BUAN 6312 - Econometrics Final Project

Team Members:
1. Prithvi Shivshankar (PXS170023)
2. Dinesh Venkata Ganta (DXG180016)
3. Ammrutha Shridharan (AXS180025)
4. Sri Kalyan Pesala (SXP180043)

Introduction

In this project, we have done an in-depth analysis in order to answer the research question "How do drunk driving laws affect traffic deaths". Initially we conduct a basic exploratory data analysis to gain an understanding of what the data is telling us. We then formulate different economic theories based on the data and choose relevant data concerned with the research hypothesis. As part of the process, we check for cases of endogeneity as we would not be able to proceed with the model selection unless those cases have been handled. We then run different regression models and come up with the best case solution to answer the question appropriately. The initial and primary model of concern we have chosen is for the case of alcohol involved vehicle fatality rates, because that forms the crux of this research. We then check the cases for vehicle fatality rates as a whole and the fatality rates for each group in order to gain a better understanding as to how the laws affect the traffic deaths. After this, forming a holistic picture, we present a conclusion.

Contents:

1. Initial exploratory data analysis

Initial summary statistics checked on every variables part of the data set.

```
    state          year         spircons        unrate          perinc     
 AL     :  7   Min.   :1982   Min.   :0.790   Min.   : 2.400   Min.   : 9514  
 AZ     :  7   1st Qu.:1983   1st Qu.:1.300   1st Qu.: 5.475   1st Qu.:12086  
 AR     :  7   Median :1985   Median :1.670   Median : 7.000   Median :13763  
 CA     :  7   Mean   :1985   Mean   :1.754   Mean   : 7.347   Mean   :13880  
 CO     :  7   3rd Qu.:1987   3rd Qu.:2.013   3rd Qu.: 8.900   3rd Qu.:15175  
 CT     :  7   Max.   :1988   Max.   :4.900   Max.   :18.000   Max.   :22193  
 (Other):294                                                                
    beertax            sobapt           mormon             mlda      
 Min.   :0.04331   Min.   : 0.0000   Min.   : 0.1000   Min.   :18.00  
 1st Qu.:0.20885   1st Qu.: 0.6268   1st Qu.: 0.2722   1st Qu.:20.00  
 Median :0.35259   Median : 1.7492   Median : 0.3931   Median :21.00  
 Mean   :0.51326   Mean   : 7.1569   Mean   : 2.8019   Mean   :20.46  
 3rd Qu.:0.65157   3rd Qu.:13.1271   3rd Qu.: 0.6293   3rd Qu.:21.00  
 Max.   :2.72076   Max.   :30.3557   Max.   :65.9165   Max.   :21.00  

      dry              yngdrv            vmiles          jaild       
 Min.   : 0.00000   Min.   :0.07314   Min.   : 4576   Min.   :0.0000  
 1st Qu.: 0.00000   1st Qu.:0.17037   1st Qu.: 7183   1st Qu.:0.0000  
 Median : 0.08681   Median :0.18539   Median : 7796   Median :0.0000  
 Mean   : 4.26707   Mean   :0.18593   Mean   : 7891   Mean   :0.2806  
 3rd Qu.: 2.42481   3rd Qu.:0.20219   3rd Qu.: 8504   3rd Qu.:1.0000  
 Max.   :45.79210   Max.   :0.28163   Max.   :26148   Max.   :1.0000  
                                                      NA's   :1       
    comserd          allmort           mrall              allnite       
 Min.   :0.0000   Min.   :  79.0   Min.   :8.212e-05   Min.   :  13.00  
 1st Qu.:0.0000   1st Qu.: 293.8   1st Qu.:1.624e-04   1st Qu.:  53.75  
 Median :0.0000   Median : 701.0   Median :1.956e-04   Median : 135.00  
 Mean   :0.1851   Mean   : 928.7   Mean   :2.040e-04   Mean   : 182.58  
 3rd Qu.:0.0000   3rd Qu.:1063.5   3rd Qu.:2.418e-04   3rd Qu.: 212.00  
 Max.   :1.0000   Max.   :5504.0   Max.   :4.218e-04   Max.   :1049.00  
 NA's   :1                                                              
     mralln             allsvn           a1517           mra1517         
 Min.   :1.716e-05   Min.   :  8.0   Min.   :  3.00   Min.   :0.0001163  
 1st Qu.:3.176e-05   1st Qu.: 35.0   1st Qu.: 25.75   1st Qu.:0.0002358  
 Median :3.707e-05   Median : 81.0   Median : 49.00   Median :0.0002956  
 Mean   :3.875e-05   Mean   :109.9   Mean   : 62.61   Mean   :0.0003034  
 3rd Qu.:4.466e-05   3rd Qu.:131.0   3rd Qu.: 77.00   3rd Qu.:0.0003575  
 Max.   :9.436e-05   Max.   :603.0   Max.   :318.00   Max.   :0.0006735  

     a1517n         mra1517n             a1820           a1820n      
 Min.   : 0.00   Min.   :0.000e+00   Min.   :  7.0   Min.   :  0.00  
 1st Qu.: 4.00   1st Qu.:3.925e-05   1st Qu.: 38.0   1st Qu.: 11.00  
 Median :10.00   Median :5.283e-05   Median : 82.0   Median : 24.00  
 Mean   :12.26   Mean   :5.977e-05   Mean   :106.7   Mean   : 33.53  
 3rd Qu.:15.25   3rd Qu.:7.239e-05   3rd Qu.:130.2   3rd Qu.: 44.00  
 Max.   :76.00   Max.   :2.571e-04   Max.   :601.0   Max.   :196.00  

    mra1820            mra1820n             a2124          mra2124         
 Min.   :0.0001855   Min.   :0.0000000   Min.   : 12.0   Min.   :0.0002000  
 1st Qu.:0.0003722   1st Qu.:0.0001056   1st Qu.: 42.0   1st Qu.:0.0003208  
 Median :0.0004524   Median :0.0001329   Median : 97.5   Median :0.0004000  
 Mean   :0.0004728   Mean   :0.0001436   Mean   :126.9   Mean   :0.0004091  
 3rd Qu.:0.0005429   3rd Qu.:0.0001721   3rd Qu.:150.5   3rd Qu.:0.0004730  
 Max.   :0.0010952   Max.   :0.0005238   Max.   :770.0   Max.   :0.0008922  

     a2124n          mra2124n             aidall            mraidall        
 Min.   :  1.00   Min.   :2.222e-05   Min.   :  24.6   Min.   :2.337e-05  
 1st Qu.: 13.00   1st Qu.:1.032e-04   1st Qu.:  90.5   1st Qu.:4.849e-05  
 Median : 30.00   Median :1.240e-04   Median : 211.6   Median :6.335e-05  
 Mean   : 41.38   Mean   :1.284e-04   Mean   : 293.3   Mean   :6.593e-05  
 3rd Qu.: 49.00   3rd Qu.:1.459e-04   3rd Qu.: 364.0   3rd Qu.:7.643e-05  
 Max.   :249.00   Max.   :3.143e-04   Max.   :2094.9   Max.   :1.772e-04  
```

```
      pop               pop1517           pop1820           pop2124
 Min.   :  479000   Min.   :  21000   Min.   :  21000   Min.   :  30000
 1st Qu.: 1545251   1st Qu.:  71750   1st Qu.:  76962   1st Qu.: 103500
 Median : 3310503   Median : 163000   Median : 170982   Median : 241000
 Mean   : 4930272   Mean   : 230816   Mean   : 249090   Mean   : 336390
 3rd Qu.: 5751735   3rd Qu.: 270500   3rd Qu.: 308311   3rd Qu.: 413000
 Max.   :28314028   Max.   :1172000   Max.   :1321004   Max.   :1892998

     miles             gspch
 Min.   :  3993   Min.   :-0.123641
 1st Qu.: 11692   1st Qu.: 0.001182
 Median : 28484   Median : 0.032413
 Mean   : 37101   Mean   : 0.025313
 3rd Qu.: 44140   3rd Qu.: 0.056501
 Max.   :241575   Max.   : 0.142361
```

