

# A Look into GDP figures

Sanjaya J Shetty

04/06/2021

## Answers

- a) and b)

$$\begin{aligned} 1) \quad l_{11} &= -2(-149.521 - (-134.178)) = \underline{30.68} \\ l_{12} &= -2(-139.747 - (-134.178)) = \underline{11.14} \\ l_{1, l_2} &= -2(-152.763 - (-134.178)) = \underline{37.17} \end{aligned}$$

$$l_{11} @ \alpha = 5\% \chi^2(1) = 3.84$$

hence  $H_0$  rejected

$$l_{12} @ \alpha = 5\% \chi^2(1) = 3.84$$

hence  $H_0$  rejected

$$l_{1, l_2} @ \alpha = 5\% \chi^2(2) = 5.99$$

hence  $H_0$  rejected.

Hence  $l_1, l_2$  &  $(l_1, l_2)$  are all significant

McFadden  $R^2$  could be used, because number of coefficient are the same,

$$R^2 = 1 - \frac{\log(L(b))}{\log(L(b_0))}$$

$$R_1^2 = 0.1216 \quad R_2^2 = 0.1220 \quad R_3^2 = 0.1467 \quad R_4^2 = 0.1460$$

Third model  $R_3^2$  performs well.

• c)

Date	PredValue
1950Q1	0.3707502
1950Q2	0.2525610
1950Q3	0.4095079
1950Q4	0.5473576
1951Q1	0.4820078
1951Q2	0.6852489
1951Q3	0.8530844
1951Q4	0.8723613
1952Q1	0.7449771
1952Q2	0.8988035
1952Q3	0.8122962
1952Q4	0.6783065
1953Q1	0.7062002
1953Q2	0.7633257
1953Q3	0.8122962
1953Q4	0.8358955
1954Q1	0.8862504
1954Q2	0.8862504
1954Q3	0.8862504
1954Q4	0.8988035
1955Q1	0.8687555
1955Q2	0.7127955
1955Q3	0.5147457
1955Q4	0.6186920
1956Q1	0.2845506
1956Q2	0.3119540
1956Q3	0.3407396
1956Q4	0.3407396
1957Q1	0.3119540
1957Q2	0.3119540
1957Q3	0.3119540
1957Q4	0.4415189
1958Q1	0.5795679
1958Q2	0.7326286
1958Q3	0.7807427
1958Q4	0.8222987
1959Q1	0.7515694
1959Q2	0.5717513
1959Q3	0.5795679
1959Q4	0.2845506
1960Q1	0.1857297
1960Q2	0.2063613
1960Q3	0.2845506
1960Q4	0.3119540
1961Q1	0.3407396
1961Q2	0.6417574
1961Q3	0.8448806
1961Q4	0.9047368
1962Q1	0.8762079
1962Q2	0.7752157
1962Q3	0.6641849
1962Q4	0.7515694

Date	PredValue
1963Q1	0.6641849
1963Q2	0.7263138
1963Q3	0.6712844
1963Q4	0.5795679
1964Q1	0.4820078
1964Q2	0.4494236
1964Q3	0.5873449
1964Q4	0.6852489
1965Q1	0.8122962
1965Q2	0.8073683
1965Q3	0.7807427
1965Q4	0.8023425
1966Q1	0.7752157
1966Q2	0.8726950
1966Q3	0.8574170
1966Q4	0.6641849
1967Q1	0.8023425
1967Q2	0.8796373
1967Q3	0.8958555
1967Q4	0.8530844
1968Q1	0.7915015
1968Q2	0.7127955
1968Q3	0.7388500
1968Q4	0.5067496
1969Q1	0.3633161
1969Q2	0.4987500
1969Q3	0.7515694
1969Q4	0.8574170
1970Q1	0.9227991
1970Q2	0.9316299
1970Q3	0.8534600
1970Q4	0.8990760
1971Q1	0.9481422
1971Q2	0.8574170
1971Q3	0.6343676
1971Q4	0.3335883
1972Q1	0.4017930
1972Q2	0.4017930
1972Q3	0.4740234
1972Q4	0.6186920
1973Q1	0.7690580
1973Q2	0.8358955
1973Q3	0.8862504
1973Q4	0.7449771
1974Q1	0.8358955
1974Q2	0.8358955
1974Q3	0.8530844
1974Q4	0.6852489
1975Q1	0.4820078
1975Q2	0.4415189
1975Q3	0.3707502
1975Q4	0.4740234

