Effect of drunk driving laws on traffic deaths

BUAN 6312.003

Applied Econometrics and Time Series Analysis

Team Members:

Sumasree Kodati (sxk171031)

Aarohi Tripathi (axt174830)

Sai Krishna Gollapally (sxg171731)

Bhowmik Dani (byd170000)

Abstract:

To understand the effect of drinking and driving laws on the number of car fatalities, this study explored the different effects these laws have in 48 states of the united states over a period of 7 years from 1982-1986. The data includes socio-economic factors like average income, unemployment, population, beer tax, legal drinking age etc. We used the panel data regression using fixed and random effects to see the true effect of these laws on the number of deaths occurred in accidents. Empirical results showed that policies based on local conditions must be used to effectively reduce drinking and driving fatality rates; that is, different measures should be adopted to target the specific conditions in various regions. The results of this study show the importance of demographic and economic characteristics as the major determinants of traffic fatalities. The results conclude that population characteristics play a major role in determining the traffic fatality rate. The regression results presented in this paper help explain the large variation in traffic fatalities across states and time.

In the following study, we direct our attention towards the effects of the drunk-drive laws in the 48 states in the Unites States of America over 7 years. Various statistical models are run to study the trends in fatality rates before and after the shall issue laws have been introduced. We also study the effects of factors like average income, unemployment rate, population etc.

Data Description:

This is an unbalanced panel data on 48 US states by year for 1982 – 1986. Each observation is a given state in a given year. The attributes in the data are explained below.

Variable	Descriptions
state	State ID (FIPS) Code
year	Year
spircons	Per Capita Pure Alcohol Consumption (Annual, Gallons)
unrate	State Unemployment Rate (%)
perinc	Per Capita Personal Income ($)
beertax	Tax on Case of Beer ($)
sobapt	% Southern Baptist
mormon	% Mormon
mlda	Minimum Legal Drinking Age (years)
dry	% Residing in Dry Counties A dry county is a county whose government forbids the sale of any kind of alcoholic beverages. Some prohibit off-premises sale, some prohibit on-premises sale, and some prohibit both.
yngdrv	% of Drivers Aged 15-24
vmiles	Ave. Mile per Driver
jaild	Mandatory Jail Sentence

comserd	Mandatory Community Service
allmort	# of Vehicle Fatalities (#VF)
mrall	Vehicle Fatality Rate (VFR)
allnite	# of Night-time VF (#NVF)
mralln	Night-time VFR (NFVR)
allsvn	# of Single VF (#SVN)
a1517	#VF, 15-17 year olds
mra1517	VFR, 15-17 year olds
a1517n	#NVF, 15-17 year olds
mra1517n	NVFR, 15-17 year olds
a1820	#VF, 18-20 year olds
a1820n	#NVF, 18-20 year olds
mra1820	VFR, 18-20 year olds
mra1820n	NVFR, 18-20 year olds
a2124	#VF, 21-24 year olds
mra2124	VFR, 21-24 year olds
a2124n	#NVF, 21-24 year olds
mra2124n	NVFR, 21-24 year olds
aidall	# of alcohol-involved VF
mraidall	Alcohol-Involved VFR
pop	Population
pop1517	Population, 15-17 year olds
pop1820	Population, 18-20 year olds
pop2124	Population, 21-24 year olds
miles	total vehicle miles (millions)
gspch	GSP Rate of Change This is a measure of economic growth

The traffic fatality rate is the number of traffic deaths in a given state in a given year, per 10,000 people living in that state in that year. Traffic fatality data were obtained from the U.S. Department of Transportation Fatal Accident Reporting System. The beer tax is the tax on a case of beer, which is an available measure of state alcohol taxes more generally. The drinking age variables are binary variables indicating whether the legal drinking age is 18, 19, or 20. The two binary punishment variables describe the state's minimum sentencing requirements for an initial drunk driving conviction: "Mandatory jail?" equals one if the state requires jail time and equals zero otherwise, and "Mandatory community service?" equals one if the state requires community service and equals zero otherwise.

Instead of diving right away into the regression analysis and results. We look at the preliminary analysis of all the variables and different trends across time periods in different states.

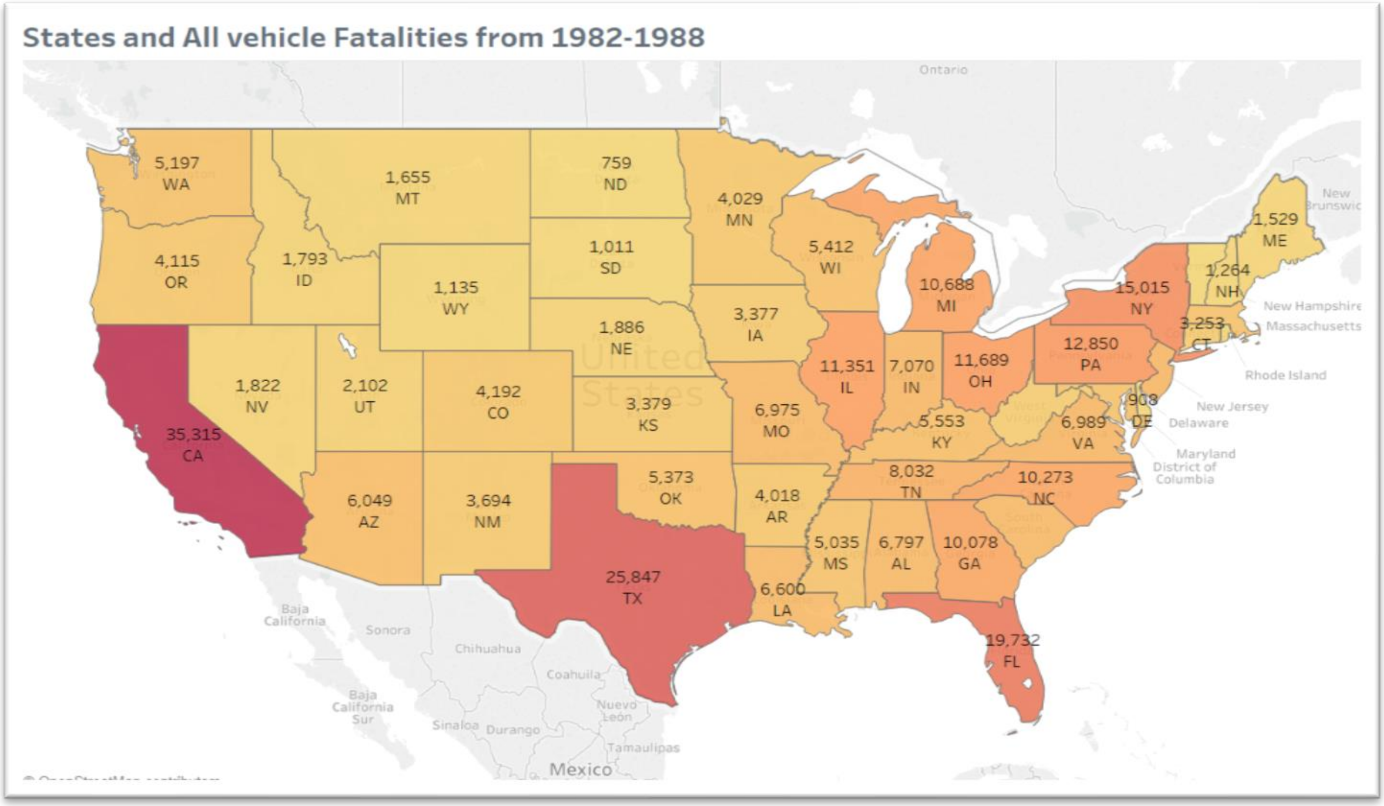
Exploratory Analysis:

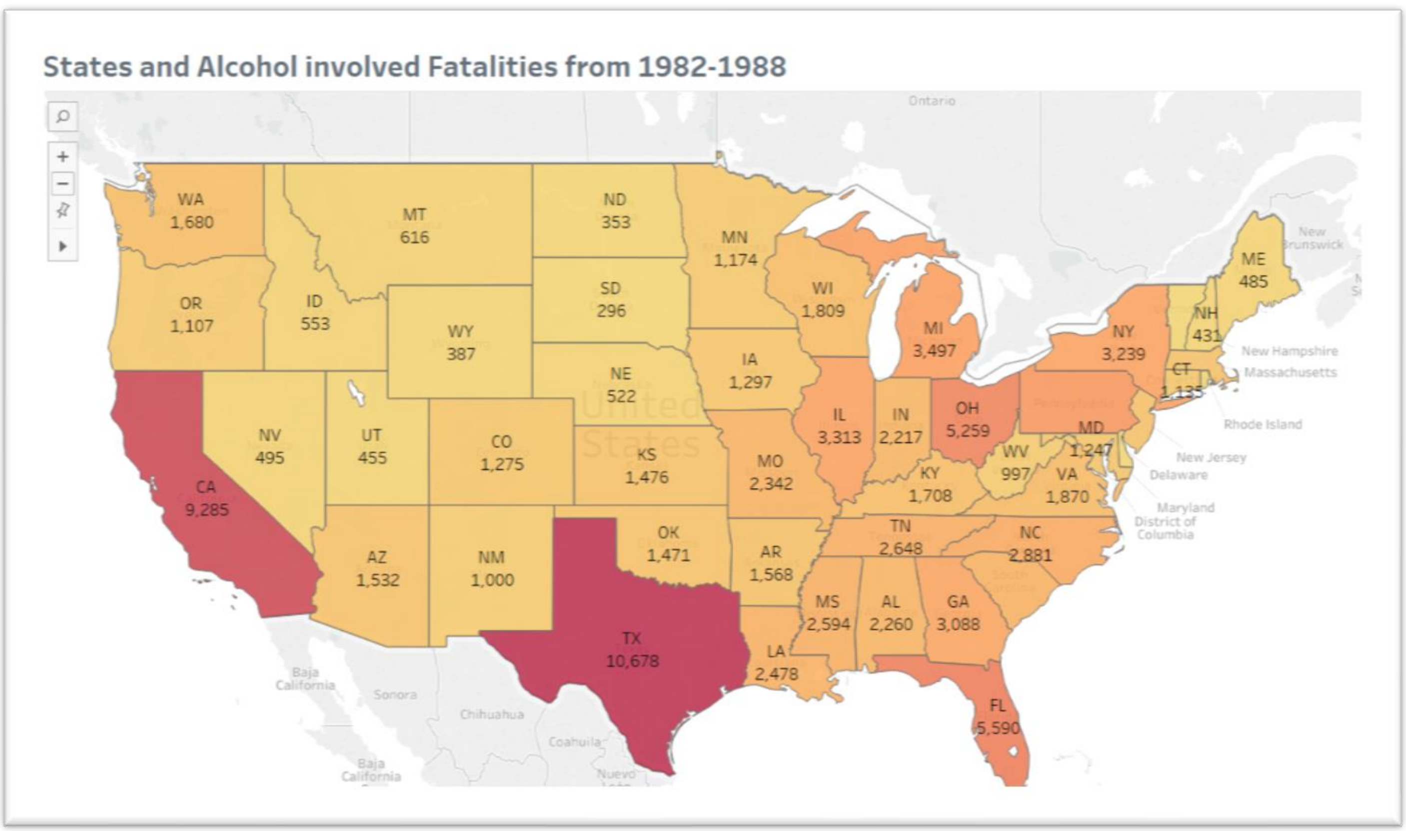
Initial summary of data:

