**BUAN 6312.002**

**Applied Econometrics**

**and Time Series Analysis**

**Group 11 Project**

**Prepared By:**

Avinash Panigrahi

Chirag Hamirani

Chirag Sardana

Edward Nonso-Anyawko

Linh Diep

Vedant Taneja

****

**Effect of Alcohol Regulation on Traffic Deaths**

Table of Contents

[Introduction 2](#_Toc57743727)

[Data Description 3](#_Toc57743728)

[Analysis 4](#_Toc57743729)

[Correlation Test 4](#_Toc57743730)

[Transformation 4](#_Toc57743731)

[Regression Analysis 5](#_Toc57743732)

[Simple Linear Regression 5](#_Toc57743733)

[Panel Regression 6](#_Toc57743734)

[Exploration by regressing with economic variables. 7](#_Toc57743735)

[Exploration with no economic variables 8](#_Toc57743736)

[Exploration of the Joint effect of Jail time and Community Service variables 8](#_Toc57743737)

[Exploration with only significant variables. 9](#_Toc57743738)

[Conclusion 10](#_Toc57743739)

[References 11](#_Toc57743740)

[Code File 12](#_Toc57743741)

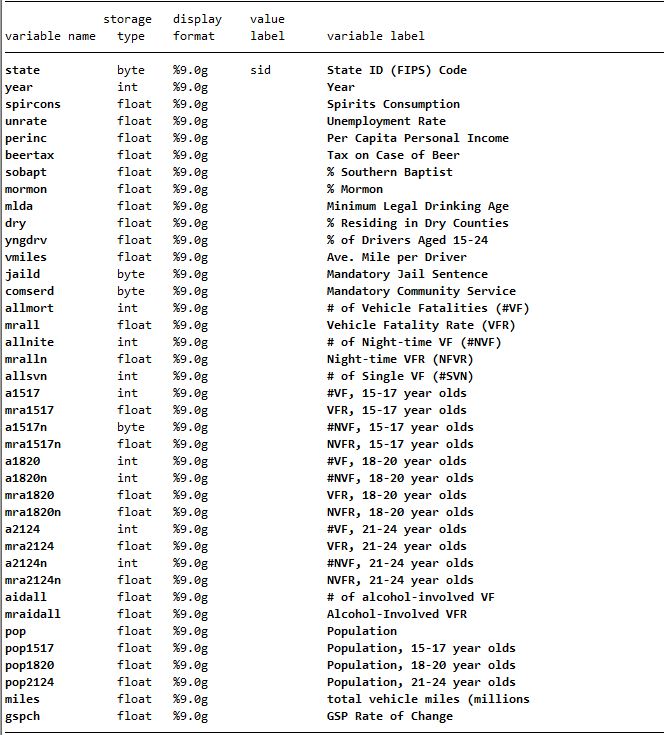
# Introduction

Vehicle fatality has been a major issue in the United States since the advent of automobiles. Whether the average citizen commuting from work, the trucker making the long drive between warehouses, or the government agent who oversees all of them, driving fatality incurs not only the tragic loss of human life but problematic disruptions and high costs. A study in 2014 conducted by the CDC on motor vehicle injuries found that American citizens “spend more than one million days in the hospital each year from crash injuries” (*CDC* *Vital Signs*, 2014). Additionally, the study found that in 2012, crash injuries cost “$18 billion in lifetime medical costs,” while “lifetime work lost because of crash injuries” was approximately $33 billion (*CDC* *Vital Signs*, 2014).

The vast majority of driving fatalities occur either because of substance abuse or negligent driving. In fact, for alcohol specifically, it was reported in 2014 that some 9,000 people were killed in alcohol-impaired driving crashes, comprising around 31% of total traffic-related deaths in the United States. With almost one-third of all traffic deaths made up by drunk driving incidents, it is little wonder that not only government organizations like the CDC but international non-profits like Mothers Against Drunk Driving have found the issue to be salient through the decades, and especially with pushes by the latter, government regulations have been enacted time and again in an effort to minimize and prevent such deaths.

