2018

Empirical Project

CONSUMPTION SMOOTHING

ANNI HONG, KELCI YOUNG, HUGO BALDNER

Abstract

In order to explore the hypothesis of Consumption smoothing we will use the data of the Dow Jones Stock, real nondurable and service consumption, and the real GDP (of the USA) from 1947 to 2007 to explore this through the models given in the Campbell_Mankiw_1989 and Hall_1978 papers.

Summarizing Existing Data(Question 1):

(a) Compute using your data the (log) growth rate of income, nondurable and service consumption, and the dow jones index:

For the purposes of time we will not show all the newly calculated data here. However, from this point onward we will be testing using this newly acquired data.

(b) Provide a set of summary statistics:

GDP Growth

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Standard Deviation
-0.0312100	0.0004114	0.0051510	0.0052230	0.0105300	0.0360500	0.009801001

ND Growth

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Standard Deviation
-0.009339	0.005324	0.008044	0.007964	0.010820	0.024990	0.005190234

DJ Growth

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Standard Deviation
-0.246100	-0.023050	0.011950	0.009407	0.049750	0.161600	0.05776443

(c) Provide a table with the correlations between the three variables of interest.

Correlation	ND Growth	DJ Growth	GDP Growth
ND Growth	1	0.2676591	0.4437897
DJ Growth	0.2676591	1	0.1933038
GDP Growth	0.4437897	0.1933038	1

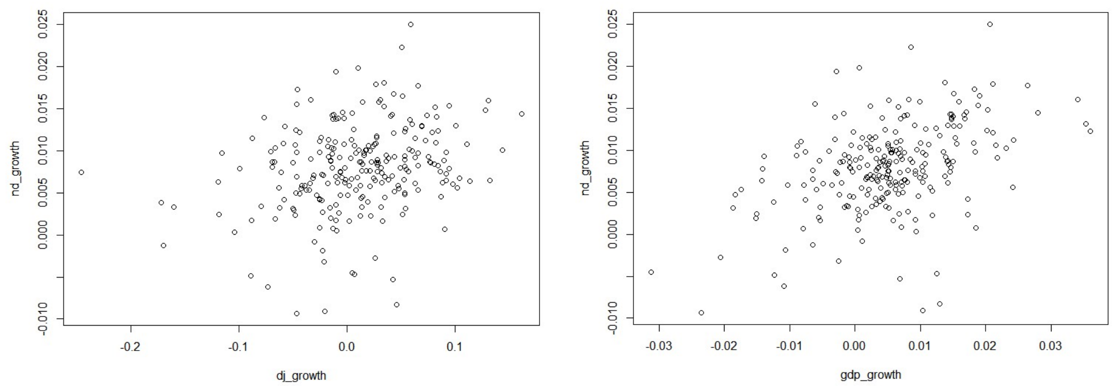
We find from this that while they tend in the same direction none of the Variables are stalwartly correlated.

(d) Provide a table with the autocorrelations of the three variables of interest, for 0 to 6 lags.

Autocorrelations (T:T-K)	GDP	ND	DJ
T	1	1	1
t-1	-0.79	-0.75	-0.80
t-2	0.11	-0.03	0.24
t-3	0.12	0.04	-0.04
t-4	-0.01	0.13	-0.01
t-5	0.02	0.00	0.04
t-6	-0.10	-0.08	-0.05

We find that while the variables are barely correlated to each other they are even less correlated to themselves. All off the autocorrelations after one period hover around 0.0, some even going negative.

(e) Provide a scatter plot of real consumption growth (y-axis) against real income growth (x-axis), and a scatter plot of real consumption growth (y-axis) against real stock market growth (x-axis).



We see from this graph once again that the correlation is dubious at best. The data points hardly implicate any sort of linear relationship. What we can pull from this is that the data represents a large variety of scenarios which makes sense considering the relatively vast time period it spans.

Testing the Martingale Property(Question 2):

Following Hall (1978 Journal of Political Economy), we test the implication that consumption growth should not be predictable. A simple way is to run an OLS regression of consumption growth on various variables known in the previous quarter, and see if they are significantly different from zero, i.e. test for b = 0 Try this regression with the new equations. What if you include additional lags of consumption, income or the stock market?
What's your conclusions?

