ChicagoBooth Honor Code:
I pledge my honor that I have not violated the Honor Code during this examination.

Signature: Name: ID:

Notes:
• Open notes and books.
• For each question, write your answer in the blank space provided.
• Manage your time carefully and answer as many questions as you can.
• The exam has 8 pages and the R output has 9 pages. Total is 17 pages. Please check to make sure that you have all the pages.
• For simplicity, ALL tests use the 5% significance level.
• Round your answer to 3 significant digits.
• This is a 2-hour exam (120 minutes).

Problem A: (30 pts) Answer briefly the following questions. Each question has two points.

1. (Questions 1 to 5) Consider the daily log returns of Qualcomm stock from May 17, 2007 to April 30, 2013. Summary statistics of the log returns are given in the R output. Is the mean log return significantly different from zero? State the null and alternative hypotheses and draw the conclusion.

2. Does the log returns have heavy tails? Perform a test and draw the conclusion.
3. What is the standard error (SE) of the mean log returns? What is the 
  \( t \)-ratio \( (t = \frac{\bar{r}}{\text{std}(\bar{r})}) \) of the sample mean \( \bar{r} \) of the log return \( r_t \)?

4. The sample PACF of the log returns shows some minor correlations 
  at lags 1 and 2 so that an AR(2) model is employed. Write down the 
  fitted AR(2) model, including the residual variance.

5. Based on the output provided, does the log returns have conditional 
  heteroscedasticity (ARCH-effect)? Why?

6. Suppose that the log return \( r_t \) of an asset is normally distributed with 
  mean 0.02 and standard error 0.4. What is the mean of the simple 
  return of the asset?

7. Describe two methods for identifying the order of an AR time series.

8. Consider model comparison. Give one method for in-sample compari-
   son and one method for out-sample comparison.

9. (Questions 9 to 10) Suppose the daily log return \( r_t \) of Stock A 
   follows the model \( r_t = 0.002 + a_t, \ a_t = \sigma_t \epsilon_t \), where \( \{ \epsilon_t \} \) is an in-
   dependent and identically distributed (iid) sequence of standardized 
   Student-\( t \) distribution with 5 degrees of freedom. In addition, \( \sigma_t^2 = 
   0.01 + 0.1a_{t-1}^2 + 0.9\sigma_{t-1}^2 \). Let \( h = 100 \) be the forecast origin with 
   \( a_h = -0.015 \) and \( \sigma_h = 0.2 \). Calculate the 1-step ahead prediction \( r_h(1) \) 
   and 1-step ahead volatility forecast.
10. Calculate the 5-step ahead prediction $r_h(5)$ and the 5-step ahead volatility forecast at the forecast origin $h$.

11. Consider a linear regression model. Describe a situation under which the conventional $R^2$ measure is not informative? Also, why is the commonly used Durbin-Watson statistic not sufficient in detecting serial correlations of time-series data?

12. Let $r_t$ be a univariate time series. Consider the following two stationary AR(2) models: (a) $r_t = \phi_0 + \phi_2 r_{t-2} + a_t$ and (b) $r_t = \phi_0 + \phi_{21} r_{t-1} + \phi_{22} r_{t-2} + a_t$, where $a_t$ is a sequence of iid random variables with mean zero and variance $\sigma_a^2$. Why is $|\phi_2| < 1$ for model (a)? Is it possible that $\phi_{21} > 1$ in model (b)? Why?

13. Consider the univariate time series model $(1 - 0.9B + 0.2B^2)r_t = 100 + (1 - 1.1B)a_t$, where $a_t$ is a sequence of iid random variables with mean zero and variance $\sigma_a^2$. Is the model stationary? Why? Is the model invertible? Why?

