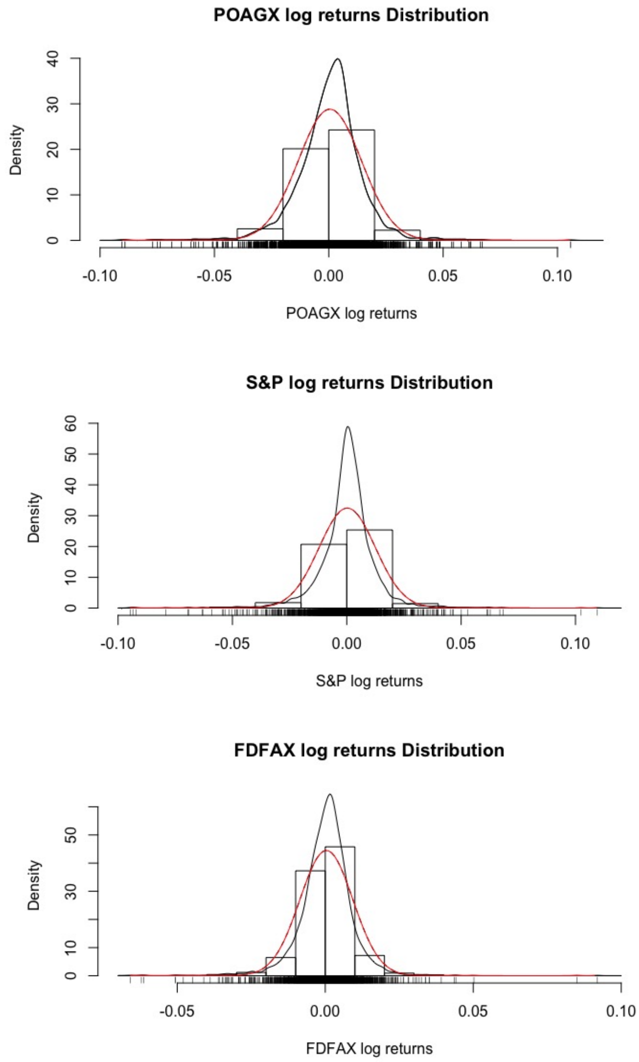


Figure 2.13: Density plots for three log returns

```
> shapiro.test(as.vector(r1))

        Shapiro-Wilk normality test

data:  as.vector(r1)
W = 0.94517, p-value < 2.2e-16

> shapiro.test(as.vector(r2))

        Shapiro-Wilk normality test

data:  as.vector(r2)
W = 0.87597, p-value < 2.2e-16

> shapiro.test(as.vector(r3))

        Shapiro-Wilk normality test

data:  as.vector(r3)
W = 0.90403, p-value < 2.2e-16
```

Figure 2.14: Shapiro-Wilk's test for normality Results for three log return series

POAGX

```
ksnormTest(as.vector(r1))

itle:
One-sample Kolmogorov-Smirnov test

est Results:
 STATISTIC:
   D: 0.4776
 P VALUE:
   Alternative Two-Sided: < 2.2e-16
   Alternative      Less: < 2.2e-16
   Alternative   Greater: < 2.2e-16
```

S&P

```
> ksnormTest(as.vector(r2))

Title:
 One-sample Kolmogorov-Smirnov test

Test Results:
  STATISTIC:
    D: 0.4789
  P VALUE:
    Alternative Two-Sided: < 2.2e-16
    Alternative      Less: < 2.2e-16
    Alternative   Greater: < 2.2e-16
```

FDFAX

```
> ksnormTest(as.vector(r3))

Title:
 One-sample Kolmogorov-Smirnov test

Test Results:
  STATISTIC:
    D: 0.484
  P VALUE:
    Alternative Two-Sided: < 2.2e-16
    Alternative      Less: < 2.2e-16
    Alternative   Greater: < 2.2e-16
```

Figure 2.15: Kolmogorov-Smirnov normality test Results for three log return series

Therefore, both of the tools in normality test suggest the returns are not normally distributed, it is also justified in the QQ plots shown in figure 2.16. The deviation away from the golden line signals the disform to normality.

One reason that the distribution would not be exactly normal is because of volatility clustering: some periods have higher volatility than others, we will discover the hereteroscadesity in the following section in more depth.

To summarize, we have establish the following observations of our three asset processes:

1. Asset prices are not stationary with obvious trends across the training period;
2. While the first moment of returns are stationary, they do behave with variance clustering phenomenon;
3. returns are nor exactly normal.

