# Final Project: US Traffic Fatalities

03/16/2020

# 1 Introduction

## 1.1 Background

Traffic fatalities are a major source of accident deaths at all ages, and almost half of drivers and more than 40% of passengers killed in vehicle crashes have been drinking [1]. The effects of drunk driving laws such as mandatory jail sentence on road fatalities is an important topic for policymakers in order to reduce the fatality rate. Our study is based on dataset "Fatalities" from the AER package in R which is a panel dataset reporting annual state level observations on U.S. traffic fatalities for the period 1982 through 1988.

There are two primary scientific questions we are interested in. The first question is whether there is an effect of mandatory jail sentence on fatality rate or not. The second question is whether there is a causal effect between jail sentence and fatality rate. In this study, we implement exploratory data analysis, fixed effect model, model diagnostics and propensity score. In the end, we draw our conclusion and present some suggestions to policymakers based on our assumptions and analysis.

## 1.2 Statistical questions of interest

To answer the primary scientific question of interests, we propose to fit a fixed effects model with fatality rate as responsible variable, economic conditions and drunk driving laws as predictor variable given state fixed effects and time fixed effects. We will then run our model diagnostic to check if the assumptions of the model hold and find whether having a mandatory jail sentence is associated with reduced traffic fatalities. Then, we use propensity score matching method to measure the causal effect between mandatory jail sentence and traffic fatality.

# 2 Analysis Plan

## 2.1 Population and study design

According to the description of the dataset, this observational study includes the fatality related information in 48 states for the period 1982 through 1988. We only focus on the vehicle fatality rate and variables related with economic condition and drunk driving laws.

## 2.2 Descriptive Analysis

First of all, we pick up all valuable variables we are interested in. To deal with missing value, we search literatures and fill in it. Then we do data pre-processing to better interpret variables. For some variables related with making policies, we draw the scatter plot and boxplots to show the relationship between each one and fatality rate respectively.

## 2.3 Fixed effects model

Fixed effects regression is a good method for controlling for omitted variables in panel data when the omitted variables vary across entities state but do not change over time. Compared with linear regression model, fixed effects regress can fit better when there are two or more time observations for entity. For the fixed effect regression model, each entity has an intercept, which represents by a set of binary variables. These binary variables absorb the influence of all omitted variables that differ from one entity to the next but are constant over time. Therefore, compared with linear regression model, fixed effected model can explain more for covariate variables.

In order to eliminate bias from unobservable that change over time but are constant over entities and control for factors that differ across entities but are constant over time, we include both individual fixed effects and time fixed effects in the model, i.e. we assume state and year as fixed effect. By combining them, the model is given as follows:

$$y_{it} = \alpha_i + \mu_t + X_{it}\beta + \varepsilon_{it}, \quad \text{for} \quad t = 1, 2, \cdots, 7 \quad \text{and} \quad i = 1, 2, \cdots, 48,$$

where $y_{it}$ is the variable observed for $i$-th state at year $t$, $X_{it}$ is the time-variant regressor vector including beetax, unemp, log(income), miles, drinkage, jail and service variables. $\beta$ is the matrix of parameters for each variable, $\alpha_i$ is the unobserved time-invariant individual effect, $\mu_t$ is the unobserved individual-invariant time effect, $\varepsilon_{it}$ is the error term.

Assumptions of T-test is $\varepsilon_{it} \overset{i.i.d}{\sim} N(0, \sigma^2)$.

Assumptions of fixed effected model specified as follows:

- $E(u_{it}|X_{i1}, X_{i2}, \cdots, X_{i7}, \alpha_i) = 0$.
- $(X_{i1}, X_{i2}, \cdots, X_{i7}, u_{i1}, u_{i2}, \cdots, u_{i7}), i = 1, 2, \cdots, 48$ are i.i.d drawn from their joint distribution.
- Large outliers are unlikely: $(X_{it}, u_{it})$ have non-zero finite fourth moments.
- There is no perfect multicollinearity.
- The error for a given state are uncorrelated over time, conditional on the regressors: specifically, $cov(u_{it}, u_{is}|X_{i1}, X_{i2}, \cdots, X_{i7}, \alpha_i) = 0$ for $t \neq s$.

