Shooting Down the "More Guns, Less Crime" Hypothesis

2023-10-02

INTRODUCTION

The impact of guns on crime in America has triggered a lot of public debate. Many strongly believe that state laws enabling citizens to carry concealed handguns had reduced crime. According to this view, gun control laws take away guns from law-abiding citizens, while would-be criminals ignore those leaving potential victims defenceless. Following this view, The National Rifle Association (NRA) and many politicians across the country advance the cause of greater freedom to carry guns. As a result, many states in the United States have passed right-to-carry laws (also known as a shall-issue laws).

A Shall-issue law is one that requires that governments issue concealed carry handgun permits to any applicant who meets the necessary criteria. These criteria are: the applicant must

- be an adult
- have no significant criminal record
- no history of mental illness
- have successfully complete a course in firearms safety training (if required by law).

If these criteria are met, the granting authority has no discretion in the awarding of the licenses, and there is no requirement of the applicant to demonstrate "good cause". Guns is a balanced panel of data on 50 US states, plus the District of Columbia (for a total of 51 "states"), by year for 1977 – 1999. Each observation is a given state in a given year. There are a total of 51 states $\times$ 23 years = 1173 observations.

Variable	Definition
year	Year (1977-1999)
violent	violent crime rate (incidents per 100,000 members of the population)
murder	murder rate (incidents per 100,000)
robbery	robbery rate (incidents per 100,000)
prisoners	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
afam	percent of state population that is African-American, ages 10 to 64.
cauc	percent of state population that is Caucasian, ages 10 to 64.
male	percent of state population that is male, ages 10 to 29.
population	state population, in millions of people.
income	real per capita personal income in the state (US dollars).
density	population per square mile of land area, divided by 1,000.
state	factor indicating state.
law	factor. Does the state have a shall carry law in effect in that year?

With the provided information we need to carry a study whether providing guns to the citizens who are eligible for the mentioned criteria will reduce the crime rate.

In the provide information we have three different crime rate: `violent`, `robbery` and `murder`, but our study mainly focus on violent crime rate but we will extend out study to see how robbery, murder and `other_violences` (violent - (robbery + murder)) are controlled using shall law and with other explanatory variables and we can see whether all 4 crime rates are travelling in same direction or in different direction.

EXPLORATORY DATA ANALYSIS

We are working on a balanced panel dataset on 50 US states, plus the District of Columbia over a 23-year period of time. The total number of observations is 1173 and there's no null values in the dataset. Before building our hypotheses and models, we will do descriptive analysis and some exploratory analysis to have an overview, as well as better insights about our data such as relationship between variables, trend, … Besides 3 dependent variables (violent, robbery, murder), we will also focus on 2 important explanatory variables - shall and prisoners to answer our main the question "Do more guns reduce crime?"

Descriptive statistics:

From Statistics we can observe that Variance for violence, robbey, other violences, prisoners are higher compared to other variables Also, Violence, Robbery, Murder,Other violences, prisoners, density data have a significant skewness and might need transformation only we can decide after looking to distribution and box plots of the same.

Correlation between variables:

From the above output we can observe that `violent` is highly correlated with `murder`, `robbery`, `other_violences`, `prisoners` which is understandable w.r.t murder and robbery but violent increases with prisoners is not a good sign to our study may be some causuality bias is expected let's see.

`violent` is moderately correlated with `afam`, `cauc`, `income` , and `density` which is quite fine for now. `violent` is least correlated with `law`, `population` , This may not be accurate as the number of shall-carry law record is much lower than no shall-carry law in this dataset.

There's also a positive correlation between density and crime rate, which makes sense to us. The percentage of white in the population is strongly negatively correlated with the percentage of black in the population. And, as time gone by, the percentage of young male in the population decrease.

