

Replication 2

Jiayi Lily Ma (jm4303)

5/10/2019

Paper's code, paper's data

```
library(dyn)

## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
dat <- read.csv("merged.csv")
d <- zoo(dat[, -1], as.Date(dat[, 1]))

y <- log(d$sales)
x <- d[, c(2, 3)]
```

Paper's base model

```
#base model
reg0 <- dyn$lm(y~lag(y,-1)+lag(y,-12))
summary(reg0)

##
## Call:
## lm(formula = dyn(y ~ lag(y, -1) + lag(y, -12)))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.209554 -0.034684  0.002482  0.040477  0.220976
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.67266    0.76355   0.881 0.381117
## lag(y, -1)   0.64345    0.07332  8.776 3.59e-13 ***
## lag(y, -12)  0.29565    0.07282  4.060 0.000118 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07985 on 76 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.7185, Adjusted R-squared:  0.7111
## F-statistic:    97 on 2 and 76 DF,  p-value: < 2.2e-16
```

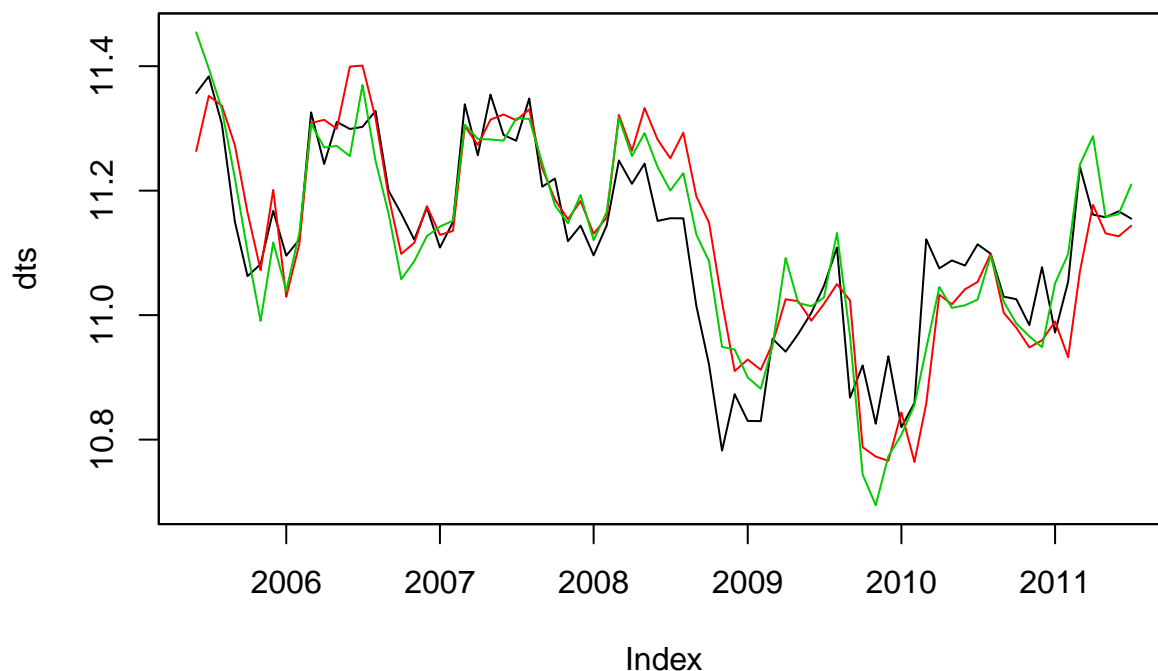