## 2.4 Causal Inference

To test the causal effect of jail on traffic fatality rate, we choose to use propensity score matching method for observational study [2][3].

### 2.4.1 Assumption

There are several assumptions to strengthen our capacity for causal inference by utilize matching based on propensity scores.

The stable unit treatment value assumption (SUTVA), which requires that the outcome of one subject is unaffected by the treatment of another; the Unitary Treatment assumption, which requires that there is only one version of the treatment; the Positivity assumption, which suggests all subjects have some probability of receiving the treatment; and the inclusion of all significant covariates, which requires all significant confounding covariates are included.

### 2.4.2 Propensity Score

Propensity score is a statistical technique that has proven useful to evaluate treatment effects when using observational data. Because there are many other predictors that may affect the likelihood of being assigned into the treated group, we use a logit model for the propensity of observations to be assigned into the treated group, since jail is a binary variable.

The propensity score model is a probit/logit model with treatment $D$ as the dependent variable and other observations $X$ as independent variables. The propensity score is the conditional (predicted) probability of receiving treatment given pre-treatment characteristics X: $p(X) = P(D = 1|X) = E[D|X]$.

The propensity score then allows matching of individuals in the control and treatment conditions with the same likelihood of receiving treatment. Thus, a pair of participants (one in the treatment, one in the control group) sharing a similar propensity score are seen as equal, even though they may differ on the specific values of the covariates (Holmes, 2014).

### 2.4.3 Matching

Then we match observations from treated and control groups based on their propensity scores. There are several matching methods available, like nearest neighbor matching, optimal matching and full matching.

The near neighbor matching procedure matches participants from the control group to participants from the treatment group based on closeness. A participant ($j$) with propensity score $P_j$ in the control sample ($I_0$) is a match for a participant ($i$) with propensity score $P_i$ in the treatment group, if the absolute difference between their propensity scores is the smallest $C(P_i) = \min_j \parallel P_i - P_j \parallel, j \in I_0$. The most traditional matching is of one participant in the control to one participant in the treatment. In those cases, we speak about 1-to-1 (1:1) matching. However, it is possible to have more than one participant from the control group to be matched with a participant in the treatment group. In those cases, we speak about an m-to-1 (m:1) matching. Having more individuals from the control group matched to every individual in the treatment group means better estimates for the counterfactual in the control group. However, this approach requires a sample size for the control group several times larger than the number of individuals in the treatment group.

### 2.4.4 ATT

Because the treatment is unbalanced, there are more observations with jail is `no` than with jail is `yes`. Therefore, we prefer ATT(average treatment effect on the treated) than ATE(Average treatment effect). ATT is the difference between the outcomes of treated and the outcomes of the treated observations if they had not been treated. The definition is $ATT = E[Y_1|X, D = 1] - E[Y_0|X, D = 1]$. The second term is a counterfactual so it is not observable and needs to be estimated.

After matching on propensity scores, we can compare the outcomes of treated and control observations to get the empirical estimation of ATT. $\hat{ATT} = \frac{1}{n_1} \sum_{i \in \{D=1\}} [y_{1,i} - \sum_j w(i, j) y_{0,j}]$. Each treated observation $i$ is matched $j$ control observations and their outcomes $y_0$ are weighed by $w$.

### 2.4.5 Sensitivity

A question about propensity score matching is: how sensitive are these results to hidden bias? Rosenbaum (2002, 2005) recommends that researchers try to answer this question by conducting a sensitivity analysis. The idea is to determine how susceptible the results presented might be to the presence of biases not identified by the researcher or removed by the matching. Rosenbaum (2002) developed methods to determine bias through several non-parametric tests such as McNemar's and Wilcoxon's signed rank test. Keele (2015) developed the package `rbounds` which estimates the sensitivity of the results to hidden bias. `rbounds` can compute sensitivity analysis straight from the package matching (Sekhon, 2011).

# 3 Results

## 3.1 Descriptive Analysis

The dataset we investigate has 336 rows and 34 columns. This dataset is a panel data set reporting annual state level observations on U.S. traffic fatalities, including the information of 48 states for the period 1982 through 1988. In order to investigate whether having a mandatory jail sentence is associated with reduced traffic fatalities. Besides fatal, state, year and jail variables, we should also choose variables which may be served as confounder into account. From literature [5], we consider unemp, income, miles, drinkage, beertax, service into account. Moreover, we pre-process the data as follows.