Variables studying in detail

Crimes across years

Let's try to plot the different types of crimes varied across years with and without shall law

```r
never_passed_law <- c("Alabama", "California", "Colorado", "Connecticut", "Delaware", "District of Colu

pre1977 <- c("Indiana", "New Hampshire", "Vermont", "Washington")

post1977_pre1990 <- c("Florida", "Georgia", "Maine", "North Dakota", "Pennsylvania", "South Dakota", "U

post_1990 <- c("Alaska", "Arizona", "Arkansas", "Idaho", "Kentucky", "Louisiana", "Mississippi", "Monta

vectors_list <- list(
  never_passed_law = never_passed_law,
  pre1977 = pre1977,
  post1977_pre1990 = post1977_pre1990,
  post_1990 = post_1990
  )

# Sample vectors with different lengths
```

```r
# Create a dataframe with elements and corresponding vector names
introd <- data.frame(
  state = unlist(vectors_list),
  introduction = rep(names(vectors_list), sapply(vectors_list, length))
)

Guns <- merge(Guns, introd, by = "state", all.x = TRUE)
```

```r
mean_v <- Guns %>%
  group_by(year = as.numeric(as.character(Guns$year)), introduction) %>%
  summarise(mean_violent = round(mean(violent), 2))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

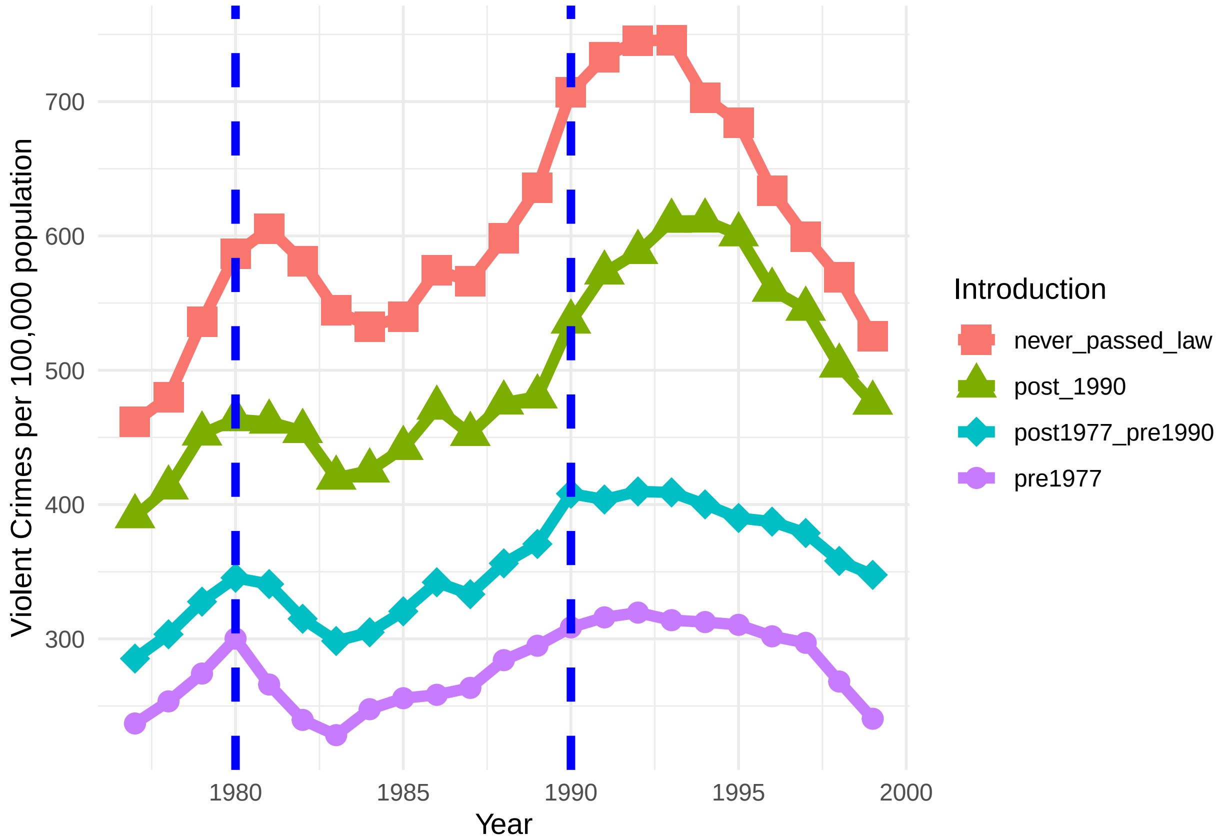
```r
mean_m <- Guns %>%
  group_by(year = as.numeric(as.character(Guns$year)), introduction) %>%
  summarise(mean_murder = round(mean(murder), 2))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

```r
mean_r <- Guns %>%
  group_by(year = as.numeric(as.character(Guns$year)), introduction) %>%
  summarise(mean_robbery = round(mean(robbery), 2))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