Date	PredValue
1976Q1	0.6783065
1976Q2	0.8958555
1976Q3	0.9293666
1976Q4	0.9101117
1977Q1	0.8358955
1977Q2	0.8171266
1977Q3	0.6563341
1977Q4	0.7449771
1978Q1	0.7193016
1978Q2	0.7449771
1978Q3	0.7388500
1978Q4	0.6783065
1979Q1	0.7062002
1979Q2	0.8073683
1979Q3	0.8269266
1979Q4	0.6712844
1980Q1	0.5717513
1980Q2	0.7574960
1980Q3	0.8073683
1980Q4	0.8796373
1981Q1	0.8612845
1981Q2	0.8897314
1981Q3	0.7752157
1981Q4	0.7455466
1982Q1	0.6858956
1982Q2	0.7919962
1982Q3	0.8690972
1982Q4	0.8990760
1983Q1	0.8726950
1983Q2	0.8897314
1983Q3	0.6995172
1983Q4	0.4336437
1984Q1	0.4740234
1984Q2	0.5067496
1984Q3	0.6111146
1984Q4	0.7388500
1985Q1	0.7861714
1985Q2	0.8073683
1985Q3	0.6995172
1985Q4	0.6343676
1986Q1	0.5638987
1986Q2	0.6927484
1986Q3	0.7972185
1986Q4	0.8726950
1987Q1	0.6927484
1987Q2	0.8023425
1987Q3	0.8612845
1987Q4	0.9355955
1988Q1	0.8796373
1988Q2	0.6417574
1988Q3	0.5394180
1988Q4	0.8073683

Date	PredValue
1989Q1	0.8490283
1989Q2	0.9202684
1989Q3	0.8988035
1989Q4	0.8358955
1990Q1	0.7193016
1990Q2	0.6262120
1990Q3	0.6262120
1990Q4	0.6563341
1991Q1	0.6186920
1991Q2	0.6490805
1991Q3	0.6111146
1991Q4	0.6712844
1992Q1	0.6034832
1992Q2	0.3941263
1992Q3	0.4987500
1992Q4	0.6343676
1993Q1	0.7515694
1993Q2	0.7752157
1993Q3	0.8406403
1993Q4	0.8222987
1994Q1	0.8023425
1994Q2	0.5394180
1994Q3	0.3407396
1994Q4	0.2845506
1995Q1	0.3479642
1995Q2	0.3858004
1995Q3	0.5631608
1995Q4	0.4573539
1996Q1	0.3552594
1996Q2	0.4250686
1996Q3	0.4573539
1996Q4	0.2116514
1997Q1	0.2063613
1997Q2	0.4415189
1997Q3	0.7861714
1997Q4	0.8073683
1998Q1	0.8269266
1998Q2	0.7574960
1998Q3	0.6783065
1998Q4	0.5473576
1999Q1	0.3407396
1999Q2	0.3407396
1999Q3	0.3119540
1999Q4	0.6186920
2000Q1	0.8122962
2000Q2	0.9202684
2000Q3	0.9374971
2000Q4	0.8829844
2001Q1	0.6186920
2001Q2	0.3782459
2001Q3	0.2286411
2001Q4	0.2845506

Date	PredValue
2002Q1	0.5147457
2002Q2	0.8122962
2002Q3	0.9314386
2002Q4	0.9463941
2003Q1	0.9582336
2003Q2	0.9374971
2003Q3	0.8829844
2003Q4	0.8530844
2004Q1	0.7193016
2004Q2	0.5950785
2004Q3	0.5227343
2004Q4	0.5873449
2005Q1	0.8122962
2005Q2	0.8490283
2005Q3	0.8650637
2005Q4	0.8612845
2006Q1	0.8612845
2006Q2	0.8762079
2006Q3	0.8222987
2006Q4	0.8612845
2007Q1	0.8448806
2007Q2	0.7062002
2007Q3	0.5147457
2007Q4	0.5873449
2008Q1	0.5552732
2008Q2	0.3479642
2008Q3	0.3119540
2008Q4	0.4740234
2009Q1	0.3707502
2009Q2	0.2230465
2009Q3	0.3051272
2009Q4	0.3335883
2010Q1	0.2717029
2010Q2	0.2717029
2010Q3	0.1227115
2010Q4	0.1092914
2011Q1	0.1809388
2011Q2	0.2525610
2011Q3	0.4095079
2011Q4	0.6186920
2012Q1	0.6852489
2012Q2	0.5552732
2012Q3	0.4900013
2012Q4	0.6262120
2013Q1	0.6563341
2013Q2	0.6186920
2013Q3	0.7388500
2013Q4	0.8122962
2014Q1	0.8314586
2014Q2	0.8490283
2014Q3	0.7861714
2014Q4	0.4740234