14. Give two univariate volatility models what can handle the leverage effect in asset returns?

15. Consider the simple model $r_t = 0.02 + a_t - 1.1a_{t-1} + 0.3a_{t-2}$, where $a_t$ is defined in the prior question. Assume that $\sigma_a^2 = 1$, $a_{100} = -0.02$ and $a_{99} = 0.01$. Calculate the 1-step and 3-step ahead point forecasts of $r_t$ at the forecast origin $t = 100$. 

Problem B. (18 pts) Consider, again, the daily log returns of Qualcomm stock from May 17, 2007 to April 30, 2013. Several volatility models are fitted to the data and the relevant R output is attached. Answer the following questions.

1. (3 points) A volatility model, called m2 in R, is entertained. Write down the fitted model, including the mean equation. Is the model adequate? Why?

2. (3 points) Another volatility model, called m3 in R, is fitted to the returns. Write down the model, including all estimated parameters.

3. (3 points) A third model, called m4 in R, is also entertained. Write down the model, including the distributional parameters.

4. (3 points) Let $\xi$ be the skew parameter in model m4. Does the estimate of $\xi$ confirm that the distribution of the log returns is skewed? Why? Perform the test to support your answer.

5. (2 points) Compare the three models m2, m3, m4. Which model is preferred? Why?

6. (2 points) To facilitate Value at Risk calculation via the RiskMetrics approach, an IGARCH(1,1) model is fitted to the daily log returns. Write down the fitted IGARCH(1,1) model.

7. (2 points) Based on the output provided at $t = 1499$. Calculate the 1-step ahead volatility forecast using the IGARCH(1,1) model.
Problem C. (17 points) Consider the monthly log return of 3M stock from January 1961 to December 2012. Use the attached R output to answer the following questions. Let $r_t$ denote the monthly log return.

1. (2 points) A simple GARCH(1,1) model is fitted to the data, called m1. Is the model adequate? If not, describe a method to improve the model.

2. (2 points) A GARCH(1,1) model is fitted to the data with Student $t$ innovations, called m2. Is the model adequate? Why?

3. (2 points) Based on the model m2, compute 95% 6-step ahead interval forecast for the 3M stock return at the forecast origin December 2012 (last data point).

4. (3 points) To study the leverage effect, a TGARCH or GJR-type of model is entertained using the APARCH model in R. Write down the fitted model, called m3.

5. (3 points) Based on the fitted model m3, is the leverage effect significant? State the null and alternative hypotheses, obtain the test statistic, and draw the conclusion.

6. (3 points) A GARCH-M model is fitted to the percentage log returns, i.e. $x_t = r_t \times 100$. Write down the fitted model.

7. (2 points) Is the risk premium significantly different from zero? Why?
Problem D. (20 points) Consider the monthly unemployment rate, seasonally adjusted, of California from January 1976 to March 2013. Since the series is close to having a unit root, we focus on the first differenced series denoted by $dca$ in the attached R analysis. To help predict the series, we use the lag-1 monthly unemployment rate of the U.S. as an explanatory variable. Denote the first differenced U.S. unemployment rate of $dus$. Use the attached output to answer the following questions:

1. (3 points) A preliminary examination of the ACF and PACF suggests an ARIMA(3,0,1) model for the $dca$ series. Write down the fitted model, including the residual variance.

2. (2 points) The fitted ARIMA(3,0,1) model does not contain a constant. Why?

3. (3 points) Model checking shows that the lag-6 ACF of the residuals of the fitted ARIMA(3,0,1) model is significant. This led to an ARIMA(3,0,6) model, which contains several insignificant MA coefficients. Write down the refined ARIMA(3,0,6) model all estimates of which are statistically significant.

4. (2 points) Provide a justification that it is ok to remove the insignificant parameters in the ARIMA(3,0,6) model. i.e., compare the full model (with all coefficients) with the refined model.

5. (2 points) Does the fitted ARIMA(3,0,6) model for the monthly California unemployment rate imply the existence of business cycles? Why?