```

#trend model
reg1 <- dyn$lm(y~lag(y,-1)+lag(y,-12)+suvs+insurance,data=dat)
summary(reg1)

##
## Call:
## lm(formula = dyn(y ~ lag(y, -1) + lag(y, -12) + suvs + insurance),
##     data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.161327 -0.043774  0.002998  0.036651  0.159219
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.45798    0.78438  -0.584 0.561081
## lag(y, -1)   0.61947    0.06318   9.805 5.09e-15 ***
## lag(y, -12)  0.42865    0.06535   6.559 6.45e-09 ***
## suvs         1.05721    0.16686   6.336 1.66e-08 ***
## insurance    -0.52966    0.15206  -3.483 0.000835 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06509 on 74 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.8179, Adjusted R-squared:  0.808
## F-statistic: 83.08 on 4 and 74 DF,  p-value: < 2.2e-16

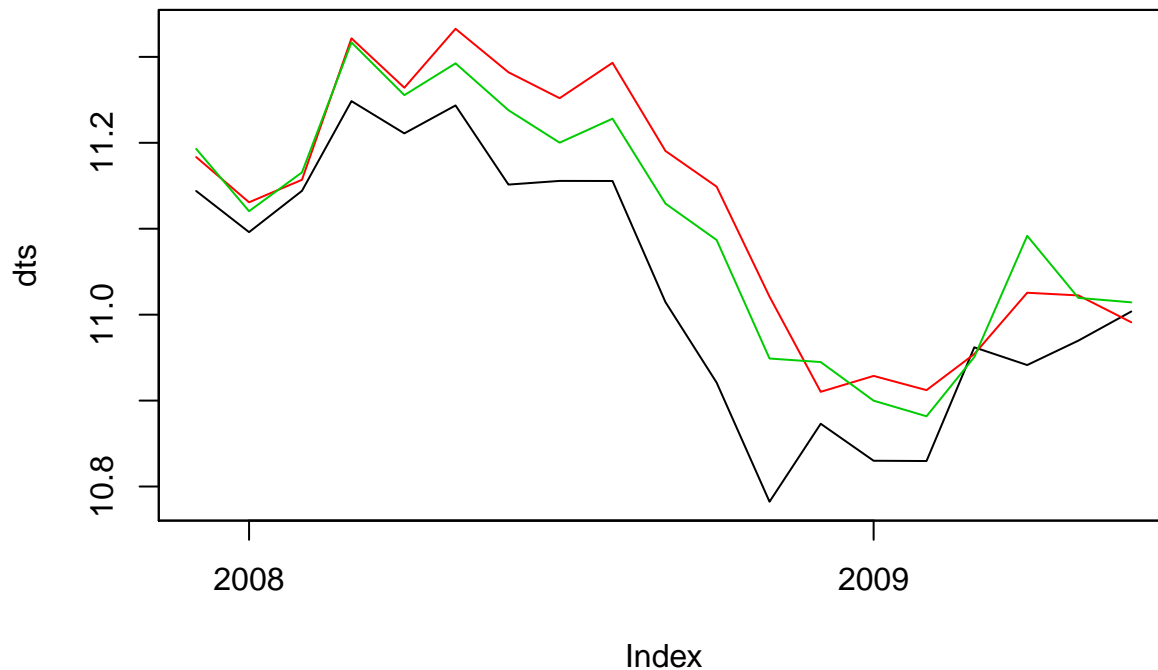
source("../oosf.R")
z <- OutOfSampleForecast12(y,x,17)
#overall fit
MaeReport(z)