Analyzing the regression on our new equations on variables known in the previous quarter we find:

For ND

Coefficients	Estimate	Standard Error	T Value	P Value
(Intercept)	0.0063837	0.0006014	10.62	< 2e-16
L(d(ND))	0.1939475	0.0631806	3.07	0.00239

For GDP

Coefficients	Estimate	Standard Error	T Value	P Value
(Intercept)	0.0072161	0.0003653	19.755	< 2e-16
L(d(GDP))	0.1367987	0.0328776	4.161	4.42e-05

For DJ

Coefficients	Estimate	Standard Error	T Value	P Value
(Intercept)	0.0077501	0.0003301	23.477	< 2e-16
L(d(DJ))	0.0192479	0.0056404	3.413	0.000755

Analyzing the regression on our new equations on variables with multiple lags (Three) we find:

Coefficients	L(d(ND))	L(d(GDP))	L(d(DJ))
T Value	2.276	3.426	3.349
P Value	0.0237	0.000723	0.000944

All these analysis's point to us rejecting the null hypothesis that b=0. With the p value being this low the null hypot hesis must be rejected as the odds
It is correct fall by a large margin. However it should be noted that the p value for DJ is larger (in comparison to the other ones) when the regression is tested with multiple lags in play.

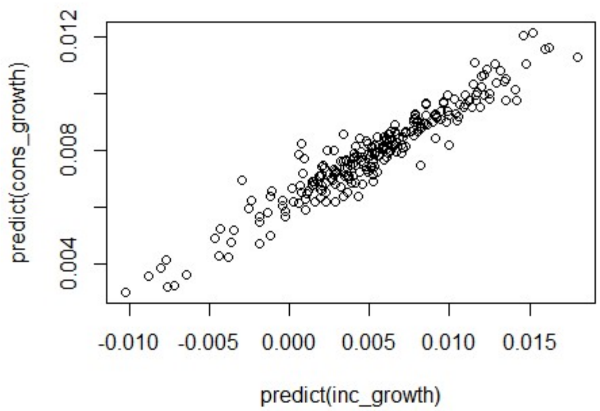
The Importance of Hand To Mouth Consumers(Question 3):

(a) Run a regression of GDP growth on lagged values of GDP growth (say the last 4 lags of GDP growth). Does lagged income growth forecast future income growth? Redo the same using lagged values of consumption growth instead of GDP growth, and using stock market growth. Conclude by specifying a set of variables Xt 1 known at time t 1, that forecasts income growth.

Coefficients	Estimate	Standard Error	T Value	P Value
Intercept	0.0039334	0.0007776	5.058	8.54e-07
L(d(GDP)) (1:4) 1	0.3066419	0.0650772	4.712	4.21e-06
L(d(GDP)) (1:4) 2	0.1320864	0.0676925	1.951	0.0522
L(d(GDP)) (1:4) 3	-0.0952191	0.0675989	-1.409	0.1603
L(d(GDP)) (1:4) 4	-0.0942912	0.0647571	-1.456	0.1467

The First two lags of GDP growth appear to have a small enough P value for the null hypothesis to be rejected. This translates into the fact that you can (in this model) use the previous two periods to more or less predict the future GDP Growth.

(b) Compare this scatter plot to the scatter plot of question 1 (e). If equation (1) is true, what should we observe for the coefficients bc and by? What does the correlation between the fitted values of consumption growth and income growth tell you?



This new graph of the first lag differences found to be most predictive in part one shows a far more correlated set of data than the graphs previously shown. This means that the new equations points bc and by being larger than zero and therefore more easily predicted by past values.

(c) Run an IV (instrumental-variable) regression. What value do you find for _? Discuss its size and significance.

Using lagged Stock growth due to it being a good predictor of GDP growth and the fact that the lagged will be in th e past it will be unrelated to current news. We use this to calculate the T, P, R^2, and lambda Values. Being 5.863, 1. 49e-08, 0.1253, and 0.3213068 respectively. This size implies that while being very significant it might not be the b est predictor for future values.

Conclusion:

We found that while analyzing the Dow Jones Stock, the Real Gross National Product of the USA, and real nondura ble and service consumption, over a large span of time displaying a variety of scenarios, that under some models y ou can at times predict future values to a degree. This is however very limited as we failed to find in either model t hat any value further than maybe 2 periods after the current can be predicted. We did however find that in the Ca mpbell/Mankiw model consumption and income growth are positively correlated.

