A Financial Time Series Analysis of Microsoft Corporation Stock Returns

Stevens Institute of Technology

FIN620-A Financial Econometrics

Instructor: Pallavi Pal

Group ABC

Zhengyuan Cai

Haoran Peng

Yuhan Sun

Juelun Wang

May 14th, 2022

a. Overview of the asset and the market for the asset the Microsoft

Microsoft Corporation was founded in 1975 and is based in Redmond, Washington. It develops, licenses, and supports software, services, devices, and solutions worldwide. Its Productivity and Business Processes segment offers Office, Exchange, SharePoint, Microsoft Teams, Office 365 Security and Compliance, and Skype for Business, as well as related Client Access Licenses (CAL); Skype, Outlook.com, OneDrive, and LinkedIn; and Dynamics 365, a set of cloud-based and on-premises business solutions for organizations and enterprise divisions. Its Intelligent Cloud segment licenses SQL, Windows Servers, Visual Studio, System Center, and related CALs; GitHub provides a collaboration platform and code hosting service for developers; and Azure, a cloud platform. It also offers support services and Microsoft consulting services to assist customers in developing, deploying, and managing Microsoft server and desktop solutions; and training and certification on Microsoft products. Its More Personal Computing segment provides Windows original equipment manufacturer (OEM) licensing and other non-volume licensing of the Windows operating system; Windows Commercial, such as volume licensing of the Windows operating system, Windows cloud services, and other Windows commercial offerings; patent licensing; Windows Internet of Things; and MSN advertising. It also offers Surface, PC accessories, PCs, tablets, gaming and entertainment consoles, and other devices; Gaming, including Xbox hardware, and Xbox content and services; video games and third-party video game royalties; and Search, including Bing and Microsoft advertising. It sells its products through OEMs, distributors, and resellers; and directly through digital marketplaces, online stores, and retail stores. It has collaborations

with Dynatrace, Inc., Morgan Stanley, Micro Focus, WPP plc, ACI Worldwide, Inc., and iCIMS, Inc., as well as strategic relationships with Avaya Holdings Corp. and wejo Limited.

b. One application for the econometric analysis

An article by Phillips, Wu, and Yu (2011; PWY) developed a new econometric methodology for real-time bubble detection. When it was applied to NASDAQ data in the 1990s, the algorithm revealed that evidence in the data supported Greenspan's declaration of "irrational exuberance" in December 1996 and that this evidence of market exuberance had existed for some 16 months prior to that declaration.

c. Properties of the time-series

i.Descriptive statistics

Daily Data

A small amount of statistical analysis on the simple daily data of Microsoft stock are shown in Table 1.1, from which the following points can be deduced. The minimum daily return is -14.74 percent, while the highest daily return is 18.60 percent. Daily returns on average are 0.07 percent.

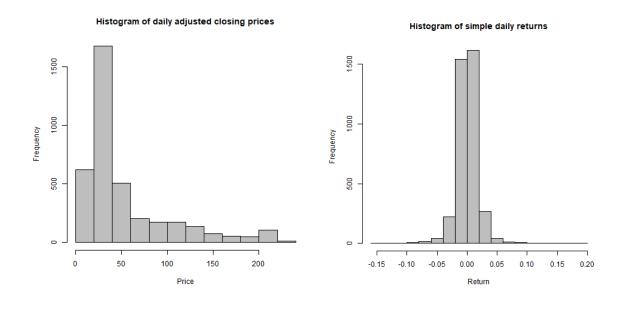
Table 1.1 Summary of the Daily Data

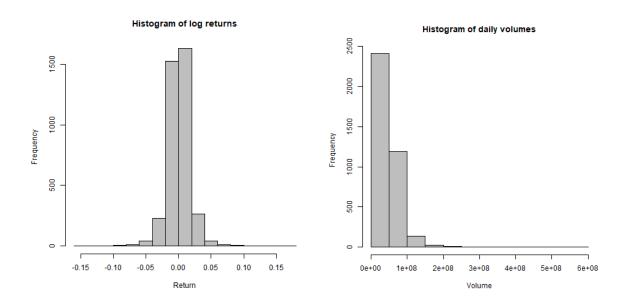
	Daily Data			
Statistics	Adjusted Closing Prices	Simple Returns	Log Return	Volumes
Min.	11.48797	-0.147390	-0.159453	7425600
1st Qu.	21.07125	-0.007295	-0.007266	27556950
Median	27.85615	0.000485	0.000551	41418000

Mean	52.19217	0.000721	0.000644	47545530
3rd Qu.	61.03758	0.008774	0.008770	58987200
Max.	228.6517	0.186047	0.170626	591052200
Standard Deviation	48.78083	0.017678	0.017640	29685430
Skewness	1.812455	0.017678	-0.057912	3.664046
Kurtosis	2.530944	11.050462	10.602632	38.23185

The histogram plot of Microsoft's simple daily returns can be seen in Figure 1.1. The skewness data in Table 1.1 and Figure 1.1 indicate that the data distribution is skewed to the right, with a skewness of 0.018. If kurtosis equals 3, the distribution is normal. The kurtosis of this distribution is 11.00, which is greater than 3. Thus, it demonstrates a "heavy-tailed" distribution.

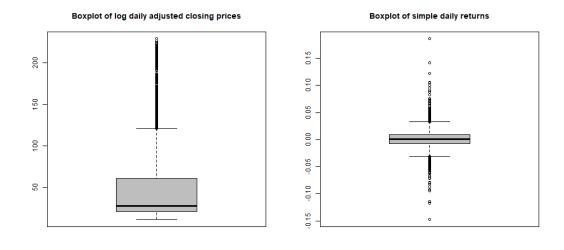
Figure 1.1 Histogram of the daily prices, simple and log returns, and volumes

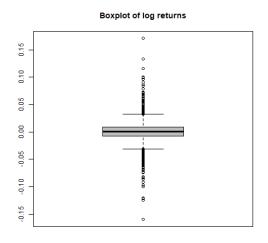


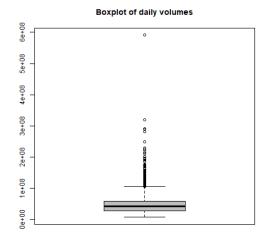


The boxplot in Figure 1.2 depicts Microsoft's simple daily returns. 0.0485 percent is the median. The first quartile has a value of -0.7295 percent, while the third quartile has a value of 0.8774 percent. In addition, 1.6069 percent is the interquartile range. Meanwhile, this distribution has several outliers.

Figure 1.2 Box plot of the daily prices, simple and log returns, and volumes







Monthly data

Several descriptive statistics on the simple monthly returns of Microsoft stock are shown in Table 1.2, from which the following statements can be inferred. The lowest monthly return touched -16.56 percent, whereas the highest monthly return reached 24.95 percent.

Furthermore, the average of monthly returns is 1.41 percent.

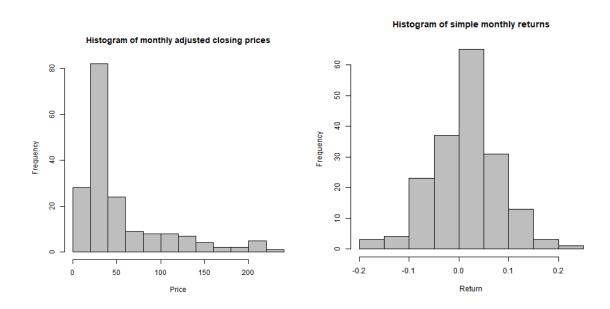
Table 1.2 Summary of the monthly data

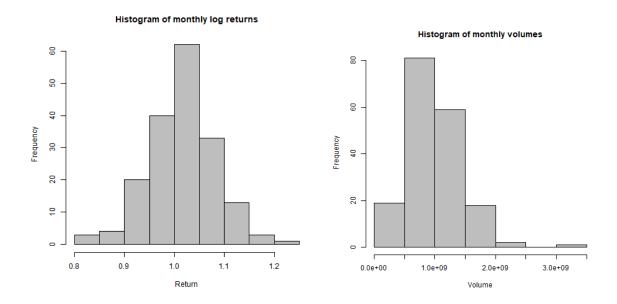
	Monthly Data				
Statistics	Adjusted Closing Prices	Simple Returns	Log Returns	Volumes	
Min.	12.246252	-0.165644	0.836643	375983900	
1st Qu.	21.092423	-0.028751	0.971476	639101400	
Median	28.314758	0.018239	1.020515	941553100	
Mean	52.673486	0.014115	1.015576	997135400	
3rd Qu.	62.086064	0.053980	1.055540	1292113000	
Max.	222.610931	0.249491	1.249491	3044579000	

Standard Deviation	49.461479	0.066706	0.066495	438141900
Skewness	1.783915	0.062257	0.054947	1.046840
Kurtosis	2.407955	0.740480	0.723427	1.912837

The histogram plot of Microsoft's simple monthly returns is shown in Figure 1.3. The skewness data in Table 1.2 and Figure 1.3 reveal that the data distribution is skewed to the right, with a skewness of 0.0623. If the kurtosis equals 3, the distribution is normal. However, the kurtosis of this distribution is 0.81, which is below 3. It has a very short tail.