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
state	336	30.1875	15.30985	1	56
year	336	1985	2.002983	1982	1988
spircons	336	1.75369	.6835745	.79	4.9
unrate	336	7.346726	2.533405	2.4	18
perinc	336	13880.18	2253.046	9513.762	22193.46
beertax	336	.513256	.4778442	.0433109	2.720764
sobapt	336	7.156925	9.762621	0	30.3557
mormon	336	2.801933	9.665279	.1	65.9165
mlda	336	20.45563	.8990255	18	21
dry	336	4.267074	9.500901	0	45.7921
yngdrv	336	.1859299	.0248736	.073137	.281625
vmiles	336	7890.754	1475.659	4576.346	26148.27
jaild	335	.280597	.449963	0	1
comserd	335	.1850746	.388939	0	1
allmort	336	928.6637	934.0515	79	5504
mrall	336	.000204	.000057	.0000821	.0004218
allnite	336	182.5833	188.4311	13	1049
mralln	336	.0000388	.000011	.0000172	.0000944
allsvn	336	109.9494	108.5397	8	603
a1517	336	62.61012	55.72909	3	318
mra1517	336	.0003034	.0000937	.0001163	.0006735
a1517n	336	12.2619	12.25341	0	76
mra1517n	336	.0000598	.000033	0	.0002571
a1820	336	106.6607	104.2236	7	601
a1820n	336	33.52679	33.23834	0	196
mra1820	336	.0004728	.0001522	.0001855	.0010952
mra1820n	336	.0001436	.0000613	0	.0005238
a2124	336	126.872	131.7886	12	770
mra2124	336	.0004091	.0001225	.0002	.0008922
a2124n	336	41.37798	42.93031	1	249
mra2124n	336	.0001284	.0000422	.0000222	.0003143
aidall	336	293.3332	303.5807	24.6	2094.9
mraidall	336	.0000659	.000026	.0000234	.0001772
pop	336	4930272	5073704	478999.7	2.83e+07
pop1517	336	230815.5	229896.3	21000.02	1172000

Map of united states with their Vehicle Fatalities and Alcohol involved Fatalities

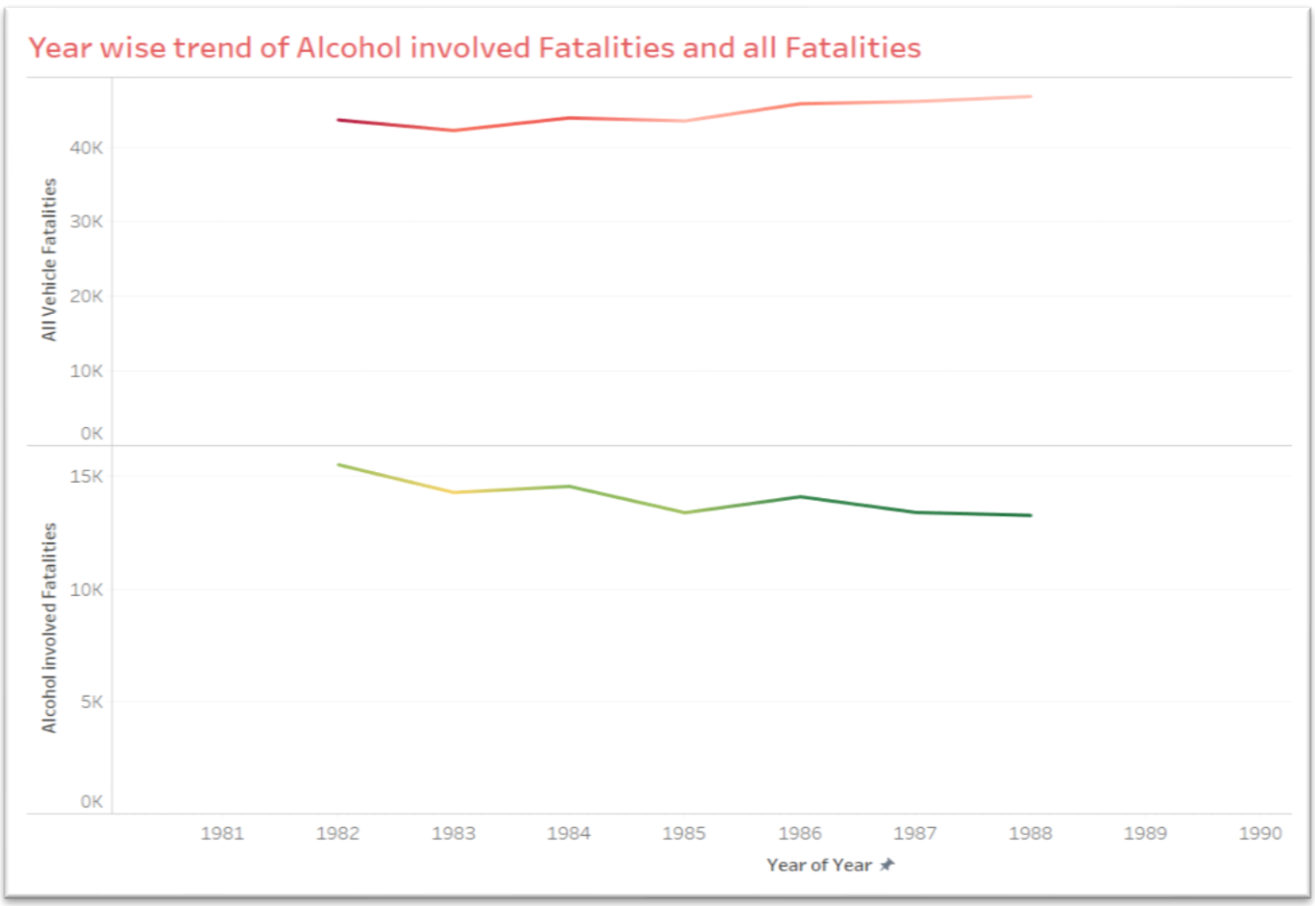




Above Two visualizations represent All vehicle fatalities and Alcohol involved fatalities according to geographic location of their occurrence from 1982-1988. Observations are as follows:

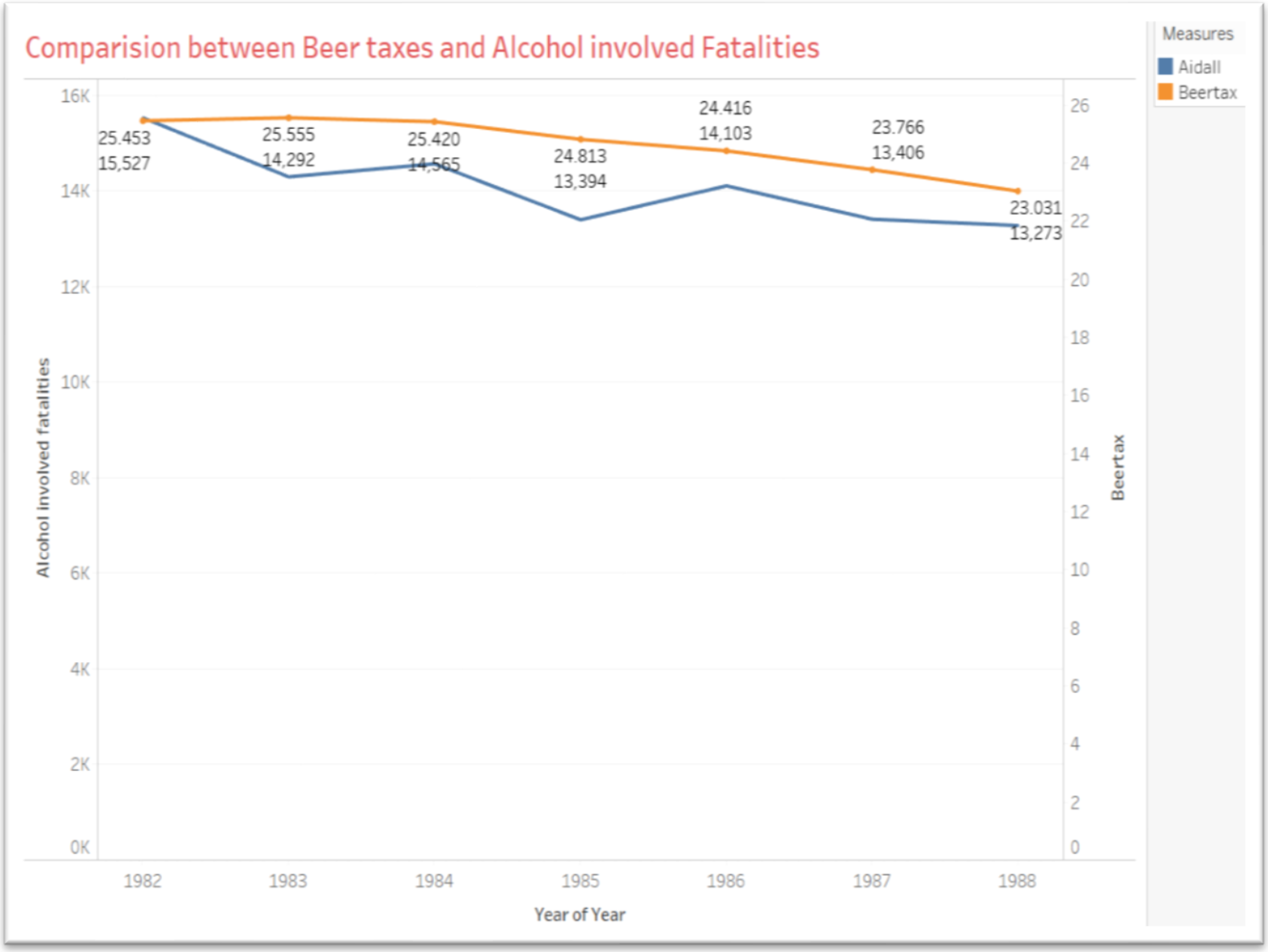
- Among all vehicle involved fatalities from 1982-1988. Top three states with highest fatalities are 1) California 2) Texas 3) Florida. We could assume California and Texas being two largest states in the united states the count of highest fatalities seems valid.
- Among alcohol involved fatalities. Texas is the state with highest number of alcohol related fatalities next being California.
- Amongst Total of 25847 fatalities in Texas approximately half are related alcohol involved fatalities.

Year-wise trend of alcohol involved fatalities and all fatalities:



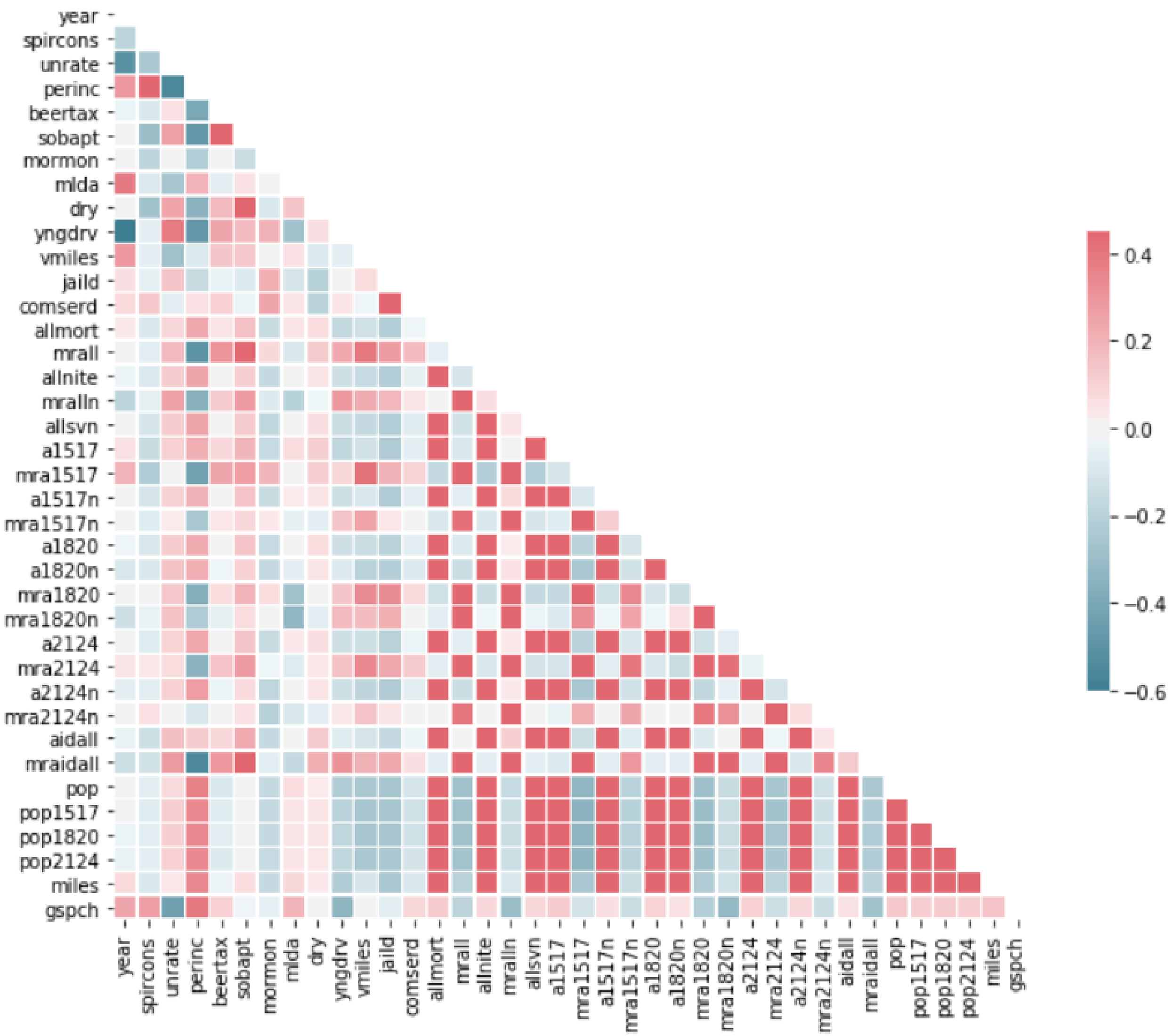
On Plotting year wise "All Vehicle Fatalities" and "Alcohol involved Fatalities" across all states from 1982-1984 gave a surprising revelation that "All Vehicles Fatalities" were increasing year by year whereas Alcohol involved Fatalities saw a downward trend. This could be attributed to the effort taken by states to control Driving under influence from 1982 – 1984 making strict laws. Proportion of Alcohol related fatalities decreased whereas other fatalities increase.

Comparison between beer-tax and alcohol involved fatalities:



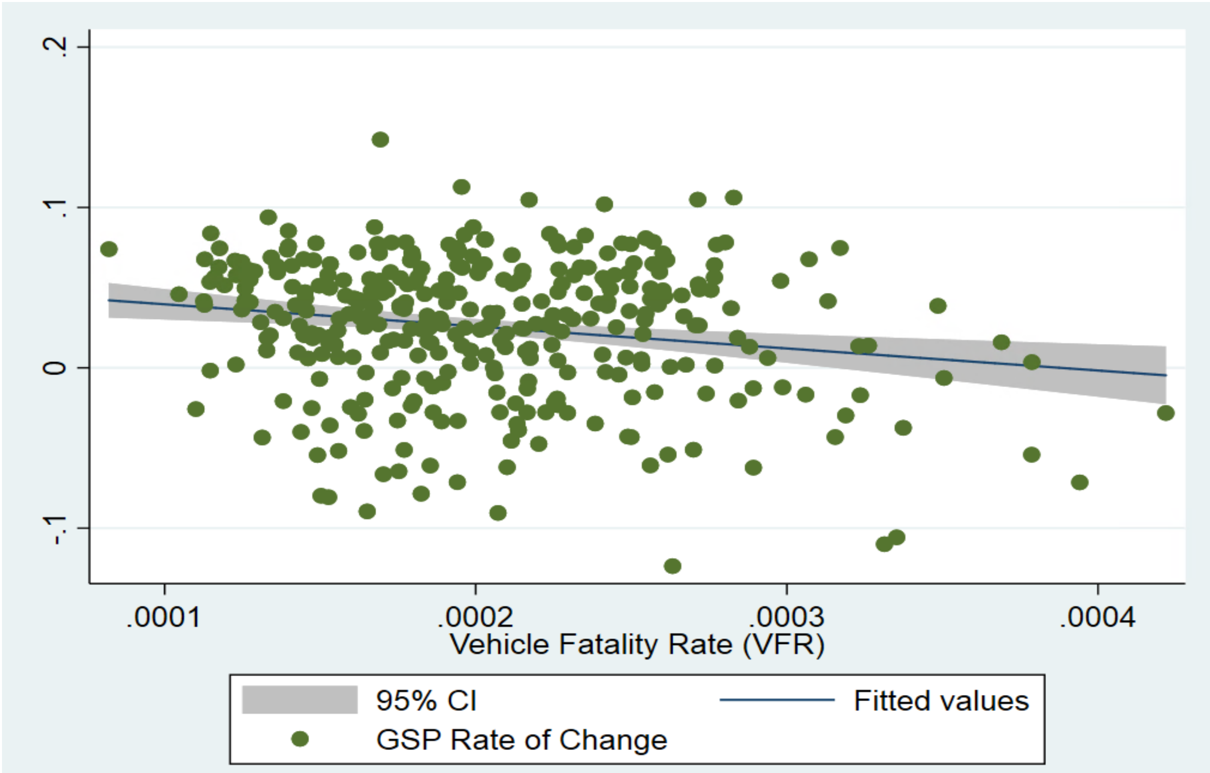
- In the figure above, beer-tax and alcohol involved fatalities across all states from the year 1982 to 1988 were plotted. One would presume that as the beer tax decreases, Alcohol involved fatalities should rise but data reveals opposite of the same.
- As we move from 1982 to 1988, the beer tax decreased and the alcohol involved fatality rates in the respective years followed a downward trend.