All Code Used:

```
library(readxl)
dat <- read_excel(file.choose())
library(dynlm)
tsdata <- ts(dat, start=1947) # make the data into a time series object
gdp <- log(dat$gdp)[2:244]
Lgdp <- log(dat$gdp)[1:243]
gdp_growth <- gdp - Lgdp

nd <- log(dat$nd)[2:244]
Lnd <- log(dat$nd)[1:243]
nd_growth <- nd - Lnd

dj <- log(dat$dj)[2:244]
Ldj <- log(dat$dj)[1:243]
dj_growth <- dj - Ldj

summary(gdp_growth)
summary(nd_growth)
summary(dj_growth)

cor(gdp_growth,nd_growth)
cor(gdp_growth,dj_growth)
cor(nd_growth,dj_growth)

autocorr_gdp <- dynlm(log(gdp)~log(L(gdp,0:6)),data=tsdata)
summary(autocorr_gdp, correlation=T)

autocorr_nd <- dynlm(log(nd)~log(L(nd,0:6)),data=tsdata)
summary(autocorr_nd, correlation=T)

autocorr_dj <- dynlm(log(dj)~log(L(dj,0:6)),data=tsdata)
summary(autocorr_dj, correlation=T)

plot(gdp_growth, nd_growth)
plot(dj_growth, nd_growth)
# question two
log(dat$gdp) -> dat[,3]
log(dat$nd) -> dat[,4]
log(dat$dj) -> dat[,5]
tsdata <- ts(dat) # make the data into a time series object

# with consumption lag:
fit <- dynlm(d(nd)~L(d(nd)), data = tsdata)
summary(fit)
#3.07  0.00239 **
# reject the null that beta = 0

# with income lag
fit <- dynlm(d(nd)~L(d(gdp)), data = tsdata)
summary(fit)
#3.518  0.00052 *

#with stocks lag
fit <- dynlm(d(nd)~L(d(dj)), data = tsdata)
summary(fit)
#3.413 0.000755 *

# with consumption lag:
fit <- dynlm(d(nd)~L(d(nd), 1:3), data = tsdata)
summary(fit)
fit <- dynlm(d(nd)~L(d(gdp), 1:3), data = tsdata)
summary(fit)
fit <- dynlm(d(nd)~L(d(dj), 1:3), data = tsdata)
summary(fit)
```