```r
ggplot(mean_v, aes(x = ymd(year, truncated = 2L), y = mean_violent, group = introduction, color = introd
  geom_line(linewidth=2) +
  geom_point(size = 5) +
  scale_shape_manual(values=c(15,17,18,20))+
  labs(col = "Introduction", x = "Year", y = "Violent Crimes per 100,000 population") +
  theme_minimal() +
  theme(legend.position = "right")+
  guides(colour = guide_legend(override.aes = list(shape = c(15, 17, 18, 20))), shape = "none") +
  geom_vline(xintercept = as.numeric(ymd("1980-01-01")), linetype="dashed",color = "blue", linewidth=1.5
  geom_vline(xintercept = as.numeric(ymd("1990-01-01")), linetype="dashed",color = "blue", linewidth=1.5
```




```
# REMOVING PRE 1977 LAW ADOPTING OBS

treat_v <- mean_v[mean_v$introduction != "pre1977", ]
treat_m <- mean_m[mean_m$introduction != "pre1977", ]
treat_r <- mean_r[mean_r$introduction != "pre1977", ]


# ADDING COLUMN TRUE/FALSE FOR LAW

treat_v$law <- ifelse(treat_v$introduction == "post1977_pre1990" & treat_v$year >1979, TRUE,
                      ifelse(treat_v$introduction == "post_1990" & treat_v$year > 1989, TRUE, F

treat_m$law <- ifelse(treat_m$introduction == "post1977_pre1990" & treat_m$year >1979, TRUE,
                      ifelse(treat_m$introduction == "post_1990" & treat_m$year > 1989, TRUE, F

treat_r$law <- ifelse(treat_r$introduction == "post1977_pre1990" & treat_r$year >1979, TRUE,
                      ifelse(treat_r$introduction == "post_1990" & treat_r$year > 1989, TRUE, F


# ADDING FIRST TREATMENT YEAR

treat_v$first_treat <- ifelse(treat_v$introduction == "post1977_pre1990", 1980,
                      ifelse(treat_v$introduction == "post_1990", 1990, Inf))

treat_m$first_treat <- ifelse(treat_m$introduction == "post1977_pre1990", 1980,
                      ifelse(treat_r$introduction == "post_1990", 1990, Inf))
```

```r
treat_r$first_treat <- ifelse(treat_r$introduction == "post1977_pre1990", 1980,
                              ifelse(treat_m$introduction == "post_1990", 1990, Inf))


# ADDING RELATIVE TIME TO TREATMENT

treat_v$rel_time <-  treat_v$year - treat_v$first_treat
treat_m$rel_time <-  treat_m$year - treat_m$first_treat
treat_r$rel_time <-  treat_r$year - treat_r$first_treat
```

```r
static <- did2s(treat_v,
                yname = "mean_violent", first_stage = ~ 0 | introduction + year,
                second_stage = ~i(law, ref=FALSE), treatment = "law",
                cluster_var = "introduction")
```

```
## Running Two-stage Difference-in-Differences
##  - first stage formula `~ 0 | introduction + year`
##  - second stage formula `~ i(law, ref = FALSE)`
##  - The indicator variable that denotes when treatment is on is `law`
##  - Standard errors will be clustered by `introduction`
```