Date	PredValue
2015Q1	0.3119540
2015Q2	0.3479642
2015Q3	NA
2015Q4	NA

	FALSE	TRUE
0	26	60
1	42	133

The hit rate = 0.6091

- d)

```
##
## Call:
## lm(formula = GrowthRate ~ LOGGDPLagOne + GrowthRateLagOne + T,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0162885 -0.0028512 -0.0005247  0.0024582  0.0152441
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -9.747e-02  3.758e-02  -2.594  0.01004 *
## LOGGDPLagOne    2.146e-02  8.147e-03   2.634  0.00896 **
## GrowthRateLagOne 6.222e-01  4.785e-02  13.003 < 2e-16 ***
## T              -6.939e-05  2.617e-05  -2.652  0.00850 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00465 on 258 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.4129, Adjusted R-squared:  0.4061
## F-statistic: 60.48 on 3 and 258 DF,  p-value: < 2.2e-16
```

- e)

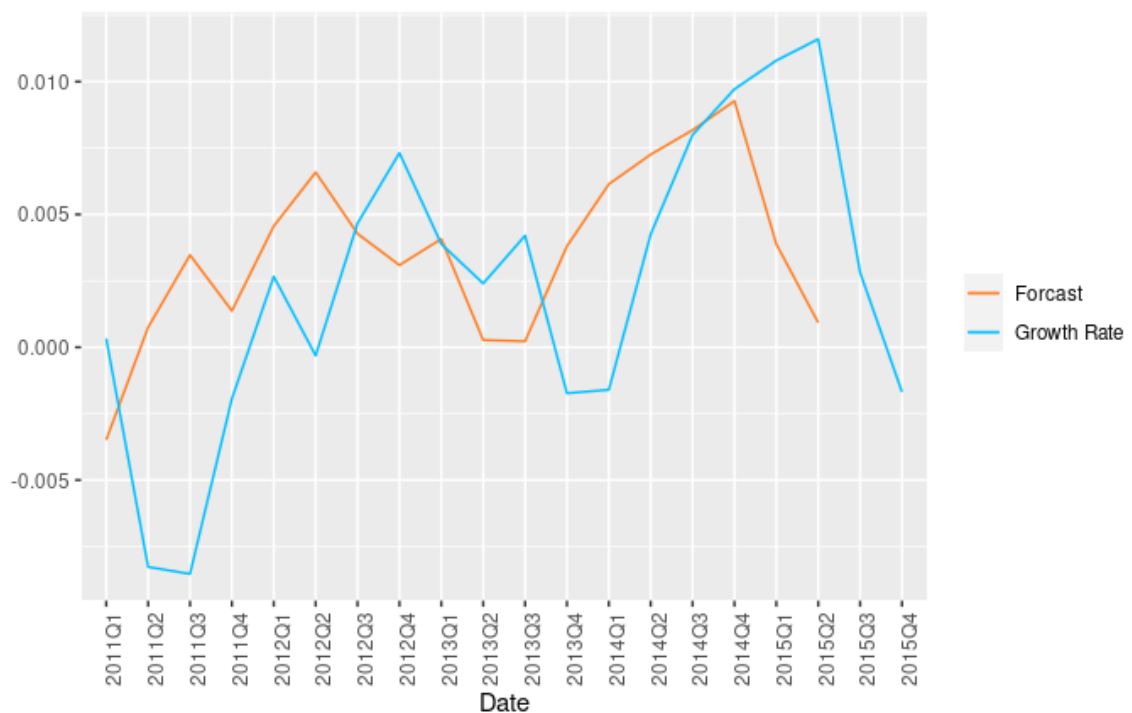
```
##
## Call:
## lm(formula = GrowthRate ~ GrowthRateLagOne + li1LagOne + li2LagOne,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0160034 -0.0025623 -0.0004464  0.0024590  0.0146213
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.305e-03  3.488e-04   3.741 0.000225 ***
## GrowthRateLagOne 6.486e-01  5.175e-02  12.533 < 2e-16 ***
## li1LagOne       1.306e-04  1.361e-04   0.960 0.338058
```



```
## li2LagOne      8.975e-05  7.093e-05  1.265 0.206885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004693 on 258 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.4021, Adjusted R-squared:  0.3951
## F-statistic: 57.83 on 3 and 258 DF,  p-value: < 2.2e-16
```