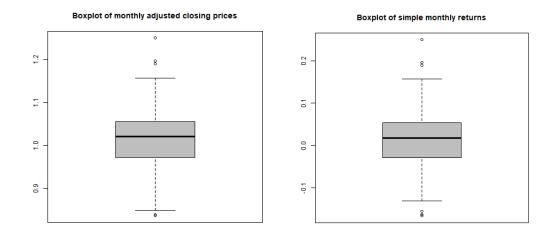
Figure 1.3 Histogram of the monthly prices, simple and log returns, and volumes

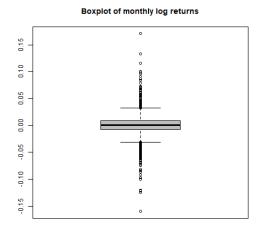


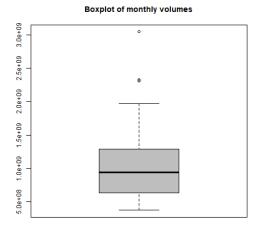


The boxplot in Figure 1.4 illustrates Microsoft's simple monthly returns. 1.8239 percent is the median. The first quartile has a value of -2.8751 percent, while the third quartile has a value of 5.3980 percent. 8.2731 percent is the interquartile range. In addition, this distribution has a few outliers.

Figure 1.4 Box plot of the monthly prices, simple and log returns, and volumes







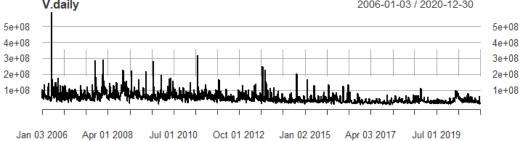
ii.Visualization

In this part, we will simply estimate the stationarity of four different time series based on their time series plot.

Microsoft's daily adjusted close price and volume are displayed in Figure 2.1. According to the graph, MSFT's price remained constant for the majority of the time before 2013. After 2013, it began an exponential ascent. Nevertheless, the volume continued to decrease over time. Particularly as stock prices skyrocketed, volume remained extremely low in comparison to the start of the period. The time series of daily prices are not stationary.

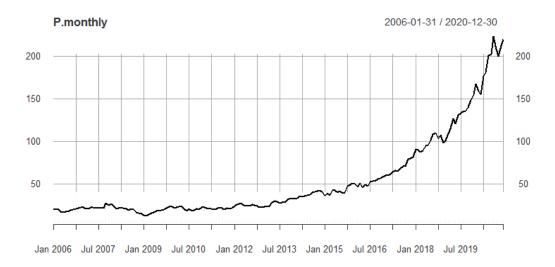
Figure 2.1 Adjusted closing price and volume of Microsoft Corp.





Microsoft's monthly price and volume are visualized in Figure 2.2. According to the figure, the monthly price performed roughly identically to the daily price. They both exhibit similar tendencies, but the monthly price has a smoother line than that of the daily price. Also, The series of monthly prices is obviously nonstationary—it has a strong upward trend.

Figure 2.2 Monthly price and volume of Microsoft Corp.



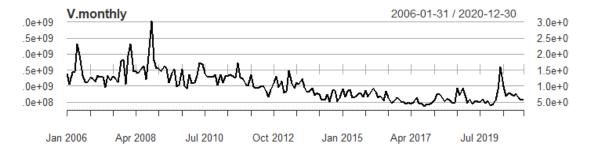
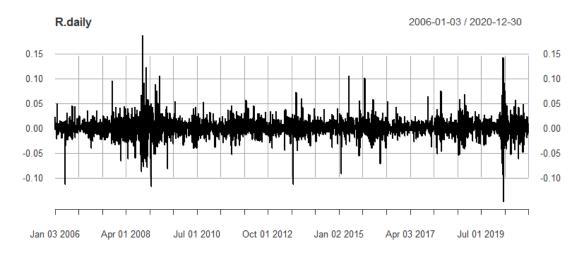


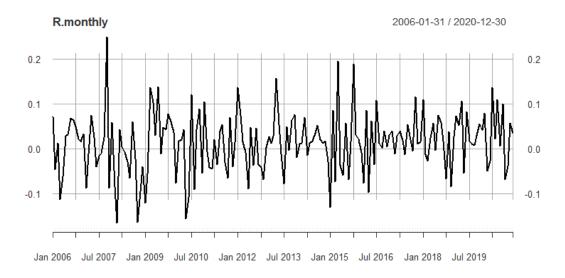
Figure 2.3 exhibits Microsoft's simple daily returns. From the figure, the daily return on the MSFT price fluctuated around some point. The time series of daily returns is basically stationary.

Figure 2.3 Simple daily returns of Microsoft Corp.



Microsoft's simple monthly returns are portrayed in Figure 2.4. Monthly returns on the MSFT price swung around a certain point, as indicated by the figure. Daily returns are essentially stationary time series.

Figure 2.4 Simple monthly returns of Microsoft Corp.



Microsoft's daily and monthly log returns are separately shown in Figure 2.5 and Figure 2.6. They are pretty much identical to the diagrams of simple returns.

Figure 2.5 Log daily returns of Microsoft Corp.

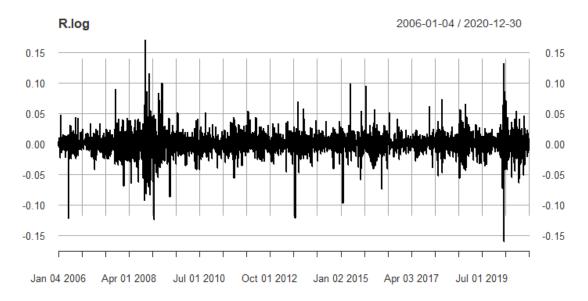
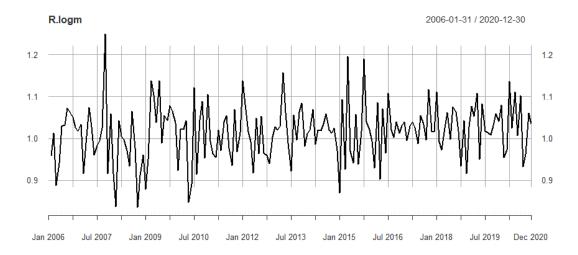


Figure 2.6 Log monthly returns of Microsoft Corp.



iii. Unit-root and seasonality tests

Numerous economic and financial time series display either trending behavior or mean nonstationarity. The most prominent example is the daily and monthly price data shown as Figure 2.1 and 2.2. Identifying the best appropriate form of the trend in the data is a critical econometric task.

Unit root tests can be used to identify whether trending data should be differentiated or regressed on deterministic time functions first to achieve stationary data.

We already found that the time series of daily and monthly prices are not stationary based on Figure 2.1 and 2.2. Although we can convert them to stationary forms, we prefer the time series which can be directly used without differencing or with more simple differencing. Hence, it only left two time series of daily and monthly simple returns. Compared with daily returns, monthly returns can perform better in long-term trends. It already filtered some unexpected shocks or surprises, which have an adverse effect on efficiency and accuracy of the model. Therefore, we decided to use monthly data as our raw data.

Unit root test for monthly returns

We speculated that the time series of monthly returns is stationary. It can pass the ADF test, however, in the Ljung-box test we found that it is independent. Thus, we convert it into the first differencing form.

The Augmented Dickey-Fuller (ADF) test's null hypothesis is that it is non-stationary and has a unit root.

```
> R.nm <- na.omit(diff(R.monthly))
> adf.test(R.nm)

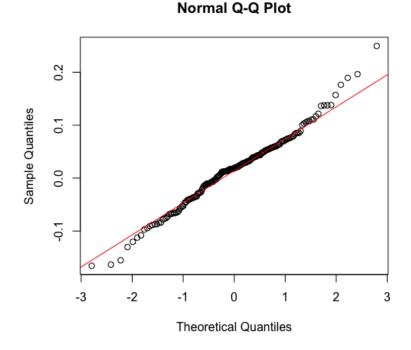
Augmented Dickey-Fuller Test

data: R.nm
Dickey-Fuller = -9.9489, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

The p-value is less than 0.05, as indicated above. This indicates that the null hypothesis should be rejected. In fact, the series of prices is stationary and does include a unit root. This can be used for the following analysis.

The Q-Q plot of the time series of simple monthly returns in Figure 3.1 demonstrates that, for the most part, the points distribution follows a straight diagonal line, except for the deviations of the two tiers. We can presume that the data in this collection are normally distributed.

Figure 3.1 The Q-Q plot of the simple monthly return series



Unit root test for monthly prices

Following that, we repeated the test using the monthly pricing. We employed the first differencing as well as seasonal differencing.

As noted, the p-value is less than 0.05. This suggests that the null hypothesis should be rejected. The price series is stationary and has a unit root. Likewise, monthly data can be used for the subsequent procedure.

d. ARIMA model of return and price

i. The ARIMA model for monthly returns

Ljung-Box Test

Prior to selecting the order of the ARIMA model, we should use the Ljung-Box test to determine whether the time series contains autocorrelation or dependence. This is because, without autocorrelation, this time series is white noise. In this case, we have no way of forecasting its performance.

The Ljung-Box model's null hypothesis is that it is independent; and there is no autocorrelation, implying that it is white noise.

Based on the preceding outcome, the p-value is less than 0.05. This does not deduce rejection of the null hypothesis. The series of simple monthly returns exhibits dependence. Therefore, it is not a source of white noise. We can forecast its future performance using the ARIMA model.

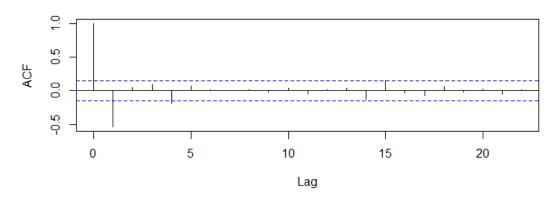
Identifying order

After confirming the feasibility of the time series we used, we started to identify the ARIMA model's order.

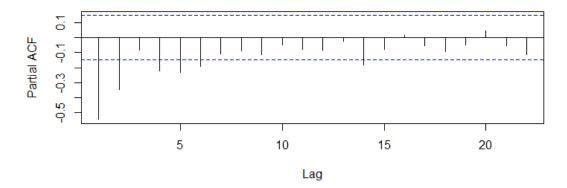
The ACF and PACF plots of the time series are shown in Figure 4.1. We notice that there is a single negative spike lag 1 in the ACF plot and a decay pattern (from below) in the PACF plot. We assume that the order can be (0, 1, 1), (6, 1, 1), (5, 1, 1), (2, 1, 1), (2, 1, 0), (5, 1, 0), and (6, 1, 0).

Figure 4.1 ACF and PACF charts





Series R.nm



After that, we used the EACF approach to estimate the AR and MA orders. From the EACF, the order (1, 1, 1), and the order (2, 2, 2) can be considered to establish the ARIMA model.