- The original data is not balanced. The value of jail and service variables in row 28 are missing. By literature[6], mandatory jail sentence and mandatory community service policy were not launched. Both of two missing values are "no".
- Because of the different population size in 48 states, investigating the traffic fatality rate is more reasonable than the number of vehicle fatalities. It is measured as the number of fatalities per 10000 inhabitants.
- Compared with the value of other variables, the magnitude of the value of variable income are pretty large. For the sake of readability of coefficients of income, log-transformation is taken.
- By observing the values of drinkage, some of them are not integers but with two decimal places, which does not make sense in practice. We present a discretized version of drinkage that classifies states into four categories of minimal drinking age: $[18, 19), [19, 20), [20, 21)$ and $[21, 22]$.

Figure 1 shows the fatality rate across the country. It can be observed that the lower rate in Mideast and Plains areas and the higher rate in Far West and Rocky Mountain areas. It may result from different laws and and lifestyles including alcohol consumption due to regional differences. Figure 2 shows that the mandatory jail sentence is overall consistent over 7 years, and it concentrate mainly on the West, which indicates the west has stricter laws.

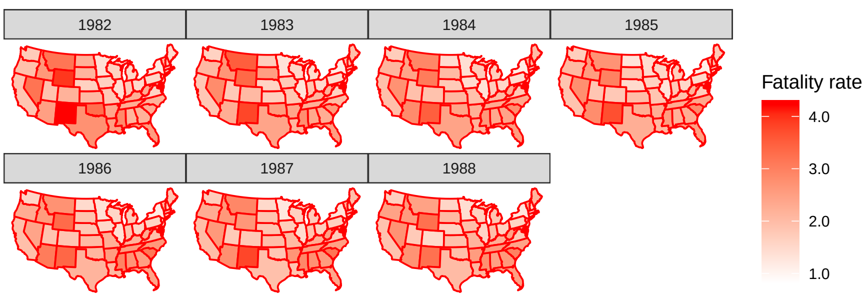


**Figure.1 Traffic fatality rate over 7 years**

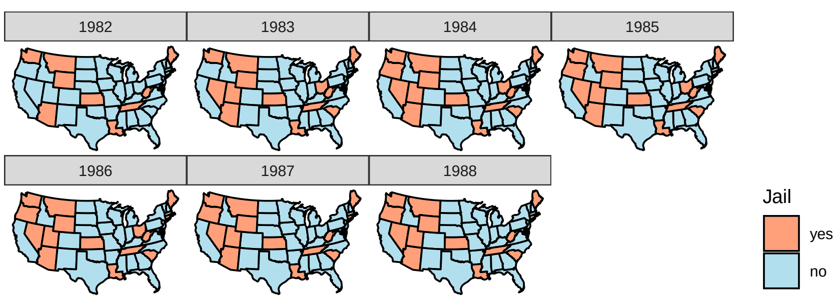


**Figure.2 Mandatory Jail Sentence over 7 Years**

After data pre-processing, we get nine variables: fatality rate, state, year, beertax, unemp, log(income), miles, drinkage, jail and service, where fatality rate represents the number of fatalities per 10000 inhabitants. Beertax represents tax on case of beer. Unemp represents unemployment rate. Log(income) represents the logarithm of real per capita income. Miles represents average miles per driver. Drinkage represents age intervals of minimal drinking age. Jail represents whether there exists mandatory jail sentence in one particular state.

Because vehicle use depends in part on whether drivers are wealthy enough to earn cars, omitting state economic conditions could result in omitted variable bias. From the Figure 3, we can see that ranging from 2.5% to 10%, the employment rate has a positive relationship with the fatality rate. In contrast, log(income) is negatively associated with the fatality rate. Therefore, we decide to choose two variables unemp and log(income) to reflect the influence of economic conditions in our model.