The initial finding was that the community service(comserd) and mandatory jail sentence(jaild) has a missing value. This data set is a case of balanced panel data. Via close analysis, the missing data value was imputed by adding the value from previous variables, since both comserd and jaild did not have much of a change over time except for very few cases. Further analysis led to the conclusion that comserd and jaild are slow changing variables over time. Apart from this, every other variable in the dataset is time-variant.

In [4]:
```
car_fatalities <- read.csv(file="car_fatalities_test.csv")
ctest <- read.csv("car_fatalities_test.csv")
x <- ctest %>% select(year,state,jaild) %>% group_by(state) %>% mutate(sumJ=sum(jaild))
table(x$sumJ)/7
```

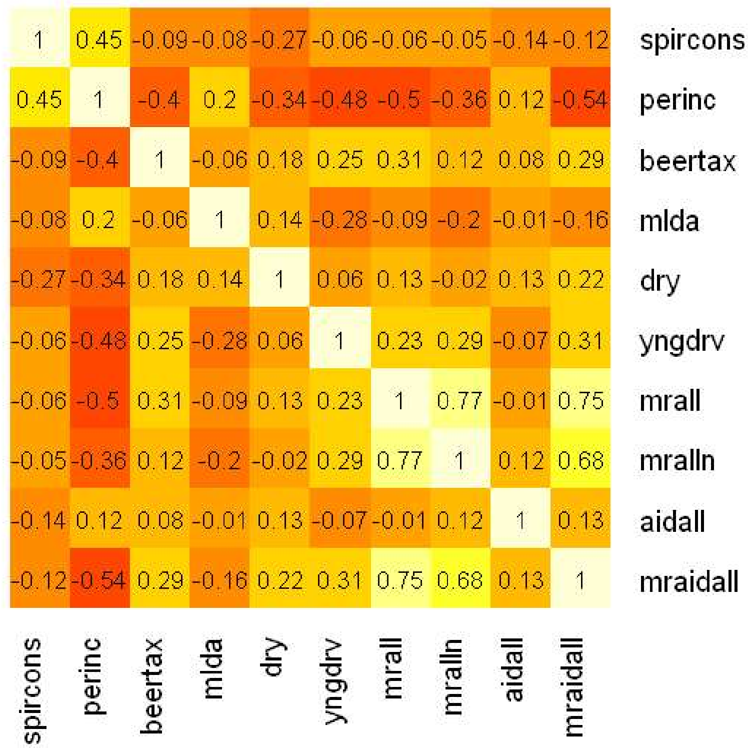
```
 0  4  5  6  7
33  2  1  3  9
```

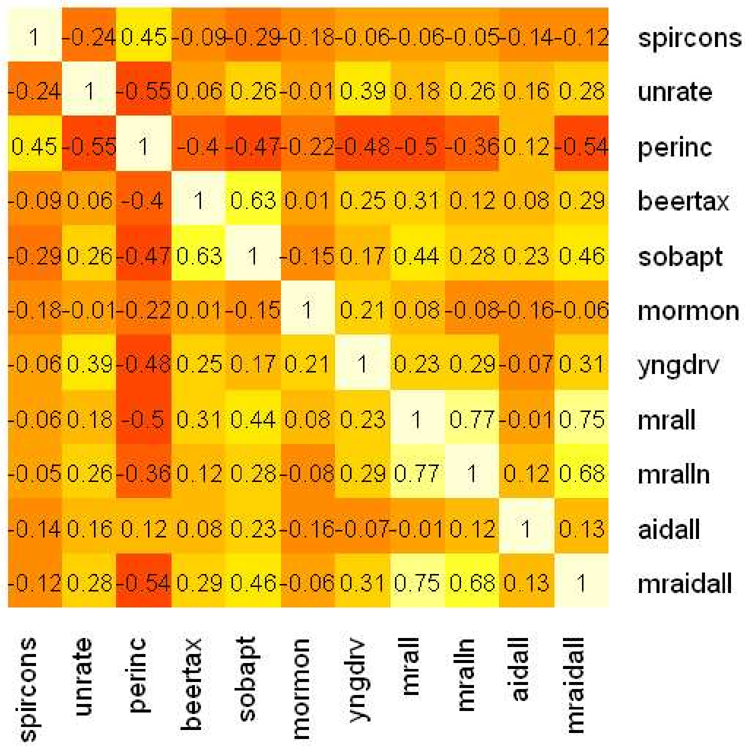
As per our findings, we see that a total of six states have cases where the jaild(mandatory jail sentence) and comserd(mandatory community service) change over time.

Further analysis of the data led to more research into the case to figure out why the percentages for Mormon and Southern Baptist were part of the dataset. From an economic standpoint, Mormons and Southern Baptists are religious by nature and are against consuming alcohol and this coincides with the economic theory, where they would be against the consumption and would push the government to take more stringent steps against its consumption.