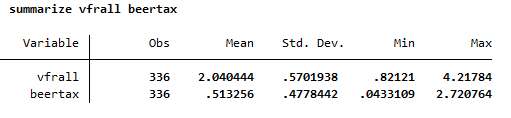
Such alcohol regulations, both as direct and indirect efforts to address alcohol-related driving fatalities, take many forms: minimum drinking ages, taxes on sales, imports, and manufacture, mandatory sentences, and so on (*Alcohol Policy*, 2019). It is, however, uncertain at times as to whether these regulations fulfill their intended purpose. Using the above data, pulled from the US Department of Transportation, the purpose of this project is to answer the following question: **how does alcohol regulation like taxes on beer affect traffic deaths**?

# Data Description

****

The dataset *car\_fatalities* is a balanced panel dataset, recording vehicle fatalities occurring in the years 1982 through 1988 in 48 states. The set consists of a total of 336 observations of 39 variables.

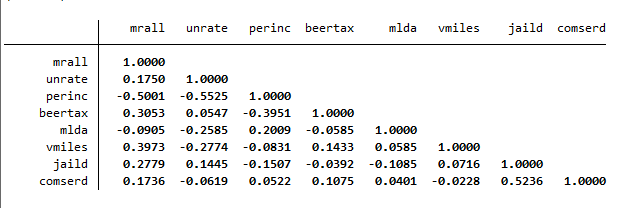
For this project, the specific variables of interest are *unrate*, *perinc*, *beertax*, *vmiles*, *mlda*, *jaild*, and *comserd*, with *mrall* (transformed into *vfrall*)being the dependent variable and *beertax* being the primary explanatory variable of interest. The additional variables were selected after conducting exploratory data analysis; a correlation matrix was generated to determine which variables to include in the regression for further insights. For the variables of primary interest, *vfrall* and *beertax*, see the summary statistics:



# Analysis

## Correlation Test

For our initial analysis, we checked correlation between the features to determine whether there is any linear correlation between the dependent and the independent variables. We find out that *beertax*(Tax on Case of Beer)is correlated with *sobapt* (% of Southern Baptist), which is related to *dry*(% residing in dry counties), this implies that states that have a higher Christian population have higher tax. Also, we found that the higher is the population, more is the demand for *allnite* drivers (No. of night time vehicle fatalities), *allnite* also seems to have a higher correlation with *miles* indicating that all night drivers travel more comparatively.

****

## Transformation

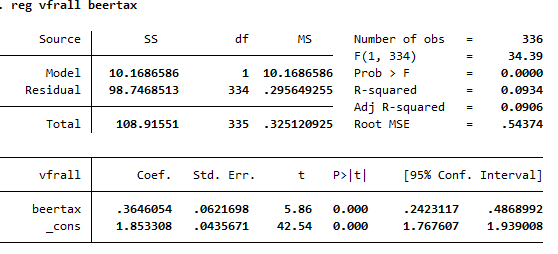
We have transformed a few variables to have an easier understanding of the model due to their units as well as to keep the variables in a standard form, so that we don’t have a bias in the model.

* The dependent variable *mrall* (Traffic Fatality Rate) which is defined in the dataset as vehicle fatality rate per 10000 people, we transform it by multiplying 10000 to show traffic deaths every 10000 people.
* We took the natural log of *perinc* because it is an income variable and its effect fall over time and distribution is right skewed.
* Used *vmiles* (Avg Miles per Driver) in units of thousands to interpret the data in a better way.
* Combined *jaild* and *comserd* into one dummy variable *punish*; if either *jaild* or *comserd* equal 1, then *punish* is 1, otherwise *punish* is 0.

## Regression Analysis

### Simple Linear Regression

To start with, we ran a normal regression model and it is quite evident that we can’t use this model to check for coefficients as the **value of R-squared is low**. Since, the data is and suitable to use for panel regression, we must look for fixed or random effects in our data.



### Panel Regression

Since we have balanced panel data and since we did not get any success with traditional regression analysis. We went ahead to perform panel data regression using 2 approaches.

