Midterm

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. Exam time: 120 minutes.
• You may use a calculator or a PC. However, turn off Internet connection and cell phones. Internet access and phone communication are strictly prohibited during the exam.
• The exam has 9 pages and the R output also has 9 pages. Please check to make sure that you have all 18 pages.
• For each question, write your answer in the blank space provided.
• Manage your time carefully and answer as many questions as you can.
• For simplicity, if not specifically given, use 5% Type-I error in hypothesis testings.
• Round your answer to 3 significant digits.

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

1. Give two situations under which serial correlations exist in observed asset returns even though the true underlying returns are serially uncorrelated.

2. (Questions 2 to 8): Consider the daily S&P 500 index for a certain period. Some analysis is attached. Let $r_t$ be the daily log return of the index. Is the expected mean return $E(r_t)$ zero? Why?
3. Does the daily log return of the S&P index follow a skew distribution? Why?

4. Does the daily log return of the S&P index have heavy tails? Why?

5. The sample ACF of \( r_t \), namely \( \hat{\rho}_0, \hat{\rho}_2, \ldots, \hat{\rho}_9 \), are given. Test the null hypothesis \( H_0 : \rho_1 = 0 \) versus the alternative hypothesis \( H_a : \rho_1 \neq 0 \), where \( \rho_1 \) is lag-1 ACF of \( r_t \). Compute the test statistic and draw the conclusion.

6. Turn to the daily log index \( p_t \). A model is fitted, called m2, in the R output. Write down the fitted model, including residual variance.

7. Use the fitted model m2 to forecast the log index at the forecast origin \( T = 1336 \). What is the 1-step ahead point forecast? Obtain a 95% interval forecast for \( p_{1337} \).

8. What is the 2-step ahead point forecast of \( p_t \) at the forecast origin \( T = 1336 \)? Use the model to derive the forecast.

9. Let \( R_t \) and \( r_t \) be the daily simple and log return, respectively, of an asset. What is the relationship between \( R_t \) and \( r_t \)? Suppose further that \( r_t \) follows a normal distribution with mean 0.05 and variance 0.04. What is the expected value of \( R_t \) for the asset?
10. Consider the monthly log return, in percentages, of the Decile 8 portfolio of Center for Research in Security Prices (CRSP) from January 1961 to December 2013 for 636 observations. A GARCH-M model is fitted to the series. Write down the fitted model.

11. Consider again the monthly log returns, in percentages, of Decile 8 portfolio. Is the risk premium statistically significant at the 5% level? Why?

12. (For questions 12-15). Consider the growth rates of the real quarterly gross domestic product (GDP) of Canada from the second quarter of 1980 to the second quarter of 2011 for 125 data points. Figure 1 shows the PACF of the GDP growth rates. Specify two possible AR models for the growth rate series and briefly justify your choices.

13. The order selection via AIC is also given. The criterion selects an AR(4) model. An AR(4) model is estimated. Write down the fitted model, including the residual variance.

14. Consider the fitted AR(4) model. Does it imply the existence of business cycles in the Canadian economy? Why?

15. If business cycles exist, compute the periods of all possible cycles.
Problem B. (27 pts) Consider the daily log returns of Apple stock starting from January 3, 2004 for 2517 observations. Let \( r_t \) be the log return series. Based on the attached R output, answer the following questions.

1. (3 points) What is the mean equation for \( r_t \)? Why?

2. (2 points) Is there any ARCH effect in \( r_t \)? Why?

3. (2 points) A simple volatility model, called \( m_1 \) in R, is entertained for \( r_t \). Is Model \( m_1 \) adequate for the log return series? Why?

4. (3 points) A refined model, called \( m_2 \) in R, is fitted. Write down the model, including the distribution of the innovations.

5. (3 points) Let \( \xi \) be the skew parameter in Model \( m_3 \). Based on the model, is the distribution of \( r_t \) skew? Perform a statistical test to support your answer.

6. (2 points) Compare the three models \( m_1, m_2, m_3 \). Which model is preferred? Why?