Raw Output:

```
> library(readxl)
> dat <- read_excel(file.choose())
> library(dynlm)
Loading required package: zoo

Attaching package: ‘zoo’

The following objects are masked from ‘package:base’:

    as.Date, as.Date.numeric

> tsdata <- ts(dat, start=1947) # make the data into a time series obje
ct
> gdp <- log(dat$gdp)[2:244]
> Lgdp <- log(dat$gdp)[1:243]
> gdp_growth <- gdp - Lgdp
>
> nd <- log(dat$nd)[2:244]
> Lnd <- log(dat$nd)[1:243]
> nd_growth <- nd - Lnd
>
> dj <- log(dat$dj)[2:244]
> Ldj <- log(dat$dj)[1:243]
> dj_growth <- dj - Ldj
>
> summary(gdp_growth)
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.0312100  0.0004114  0.0051510  0.0052230  0.0105300  0.0360500
> summary(nd_growth)
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.009339  0.005324  0.008044  0.007964  0.010820  0.024990
> summary(dj_growth)
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.246100 -0.023050  0.011950  0.009407  0.049750  0.161600
>
> cor(gdp_growth,nd_growth)
[1] 0.4437897
> cor(gdp_growth,dj_growth)
[1] 0.1933038
> cor(nd_growth,dj_growth)
[1] 0.2676591
>
> autocorr_gdp <- dynlm(log(gdp)~log(L(gdp,0:6)),data=tsdata)
> summary(autocorr_gdp, correlation=T)

Time series regression with "ts" data:
Start = 1953, End = 2190

Call:
dynlm(formula = log(gdp) ~ log(L(gdp, 0:6)), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-9.615e-16 -1.008e-17  3.130e-18  1.478e-17  9.921e-17

Coefficients:
                    Estimate Std. Error    t value Pr(>|t|)
(Intercept)       -2.303e-16  1.011e-17 -2.278e+01   <2e-16 ***
log(L(gdp, 0:6))0  1.000e+00  4.864e-16  2.056e+15   <2e-16 ***
log(L(gdp, 0:6))1 -1.345e-16  7.943e-16 -1.690e-01    0.866
log(L(gdp, 0:6))2 -9.111e-16  7.988e-16 -1.140e+00    0.255
log(L(gdp, 0:6))3  6.237e-16  8.047e-16  7.750e-01    0.439
log(L(gdp, 0:6))4 -4.086e-16  7.974e-16 -5.120e-01    0.609
log(L(gdp, 0:6))5  1.141e-15  7.930e-16  1.439e+00    0.151
log(L(gdp, 0:6))6 -7.238e-16  4.834e-16 -1.497e+00    0.136
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.784e-17 on 230 degrees of freedom
Multiple R-squared:      1,	Adjusted R-squared:      1
F-statistic: 1.007e+33 on 7 and 230 DF,  p-value: < 2.2e-16

Correlation of Coefficients:
                  (Intercept) log(L(gdp, 0:6))0 log(L(gdp, 0:6))1 log(L
(gdp, 0:6))2 log(L(gdp, 0:6))3 log(L(gdp, 0:6))4 log(L(gdp, 0:6))5
log(L(gdp, 0:6))0 -0.24
log(L(gdp, 0:6))1  0.05       -0.79
log(L(gdp, 0:6))2  0.02        0.11              -0.57
log(L(gdp, 0:6))3  0.00        0.12              -0.03             -0.55
log(L(gdp, 0:6))4 -0.02       -0.01               0.09             -0.04
-0.55
log(L(gdp, 0:6))5 -0.05        0.02              -0.02              0.09
-0.03             -0.57
log(L(gdp, 0:6))6  0.21       -0.10               0.02             -0.01
0.12              0.11             -0.79

Warning message:
In summary.lm(autocorr_gdp, correlation = T) :
  essentially perfect fit: summary may be unreliable
>
> autocorr_nd <- dynlm(log(nd)~log(L(nd,0:6)),data=tsdata)
> summary(autocorr_nd, correlation=T)

Time series regression with "ts" data:
Start = 1953, End = 2190

Call:
dynlm(formula = log(nd) ~ log(L(nd, 0:6)), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-4.231e-16 -6.940e-17 -7.700e-18  5.380e-17  4.855e-15

Coefficients:
                   Estimate Std. Error    t value Pr(>|t|)
(Intercept)       1.151e-15  9.718e-17  1.185e+01   <2e-16 ***
log(L(nd, 0:6))0  1.000e+00  4.571e-15  2.187e+14   <2e-16 ***
log(L(nd, 0:6))1 -1.085e-14  6.880e-15 -1.577e+00   0.1163
log(L(nd, 0:6))2  1.645e-14  6.853e-15  2.400e+00   0.0172 *
log(L(nd, 0:6))3 -4.516e-15  6.742e-15 -6.700e-01   0.5036
log(L(nd, 0:6))4 -1.655e-14  6.755e-15 -2.450e+00   0.0150 *
log(L(nd, 0:6))5  8.582e-15  6.688e-15  1.283e+00   0.2007
log(L(nd, 0:6))6  5.374e-15  4.391e-15  1.224e+00   0.2222
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.387e-16 on 230 degrees of freedom
Multiple R-squared:      1,	Adjusted R-squared:      1
F-statistic: 9.351e+31 on 7 and 230 DF,  p-value: < 2.2e-16

Correlation of Coefficients:
                 (Intercept) log(L(nd, 0:6))0 log(L(nd, 0:6))1 log(L(nd
, 0:6))2 log(L(nd, 0:6))3 log(L(nd, 0:6))4 log(L(nd, 0:6))5
log(L(nd, 0:6))0 -0.34
log(L(nd, 0:6))1  0.04       -0.75
log(L(nd, 0:6))2  0.05       -0.03             -0.47
log(L(nd, 0:6))3  0.02        0.04             -0.03            -0.50
log(L(nd, 0:6))4 -0.04        0.13             -0.08            -0.03
-0.49
log(L(nd, 0:6))5 -0.06        0.00              0.07            -0.06
-0.01            -0.51
log(L(nd, 0:6))6  0.32       -0.08              0.01             0.11
0.01             0.01             -0.75

Warning message:
In summary.lm(autocorr_nd, correlation = T) :
  essentially perfect fit: summary may be unreliable
>
> autocorr_dj <- dynlm(log(dj)~log(L(dj,0:6)),data=tsdata)
> summary(autocorr_dj, correlation=T)

Time series regression with "ts" data:
Start = 1953, End = 2190

Call:
dynlm(formula = log(dj) ~ log(L(dj, 0:6)), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-4.655e-16 -9.470e-17 -2.170e-17  2.040e-17  7.023e-15

Coefficients:
                   Estimate Std. Error    t value Pr(>|t|)
(Intercept)       6.448e-15  4.163e-16  1.549e+01   <2e-16 ***
