**Project Report**

# The Impact of
# 'Shall-Carry' law and guns on Crime Rate in the U.S.

**Applied Econometrics and Time Series Analysis**

**Professor Moran Blueshtein**

Submitted on
Apr 20, 2018

The University of Texas at Dallas
Naveen Jindal School of Management

by

Saundarya Badhrinathan

# Table of Contents

## Introduction:

The impact of incarceration rate and guns on crime has always engendered many public debates in America. Many states have adopted the right-to-carry laws, also known as the shall-issue laws, as an effort toward regulation of guns. A shall-carry law requires that government issue concealed carry handgun permits to any applicant who meets the necessary criteria. These criteria are that the applicant must be an adult, have no history of mental illness and no significant criminal record and should successfully complete a course in firearms safety training (if required by law). If these criteria are met, then the granting authority has no discretion in awarding the license. Many believe this would lead to lower crime rates as restrictions would be placed to own a gun. Whereas, many others who are against this law suggest that crime rate would increase as criminals would be more willing to commit a crime since there would be fewer guns issues, which would have been a threat to them.

The objective of this report is to evaluate the effect of higher incarceration rate and the effect of the shall-issue law on the reported crime rate, specifically robbery rate, murder rate and violent crime rate (all measured as incidents per 100,000 members of the population). A balanced panel of data of the 50 states of the United States, plus the District of Columbia, from 1977 through 1999 is used for the analysis.

## Data understanding:

We started out by performing a univariate analysis to gain an understanding about the variables in the dataset.

Missing data: There were no missing data to be dealt with.

### Univariate Analysis:

```
. summarize vio   mur     rob     incarc_rate     pb1064  pw1064  pm1029  pop
>    avginc  density shall
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| vio | 1,173 | 503.0747 | 334.2772 | 47 | 2921.8 |
| mur | 1,173 | 7.665132 | 7.52271 | .2 | 80.6 |
| rob | 1,173 | 161.8202 | 170.51 | 6.4 | 1635.1 |
| incarc_rate | 1,173 | 226.5797 | 178.8881 | 19 | 1913 |
| pb1064 | 1,173 | 5.336217 | 4.885688 | .2482066 | 26.97957 |
| pw1064 | 1,173 | 62.94543 | 9.761527 | 21.78043 | 76.52575 |
| pm1029 | 1,173 | 16.08113 | 1.732143 | 12.21368 | 22.35269 |
| pop | 1,173 | 4.816341 | 5.252115 | .402753 | 33.14512 |
| avginc | 1,173 | 13.7248 | 2.554543 | 8.554884 | 23.64671 |
| density | 1,173 | .3520382 | 1.355472 | .0007071 | 11.10212 |
| shall | 1,173 | .2429668 | .4290581 | 0 | 1 |

## Corellation matrix:

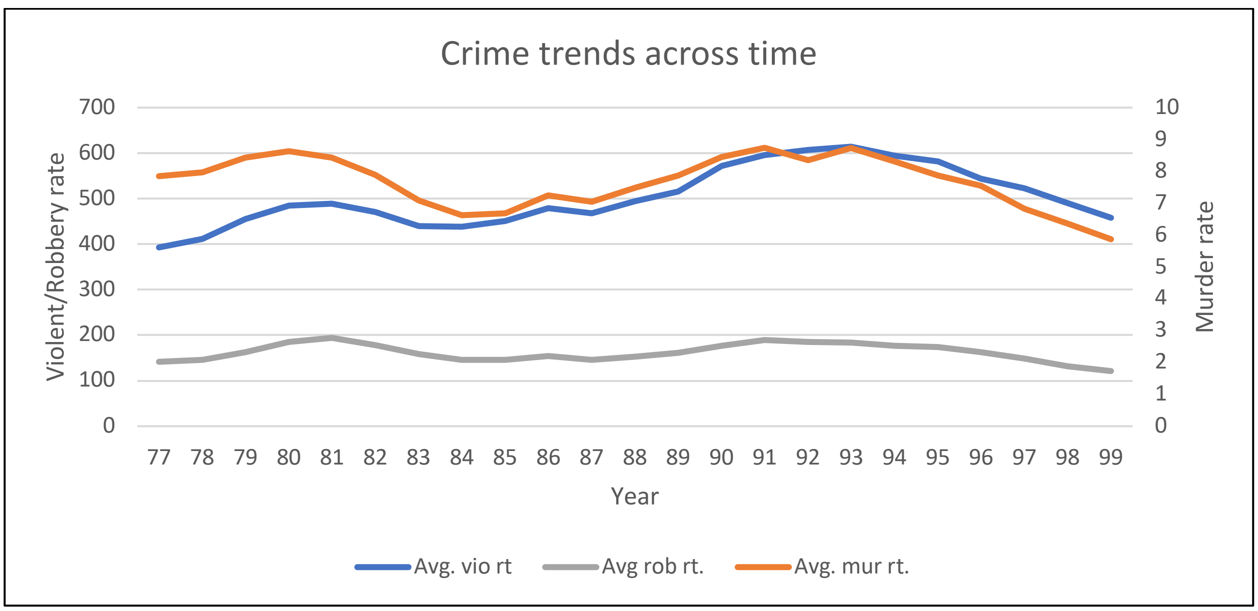
We created a corelation matrix to understand the relationship between variables. On a broad level all types of crimes are strongly corealted to each other. Percentage of white population is negatively corelated with crime and percentage of black population is positively corelated with crime. Population density has a strong corelation with crime whereas averageincome has a weak corelation.