```r
fixest::esttable(static)
```

```
##                         static
## Dependent Var.:   mean_violent
##
## law = TRUE      -31.98 (24.72)
## _______________ ______________
## S.E. type               Custom
## Observations                69
## R2                     0.15938
## Adj. R2                0.15938
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
es <- did2s(treat_v,
            yname = "mean_violent", first_stage = ~ 0 | introduction + year,
            second_stage = ~i(rel_time, ref= -c(1, Inf)), treatment = "law",
            cluster_var = "introduction")
```

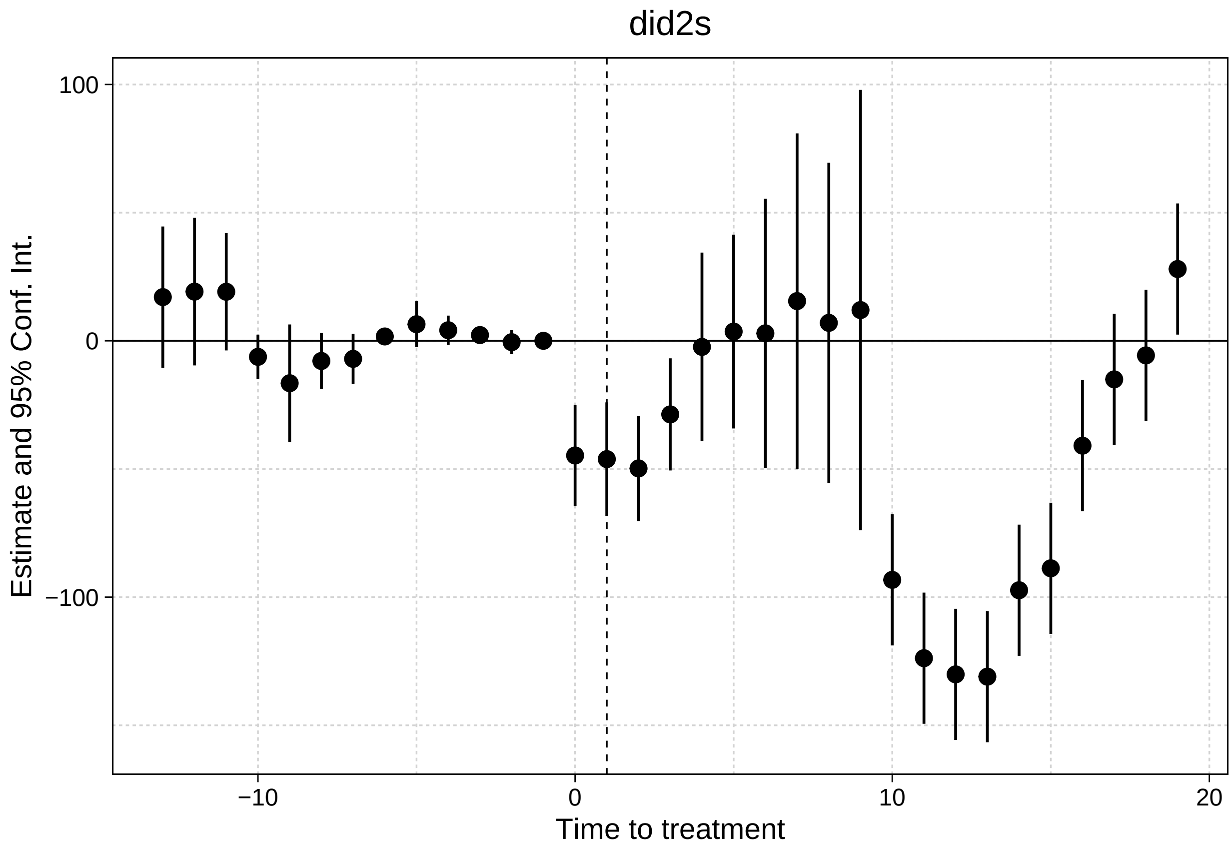
```
## Running Two-stage Difference-in-Differences
##  - first stage formula `~ 0 | introduction + year`
##  - second stage formula `~ i(rel_time, ref = -c(1, Inf))`
##  - The indicator variable that denotes when treatment is on is `law`
##  - Standard errors will be clustered by `introduction`
```