Therefore, we will have to use appropriate models to capture these patterns in the following chapter.

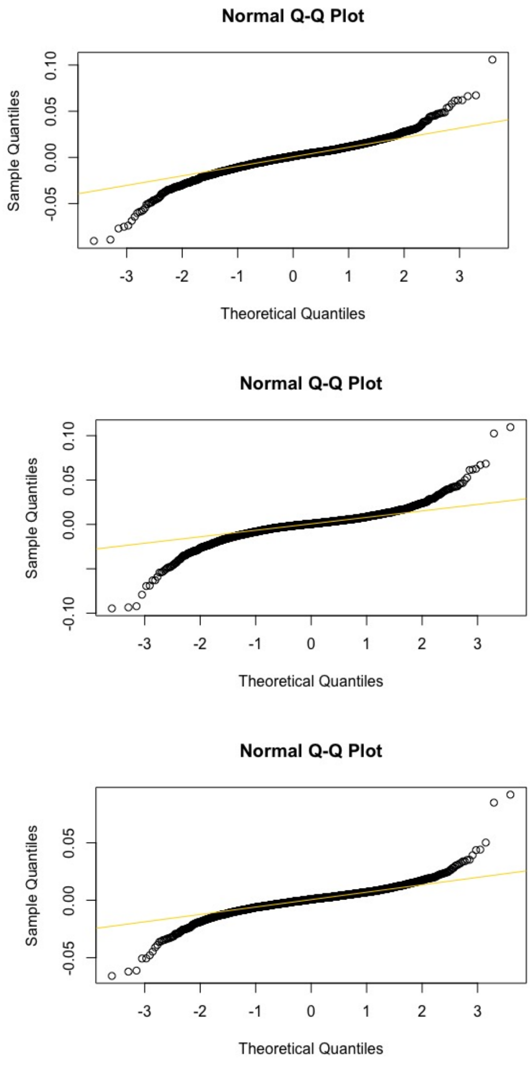


Figure 2.16: Density plots for three log returns

CHAPTER 3

Time Series Modelling

The first observation at the end of last chapter suggests an ARIMA model to capture the correlation among lags (i.e. trends feature in returns), while the remaining two observations suggest a model to capture the serial correlation in variance.

3.1 AR,MA, ARMA AND ARIMA MODELS

3.1.1 *ARIMA*

A time series $\{X_t\}_t$ is said to be an ARIMA process if, when differentiated finitely many times, it becomes an ARMA time series. More precisely, one says that $\{X_t\}_t$ is an ARIMA(p, d, q) if its becomes an ARMA(p, q) after d differences. Mathematically, ARIMA model states that:

$$X \sim ARIMA(p,d,q) \Leftrightarrow \nabla^d X \sim ARMA(p,q)$$

where $\nabla = 1 - B$ is the first difference operator.

3.1.2 *ARMA*

A time series $\{X_t\}_t$ is said to be an auto-regressivemoving average time series of order p and q (i.e. $X \sim ARMA(p,q)$) if there exists a white noise $\{W_t\}_t$ such that:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$

for some real numbers $\phi_1, \ldots, \phi_p$, and $\theta_1, \ldots, \theta_q$.

3.1.3 *AR*

A mean-zero time series $\{X_t\}_t$ is said to be auto-regressive of order p (with respect to a white noise $\{W_t\}_t$) if:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} \cdots + \phi_p X_{tp} + W_t$$

for some set of real numbers $\phi_1, \ldots, \phi_p$. More generally, we say that $\{X_t\}_t$ is auto-regressive of order p if there exists a number μ_X such that the $\{X_t - \mu_t\}_t$ is auto-regressive of order p in the sense given above.

3.1.4 *MA*

A time series $\{X_t\}_t$ is said to be a moving average time series of order q (with respect to a white noise$\{W_t\}_t$) if:

$$X_t = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$

for some real numbers $\theta_1, \ldots, \theta_q$. In such a case we use the notation $X \sim MA(q)$.

From the very definition of the MA process, one could figure out that the random variables X_{t+s} and X_t are independent because of the disjoint white noise term. Graphically, an MA process is often characterized by the vanishing after lags greater than q in ACF plots.

3.2 FINDING THE ORDER OF P AND Q

The determination of the order of the model, (p and q to be more precise) has always been of interest. In some sense, one can interpret AR as a special category of linear regression models, therefore the parsimonious balance decision criteria for model comparison also suits for the model choice in ARIMA.

