UC DAVIS
DEPARTMENT OF STATISTICS

STA 207

Final Presentation: US Traffic Fatalities

Team 7:
Yahui Li
Bingdao Chen
Zihan Wang
Jian Shi

3/9/2020

Contents

Contents

- Introduction

Background

Traffic fatalities are a major source of accidental deaths in the United States. In fatal accidents, nearly half of driver deaths and 40% of passenger deaths are the results of drunk driving.

(Corrected by Connor)

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.

Question of Interest

- What are the effects of alcohol taxes and drunk driving laws on traffic fatalities?
- Whether there is a causal effect of jail sentence on fatality rate.

Data

Source: “Fatalities” in R package “AER”.

Description: US traffic fatalities panel data for the 48 US states, annually from 1982 to 1988.

Basic Information: 336 observations and 34 variables.

Exploratory Analysis

Data pre-processing:

- Missing value: Two missing values in jail and service respectively, we made them up by reference [3]. Now it is balanced.
- Response variable: Fatality rate: the number of fatalities per 10000 inhabitants. frate <- fatal/pop*10000
- Transformation: income -> log(income).
- Classification: Classify drinkage into four categories: [18, 19), [19, 20), [20, 21) and [21, 22].

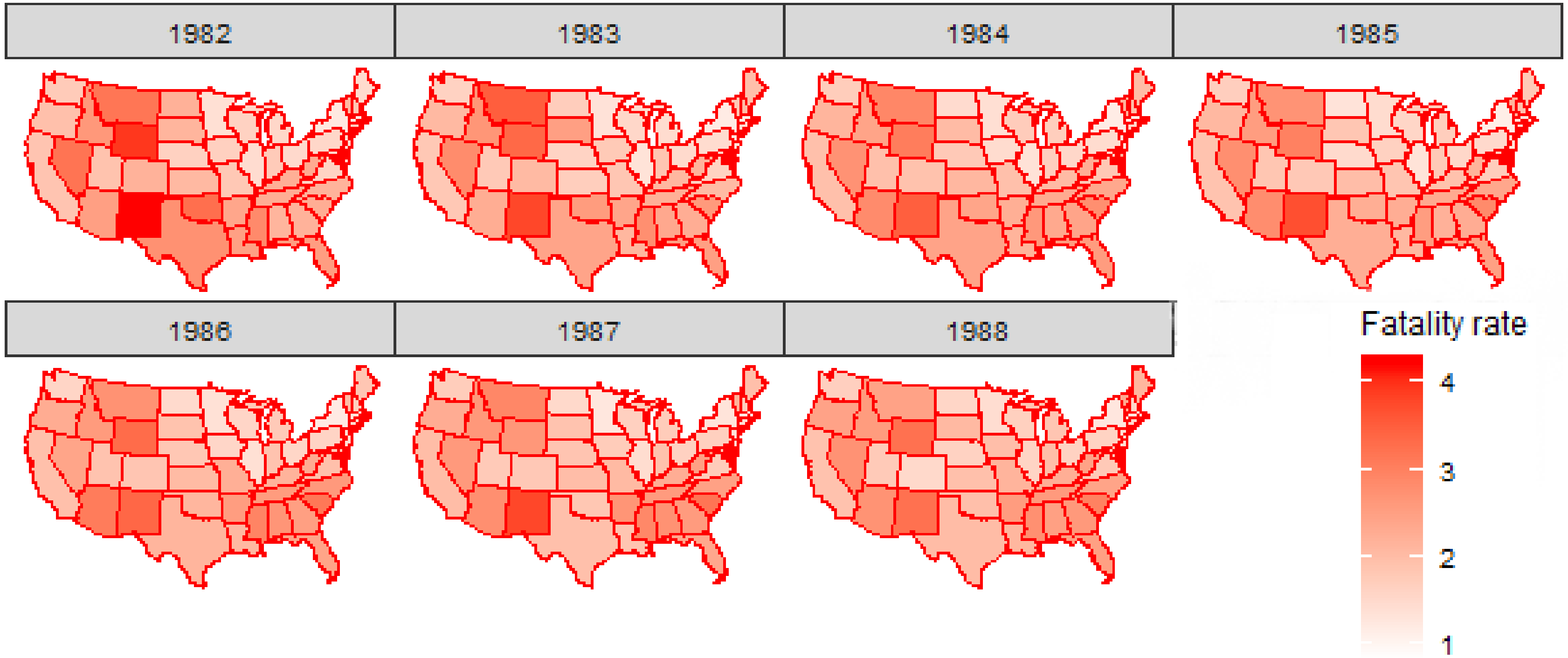
Chosen variables:

fatality rate, state, year, beertax, drinkage, jail, service,

unemp, log(income), miles

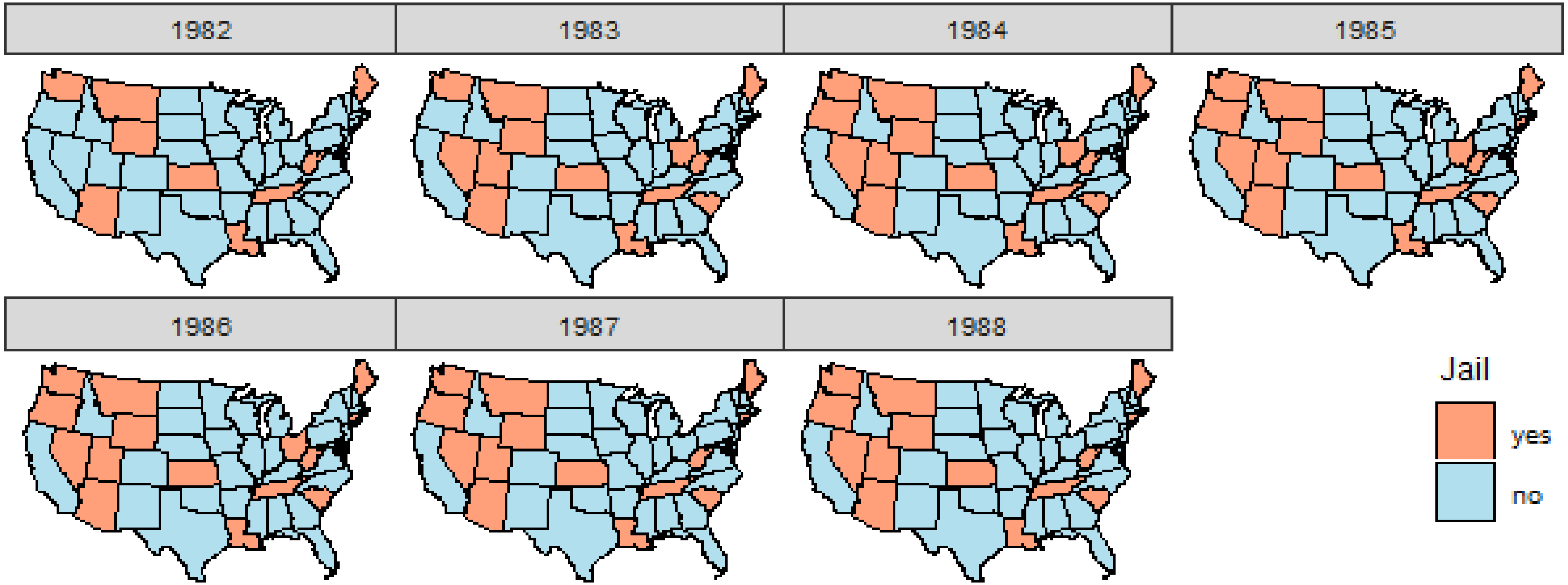
omitting economic conditions could result in omitted variable bias. (Suggestion from Siyao Wang)

Exploratory Analysis



In average, New Mexico, Wyoming, Montana have the fatality rate of top 3.
Rhode Island, Massachusetts, New York have the fatality rate of bottom 3.
(Suggestion from Yunan Hou)

Exploratory Analysis



There were more states which do not have a jail sentence.

Only several states changed the law through these years. (Suggestion from Yunan Hou)

Exploratory Analysis

Contents

Fixed effect model

Fixed effects regression is a method for controlling for omitted variables in panel data.

With entity fixed effects, this panel data set lets us control for unobserved variables that differ from one state to the next, such as prevailing cultural attitudes toward drinking and driving, but do not change over time.