6. (3 points) Let $y$ be the $dca$ series and $x$ be the lag-1 $dus$ series. Write down the regression model with time series error between $y$ and $x$. 
7. (2 points) Based on the regression model with time series errors, is the lag-1 U.S. unemployment rate helpful in predicting the California unemployment rate? Why?

8. (3 points) Compare the ARIMA(3,0,6) model and the regression model with time series errors. Which model is preferred? Why?

**Problem E.** (15 points) Consider the quarterly earnings per share of FedEx starting from the second quarter of 1992 to the fourth quarter of 2006. We analyze the log earnings per share, denoted by $y_t$. Answer the following questions.

1. (3 points) Write down the fitted time series model $m_1$ for the $y_t$ series, including the residual variance.

2. (2 points) Is the model adequate? Why?

3. (2 points) Write down the fitted time series model $m_2$ for $y_t$, including the residual variance.

4. (3 points) Let $\theta_3$ denote the coefficient of lag-3 of MA polynomial. Test $H_0 : \theta_3 = 0$ versus $H_a : \theta_3 \neq 0$. Calculate the test statistic and draw the conclusion.
5. (3 points) Write down the fitted time series model $m3$ for $y_t$, including the residual variance.

6. (2 points) Among the three models $m1$, $m2$, $m3$, which model is preferred? Why?
```r
# Part A

```r
dim(QCOM)
[1] 1500 6
```r
rt=diff(log(as.numeric(QCOM$QCOM.Adjusted)))
```r
require(fBasics)
```r
basicStats(rt)
rt
nobs 1499.000000
NAs 0.000000
Mean 0.000272
SE Mean ??
LCL Mean -0.000834
UCL Mean 0.001377
Variance 0.000476
Stdev 0.021824
Skewness -0.048756
Kurtosis 7.179616
```r
acf(rt); pacf(rt) ## plots are not shown.
```r
m1=arima(rt,order=c(2,0,0))
m1
Call: arima(x = rt, order = c(2, 0, 0))
```r
Coefficients:
```
         ar1     ar2  intercept
     -0.0830 -0.0708  3e-04
```
s.e. 0.0258 0.0257 5e-04
```r
sigma^2 estimated as 0.0004707:  log likelihood=3615.1, aic=-7222.21
```r
Box.test(m1$residuals^2,lag=10,type="Ljung")
```r
Box-Ljung test
```r
data: m1$residuals^2
X-squared = 161.0606, df = 10, p-value < 2.2e-16
```r

# Part B

