Whether having a mandatory jail sentence is associated with reduced traffic fatalities

Team ID: 3

Name (writing): Kieran Zhou

Name (writing): Liela Meng

Name (coding): Yunan Hou

Name (coding): Zhen Li

Github repo link:https://github.com/maoqunqun/STA207_project3.git

1.Introduction

In 2018, there were 36,560 people killed in traffic crashes in the United States (NHTSA, 2019). In order to reduce the traffic fatalities, many studies have made suggestions on alcohol-control policies to federal and state governments. Though alcohol involvement is the most concerned factor, other potential factors may impact the fatality rate in the United States.

This study investigates the impact of both alcohols involved factors and non-alcohol involved factors on traffic fatality rate based on the traffic fatalities panel data for 48 states from 1982 to 1988. In this paper, the specific causes and the age of death under vehicle fatalities are not concerned. This paper focuses on the rate of vehicle fatalities and aims to find the factors which cause the reduction or increase of the vehicle fatalities rate. In particular, whether punishment (mandatory jail sentence or community service) is useful is also concerned in this paper.

This study provides evidence that higher spirits consumption is associated with increases in vehicle fatalities rate, and higher beer taxes are associated with reductions in vehicle fatalities rate. Interestingly, per capita, personal income, and employment rate are positively associated with vehicle fatalities rate. This indicates that higher personal income and employment rates may also be associated with increases in alcohol consumption.

2.Data analysis

2.1 Analysis plan

2.1.1 Exploratory data analysis

The data this paper uses contains 16 distinct variables that may affect the vehicle fatalities rate. In this data, whether there was mandatory community service in California in 1988 is unclear. After an investigation, the answer “no” is filled in, and hence the omitted value bias is avoided.

According to the interest of this paper, percent of southern baptists, percent of Mormon, miles, and GSP rate of change are not included in the analysis as they can not be controlled by policies. For the rest of the variables, previous research has shown the significant influence of unemployment, income, spirits consumption, and beer tax.

Though the governments made efforts setting several policies to reduce alcohol-involved driving (Ruhm, 1996), there is no pattern on the changes of total vehicle fatalities rate. Based on this situation, an investigation on alcohol-involved vehicle fatalities rate is also made in this paper, in order to analyze whether the punishment (mandatory jail sentence or community service) and other related policies are effective.

Figure 1 and figure 2 shows confidence intervals of estimated means grouped by state and year, respectively. Those figures indicate that the vehicle fatalities rates vary a lot in different states, and vehicle fatalities rates show differences among distinct years as well. Thus, the state and year fixed effects terms are considered in the model.

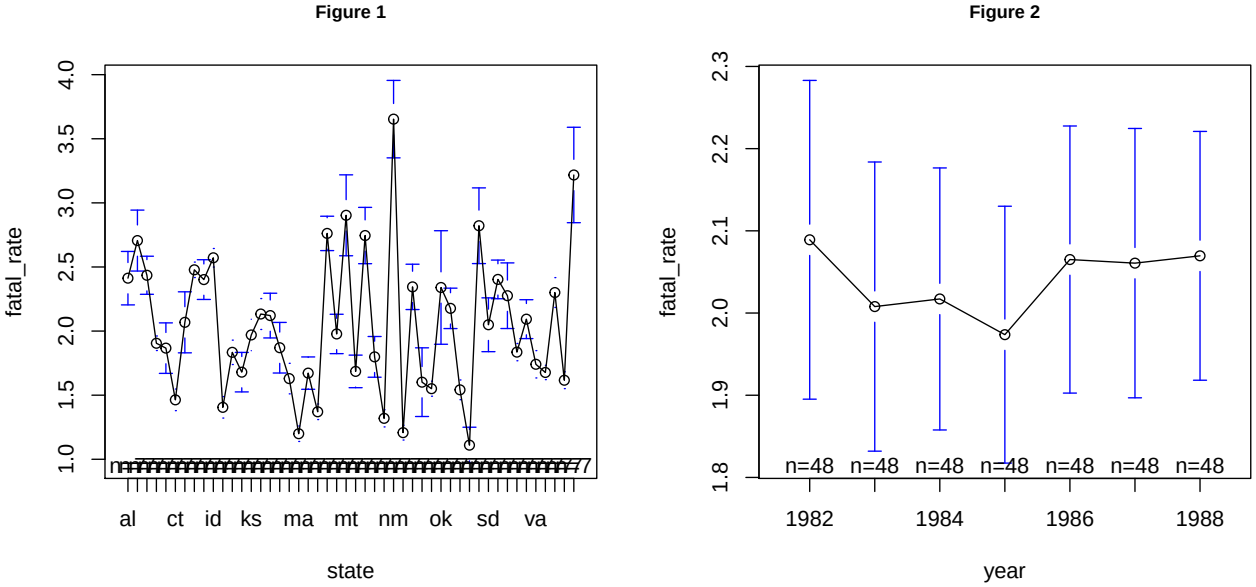

Figure 1: Heterogeniety across state - Mean of seven years' fatal rate for each state. Figure 2: Heterogeniety across year - Mean of each states' fatal rate for each year.