7. (3 points) To estimate the potential leverage effect in \( r_t \), we consider an APARCH(1,1) model with \( \delta = 2.0 \). Write down the fitted model, including the innovation distribution.
8. (4 points) The average volatility of $r_t$ via the APARCH model is 0.02248 and the approximate 99th quantile of $r_t$ is 0.061758 resulting in an $a_t = 0.06$. To see the impact of leverage effect, (a) compute the volatility $\sigma_t$ if $a_{t-1} = 0.06$ and $\sigma_{t-1} = 0.02248$, (b) compute the volatility $\sigma_t$ if $a_{t-1} = -0.06$ and $\sigma_{t-1} = 0.02248$, and finally, (c) compute volatility ratio $[(b)/(a)]$.

9. (2 points) An IGARCH model with normal innovations is also fitted for the Apple log return $r_t$. Write down the fitted model.

10. (3 points) Using the fitted IGARCH(1,1) model and the information provided, compute the volatility $\sigma_{2518}$ for the Apple log return.

**Problem C.** (17 points) Consider the monthly log return of Decile 1 portfolio of CRSP from January 1961 to December 2013 for 636 observations. Let $d1_t$ denote the monthly log return. Several volatility models were fitted. Use the attached R output to answer the following questions.

1. (4 points) Both the GARCH(1,1) model with Gaussian innovations, $g1$, and the GARCH(1,1) model with Student-$t$ innovations, $g2$, were rejected based on model checking. A refined model, called $g3$, is entertained. Write down the fitted model, including the mean equation and the innovation distribution.
2. (2 points) Based on the fitted model $g_3$, is the distribution of the log returns skew? Why?

3. (3 points) Based on the model $g_3$, compute a 95% 4-step ahead interval forecast for the log return of Decile 1 portfolio at the forecast origin December 2013.

4. (3 points) To study the leverage effect, a TGARCH or GJR-type of model is entertained. Denote the model by $g_4$. Based on the model, is the leverage effect significant? State the null and alternative hypotheses, obtain the test statistic, and draw the conclusion.

5. (3 points) Based on the model $g_4$, compute a 95% 4-step ahead interval forecast for the log return of Decile 1 portfolio at the forecast origin December 2013.

6. (2 points) Compare the two 95% interval forecasts. Briefly state the impact of leverage effect?
Problem D. (14 points) Consider the monthly U.S. heating oil price and the natural gas price from November 1993 to August 2012. Use the attached R output to answer the following questions:

1. (2 points) Focus on the logarithm of the heating oil price. Preliminary analysis shows that the log price has a unit root so that the growth rate is used in model specification. The AIC selects an AR(1) for the growth rate. Therefore, an ARIMA(1,1,0) model is entertained for the log heating price. Write down the fitted model, including residual variance.

2. (2 points) Since the fitted AR(1) coefficient is not large, we also entertained an exponential smoothing model. Write down the fitted model, including the residual variance.

3. (2 points) Model checking shows that the prior two models fit the data reasonably well. Based on in-sample fit, which model is preferred? Why?

4. (2 points) The two models were used in out-of-sample forecasting. Based on the out-of-sample performance, which model is preferred? Why?

5. (3 points) Next, to make use of the information in the natural gas price, we consider a simple linear regression between the log heating oil price and log natural gas price. The residuals of the regression model shows strong serial correlations. To avoid spurious regression, let $y_t$ and $x_t$ be the growth rate of heating oil price and natural gas price, respectively. White down the simple linear regression for the two growth rate series. What is the $R^2$ of the model?
6. (3 points) The residuals of the prior simple linear regression contains significant lag-1 serial correlation so that a regression model with time series errors is fitted. Write down the fitted model.

**Problem E.** (12 points) Consider the quarterly earnings per share of Procter & Gamble from 1983.II to 2012.III. Figure 2 shows the time plot of the earnings. From the plot, there was a negative earnings in the 80s and two large jumps occurred around 2010. For simplicity, we analyze the earnings $x_t$ directly. Sample autocorrelations of differenced data suggest the Airline model.