```r
require(fGarch)
m2=garchFit(~arma(2,0)+garch(1,1),data=rt,trace=F)
```r
summary(m2)
```r
Title: GARCH Modelling
Call:
garchFit(formula = ~arma(2, 0) + garch(1, 1), data = rt, trace = F)
```r
Mean and Variance Equation:
```
data ~ arma(2, 0) + garch(1, 1)  [data = rt]
```r
Conditional Distribution: norm
Error Analysis:

|     | Estimate  | Std. Error | t value | Pr(>|t|) |
|-----|-----------|------------|---------|----------|
| mu  | 1.129e-03 | 4.794e-04  | 2.355   | 0.01852 *|
| ar1 | -2.668e-02| 2.869e-02  | -0.930  | 0.35235  |
| ar2 | -8.053e-02| 2.832e-02  | -2.843  | 0.00447 **|
| omega| 1.390e-05 | 3.543e-06  | 3.923   | 8.74e-05 ***|
| alpha1 | 8.041e-02 | 1.727e-02  | 4.656   | 3.22e-06 ***|
| beta1 | 8.917e-01 | 2.061e-02  | 43.259  | < 2e-16 ***|

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera</td>
<td>R Chi^2</td>
<td>4671.285</td>
</tr>
<tr>
<td>Shapiro-Wilk</td>
<td>R W</td>
<td>0.9435465</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>R Q(10)</td>
<td>10.49343</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>R Q(20)</td>
<td>14.87801</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>R^2 Q(10)</td>
<td>1.992025</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>R^2 Q(20)</td>
<td>3.762941</td>
</tr>
</tbody>
</table>

Information Criterion Statistics:

AIC  | BIC  | SIC  | HQIC |
-----|------|------|------|
-5.002726 | -4.981461 | -5.002758 | -4.994804 |

> m3=garchFit(~arma(2,0)+garch(1,1),data=rt,trace=F,cond.dist="std")
> summary(m3)

Title: GARCH Modelling

Call:
garchFit(formula = ~arma(2, 0) + garch(1, 1), data = rt, cond.dist = "std",
          trace = F)

Mean and Variance Equation:
data ~ arma(2, 0) + garch(1, 1) [data = rt]

Conditional Distribution: std

Error Analysis:

|     | Estimate  | Std. Error | t value | Pr(>|t|) |
|-----|-----------|------------|---------|----------|
| mu  | 7.932e-04 | 4.102e-04  | 1.934   | 0.0532 . |
| ar1 | -2.555e-02| 2.588e-02  | -0.987  | 0.3235  |
| ar2 | -5.005e-02| 2.538e-02  | -1.972  | 0.0486 *|
| omega| 5.722e-06 | 2.629e-06  | 2.176   | 0.0295 *|
| alpha1 | 7.306e-02 | 1.843e-02  | 3.964   | 7.37e-05 ***|
| beta1 | 9.171e-01 | 1.981e-02  | 46.291  | < 2e-16 ***|
| shape | 4.992e+00 | 6.054e-01  | 8.245   | 2.22e-16 ***|

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box</td>
<td>R Q(10)</td>
<td>9.027788</td>
</tr>
</tbody>
</table>
Ljung-Box Test R Q(20) 13.22441 0.8675492
Ljung-Box Test R^2 Q(10) 1.897581 0.9970665
Ljung-Box Test R^2 Q(20) 3.406239 0.9999878

Information Criterion Statistics:
<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>SIC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>-5.139432</td>
<td>-5.114624</td>
<td>-5.139475</td>
<td>-5.130190</td>
</tr>
</tbody>
</table>

> plot(m3)
> m4=garchFit(~arma(2,0)+garch(1,1),data=rt,trace=F,cond.dist="sstd")
> summary(m4)
Title: GARCH Modelling
Call:
garchFit(formula = ~arma(2, 0) + garch(1, 1), data = rt, cond.dist = "sstd",
trace = F)
Mean and Variance Equation:
data ~ arma(2, 0) + garch(1, 1)  [data = rt]

Conditional Distribution: sstd

Error Analysis:

| Estimate  | Std. Error  | t value  | Pr(>|t|) |
|-----------|-------------|----------|----------|
| mu        | 1.014e-03   | 4.408e-04| 2.300    | 0.0215 * |
| ar1       | -2.100e-02  | 2.593e-02| -0.810   | 0.4180   |
| ar2       | -4.670e-02  | 2.542e-02| -1.837   | 0.0662 . |
| omega     | 5.691e-06   | 2.594e-06| 2.194    | 0.0283 * |
| alpha1    | 7.405e-02   | 1.843e-02| 4.018    | 5.87e-05 ***|
| beta1     | 9.170e-01   | 1.937e-02| 47.345   | < 2e-16 ***|
| skew      | 1.050e+00   | 3.836e-02| 27.376   | < 2e-16 ***|