• f)

```
##
## Breusch-Godfrey test for serial correlation of order up to 1
##
## data: Error ~ GrowthRateLagOne + li1LagOne + li2LagOne + ErrorLagOne
## LM test = 111.44, df = 1, p-value < 2.2e-16
```



• g)

## Method

### Load the Data

```
data <- read.table('GDP_DATA.txt', sep = ',', header = T)
```

```
head(data)
```

```
##      Date      GDP GDPIMPR   LOGGDP   GrowthRate li1 li2 T
## 1 1950Q1  94.300      NA 4.546481      NA      0  0 0
## 2 1950Q2  95.200      1 4.555980 0.0094987522  0  0 1
## 3 1950Q3  97.663      1 4.581523 0.0255428352  3  1 2
## 4 1950Q4  99.728      1 4.602446 0.0209237031  4  2 3
## 5 1951Q1 100.445      1 4.609610 0.0071638339  2  1 4
## 6 1951Q2 100.406      0 4.609222 -0.0003883476  1  3 5
```

```

data$li1LagOne <- dplyr::lead(data$li1, n = 1)

data$li1LagTwo <- dplyr::lead(data$li1, n= 2)

data$li2LagOne <- dplyr::lead(data$li2, n = 1)

data$li2LagTwo <- dplyr::lead(data$li2, n= 2)

data$LOGGDPLagOne <- dplyr::lead(data$LOGGDP, n = 1)

data$GrowthRateLagOne <- dplyr::lead(data$GrowthRate, n =1)

```

```

data$PredValue <- exp((0.746-0.425*data$li1LagTwo) - (0.131*data$li2LagOne))/(1+exp((0.746-0.425*data$li1LagTwo) - (0.131*data$li2LagOne)))

```

```

head(data)

```

```

##      Date      GDP GDPIMPR   LOGGDP   GrowthRate li1 li2 T li1LagOne li1LagTwo
## 1 1950Q1  94.300      NA 4.546481      NA      0  0  0          0          3
## 2 1950Q2  95.200      1 4.555980  0.0094987522  0  0  1          3          4
## 3 1950Q3  97.663      1 4.581523  0.0255428352  3  1  2          4          2
## 4 1950Q4  99.728      1 4.602446  0.0209237031  4  2  3          2          1
## 5 1951Q1 100.445      1 4.609610  0.0071638339  2  1  4          1          1
## 6 1951Q2 100.406      0 4.609222 -0.0003883476  1  3  5          1         -1
##      li2LagOne li2LagTwo LOGGDPLagOne GrowthRateLagOne PredValue
## 1          0          1      4.555980      0.0094987522 0.3707502
## 2          1          2      4.581523      0.0255428352 0.2525610
## 3          2          1      4.602446      0.0209237031 0.4095079
## 4          1          3      4.609610      0.0071638339 0.5473576
## 5          3          3      4.609222     -0.0003883476 0.4820078
## 6          3          2      4.612007      0.0027847968 0.6852489

```

```

table <- table(ActualValue = data$GDPIMPR, PredictedValue = data$PredValue >0.5)

```

```

table

```

```

##           PredictedValue
## ActualValue FALSE TRUE
##           0      26   60
##           1      42  133