2. Understanding the data

In [19]:
```
correlation_Car_test <- car_fatalities %>% dplyr::select(-X,-jaild,-comserd,-state,-year
,-pop,-pop1517,-pop1820,-pop2124,-vmiles,-allmort,-allnite,-allsvn,-a1517,-a1517n,-a1820
,-a1820n,-a2124,-a2124n,-mra1517n,-mra1820n,-mra2124n,-miles,-gspch,-mra1517,-mra1820,-m
ra2124,-sobapt,-mormon,-unrate)
correlation_Car_test1 <- car_fatalities %>% dplyr::select(-X,-jaild,-comserd,-state,-yea
r,-pop,-pop1517,-pop1820,-pop2124,-vmiles,-allmort,-allnite,-allsvn,-a1517,-a1517n,-a182
0,-a1820n,-a2124,-a2124n,-mra1517n,-mra1820n,-mra2124n,-miles,-gspch,-mra1517,-mra1820,-
mra2124,-dry,-mlda)

library(gplots)
heatmap.2(cor(correlation_Car_test), Rowv = FALSE, Colv = FALSE, dendrogram = "none",
          cellnote = round(cor(correlation_Car_test),2),
          notecol = "black", key = FALSE, trace = 'none', margins = c(10,10))
heatmap.2(cor(correlation_Car_test1), Rowv = FALSE, Colv = FALSE, dendrogram = "none",
          cellnote = round(cor(correlation_Car_test1),2),
          notecol = "black", key = FALSE, trace = 'none', margins = c(10,10))
```

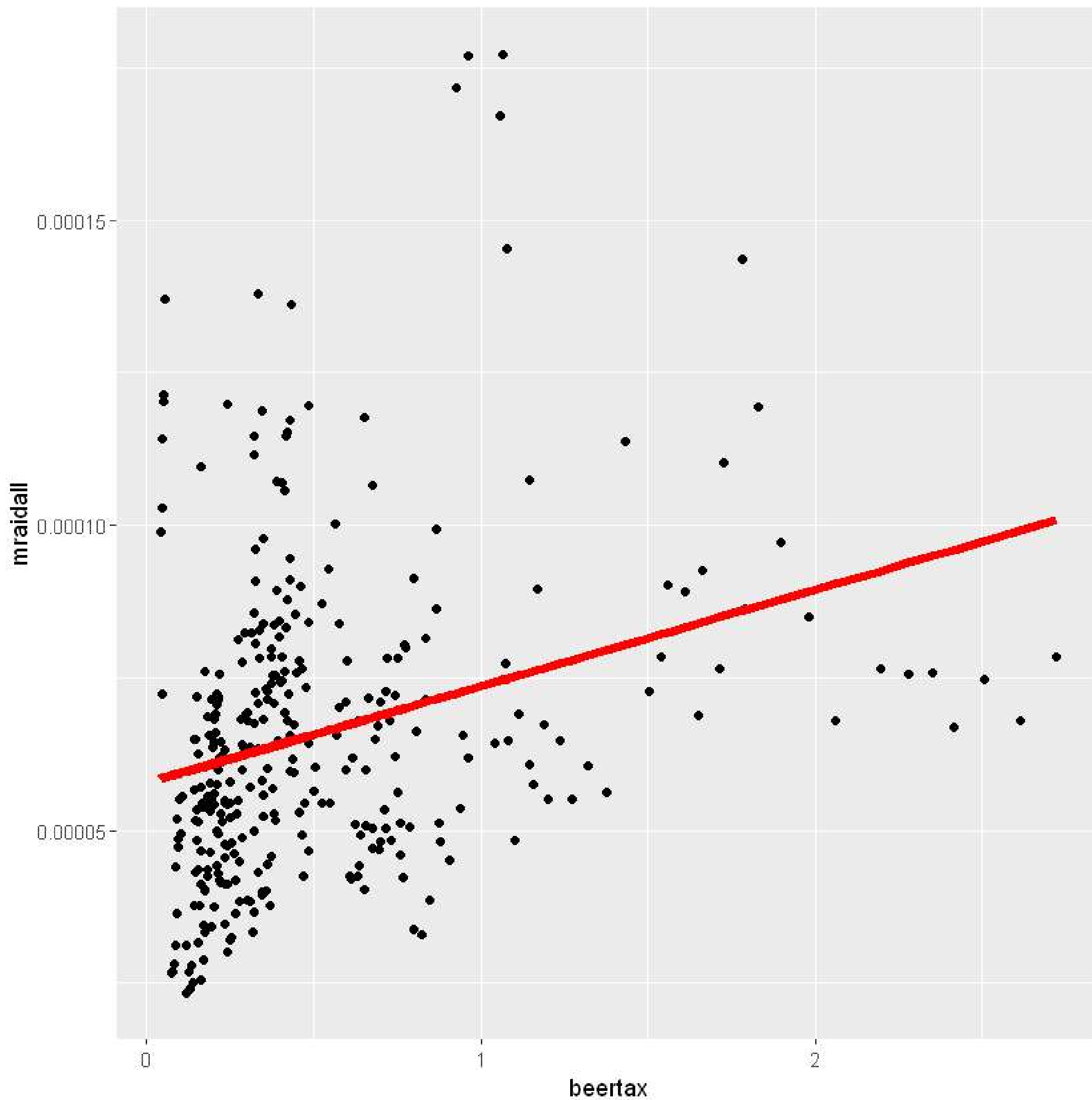

1 0.45 -0.09 -0.08 -0.27 -0.06 -0.06 -0.05 -0.14 -0.12 spircons
0.45 1 -0.4 0.2 -0.34 -0.48 -0.5 -0.36 0.12 -0.54 perinc
-0.09 -0.4 1 -0.06 0.18 0.25 0.31 0.12 0.08 0.29 beertax
-0.08 0.2 -0.06 1 0.14 -0.28 -0.09 -0.2 -0.01 -0.16 mlda
-0.27 -0.34 0.18 0.14 1 0.06 0.13 -0.02 0.13 0.22 dry
-0.06 -0.48 0.25 -0.28 0.06 1 0.23 0.29 -0.07 0.31 yngdrv
-0.06 -0.5 0.31 -0.09 0.13 0.23 1 0.77 -0.01 0.75 mrall
-0.05 -0.36 0.12 -0.2 -0.02 0.29 0.77 1 0.12 0.68 mralln
-0.14 0.12 0.08 -0.01 0.13 -0.07 -0.01 0.12 1 0.13 aidall
-0.12 -0.54 0.29 -0.16 0.22 0.31 0.75 0.68 0.13 1 mraidall
spircons
perinc
beertax
mlda
dry
yngdrv
mrall
mralln
aidall
mraidall



The correlation matrix was checked and in each case, the results found were consistent with the economic theory and based on this, the cases with high correlation were not considered. In order to gain a better understanding of the data to well answer the research question, "How do drunk driving laws affect traffic deaths?", an initial scatter plot between the beer tax(beertax) variable and the alcohol involved vehicle fatality rate(mraidall) along with the linear regression line was done.