```
> eacf(R.monthly, ar.max = 10, ma.max = 10)
AR/MA
            0
          Х
             0
               0
          Х
             0
        Х
               0
          0
            0
        Х
               0
                  0
5
        Х
          0
             0
               0
        0
          0
             0
               0
        0
          0 0
                 0
               0
     x \circ o \circ
               0
                 0
   0 \ X \ X \ X \ 0 \ 0 \ 0 \ 0
10 o x x x o o o o o o o
```

At present, there are nine candidates. We generated nine models with nine distinct orders and then compared their AIC values to determine which order was the best fit.

```
> m1 = arima(R.monthly, order = c(0, 1, 1))
> m2 = arima(R.monthly, order = c(6, 1, 1))
> m3 = arima(R.monthly, order = c(5, 1, 1))
> m4 = arima(R.monthly, order = c(2, 1, 1))
> m5 = arima(R.monthly, order = c(2, 1, 0))
> m6 = arima(R.monthly, order = c(5, 1, 0))
> m7 = arima(R.monthly, order = c(6, 1, 0))
> m8 = arima(R.monthly, order = c(1, 1, 1))
> m9 = arima(R.monthly, order = c(2, 1, 2))
> m1$aic
[1] -452.9416
> m2$aic
[1] -449.47
> m3$aic
[1] -451.399
> m4$aic
[1] -449.7976
> m5$aic
[1] -407.8564
> m6$aic
[1] -424.3293
> m7$aic
[1] -429.5187
> m8$aic
[1] -451.6247
> m9$aic
[1] -447.805
> m1
```

Fitted model: $x_t = a_t - 0.9818a_{t-1}$

The ARIMA(0, 1, 1) has the smallest AIC value. Hence, we preferred to use it to continue the next steps.

Model checking

```
> checkresiduals(m1)
        Ljung-Box test

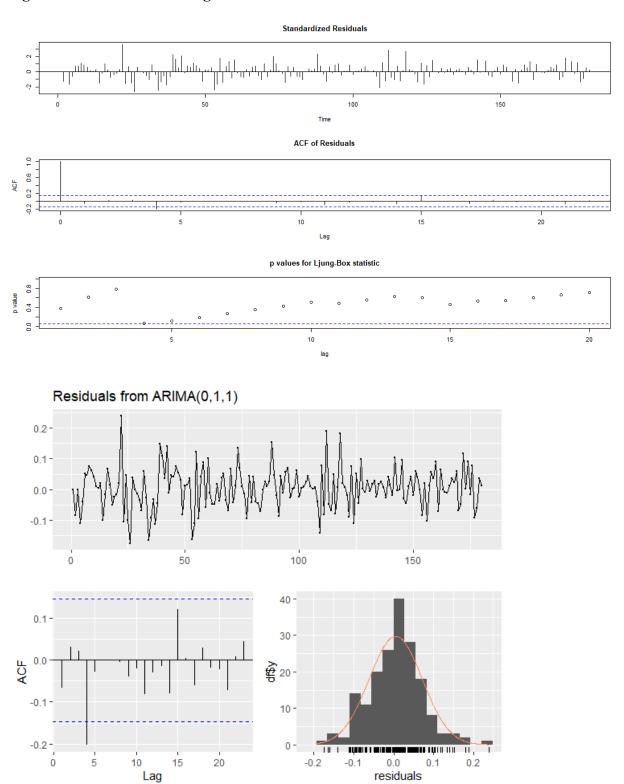
data: Residuals from ARIMA(0,1,1)
Q* = 9.1826, df = 9, p-value = 0.4206

Model df: 1. Total lags used: 10
```

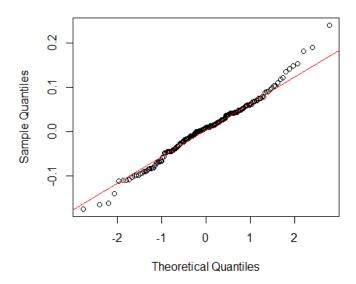
p-value = 0.4206 > 0.05, based on the above results. This demonstrates that the residuals are consistent with the assumption of uncorrelated white noise. This model was found to be valid in residuals checking.

The residuals checking procedure is informed in Figure 4.2. If the model is fitted correctly, the residual series should be random and devoid of autocorrelations. The residuals follow the normal distribution as seen by the histogram and the qq-plot. As the ACF plot indicates, the residuals are uncorrelated, with no points crossing the dotted line. The p-values of residuals are far away from the 5% confidence level.

Figure 4.2 Residuals checking



Normal Q-Q Plot



Forecast result

The ARIMA(0, 1, 1) model was used to forecast the future ten months' returns. The predicted monthly return is 2.189 percent. The information of the confidence interval is shown in Table 4.1.

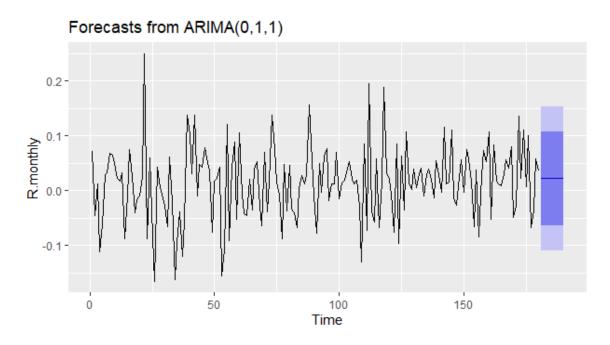
Table 4.1 The forecast value of next ten months returns

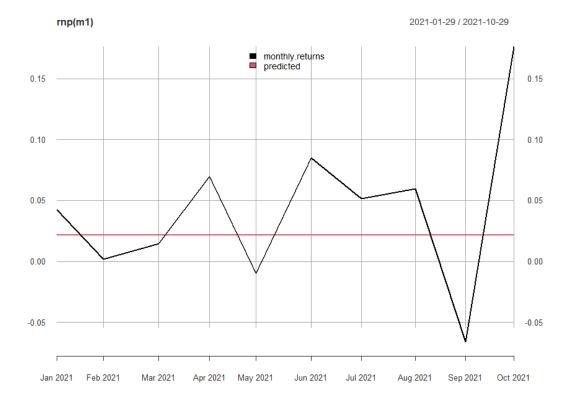
	LCL	Prediction(R)	UCL
181	-0.1092349	0.02188899	0.1530128
182	-0.1092567	0.02188899	0.1530346
183	-0.1092785	0.02188899	0.1530565
184	-0.1093003	0.02188899	0.1530783
185	-0.1093221	0.02188899	0.1531001
186	-0.1093439	0.02188899	0.1531219
187	-0.1093657	0.02188899	0.1531437
188	-0.1093875	0.02188899	0.1531655

189	-0.1094093	0.02188899	0.1531873
190	-0.1094311	0.02188899	0.1532091

We plotted this result in Figure 4.3. The prediction is basically consistent with historical data. The real return fluctuated around the predicted return. Accordingly, the model we constructed performed well.

Figure 4.3 Forecasts from ARMA





ii. The ARIMA model for monthly prices

Following that, we began developing the ARIMA model for monthly prices.

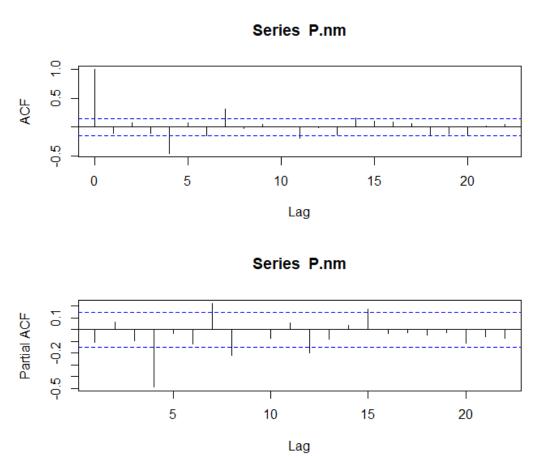
Ljung-Box Test

The previous result revealed that the p-value is less than 0.05. This does not imply that the null hypothesis is rejected. The time series of simple monthly returns reflects dependence and is not a source of white noise. We can forecast its future performance using this approach.

Identifying order

The ACF and PACF plots of the time series can be seen in Figure 5.1. Both the ACF and PACF plots look completely random, their trends are not clear. Lag 12, which in the PACF plot is significant, is not significant in the ACF plot. We experimented with AR(1) and MA(1) separately for seasonal order.

Figure 5.1 ACF and PACF charts



After that, we used the EACF approach to identify non-seasonal AR and MA orders. From the EACF, the order (1, 1, 7), (2, 1, 7), (3, 1, 7) can be considered for the model.