```

```

((table[1,1]+table[2,2])/sum(table))*100

```

```

## [1] 60.91954

```

```

model1 <- lm(formula = GrowthRate~LOGGDPLagOne+GrowthRateLagOne+T,data = data)

```

```

summary(model1)

```

```

##
## Call:
## lm(formula = GrowthRate ~ LOGGDPLagOne + GrowthRateLagOne + T,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max

```

```
## -0.0162885 -0.0028512 -0.0005247 0.0024582 0.0152441
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -9.747e-02  3.758e-02  -2.594  0.01004 *
## LOGGDPLagOne    2.146e-02  8.147e-03   2.634  0.00896 **
## GrowthRateLagOne 6.222e-01  4.785e-02  13.003 < 2e-16 ***
## T              -6.939e-05  2.617e-05  -2.652  0.00850 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00465 on 258 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.4129, Adjusted R-squared:  0.4061
## F-statistic: 60.48 on 3 and 258 DF,  p-value: < 2.2e-16
model2 <- lm( formula = GrowthRate~GrowthRateLagOne+li1LagOne+li2LagOne, data = data)
summary(model2)
```

```
##
## Call:
## lm(formula = GrowthRate ~ GrowthRateLagOne + li1LagOne + li2LagOne,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0160034 -0.0025623 -0.0004464  0.0024590  0.0146213
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.305e-03  3.488e-04   3.741 0.000225 ***
## GrowthRateLagOne 6.486e-01  5.175e-02  12.533 < 2e-16 ***
## li1LagOne       1.306e-04  1.361e-04   0.960 0.338058
## li2LagOne       8.975e-05  7.093e-05   1.265 0.206885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004693 on 258 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.4021, Adjusted R-squared:  0.3951
## F-statistic: 57.83 on 3 and 258 DF,  p-value: < 2.2e-16
model3 <- lm( formula = GrowthRate~GrowthRateLagOne+li1LagTwo+li2LagOne, data = data)
summary(model3)
```

```
##
## Call:
## lm(formula = GrowthRate ~ GrowthRateLagOne + li1LagTwo + li2LagOne,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0155389 -0.0026804 -0.0003973  0.0026503  0.0146979
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.292e-03  3.431e-04   3.765 0.000206 ***
## GrowthRateLagOne 6.472e-01  4.914e-02  13.170 < 2e-16 ***
## li1LagTwo     2.061e-04  1.294e-04   1.593 0.112382
## li2LagOne     9.178e-05  7.098e-05   1.293 0.197116
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004685 on 257 degrees of freedom
## (3 observations deleted due to missingness)
## Multiple R-squared:  0.4062, Adjusted R-squared:  0.3993
## F-statistic: 58.61 on 3 and 257 DF, p-value: < 2.2e-16
model4 <- lm( formula = GrowthRate~GrowthRateLagOne+li1LagOne+li2LagTwo, data = data)
summary(model4)
```

```
##
## Call:
## lm(formula = GrowthRate ~ GrowthRateLagOne + li1LagOne + li2LagTwo,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.015963 -0.002538 -0.000421  0.002538  0.014474
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.304e-03  3.514e-04   3.711 0.000253 ***
## GrowthRateLagOne 6.462e-01  5.180e-02  12.474 < 2e-16 ***
## li1LagOne     1.273e-04  1.363e-04   0.934 0.351414
## li2LagTwo     8.749e-05  7.113e-05   1.230 0.219806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0047 on 257 degrees of freedom
## (3 observations deleted due to missingness)
## Multiple R-squared:  0.4025, Adjusted R-squared:  0.3955
## F-statistic: 57.7 on 3 and 257 DF, p-value: < 2.2e-16
model5 <- lm( formula = GrowthRate~GrowthRateLagOne+li1LagTwo+li2LagTwo, data = data)
summary(model5)
```

```
##
## Call:
## lm(formula = GrowthRate ~ GrowthRateLagOne + li1LagTwo + li2LagTwo,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0154964 -0.0027235 -0.0003078  0.0027173  0.0145271
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.302e-03  3.441e-04   3.783 0.000193 ***
## GrowthRateLagOne 6.440e-01  4.915e-02  13.102 < 2e-16 ***
## li1LagTwo      2.054e-04  1.293e-04   1.589 0.113389
## li2LagTwo      9.205e-05  7.079e-05   1.300 0.194665
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004685 on 257 degrees of freedom
## (3 observations deleted due to missingness)
## Multiple R-squared:  0.4063, Adjusted R-squared:  0.3994
## F-statistic: 58.62 on 3 and 257 DF, p-value: < 2.2e-16

dataResidual <- data.frame(model2$residuals)

rownames(dataResidual) <- 1:262
colnames(dataResidual) <- 'Error'

dataResidual <- tibble::add_row(dataResidual, Error = NA)
dataResidual <- tibble::add_row(dataResidual, Error = NA)

data <- cbind(data, dataResidual)

data$ErrorLagOne <- dplyr::lead(data$Error, n = 1)

lmtest::bptest(Error ~ GrowthRateLagOne + li1LagOne + li2LagOne +
  ErrorLagOne, data = data, order = 1)

##
## Breusch-Godfrey test for serial correlation of order up to 1
##
## data: Error ~ GrowthRateLagOne + li1LagOne + li2LagOne + ErrorLagOne
## LM test = 111.44, df = 1, p-value < 2.2e-16

dataForecast <- data.frame(model2$fitted.values)

rownames(dataForecast) <- 1:262
colnames(dataForecast) <- 'Forecast'

dataForecast <- tibble::add_row(dataForecast, Forecast = NA)
dataForecast <- tibble::add_row(dataForecast, Forecast = NA)

data <- cbind(data, dataForecast)

head(data)
```