1. (2 points) Write down the fitted time series model $m_1$ for the $x_t$ series, including the residual variance.

2. (2 points) The fitted model show a large outlier at $t = 104$. Define an indicator variable for this particular data point.

3. (2 points) As a matter of fact, there are several outliers. The model $m_4$ contains three large outliers. Model checking shows that the ACF of the residuals has a significant correlation at lag 3 so that a refined model is entertained. The resulting model is denoted by $m_5$. Is the lag-3 MA coefficient $\theta_3$ of Model $m_5$ significantly different from zero? Why?

4. (6 points) Finally, an additional outlier is found and an insignificant parameter is also detected. The final model for $x_t$ is Model $m_7$. Write down the fitted model, including residual variance.
Figure 1: The sample partial autocorrelation function of the quarterly growth rates of Canadian gross domestic product from 1980.II to 2011.II.

Figure 2: Quarterly earnings per share of Procter & Gamble stock from 1983.II to 2012.III
### Problem A

```r
gtSymbos("^GSPC", from="XXXX", to="XXXX")
sp = log(as.numeric(GSPC[,6]))
rtn = diff(sp)
require(fBasics)
basicStats(rtn)
```

```
  nobs        1335.000000
  Minimum      -0.068958
  Maximum       0.068366
  Mean         0.000523
  SE Mean       0.000329
  LCL Mean     -0.000123
  UCL Mean      0.001169
  Variance     0.000145
  Stdev        0.012032
  Skewness     -0.281485
  Kurtosis      4.239532
```

```r
m1 = acf(rtn)
m2 = arima(sp, order = c(0, 1, 1))
```

```
  Call: arima(x = sp, order = c(0, 1, 1))
  Coefficients:
     ma1
    -0.0788
  s.e.   0.0265

  sigma^2 estimated as 0.000144:  log likelihood = 4010.25,  aic = -8016.49
```

```r
m2$residuals[,1336]
```

```
[1] -0.008005722
```

```
# Decile 8

idx = c(1:6360)[da[,2] == 8]
d8 = log(da[idx, 3] + 1)
plot(d8, type = 'l')
source("garchM.R")
d8 = d8 * 100
g5 = garchM(d8)
Maximized log-likehood:  -2064.533
```

Coefficient(s):
Estimate               Std. Error         t value     Pr(>|t|)
mu    -1.2634761           1.1378947          -1.11036     0.266843
gamma  0.0625708           0.0289221           2.16342     0.030509  *
omega  4.0752975           1.6008262           2.54575     0.010904  *
alpha  0.0700690           0.0263922           2.65491     0.007933  **
beta   0.8288135           0.0534552          15.50483    < 2e-16  ***

########### Canadian GDP ###
> dim(qgdp)
[1] 126 5
> y=log(qgdp[,3:5])
> head(y)
     uk     ca     us
1 12.05778 13.34518 15.59190
...
6 12.02211 13.38773 15.60021
> ca=diff(y$ca)
> pacf(ca) ### See Figure 1 of the exam.
> m0=ar(ca,method="mle")
> m0$order
[1] 4
> m1=arima(ca,order=c(4,0,0))
> m1
Call: arima(x = ca, order = c(4, 0, 0))
Coefficients:
                ar1          ar2          ar3          ar4      intercept
       0.6485034   -0.1757386   0.2334611   -0.2067754    0.0060327
              s.e.  0.0880205    0.1036993    0.1032314    0.0899141    0.0010903
sigma^2 estimated as 3.898e-05: log likelihood = 456.85, aic = -901.