Initial Correlation matrix:

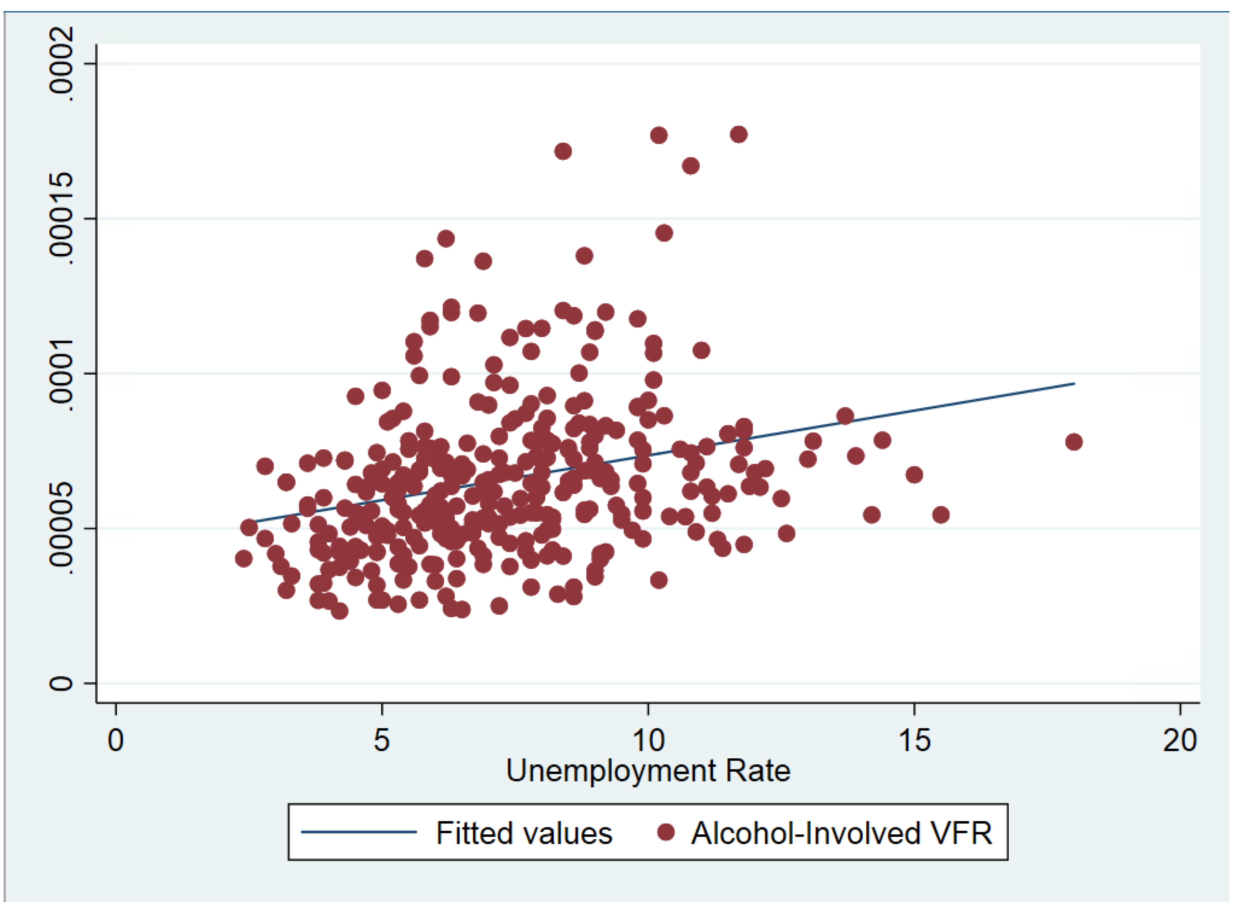


- The above graph depicts a correlation matrix which shows the correlation between all the variables in the data with every other variable.
- It is very clear from the matrix that there is high correlation between the variables like the total number of vehicle fatalities, total number of night time fatalities and total number of alcohol involved fatalities and also between their respective rates. Hence we do not include all these variables in the regression at once.
- Instead we analyze them separately and see the factors that are affecting them. We perform regression analysis for Total vehicle fatality rate(mrall) , alcohol involved fatality rate (mraidall) and night time fatality rate (mralln).
- There is also high population between the total population and the population of 15-17 year old, 18-20 year old and 21-24 year old.

Relation between growth rate and vehicle fatality rate



Relation between unemployment rate and alcohol involved fatality rate



Regression Analysis:

Now that we have looked at the descriptive statistics we need to need to validate whether the inferences we made from the plots are consistent with outputs of our regression models. For that, we ran several models to draw a conclusion of which model best describes our dataset.

We have plotted histograms of violent crime rate, robbery rate and murder rate to understand their distribution. We can see that all three rates are right skewed, which means their variance is not constant. So, to control for that, we have taken log on all three variables.

Vehicle Fatality Rate:

Pooled OLS:

Code:

```
d2 <- plm (mrall ~ perinc + beertax + spircons + unrate + mormon + yngdrv + dry + punish + drinkagec +
pop1517 + pop1820 + pop2124,index=c("state", "year"), model="pooling", data=table1)
```

Result:

```
Unbalanced Panel: n = 48, T = 6-7, N = 335

Residuals:
       Min.     1st Qu.      Median     3rd Qu.        Max.
-9.6499e-05 -2.8466e-05 -6.3389e-06  2.2997e-05  1.9325e-04

Coefficients:
                    Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)       4.0128e-04  4.0918e-05  9.8069 < 2.2e-16 ***
perinc           -1.5807e-08  2.0822e-09 -7.5912 3.500e-13 ***
beertax           1.2230e-05  5.9752e-06  2.0467  0.041501 *
spircons          1.9122e-05  4.5071e-06  4.2427 2.895e-05 ***
unrate           -3.7997e-06  1.4345e-06 -2.6489  0.008476 **
mormon           -5.9157e-07  2.7480e-07 -2.1528  0.032085 *
yngdrv            5.1039e-05  1.2173e-04  0.4193  0.675297
dry               2.8284e-07  2.9700e-07  0.9523  0.341651
punish1           2.9955e-05  5.7407e-06  5.2181 3.257e-07 ***
drinkagec[18,19) -2.9517e-05  1.2408e-05 -2.3788  0.017956 *
drinkagec[19,20)  6.5344e-06  6.7036e-06  0.9748  0.330412
drinkagec[20,21) -4.2647e-05  9.1689e-06 -4.6513 4.833e-06 ***
pop1517           2.4612e-10  1.7900e-10  1.3750  0.170104
pop1820          -8.3630e-10  3.3397e-10 -2.5041  0.012772 *
pop2124           4.3654e-10  1.6029e-10  2.7235  0.006814 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.089e-06
Residual Sum of Squares: 6.1106e-07
R-Squared:      0.43886
Adj. R-Squared: 0.41431
F-statistic: 17.8764 on 14 and 320 DF, p-value: < 2.22e-16
```

A Pooled OLS with Least Squared Standard Error estimates was run.

We choose to not include the variables like night time fatalities, alcohol involved fatalities and their rate as they are all highly correlated to the y variable here which is the total vehicle fatality rate and does not add any new significance to explaining the variation in vehicle fatality rate. Total number of vehicle fatalities include the number of night time fatalities and the alcohol involved fatalities.

Results:

- The regression results show that most of the variables are significant at 5% or less significant level. However, the coefficient of beertax has a positive sign saying that the total fatality rate increases with the increase in the beertax which is unexpected.
- The per capita income has a negative effect here i.e. as income increases, the number of alcohol involved fatality rate decreases. This seems to be plausible because high per capita income refers to high education which could make people aware of road safety rules and act responsibly on public roads.
- Punish1 is a variable which takes the value 1 if there is any one of the punishment like jail or community service. This has a positive effect on the fatality rate i.e. the fatality rate is high when there is a mandatory punishment rule than when there is no such rule.
- The coefficients of variables of population between age 18-20 and 21-24 are highly significant and pop18-20 has a negative effect whereas pop2124 has a positive effect on the fatality rate.

We also performed Breusch-Pagan Test in to check if there is any heteroskedasticity in the data.

Heteroskedasticity: The variance of an explanatory variable increases the variance of the error term increases.

Implications

1. Least square estimators are no longer best estimators although unbiased and consistent
2. The standard errors are incorrect hence incorrect confidence interval and hypothesis testing

The following results show the results of the Breusch-Pagan test:

```
> lmtest::bptest(d2)

	studentized Breusch-Pagan test

data:  d2
BP = 41.958, df = 12, p-value = 3.386e-05
```

The null hypothesis and alternative hypothesis in a Breusch-Pagan test are:

H0: There is no heteroskedasticity

H1: There is heteroskedasticity in the data

- Here, we can see that the p-value 3.386e-05 < 0.05. So, we can reject the null hypothesis at 5% significance level and conclude that there is heteroskedasticity in the data.
- In order the remove the heteroskedasticity in the pooled OLS, White's robust standard errors is calculated.