**Figure.3 Economy condition vs. Fatality rate scatter plot**

Besides economic conditions, we are interested in driving and alcohol policy in order to make some suggestions to policymakers. We investigate the relationship between fatality rate and beertax, jail, drinkage respectively. From Figure 4, the result is contrary to our expectations, alcohol taxes, mandatory jail sentence, mandatory community service and higher minimal drinking age are supposed to lower the rate of traffic fatalities. This is possibly due to omitted variable bias, for example, economic conditions. We can't make a conclusion just from these figures. It indicates that we need to consider other covariates into study.

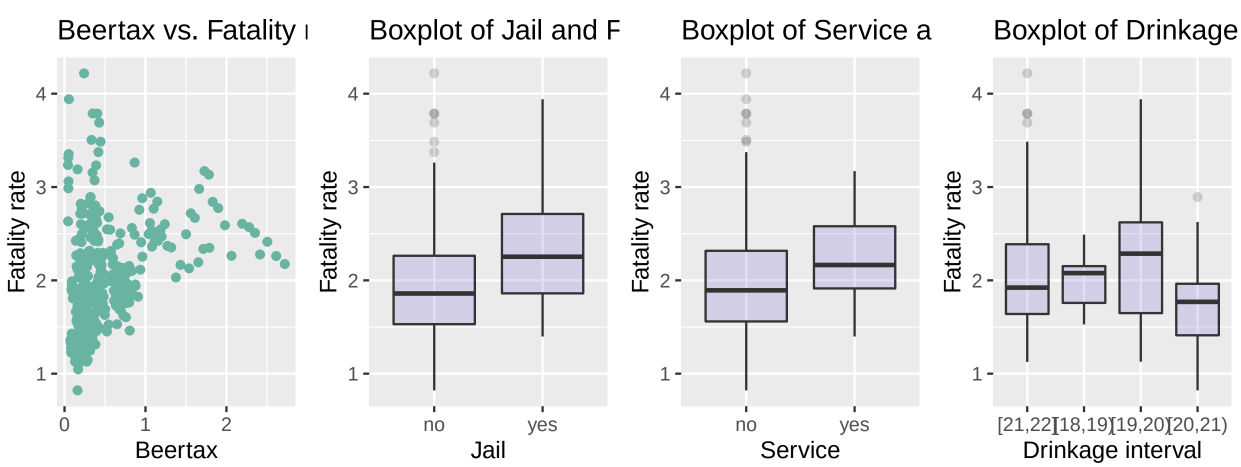


**Figure.4 The Relationships between Fatality Rate and Other Variables**

## 3.2 Fixed effects model

After fitting fixed effect panel model with R, we can get the model is,

$$\hat{fatality\ rate} = StateEffect + TimeEffect - 0.4607 * Beertax - 0.0625 * Unemp + 1.7879 * Log(income)$$

The p-value of F-test is less than 2.22 x $10^{-16}$, which means there is a significant linear association between fatality rate and predictor variables. Since the number of observations for every state and years periods are not the same, we group observations by jail sentence variable. The result shows below by Table 1:

| | Estimate Effect | Standard Error | t-value | p-value |
|---|---|---|---|---|
| Jailyes | 0.0136 | 0.1230 | 0.1133 | 0.9099 |
| Beertax | -0.4607 | -0.1673 | -2.7530 | 0.0063 |
| Unemp | -0.0625 | 0.011 | -5.5960 | 5.318 x $10^{-8}$ |
| Log(income) | 1.7879 | 0.3641 | 4.9111 | 1.559 x $10^{-6}$ |

**Table.1 Summary of fixed effects model**

From table 1, it shows each estimate effect on traffic fatality rate. For each variables, we have the null hypothesis $H_0 : \beta_i = 0 \ \ for \ \ i = 1, 2, \cdots, 8$ and the alternative hypothesis $H_a : \beta_i \neq 0 \ \ for \ \ i = 1, 2, \cdots, 8$. By t-test for every $\beta$, we can get that p values of beertax, unemployment rate and income are less than 0.05. It means that beer tax, unemployment rate and income are statistically significant. The estimate effect of unemployment rate is -0.06 and beer tax effect is -0.46, which means that the higher the unemployment rate or beer tax, the lower the traffic fatality rate. The estimate effect of log(income) is 1.80, which is proportion to traffic fatality rate. The main goal for the project is to test whether there is statistically significant for mandatory jail sentence variable or not. If there is a policy of mandatory jail sentence, it seems that the traffic fatality rate is relatively higher since the estimate is 0.01. However, the p value of this variable is greater than 0.05. Therefore, to find the causal effect of mandatory jail sentence, we need to use other method in the following part.