########## Problem B ####
> aapl=log(da$rtn+1)
> t.test(aapl)

One Sample t-test
data:  aapl
t = 3.4145, df = 2516, p-value = 0.0006491
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  0.0006756272 0.0024984813

> Box.test(aapl,lag=15,type='Ljung')

Box-Ljung test
data:  aapl
X-squared = 22.8513, df = 15, p-value = 0.08735

> at=aapl-mean(aapl)
> Box.test(at^2,lag=10,type='Ljung')
Box-Ljung test

data: at^2
X-squared = 432.3742, df = 10, p-value < 2.2e-16

> m1=garchFit(~garch(1,1),data=aapl,trace=F)
> summary(m1)

Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 1), data = aapl, trace = F)

Mean and Variance Equation: data ~ garch(1, 1) [data = aapl]
Conditional Distribution: norm

Error Analysis:

| Estimate  | Std. Error  | t value | Pr(>|t|) |
|-----------|-------------|---------|----------|
| mu        | 2.297e-03   | 4.025e-04 | 5.706   | 1.16e-08 *** |
| omega     | 6.855e-06   | 2.291e-06 | 2.992   | 0.00277 ** |
| alpha1    | 5.635e-02   | 9.243e-03 | 6.097   | 1.08e-09 *** |
| beta1     | 9.320e-01   | 1.153e-02 | 80.859  | < 2e-16 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera Test</td>
<td>R Chi^2</td>
<td>880.0706</td>
</tr>
<tr>
<td>Shapiro-Wilk Test</td>
<td>R W</td>
<td>0.9743081</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(10)</td>
<td>14.85476</td>
<td>0.137473</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>19.38376</td>
<td>0.4970216</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>5.644311</td>
<td>0.8442093</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>12.05117</td>
<td>0.9143034</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>m1</td>
<td>-4.826153</td>
<td>-4.816887</td>
<td>-4.826158</td>
<td>-4.822790</td>
</tr>
<tr>
<td>m2</td>
<td>-4.826153</td>
<td>-4.816887</td>
<td>-4.826158</td>
<td>-4.822790</td>
</tr>
</tbody>
</table>

> m2=garchFit(~garch(1,1),data=aapl,trace=F,cond.dist="std")
> summary(m2)

Call: garchFit(formula = ~garch(1, 1), data = aapl, cond.dist="std", trace = F)

Mean and Variance Equation: data ~ garch(1, 1) [data = aapl]
Conditional Distribution: std

Error Analysis:

| Estimate  | Std. Error  | t value | Pr(>|t|) |
|-----------|-------------|---------|----------|
| mu        | 1.879e-03   | 3.676e-04 | 5.111   | 3.21e-07 *** |
| omega     | 5.970e-06   | 2.443e-06 | 2.444   | 0.0145 *    |
| alpha1    | 5.221e-02   | 1.131e-02 | 4.616   | 3.90e-06 *** |
| beta1     | 9.378e-01   | 1.339e-02 | 70.008  | < 2e-16 *** |
| shape     | 5.319e+00   | 5.494e-01 | 9.682   | < 2e-16 *** |

---

Standardised Residuals Tests:
Statistic p-Value
Ljung-Box Test R Q(10) 14.73618 0.1419803
Ljung-Box Test R Q(20) 19.45028 0.492754
Ljung-Box Test R^2 Q(10) 6.34873 0.7851631
Ljung-Box Test R^2 Q(20) 12.69443 0.8901063

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>m3</td>
<td>-4.901877</td>
<td>-4.890294</td>
<td>-4.901885</td>
<td>-4.897673</td>
</tr>
</tbody>
</table>

> m3=garchFit(~garch(1,1),data=aapl,trace=F,cond.dist="sstd")
> summary(m3)

Call: garchFit(formula="garch(1, 1)", data=aapl, cond.dist="sstd", trace=F)

Mean and Variance Equation: data ~ garch(1, 1) [data = aapl]
Conditional Distribution: sstd

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | 2.065e-03  | 3.997e-04 | 5.168     | 2.37e-07 *** |
| omega    | 6.160e-06  | 2.497e-06 | 2.467     | 0.0136 *     |
| alpha1   | 5.345e-02  | 1.152e-02 | 4.639     | 3.51e-06 *** |
| beta1    | 9.364e-01  | 1.359e-02 | 68.903    | <2e-16 ***   |
| skew     | 1.033e+00  | 2.846e-02 | 36.307    | <2e-16 ***   |
| shape    | 5.312e+00  | 5.506e-01 | 9.648     | <2e-16 ***   |