```r
es
```

```
## OLS estimation, Dep. Var.: mean_violent
## Observations: 69
```

```
## Standard-errors: Custom
##                  Estimate Std. Error   t value   Pr(>|t|)
## rel_time::-13   17.056368   14.05556   1.213496 2.3263e-01
## rel_time::-12   19.189701   14.68781   1.306505 1.9944e-01
## rel_time::-11   19.139701   11.67887   1.638832 1.0973e-01
## rel_time::-10   -6.253077    4.42159  -1.414214 1.6566e-01
## rel_time::-9   -16.543077   11.69772  -1.414214 1.6566e-01
## rel_time::-8    -7.868077    5.56357  -1.414214 1.6566e-01
## rel_time::-7    -7.038077    4.97667  -1.414214 1.6566e-01
## rel_time::-6     1.696923    1.19991   1.414214 1.6566e-01
## rel_time::-5     6.471923    4.57634   1.414214 1.6566e-01
## rel_time::-4     4.121923    2.91464   1.414214 1.6566e-01
## rel_time::-3     2.244017    1.62902   1.377522 1.7663e-01
## rel_time::-2    -0.534316    2.39170  -0.223404 8.2445e-01
## rel_time::0    -44.737244   10.01166  -4.468513 7.1856e-05 ***
## rel_time::1    -46.157244   11.29554  -4.086326 2.2597e-04 ***
## rel_time::2    -49.784744   10.47165  -4.754239 3.0065e-05 ***
## rel_time::3    -28.709744   11.16083  -2.572365 1.4249e-02 *
## rel_time::4     -2.377244   18.77646  -0.126608 8.9994e-01
## rel_time::5      3.610256   19.28966   0.187160 8.5256e-01
## rel_time::6      2.915256   26.78241   0.108850 9.1391e-01
## rel_time::7     15.492756   33.38970   0.463998 6.4537e-01
## rel_time::8      7.007756   31.86302   0.219934 8.2713e-01
## rel_time::9     11.995256   43.82565   0.273704 7.8583e-01
## rel_time::10   -93.261410   13.05455  -7.143977 1.8240e-08 ***
## rel_time::11  -123.821410   13.05455  -9.484923 1.9006e-11 ***
## rel_time::12  -130.131410   13.05455  -9.968279 5.0028e-12 ***
## rel_time::13  -131.021410   13.05455 -10.036455 4.1545e-12 ***
## rel_time::14   -97.311410   13.05455  -7.454213 7.0963e-09 ***
## rel_time::15   -88.751410   13.05455  -6.798503 5.2679e-08 ***
## rel_time::16   -40.911410   13.05455  -3.133881 3.3692e-03 **
## rel_time::17   -15.041410   13.05455  -1.152197 2.5663e-01
## rel_time::18    -5.691410   13.05455  -0.435971 6.6539e-01
## rel_time::19    28.008590   13.05455   2.145504 3.8542e-02 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 16.6   Adj. R2: 0.650932
```