Criteria such as AIC and BIC assess the trade of between the number of parameters in the model(model accuracy) against efficiency(model simplicity and computational power). Yet, for time series particularly, ACF and PACF plots provide a good suggestion about the order of parameters to be used in the model, and are considered as more powerful tools in some cases. (Carmona, 2014)

3.3 ARIMA FOR RETURN SERIES

In this section, we are going to fit in the ARIMA model for all of the three processes. ACF and PACF plots provide a good suggestion about the order of parameters to be used in the model, we can easily find the model by auto.arima functions in R. The results with parameters and information criteria computes are shown in figure 3.1.

```
> arima.r1
Series: r1
ARIMA(2,0,0) with non-zero mean

Coefficients:
          ar1      ar2   mean
      -0.0081  -0.0306  5e-04
s.e.   0.0181   0.0181  2e-04

sigma^2 estimated as 0.0001916:  log likelihood=8698.93
AIC=-17389.86   AICc=-17389.85   BIC=-17365.78
> arima.r2
Series: r2
ARIMA(0,0,2) with zero mean

Coefficients:
          ma1      ma2
      -0.1061  -0.0550
s.e.   0.0182   0.0188

sigma^2 estimated as 0.000149:  log likelihood=9081.07
AIC=-18156.14   AICc=-18156.13   BIC=-18138.08
> arima.r3
Series: r3
ARIMA(2,0,1) with non-zero mean

Coefficients:
          ar1      ar2     ma1   mean
      -0.4820  -0.1071  0.3934  4e-04
s.e.   0.1399   0.0194  0.1400  1e-04

sigma^2 estimated as 7.933e-05:  log likelihood=10040.05
AIC=-20070.09   AICc=-20070.07   BIC=-20039.99
```

Figure 3.1: Arima models for three training returns

The return of POAGX is AR(2), and MA(2) for S&P and ARMA(2,1) for FDFAX respectively. We can compare the order selection criteria shown in figure 3.2 with the instructions from ACF and PACF plots figure 2.6, figure 2.7 and figure 2.8 . Yet, the ACF plots and PACF plots are far from simulated examples that it is not easy to read the order from the cut offs or vanishings directly.

Conditional Mean Model	ACF	PACF
AR(p)	Tails off gradually	Cuts off after p lags
MA(q)	Cuts off after q lags	Tails off gradually
ARMA(p,q)	Tails off gradually	Tails off gradually

Figure 3.2: ACF and PACF signal for order

After fitting the model, we also compare and contrast the ACF and PACF of the original series with the residual series.