log(L(dj, 0:6))0  1.000e+00  5.679e-16  1.761e+15   <2e-16 ***
log(L(dj, 0:6))1 -6.331e-16  9.514e-16 -6.650e-01   0.5064
log(L(dj, 0:6))2  1.744e-15  9.715e-16  1.795e+00   0.0739 .
log(L(dj, 0:6))3 -1.923e-15  9.715e-16 -1.980e+00   0.0489 *
log(L(dj, 0:6))4  7.224e-16  9.709e-16  7.440e-01   0.4576
log(L(dj, 0:6))5  4.796e-16  9.435e-16  5.080e-01   0.6117
log(L(dj, 0:6))6 -3.748e-16  5.650e-16 -6.630e-01   0.5078
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.755e-16 on 230 degrees of freedom
Multiple R-squared:      1,	Adjusted R-squared:      1
F-statistic: 5.747e+31 on 7 and 230 DF,  p-value: < 2.2e-16

Correlation of Coefficients:
```

```
              (Intercept) log(L(dj, 0:6))0 log(L(dj, 0:6))1 log(L(dj
, 0:6))2 log(L(dj, 0:6))3 log(L(dj, 0:6))4 log(L(dj, 0:6))5
log(L(dj, 0:6))0 -0.05
log(L(dj, 0:6))1  0.00       -0.80
log(L(dj, 0:6))2  0.01        0.24             -0.65
log(L(dj, 0:6))3  0.00       -0.04              0.17             -0.64
log(L(dj, 0:6))4  0.00       -0.01             -0.02              0.17
-0.65
log(L(dj, 0:6))5  0.01        0.04             -0.03             -0.02
0.18             -0.65
log(L(dj, 0:6))6 -0.07       -0.05              0.04              0.00
-0.05             0.24             -0.80

Warning message:
In summary.lm(autocorr_dj, correlation = T) :
  essentially perfect fit: summary may be unreliable
>
> plot(gdp_growth, nd_growth)
> plot(dj_growth, nd_growth)
> # question two
> log(dat$gdp) -> dat[,3]
> log(dat$nd) -> dat[,4]
> log(dat$dj) -> dat[,5]
> tsdata <- ts(dat) # make the data into a time series object
>
> # with consumption lag:
> fit <- dynlm(d(nd)~L(d(nd)), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 3, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(nd)), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.017747 -0.002374 -0.000086  0.002806  0.017457

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0063837  0.0006014   10.62  < 2e-16 ***
L(d(nd))    0.1939475  0.0631806    3.07  0.00239 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.00509 on 240 degrees of freedom
Multiple R-squared:  0.03778,  Adjusted R-squared:  0.03377
F-statistic: 9.423 on 1 and 240 DF,  p-value: 0.002389

> #3.07  0.00239 **
> # reject the null that beta = 0
>
> # with income lag
> fit <- dynlm(d(nd)~L(d(gdp)), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 3, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(gdp)), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.020137 -0.002449 -0.000179  0.002731  0.015492

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0072161  0.0003653  19.755  < 2e-16 ***
L(d(gdp))   0.1367987  0.0328776   4.161 4.42e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.005011 on 240 degrees of freedom
Multiple R-squared:  0.06728,  Adjusted R-squared:  0.0634
F-statistic: 17.31 on 1 and 240 DF,  p-value: 4.419e-05

> #3.518  0.00052 *
>
> #with stocks lag
> fit <- dynlm(d(nd)~L(d(dj)), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 3, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(dj)), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0174226 -0.0030520 -0.0000938  0.0030327  0.0175158

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0077501  0.0003301  23.477  < 2e-16 ***
L(d(dj))    0.0192479  0.0056404   3.413 0.000755 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.005068 on 240 degrees of freedom
Multiple R-squared:  0.04628,  Adjusted R-squared:  0.0423
F-statistic: 11.65 on 1 and 240 DF,  p-value: 0.0007552

> #3.413 0.000755 *
>
> # with consumption lag:
> fit <- dynlm(d(nd)~L(d(nd), 1:3), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 5, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(nd), 1:3), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0192882 -0.0024831 -0.0003253  0.0026527  0.0151942

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.0050131  0.0007934   6.318 1.31e-09 ***
L(d(nd), 1:3)1 0.1440835  0.0632963   2.276   0.0237 *
L(d(nd), 1:3)2 0.1409954  0.0634975   2.220   0.0273 *
L(d(nd), 1:3)3 0.0924505  0.0629285   1.469   0.1431
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.004917 on 236 degrees of freedom
Multiple R-squared:  0.06956,  Adjusted R-squared:  0.05773
F-statistic: 5.881 on 3 and 236 DF,  p-value: 0.0006901

> fit <- dynlm(d(nd)~L(d(gdp), 1:3), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 5, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(gdp), 1:3), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0197119 -0.0024216 -0.0002899  0.0028510  0.0158562

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.0073275  0.0003966  18.475  < 2e-16 ***
L(d(gdp), 1:3)1  0.1183671  0.0345536   3.426 0.000723 ***
L(d(gdp), 1:3)2  0.0419824  0.0360217   1.165 0.245003
L(d(gdp), 1:3)3 -0.0287814  0.0344207  -0.836 0.403907
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.004923 on 236 degrees of freedom
Multiple R-squared:  0.06708,  Adjusted R-squared:  0.05522
F-statistic: 5.657 on 3 and 236 DF,  p-value: 0.0009289

> fit <- dynlm(d(nd)~L(d(dj), 1:3), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 5, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(dj), 1:3), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0169554 -0.0027712 -0.0001608  0.0027245  0.0172621
```