## 3.3 Model Diagnostics

### 3.3.1 Zero-mean and equal variance (Homoscedasticity)

From the Residuals vs Fitted Values scatterplot (Figure 5), these points are uniformly distributed on both sides of x-axis, which means that our model satisfies the zero mean assumption and does not violate the equal variance assumption.

### 3.3.2 Independence

In the United States, states generally differ from each other. Every state make their own law, even though some of their situations are similar. Besides, we have chosen unobserved time-invariant individual effect as one of the variables to control the influence of different states. In addition, we think that the previous outcomes will not affect the future. Therefore, in our model, $(X_{i1}, X_{i2}, \cdots, X_{i7}, u_{i1}, u_{i2}, \cdots, u_{i7}), i = 1, 2, \cdots, 48$ are independently and identically distributed.

### 3.3.3 Normality

Through Q-Q Plot (Figure 5), we can see that the residuals are slightly heavy tailed compared with normal distribution. Since the p values of our variables are extremely small, slightly heavy tailed probability distribution has little influence to our result. Moreover, we do not know how to solve this situation that our model violates normality assumption.

#### 3.3.4 Influential observations

To derive the leverage of the observations, a half-normal plot (Figure 5) is drawn to sort the observations by their leverage. And we can find that there is only one influential observations in the dataset. In order to maintain the balance of our dataset, we decide to reserve this point.

#### 3.3.5 Multicollinearity

A popular way of diagnosing multicollinearity is through the calculation of variance inflation factors (VIFs). The VIF score indicates the proportion by which the variance of an estimator increases due to the inclusion of a particular covariate. From Table.2, the VIFs of the variables in the model ranges from 1.0 to 1.6, which means that multicollinearity does not exist in our model. Therefore, the model coefficients are stable and unbiased.

**Table.2 Variance Inflation Factors of Variables**

| beertax | drinkage | jailyes | unemp | miles | log(income) | serviceyes |
|---|---|---|---|---|---|---|
| 1.0505 | 1.0350 | 4.1313 | 1.5182 | 1.0148 | 1.5483 | 4.1642 |

#### 3.3.6 Uncorrelation

There is only one error for a given state each year, so the correlation value for a given state over time does not exist. Therefore, we can just assume that the model meets the uncorrelation assumption.