```
> eacf(P.monthly, ar.max = 10, ma.max = 10)
AR/MA
   0 1 2 3 4 5 6 7 8 9 10
 x x x x x x x x x x x
1 x o o o o o x o o o o
2
  X X O X X O X O O O X
3
  x o x x x o x o o o x
  x x x o o o x o o o x
5 x o x o x o x o o o o
6
 x x o o x o x x o o o
  0 0 0 0 0 X X 0 0 0 0
8 xoxxxoxoooo
9 x o o o x o x o o o o
10 x x x o o o x x x o o
```

Below are some orders that we have tried. We built nine models with nine different orders and then compared their AIC values to decide which order was the most acceptable one.

```
> m11 = arima(P.monthly, order = c(1, 1, 7), seasonal = c(0, 1, 1))
> m12 = arima(P.monthly, order = c(1, 1, 7), seasonal = c(1, 1, 0))
> m13 = arima(P.monthly, order = c(2, 1, 7), seasonal = c(0, 1, 1))
> m14 = arima(P.monthly, order = c(2, 1, 7), seasonal = c(1, 1, 0))
> m15 = arima(P.monthly, order = c(3, 1, 7), seasonal = c(0, 1, 1))
> m16 = arima(P.monthly, order = c(3, 1, 7), seasonal = c(1, 1, 0))
> m17 = arima(P.monthly, order = c(2, 1, 2), seasonal = c(1, 1, 0))
> m18 = arima(P.monthly, order = c(3, 1, 2), seasonal = c(1, 1, 0))
> m19 = arima(P.monthly, order = c(3, 1, 3), seasonal = c(1, 1, 0))
> m11$aic
[1] 997.492
> m12$aic
[1] 1006.576
> m13$aic
[1] 998.3048
> m14$aic
[1] 1005.255
> m15$aic
[1] 1000.268
> m16$aic
[1] 1006.215
> m17$aic
[1] 1009.836
> m18$aic
[1] 1009.313
> m19$aic
[1] 1008.966
```

ARIMA(1, 1, 7),(0, 1, 1) has the smallest AIC value. However, its residuals are not independent, which cannot pass the model checking. On account of this, we selected the second smallest one, ARIMA(2, 1, 7),(0, 1, 1).

We obtained nine non-seasonal coefficients and one seasonal coefficient during the model's construction. Nonetheless, with the exception of MA(7), the bulk of non-seasonal coefficients are non-significant.

Fitted model: $x_t = 0.4565a_{t-7} - 0.9263sa_{t-1}$

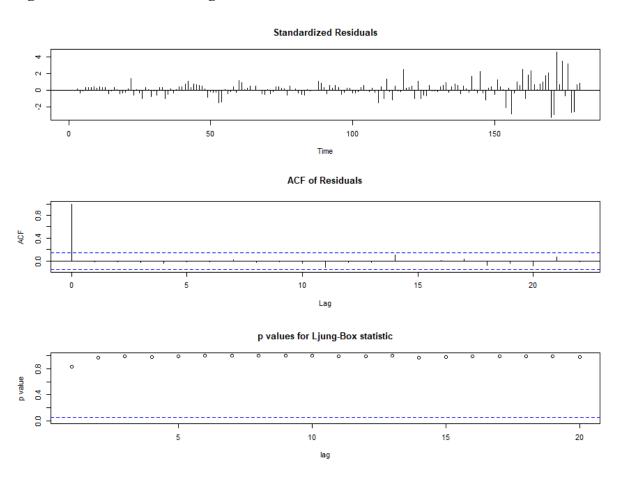
Model checking

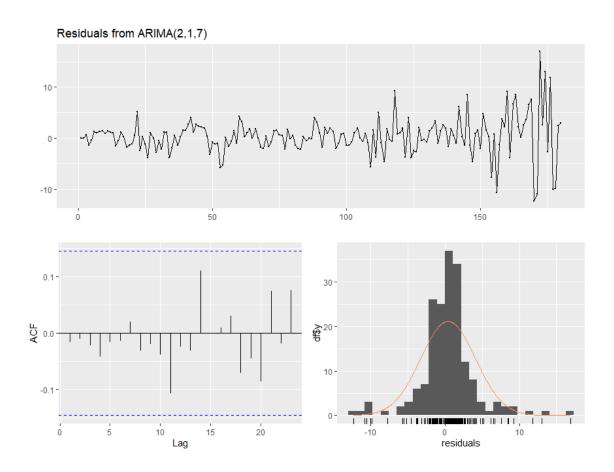
p-value = 0.2957 > 0.05, based on the above results. This demonstrates that the residuals are consistent with the assumption of uncorrelated white noise. This model was found to be valid in residuals checking.

The residuals checking procedure can be informed in Figure 5.2. If the model is fitted properly, the residual series should be random and devoid of autocorrelation. The residuals follow the normal distribution as seen by the histogram and the qq-plot. As indicated by the

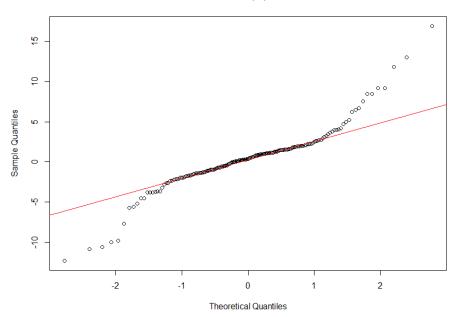
ACF plot, the residuals are uncorrelated, with no points crossing the dotted line. The p-values of residuals deviates considerably from the 5% confidence level.

Figure 5.2 Residuals checking









Forecast result

The ARIMA(2, 1, 7), (0, 1, 1) model was utilized to forecast the next ten months' prices.

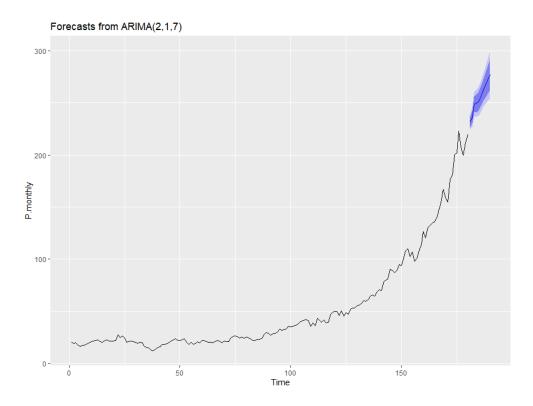
The average estimated monthly price is \$255.15203. The confidence intervals are included in Table 5.1.

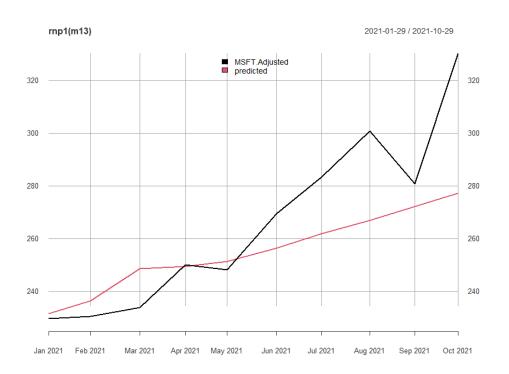
Table 5.1 The forecast value of next ten months prices

	LCL	Prediction(P)	UCL
181	224.1656	231.4383	238.7110
182	227.0030	236.4078	245.8126
183	237.1681	248.5378	259.9076
184	236.6613	249.3648	262.0683
185	237.5515	251.3883	265.2251
186	241.5612	256.3261	271.0910
187	246.2591	261.9291	277.5991
188	249.0386	266.9404	284.8422
189	251.9770	272.0853	292.1935
190	254.4977	277.1024	299.7070

On Figure 5.3, we plotted this result. The prediction is essentially accurate when compared to previous data. The actual return changed within a few percentage points of the expected return. Thus, our model performed admirably.

Figure 5.3 Forecasts from ARMA





e. Multivariate analysis: VAR

i. Selection of the second variate

The stock price is affected by a multitude of factors, including the GDP, inflation, market sentiment, and the business performance of the company. Among these elements, Microsoft Corporation's stock price may be more closely related to the overall market performance.

The S&P 500 Index is a market-capitalization-weighted index of 500 leading publicly traded companies in the U.S. Because of its depth and diversity, the S&P 500 is widely regarded as one of the best gauges of large U.S. stocks, and even the entire equity market. For our second variable, we assessed that the S&P 500 index would be a good fit. We substituted the S&P 500 ETF (SPY) for it because its quote is an index, not a price.

ii. The VAR model for monthly returns

Test for cross-autocorrelations

On the monthly returns of MSFT and SPY, we employed first differencing.

There exist cross-autocorrelations between MSFT's and SPY's monthly change of return using the CCF approach. As illustrated in Figure 6.1, each CCF plot contains spikes, indicating that the spike point exhibits statistically significant autocorrelation. Additionally, Figure 6.2 helps substantiate this claim.

Figure 6.1 The CCF plot

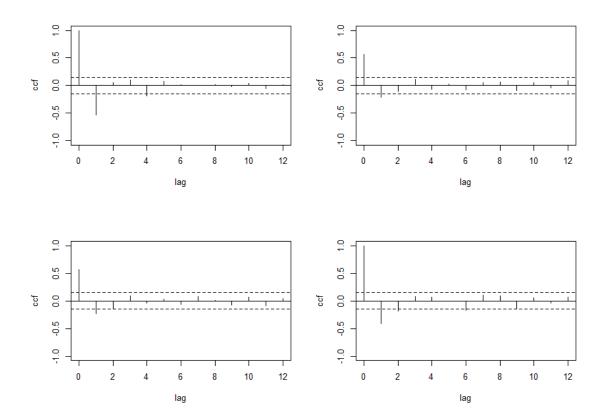
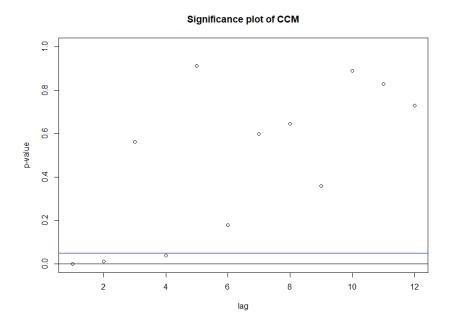
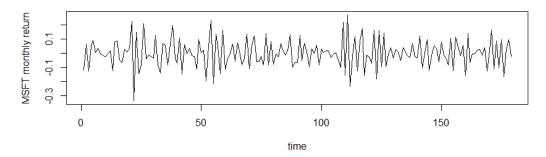


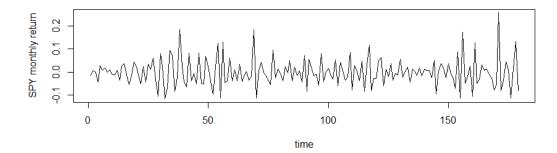
Figure 6.2 The significance plot of CCM for MSFT's and SPY's monthly change of return



The monthly return on MSFT and SPY is depicted in Figure 6.3. Microsoft's monthly return movements, as illustrated in the graph, are extremely comparable to those of the SPY.

Figure 6.3 The plot of MSFT's and SPY's monthly change of return





Identifying order

The function indicates that the ideal order is six for the AIC method, three for the BIC method, and five for the HQ method. With lag 6, we used the AIC technique.