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.007730   0.000325  23.787  < 2e-16 ***
L(d(dj), 1:3)1  0.019387   0.005789   3.349 0.000944 ***
L(d(dj), 1:3)2 -0.002687   0.006106  -0.440 0.660313
L(d(dj), 1:3)3  0.015020   0.005777   2.600 0.009914 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.004899 on 236 degrees of freedom
Multiple R-squared:  0.07638,	Adjusted R-squared:  0.06464
F-statistic: 6.505 on 3 and 236 DF,  p-value: 0.0003021

> #question3
> #a)
> fit <- dynlm(d(gdp)~L(d(gdp), 1:4), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 6, End = 244

Call:
dynlm(formula = d(gdp) ~ L(d(gdp), 1:4), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.031631 -0.004864 -0.000083  0.005020  0.035247

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.0039334  0.0007776   5.058 8.54e-07 ***
L(d(gdp), 1:4)1  0.3066419  0.0650772   4.712 4.21e-06 ***
L(d(gdp), 1:4)2  0.1320864  0.0676925   1.951   0.0522 .
L(d(gdp), 1:4)3 -0.0952191  0.0675989  -1.409   0.1603
L(d(gdp), 1:4)4 -0.0942912  0.0647571  -1.456   0.1467
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.009175 on 234 degrees of freedom
Multiple R-squared:  0.1413,	Adjusted R-squared:  0.1267
F-statistic:  9.63 on 4 and 234 DF,  p-value: 3.168e-07

> fit <- dynlm(d(gdp)~L(d(nd), 1:4), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 6, End = 244

Call:
dynlm(formula = d(gdp) ~ L(d(nd), 1:4), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.033220 -0.004706  0.000092  0.004869  0.032915

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.002351   0.001546  -1.521  0.12966
L(d(nd), 1:4)1  0.789583   0.116693   6.766 1.05e-10 ***
L(d(nd), 1:4)2  0.329833   0.114646   2.877  0.00439 **
L(d(nd), 1:4)3 -0.107425   0.115005  -0.934  0.35122
L(d(nd), 1:4)4 -0.064639   0.113217  -0.571  0.56859
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008804 on 234 degrees of freedom
Multiple R-squared:  0.2094,	Adjusted R-squared:  0.1958
F-statistic: 15.49 on 4 and 234 DF,  p-value: 2.952e-11