```
. corr vio mur rob incarc_rate pb1064 pw1064 pm1029 pop avginc density shall
(obs=1,173)
```

| | vio | mur | rob | incarc~e | pb1064 | pw1064 | pm1029 | pop | avginc | density | shall |
|---|---|---|---|---|---|---|---|---|---|---|---|
| vio | 1.0000 | | | | | | | | | | |
| mur | 0.8265 | 1.0000 | | | | | | | | | |
| rob | 0.9071 | 0.7976 | 1.0000 | | | | | | | | |
| incarc_rate | 0.7027 | 0.7096 | 0.5668 | 1.0000 | | | | | | | |
| pb1064 | 0.5698 | 0.6018 | 0.5812 | 0.5308 | 1.0000 | | | | | | |
| pw1064 | -0.5730 | -0.6154 | -0.5842 | -0.5271 | -0.9820 | 1.0000 | | | | | |
| pm1029 | -0.1696 | 0.0150 | -0.0860 | -0.4463 | 0.0162 | -0.0126 | 1.0000 | | | | |
| pop | 0.3190 | 0.0999 | 0.3172 | 0.0953 | 0.0581 | -0.0654 | -0.0975 | 1.0000 | | | |
| avginc | 0.4080 | 0.2206 | 0.4148 | 0.4615 | 0.2627 | -0.1912 | -0.5279 | 0.2152 | 1.0000 | | |
| density | 0.6647 | 0.7486 | 0.7818 | 0.5593 | 0.5432 | -0.5551 | -0.0637 | -0.0780 | 0.3433 | 1.0000 | |
| shall | -0.2069 | -0.1794 | -0.2125 | 0.0424 | -0.1839 | 0.2123 | -0.2772 | -0.1244 | -0.0000 | -0.1126 | 1.0000 |

# Exploratory Data Analysis:

## Crimes across time:

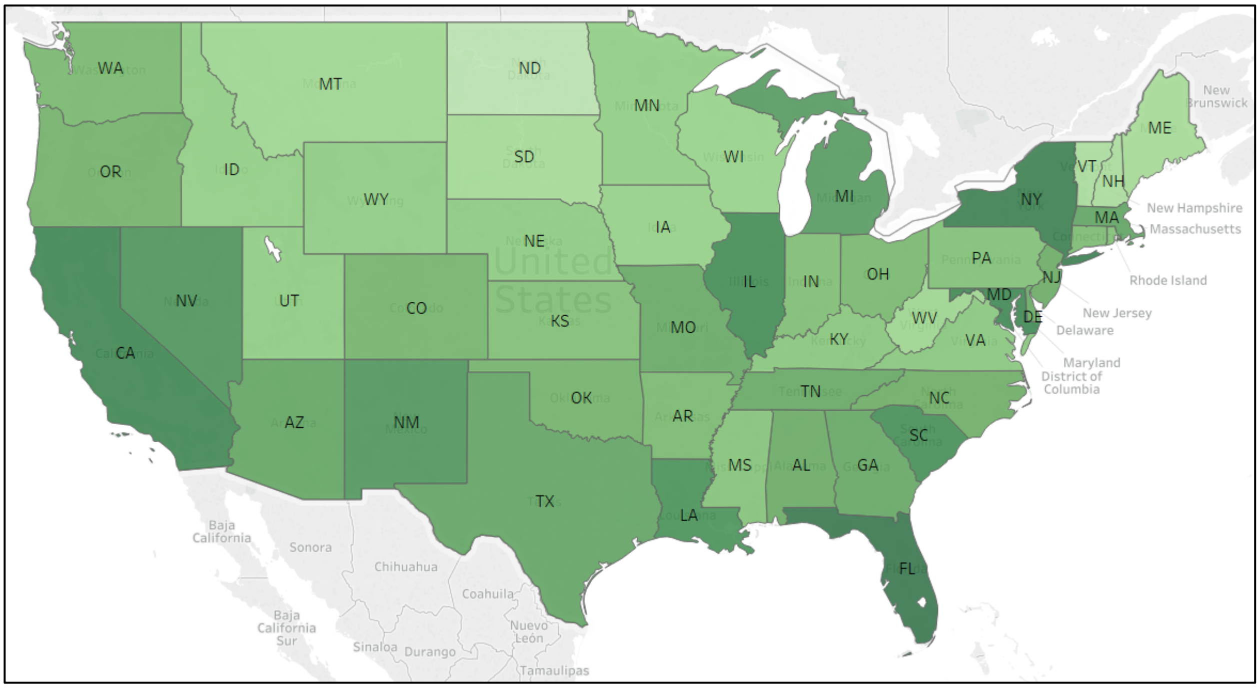



First step was to understand the distribution of data across time. Here we can see that the three nationwide averages for violent crime rate, Robbery rate and murder rate all follow a similar trend across time. The distribution looks like bimodal, with a peak in crime in the years 1980 and 1991 with a trough at 1984 and continuous decrease post 1993.