```
> VARorder(gt)
selected order:
selected order: bic =
selected order:
                hq
Summary table:
                                          M(p)
                                                p-value
       0 -10.6258 -10.6258 -10.6258
                                        0.0000
                                                 0.0000
 [2,]
[3,]
[4,]
[5,]
                                                 0.0000
       1 -11.2810 -11.2098 -11.2522
                                      113.7335
       2 -11.5698 -11.4273 -11.5120
                                       53.5190
                                                 0.0000
       3 -11.6565 -11.4428 -11.5699
                                       20.8303
                                                 0.0003
       4
         -11.7019
                   -11.4170
                            -11.5863
                                       14.0911
                                                 0.0070
       5 -11.7758
                  -11.4196 -11.6314
                                       18.3222
                                                 0.0011
       6 -11.7987
                  -11.3714 -11.6254
                                       10.3178
                                                 0.0354
       7
        -11.7811 -11.2825 -11.5789
                                        4.0689
                                                 0.3968
       8 -11.7687 -11.1989 -11.5376
                                        4.7950
                                                 0.3090
       9 -11.7852 -11.1442 -11.5253
                                        8.9732
                                                 0.0618
```

Cointegration test

We used a parameter that lag is 6 to proceed with the cointegration test.

```
#####################
# Johansen-Procedure #
####################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.3866772 0.3059433
Values of teststatistic and critical values of test:
Eigenvectors, normalised to first column:
(These are the cointegration relations)
                      MSFT.monthly.return.l6 SPY.monthly.return.l6
MSFT.monthly.return.l6 1.000000 1.000000
SPY.monthly.return.l6 -0.3623889 -2.892432
SPY. monthly. return. 16
Weights W:
(This is the loading matrix)
                     MSFT.monthly.return.l6 SPY.monthly.return.l6
MSFT.monthly.return.d -4.902923 0.1214066
SPY.monthly.return.d -1.941784 1.3264847
```

Cointegration is the null hypothesis. The test, for the first time, rejects the null hypothesis. Therefore, there is no cointegration relationship between the monthly returns of MSFT and SPY.

iii. The VAR model for monthly prices

Test for cross-autocorrelations

On the monthly prices of MSFT and SPY, we adopted the first differencing method.

There exist cross-autocorrelations between the monthly price changes of MSFT and SPY using the CCF method. As described in Figure 7.1, each CCF plot contains spikes, implying that the spike point exhibits statistically significant autocorrelation. Likewise, Figure 7.2 can back up this claim.

Figure 7.1 The CCF plot

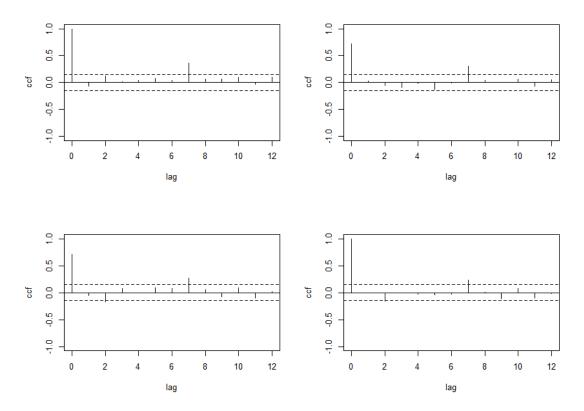


Figure 7.2 The significance plot of CCM for MSFT's and SPY's monthly change of prices

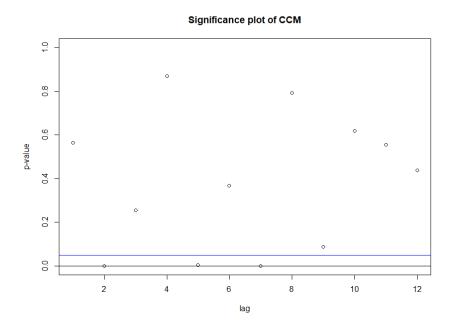
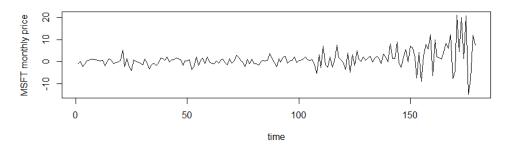
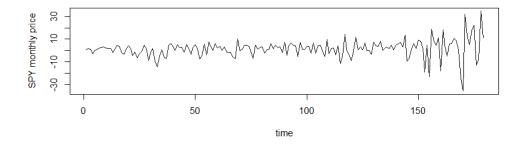


Figure 7.3 compares the monthly price movements of MSFT and SPY. As highlighted in the graph, Microsoft's monthly price fluctuations are quite similar to those of the SPY.

Figure 7.3 The plot of MSFT's and SPY's monthly change of prices





Identifying order

The function indicates that the appropriate order is ten in the AIC method, two in the BIC method, and seven in the HQ method. We used the AIC approach with lag 10.

```
> VARorder(gt1)
selected order:
                        10
                 aic =
selected order:
                 bic
selected order:
                 hq =
Summary table:
                                   M(p)
                                         p-value
             AIC
                    BIC
                             HQ
       Ö
                                 0.0000
                                          0.0000
         6.4622
                 6.4622
                         6.4622
 [2,]
[3,]
       1
         6.4881
                 6.5593
                        6.5170
                                 3.0605
                                          0.5478
       2 6.2967
                 6.4391
                        6.3544
                                37.8960
                                          0.0000
 [4,]
       3
         6.2821
                 6.4958
                        6.3688
                                 9.3911
                                          0.0520
 [5,]
       4
                        6.4062
         6.2906
                 6.5756
                                 5.6579
                                          0.2262
 [6,]
[7,]
       5
         6.2697
                 6.6258
                         6.4141
                                10.1477
       6
         6.2655
                 6.6929
                        6.4388
                                 7.4487
                                          0.1140
 [8,]
       7
         6.1008 6.5993 6.3029
                                31.5219
       8 6.0941 6.6639
                        6.3252
 [9,]
                                 7.6224
                                          0.1064
[10,]
       9 6.0723
                 6.7133
                        6.3322
                                 9.7493
                                          0.0449
      10
         6.0351
                 6.7474 6.3239
                                11.8257
                                          0.0187
[12,] 11 6.0538 6.8373 6.3715
                                 3.7095
                                          0.4468
[13,] 12 6.0875 6.9422 6.4341
                                 1.5399
```

Cointegration test

We used a parameter that lag is 10 to proceed with the cointegration test.