```r
# fixest::iplot(es_mod)
# Vanilla option (above) is fine, but we can tweak a bit...
es |>
  ggiplot(
    main     = "did2s",
    xlab     = "Time to treatment",    # Drop any leads/lags greater than |9|
    ref.line = 1
  )
```

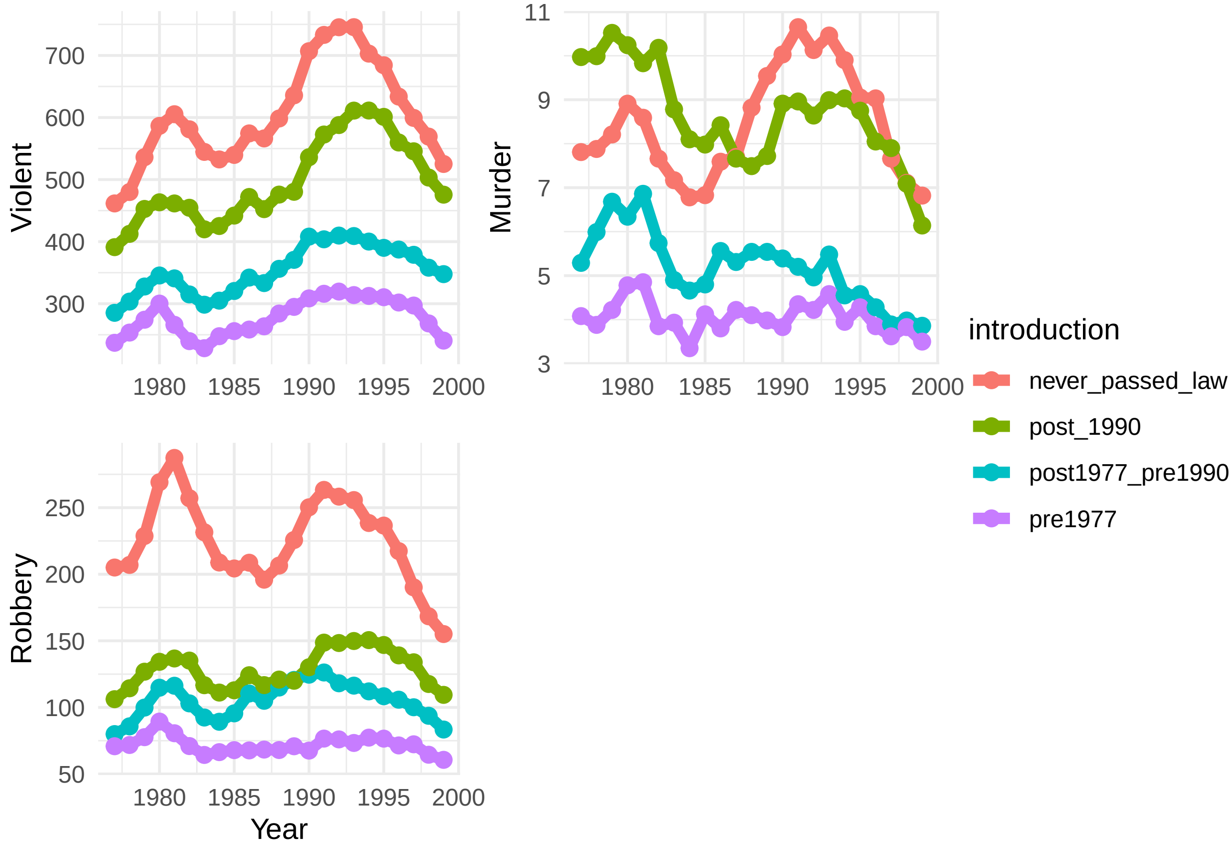



```
p_v <- ggplot(mean_v, aes(x = year, y = mean_violent, color = introduction)) +
  geom_line(linewidth=2) +
  geom_point(size = 2.5) +
  labs(y = "Violent", x = element_blank()) +
  theme_minimal()

p_m <- ggplot(mean_m, aes(x = year, y = mean_murder, color = introduction)) +
  geom_line(linewidth=2) +
  geom_point(size = 2.5) +
  labs(y = "Murder", x = element_blank()) +
  theme_minimal()

p_r <- ggplot(mean_r, aes(x = year, y = mean_robbery, color = introduction)) +
  geom_line(linewidth=2) +
  geom_point(size = 2.5) +
  labs(y = "Robbery", x = "Year") +
  theme_minimal()
```

```
ggarrange(p_v, p_m, p_r, ncol=2, nrow=2, common.legend = TRUE, legend="right")
```


Violent
700
600
500
400
300
1980
1985
1990
1995
2000
Murder
11
9
7
5
3
Robbery
250
200
150
100
50
Year
introduction
never_passed_law
post_1990
post1977_pre1990
pre1977