---

Standardised Residuals Tests:

Statistic p-Value
Ljung-Box Test R Q(10) 14.7672 0.1407828
Ljung-Box Test R Q(20) 19.40163 0.4958739
Ljung-Box Test R^2 Q(10) 6.163352 0.8013575
Ljung-Box Test R^2 Q(20) 12.55339 0.8957141

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>m4</td>
<td>-4.901640</td>
<td>-4.887741</td>
<td>-4.901651</td>
<td>-4.896596</td>
</tr>
</tbody>
</table>

> m4=garchFit(~aparch(1,1),data=aapl,trace=F,cond.dist="std",delta=2,include.delta=F)
> summary(m4)

Call: garchFit(formula="aparch(1, 1)", data=aapl, delta=2, cond.dist="std", includedelta = F, trace = F)

Mean and Variance Equation: data ~ aparch(1, 1)
Conditional Distribution: std

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | 1.758e-03  | 3.644e-04 | 4.824     | 1.41e-06 *** |
| omega    | 1.087e-05  | 3.770e-06 | 2.882     | 0.00395 **   |
| alpha1   | 6.487e-02  | 1.368e-02 | 4.742     | 2.12e-06 *** |
\begin{verbatim}
> gamma1 3.046e-01 7.254e-02 4.198 2.69e-05 ***
beta1 9.119e-01 1.806e-02 50.496 < 2e-16 ***
shape 5.463e+00 5.788e-01 9.438 < 2e-16 ***
---
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>15.83973</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>20.34165</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>3.01369</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>8.252486</td>
</tr>
</tbody>
</table>

> mean(m4@sigma.t)
[1] 0.02247951

> m5=Igarch(aapl,include.mean=T)
Estimates: 0.002121092 0.9606989
Maximized log-likelyhood: -6063.435
Coefficient(s):

| Estimate       | Std. Error | t value | Pr(>|t|) |
|----------------|------------|---------|----------|
| mu             | 0.002121092| 0.000403833| 5.25239 | 1.5014e-07 *** |
| beta           | 0.960698868| 0.004745300| 202.45272 | < 2.22e-16 *** |
---

> names(m5)
[1] "par" "volatility"

> length(aapl)
[1] 2517

> aapl[2517]
[1] 0.01165383

> m5$volatility[2517]
[1] 0.01324704

Problem C

> d1=log(da[idx,3]+1) ### Decile 1 log returns
> g1=garchFit(~garch(1,1),data=d1,trace=F)
> summary(g1)
Conditional Distribution: norm

> g2=garchFit(~garch(1,1),data=d1,trace=F,cond.dist="std")
> summary(g2)
Conditional Distribution: std

> g3=garchFit(~garch(1,1),data=d1,trace=F,cond.dist="sstd")
> summary(g3)
Call: garchFit(formula=~garch(1,1),data=d1,cond.dist="sstd", trace=F)
Mean and Variance Equation: data ~ garch(1, 1) [data = d1]
Conditional Distribution: sstd

Error Analysis:

| Estimate       | Std. Error | t value | Pr(>|t|) |
|----------------|------------|---------|----------|
\end{verbatim}
mu  7.968e-03  1.459e-03  5.461  4.74e-08 ***
omega  8.496e-05  3.871e-05  2.195  0.028177 *
alpha1  1.379e-01  3.377e-02  4.083  4.45e-05 ***
beta1  8.243e-01  3.667e-02  22.478  < 2e-16 ***
skew  7.837e-01  4.706e-02  16.655  < 2e-16 ***
shape  7.069e+00  1.852e+00  3.816  0.000135 ***

---

Statistic p-Value
Ljung-Box Test   R Q(10) 10.01715  0.4389901
Ljung-Box Test   R Q(20) 15.1446  0.7680746
Ljung-Box Test   R^2 Q(10) 5.619747  0.8461352
Ljung-Box Test   R^2 Q(20) 9.377167  0.9781127