| shape     | 4.912e+00   | 5.943e-01| 8.265    | 2.22e-16 ***|

> source("Igarch.R")
> m5=Igarch(rt)
Estimates: 0.9772298
Maximized log-likehood: -3712.597

Coefficient(s):

| Estimate  | Std. Error  | t value  | Pr(>|t|) |
|-----------|-------------|----------|----------|
| beta      | 0.97722984  | 0.00744395| 131.278  | < 2.22e-16 ***|

> rt[1499]
[1] -0.0001623245
> m5$volatility[1499]
[1] 0.01421029

******* Part C ***********************************************
> da=read.table("m-3m2dx-6112.txt",header=T)
> head(da)
  PERMNO  date       mmm  vwretd   sprtrn
 1 22592 19610131  0.013536  0.063879  0.063156
 6 22592 19610630 -0.012251 -0.028499 -0.028846
> rt=log(da$mmm+1) ### Compute log return
> m1=garchFit(~garch(1,1),data=rt,trace=F)
> summary(m1)
Title: GARCH Modelling
Call:
garchFit(formula = ~garch(1, 1), data = rt, trace = F)

Mean and Variance Equation:
data ~ garch(1, 1) [data = rt]

Conditional Distribution: norm

Error Analysis:

|     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----|----------|------------|---------|----------|
| mu  | 0.0073972| 0.0023007  | 3.215   | 0.0013   **|
| omega| 0.0007718| 0.0003182  | 2.425   | 0.0153   *|
| alpha| 0.0965805| 0.0344871  | 2.800   | 0.0051   **|
| beta 1| 0.6948362| 0.1061987  | 6.543   | 6.04e-11 ***|

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera Test</td>
<td>R Chi^2</td>
<td>102.5758</td>
</tr>
<tr>
<td>Shapiro-Wilk Test</td>
<td>R W</td>
<td>0.9852546</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(10)</td>
<td>10.93472</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(20)</td>
<td>22.0489</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R^2 Q(10)</td>
<td>3.910077</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R^2 Q(20)</td>
<td>14.99407</td>
</tr>
</tbody>
</table>

Information Criterion Statistics:

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>SIC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.779438</td>
<td>-2.751001</td>
<td>-2.779519</td>
<td>-2.768387</td>
</tr>
</tbody>
</table>

> m2=garchFit(~garch(1,1),data=rt,trace=F,cond.dist="std")
> summary(m2)
Title: GARCH Modelling
Call:
garchFit(formula = ~garch(1, 1), data = rt, trace = F, cond.dist="std")

Mean and Variance Equation:
data ~ garch(1, 1) [data = rt]

Conditional Distribution: std
Error Analysis:

|   | Estimate  | Std. Error | t value | Pr(>|t|) |
|---|-----------|------------|---------|----------|
| mu| 0.0080579 | 0.0022168  | 3.635   | 0.000278 *** |
| omega| 0.0006094 | 0.0003576  | 1.705   | 0.088287 .  |
| alpha1 | 0.0892538 | 0.0387345  | 2.304   | 0.021209 *  |
| beta1 | 0.7455437 | 0.1235584  | 6.034   | 1.6e-09 *** |
| shape | 7.7557612 | 2.2044862  | 3.518   | 0.000435 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box Test R Q(10)</td>
<td>10.90092</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>21.95176</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>3.885065</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>14.57763</td>
</tr>
</tbody>
</table>

Information Criterion Statistics:

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>SIC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.813309</td>
<td>-2.777762</td>
<td>-2.813436</td>
<td>-2.799496</td>
</tr>
</tbody>
</table>

> plot(m2)
> predict(m2,6)

    | meanForecast | meanError | standardDeviation |
|---|--------------|-----------|-------------------|
| 1 | 0.008057917  | 0.05320784| 0.05320784        |
| 6 | 0.008057917  | 0.05780348| 0.05780348        |

> m3=garchFit(~aparch(1,1),data=rt,trace=F,delta=2,include.delta=F,cond.dist="std")