	Date	GDP	GDPIMPR	LOGGDP	GrowthRate	li1	li2	T	li1LagOne	li1LagTwo
## 1	1950Q1	94.300	NA	4.546481	NA	0	0	0	0	3
## 2	1950Q2	95.200	1	4.555980	0.0094987522	0	0	1	3	4
## 3	1950Q3	97.663	1	4.581523	0.0255428352	3	1	2	4	2
## 4	1950Q4	99.728	1	4.602446	0.0209237031	4	2	3	2	1
## 5	1951Q1	100.445	1	4.609610	0.0071638339	2	1	4	1	1
## 6	1951Q2	100.406	0	4.609222	-0.0003883476	1	3	5	1	-1
##	li2LagOne	li2LagTwo	LOGGDLagOne	GrowthRateLagOne	PredValue	Error				
## 1	0	1	4.555980	0.0094987522	0.3707502	-0.008854881				

```
## 2      1      2      4.581523      0.0255428352 0.2525610 0.009964787
## 3      2      1      4.602446      0.0209237031 0.4095079 0.014621297
## 4      1      3      4.609610      0.0071638339 0.5473576 0.005710901
## 5      3      3      4.609222     -0.0003883476 0.4820078 -0.003899377
## 6      3      2      4.612007      0.0027847968 0.6852489 0.001160494
```

```
##      ErrorLagOne      Forecast
## 1 0.0099647866 0.018353633
## 2 0.0146212966 0.015578049
## 3 0.0057109010 0.006302407
## 4 -0.0038993775 0.001452933
## 5 0.0011604935 0.003511030
## 6 -0.0005007337 0.001624303
```

```
dataGraphing <- data[,c("Date", "GrowthRate", "Forecast")]
```

```
dataGraphing <- tail(dataGraphing, 20)
```

```
head(dataGraphing)
```

```
##      Date      GrowthRate      Forecast
## 245 2011Q1 0.0003231626 -0.0034871148
## 246 2011Q2 -0.0082712134 0.0007503807
## 247 2011Q3 -0.0085337293 0.0034721465
## 248 2011Q4 -0.0019373826 0.0013699363
## 249 2012Q1 0.0026618100 0.0045612119
## 250 2012Q2 -0.0003149040 0.0065867908
```

```
library(ggplot2)
ggplot(data = dataGraphing, aes(x = Date, group = 1))+
  geom_line(aes(y = Forecast, col = 'Forecast'))+
  geom_line(aes(y = GrowthRate, col = 'Growth Rate'))+
  labs(x = "Date", y = "")+
  theme(axis.text.x = element_text(angle = 90), legend.title = element_blank())+
  scale_color_manual(breaks = c("Forecast", "Growth Rate"),
                    values = c('chocolate1', "deepskyblue1"))
```