Since the three distributions across time are similar we will be summing up the values across year to create a crime score per year for the remaining exploratory data analysis.

## Crimes across states:

### Violent Crimes:




Average violent crime rate was calculated across time and the states were plotted.

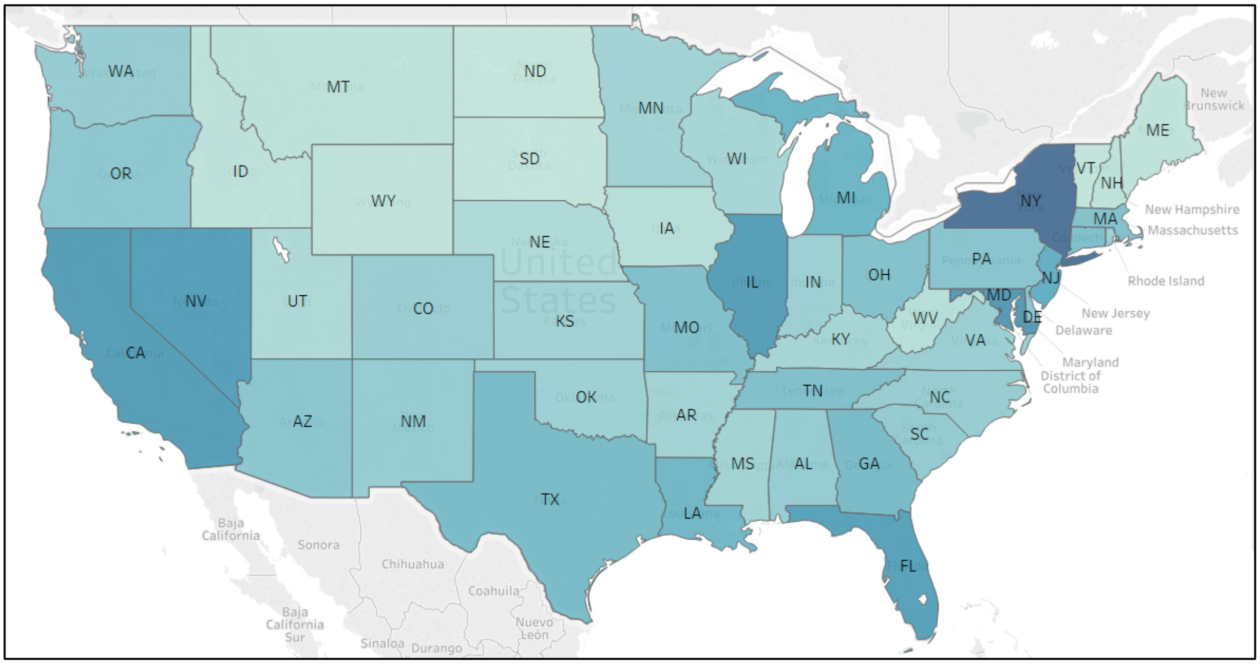
Top states:

1. DC
2. Florida
3. New York

Bottom state:

1. North Dakota
2. New Hampshire
3. Vermont

## Robbery:




Average robbery rate was calculated across time and the states were plotted.

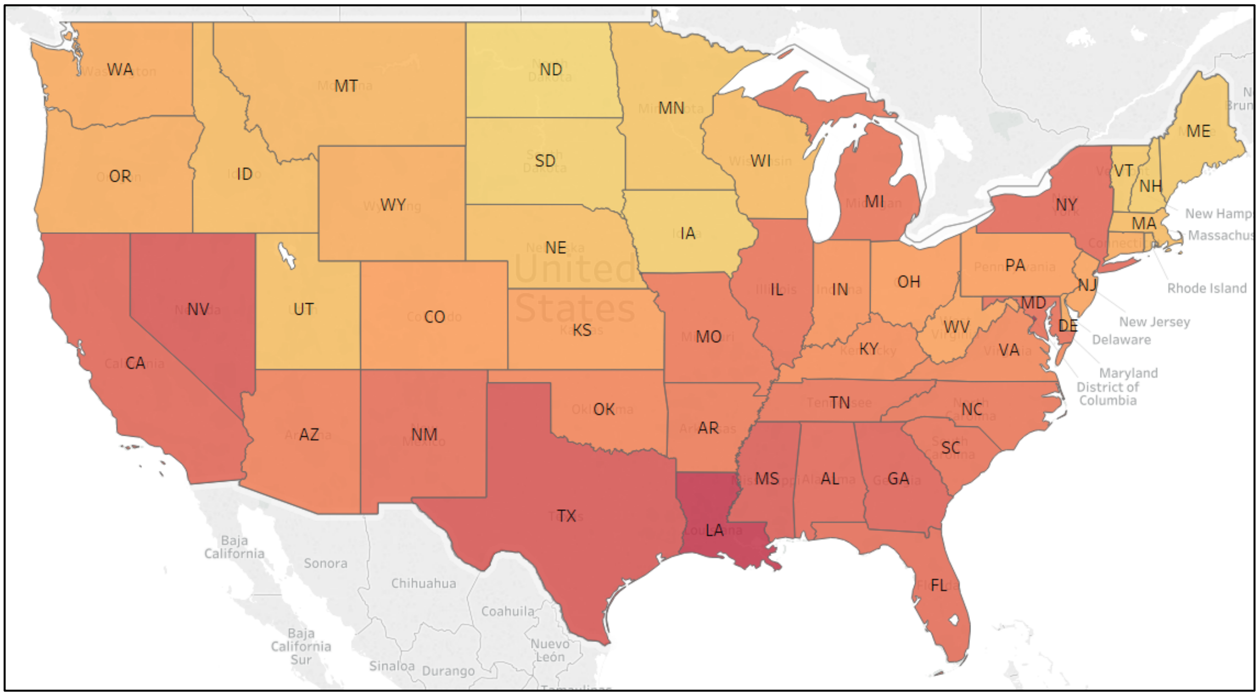
Top states:

1. DC
2. Maryland
3. New York

Bottom state:

1. North Dakota