> summary(m3)

Title: GARCH Modelling
Call:
garchFit(formula = ~aparch(1, 1), data = rt, delta = 2, cond.dist = "std",
include.delta = F, trace = F)

Mean and Variance Equation:
data ~ aparch(1, 1) [data = rt]

Conditional Distribution: std

Error Analysis:

|   | Estimate  | Std. Error | t value | Pr(>|t|) |
|---|-----------|------------|---------|----------|
| mu| 0.0070961 | 0.0022293  | 3.183   | 0.001457 ** |
| omega| 0.0006575 | 0.0003185  | 2.065   | 0.038950 *  |
| alpha1 | 0.0397115 | 0.0569406  | 0.697   | 0.485541   |
| gamma1 | 1.0000000 | 1.3814017  | 0.724   | 0.469126   |
| beta1 | 0.7403939 | 0.1082974  | 6.837   | 8.11e-12 *** |
| shape | 8.1078458 | 2.3149477  | 3.502   | 0.000461 *** |

> source("garchM.R")
> rt=rt*100
> m4=garchM(rt)
Maximized log-likehood: 2002.437
Coefficient(s):
   Estimate Std. Error t value Pr(>|t|)  
mu 0.7393842 0.2303481 3.20986 0.0013280 **
gamma 0.0156864 0.1637841 0.09578 0.9236992
omega 7.7189315 3.1838789 2.42438 0.0153346 *
alpha 0.0967951 0.0345582 2.80093 0.0050955 **
beta 0.6946563 0.1063378 6.53255 6.4661e-11 ***
---

#### Part D ##########################################
> da=read.table("m-unemp-ca.txt",header=T)
> ca=da$VALUE
> da1=read.table("m-unrate76.txt",header=T)
> us=da1$rate

> dca=diff(ca) ### Difference of the California ratee
> acf(dca); pacf(dca)
> t.test(dca)
One Sample t-test
data: dca
 t = 0.035, df = 445, p-value = 0.9721

> m1=arima(dca,order=c(3,0,1),include.mean=F)
> m1
Call: arima(x = dca, order = c(3, 0, 1), include.mean = F)
Coefficients:
ar1  ar2  ar3  ma1
 1.0061 0.2504 -0.3275 -0.4535
s.e. 0.1368 0.0977 0.0611 0.1378
sigma^2 estimated as 0.004629: log likelihood=565.05, aic=-1120.11

> m2=arima(dca,order=c(3,0,6),include.mean=F)
> m2
Call: arima(x = dca, order = c(3, 0, 6), include.mean = F)
Coefficients:
ar1  ar2  ar3  ma1  ma2  ma3  ma4  ma5
 0.4385 0.7998 -0.3328 0.0772 -0.2012 0.3098 -0.1192 -0.0621
s.e. 0.2518 0.1787 0.2163 0.2496 0.1934 0.0882 0.1111 0.0586
ma6
-0.1593
s.e. 0.0514
sigma^2 estimated as 0.004403: log likelihood = 576.03, aic = -1132.07

> f1=c(NA,NA,NA,0,0,NA,0,0,NA)
> m3=arima(dca,order=c(3,0,6),include.mean=F,fixed=f1)
> m3
Call: arima(x = dca, order = c(3, 0, 6), include.mean = F, fixed = f1)

Coefficients:
           ar1     ar2     ar3    ma1    ma2     ma3    ma4   ma5    ma6
  0.5186  0.5504 -0.2086   0.00   0.2484   0.00 -0.1337

s.e.  0.0469  0.0486  0.0549   0.00   0.0619   0.00  0.0499

sigma^2 estimated as 0.004432:  log likelihood = 574.61,  aic = -1137.23
> tsdiag(m3,gof=24) ### The model fares well
> p1=c(1,-m3$coef[1:3])
> roots=polyroot(p1)
> roots
[1] 1.142611+0i -1.432688-0i 2.928975+0i
>
> dus=diff(us) ### Difference of the US unemployment rate
> length(dca)
[1] 446
> y=dca[2:446]; x=dus[1:445]
> m4=lm(y~-1+x)
> summary(m4)

Call: lm(formula = y ~ -1 + x)

Coefficients:           Estimate Std. Error t value Pr(>|t|)
  x         0.4142     0.0314  13.19  <2e-16 ***
---
Residual standard error: 0.1144 on 444 degrees of freedom
Multiple R-squared: 0.2816,  Adjusted R-squared: 0.2799

> Box.test(m4$residuals,lag=12,type="Ljung")

Box-Ljung test

data:  m4$residuals
X-squared = 631.3505,  df = 12,  p-value < 2.2e-16

> m5=ar(m4$residuals,method="mle")
> m5$order
[1] 4
> m6=arima(dca,order=c(4,0,0),include.mean=F,xreg=dus)
> m6
Call: arima(x = dca, order = c(4, 0, 0), xreg = dus, include.mean = F)

Coefficients:
            ar1     ar2     ar3     ar4     dus
  0.4867    0.5564    0.0079 -0.1849  0.0708

s.e.  0.0486  0.0522  0.0537  0.0471  0.0173
sigma^2 estimated as 0.004394: log likelihood = 576.66, aic = -1141.33