```



```
##   mae.base mae.trends mae.delta
## 0.06343984 0.05667658 0.10660890
```

```
MaeReport(z, "2007-12-01", "2009-06-30")
```



```
##   mae.base mae.trends mae.delta
## 0.08869325 0.06965812 0.21461753
```

Paper's code, our data 2014-2019

```
merged1419 <- read.csv("merged_14_19.csv")
d1419 <- zoo(merged1419[, -1], as.Date(merged1419[, 1]))
```

```
y1419 <- log(d1419$sales)
x1419 <- d1419[, c(2, 3)]
```

```
#base model
```

```
reg0_14_19 <- dyn$lm(y1419 ~ lag(y1419, -1) + lag(y1419, -12), data=merged1419)
summary(reg0_14_19)
```

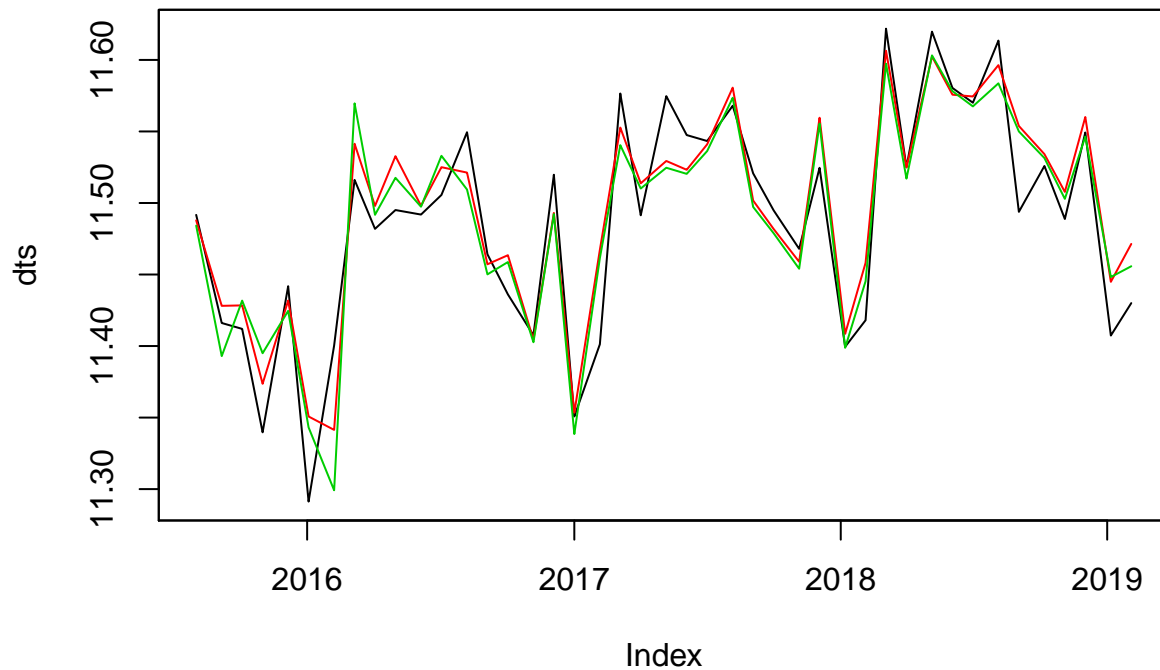
```
##
## Call:
## lm(formula = dyn(y1419 ~ lag(y1419, -1) + lag(y1419, -12)), data = merged1419)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.056864 -0.011243  0.004109  0.017806  0.042493
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.12923    0.60807   3.502  0.00106 **
## lag(y1419, -1) -0.07381    0.05181  -1.425  0.16119
## lag(y1419, -12)  0.89178    0.05061  17.621 < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02472 on 45 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.8912, Adjusted R-squared:  0.8863
## F-statistic: 184.3 on 2 and 45 DF,  p-value: < 2.2e-16

#trend model
reg1_14_19 <- dyn$lm(y1419~lag(y1419,-1)+lag(y1419,-12)+suvs+insurance,data=merged1419)
summary(reg1_14_19)

##
## Call:
## lm(formula = dyn(y1419 ~ lag(y1419, -1) + lag(y1419, -12) + suvs +
##      insurance), data = merged1419)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.051344 -0.013505  0.004453  0.014885  0.052805
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9894296   0.7522033   1.315   0.195
## lag(y1419, -1) -0.0544858   0.0508947  -1.071   0.290
## lag(y1419, -12)  0.9757246   0.0595366  16.389 <2e-16 ***
## suvs           -0.0019046   0.0008538  -2.231   0.031 *
## insurance       0.0014033   0.0009143   1.535   0.132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02371 on 43 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.9044, Adjusted R-squared:  0.8955
## F-statistic: 101.7 on 4 and 43 DF,  p-value: < 2.2e-16

source("../oosf.R")
z1419 <- OutOfSampleForecast12(y1419,x1419,17)
MaeReport(z1419)
```



```
##      mae.base  mae.trends  mae.delta
## 0.02215164  0.02470196 -0.11513004
```