Pooled OLS with clustered robust standard errors:

```
> #White Robust Standard Errors
> coeftest(d2, vcov. = vcovHC(d2, type = 'HC1', cluster = 'group'))


t test of coefficients:

               Estimate  Std. Error t value  Pr(>|t|)    
(Intercept)  3.4406e-04  1.5042e-04  2.2873  0.022824 *  
perinc      -1.5990e-08  3.5866e-09 -4.4583 1.142e-05 ***
beertax      1.2374e-05  1.0035e-05  1.2330  0.218480    
spircons     1.4407e-05  8.8777e-06  1.6229  0.105596    
unrate      -3.9821e-06  2.2498e-06 -1.7700  0.077671 .  
mormon      -5.3239e-07  7.8076e-07 -0.6819  0.495800    
yngdrv       4.9807e-05  1.8666e-04  0.2668  0.789769    
dry          3.2274e-07  5.1758e-07  0.6236  0.533357    
punish1      3.1660e-05  1.4126e-05  2.2412  0.025692 *  
mlda         3.0640e-06  7.3687e-06  0.4158  0.677828    
pop1517      3.4002e-10  2.2605e-10  1.5042  0.133520    
pop1820     -9.9152e-10  3.8317e-10 -2.5877  0.010100 *  
pop2124      4.9349e-10  1.5888e-10  3.1061  0.002064 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Results:

- The above results show that the variables per capita income, punishment and the population of age 18-24 and 21-24 are significant at 5% significance level.
- However, this says that the effect of these variables is very minute. It says that for every 1% increase in the average per capita income the vehicle fatality rate decreases by 0.000016. This is very minute but the effect is significant.
- The variable ‘punish’ tell whether or not a person who committed drunk and drive offence is punished or not (either ‘jaild’ =1 or ‘comserv’ =1. This effect is interesting because it has a positive coefficient estimate. One would usually assume that the fatality rate would decrease if there the police take strict action against the people who committed it. But this explains that the vehicle fatality rate is slightly higher in case of punishment compared to the case where is no punishment.
- But Pooled OLS model will not avoid the problem of observed and unobserved heterogeneity. Which means this model will not take state specific variables (Observed and unobserved) into the consideration. Like, cultural attitude of people towards robbery, effectiveness of crime prevention departments etc.
- We can avoid heterogeneity by using “Fixed Effects” model.

Panel Regression Using Entity Fixed Effects Model:

Code:

```
>FixedEffectModel <- plm(mrall ~ unrate + spircons + perinc + beertax + mormon +
                    yngdrv + dry + vmiles + punish +drinkagec + gspch + pop1517 +
                    pop1820 + pop2124,index=c("state","year"), model="within",
                    data=context1)

> #White's clustered robust errors
> coeftest(FixedEffectModel, vcov. = vcovHC(FixedEffectModel, type = 'HC1', cluster
= "group"))
```

```
t test of coefficients:

                    Estimate  Std. Error t value  Pr(>|t|)    
unrate            -3.4746e-06  1.1101e-06 -3.1300  0.001939 ** 
spircons           8.5328e-05  1.4669e-05  5.8167 1.687e-08 ***
perinc             1.0855e-08  3.8122e-09  2.8473  0.004746 ** 
beertax           -4.0697e-05  2.7626e-05 -1.4731  0.141878    
mormon             4.4149e-07  4.8770e-06  0.0905  0.927938    
yngdrv             5.0292e-05  7.3376e-05  0.6854  0.493673    
dry                2.9526e-06  1.2053e-06  2.4498  0.014927 *  
vmiles             1.0525e-09  6.8883e-10  1.5279  0.127699    
punish1            6.0380e-07  1.0371e-05  0.0582  0.953617    
drinkagec[18,19) -4.6643e-06  8.2882e-06 -0.5628  0.574064    
drinkagec[19,20) -4.0621e-06  4.6751e-06 -0.8689  0.385687    
drinkagec[20,21) -9.3040e-08  3.6190e-06 -0.0257  0.979509    
gspch             -3.7987e-05  2.5524e-05 -1.4883  0.137838    
pop1517            1.9702e-10  9.3759e-11  2.1013  0.036534 *  
pop1820           -3.3094e-11  1.3109e-10 -0.2524  0.800887    
pop2124           -4.1101e-11  5.4777e-11 -0.7503  0.453700    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results:

- The above results show that the variables ‘unrate’, ‘spircons’, ‘perinc’ , ‘dry’ and ‘pop1517’ are significant at 5% significance level.
- Spircons refer to the annual pure alcohol consumption per head in gallons. It indicates that for every for 1 Gallon increase in it, there is 0.000085 increase in the vehicle fatality rate. The increase in the per capita personal income increases the vehicle fatality rate by a minute amount. But the effect is significant at 1% significance level. It makes complete sense because, as income increases people tend to buy personal vehicle which could result in the number of accidents occurring.
- Beertax has a negative effect on the vehicle fatality rate i.e. as the beertax goes up, the vehicle fatality rate decreases. But this effect is not significant.
- The variable ‘dry’ refers to the percentage of population residing in the dry counties. It’s estimate says that the increase in the % of people residing in the dry counties causes a certain increase in the vehicle fatality rate. This effect is not so obvious but it’s significant. This gives an intuition that alcohol has no major effect on the number of accident deaths that are occurring.

Time and Entity fixed effects model:

One of the disadvantages of the fixed effects model is that it does not capture the effects of time invariant variables and slow changing variables. Hence, we should try and run a fixed effects model with both time and entity.

Code:

```
> FixedtimeEntity <- plm(mrall ~ unrate + spircons + perinc + beertax + mormon +
                     yngdrv + dry + vmiles + punish +drinkagec + gspch + pop1517 +
                     pop1820 + pop2124 + factor(year),index=c("state","year"),
                     model="within", data=context1)

> #White's clustered robust errors
> coeftest(FixedtimeEntity, vcov. = vcovHC(FixedtimeEntity, type = 'HC1', cluster =
"group"))
```

```
t test of coefficients:

                    Estimate  Std. Error t value  Pr(>|t|)    
unrate           -5.3106e-06  1.2958e-06 -4.0982 5.538e-05 ***
spircons          8.3251e-05  1.3446e-05  6.1914 2.260e-09 ***
perinc            8.6787e-09  3.6147e-09  2.4010 0.0170417 *  
beertax          -3.9529e-05  2.8537e-05 -1.3852 0.1671639    
mormon            1.0988e-06  4.5732e-06  0.2403 0.8103006    
yngdrv           -9.7731e-06  1.0234e-04 -0.0955 0.9239920    
dry               1.7883e-06  1.0245e-06  1.7455 0.0820562 .  
vmiles            1.0674e-09  6.9334e-10  1.5395 0.1248752    
punish1           3.0569e-06  1.0329e-05  0.2960 0.7674890    
drinkagec[18,19) -2.4790e-06  7.8338e-06 -0.3164 0.7519119    
drinkagec[19,20) -3.5355e-06  4.6364e-06 -0.7625 0.4464115    
drinkagec[20,21)  1.2698e-06  4.0423e-06  0.3141 0.7536756    
gspch             3.5112e-05  3.6696e-05  0.9568 0.3395285    
pop1517           7.3140e-11  8.9893e-11  0.8136 0.4165894    
pop1820          -1.1437e-10  1.6824e-10 -0.6798 0.4972426    
pop2124           1.1207e-10  8.2023e-11  1.3663 0.1729917    
factor(year)1983 -7.0468e-06  4.2291e-06 -1.6663 0.0968445 .  
factor(year)1984 -1.9039e-05  5.8224e-06 -3.2699 0.0012182 ** 
factor(year)1985 -2.0381e-05  5.9977e-06 -3.3981 0.0007829 ***
factor(year)1986 -4.2033e-06  8.5290e-06 -0.4928 0.6225458    
factor(year)1987 -9.1757e-06  9.9775e-06 -0.9196 0.3585946    
factor(year)1988 -1.4534e-05  1.2331e-05 -1.1786 0.2396033    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results:

- From the above results, you can see that the time effects are significant but not all of them. Only the years 1983, 84 and 85 are significant. This represents that the groups do change over time.
- Here the other explanatory variables like spircons, unrate and perinc are significant and result in slight increase in the vehicle fatality rate in any given state.
- For every 1 Gallon increase in the pure alcohol consumption per head results in 0.000083 increase in the vehicle fatality rate and for every $1000 increase in the per capita income results in increase in the vehicle fatality rate by 0.000086.

Random Effects Model:

Random effects model is used when the sample is a small part (i.e. fraction) of the population. Our data is about the vehicle fatalities in 48 states across US out of the 50 states. Hence random effects model is not recommended. To confirm this, we perform Hausman's test to see if there is endogeneity present in the data.