* **Fixed Effects**: There are unique, time constant attributes of individuals that may or may not correlate with individual regressors.
* **Random Effects**: There are unique, time constant attributes of individuals that dont correlate with individual regressors

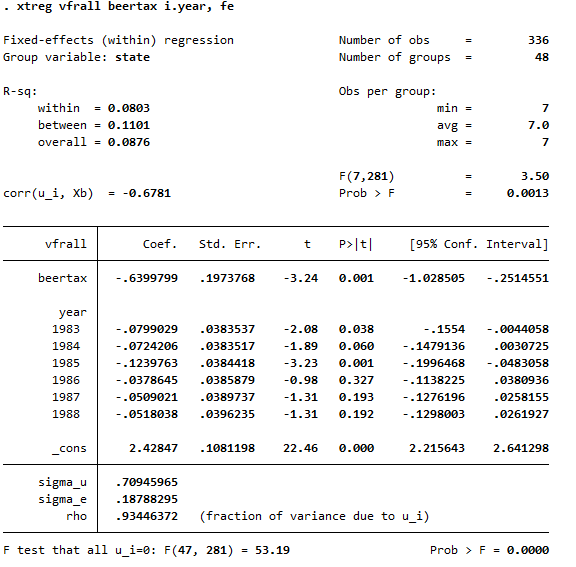
Then, taking the estimates from these two models, we used a Hausman test to determine which would be the most appropriate to draw conclusions from. As the results of the Hausman test showed a *prob>chi2* of 0.0000, which is less than 5, the fixed effects model appears to be the more appropriate of the two.

Among all the variables used for **fe** model, *beertax* was the primary interest variable. We ran a **panel regression with fixed effects** to see how it varies with the dependent variable *vfrall.*

We see that the effect of *beertax* is very much significant. The inference that we derived by running this model is as follows:

* Whenever the tax on the increases by $1, the number of traffic related deaths will likely decrease by .64 per 10,000.

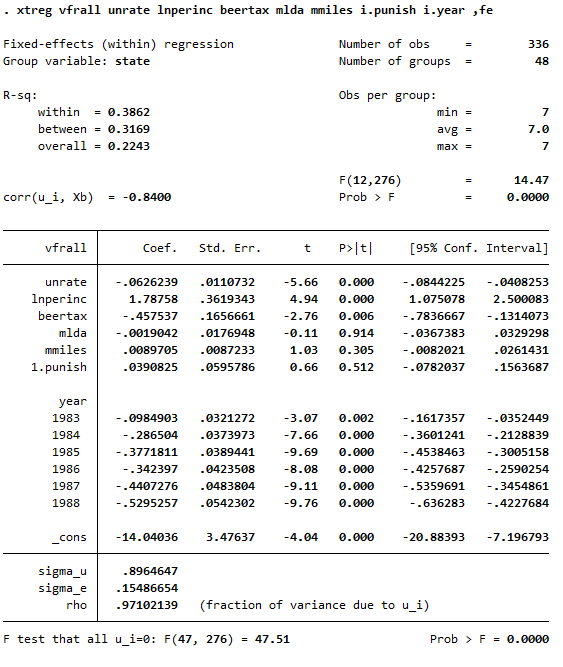
The fixed effects model using *beertax* is as follows:



### Exploration by regressing with economic variables.

After finding out the desired results on how *beertax* affects the fatality rate. We tried to explore more by performing panel regression with fixed effects keeping state and time fixed.

We were able to discover that *unrate,* *lnperinc (natural log of personal income per capital), and beertax* are significant at 1% while other variables are not that significant.



We got the following inferences:

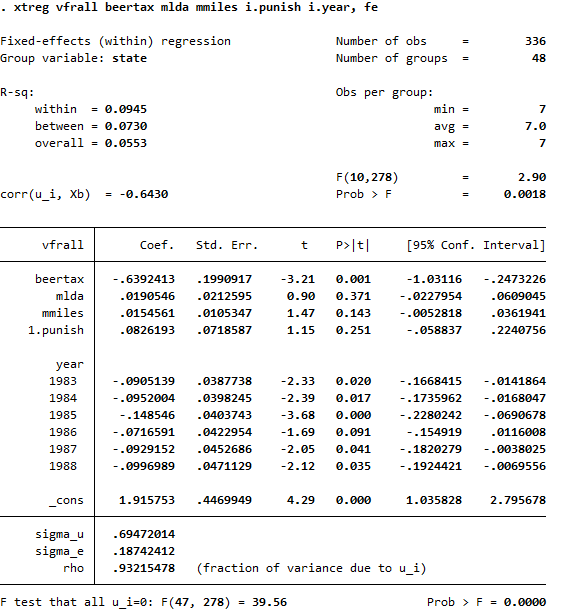
* Whenever there is a 1% increase in the *unemployment rate*, we can see the number of traffic related deaths likely to decrease by 0.0626 per 10000.
* Whenever there is a 1% increase in the *personal income*, we can see the number of traffic related deaths likely to increase 1.78 per 10000.
* Whenever there is a $1 increase in the *tax of beer*, we can see the number of traffic related deaths likely to decrease by 0.46 per 10000.
* Whenever there is a 1-year increase in the *minimum drinking age*, we can see the number of traffic related deaths likely to decrease by 0.0019 per 10000.
* When average miles increase by 1000 miles in a year, the number of traffic related deaths likely to increase by 0.009 per 10000.
* When the state has *jail sentence* or *community service* mandatory, the number of traffic related deaths likely to increase by 0.04 per 10000.