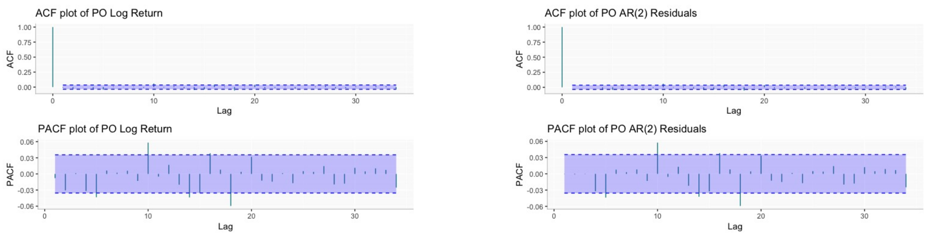


Figure 3.3: Correlogram for POAGX: log return v.s.Residuals

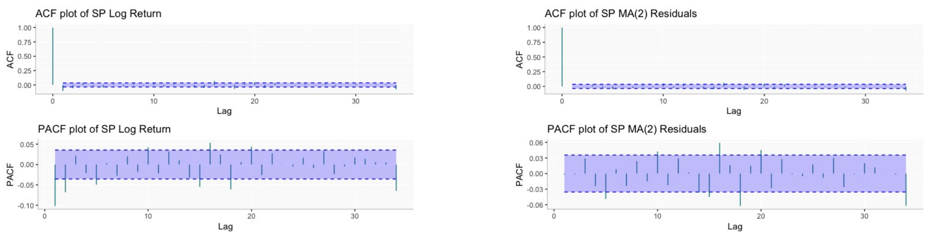


Figure 3.4: Correlogram for S&P: log return v.s.Residuals

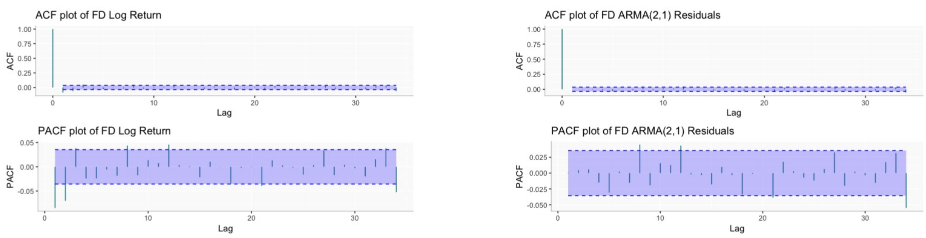


Figure 3.5: Correlogram for FDFAX: log return v.s.Residuals

A look at the above plots for residuals seem to indicate that the auto-correlation function of these raw residuals is not much different from the auto-correlation of a white noise, suggesting that our ARIMA models perform good to capture the auto-correlation among lags to reveal the trend pattern. Yet, we do know from the serial plots before that conditional heteroskedasticity exist, which we also need to address.

3.4 GARCH MODEL FOR VOLATILITIES

3.4.1 *Motivation*

Time series plot of the residuals of the ARIMA model as fitted to the log returns computed from the asset returns (left) and their Q-Q plot against the Gaussian distribution (right) are shown belown which motivates us to incorporate the GARCH model.

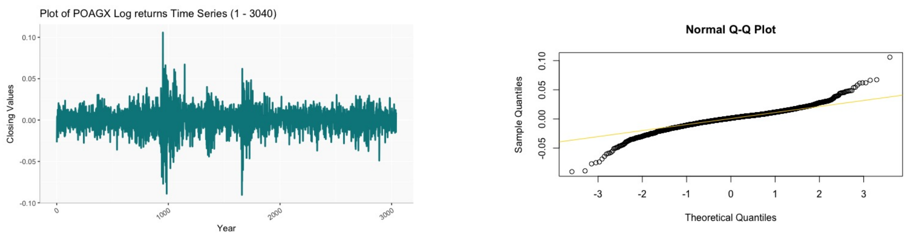


Figure 3.6: Time series plot of the residuals and their Q-Q plot (right) from POAGX

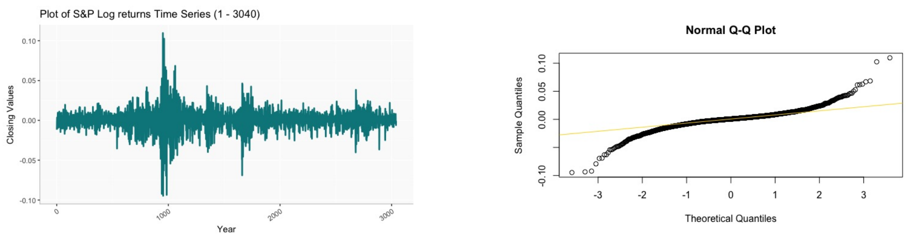


Figure 3.7: Time series plot of the residuals and their Q-Q plot (right) from S&P

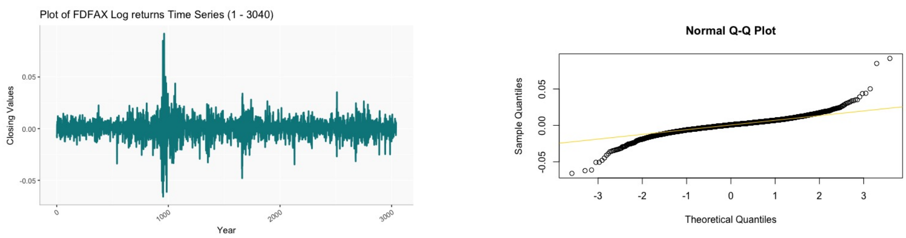


Figure 3.8: Time series plot of the residuals and their Q-Q plot (right) from FDFAX

From the left panel, it is clear that too many measurements end up several standard deviation away from the mean, especially around time 2008, during the financial crisis period. The marginal distribution of these residuals is presumably not normal. This is confirmed by the normal Q-Q plot of these residuals reproduced in the right pane of these figures, which shows that the distribution of the residuals has heavy tails for all of the three asset processes. So even if the ARIMA model was able to capture the serial correlation contained in the log returns, we had no confidence about the independence in the residuals.