```
> phtest(FixedtimeEntity, RandomEffectModel)
```

Hypothesis Testing:

H0: There is no endogeneity in the data, hence we can run random effects model.

H1: There is endogeneity in the data, hence, we need to run fixed effects model.

Hausman Test Results:

```
data:  mrall ~ unrate + spircons + perinc + beertax + mormon + yngdrv +  ...
chisq = 39.16, df = 7, p-value = 1.822e-06
alternative hypothesis: one model is inconsistent
```

- The results are synonymous with our expectation that we cannot run random effects model on this dataset as the p-value is <0.05. Hence, we can reject the null hypothesis and conclude that the fixed effects model is the most appropriate one.

Alcohol Involved Fatality Rate:

Here, we analyze the effect of the socio-economic factors and the drunk driving laws on the alcohol involved fatality rates. Firstly, we performed pooled regression followed by fixed effects model and time and entity fixed effects model.

Pooled OLS with beertax as explanatory variable:

```
> #Pooled OLS of Number of fatalities against beertax alone
> #To see if beertax explains any variation in the number of fatalities occurred
> d21 <- plm (mraidall ~ beertax, index=c("state","year"), model="pooling",
              data=table3)
```

```
Coefficients:
              Estimate Std. Error t-value  Pr(>|t|)
(Intercept) 5.7904e-05 2.0008e-06 28.9396 < 2.2e-16 ***
beertax     1.5726e-05 2.8510e-06  5.5158 6.982e-08 ***
```

Results:

- In the above regression results, we are trying to see if beertax alone explains any variation in the number of alcohol fatalities occurred over a period of 7 years in 48 states.
- It shows that beertax is highly significant, but the effect is not normal i.e. for every 1 unit increase in the beertax, there is certain increase in the alcohol involved fatality rate. We have seen this in the exploratory data analysis. But we will also run a pooled OLS along with other explanatory variables.

Pooled OLS:

```
> d23<-plm(mraidall ~ perinc + beertax + spircons + unrate + mormon + yngdrv + dry +
                      punish + drinkagec + pop1517 + pop1820 + pop2124,
                      index=c("state","year"), model="pooling", data=table3)
```

```
Unbalanced Panel: n = 48, T = 6-7, N = 335

Residuals:
      Min.     1st Qu.      Median     3rd Qu.        Max.
-4.5191e-05 -1.2901e-05 -2.4213e-06  9.0325e-06  9.3944e-05

Coefficients:
                     Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)        1.3525e-04  1.8649e-05  7.2524 3.103e-12 ***
perinc            -6.8687e-09  9.4900e-10 -7.2378 3.406e-12 ***
beertax            2.3593e-06  2.7233e-06  0.8663  0.386959
spircons           5.4193e-06  2.0542e-06  2.6382  0.008742 **
unrate            -1.2538e-06  6.5377e-07 -1.9178  0.056021 .
mormon            -6.7672e-07  1.2524e-07 -5.4033 1.280e-07 ***
yngdrv             1.1989e-04  5.5481e-05  2.1609  0.031443 *
dry                2.9396e-07  1.3536e-07  2.1716  0.030617 *
punish1            1.2676e-05  2.6164e-06  4.8449 1.978e-06 ***
drinkagec[18,19)  -9.0803e-06  5.6552e-06 -1.6056  0.109340
drinkagec[19,20)   5.5002e-06  3.0552e-06  1.8002  0.072764 .
drinkagec[20,21)  -9.8761e-06  4.1788e-06 -2.3634  0.018708 *
pop1517            6.4327e-11  8.1581e-11  0.7885  0.430981
pop1820           -2.3020e-10  1.5221e-10 -1.5124  0.131418
pop2124            1.2351e-10  7.3053e-11  1.6906  0.091884 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    2.2542e-07
Residual Sum of Squares: 1.2693e-07
R-Squared:      0.43692
Adj. R-Squared: 0.41228
F-statistic: 17.7356 on 14 and 320 DF, p-value: < 2.22e-16
```

Results:

- The regression results show that most of the variables are significant. However, the beertax which was significant in the previous regression has become insignificant now. This could be due to the fact that it was overestimated in the previous regression.
- The per capita income has a negative effect here i.e. as income increases, the number of alcohol involved fatality rate decreases. This seems to be plausible because high per capita income refers to high education which could make people aware of road safety rules and act responsibly on public roads.
- Punish1 is a variable which takes the value 1 if there is any one of the punishment like jail or community service. This has a positive effect on the fatality rate i.e. the fatality rate is high when there is a mandatory punishment rule than when there is no such rule.

We also performed Breusch-Pagan Test in to check if there is any heteroskedasticity in the data.

Heteroskedasticity: The variance of an explanatory variable increases the variance of the error term increases.

Implications

1. Least square estimators are no longer best estimators although unbiased and consistent
2. The standard errors are incorrect hence incorrect confidence interval and hypothesis testing

The following results show the results of the Breusch-Pagan test:

```
> lmtest::bptest(d23)

        studentized Breusch-Pagan test

data:  d23
BP = 30.202, df = 14, p-value = 0.007159
```

The null hypothesis and alternative hypothesis in a Breusch-Pagan test are:

H0: There is no heteroskedasticity

H1: There is heteroskedasticity in the data

- Here, we can see that the p-value 0.007159 < 0.05. So, we can reject the null hypothesis at 5% significance level and conclude that there is heteroskedasticity in the data.
- In order the remove the heteroskedasticity in the pooled OLS, White's robust standard errors is calculated.

Pooled OLS with clustered robust errors:

```
t test of coefficients:

                    Estimate  Std. Error t value  Pr(>|t|)    
(Intercept)       1.3525e-04  3.3500e-05  4.0373 6.772e-05 ***
perinc           -6.8687e-09  1.8055e-09 -3.8043 0.0001703 ***
beertax           2.3593e-06  3.8230e-06  0.6171 0.5375890    
spircons          5.4193e-06  2.2836e-06  2.3732 0.0182275 *  
unrate           -1.2538e-06  1.1421e-06 -1.0978 0.2731168    
mormon           -6.7672e-07  1.9929e-07 -3.3957 0.0007709 ***
yngdrv            1.1989e-04  9.4218e-05  1.2725 0.2041283    
dry               2.9396e-07  2.5931e-07  1.1336 0.2578086    
punish1           1.2676e-05  4.8621e-06  2.6072 0.0095571 ** 
drinkagec[18,19) -9.0803e-06  4.9739e-06 -1.8256 0.0688436 .  
drinkagec[19,20)  5.5002e-06  6.7960e-06  0.8093 0.4189243    
drinkagec[20,21) -9.8761e-06  4.9705e-06 -1.9869 0.0477813 *  
pop1517           6.4327e-11  8.9329e-11  0.7201 0.4719790    
pop1820          -2.3020e-10  1.5227e-10 -1.5118 0.1315621    
pop2124           1.2351e-10  6.8136e-11  1.8126 0.0708267 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- We cannot be sure that coefficients in pooled OLS give us the correct estimate because in pooled OLS we cannot account or control for unobserved heterogeneity, which makes the estimate biased and inconsistent.

Fixed Effects Model:

Code:

```
FixedEffectModel <- plm(mraidall ~ unrate + spircons + perinc + beertax + mormon +
                    yngdrv + dry + vmiles + punish + drinkagec + gspch + pop1517 +
                    pop1820 + pop2124,index=c("state","year"), model="within",
                    data=context33)
```

White's clustered robust standard errors:

```
> coeftest(FixedEffectModel, vcov. = vcovHC(FixedEffectModel, type = 'HC1', cluster
= "group"))
```

```
t test of coefficients:

                    Estimate  Std. Error t value  Pr(>|t|)    
unrate           -1.5821e-06  7.9942e-07 -1.9791 0.0488170 *  
spircons          3.3594e-05  8.8559e-06  3.7934 0.0001833 ***
perinc            3.6026e-10  1.7840e-09  0.2019 0.8401124    
beertax          -2.9764e-05  2.4093e-05 -1.2354 0.2177616    
mormon           -2.8562e-06  2.8140e-06 -1.0150 0.3110039    
yngdrv            1.0213e-04  5.4700e-05  1.8670 0.0629770 .  
dry               3.0488e-07  2.1146e-06  0.1442 0.8854646    
vmiles           -3.3503e-10  3.6033e-10 -0.9298 0.3533143    
punish1           4.6049e-06  7.6221e-06  0.6041 0.5462530    
drinkagec[18,19) -3.3870e-06  5.4873e-06 -0.6172 0.5375904    
drinkagec[19,20)  6.1346e-07  3.0635e-06  0.2002 0.8414359    
drinkagec[20,21) -1.4257e-06  2.4322e-06 -0.5862 0.5582457    
gspch            -1.5741e-05  2.3045e-05 -0.6831 0.4951499    
pop1517          -6.9261e-11  8.0614e-11 -0.8592 0.3910076    
pop1820           1.0067e-10  1.0932e-10  0.9208 0.3579554    
pop2124          -2.7581e-11  4.3402e-11 -0.6355 0.5256499    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results:

- The variables unemployment rate is significant at 5% significance level and it has a negative effect on the alcohol involved fatality rates i.e. as the unemployment increases the fatality rate increases by a certain amount.
- The variable spircons (annual pure alcohol consumption per head) is highly significant just like it was in the total vehicle fatality rates earlier.
- The percentage of young drivers 'yngdrv' is also significant at 10% significance level.
- One of the disadvantages of the fixed effects model is that it does not capture the effects of time invariant variables and slow changing variables. Hence, we should try and run a fixed effects model with both time and entity.

Time and Entity Fixed Effects Model:

Code:

```
> FixedtimeEntity <- plm(mraidall ~ unrate + spircons + perinc + beertax + mormon +
                         yngdrv + dry + vmiles + punish +drinkagec + gspch + pop1517 +
                         pop1820 + pop2124 +  factor(year),index=c("state","year"),
                         model="within", data=context33)
```

White's clustered robust standard errors:

```
> coeftest(FixedtimeEntity, vcov. = vcovHC(FixedtimeEntity, type = 'HC1', cluster =
"group"))
```

```
t test of coefficients:

                     Estimate  Std. Error t value  Pr(>|t|)    
unrate            -2.4828e-06  1.1797e-06 -2.1046 0.0362705 *  
spircons           2.9035e-05  8.6243e-06  3.3667 0.0008737 ***
perinc            -7.3642e-10  2.1117e-09 -0.3487 0.7275711    
beertax           -3.0893e-05  2.2428e-05 -1.3774 0.1695411    
mormon            -2.8240e-06  2.2159e-06 -1.2744 0.2036245    
yngdrv             5.0524e-05  5.6198e-05  0.8990 0.3694521    
dry               -4.3939e-07  2.0202e-06 -0.2175 0.8279826    
vmiles            -3.8750e-10  4.1503e-10 -0.9337 0.3513270    
punish1            7.0752e-06  7.5841e-06  0.9329 0.3517245    
drinkagec[18,19)  -1.9336e-06  5.2848e-06 -0.3659 0.7147433    
drinkagec[19,20)   1.0188e-06  3.3897e-06  0.3005 0.7639956    
drinkagec[20,21)  -3.8715e-07  2.7599e-06 -0.1403 0.8885507    
gspch              5.5605e-05  4.1599e-05  1.3367 0.1824634    
pop1517           -1.7649e-10  8.4847e-11 -2.0801 0.0384736 *  
pop1820            1.3359e-10  9.0923e-11  1.4692 0.1429547    
pop2124            1.7176e-11  5.9823e-11  0.2871 0.7742505    
factor(year)1983  -9.7661e-06  3.6142e-06 -2.7022 0.0073339 ** 
factor(year)1984  -1.5906e-05  3.5245e-06 -4.5131 9.613e-06 ***
factor(year)1985  -1.6898e-05  4.7466e-06 -3.5601 0.0004393 ***
factor(year)1986  -8.4475e-06  7.3540e-06 -1.1487 0.2517221    
factor(year)1987  -1.2023e-05  9.1767e-06 -1.3102 0.1912639    
factor(year)1988  -1.2920e-05  1.0355e-05 -1.2478 0.2132196    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results:

- The above regression result shows that the variables unemployment rate and annual pure alcohol consumption per capita are significant at 5% and 1% significance levels respectively.
- Along with them, the variable pop1517 (the percentage of population of people aged between 15-17) also turned to be significant at 5% significance level.
- From the above results, you can see that the time effects are significant but not all of them. Only the years 1983, 84 and 85 are significant. This represents that the groups do change over time. These coefficients have a negative sign indicating that the alcohol involved fatality rate decreased in these years compared to what it was in the year 1982.

Random Effects Model:

Code:

```
> RandomEffectModel<- plm(mraidall ~ unrate + spircons + beertax + mormon + yngdrv
                      + dry + vmiles  + punish +drinkagec + gspch + pop1517 +
                      pop1820 + pop2124 ,index=c("state","year"),model="random",
                      data=context33)
```

Random effects model is used when the sample is a small part (i.e. fraction) of the population. Our data is about the vehicle fatalities in 48 states across US out of the 50 states. Hence random effects model is not recommended. To confirm this, we perform Hausman's test to see if there is endogeneity present in the data.

```
> phtest(FixedtimeEntity, RandomEffectModel)
```

Hypothesis Testing:

H0: There is no endogeneity in the data, hence we can run random effects model.

H1: There is endogeneity in the data, hence, we need to run fixed effects model.

Hausman Test Results:

```
data:  mraidall ~ unrate + spircons + perinc + beertax + mormon + yngdrv +  ...
chisq = 24.589, df = 7, p-value = 0.0008973
alternative hypothesis: one model is inconsistent
```

- The results are synonymous with our expectation that we cannot run random effects model on this dataset as the p-value is <0.05. Hence, we can reject the null hypothesis and conclude that the fixed effects model is the most appropriate one.

Night Time Fatality Rate:

Pooled OLS Model:

```
. reg mralln spircons unrate perinc beertax sobapt mormon dry vmiles jaild comserd pop pop1517 pop1820 pop2124 miles gspch, vce(cluster state)

Linear regression                               Number of obs     =        335
                                                F(10, 47)         =          .
                                                Prob > F          =          .
                                                R-squared         =     0.3445
                                                Root MSE          =    9.1e-06

                                  (Std. Err. adjusted for 48 clusters in state)
```

mralln	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
spircons	3.14e-06	1.49e-06	2.10	0.041	1.38e-07	6.14e-06
unrate	1.41e-07	4.64e-07	0.30	0.763	-7.92e-07	1.07e-06
perinc	-1.80e-09	7.06e-10	-2.55	0.014	-3.22e-09	-3.81e-10
beertax	-4.79e-06	2.37e-06	-2.02	0.049	-9.56e-06	-1.93e-08
sobapt	4.15e-07	1.63e-07	2.54	0.014	8.68e-08	7.43e-07
mormon	-1.39e-07	9.87e-08	-1.40	0.167	-3.37e-07	5.99e-08
dry	-2.92e-07	1.18e-07	-2.47	0.017	-5.31e-07	-5.40e-08
vmiles	7.12e-10	1.10e-09	0.65	0.522	-1.51e-09	2.93e-09
jaild	2.83e-06	3.51e-06	0.81	0.424	-4.23e-06	9.89e-06
comserd	3.64e-07	3.29e-06	0.11	0.912	-6.25e-06	6.98e-06
pop	-1.26e-12	2.47e-12	-0.51	0.612	-6.23e-12	3.71e-12
pop1517	3.68e-11	3.38e-11	1.09	0.281	-3.11e-11	1.05e-10
pop1820	-7.32e-11	7.10e-11	-1.03	0.307	-2.16e-10	6.95e-11
pop2124	2.97e-11	3.54e-11	0.84	0.407	-4.16e-11	1.01e-10
miles	1.78e-10	9.13e-11	1.95	0.058	-5.89e-12	3.61e-10
gspch	-.0000422	.0000208	-2.03	0.048	-.000084	-3.17e-07
_cons	.0000523	.0000141	3.72	0.001	.000024	.0000807

Results:

- In a pooled OLS, each observation is assumed as an independent observation rather than panel data of different entities/ groups. The explanatory variables such as unemployment rate, mandatory jail time , mandatory community service are insignificant at 10% level.
- For every 1 percentage increase in the southern Baptist population increases the night time fatality by a minute difference. This effect is a bit abnormal.
- The coefficient of Beer tax has a negative sign which signifies that the night time fatality rate decreases with the increase in the beertax which is plausible to believe. For every 1 unit increase in the beer tax, there will be approximately 0.0000048 units decrease in the night time fatality rate and it is significant at 5% significance level.

Fixed Effects Model:

```
. xtreg mralln spircons unrate perinc beertax sobapt mormon dry vmiles jaild comserd pop pop1517 pop1820 pop2124 miles gspch, fe vce(cluster state)

Fixed-effects (within) regression               Number of obs      =       335
Group variable: state                           Number of groups   =        48

R-sq:                                           Obs per group:
     within  = 0.1619                                         min =         6
     between = 0.0429                                         avg =       7.0
     overall = 0.0381                                         max =         7

                                                F(11,47)           =         .
corr(u_i, Xb)  = -0.9074                        Prob > F           =         .

                                (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      mralln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    spircons |   .0000146   5.34e-06     2.74   0.009     3.91e-06    .0000254
      unrate |  -8.19e-08   2.97e-07    -0.28   0.784    -6.80e-07    5.16e-07
      perinc |   1.45e-09   9.66e-10     1.50   0.141    -4.96e-10    3.39e-09
     beertax |  -.0000104   9.61e-06    -1.08   0.284    -.0000298    8.92e-06
      sobapt |   1.03e-06   2.68e-06     0.39   0.701    -4.36e-06    6.43e-06
      mormon |  -4.11e-07   2.10e-06    -0.20   0.846    -4.64e-06    3.81e-06
         dry |   4.77e-07   5.71e-07     0.84   0.408    -6.71e-07    1.62e-06
      vmiles |  -1.12e-09   5.17e-10    -2.18   0.035    -2.16e-09   -8.54e-11
       jaild |   6.75e-07   6.43e-07     1.05   0.299    -6.18e-07    1.97e-06
     comserd |  -4.77e-06   3.67e-06    -1.30   0.201    -.0000122    2.62e-06
         pop |  -1.20e-11   4.32e-12    -2.76   0.008    -2.06e-11   -3.26e-12
     pop1517 |   7.12e-11   2.87e-11     2.49   0.017     1.36e-11    1.29e-10
     pop1820 |   2.64e-11   3.53e-11     0.75   0.459    -4.47e-11    9.75e-11
     pop2124 |   1.03e-12   1.92e-11     0.05   0.958    -3.77e-11    3.97e-11
       miles |   5.25e-10   2.62e-10     2.00   0.051    -2.02e-12    1.05e-09
       gspch |  -.0000165   .0000131    -1.26   0.214    -.0000429    9.86e-06
       _cons |   .0000165   .0000369     0.45   0.657    -.0000578    .0000908
-------------+----------------------------------------------------------------
     sigma_u |  .00002233
     sigma_e |  6.193e-06
         rho |  .92860335   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Results:

- Mandatory jail time and mandatory community service are not significant even at 10% level. Seems like there is no effect of change in southern Baptist or Mormon population on night time fatality rate.
- Population between 15-17 years is a significant variable with positive sign which was insignificant in the previous pooled OLS model. As population between 15-17 years increases there is a certain increase in night time fatality rate but the coefficient is very small.
- Per capita annual consumption of pure alcohol has positive effect on the night time fatality rate. As the variable 'spircons' increases the night time fatality rate increases.
- The entity fixed effect model doesn't capture the effect of time invariant variables and slow changing variables. Therefore we go for Time and Entity Fixed Effects Model.

Time and Entity Fixed Effects Model:

```
. xtreg mralln spircons unrate perinc beertax sobapt mormon dry vmiles jaild comserd pop pop1517 pop1820 pop2124 miles gspch i.year, fe vce(cluster state)

Fixed-effects (within) regression               Number of obs     =        335
Group variable: state                           Number of groups  =         48

R-sq:                                           Obs per group:
     within  = 0.2533                                         min =          6
     between = 0.0113                                         avg =        7.0
     overall = 0.0007                                         max =          7

                                                F(17,47)          =          .
corr(u_i, Xb)  = -0.8351                        Prob > F          =          .