2.1.2 Model

To analyze these fatalities panel data, both random and fixed effects models are required to check. Ruhm (1996) suggests using the fixed-effects model to analyze this data. Besides, by running the Hausman Test, the result suggests fixed-effects model is more suitable for given dataset.

$Y_{i,t} = \mu_0 + \alpha_i + \beta_t + X_{1,i,t} + X_{2,i,t} + X_{3,i,t} + X_{4,i,t} + \beta_{jail,i,t} + \varepsilon_{i,t}$, $i = 1, 2, \ldots, 48$, $t = 1, 2, \ldots, 7$, where i denotes state and t denotes year. μ_0 is total average of fatal rate for all 48 states in 7 years. α_i is state fixed term and β_t is year fixed term. $X_{1,i,t}$, $X_{2,i,t}$, $X_{3,i,t}$, $X_{4,i,t}$ represent unemployment, income, spirits consumption, and beer tax. β_{jail} is a factor variable represents whether mandatory jail for drunk driving exists or not.

In the fixed effects model, there are four assumptions (Hanck et al., 2019) as follows.

- The error term u_{it} has conditional mean zero, that is, $E(u_{it}|X_{i1}, X_{i2}, \ldots, X_{it}) = 0$
- $(X_{i1}, X_{i2}, \ldots, X_{i3}, u_{i1}, u_{i2}, \ldots, u_{iT}), i = 1, 2, \ldots, n$ are i.i.d. draws from their joint distribution.
- Large outliers are unlikely, i.e., X_{it}, u_{it} have nonzero finite fourth moments.
- There is no perfect multicollinearity.

2.1.3 Hypothesis and hypothesis testing

- Mandatory jail sentence

As the mandatory jail sentence is the most concerned factor in this study, a null hypothesis that mandatory jail sentence does not reduce the fatalities rate is proposed as follows.

$H_0 : Y_{i,t} = \mu_0 + \alpha_i + \beta_t + X_{1,i,t} + X_{2,i,t} + X_{3,i,t} + X_{4,i,t} + \varepsilon_{i,t}$ vs. $H_a : Y_{i,t} = \mu_0 + \alpha_i + \beta_t + X_{1,i,t} + X_{2,i,t} + X_{3,i,t} + X_{4,i,t} + \beta_{jail,i,t} + \varepsilon_{i,t}$.

The null hypothesis cannot be rejected (p-value = 0.437) at significant level equals to 0.05 with F-statistic equals to 0.605, which means there is no significant evidence that mandatory jail sentence can reduce the vehicle fatalities rate.

- Time fixed effects

$H_0 : \beta_t = 0$ vs. $H_a : \beta_t \neq 0$, $t = 1, 2, \ldots, 7$.

Time fixed effects control the variables that are constant across state but vary over time (Hanck, 2019). The Lagrange Multiplier test is used to check whether the the model contains time fixed effects. The null hypothesis that time fixed effects term equals to zero is rejected with p-value equals to 0.005. Thus, time fixed effects should be kept in the model.

- State fixed effects

$H_0 : \alpha_i = 0$ vs. $H_a : \alpha_i \neq 0$, $i = 1, 2, \ldots, 48$.

Similarly, the Lagrange Multiplier test is also applied to check whether state fixed effects term should be kept. The result also indicates that there is significant evidence (p-value < 0.001) that state fixed effects term cannot be dropped.

2.2 Analysis result

2.2.1 Model fitting result

After drop Jail variable by F test, the final model is shown below:

$Y_{i,t} = \mu_0 + \alpha_i + \beta_t + X_{1,i,t} + X_{2,i,t} + X_{3,i,t} + X_{4,i,t} + \varepsilon_{i,t}$, $i = 1, 2, \ldots, 48$, $t = 1, 2, \ldots, 7$.