With time fixed effects, it also allow us to control for variables that very through time, like improvements in the safety of new cars, but do not vary across state. [1]

(Suggestion from Connor)

Model formula

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

y_{it} represents the variable observed for ith state at year t

X_{it} represents the time variant regressor vector including beetax, unemp, log(income), miles, drinkage, jail and service variables.

β represents the matrix of parameters for each variable.

α_i represents the unobserved time invariant individual effect.

μ_t represents the unobserved individual invariant time effect.

ε_{it} represents the error term.

Fitted Model

```
fem_full <- plm(frate ~ beertax + unemp + log(income) + miles + drinkagec + jail +service,
          index = c("state", "year"), model = "within",  effect = "twoways",  data = Fatalities)
summary(fem_full)
```

```
Coefficients:
                    Estimate  Std. Error t-value  Pr(>|t|)
beertax          -4.5027e-01  1.6929e-01 -2.6598  0.008281 **
unemp            -6.3043e-02  1.1171e-02 -5.6432 4.174e-08 ***
log(income)       1.8149e+00  3.8016e-01  4.7740 2.949e-06 ***
miles             8.2262e-06  8.8213e-06  0.9325  0.351884
drinkagec[18,19)  2.7509e-02  6.5701e-02  0.4187  0.675769
drinkagec[19,20) -1.9096e-02  3.9939e-02 -0.4781  0.632940
drinkagec[20,21)  3.0875e-02  4.5129e-02  0.6842  0.494457
jailyes           1.2644e-02  1.2027e-01  0.1051  0.916349
serviceyes        3.4135e-02  1.3804e-01  0.2473  0.804871
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    10.29
Residual Sum of Squares: 6.585
R-Squared:      0.36008
Adj. R-Squared: 0.21475
F-statistic: 17.0684 on 9 and 273 DF, p-value: < 2.22e-16
```

The result shows that only beer tax, unemployment rate and income have statistically significant effect on fatalities.

The effect of dunk driving laws is not significant.

Model Diagnostic

- Independence

Every states make their own law, even though some of their situations are similar.

We have chosen unobserved time-invariant individual effect as one of the variables to control the influence of different states.

We think that the previous outcomes will not affect the future.

- 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.