If the ARIMA models were able to capture the patterns entirely from the return process, then the residuals in these models should behave like white noise. Since squares of random variables are independent whenever the original random variables are independent, then the squares of the residuals should also behave randomly. The plot of the auto-correlation function of the squares of the time series of residuals which we reproduce in Fig. 3.14 confirm the dependincies remaining in the residuals. Astonishingly, the autocorrelgram of the squared error terms are highly correlated with each other.

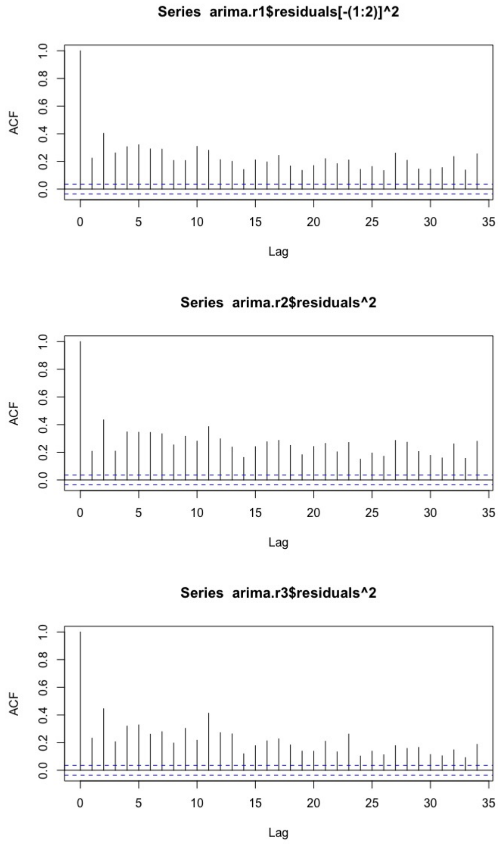


Figure 3.9: ACF plot of the residuals square

3.4.2 *Models: ARCH and GARCH*

An autoregressive model for the variance process, ($ARCH$ model) accounts for the conditional changes in the variance. Mathematically, a series $\{\epsilon_t\}_t$ is defined to be first-order autoregressive conditional heteroskedastic, denoted as $ARCH(1)$, if

$$\epsilon_t = w_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}$$

where w_t is white noise with zero mean and unit variance. Taking a square of the above equation, one get the calculated variance:

$$\begin{aligned} Var(\epsilon_t) &= E[\epsilon_t^2] \\ &= \alpha_0 + \alpha_1 Var(\epsilon_{t-1}) \end{aligned}$$

If we compare $ARCH(1)$ with the AR(1) process $x_t = \alpha_0 + \alpha_1 x_{t-1} + w_t$, we see that the variance of an ARCH(1) process behaves just like an AR(1) model. Hence, in model fitting, a decay in the autocorrelations of the squared residuals indicates a good signal to use an ARCH model.

The $ARCH(1)$ model, or more generally, $ARCH(p)$ model can be extended to the generalised $ARCH$ model, denoted as $GARCH(q, p)$, which has the $ARCH(p)$ model as the special case $ARCH(0, \text{p})$. A series ϵ_t is $GARCH(q, p)$ if

$$\epsilon_t = w_t \sqrt{h_t}$$

where

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

Reflecting on the ideas behind the model, we see that ARCH and GARCH models are to solve the problem of un-random variance. In order to forecast the return volatily, which is usually of the interest of quanlitative financial practitioners, one of the best estimates in prior to these models may be the naive rolling window sample variance to forecast the upcoming variance according to CLT. Yet, such estimates allocate weight equally among the historical occurances. Intuitively, one may wish to put more weight in the more recent occurrence compared with the far previous history. ARCH and GARCH are the models to realize that, providing quantitatively-sound weight schemes for historical variance. Looking closely at the GARCH model, it says that the forecast for the next instantaneous variance would be a weighted sum of historical randomness(long term variance for up to p periods plus the instantaneous variance for q periods(Financier, 2018).

3.4.3 *GARCH Model on Assets*

The results after we fit in GARCH models are reproduced in Fig. 3.12. These plots of the auto-correlation functions of the residuals and squared residuals of the fitted GARCH(1,1) model show that GARCH were able to practically completely removed the serial correlation existed in ARIMA model residuals.