Table 1 (Summary of tests for Time and State fixed effects)

	Chi-square	degree of freedom	P-value
Time Fixed Effects	4.43	1	0.03523
State Fixed Effects	330.34	1	2.2e-16

Table 2 (Summary of fixed-effect model)

	Estimate	Standard Error	T-value	P-value
Unemployment Rate	0.055	0.010	-5.3184	<0.001
Income	0.00008	0.00002	4.3111	<0.001
Spirits	0.806	0.113	7.1501	<0.001
Beer Tax	-0.448	0.155	-2.8916	0.004
Jail	0.044	0.056	0.7781	0.4372

2.2.2 Assumption satisfaction

- Independence

In order to apply F test, the assumptions of error term are also required to be checked.

Due to the characteristics of panel data, observations are not required to be independent in years (Hanck, 2019). In this fatalities panel data, the observations within each state are independent of those within other states, hence the independence assumption is justified.

- Constant variance

Figure 4 provides the residual scatter plot which shows that residuals of the model uniformly gather around zero. The result indicates that the constant variance assumption is not violated.

- Normality

Figure 5 and 6 shows the histogram and Q-Q plot of residuals of the model. The histogram shows an approximately bell shape of the distribution, and the Q-Q plot also shows no obvious heavy/light tail or skewness. Thus error term of the model follows normal distribution.

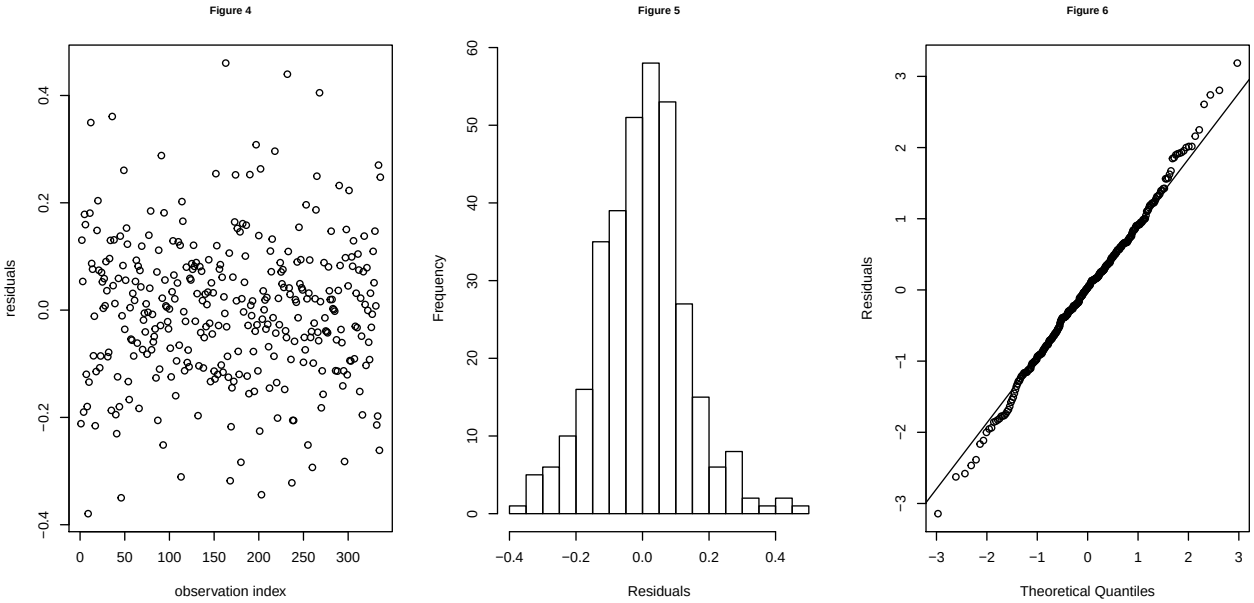

Figure 4: Residuals vs observation. Firure 5: Histogram of residuals of the model. Figure 6: QQ plot of residuals of the model.

- Summary

The model is diagnosed to ensure the assumptions are not violated. The first assumption is justified as the correlations between the error term, and concerned variables are very small, which indicates that they are uncorrelated. As the observations in these 48 states are independent and collected from the same distribution, the second assumption is justified as well. Figure 3 shows that there is no extreme outlier in this dataset; hence the third assumption is not violated. Also, as none of VIFs is large, it indicates that there is no perfect multicollinearity in the model. Hence the fourth assumption is justified.