In [114]:
```
p1<-lm(mraidall~beertax,data = car_fatalities)
coeftest(p1,vcov.=vcovHC)
plot1 <- ggplot(car_fatalities, aes(x=beertax,y=mraidall)) + geom_point()
plot1 <- plot1 + geom_line(aes(y=predict(lm(mraidall~beertax,data=car_fatalities))),colo
r="red",size=2)
plot1
```

```
t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)    
(Intercept) 5.7812e-05 1.9083e-06 30.2948 < 2.2e-16 ***
beertax     1.5811e-05 3.0433e-06  5.1953 3.563e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



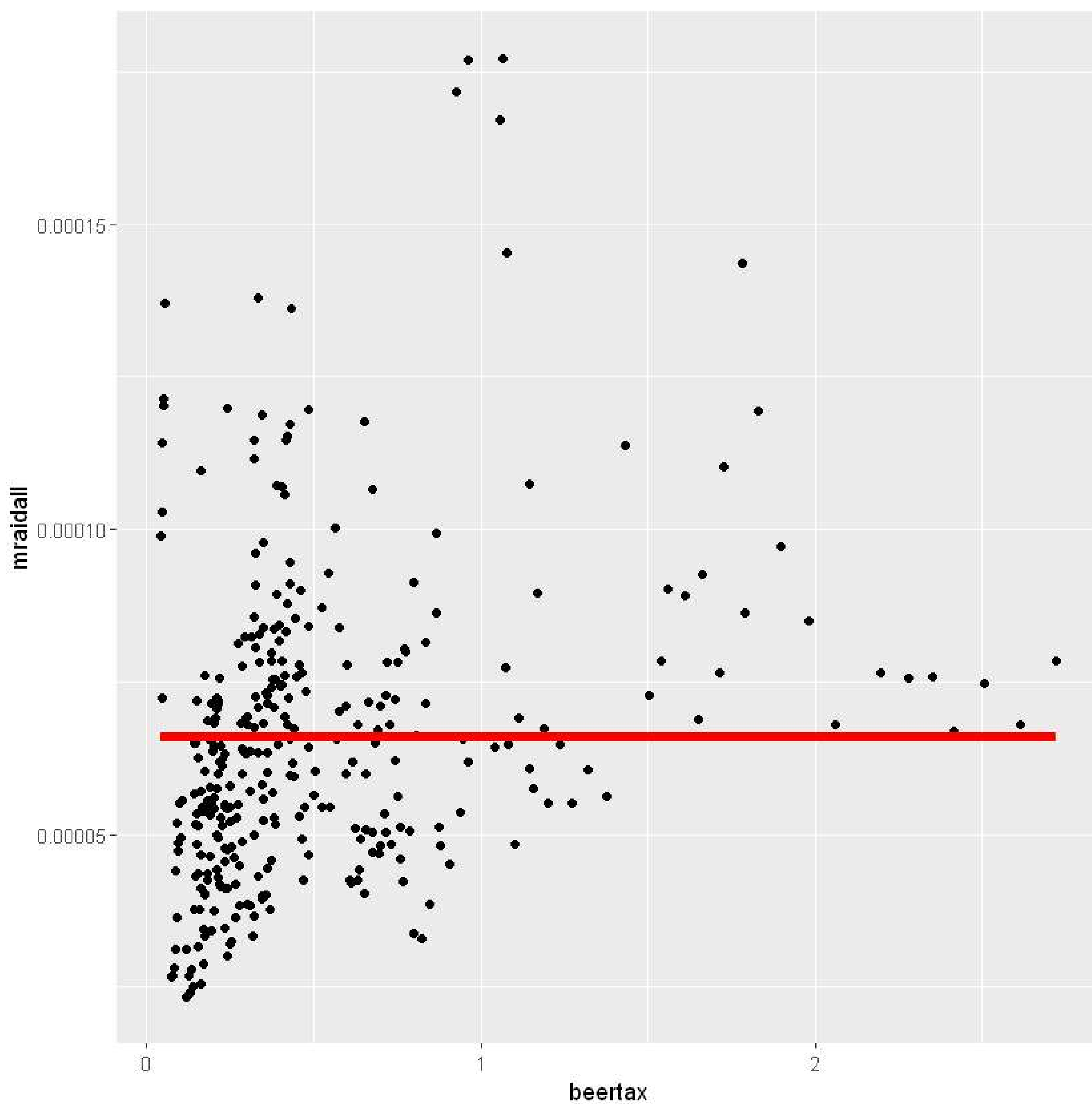
The plot shows that as beertax increases, on an average, the alcohol involved vehicle fatality rate increases. This is not consistent with the economic theory because as beer tax increases, there should be a decrease in alchohol induced vehicle fatality rates. This could be due to the case of endogeneity and simultaneous causality bias, where both the beertax and the alcohol vehicle fatality rates are being determined within the system. Beer taxes could be imposed in states where there are more number of alcohol related traffic deaths and more number of people drive. Another factor to be considered is the case of omitted variable bias, where the qualitative variable, that is, the cultural attitude towards drinking, has not been taken into account. In oder to deal with the omitted variable bias, a suitable set of instrumental variables are chosen. In this case, the mormon(% of Mormon) and the sobapt(% of southern Baptist) could be suitable instrumental variables as they, to an extent, capture the effect of the cultural attitude towards drinking because those two religious groups as discussed earlier are against alcohol consumption. To choose the right instrumental variable, the conditions to be met are:

- **They should have a high correlation with the endogenous regressor, which is true in this case, as the two religious groups would push the goverment to take more stringent steps against alcohol and in the process, there could be an increase in the beer tax**
- **The instrumental variables should not be correlated with the error term**
- **The instrumental variables should not have a direct effect on the dependant variable, which coincides with our case because the percentage of Mormons and Southern Baptists do not directly influence the alcohol involved vehicle fatailty rate**

In [5]:

```
car_fatalities <- car_fatalities %>% pdata.frame(index=c('state','year'))
summary(plm(mraidall~beertax | mormon + sobapt,data = car_fatalities))
plot3 <- ggplot(car_fatalities, aes(x=beertax,y=mraidall)) + geom_point()
plot3 <- plot3 + geom_line(aes(y=predict(lm(mraidall~beertax | mormon + sobapt,data=car_
fatalities))),color="red",size=2)
plot3
```

```
Oneway (individual) effect Within Model
Instrumental variable estimation
   (Balestra-Varadharajan-Krishnakumar's transformation)

Call:
plm(formula = mraidall ~ beertax | mormon + sobapt, data = car_fatalities)

Balanced Panel: n = 48, T = 7, N = 336

Residuals:
       Min.      1st Qu.       Median      3rd Qu.         Max.
-9.3555e-05 -5.9016e-06 -2.6175e-07  4.5899e-06  4.7739e-05

Coefficients:
           Estimate  Std. Error t-value Pr(>|t|)
beertax -1.4271e-06  2.6313e-05 -0.0542   0.9568

Total Sum of Squares:    5.3896e-08
Residual Sum of Squares: 5.3879e-08
R-Squared:      0.00080499
Adj. R-Squared: -0.16631
F-statistic: 0.090133 on 1 and 287 DF, p-value: 0.76423
```



As per the result, it is observed that the beertax now has a negative effect on the alcohol involved vehicle fatality rate, which is consistent with the economic theory. But, the variable is highly insignificant and the null hypothesis that beer tax does not affect the alcohol involved vehicle fatality rate cannot be rejected. Thus, the instrumental variables would not be the right course of action to overcome the endogeneity in the model.

3. Variable selection for each model

In order to select the variables for each model, a specific analysis was performed as to whether the data variable in each case aligns with the economic theory. The main variable in this case, is the beertax. The other 2 important variables would be the jaild and the comserd variable. But, for those 2 variables, a case of insignificance is expected as they are slow changing variables over time.And if the variables are significant then we can expect a interaction effect to see the synergy effect of both the laws. The other variables taken into account are the unemployment rate, the per capita alcohol consumption, the minimum legal drinking age, the percentage of young drivers and the measure of economic growth. All the variables have some kind of effect on the traffic fatality rates as the per alcohol consumption is expected to increase fataility rates. The minimum legal drinking age would discourage more number of youth from drinking as they would be the ones who possibly drive recklessly and cause more accidents. For this reason, the percentage of young drivers is also considered. The measure of economic growth is another factor to be considered as the higher the economic growth, better would be the infrastructure of the city, which could lead to lesser number of accidents. We have included dry variables to see how faltality rates changes in the regions where no alcohol is consumed. We thought about adding a quadratic term for spircons, but we can observe from the summary statistics that their range is not that high. Hence, it is not required.

4. Model selection

In [70]:
```
model1 <- plm(mraidall~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(ja
ild*comserd)+pop1820+gspch,model="pooling",data=pcar)
coeftest(model1,method=vcovHC)
```

```
t test of coefficients:

                      Estimate  Std. Error t value  Pr(>|t|)    
(Intercept)         2.2261e-05  3.4938e-05  0.6372 0.5244778    
beertax             1.1856e-05  2.7147e-06  4.3674 1.696e-05 ***
spircons            2.6018e-06  1.9221e-06  1.3536 0.1768131    
unrate              1.0495e-06  6.3598e-07  1.6501 0.0998852 .  
mlda               -1.3980e-06  1.4331e-06 -0.9755 0.3300484    
dry                 6.6522e-07  1.4195e-07  4.6863 4.107e-06 ***
yngdrv              1.3132e-04  5.6913e-05  2.3074 0.0216641 *  
vmiles              3.4743e-09  8.8212e-10  3.9385 0.0001005 ***
jaild               1.7094e-05  3.9176e-06  4.3633 1.726e-05 ***
comserd             4.8502e-06  5.9524e-06  0.8148 0.4157709    
I(jaild * comserd) -1.1758e-05  7.4743e-06 -1.5732 0.1166583    
pop1820            -7.8439e-12  5.3006e-12 -1.4798 0.1399021    
gspch              -1.1093e-04  3.3344e-05 -3.3269 0.0009794 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Initially, a pooled OLS model with white correction is run, where, the alcohol involved vehicle fatality rate variable is regressed on a set of explanatory variables consistent from an economic standpoint. Most of the variables, in this case, are highly insignificant. But for the beertax variable, although the coefficient on beertax is positive- which is inconsistent with the economic theory- it is highly significant. As discussed earlier, there is a case of endogeneity, possibly caused by the omitted variable bias or the simultaneous causality bias. Thus, rather than considering a pooled OLS regression model, the fixed effects model would make better sense as it would ignore the case of unobserved heterogeneity and endogeneity.

In [117]:

```
model2 <- plm(mraidall~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(ja
ild*comserd)+dry+pop1820,model="within",data=pcar)
coeftest(model2,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value  Pr(>|t|)    
beertax  -2.3641e-05  1.3353e-05 -1.7706  0.077731 .  
spircons  3.1159e-05  7.5356e-06  4.1349 4.707e-05 ***
unrate   -1.8283e-06  6.2205e-07 -2.9392  0.003567 ** 
mlda      2.1761e-08  1.4115e-06  0.0154  0.987711    
dry       4.3771e-07  1.0507e-06  0.4166  0.677308    
yngdrv    1.0282e-04  6.1504e-05  1.6717  0.095709 .  
vmiles   -4.1396e-10  7.0838e-10 -0.5844  0.559443    
jaild     2.0564e-05  9.7230e-06  2.1150  0.035323 *  
comserd  -2.0132e-05  1.1183e-05 -1.8003  0.072903 .  
pop1820   2.5956e-11  7.0211e-11  0.3697  0.711900    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

After running the Fixed effects model, a lot of variables turn insignificant and those are removed from the model. Suprsisingly, minimum legal drinking age is insignificant in the mode, which is a very important variable according to economic theory. The beertax variable now has a negative effect, which is consistent with the economic theory. As expected comserd is reducing the fatality rate and it is significant at 10%. Suprisingly, jaild is having a positive coefficient which is not expected, and the variable is significant at 5%. The next step would be to run the model with only the variables that are significant in the above fixed effects model.

In [118]:

```
model3 <- plm(mraidall~beertax+spircons+unrate+yngdrv+jaild+comserd,model="within",data=
pcar)
coeftest(model3,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value  Pr(>|t|)    
beertax  -2.2430e-05  1.3170e-05 -1.7031  0.089646 .  
spircons  3.1762e-05  7.1113e-06  4.4664 1.152e-05 ***
unrate   -1.6769e-06  5.4951e-07 -3.0516  0.002493 ** 
yngdrv    1.1062e-04  5.9371e-05  1.8632  0.063470 .  
jaild     2.0794e-05  9.6204e-06  2.1614  0.031506 *  
comserd  -2.0204e-05  1.1104e-05 -1.8194  0.069907 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

A model with the time and entity fixed effects is then run in order to check the effects across time and entity. An F-test is conducted to check whether it is better to have the time and entity fixed effects.