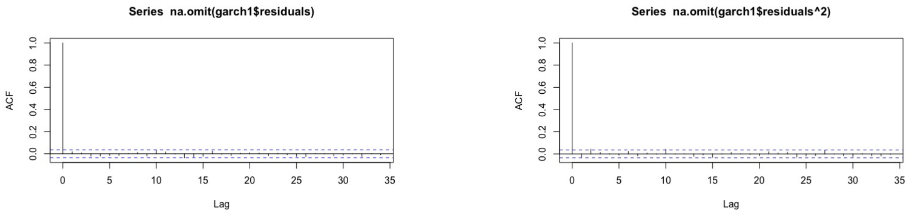


Figure 3.10: POAGX: ACF plots of residuals and residuals square from GARCH(1,1) models

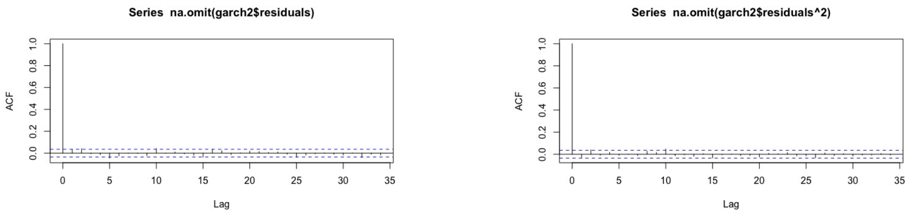


Figure 3.11: S&P: ACF plots of residuals and residuals square from GARCH(1,1) models

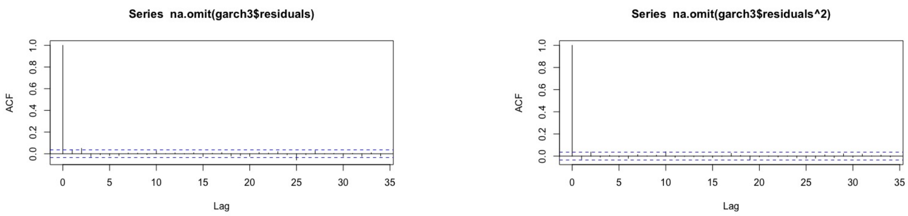


Figure 3.12: FDFAX: ACF plots of residuals and residuals square from GARCH(1,1) models

As one of the most distinctive feature of the GARCH model is to distinguish unconditional variance and produce the conditional variance, it is insightful to look at the model estimated instantaneous conditional variance form historical data at each time point in the dark line. In figure 3.13, the instantaneous conditional sigma estimated in the fitting of a GARCH(1,1) model to the ARMA residuals from the POAGX returns is shown in dark line. In companion with it, is a rolling standard deviation estimate with weight equally distributed to the previous 60 past volatility in red line. We see in general conherence in these two estimates, while the GARCH estimate gives more sharp and quick responds in corresponding period, while the rolling average is smoother as expected. In the down panal shows the orginal log return series as numerical vector (therefore, the index is not time perspecive) with 2 conditional standard deviation superimposed. GARCH is powerful and sophisicated to predict the confidence interval for the true sigma with more than 90% confidence.

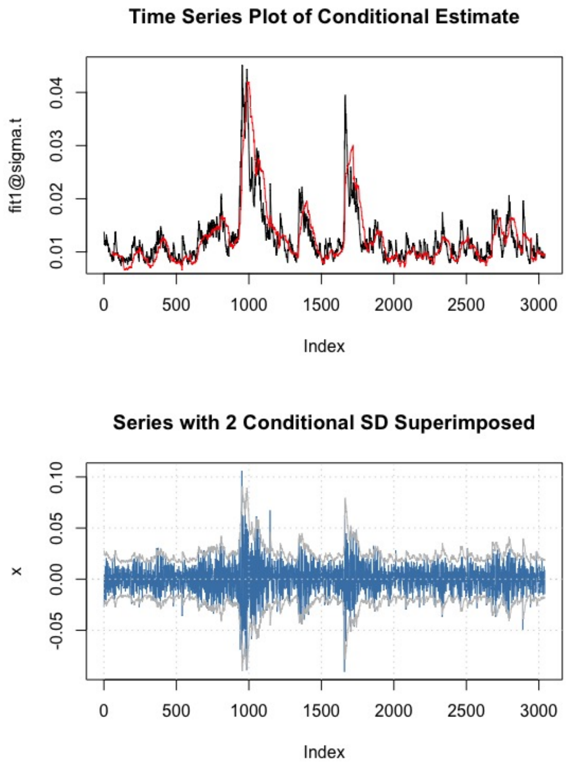


Figure 3.13: Conditional standard deviation as estimated in the fitting of a GARCH(1,1) model to the ARMA residuals from the POAGX returns

To visualize the quality of the fit of the GARCH model, we draw the scatter plot 3.14 of the fitted values against the actual values to which the GARCH model was fitted. The fact that the points are found around a straight lines for all of the three assets is a good indication of the quality of the fit.