```
#####################
# Johansen-Procedure #
#######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend
Eigenvalues (lambda):
[1] 0.0807927151 0.0001319828
Values of teststatistic and critical values of test:
          test 10pct 5pct 1pct
r <= 1 | 0.02 6.50 8.18 11.65
r = 0 | 14.24 12.91 14.90 19.19
Eigenvectors, normalised to first column:
(These are the cointegration relations)
                       MSFT.monthly.price.l10 SPY.monthly.price.l10
MSFT.monthly.price.l10 1.000000 1.00000000 SPY.monthly.price.l10 -1.558268 -0.02565205
SPY.monthly.price.l10
Weights W:
(This is the loading matrix)
                    MSFT.monthly.price.l10 SPY.monthly.price.l10
MSFT.monthly.price.d 0.1437529 0.02700818
SPY.monthly.price.d 0.8070325 0.02804709
```

Cointegration is the null hypothesis. When r = 1, the test fails for the first time to reject the null hypothesis. Therefore, there is just one cointegration relationship.

As shown in Table 7.1, the following result is obtained. The table's first column contains variables with different lags. The second and final column represents the coefficients in the functions of MSFT and SPY respectively.

Table 7.1 The result of the VAR model for monthly change of price

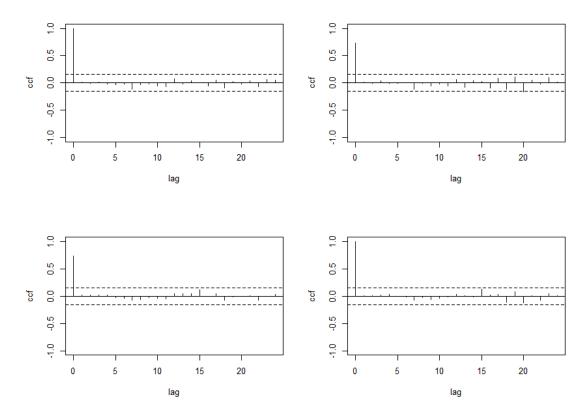
variables	MSFT	SPY
Const	0.43530424	1.069307556
$MSFT_1$	-0.42296832	-0.592519121
SPY_1	0.13554429	0.139489621

$MSFT_2$	0.32454488	-0.269486967
SPY_2	-0.18663300	-0.115597613
MSFT ₃	0.25834552	0.189303788
SPY_3	-0.17811104	-0.101095404
$MSFT_4$	-0.08658858	0.003936571
SPY_4	-0.03583365	-0.062743571
MSFT ₅	-0.02270406	0.139544447
SPY ₅	-0.07017737	-0.131170452
MSFT ₆	0.04548276	0.671430960
SPY_6	-0.00982154	-0.279540379
MSFT ₇	0.29349637	0.640363310
SPY_7	0.14178895	0.037953123
MSFT_8	0.26347081	0.298121044
SPY_8	0.02585900	-0.045394317
MSFT ₉	0.29377424	-0.178488886
SPY_9	-0.07493452	0.006348264
$MSFT_{10}$	0.22390861	-0.067124920
SPY_{10}	0.02761945	0.293458393

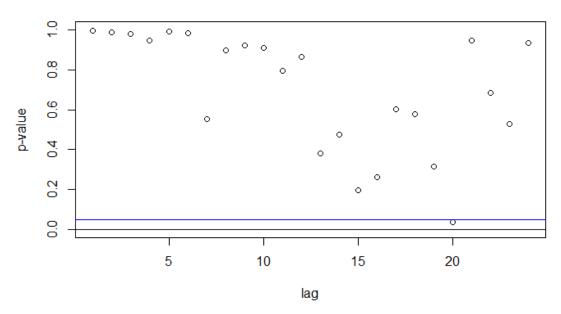
Model checking

The CCF plot shown at the top of Figure 7.4 illustrates that there are no cross-correlations between the residuals. Because all values except for lag zero are within the dotted lines. All p-values are more than the dash line in the Ljung-Box test presented at the bottom of Figure 7.4. This suggests that the residuals lack autocorrelations and are hence considered white noise. Thus, this model is valid in residuals checking.

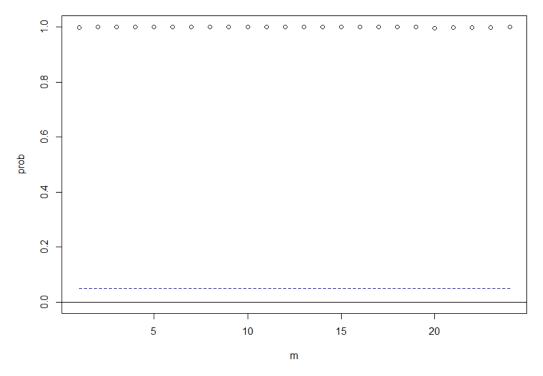
Figure 7.4 Residuals checking



Significance plot of CCM



p-values of Ljung-Box statistics



Forecast result

We can conclude from the result that the prediction follows the same trend for both MSFT and SPY's data. For MSFT, the monthly mean change of price is 6.6404, while the minimum and maximum value are -2.1847 and 16.6312, respectively. For SPY, the monthly mean change of price is 0.7545, while the minimum and maximum value are -22.056 and 23.259, respectively.

Table 7.2 The forecast value of monthly change of price

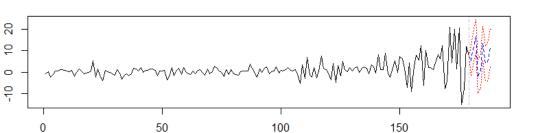
MSFT m	onthly chang	e of price	SPY mo	onthly change	of price
Prediction	Standard errors	Root mean square errors	Prediction	Standard errors	Root mean square errors

1	5.2618	3.424	3.620	-13.032	6.962	7.359
2	10.8596	3.564	4.590	13.858	7.115	8.315
3	16.6312	3.730	4.929	2.170	7.272	8.501
4	-2.1847	3.767	4.073	-22.056	7.277	7.314
5	0.4801	3.771	3.802	-13.681	7.286	7.362
6	13.2934	3.810	4.131	23.259	7.303	7.449
7	4.1631	3.817	3.882	12.353	7.458	8.673
8	4.6187	4.249	6.921	3.165	7.875	10.810

In Figure 7.5, the simulated results for MSFT and SPY are displayed individually.

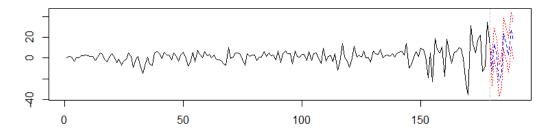
According to these forecasts, the MSFT and SPY have highly comparable changing tendencies that exhibit partial auto-correlation.

Figure 7.5 Forecasts from VAR model for monthly price



Forecast of series MSFT.monthly.price

Forecast of series SPY.monthly.price



f. Conditional variance analysis: Various types of GARCH models

i. The GARCH model for monthly returns

Test for ARCH effects

The null hypothesis is that the ACF of the squared residuals of mean series is zero for the first m lags. Due to the p-value of 0.0001584, we should reject the null hypothesis, indicating that the series has ARCH effects.

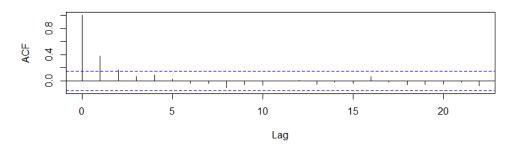
Identifying order

ACF: lag 1 or lag 2.

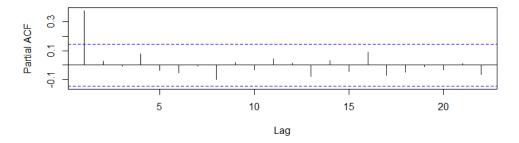
The PACF (following) of the squared values has a single spike at lag 1 suggesting an AR(1) model for the squared series. The ACF of the squared series follows an ARMA pattern because of both the ACF and PACF taper. This suggests a GARCH(1,1) model.

Figure 8.1 The ACF and PACF plots of squared series





Series R.nm^2



```
> eacf(R.nm^2)
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
      0 0 0 0 0 0 0 0
        0
          0 0 0
                0
                  0
                    0
                               0
        0
          0
            0 0
                0
                    0
                      0
      0
        0
          0 0 0 0
                  0
          0 0 0 0
    X X
        0
                  0
                    0 0
 x x x x o o o o o o o
 x x o x x o o o o o o
 0 X X X X 0 0 0 0 0 0
 eacf(abs(R.nm))
  0
   1
      2 3 4 5 6 7
                  8 9 10 11 12 13
      0 0 0 0 0 0
                  0
                    0
                      0
      0 0 0 0 0 0
                  0
                    0
                      0
                         0
   0 0 0 0 0 0 0
 0
                  0
                    0
   X  X  0  0  0  0  0
    x \ x \ x \ 0 \ 0 \ 0 \ 0
6 x x x x x x o o o o
                    0 0
                         0
                            0
                               0
   x x x x 0 0 0 0 0 0
```

We constructed five models with order (1, 1), (2, 1), (1, 2), (1, 0), and (2, 0) separately.