> predict(g3,4)

meanForecast meanError standardDeviation
1  0.007967509  0.03286457  0.03286457
2  0.007967509  0.03352915  0.03352915
3  0.007967509  0.03415640  0.03415640
4  0.007967509  0.03474924  0.03474924

> g4=garchFit(~garch(1,1),data=d1,trace=F,cond.dist="sstd",leverage=T)
> summary(g4)

Call: garchFit(formula =~garch(1,1),data=d1,cond.dist="sstd", leverage=T,trace=F)

Mean and Variance Equation: data ~ garch(1, 1) [data=d1]
Conditional Distribution: sstd

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | 7.365e-03  | 1.483e-03 | 4.965 | 6.87e-07 *** |
| omega    | 1.144e-04  | 4.939e-05 | 2.317 | 0.020512 *   |
| alpha1   | 1.200e-01  | 3.491e-02 | 3.438 | 0.000585 *** |
| gamma1   | 3.016e-01  | 1.495e-01 | 2.018 | 0.043634 *   |
| beta1    | 8.107e-01  | 3.842e-02 | 21.102 | < 2e-16 *** |
| skew     | 7.850e-01  | 4.707e-02 | 16.677 | < 2e-16 *** |
| shape    | 7.177e+00  | 1.898e+00 | 3.782 | 0.000156 *** |

> predict(g4,4)

meanForecast meanError standardDeviation
1  0.007365254  0.03140876  0.03140876
2  0.007365254  0.03213370  0.03213370
3  0.007365254  0.03279401  0.03279401
4  0.007365254  0.03339684  0.03339684

Problem D

> da=read.table("m-gasoil.txt",header=T)
> hp=da$hoil; ng=da$gasp
> lhp=log(hp)
> ghp=diff(lhp)

15
> m1=ar(ghp, method="mle")
> m1$order
[1] 1
> m2=arima(lhp, order=c(1,1,0)) ### model for log(heating oil price)
> m2
Call:arima(x = lhp, order = c(1, 1, 0))
Coefficients:
   ar1
0.2029
s.e. 0.0657

sigma^2 estimated as 0.007063: log likelihood = 237.92, aic = -471.83
> m3=arima(lhp, order=c(0,1,1))
> m3
Call:arima(x = lhp, order = c(0, 1, 1))
Coefficients:
   ma1
0.1833
s.e. 0.0608

sigma^2 estimated as 0.007091: log likelihood = 237.48, aic = -470.96
> backtest(m2, lhp, 200, 1)
[1] "RMSE of out-of-sample forecasts"
[1] 0.05218009
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.04500329
> backtest(m3, lhp, 200, 1)
[1] "RMSE of out-of-sample forecasts"
[1] 0.05219218
[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.04506306
> lng=log(ng)
> m3=lm(lhp~lng)
> summary(m3)
Call: lm(formula = lhp ~ lng)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.09767  0.08862  -12.39 <2e-16 ***
    lng      0.85132  0.06194   13.74 <2e-16 ***
---
Residual standard error: 0.4997 on 224 degrees of freedom
Multiple R-squared:  0.4575,   Adjusted R-squared:  0.4551
> acf(m3$residuals)
> gng=diff(lng)
> m3a=lm(ghp~-1+gng)
> summary(m3a)
Call: lm(formula = ghp ~ -1 + gng)
Coefficients:
   Estimate Std. Error t value Pr(>|t|)
gng 0.21003    0.03726   5.637 5.18e-08 ***
---
Residual standard error: 0.08049 on 224 degrees of freedom
Multiple R-squared: 0.1242, Adjusted R-squared: 0.1203

> acf(m3a$residuals)
> m4=aRIMA(ghp, order=c(1,0,0), xreg=gng, include.mean=F)
> m4
Call:aRIMA(x = ghp, order = c(1,0,0), xreg = gng, include.means = F)
Coefficients:
          ar1   gng
ar1   0.1919  0.2018
s.e.  0.0660  0.0365
sigma^2 estimated as 0.006215: log likelihood = 252.31, aic = -498.63