> fit <- dynlm(d(gdp)~L(d(dj), 1:4), data = tsdata)
> summary(fit)

Time series regression with "ts" data:
Start = 6, End = 244

Call:
dynlm(formula = d(gdp) ~ L(d(dj), 1:4), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.029077 -0.005193 -0.000362  0.005589  0.039056

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.0043667  0.0006088   7.172 9.64e-12 ***
L(d(dj), 1:4)1 0.0507556  0.0107632   4.716 4.14e-06 ***
L(d(dj), 1:4)2 0.0251828  0.0113770   2.213   0.0278 *
L(d(dj), 1:4)3 0.0072323  0.0113704   0.636   0.5254
L(d(dj), 1:4)4 0.0102091  0.0107293   0.952   0.3423
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.009096 on 234 degrees of freedom
Multiple R-squared:  0.1561,	Adjusted R-squared:  0.1417
F-statistic: 10.82 on 4 and 234 DF,  p-value: 4.58e-08

> R
Error: object 'R' not found
> # it seems like the first one or two lags is/are predictive of the gd
p growth but not beyond that
> # variables that are significant: first lag of gdp growth, second lag
of gdp growth, first lag of consumption growth,
> # second lag of consumption growth, 1 through 3 lags of the stock gro
wth
>
> #b)
> cons_growth <- dynlm(d(nd)~L(d(gdp)) + L(d(nd)) + L(d(dj)), data = ts
data)
> inc_growth <- dynlm(d(gdp)~L(d(gdp)) + L(d(nd)) + L(d(dj)), data = ts
data)
> summary(cons_growth)

Time series regression with "ts" data:
Start = 3, End = 244

Call:
dynlm(formula = d(nd) ~ L(d(gdp)) + L(d(nd)) + L(d(dj)), data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0191299 -0.0026183 -0.0002419  0.0029003  0.0157962

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0067427  0.0005919  11.391  < 2e-16 ***
L(d(gdp))   0.1058875  0.0363448   2.913  0.00392 **
L(d(nd))    0.0625956  0.0700147   0.894  0.37221
L(d(dj))    0.0142829  0.0057353   2.490  0.01345 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.004947 on 238 degrees of freedom
Multiple R-squared:  0.09861,	Adjusted R-squared:  0.08725
F-statistic: 8.679 on 3 and 238 DF,  p-value: 1.733e-05

> plot(predict(inc_growth), predict(cons_growth))
> # they are highly correlated here versus that the random scatter in p
art e)
>
>
> #c)
> # i chose to use lagged stock growth as the instrumental variable bec
ause it happened at t-1 so it's unrelated to th news
> # but it's still a good predictor for gdp growth
> fir_stage <- dynlm(d(gdp)~L(d(dj)), data = tsdata)
> x_hat_gdp_growth <- predict(fir_stage)
> summary(fir_stage) # using the first lag of stock index as instrument
al variable, super significant but not very predictive, is that cool?

Time series regression with "ts" data:
Start = 3, End = 244

Call:
dynlm(formula = d(gdp) ~ L(d(dj)), data = tsdata)

Residuals:
      Min        1Q    Median        3Q       Max
-0.028798 -0.004815 -0.000662  0.004958  0.036619

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0047015  0.0005979   7.863 1.27e-13 ***
L(d(dj))    0.0599052  0.0102167   5.863 1.49e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.009179 on 240 degrees of freedom
Multiple R-squared:  0.1253,	Adjusted R-squared:  0.1217
F-statistic: 34.38 on 1 and 240 DF,  p-value: 1.492e-08

> sec_stage <- lm(nd_growth[2:243]~x_hat_gdp_growth, data = tsdata)
```

```
> summary(sec_stage)

Call:
lm(formula = nd_growth[2:243] ~ x_hat_gdp_growth, data = tsdata)

Residuals:
       Min         1Q     Median         3Q        Max 
-0.0174226 -0.0030520 -0.0000938  0.0030327  0.0175158 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      0.0062395  0.0005935  10.513  < 2e-16 ***
x_hat_gdp_growth 0.3213068  0.0941548   3.413 0.000755 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.005068 on 240 degrees of freedom
Multiple R-squared:  0.04628,	Adjusted R-squared:  0.0423 
F-statistic: 11.65 on 1 and 240 DF,  p-value: 0.0007552
```