2. South Dakota
3. Vermont

## Murder:




Average murder rate was calculated across time and the states were plotted.

Top states:

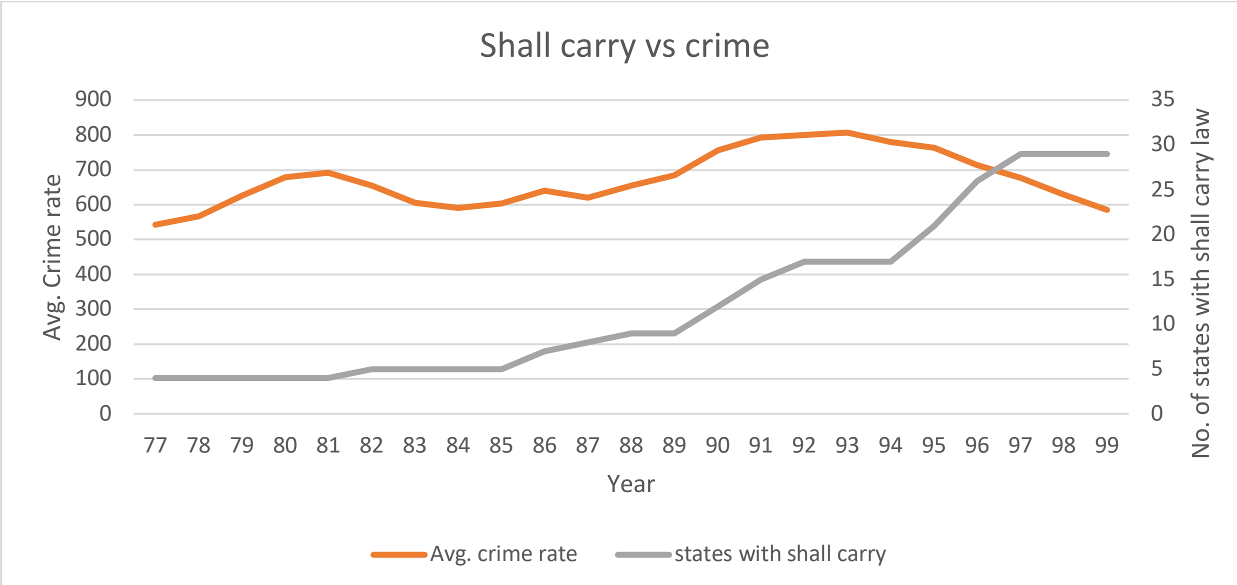
1. DC
2. Louisiana
3. Texas

Bottom state:

1. North Dakota
2. South Dakota
3. Iowa

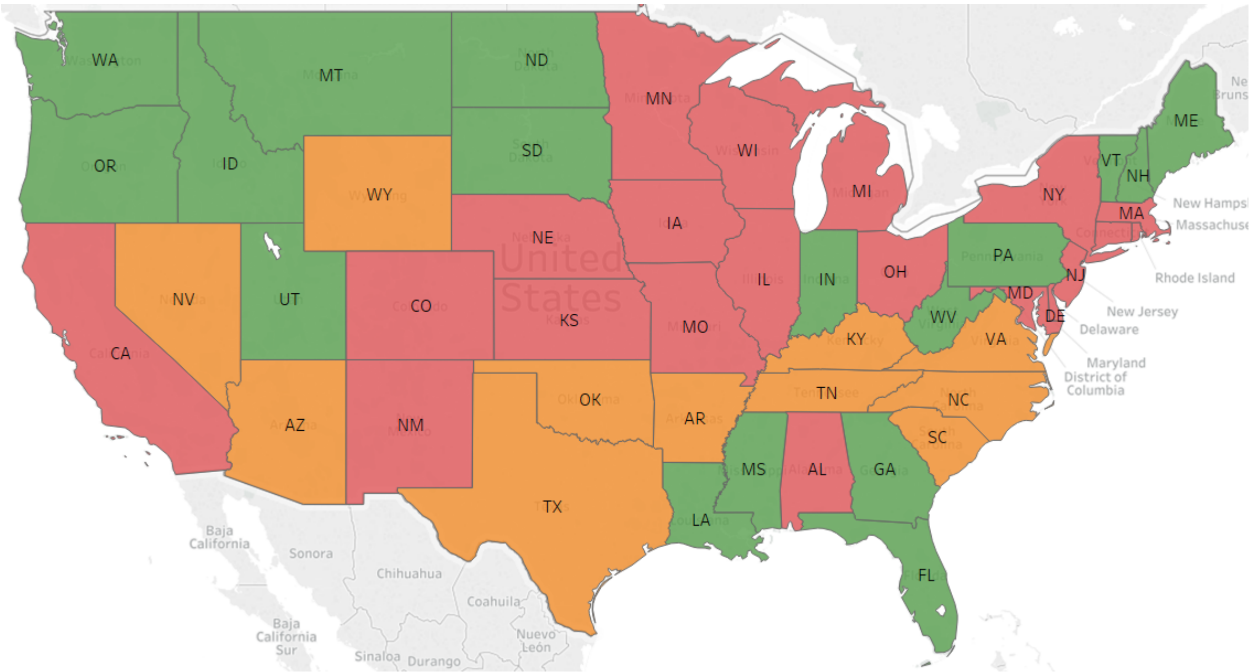
Overall DC is among the top states for all three types of crimes and North Dakota is among the bottom three for all types of crimes. Most of the sun belt states have higher crime rate compared to the rest of USA.

## Shall carry law vs crime:




Nation’s average crime rate and the number of states with shall carry law doesn’t show a clear trend. In the years from 1989 to 1992 there is an increase in the number of states