rescale Google trend's data using $\log(x/100)$, data 2014-2019

```
head(merged1419)
```

```
##      Period sales suvs insurance
## 1 2014/03/02 90488   65         76
## 2 2014/04/06 87959   62         70
## 3 2014/05/04 93239   60         69
## 4 2014/06/01 86715   59         72
## 5 2014/07/06 91275   63         70
## 6 2014/08/03 92624   65         73
```

```
suvs_rescaled <- log(merged1419$suvs/100)
insurance_rescaled <- log(merged1419$insurance/100)
```

```
merged1419$suvs_rescaled <- suvs_rescaled
merged1419$insurance_rescaled <- insurance_rescaled
head(merged1419)
```

```
##      Period sales suvs insurance suvs_rescaled insurance_rescaled
## 1 2014/03/02 90488   65         76   -0.4307829   -0.2744368
## 2 2014/04/06 87959   62         70   -0.4780358   -0.3566749
## 3 2014/05/04 93239   60         69   -0.5108256   -0.3710637
## 4 2014/06/01 86715   59         72   -0.5276327   -0.3285041
## 5 2014/07/06 91275   63         70   -0.4620355   -0.3566749
## 6 2014/08/03 92624   65         73   -0.4307829   -0.3147107
```

```
d1419 <- zoo(merged1419[, -1], as.Date(merged1419[, 1]))
```

```

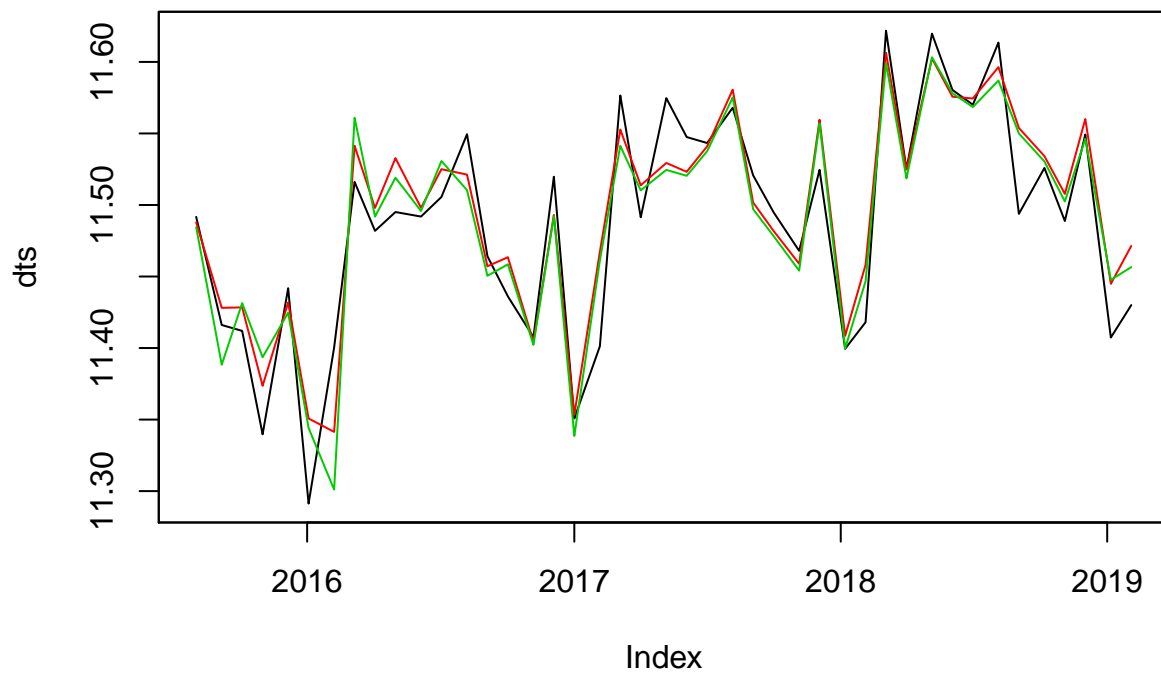
y1419 <- log(d1419$sales)
x1419 <- d1419[,c(4,5)]

#trend model
reg1419_rescaled <- dyn$lm(y1419~lag(y1419,-1)+lag(y1419,-12)+suvs_rescaled+insurance_rescaled,data=merged1419)
summary(reg1419_rescaled)

##
## Call:
## lm(formula = dyn(y1419 ~ lag(y1419, -1) + lag(y1419, -12) + suvs_rescaled +
##      insurance_rescaled), data = merged1419)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.050783 -0.013022  0.004092  0.014817  0.053273
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.96517    0.75442   1.279   0.2076
## lag(y1419, -1)   -0.05388    0.05073  -1.062   0.2942
## lag(y1419, -12)   0.97297    0.05842  16.656 <2e-16 ***
## suvs_rescaled    -0.14685    0.06659  -2.205   0.0328 *
## insurance_rescaled 0.11111    0.07242   1.534   0.1323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02365 on 43 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.9048, Adjusted R-squared:  0.896
## F-statistic: 102.2 on 4 and 43 DF,  p-value: < 2.2e-16

source("../oosf.R")
z1419_rescaled <- OutOfSampleForecast12(y1419,x1419,17)
MaeReport(z1419_rescaled)

```



```
##      mae.base  mae.trends  mae.delta
## 0.02215164  0.02431243 -0.09754551
```