                                  (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      mralln |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    spircons |   9.00e-06   4.57e-06     1.97   0.055    -1.91e-07    .0000182
      unrate |  -9.08e-07   4.94e-07    -1.84   0.072    -1.90e-06    8.53e-08
      perinc |   3.76e-10   9.47e-10     0.40   0.693    -1.53e-09    2.28e-09
     beertax |  -.0000109     .00001    -1.09   0.282    -.0000311    9.27e-06
      sobapt |  -2.49e-07   2.25e-06    -0.11   0.913    -4.78e-06    4.28e-06
      mormon |  -8.25e-07   1.72e-06    -0.48   0.633    -4.28e-06    2.63e-06
         dry |   2.46e-07   4.89e-07     0.50   0.618    -7.38e-07    1.23e-06
      vmiles |  -7.02e-10   5.42e-10    -1.30   0.202    -1.79e-09    3.88e-10
       jaild |   1.32e-06   6.72e-07     1.96   0.056    -3.36e-08    2.67e-06
     comserd |  -3.86e-06   3.57e-06    -1.08   0.285     -.000011    3.32e-06
         pop |  -5.00e-12   5.10e-12    -0.98   0.332    -1.53e-11    5.26e-12
     pop1517 |  -4.65e-12   3.16e-11    -0.15   0.883    -6.82e-11    5.89e-11
     pop1820 |   2.55e-12   5.66e-11     0.04   0.964    -1.11e-10    1.16e-10
     pop2124 |   3.29e-11   3.56e-11     0.92   0.361    -3.88e-11    1.05e-10
       miles |   2.97e-10   2.78e-10     1.07   0.291    -2.62e-10    8.56e-10
       gspch |    .000018   .0000265     0.68   0.500    -.0000353    .0000714
             |
        year |
        1983 |  -4.38e-06   2.64e-06    -1.66   0.104    -9.70e-06    9.38e-07
        1984 |  -9.29e-06   2.74e-06    -3.39   0.001    -.0000148   -3.78e-06
        1985 |  -9.29e-06   2.27e-06    -4.10   0.000    -.0000139   -4.73e-06
        1986 |  -5.95e-06   2.85e-06    -2.09   0.042    -.0000117   -2.24e-07
        1987 |  -7.45e-06   3.36e-06    -2.21   0.032    -.0000142   -6.83e-07
        1988 |  -9.17e-06   3.51e-06    -2.62   0.012    -.0000162   -2.12e-06
             |
       _cons |    .000048   .0000323     1.48   0.144     -.000017    .0001131
-------------+----------------------------------------------------------------
     sigma_u |  .00001761
     sigma_e |  5.911e-06
         rho |  .89877393   (fraction of variance due to u_i)
```

Results:

- From the above results, you can see that the all the time effects are significant at 5% significance level except for 1983 which is significant at 10% significance level. This represents that the groups do change over time. These coefficients have a negative sign indicating that the alcohol involved fatality rate decreased in these years compared to what it was in the year 1982.
- Unemployment rate, Per capita annual pure alcohol consumption and jaild are the other explanatory variables that are significant at 5% significance level. The rest of the variables are insignificant.
- These results show that the drunk and drive laws and policies hardly have any effect on the vehicle fatality rates during night time.

Random Effects Model:

Random effects model is used when the sample is a small part (i.e. fraction) of the population. Our data is about the vehicle fatalities in 48 states across US out of the 50 states. Hence random effects model is not recommended. To confirm this, we perform Hausman's test to see if there is endogeneity present in the data.

```
. xtreg mralln spircons unrate perinc beertax sobapt mormon dry vmiles jaild comserd pop pop1517 pop1820 pop2124 miles gspch, re cluster(state)

Random-effects GLS regression                   Number of obs     =        335
Group variable: state                           Number of groups  =         48

R-sq:                                           Obs per group:
     within  = 0.0947                                         min =          6
     between = 0.3362                                         avg =        7.0
     overall = 0.2615                                         max =          7

                                                Wald chi2(9)      =          .
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =          .

                                 (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      mralln |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    spircons |   4.25e-06   1.99e-06     2.13   0.033     3.42e-07    8.16e-06
      unrate |  -3.90e-07   3.17e-07    -1.23   0.218    -1.01e-06    2.31e-07
      perinc |  -1.20e-09   4.73e-10    -2.53   0.011    -2.13e-09   -2.71e-10
     beertax |  -6.24e-06   2.77e-06    -2.25   0.024    -.0000117   -8.10e-07
      sobapt |   5.77e-07   1.68e-07     3.44   0.001     2.48e-07    9.05e-07
      mormon |  -3.58e-08   1.07e-07    -0.33   0.739    -2.46e-07    1.75e-07
         dry |  -2.86e-07   1.29e-07    -2.21   0.027    -5.39e-07   -3.19e-08
      vmiles |  -4.59e-10   3.58e-10    -1.28   0.200    -1.16e-09    2.43e-10
       jaild |   2.60e-06   3.15e-06     0.82   0.410    -3.58e-06    8.77e-06
     comserd |  -3.08e-06   3.70e-06    -0.83   0.404    -.0000103    4.17e-06
         pop |  -4.44e-12   1.87e-12    -2.37   0.018    -8.11e-12   -7.77e-13
     pop1517 |   3.27e-11   3.00e-11     1.09   0.275    -2.61e-11    9.16e-11
     pop1820 |  -1.41e-11   4.20e-11    -0.34   0.736    -9.65e-11    6.82e-11
     pop2124 |   2.77e-11   2.47e-11     1.12   0.262    -2.07e-11    7.61e-11
       miles |   2.32e-10   9.00e-11     2.57   0.010     5.52e-11    4.08e-10
       gspch |  -.0000284   .0000135    -2.10   0.036    -.0000549   -1.92e-06
       _cons |   .0000554   9.49e-06     5.83   0.000     .0000368     .000074
-------------+----------------------------------------------------------------
     sigma_u |  7.091e-06
     sigma_e |  6.193e-06
         rho |  .56727986   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- In the Random effects model on comparison to Fixed effects model most of the variables such as perinc , sobapt , Mormon , growth rate, dry are significant at 5 % level .
- Percentage of southern Baptist variable has a positive coefficient and significant which implies night time fatality rate increases in states with high southern Baptist population.
- Beer tax and perinc have negative sign , as beer taxes and per capita income increases there is a decrease in night time fatality rates.

```
. hausman fe re , sigmamore

Note: the rank of the differenced variance matrix (9) does not equal the number of coefficients being tested (16); be sure this is what you expect, or there may be problems computing the test.
      Examine the output of your estimators for anything unexpected and possibly consider scaling your variables so that the coefficients are on a similar scale.

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      fe           re           Difference          S.E.
-------------+----------------------------------------------------------------
    spircons |   .0000146     4.25e-06        .0000104        3.21e-06
      unrate |  -8.19e-08    -3.90e-07        3.09e-07        2.43e-07
      perinc |   1.45e-09    -1.20e-09        2.65e-09        6.80e-10
     beertax |  -.0000104    -6.24e-06       -4.18e-06        7.32e-06
      sobapt |   1.03e-06     5.77e-07        4.58e-07        2.38e-06
      mormon |  -4.11e-07    -3.58e-08       -3.75e-07        1.75e-06
         dry |   4.77e-07    -2.86e-07        7.62e-07        5.20e-07
      vmiles |  -1.12e-09    -4.59e-10       -6.66e-10        6.42e-10
       jaild |   6.75e-07     2.60e-06       -1.92e-06        4.24e-06
     comserd |  -4.77e-06    -3.08e-06       -1.68e-06        4.88e-06
         pop |  -1.20e-11    -4.44e-12       -7.51e-12        6.65e-12
     pop1517 |   7.12e-11     3.27e-11        3.85e-11        2.24e-11
     pop1820 |   2.64e-11    -1.41e-11        4.06e-11        3.83e-11
     pop2124 |   1.03e-12     2.77e-11       -2.66e-11        1.49e-11
       miles |   5.25e-10     2.32e-10        2.93e-10        3.81e-10
       gspch |  -.0000165    -.0000284        .0000119        4.77e-06
------------------------------------------------------------------------------
                        b = consistent under Ho and Ha; obtained from xtreg
         B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       21.48
                Prob>chi2 =      0.0107
```

- As general Hausman test fails on these models had to use sigma more option because of clash between finite samples and asymptotic assumption of Hausman test. Based on the p-value which is less than 0.05, we can safely reject the null hypothesis at 5% significance level and can conclude that we can use fixed effects model for our analysis.

Discussion:

In Pooled OLS model, it does not account for fixed effects of states and time effects. To control the unobserved heterogeneity such as cultural attitude, alcoholism and crime prevention programs we move to Entity and Time Fixed Effect models. The estimated parameters of beertax in pooling method shows a certain effect on the fatality rates which vanishes when we move to the fixed effects model. However, this initial effect is due to omitted variable bias and the effect disappears when state and time effects are included in the model.

Limitations of entity and Time fixed effect models

- There might be unobserved heterogeneity in the regression model that vary between states and over time. For example, other strategies that are related to the of drunk and drive laws implementation and that affect the road accidents and deaths. There is a serious risk of simultaneous causality bias.

Conclusion:

Using a panel data of 48 states of Unites States from 1982 – 1988 and an extensive set of control variables, we analyzed the impact of socio-economic factors, political factors, laws and regulations on traffic fatalities using panel regression with state and time fixed effects. The hypothesis results say that certain demographic characteristics play an important role in suppressing the positive effects of the drunk and drive laws and policies. From the analysis, we can conclude that the socio-economic factors like population, unemployment rate, per capita annual pure alcohol consumption, percentage of young drivers have huge impact on the traffic fatalities occurred. The drunk and drive policies have hardly any impact on the number of traffic deaths. This analysis lays foundation for several suggestions to mitigate the number of traffic deaths like introducing stringent drunk and drive laws specific to a state with high population growth and high crime rates. Another suggestion could be introducing weather related driving rules like lower maximum speed limit during a bad weather condition in certain states with high rainfall and chances of snow.