Model Diagnostic

- Zero-mean and equal variance

The Residuals vs Fitted Values plot

Zero mean and equal variance

- Normality

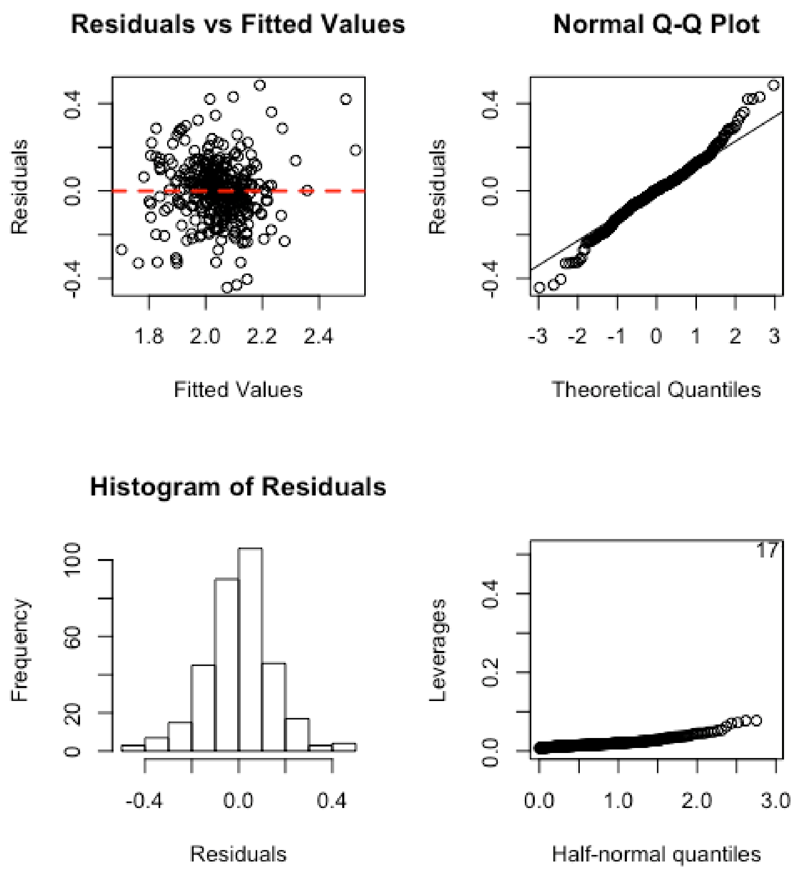
The Q-Q Plot and Histogram of Residuals

Slightly heavy tailed

- No large outliers

The half-normal plot

Only one influential observations



Model Diagnostic

- No perfect multicollinearity

The VIF score indicates the proportion by which the variance of an estimator increases due to the inclusion of a particular covariate.

Therefore, the model coefficients are stable and unbiased.

Variance Inflation Factors of Variables Table

beertax	drinkage	jailyes	breathyes	unemp	miles	log(income)	serviceyes
1.0505	1.0350	4.1313	1.0316	1.5182	1.0148	1.5483	4.1642

Contents

Instrument Variable

In this case, we choose beertax, unemp, log(income) and drinkage as instrument variables when studying the casual effect between jail sentence and fatality rate.

We set jail sentence `yes` as treatment group 1, and jail sentence `no` as control group 0. Note it is unbalanced: 94 cases in group 1 but 242 in group2.

Summary for all data						
Mean	n	Beertax	unemp	Log(income)	drinkage	frate
Treated	94	0.4846	7.9383	9.4855	20.2979	2.2946
Control	242	0.5244	7.1169	9.5411	20.5169	1.9417

The instrument variables has a correlation with jail sentence.

Propensity Score

The propensity score model is a probit/logit model with D as the dependent variable and x as independent variables.

$$p(X) = P(D=1|X) = E[D|X].$$

The propensity score is the conditional (predicted) probability of receiving treatment given pre-treatment characteristics x.

In analysis, we build a logistic regression on treatment jail sentence with instrument variables, then the fitted value is the propensity score in each case.

```
ps <- glm(jail ~  beertax+unemp+log_income+drinkage, data = data, family = binomial())

data$psvalue <- predict(ps, type = "response")
```

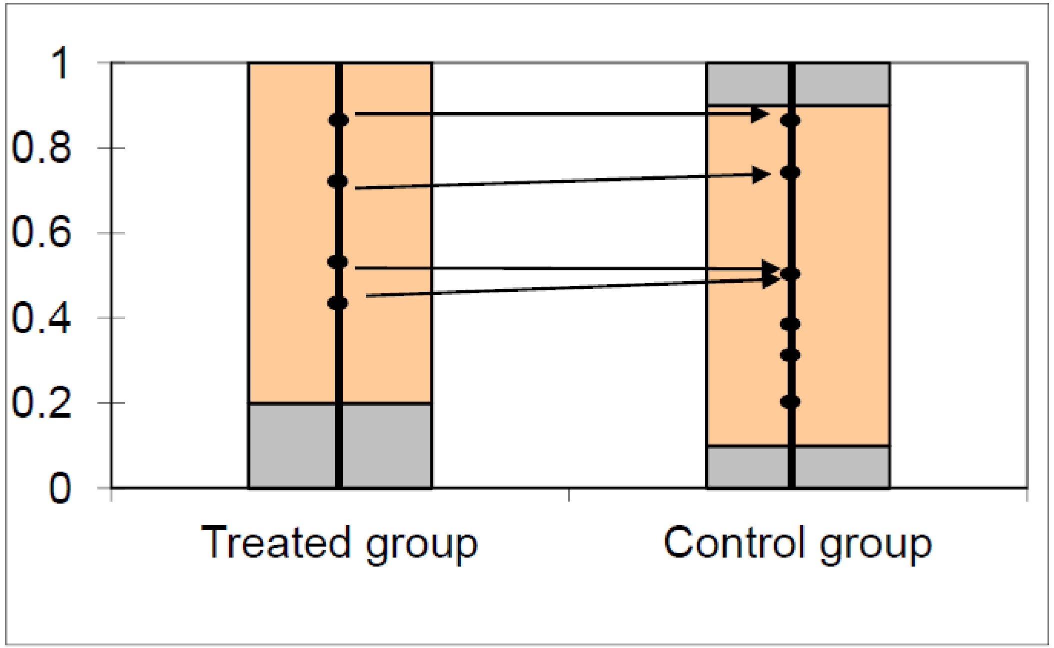
Matching

Matching methods: for each treated observation i, we need to find matches of control observation(s) j with similar characteristics based on their propensity scores.