```
g01 <- garch(R.nm, order = c(1, 1))
g02 <- garch(R.nm, order = c(2, 1))
g03 <- garch(R.nm, order = c(1, 2))
g04 <- garch(R.nm, order = c(1, 0))
g05 <- garch(R.nm, order = c(2, 0))
> AIC(g01)
[1] -343.8483
> AIC(g02)
[1] -338.6735
> AIC(g03)
[1] -337.9566
> AIC(g04)
[1] -323.0878
> AIC(g05)
[1] -318.6138
```

The AIC value of the model does not considerably increase as the parameters are increased.

The most basic model, GARCH(1,1), is theoretically the best suitable for modeling.

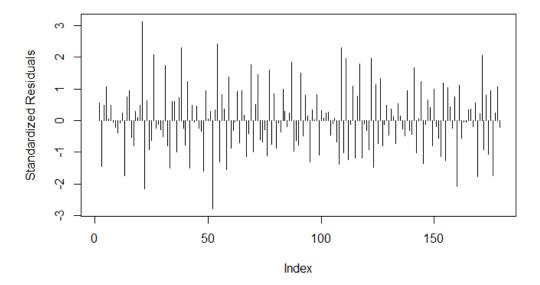
```
> summary(g01)
call:
garch(x = R.nm, order = c(1, 1))
Model:
GARCH(1,1)
Residuals:
   Min 1Q Median 3Q
                                   Max
-2.7765 -0.6926 -0.0468 0.6428 3.1173
Coefficient(s):
   Estimate Std. Error t value Pr(>|t|)
                         2.795 0.00519 **
a0 0.003586 0.001283
a1 0.390195 0.167328
b1 0.231108 0.205494
                           2.332 0.01971 *
1.125 0.26074
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '1
Diagnostic Tests:
        Jarque Bera Test
data: Residuals
X-squared = 1.527, df = 2, p-value = 0.466
        Box-Ljung test
data: Squared.Residuals
X-squared = 0.40512, df = 1, p-value = 0.5245
```

The above summary provides the Jarque Bera Test for the null hypothesis that the residuals are normally distributed and the familiar Ljung-Box Tests. Ideally all p-values are above 0.05, implying that there is no serial correlation in the residuals.

Model checking

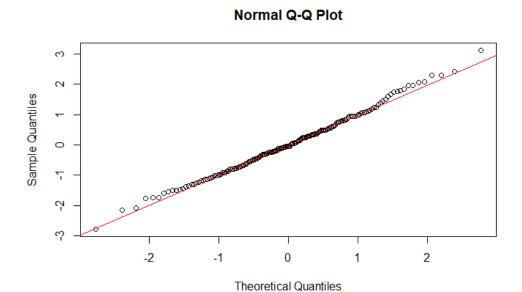
The residual series' time series diagram is shown in Figure 8.2. As can be seen, the residual series exhibits no clear clustering of fluctuations.

Figure 8.2 The plot of standardized residuals



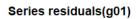
The QQ plot contains almost no departures from a straight line pattern, informing that the residuals are normally distributed.

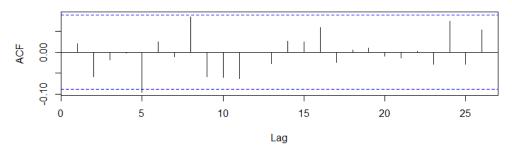
Figure 8.3 The qq plot of residuals



By comparing the ACFs of the standardized residuals to the ACFs of the squared standardized residuals, it is demonstrated that the GARCH model effectively explains the return series.

Figure 8.4 The ACF plot and Generalized Portmanteau Test of residuals





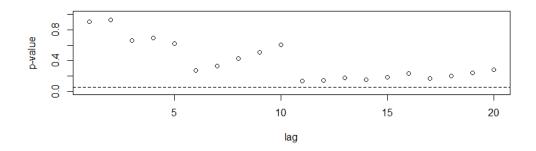
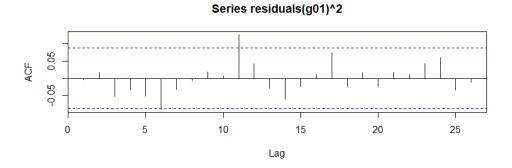


Figure 8.5 The ACF plot and Generalized Portmanteau Test of squared residuals



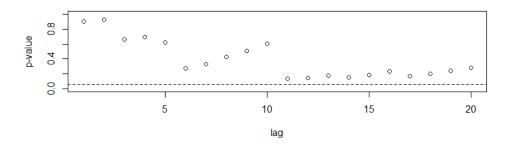
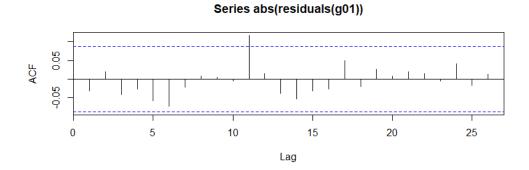
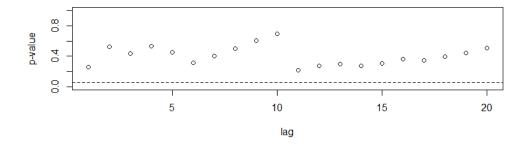


Figure 8.6 The ACF plot and Generalized Portmanteau Test of absolute residuals

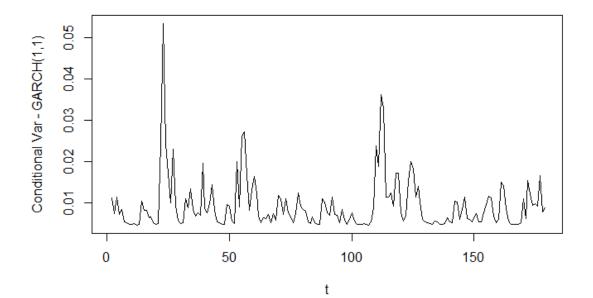




Forecast result

Figure 8.7 is a simulation of 200 observations from a GARCH(1,1) process. The conditional variance processes a relatively long-term persistence, at least compared to its behavior under an ARCH model. In particular, notice that the conditional variance is less "bursty" than for the ARCH(1) process.

Figure 8.7 The plot of forecast



g. Value-at-Risk analysis

The value at risk of log return is 9.95% at 95% confidence level, 14.07% at 99% confidence level, and 18.69% at 99.9% confidence level. The expected loss if it exceeds the VaR value (expected shortfall) is 12.48%, 16.12%, and 20.36% respectively.

The predictions of the GARCH (1, 1) model are listed below.

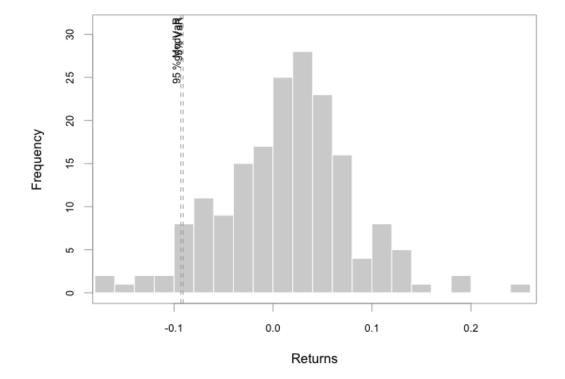
```
> pm1=predict(gfit1,10)
> pm1
  meanForecast meanError standardDeviation
   -0.01448542 0.06287876
                               0.06287876
1
2 -0.01448542 0.06325265
                               0.06325265
3 -0.01448542 0.06357553
                               0.06357553
4 -0.01448542 0.06385457
                               0.06385457
5 -0.01448542 0.06409588
                               0.06409588
6 -0.01448542 0.06430468
                              0.06430468
7 -0.01448542 0.06448544
                               0.06448544
8 -0.01448542 0.06464200
                               0.06464200
9 -0.01448542 0.06477763
                               0.06477763
10 -0.01448542 0.06489517
                               0.06489517
```

> RMeasure(-0.01448542, 0.06287876)

Risk Measures for selected probabilities: prob VaR ES [1,] 0.9500 0.08894094 0.1152154 [2,] 0.9900 0.13179245 0.1530999 [3,] 0.9990 0.17982456 0.1972330 [4,] 0.9999 0.21936173 0.2344189

The VaR is 8.89%, 1.18%, and 17.98% according to the GARCH model. The expected shortfall is 11.52%, 15.31%, and 19.72% respectively.

Figure 9.1 The distribution of returns



h. Conclusion and managerial implications

Microsoft Corporation's stock price has increased in recent years. With the increase in price, volatility hit a new level as well. Microsoft Corporation's stock price is affected by several factors, such as the global financial markets performance.

According to our analysis, the monthly price has seasonal and time trending, the monthly simple return has time trending. ARIMA (2, 1, 7) (0, 1, 1) for price, ARIMA (0, 1, 1) for return. In the multivariate analysis, we used SPY as the second variate. Based on the criteria of AIC, we selected a lag of 10. Then we used Johansen Procedure to test cointegration of two variates. It passed the test. They have an integration effect. In the conditional variance analysis, we tested the ARCH effect of returns, then identified the GARCH model's order as (1, 1).

The VaR of log return is 9.95% at 95% confidence level, 14.07% at 99% confidence level, and 18.69% at 99.9% confidence level. The expected loss if it exceeds the VaR value (expected shortfall) is 12.48%, 16.12%, and 20.36% respectively. Over the GARCH model, the value at risk is 8.89%, 1.18%, and 17.98% respectively. The expected shortfall is 11.52%, 15.31%, and 19.72% respectively.

Using only historical data, the ARIMA model can forecast the future trend of a stock's performance in technical analysis. Due to the ARIMA model's reliance on its own past data, the predictive accuracy is limited. The stock price will either exceed or underperform expectations due to unanticipated surprises or shocks. The VAR model can solve this challenge. In economics, leading indicators can reflect the performance of the entire economy. If we identify a variable that serves as a leading indication of our screened stock price, a fitted VAR model can be used to generate an accurate forecast. The GARCH model is more complex

than the former two models. Unlike the ARIMA and VAR models, it considers the variance of the time series in addition to the mean.

We measured the VaR of Microsoft. It presented a relatively high risk. If a portfolio manager intends to use it to control risk, he must conduct further research on risk management.

Appendix

iGARCH

LogLikelihood: 195.017