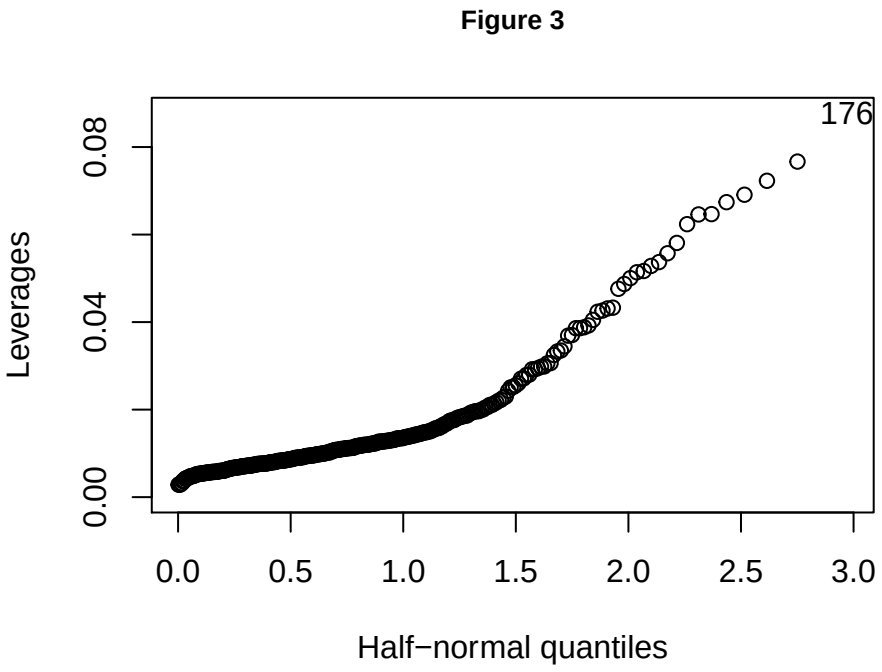

Figure 3: Leverage of the observations

3.Discussion

3.1 Conclusions and problems for further investigation

- This study shows more evidence that alcohol-involved driving is possibly one of the most severe causes of vehicle fatalities. The higher spirits consumption is positively associated with total vehicle fatalities. In particular, spirits consumption is associated with alcohol-involved fatalities. In order to reduce traffic fatalities, this study suggests higher beer taxes. In addition, the mandatory jail sentence and mandatory community service are not necessarily required to reduce the alcohol-involved traffic fatalities as there is no significant evidence shown that the alcohol involved vehicle fatalities rate can be reduced by either panel policies.
- New evidence on non-alcohol involved effects is also provided. Though the result indicates higher personal income and a higher employment rate are associated with higher vehicle fatalities rate, it does not mean that governments should consider to lower citizens' personal income. There may be some confounders that confuse the association between income and vehicle fatalities rate. For example, higher personal income and employment rates may also be related to the increasing number of vehicles. This investigation only shows evidence that these non-alcohol involved factors are associated with vehicle fatalities rate as well, however, for further causal inference on these non-alcohol involved factors, more data and investigations are required to improve the related policies.

3.2 Causal inference

- Ignorability assumption is violated

For observational study, overt bias may occur since variables are not fully controlled. According to reference, if unobserved variables are confounders, then it leads to hidden bias, hence ignorability assumption is violated.

In this report, the adjusted r-square of the final fixed-effect model is 0.3, which reflects the explanatory power of the given regression model is relatively weak. Therefore, it is reasonable to conclude that there exist other variables, which can be confounders that affect the traffic fatalities rate as well, and are not included in the model as explanatory variables. Thus, the ignorability of treatment assignment assumption is violated.

- SUTVA assumption holds

This paper assumes that for each model, there is treatment effect on individual observation and the effect does not spill over onto other units. Considering that the effect of personal income, spirits consumption, beer tax, and unemployment rate on each state does not impact any other state in this dataset, the stable unit treatment value assumption holds.

- Multicollinearity

The presence of perfect multicollinearity makes it impossible to draw causal inference between any variable and the observable variable (Nichols, 2007). Due to a relatively high correlation (0.55) between personal income and the unemployment rate, there may be other variables which impact both personal income and vehicle fatalities rate. Thus, though there is strong evidence that personal income is positively associated with vehicle fatalities rate, it is still not safe to conclude that high personal income can cause high traffic fatalities rate. There may be other reasons which can explain this association as well. For example, educational background can also be involved in the association, which means people who have high personal income but with low education background may ignore the dangers of alcohol involved driving.

4.Reference

National Center for Statistics and Analysis. (2019). 2018 fatal motor vehicle crashes: Overview, (Traffic Safety Facts Research Note. Report No. DOT HS 812 826).

C. J., Ruhm. (1996), Alcohol policies and highway vehicle fatalities, Journal of Health Economics 15, 435-454.

C., Hanck et al. (2019), The fixed effects regression assumptions and standard errors for fixed effects regression, Introduction to econometrics with R, 227-228.

A., Nichols,(2007), Causal inference with observational data,Stata Journal 7(4), 507–541.