### Exploration with no economic variables

In this regression model we removed the economic variables *unrate* & *lnperinc.*  The following changes in output coefficients follows:

* Whenever there is a $1 increase in the *tax of beer*, we can see the number of traffic related deaths likely to decrease by 0.64 per 10000.
* Whenever there is a 1-year increase in the *minimum drinking age*, we can see the number of traffic related deaths likely to increase by 0.0019 per 10000.
* When average miles increase by 1000 miles in a year, the number of traffic related deaths likely to increase by 0.015 per 10000.
* When the state has *jail sentence* or *community service* mandatory, the number of traffic related deaths likely to increase by 0.08 per 10000.

### Exploration of the Joint effect of Jail time and Community Service variables

* In this regression, we take the joint effect of the mandatory jail sentence or mandatory community service into consideration.
* In brief we included a dummy variable called *punish* to describe the two binary variables. The new variable measures whether there are any laws that dictate punishment.
* Either one of the punishment variables are included in the laws, the new variable is supposed to tell us if there is a punishment associated with drunk driving for a given state.
* We don’t see much change in the coefficient of *beertax* or other variables.
* Hence, we conclude that this model is not **sensitive to the joint effect of the two stated punishments**.

### Exploration with only significant variables.

In this model we kept only those variables which were significant in the model where we ran panel regression without economic variables. We found that the three of these variables increase by a little because they inherit the effect of the other 4 variables in the previously ran model. The three variables seem to retain their significance and we can see that this model has the highest observed R square among all the models we ran till now.