```
> ### igarch
> spec = ugarchspec(variance.model=list(model="iGARCH", garchOrder=c(1,1)),
                     mean.model=list(armaOrder=c(1,0), include.mean=TRUE),
                     distribution.model="norm" )
> def.fit = ugarchfit(spec = spec, data = R.nm)
> print(def.fit)
         GARCH Model Fit *
Conditional Variance Dynamics
-----
GARCH Model : iGARCH(1,1)
Mean Model : ARFIMA(1,0,0)
Distribution : norm
Optimal Parameters
      Estimate Std. Error t value Pr(>|t|)
    0.000344 0.004032 0.085362 0.93197
ar1 -0.549474 0.066830 -8.221983 0.00000
omega 0.000000 0.000025 0.003713 0.99704
alpha1 0.016047 0.051396 0.312218 0.75487
beta1 0.983953 NA NA NA
Robust Standard Errors:
      Estimate Std. Error t value Pr(>|t|)
        0.000344 0.005551 0.061997 0.95056
ar1 -0.549474 0.106373 -5.165547 0.00000
omega 0.000000 0.000079 0.001185 0.99906
alpha1 0.016047 0.185478 0.086515 0.93106
beta1 0.983953 NA NA NA
```

Information Criteria

Akaike -2.1343 Bayes -2.0630 Shibata -2.1352 Hannan-Quinn -2.1054

Weighted Ljung-Box Test on Standardized Residuals

| statistic | p-value | Lag[1] | 6.08 | 1.367e-02 | Lag[2*(p+q)+(p+q)-1][2] | 11.52 | 2.248e-11 | Lag[4*(p+q)+(p+q)-1][5] | 18.74 | 9.985e-08

d.o.f=1

H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

statistic p-value
Lag[1] 0.4821 0.4875
Lag[2*(p+q)+(p+q)-1][5] 3.4299 0.3341
Lag[4*(p+q)+(p+q)-1][9] 5.2112 0.3990
d.o.f=2

Weighted ARCH LM Tests

ARCH Lag[3] 0.02122 0.500 2.000 0.8842 ARCH Lag[5] 3.12336 1.440 1.667 0.2724 ARCH Lag[7] 3.67895 2.315 1.543 0.3952

Nyblom stability test

Joint Statistic: 0.5759 Individual Statistics: mu 0.005799 ar1 0.149809

omega 0.278031 alpha1 0.069901

Asymptotic Critical Values (10% 5% 1%) Joint Statistic: 1.07 1.24 1.6 Individual Statistic: 0.35 0.47 0.75

eGARCH

LogLikelihood: 236.9458

```
> ### egarch
> spec = ugarchspec(variance.model=list(model="eGARCH", garchOrder=c(1,1)),
                mean.model=list(armaOrder=c(1,2), include.mean=TRUE),
                distribution.model="norm" )
> def.fit = ugarchfit(spec = spec, data = R.nm)
> print(def.fit)
* GARCH Model Fit *
Conditional Variance Dynamics
_____
GARCH Model : eGARCH(1,1)
Mean Model : ARFIMA(1,0,2)
Distribution : norm
Optimal Parameters
______
      Estimate Std. Error t value Pr(>|t|)
               0.000000 -4890.4
     -0.000209
mu
               0.000162 -6082.5
ar1
    -0.982786
     0.015162 0.000004 4249.8
-0.947236 0.000111 -8507.7
ma1
ma2
omega -0.620030 0.000086 -7189.1
alpha1 -0.194142 0.000035 -5612.9
beta1 0.887408 0.000114 7806.4
                                     0
Robust Standard Errors:
      Estimate Std. Error t value Pr(>|t|)
     -0.000209 0.000023 -8.9596 0
mu
ar1 -0.982786 0.025044 -39.2423
ma1 0.015162 0.001748 8.6740
ma2 -0.947236 0.020675 -45.8162
omega -0.620030 0.005895 -105.1762
alpha1 -0.194142 0.010602 -18.3117
                                    0
beta1 0.887408 0.141157 6.2867
                                     0
gamma1 -0.369327 0.020417 -18.0896
                                     0
```

```
Information Criteria
Akaike -2.5581
Bayes -2.4156
Shibata -2.5618
Hannan-Quinn -2.5003
Weighted Ljung-Box Test on Standardized Residuals
-----
                   statistic p-value
                      2.214 0.136756
Lag[1]
Lag[2*(p+q)+(p+q)-1][8]
                       6.146 0.006935
Lag[4*(p+q)+(p+q)-1][14] 9.017 0.219418
d.o.f=3
H0 : No serial correlation
Weighted Ljung-Box Test on Standardized Squared Residuals
------
        statistic p-value
Lag[1] 1.278 0.25834
Lag[2*(p+q)+(p+q)-1][5] 7.549 0.03809
Lag[4*(p+q)+(p+q)-1][9] 11.286 0.02626
d.o.f=2
Weighted ARCH LM Tests
-----
         Statistic Shape Scale P-Value
ARCH Lag[3] 8.245 0.500 2.000 0.004087
ARCH Lag[5] 9.087 1.440 1.667 0.010961
ARCH Lag[7] 10.876 2.315 1.543 0.011478
Nyblom stability test
-----
Joint Statistic: 2.9156
Individual Statistics:
mu 0.02736
ar1 0.02791
ma1 0.02750
ma2
     0.02846
omega 0.02992
alpha1 0.02745
beta1 0.24594
gamma1 0.02719
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 0.76319 0.4464
Negative Sign Bias 0.07988 0.9364
Positive Sign Bias 0.38798 0.6985
```

Adjusted Pearson Goodness-of-Fit Test:

Joint Effect 0.84789 0.8380

	group	statistic	p-value(g-1)		
1	20	19.10	0.4504		
2	30	26.64	0.5910		
3	40	33.96	0.6987		
4	50	41.39	0.7716		

fGARCH

*		*
*	GARCH Model Fit	*
*		*

Conditional Variance Dynamics

GARCH Model : fGARCH(1,1) fGARCH Sub-Model : TGARCH Mean Model : ARFIMA(1,0,0)

Distribution : norm

Optimal Parameters

	Estimate	Std. Error	t value	Pr(> t)
mu	-0.002528	0.004250	-0.59479	0.551981
ar1	-0.554807	0.069004	-8.04018	0.000000
omega	0.014158	0.012939	1.09423	0.273854
alpha1	0.060933	0.076025	0.80148	0.422855
beta1	0.780943	0.201417	3.87725	0.000106
etal1	0.988981	1 384814	0.71416	0.475127

Robust Standard Errors:

Estimate Std. Error t value Pr(>It1)
mu -0.002528 0.003318 -0.76180 0.446181
ar1 -0.554807 0.054651 -10.15184 0.000000
omega 0.014158 0.013491 1.04943 0.293978
alpha1 0.060933 0.092156 0.66119 0.508489
beta1 0.780943 0.232045 3.36549 0.000764
eta11 0.988981 1.885031 0.52465 0.599827

LogLikelihood : 196.7345

```
Information Criteria
_____
Akaike
          -2.1311
Bayes -2.0243
Shibata -2.1333
Hannan-Quinn -2.0878
Weighted Ljung-Box Test on Standardized Residuals
-----
                     statistic p-value
Lag[1]
                        4.917 2.660e-02
Lag[2*(p+q)+(p+q)-1][2] 10.722 1.945e-10
Lag[4*(p+q)+(p+q)-1][5] 17.813 2.784e-07
d.o.f=1
H0: No serial correlation
Weighted Ljung-Box Test on Standardized Squared Residuals
           statistic p-value
Lag[1] 0.4596 0.4978

Lag[2*(p+q)+(p+q)-1][5] 2.1772 0.5772

Lag[4*(p+q)+(p+q)-1][9] 3.3610 0.6982
d.o.f=2
Weighted ARCH LM Tests
-----
     Statistic Shape Scale P-Value
ARCH Lag[3] 0.1402 0.500 2.000 0.7080
ARCH Lag[5] 1.6315 1.440 1.667 0.5584
ARCH Lag[7] 2.0702 2.315 1.543 0.7025
Nyblom stability test
-----
Joint Statistic: 1.0231
Individual Statistics:
     0.02919
ar1 0.10031
omega 0.48148
alpha1 0.20794
betal 0.47951
eta11 0.24061
Asymptotic Critical Values (10% 5% 1%)
Joint Statistic: 1.49 1.68 2.12
Individual Statistic: 0.35 0.47 0.75
Sign Bias Test
-----
           t-value prob sig
```

Adjusted Pearson Goodness-of-Fit Test:

Negative Sign Bias 0.58426 0.5598 Positive Sign Bias 0.09224 0.9266 Joint Effect 0.36306 0.9478

0.22035 0.8259

_			
	group	statistic	p-value(g-1)
1	20	14.85	0.7318
2	30	19.94	0.8948
3	40	33.96	0.6987
4	50	29.66	0.9869

Sign Bias