In [119]:
```
model4 <- plm(mraidall~beertax+spircons+unrate+yngdrv+jaild+comserd+as.factor(year),mode
l="within",data=pcar)
coeftest(model4,method=vcovHC)
```

```
t test of coefficients:

                      Estimate  Std. Error t value  Pr(>|t|)    
beertax            -2.1947e-05  1.3019e-05 -1.6858 0.0929653 .  
spircons            2.6229e-05  9.3963e-06  2.7914 0.0056143 ** 
unrate             -3.0981e-06  7.5267e-07 -4.1161 5.094e-05 ***
yngdrv              4.1789e-05  6.7415e-05  0.6199 0.5358537    
jaild               2.3843e-05  9.4491e-06  2.5234 0.0121856 *  
comserd            -1.9843e-05  1.0848e-05 -1.8293 0.0684364 .  
as.factor(year)1983 -6.4371e-06  2.5933e-06 -2.4821 0.0136546 *  
as.factor(year)1984 -1.1875e-05  3.3660e-06 -3.5279 0.0004905 ***
as.factor(year)1985 -1.5098e-05  3.8506e-06 -3.9209 0.0001113 ***
as.factor(year)1986 -9.3923e-06  4.8126e-06 -1.9516 0.0519963 .  
as.factor(year)1987 -1.3722e-05  5.5475e-06 -2.4736 0.0139786 *  
as.factor(year)1988 -1.4986e-05  6.5288e-06 -2.2953 0.0224651 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In [59]:
```
waldtest(model3,model4)
```

Res.Df	Df	Chisq	Pr(>Chisq)
282	NA	NA	NA
276	6	23.21943	0.0007262204

The F-test shows a very low p-value, which is highly significant in this case and thus, it can be concluded that the time and entity effects is a better fit.

5. Analysis of vehicle fatality rate as a whole

In [120]:
```
model5 <- plm(mrall~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(jaild
*comserd)+pop1820+gspch,model="pooling",data=pcar)
coeftest(model5,method=vcovHC)
```

```
t test of coefficients:

                     Estimate  Std. Error t value  Pr(>|t|)    
(Intercept)       -1.5807e-06  7.3866e-05 -0.0214 0.9829399    
beertax            2.4622e-05  5.7395e-06  4.2900 2.364e-05 ***
spircons           7.0809e-06  4.0637e-06  1.7425 0.0823782 .  
unrate             3.3105e-06  1.3446e-06  2.4621 0.0143353 *  
mlda              -2.5544e-07  3.0299e-06 -0.0843 0.9328670    
dry                1.0189e-06  3.0011e-07  3.3952 0.0007714 ***
yngdrv             1.9424e-04  1.2033e-04  1.6143 0.1074409    
vmiles             1.5149e-08  1.8650e-09  8.1228 9.748e-15 ***
jaild              3.1212e-05  8.2827e-06  3.7683 0.0001953 ***
comserd            2.7544e-05  1.2585e-05  2.1887 0.0293328 *  
I(jaild * comserd) -2.4849e-05  1.5802e-05 -1.5725 0.1168173    
pop1820           -1.9993e-11  1.1207e-11 -1.7840 0.0753550 .  
gspch             -1.5051e-04  7.0496e-05 -2.1350 0.0335153 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Analysing the different effects of each of the explanaory variables consistent with an economic theory, again it is observed that most of the variables are insignificant and for this purpose, a fixed effects model is preferred.

In [121]:
```
model6 <- plm(mrall~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(jaild
*comserd)+pop1820+gspch,model="within",data=pcar)
coeftest(model6,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value  Pr(>|t|)    
beertax  -5.1319e-05  1.7153e-05 -2.9918  0.003023 ** 
spircons  6.7132e-05  9.6688e-06  6.9432 2.725e-11 ***
unrate   -6.4221e-06  8.7508e-07 -7.3388 2.398e-12 ***
mlda      3.3170e-06  1.8092e-06  1.8334  0.067822 .  
dry       2.7731e-06  1.3453e-06  2.0614  0.040201 *  
yngdrv    6.9788e-05  7.8799e-05  0.8856  0.376580    
vmiles    1.3518e-09  9.1313e-10  1.4804  0.139890    
jaild     2.4427e-06  1.2506e-05  0.1953  0.845285    
comserd   1.4497e-06  1.4326e-05  0.1012  0.919468    
pop1820  -2.4224e-11  9.0100e-11 -0.2689  0.788241    
gspch    -6.4062e-05  3.3384e-05 -1.9189  0.056021 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The beertax in this case is significant at a 5% level. The effect is again observed in a fixed effects model for only the variables that are significant in this case.

In [122]:
```
model7 <- plm(mrall~beertax+spircons+unrate+dry+mlda+gspch,model="within",data=pcar)
coeftest(model7,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value  Pr(>|t|)    
beertax  -5.2092e-05  1.6750e-05 -3.1099  0.002063 ** 
spircons  6.9184e-05  7.8324e-06  8.8331 < 2.2e-16 ***
unrate   -6.5217e-06  7.7299e-07 -8.4369 1.721e-15 ***
dry       2.7418e-06  1.3324e-06  2.0578  0.040530 *  
mlda      3.2415e-06  1.8007e-06  1.8002  0.072906 .  
gspch    -5.7418e-05  3.2565e-05 -1.7632  0.078949 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The time and entity effects is then run and then compared using an f-test with the above model.