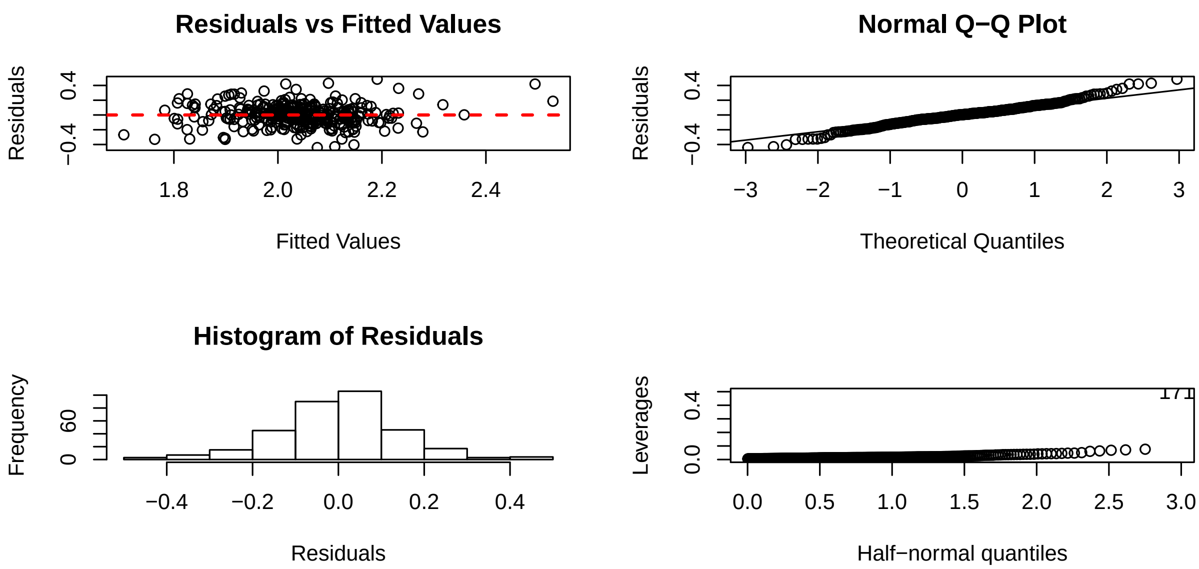


**Figure.5 Plots of Model Diagnostics**

## 3.4 Causal Inference

In this case, we choose beer tax, unemployment rate, logistic income, miles and drinkage as instrument variables when studying the casual effect between jail sentence and fatality rate, since we are interested in the influence from economy condition as well as other drunk driving law.

We set jail sentence `yes` as treatment group, and jail sentence `no` as control group. Note it is unbalanced for jail sentence: 94 cases in treatment group but 242 in control group.

From Table 3, the mean value of each instrument variable seems different from each other in treatment group and control group.

**Table.3 Summary for all data**

| Mean | n | beertax | unemp | log(income) | miles | drinkage | frate |
|---|---|---|---|---|---|---|---|
| Treated | 94 | 0.4846 | 7.9383 | 9.4855 | 8057.95 | 20.2979 | 2.2946 |
| Control | 242 | 0.5244 | 7.1169 | 9.5411 | 7825.81 | 20.5169 | 1.9417 |

By Chi-square test using XBalance in R, the result indicates that at least one of the variables included in the model is creating an imbalance between the two groups.

### 3.4.1 Assumption

A quick inspection of the distributions of propensity scores between treatment and control in Figure 6 shows that the Positivity assumption appears to be satisfied. Unitary Treatment assumption is satisfied, only if each state has that has a mandatory jail sentence has a similar length for each jail sentence. However, it is unclear from the current data if this is the case. There are many potentially important covariates we do not have data on, causing problems for the assumption requiring inclusion of all significant covariates. Furthermore, some of these potentially relevant covariates have some influence on whether the STUVA is satisfied. For example, if the sentence lengths for mandatory jail time differ drastically between states that have these laws, this could affect not only the Unitary Treatment assumption, but the STUVA as well, for this might influence the both the driving and drinking habits of people near state lines. The existence of dry counties in the U.S. adds further complexity to this problem.

### 3.4.2 Propensity Score

We build on a logistic regression on treatment `jail` with other instrument variables. The fitted value is the propensity score in each case.

Once the propensity scores have been calculated, a graphical approach can be used to assess the distributional similarity between score distributions. This graphical approach uses back to back histograms such as those created through the package `Hmisc` (Harrell, 2015) in R. Figure 6 presents the histograms `Hmisc` generates. The result shows there is some difference in the shape between two distributions.