Table

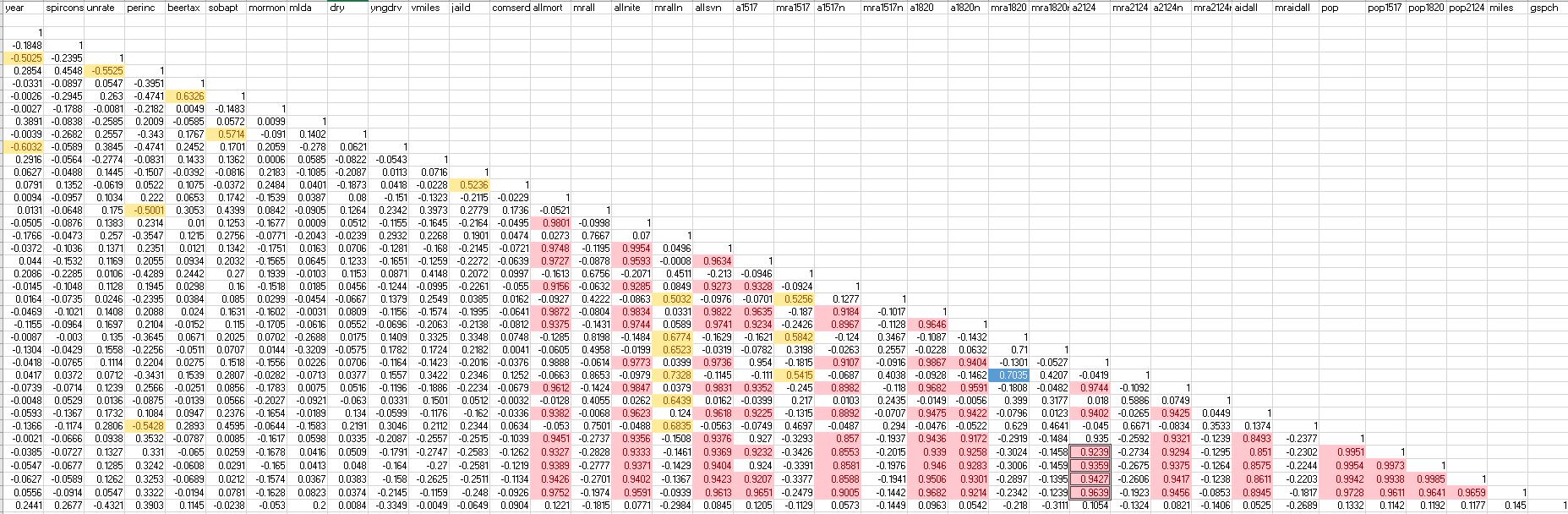
Description automatically generated

# Conclusion

Based on our analysis we conclude that drunk driving laws affect the traffic related casualties. Our models show that variable beer tax significantly affect the traffic deaths. On further exploration we discovered that two other variables unemployment rate and personal income along with beer tax affect the number of deaths in traffic related accidents. Unemployment rate and beer tax decreases number of deaths while personal income increases. When tax on beer increases by 1 dollar, the number of deaths in a traffic accident decreases by 0.45 per 10000 people. We can deduce convincingly that increasing the beer tax will likely lead to the decrease in traffic related fatalities. In addition to that, we also found that mandatory jail sentence and community service were not significant. Any drunk driving laws that leads to the increase in beer tax would reduce the traffic fatality rate.

However, our model is far from being perfect may be because of high possibility that biased omitted variables might exist or other factors which might have affected the fatality rate but are absent in the dataset. As beer tax is the only explanatory variable that represent drunk driving laws in this dataset, inclusion of other variables which is affects drunk driving laws will greatly improve the R square and in turn improve the understandability of the model.

# References

****National Highway Traffic Safety Administration. (2020). *Fatality Analysis Reporting System Data*. [Data file].

Retrieved from <https://www.nhtsa.gov/node/97996/251>

NIH, N. (2019). Alcohol Policy.

Retrieved November 2, 2020, from <https://www.niaaa.nih.gov/alcohols-effects-health/alcohol-policy>

Unknown, U. (2014, October 07).

Motor Vehicle Crash Injuries.

Retrieved November 2, 2020, from <https://www.cdc.gov/vitalsigns/crash-injuries/index.html>

## Code File

clear

\*\* Make sure to open the dataset.

use "H:\Vedant\Econ\car\_fatalities.dta"

\*EDA

summarize

correlate mrall unrate perinc beertax mlda vmiles jaild comserd

describe

\*variable transformation

gen vfrall=mrall\*10000

gen mmiles=vmiles/1000

gen lnperinc=ln(perinc)

gen punish=0

replace punish = 1 if jaild==1

replace punish =1 if comserd==1

summarize vfrall beertax

\*Correlation Test

corr mrall unrate perinc beertax mlda vmiles jaild comserd

\*dry Regression model 1

reg vfrall beertax

\*PANEL DATA (Fixed effect State fixed) model 2

xtset state year

xtreg vfrall unrate lnperinc beertax mlda mmiles i.jaild i.comserd, fe

estimates store fixed

areg vfrall unrate lnperinc beertax mlda mmiles i.jaild i.comserd, absorb(state)

\*PANEL DATA (Random effect)

xtreg vfrall unrate lnperinc beertax mlda mmiles i.jaild i.comserd, re ( screenshot needed)

estimates store random

\*HAUSMAN TEST

hausman fixed random

\*Moving forward with FIXED EFFECTS MODEL for beertax

xtreg vfrall beertax,fe

\*Exploratory models - examining correlation and other possible explanatory variables.

\*PANEL DATA (Fixed effect State and Time fixed)

\* model 3

xtreg vfrall beertax i.year ,fe

xtreg vfrall unrate lnperinc beertax mlda mmiles i.punish i.year, absorb(state)

\* model 4 not containing any economic variables

xtreg vfrall beertax mlda mmiles i.punish i.year, fe