In [123]:
```
model8 <- plm(mrall~beertax+spircons+unrate+dry+mlda+gspch+as.factor(year),model="withi
n",data=pcar)
coeftest(model8,method=vcovHC)
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)    
beertax             -4.6023e-05  1.5845e-05 -2.9045  0.003976 ** 
spircons             7.7477e-05  1.1649e-05  6.6510 1.557e-10 ***
unrate              -7.2430e-06  1.0304e-06 -7.0296 1.626e-11 ***
dry                  2.0363e-06  1.2424e-06  1.6390  0.102344    
mlda                 1.8644e-06  1.7191e-06  1.0845  0.279077    
gspch                4.7512e-05  4.3780e-05  1.0853  0.278758    
as.factor(year)1983 -6.9209e-06  3.7940e-06 -1.8241  0.069213 .  
as.factor(year)1984 -1.9144e-05  4.6692e-06 -4.1000 5.438e-05 ***
as.factor(year)1985 -1.9271e-05  4.3555e-06 -4.4244 1.394e-05 ***
as.factor(year)1986 -2.8347e-06  5.3160e-06 -0.5332  0.594290    
as.factor(year)1987 -6.3900e-06  6.0582e-06 -1.0548  0.292450    
as.factor(year)1988 -8.8373e-06  6.9139e-06 -1.2782  0.202258    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In [97]:
```
waldtest(model7, model8)
```

Res.Df	Df	Chisq	Pr(>Chisq)
282	NA	NA	NA
276	6	50.62849	3.516748e-09

There is certain level of significant evidence to show that the time and entity effects is a better model.

We are running the time and entity effects with only significant variables

In [124]:
```
model9 <- plm(mrall~beertax+spircons+unrate+as.factor(year),model="within",data=pcar)
coeftest(model9,method=vcovHC)
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)    
beertax             -4.4980e-05  1.5894e-05 -2.8301  0.004992 ** 
spircons             7.4881e-05  1.1555e-05  6.4807 4.123e-10 ***
unrate              -7.8409e-06  9.1546e-07 -8.5650 7.432e-16 ***
as.factor(year)1983 -4.6127e-06  3.1148e-06 -1.4809  0.139759    
as.factor(year)1984 -1.6621e-05  3.8703e-06 -4.2944 2.420e-05 ***
as.factor(year)1985 -1.8554e-05  4.2807e-06 -4.3343 2.044e-05 ***
as.factor(year)1986 -1.6766e-06  5.2622e-06 -0.3186  0.750265    
as.factor(year)1987 -5.2780e-06  5.9973e-06 -0.8801  0.379585    
as.factor(year)1988 -7.7179e-06  6.8702e-06 -1.1234  0.262236    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Vehicle fatality rate across age groups

In [100]:
```
model_10 <- plm(mra2124~beertax+spircons+perinc+vmiles+gspch+jaild+comserd+unrate+as.fac
tor(year), model="within", data = pcar)
coeftest(model_2124_g,method=vcovHC)
```

```
t test of coefficients:

                       Estimate  Std. Error t value Pr(>|t|)
beertax             -7.4588e-05  6.9397e-05 -1.0748 0.283411
spircons             1.5632e-04  5.0440e-05  3.0991 0.002143 **
perinc               1.3820e-08  8.9365e-09  1.5464 0.123154
vmiles               1.2273e-10  3.6697e-09  0.0334 0.973345
gspch                3.5186e-04  1.9051e-04  1.8470 0.065832 .
jaild                3.8218e-05  5.0104e-05  0.7628 0.446258
comserd             -8.4668e-05  5.7412e-05 -1.4747 0.141433
unrate              -1.2342e-05  4.9862e-06 -2.4751 0.013924 *
as.factor(year)1983 -3.5801e-05  1.6431e-05 -2.1789 0.030193 *
as.factor(year)1984 -6.4455e-05  2.0193e-05 -3.1920 0.001578 **
as.factor(year)1985 -4.3016e-05  1.8950e-05 -2.2700 0.023983 *
as.factor(year)1986 -1.9654e-06  2.3206e-05 -0.0847 0.932565
as.factor(year)1987 -1.8264e-05  2.6409e-05 -0.6916 0.489780
as.factor(year)1988 -4.9676e-05  3.0298e-05 -1.6396 0.102243
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- **Here in the above model we can see that jail and community services don't fear the society, which is a surprising observation as the variables are highly insignificant.**
- **gscph is highly significant, with which we can say the society is misusing the infrastructure advancements(This age group has high significance when compared to other age groups, so we need to have strict policies with this age group when the economy growth is high)**
- **if the unemployment rate is more, we observe a decrease in VFR, assuming there is less usage of personal commute. There is a possibility that this age group is using more public transport which is considered safe.**

In [101]:
```
model_11 <- plm(mra1820~beertax+spircons+perinc+as.factor(year), model="within", data =
 pcar)
coeftest(model_1820_g1,method=vcovHC)
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)
beertax             -1.4109e-04  8.0907e-05 -1.7439  0.082277 .
spircons             3.0356e-04  5.6500e-05  5.3727 1.638e-07 ***
perinc               2.1690e-08  9.0107e-09  2.4071  0.016729 *
as.factor(year)1983 -4.2361e-05  1.5876e-05 -2.6682  0.008072 **
as.factor(year)1984 -3.0460e-05  1.7293e-05 -1.7614  0.079258 .
as.factor(year)1985 -6.0348e-05  1.9444e-05 -3.1037  0.002108 **
as.factor(year)1986  1.6398e-05  2.4731e-05  0.6630  0.507854
as.factor(year)1987 -3.5783e-06  2.7926e-05 -0.1281  0.898134
as.factor(year)1988  5.2799e-05  3.1502e-05  1.6760  0.094849 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- **Beer tax imposition has a very good effect on decreasing the vehicle fatality rate and the variable is significant at 10%.**
- **Here Spircons is highly significant as expected and this age group is more prone to accidents.**
- **This age group is misusing the advancement in the society which is observed in perinc.**
- **To control VFR in this age group, we have to build a policy around beertax for which they will respond more.**