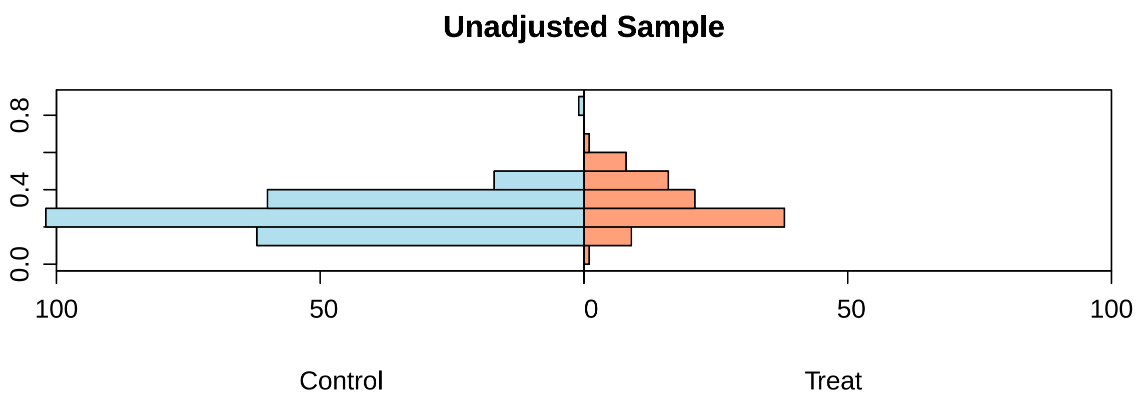


**Figure.6 Back to back histograms of distribution on propensity score by whole data**

### 3.4.3 Matching

Packages such as `MatchIt` provide summary tables that include means and standard deviations for the two groups both before and after the matching was completed. It also includes percent improvement, and finally, it provides a summary of the number of individuals included in the final sample, and cases that were not matched. We use the near neighbor matching method, which means each case in treatment group will be matched with a case in control group which has the closest propensity score.

From Table 4, the mean value of each instrument variable are closer to each other in different group. The chi-square test indicates no significance, thus suggesting equivalence between the groups.

**Table.4 Summary for matched data**

| Mean | n | beertax | unemp | log(income) | miles | drinkage | frate |
|---|---|---|---|---|---|---|---|
| Treated | 94 | 0.4846 | 7.9383 | 9.4855 | 8057.95 | 20.2979 | 2.2946 |
| Control | 94 | 0.4789 | 7.8426 | 9.4876 | 7732.10 | 20.3697 | 2.0384 |

As can be observed in figure 7, there is a remarkable improvement in the match between the two distributions of propensity scores after the match compared to Figure 6. This match suggests that the two groups are much more similar in terms of their propensity scores, and thus, the selection bias has been reduced substantially.

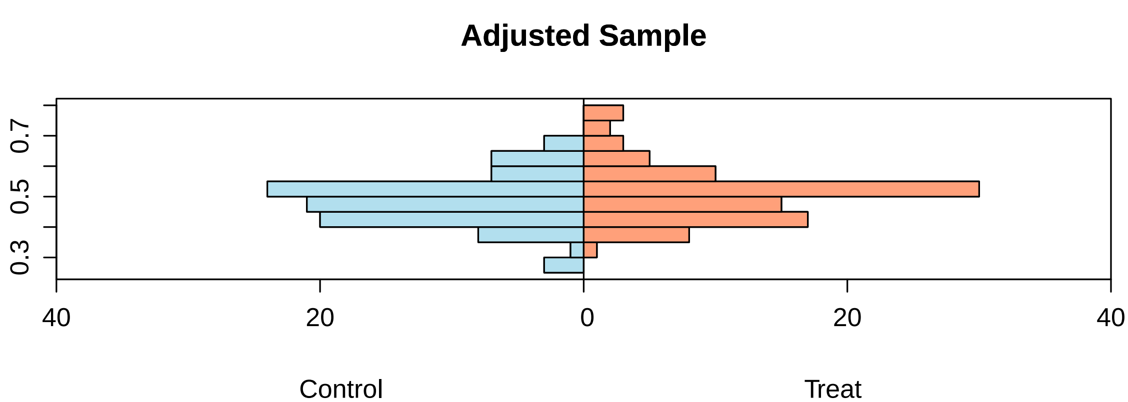

**Figure.7 Back to back histograms of distribution on propensity by matched data**

### 3.4.4 ATT

The paired T test shows that there is a significant effect of jail sentence on fatality, with the p value as 0.0014 and the 95 percent confidence interval of difference between two groups as (0.1020, 0.4103).

The estimation of ATT (average treatment effect on the treated) is 0.2561, which means the fatality rate is 0.2561 higher when jail sentence is `yes` than when jail sentence is `no`. Therefore, we draw a conclusion that there is a casual effect of jail sentence on fatality rate, and the mandatory jail sentence law causes a higher fatality rate.