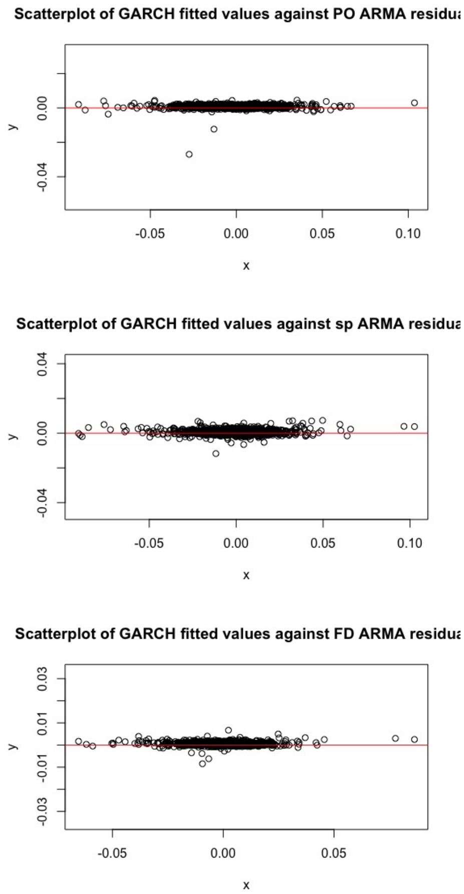


Figure 3.14: ACF plot of the residuals square

For simplicity reasons, the same set of plots are not included in this thesis, as evidence are similar to support the power of GARCH model.

3.5 FORECASTING

3.5.1 *Hybrid ARIMA and GARCH model*

As we have illustrated, returns will often possess an ARMA (mean) structure with GARCH (volatility) errors.

The regression equation for a simple AR-GARCH model is

$$r_t = \mu_t + \sigma_t \epsilon_t .$$

where the mean term μ_t is an ARMA process

$$\mu_t = \mu + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \cdots + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots .$$

and σ_t has GARCH behavior:

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$

For example, referring to the model output 3.15, the the hybrid model for the first assets, POAGX are:

```
Coefficient(s):
           mu            ar1            ar2          omega         alpha1          beta1
 0.0007832601   0.0134686061  -0.0378046939   0.0000038865   0.0871868650   0.8887108620

Std. Errors:
 based on Hessian

Error Analysis:
            Estimate    Std. Error  t value             Pr(>|t|)
mu      0.0007832601  0.0001940765    4.036 0.000054409126024391 ***
ar1     0.0134686061  0.0192306396    0.700               0.4837
ar2    -0.0378046939  0.0189071838   -1.999               0.0456 *
omega   0.0000038865  0.0000008383    4.636 0.000003549576055306 ***
alpha1  0.0871868650  0.0107955175    8.076 0.000000000000000666 ***
beta1   0.8887108620  0.0136049438   65.323 < 0.0000000000000002 ***
---
```

Figure 3.15: Hybrid model for POAGX returns

$$\mu_t = \mu + \phi_1 r_{t-1} + \phi_2 r_{t-2} + +\sigma_t \epsilon_t$$
$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$
$$\mu_t = 0.00078 + 0.01347 r_{t-1} + 0.03780 r_{t-2} + \epsilon_t \sigma_t$$
$$\sigma_t^2 = 0.00000 + 0.08719 r_{t-1}^2 + 0.88871 \sigma_{t-1}^2$$

With the models identified and parameters estimated, now we are at the point of forecasting. The procedures are as follows, we first estimate the mean returns for the coming period, according to the ARMA model; in parallel, we have our volatility forecast from GARCH model and from the fitted value from ARMA forecast:

$$
\begin{aligned}
\hat{\sigma}_{t+1}^2 &= \hat{\omega} + \hat{\alpha} r_t^2 + \hat{\beta}\sigma_t^2 \\
\hat{r}_{t+1}^2 &= \hat{\sigma}_{t+1}^2 \\
\hat{\sigma}_{t+2}^2 &= \hat{\omega} + \hat{\alpha} r_{t+1}^2 + \hat{\beta}\hat{\sigma}_{t+1}^2 \\
\hat{r}_{t+2}^2 &= \hat{\sigma}_{t+2}^2 \\
&\ldots
\end{aligned}
$$

With this rational, we have shown the prediction result of POAGX returns in figure 3.16 with the realized return value in transparent red, the 30 day ahead forecast in red in the right tail of the index, with upper and lower confidence interval bounded, we see that the boundary is able to cover the realizations in general.