Problem E
> da=read.table("q-pg-earnings.txt",header=T)
> pg=da[,2]
> acf(pg); acf(diff(pg)); acf(diff(diff(pg),4))
> m1=aRIMA(pg, order=c(0,1,1), seasonal=list(order=c(0,1,1),period=4))
> m1
Call:aRIMA(x = pg, order = c(0,1,1), seasonal = list(order = c(0,1,1), period = 4))
Coefficients:
       ma1    sma1
ma1   -0.7098  -0.5198
s.e.   0.0736   0.2182
sigma^2 estimated as 0.006728: log likelihood = 121.12, aic = -236.24

> which.max(m1$residuals)
[1] 104
> length(pg)
[1] 118
> I104=rep(0,118); I104[104]=1
> m2=aRIMA(pg, order=c(0,1,1), seasonal=list(order=c(0,1,1),period=4), xreg=I104)
> m2
Coefficients:
       ma1    sma1    I104
ma1   -0.7136  -0.6427    0.4498
s.e.   0.0699   0.0900   0.0586
sigma^2 estimated as 0.004327: log likelihood = 141.58, aic = -275.16
> which.max(m2$residuals)
[1] 108
> I108=rep(0,118); I108[108]=1

17
> X=cbind(I104,I108)
> m3=arima(pg,order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4),xreg=X)
> m3
Coefficients:
   ma1   sma1  I104  I108
ma1  -0.4855 -0.628  0.5356  0.4466
s.e.  0.1029  0.077  0.0344  0.0349
sigma^2 estimated as 0.001809:  log likelihood = 189.08,  aic = -368.16
> which.min(m3$residuals)
[1] 18
> I18=rep(0,118); I18[18]=1
> X=cbind(X,I18)
> m4=arima(pg,order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4),xreg=X)
> m4
Coefficients:
   ma1   sma1  I104  I108  I18
ma1  -0.3301 -0.5322  0.5303  0.4421 -0.1711
s.e.  0.1220  0.0858  0.0273  0.0277  0.0261
sigma^2 estimated as 0.001341:  log likelihood = 205.69,  aic = -399.39
> m5=arima(pg,order=c(0,1,3),seasonal=list(order=c(0,1,1),period=4),xreg=X)
> m5
Coefficients:
   ma1   ma2  ma3   sma1  I104  I108  I18
ma1  -0.4113  0.3465 -0.7428 -0.2694  0.4790  0.4419 -0.1754
s.e.  0.0901  0.0889  0.1009  0.1443  0.0165  0.0176  0.0162
sigma^2 estimated as 0.001148:  log likelihood = 213.08,  aic = -410.17
> which.max(m5$residuals)
[1] 102
> I102=rep(0,118); I102[102]=1
> X=cbind(X,I102)
> m6=arima(pg,order=c(0,1,3),seasonal=list(order=c(0,1,1),period=4),xreg=X)
> m6
Coefficients:
   ma1   ma2  ma3   sma1  I104  I108  I18  I102
ma1  -0.0741 -0.2757 -0.3232 -0.3347  0.5389  0.4195 -0.1701  0.1501
s.e.  0.0981  0.1022  0.0916  0.1081  0.0163  0.0176  0.0139  0.0144
sigma^2 estimated as 0.0006837:  log likelihood = 242.17,  aic = -466.34
> c1=c(0,NA,NA,NA,NA,NA,NA,NA)
> m7=arima(pg,order=c(0,1,3),seasonal=list(order=c(0,1,1),period=4),xreg=X,fixed=c1)
> m7
Coefficients:
   ma1   ma2  ma3   sma1  I104  I108  I18  I102
ma1   0  -0.2757 -0.3232 -0.3347  0.5389  0.4195 -0.1701  0.1501
s.e.  0.1022  0.0916  0.1081  0.0163  0.0158  0.0139  0.0144
sigma^2 estimated as 0.0006877:  log likelihood = 241.89,  aic = -467.77