### 3.4.5 Sentivity

**Table.5 Sensitivity analysis using Wilcoxon's rank sign test**

| Gamma | Lower Bound | Upper Bound |
|---|---|---|
| 1.0 | 0.0003 | 0.0003 |
| 1.1 | 0.0001 | 0.0012 |
| 1.2 | 0.0000 | 0.0036 |
| 1.3 | 0.0000 | 0.0091 |
| 1.4 | 0.0000 | 0.0194 |
| 1.5 | 0.0000 | 0.0364 |

The value of Gamma is interpreted as the odds of treatment assignment hidden bias. A change in the odds lower/upper bounds from significant to non-significant (or vice-versa) indicates by how much the odds need to change before the statistical significance of the outcome shifts. For example, in Table 5, the upper bound estimate changes from significant (0.0091) to non-significant (0.0194) when gamma is 1.4. That is, a change of 0.4 in the odds will produce a change in the non-significance value. Rosenbaum (2002) defines a study as sensitive if values of Gamma close to 1 lead to changes in significance compared to those that could be obtained if the study is free of bias. Thus results will be more robust to hidden bias, if a very large change in the odds is needed before a change in statistical significance happens.

# 4 Conclusion

By fixed effect model, neither stiff punishments nor increases in the minimal legal drinking age have important effects on fatalities. In contrast, there is some evidence that increasing alcohol taxes, as measured by the real tax on beer, does reduce traffic deaths. Good economic conditions are associated with higher fatalities, perhaps because of increased traffic density when the unemployment rate is low or greater alcohol consumption when income is high[5]. Miles also has a significant positive estimate coefficient, and there is no doubt that the more people driving, the more possibility they would run into a car accident.

For casual inference, there is significant evidence to show a mandatory jail sentence has causal effect on fatality rate. What is worse, the mandatory jail is slightly related to a higher fatality rate. It may be interpreted by a negative mentality on this policy, or a low cost to violate this regulation.

Here are suggestions for policymakers:

(a) Raising beer tax properly can effectively reduce the vehicle fatality rate. In detail, the effect of a \$2.17 increase (in 1988 dollars) in the beer tax is a decrease in the expected fatality rate by 1 death per 10,000.

(b) There is significant casual effect of jail on a higher vehicle fatality rate. The policymaker should introspect on the rationalization of law in drinking and driving.

# 5 Discussion

The distribution of residuals is a little heavy-tailed and violates the normality assumption. However, if we use any transformation on response variable, the model will be less interpretable especially for observational study. Thus, the t test is less reliable in this case. We need further method like non-parameter test to avoid normality.

The p-value in linear regression model only means association, so we can't make a causal statement for quantitative variable such as beer tax. We need further model to study the causal effect of beer tax on fatality rate.

There are other matching methods like the optimal matching and full matching. In further research, we can compare with different matching methods and choose the best one.

# 6 Appendix. Reference

[1] Zobeck, Terry S., Frederick S. Stinson, Bridget F. Grant, and Darryl Bertolucci. 1993. Trends in Alcohol-Related Fatal Traffic Crashes, United States: 1979-91, Washington, D.C.: National Institue of Alcohol Abuse and Alcoholism, Surveeillance Report #26, November.

[2] Ani Katchova. 2013. Propensity Score Matching. https://sites.google.com/site/econometricsacademy/econometrics-models/propensity-score-matching

[3] Antonio Olmos; Priyalatha Govindasamy. Propensity Scores: A Practical Introduction Using R. Journal of MultiDisciplinary Evaluation Volume 11, Issue 25, 2015.

[4] Jasjeet Singh Sekhon. Package 'Matching'. https://cran.r-project.org/web/packages/Matching/Matching.pdf

[5] James H. Stock; Mark W. Watson. 2007. Introduction to Econometrics, 2nd Edition. Pearson.

[6] NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION. Digest of State Alcohol Highway Safety-Related Legislation. https://nhtsa.dr.del1.nhtsa.gov/Driving-Safety/Impaired-Driving/Digest-of-State-Alcohol-Highway-Safety%E2%80%93Related-Legislation

# 7 Appendix II. Group Partners

This Document is the project 3 of Team 7 in STA 207, Winter 2020.

1. Bingdao Chen bdchen@ucdavis.edu contribution: descriptive analysis and model establishment
2. Yahui Li yhuli@ucdavis.edu contribution: casual inference and conclusion
3. Zihan Wang zihwang@ucdavis.edu contribution: fixed effects model
4. Jian Shi jnshi@ucdavis.edu contribution: model diagnose

The repository in Github is on https://github.com/yhli097/STA207_Final_Project.git