```r
> tsdiag(m6,gof=24); source("backtest.R")
> backtest(m3,dca,400,1, fixed=f1, inc.mean=F)
[1] "RMSE of out-of-sample forecasts"
[1] 0.08590367
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.06957228
> backtest(m6,y,399,1,xre=dus)
[1] "RMSE of out-of-sample forecasts"
[1] 0.08242214
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.06562168
>
#### Part E #########################################
```
```r
> da=read.table("q-earn-fdx.txt",header=T)
> head(da)
   day mon year earnings
 1   14    7 1992      0.17
 2   16    9 1992      0.05
> xt=da[,4]; plot(xt,type='l')
> yt=log(xt); plot(yt,type='l')
> acf(yt); acf(diff(yt)); acf(diff(diff(yt),4))
> m1=arima(yt,order=c(0,1,2),seasonal=list(order=c(0,1,1),period=4))
> m1
Call: arima(x = yt, order = c(0, 1, 2), seasonal = list(order = c(0, 1, 1),
period = 4))

Coefficients:
  ma1    ma2    sma1
-0.7054  0.4232 -0.5754
s.e. 0.1302  0.1208  0.1289

sigma^2 estimated as 0.06357: log likelihood = -3.23, aic = 14.45

> Box.test(m1$residuals,lag=12,type='Ljung')
Box-Ljung test
data: m1$residuals
X-squared = 15.29, df = 12, p-value = 0.226

> m2=arima(yt,order=c(0,1,3),seasonal=list(order=c(0,1,1),period=4))
> m2
Call: arima(x = yt, order = c(0, 1, 3), seasonal = list(order = c(0, 1, 1),
period = 4))

Coefficients:
  ma1    ma2    ma3    sma1
-0.7841  0.5975 -0.3977 -0.3761
s.e. 0.1341  0.1601  0.2064  0.2110

```
\[
\sigma^2 \text{ estimated as } 0.05821: \ \log \text{ likelihood} = -1.01, \ \text{aic} = 12.02
\]

\[
> \text{Box.test(m2$residuals,lag=12,type='Ljung')}
\]

Box-Ljung test

data: m2$residuals
X-squared = 7.8647, df = 12, p-value = 0.7956

\[
> \text{m3=arima(yt,order=c(1,1,0),seasonal=list(order=c(0,1,1),period=4))}
\]

\[
> \text{m3}
\]

Call:
arima(x=yt,order=c(1,1,0), seasonal=list(order=c(0,1,1),period = 4))

Coefficients:

\[
\begin{array}{ll}
ar1 & \text{sma1} \\
-0.7572 & -0.5966 \\
\text{s.e.} & 0.0999 0.1156 \\
\end{array}
\]

\[
\sigma^2 \text{ estimated as } 0.05912: \ \log \text{ likelihood} = -1.35, \ \text{aic} = 8.7
\]

\[
> \text{Box.test(m3$residuals,lag=12,type='Ljung')}
\]

Box-Ljung test

data: m3$residuals
X-squared = 9.8561, df = 12, p-value = 0.6286