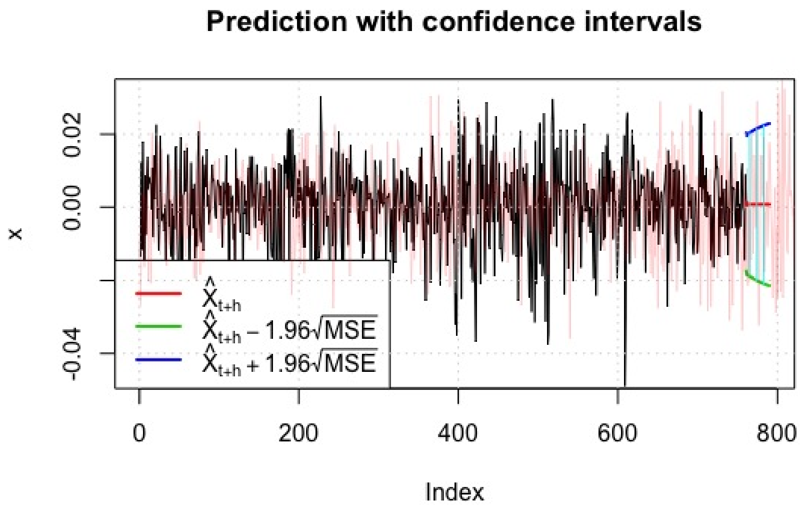


Figure 3.16: ARMA-GARCH Prediciton for POAGX returns

CHAPTER 4

Conclusion

4.1 Summary

In this project, we investigate the time series theories and models on three asset processes. We examine three different assets including a stock index, a diversified and un-diversified portfolio mutual funds repectively. We examine various ideas in time series and check the staionarity, independnece, correlation and normality. After that, we observe the data patterns and fit in different models for trends and also fit in GARCH models for capturing the variance variability.

From a hands-on point of view, we appreciate the concepts and models developed in time-series arena. We see that, time series models, in general, are able to capture a large part of the pattern inhibited in the data structure. Especially with GARCH model for capturing the variance clustering phenomenoun in usual financial data series.

We acknowledge the fact that the time is limited for us to extend further to consider and examine the long term-memories of financial data series by polynomial models and fractionally integrated models. And we may include cross-sectional analyses such as incorporating factor models in our three assets in the future.

II

Appendices

Bibliography

Lo, Andrew W and A Craig MacKinlay (1988). "Stock market prices do not follow random walks: Evidence from a simple specification test". In: *The review of financial studies* 1.1, pp. 41–66.

Belaire-Franch, Jorge and Dulce Contreras (2002). "How to compute the BDS test: a software comparison". In: *Journal of Applied Econometrics* 17.6, pp. 691–699.

Hull, John C. (2009). *Options, Futures, and Other Derivatives*. Dorling Kindersley.

Metcalfe, Andrew V and Paul SP Cowpertwait (2009). *Introductory time series with R*.

Carmona, René (2014). *Statistical analysis of financial data in R*. Vol. 2. Springer.

Simaan, Majeed (2017). "Financial Time Series Analysis using R". In: *Interactive Brokers Webinar Series*.

Brownlee, Jason (2018). *How to Check if Time Series Data is Stationary with Python*. URL: https://machinelearningmastery.com/time-series-data-stationary-python/ (visited on 2018).

Eulogio, Raul (2018). *Performing a Time-Series Analysis on the SP 500 Stock Index*. URL: https://www.datascience.com/blog/stock-price-time-series-arima (visited on 2018).

Financier, Quantum (2018). *Basic Introduction to GARCH and EGARCH (part 1)*. URL: https://quantumfinancier.wordpress.com/2010/09/12/381/ (visited on 2018).

Kuan, Chung-Ming (2018). *Lecture Slides diagnostic tests*. URL: http://homepage.ntu.edu.tw/~ckuan/pdf/Lec-DiagTest.pdf (visited on 2018).

R Programming (2018). URL: http://cran.r-project.orgf (visited on 2018).