In [102]:
```
model_12 <- plm(mra1517~beertax+spircons+perinc+vmiles+gspch+jaild+comserd+as.factor(yea
r), model="within", data = pcar)
coeftest(model_1517_g,method=vcovHC)
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)
beertax             -4.3497e-05  6.1480e-05 -0.7075 0.4798622
spircons             1.4860e-04  4.2393e-05  3.5052 0.0005326 ***
perinc               6.1589e-09  7.0579e-09  0.8726 0.3836279
vmiles               7.6278e-09  3.2514e-09  2.3460 0.0196868 *
gspch                1.4884e-04  1.5582e-04  0.9552 0.3403070
jaild                1.1314e-06  4.4330e-05  0.0255 0.9796574
comserd              3.3452e-05  5.0834e-05  0.6581 0.5110436
as.factor(year)1983 -1.0251e-05  1.4241e-05 -0.7199 0.4722254
as.factor(year)1984 -1.0867e-05  1.7689e-05 -0.6143 0.5395043
as.factor(year)1985  2.0446e-06  1.5770e-05  0.1296 0.8969398
as.factor(year)1986  5.7134e-05  1.9446e-05  2.9381 0.0035822 **
as.factor(year)1987  6.1221e-05  2.1636e-05  2.8296 0.0050035 **
as.factor(year)1988  5.5997e-05  2.4484e-05  2.2871 0.0229491 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- **It is surprising that spircons is highly significant because this age group is not eligible to drink ,so we need to have strict guidelines around alcohol consumption in this particular age group.**

Analysis of Night vehicle fatality rate as a whole

In [106]:
```
model13 <- plm(mralln~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(jai
ld*comserd)+pop1820+gspch,model="pooling",data=pcar)
coeftest(model13,method=vcovHC)
```

```
t test of coefficients:

                      Estimate  Std. Error t value  Pr(>|t|)
(Intercept)         1.7826e-05  1.5787e-05  1.1292  0.259667
beertax             1.6192e-06  1.2267e-06  1.3200  0.187766
spircons            1.1372e-06  8.6853e-07  1.3093  0.191365
unrate              6.9884e-07  2.8738e-07  2.4318  0.015566 *
mlda               -7.5285e-07  6.4758e-07 -1.1626  0.245862
dry                -1.9082e-09  6.4142e-08 -0.0297  0.976285
yngdrv              7.1317e-05  2.5717e-05  2.7731  0.005874 **
vmiles              1.9131e-09  3.9860e-10  4.7997 2.433e-06 ***
jaild               5.1750e-06  1.7702e-06  2.9234  0.003707 **
comserd             4.3825e-06  2.6897e-06  1.6294  0.104204
I(jaild * comserd) -7.6574e-06  3.3773e-06 -2.2673  0.024034 *
pop1820            -1.9661e-13  2.3951e-12 -0.0821  0.934628
gspch              -4.3343e-05  1.5067e-05 -2.8767  0.004285 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In [107]:
```
model14 <- plm(mralln~beertax+spircons+unrate+mlda+dry+yngdrv+vmiles+jaild+comserd+I(jai
ld*comserd)+pop1820+gspch,model="within",data=pcar)
coeftest(model14,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value Pr(>|t|)
beertax  -1.0260e-05  6.5657e-06 -1.5626 0.119279
spircons  1.3907e-05  3.7009e-06  3.7578 0.000209 ***
unrate   -4.2106e-07  3.3495e-07 -1.2571 0.209789
mlda      3.5223e-07  6.9251e-07  0.5086 0.611419
dry       5.6524e-07  5.1492e-07  1.0977 0.273275
yngdrv    1.7886e-05  3.0162e-05  0.5930 0.553656
vmiles   -1.9648e-10  3.4952e-10 -0.5622 0.574464
jaild     5.1897e-07  4.7868e-06  0.1084 0.913744
comserd  -3.8399e-06  5.4833e-06 -0.7003 0.484339
pop1820   4.6879e-13  3.4487e-11  0.0136 0.989164
gspch    -2.2119e-05  1.2778e-05 -1.7310 0.084574 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In [108]:
```
model15 <- plm(mralln~spircons+gspch,model="within",data=pcar)
coeftest(model15,method=vcovHC)
```

```
t test of coefficients:

            Estimate  Std. Error t value  Pr(>|t|)
spircons  1.2650e-05  2.3527e-06  5.3768 1.577e-07 ***
gspch    -1.8319e-05  1.1013e-05 -1.6634   0.09732 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In [109]:
```
model16 <- plm(mralln~spircons+gspch+as.factor(year),model="within",data=pcar)
coeftest(model16,method=vcovHC)
```

```
t test of coefficients:

                        Estimate  Std. Error t value  Pr(>|t|)
spircons              1.3176e-05  4.3268e-06  3.0453 0.0025458 **
gspch                 3.4979e-05  1.5261e-05  2.2921 0.0226391 *
as.factor(year)1983 -5.1748e-06  1.4516e-06 -3.5649 0.0004279 ***
as.factor(year)1984 -8.5724e-06  1.8138e-06 -4.7263 3.626e-06 ***
as.factor(year)1985 -7.4437e-06  1.5459e-06 -4.8152 2.410e-06 ***
as.factor(year)1986 -4.0239e-06  1.8530e-06 -2.1716 0.0307247 *
as.factor(year)1987 -4.6468e-06  1.9668e-06 -2.3626 0.0188322 *
as.factor(year)1988 -4.9499e-06  2.1683e-06 -2.2828 0.0231914 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

if the alcohol consumption is more, then vehicle fatality rate increases at night, which is as expected.

In [110]:
```
waldtest(model15,model16)
```

Res.Df	Df	Chisq	Pr(>Chisq)
286	NA	NA	NA
280	6	35.10342	4.115173e-06

There is certain level of significant evidence to show that the time and entity effects is a better model.

7. Conclusion

As a whole, we can conclude that beertax does have a negative effect on traffic fatality rates, but, taking the case of the jaild(Mandatory jail sentence) and the comserd(Mandatory community serivce) variables, which are pretty important in answering our research hypothesis as to how drunk driving laws affect traffic deaths, we find that both variables are not at all significant in any case because as discussed earlier, they are slow changing variables over time. We were not able to get a proper effect in the case of the pooled OLS model and even in the case after running a fixed effects model, we do not get a significant result. This is partly attributed to the disadvantage held with the fixed effects model, where, slow changing variables over time does not give a significant result. We have run models initially for the case of alcohol involved vehicle fatality rates, which shows us that beertax, unemployment rate, mandatory jail time and community service does have a negtaive effect on fatailty rate, but the model is more significant when time and entity effects are taken into consideration. This is in no way a perfect estimate and there never would be one. But, we can conclude that drunk driving laws do seem to have some kind of a negative effect on the fatality rates. However, the clear cut effect by beertax is the only one that gave us a significant value, while other variables such as the minimum legal drinking age did not give us a very significant result. This shows that other factors and variables would be needed to estimate a proper effect. This is what the data tells us.