Several matching methods are available, in this case, we simply use the nearest neighbor method.

The right graph is from[2].

Nearest neighbor matching




```
library(MatchIt)

m.nn <- matchit(jail ~  beertax+unemp+log_income+drinkage, data = data, method = "nearest",ratio = 1)
```

Assumption [2]

- No general equilibrium effects

Treatment does not indirectly affect the control observations.

The case where jail is `yes` will not influence jail is `no`.

- Matching assumption

For each treated observation, there is a matched control observation with similar x.

- Conditional independence assumption (SUTVA)

For observational studies, the outcomes are independent of treatment, conditional on x. i.e. $y_1, y_2 \perp D|x$. In null hypothesis, we assume there is no effect of jail on fatality.

- Balancing condition

Assignment to treatment is independent of the x characteristics, given the same propensity score. i.e. $D \perp x|p(x)$. It can be checked in logistic regression.

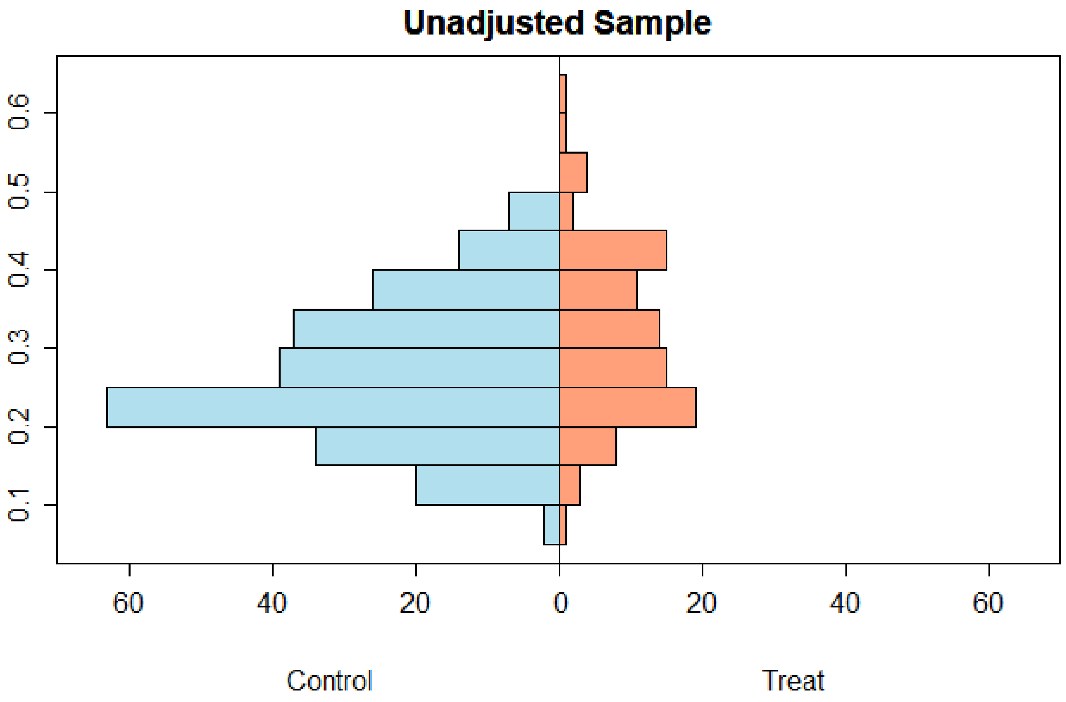
Result

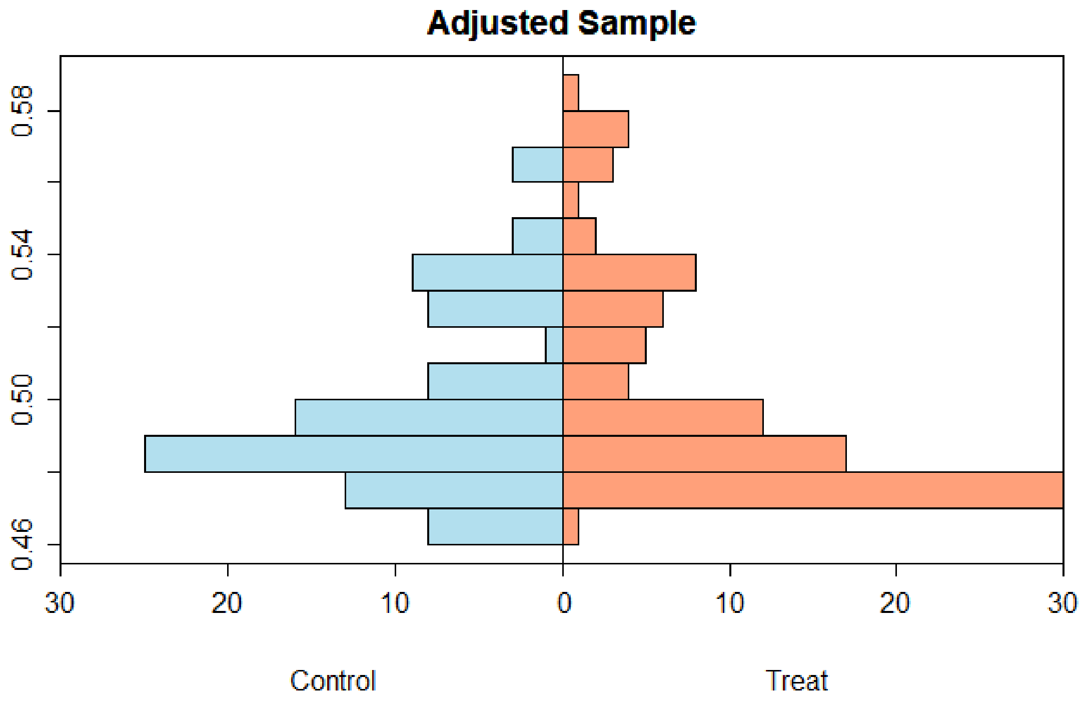
Summary for all data						
Mean	n	Beertax	unemp	Log(income)	drinkage	frate
Treated	94	0.4846	7.9383	9.4855	20.2979	2.2946
Control	242	0.5244	7.1169	9.5411	20.5169	1.9417

Summary for matched data						
Mean	n	Beertax	unemp	Log(income)	drinkage	frate
Treated	94	0.4846	7.9383	9.4855	20.2979	2.2946
Control	94	0.4865	7.7521	9.4912	20.3971	2.0463

The mean of each instrument variable is closer between treated and control group.

Result





chisq test: 0.005 (unbalanced)

chisq test: 0.967 (balanced)

We get a more balanced sample on jail sentence after matching.

Result

We used paired t test on the balanced data to test the effect of jail sentence on fatality rate.[4]

```
t.test(matched.cases$yT,matched.cases$yC, paired = TRUE)
```

```
        Paired t-test

data:  matched.cases$yT and matched.cases$yC
t = 3.2816, df = 93, p-value = 0.001454
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.09803828 0.39852410
sample estimates:
mean of the differences
              0.2482812
```

The result shows that there is a significant effect of jail sentence on fatality rate.

On average, the fatality rate is 0.25 higher when jail sentence is `yes` than jail sentence is `no`.

So we draw a conclusion that there is a casual effect of jail sentence on fatality rate.

However, the mandatory jail sentence may lead the increase of fatality rate.

Note: the p-value depends on the choice of instrument variables and the method of matching. In other case, the result can be insignificant.

Contents

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, is correlated with reducing 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[1].

Conclusion

For casual inference,

The mandatory jail sentence has a significant causal effect on fatality rate.

However, 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.

Suggestion

Here are suggestions for policymakers:

(a) There is a significant positive correlation between beer tax and 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 vehicle fatality rate, but carrying out the mandatory jail sentence can lead to the increase of fatality rate. The policymaker should introspect on the rationalization of law in drinking and driving.

Reference

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

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

[3] 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

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

(Suggestion from Connor)

Q & A Session

Thank you very much!