which passed the shall carry law however there is an increase in the nationwide average for crime. Where as from 1994 to 1997 the increase in number of states with the shall carry law is corelated to the decrease in crime. 1997 to 1999 where no new states implemented the law, there is still a decrease in the number of crimes, suggesting other variables are driving crime rate.

Grouping states based on adoption of Shall carry law:

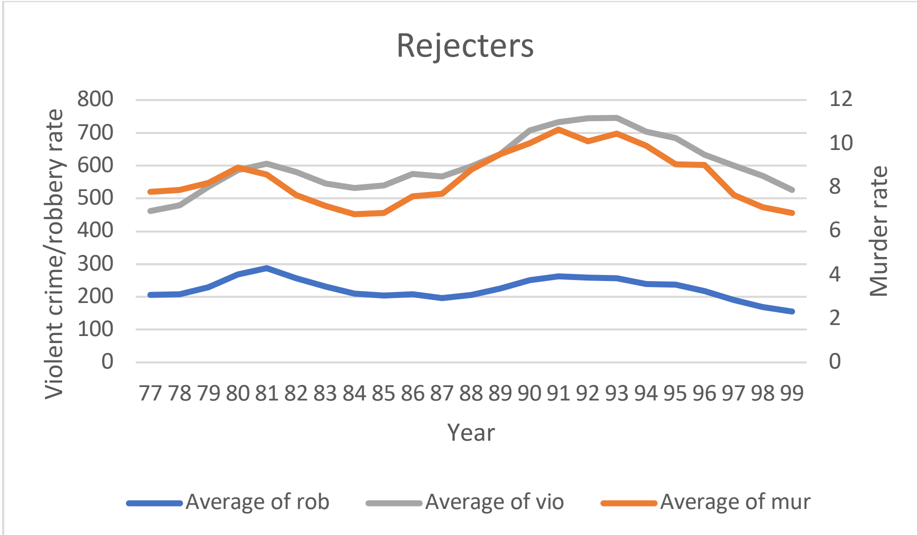


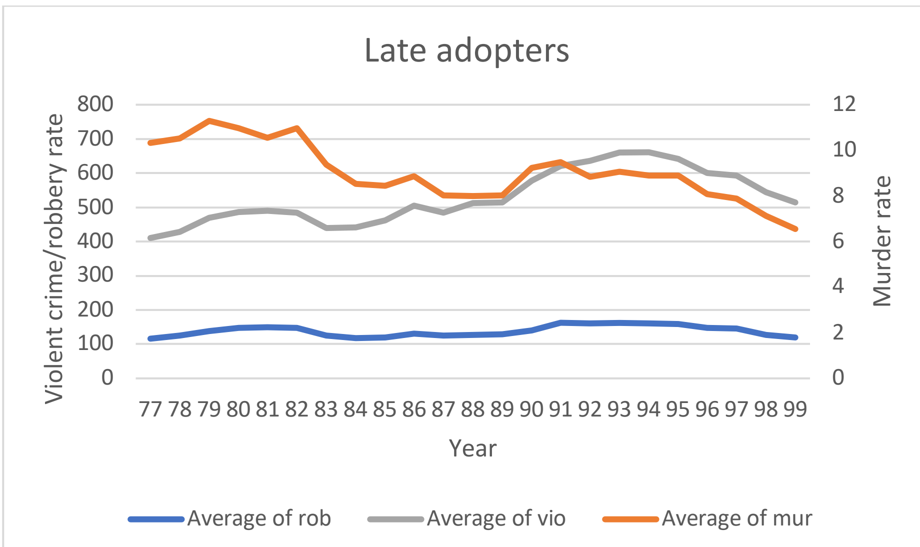

| State | Average of shall | Percentile Rank | Group |
|---|---|---|---|
| AL | 0 | 0 | 1 |
| CA | 0 | 0 | 1 |
| CO | 0 | 0 | 1 |
| CT | 0 | 0 | 1 |
| DC | 0 | 0 | 1 |
| DE | 0 | 0 | 1 |
| HI | 0 | 0 | 1 |
| IA | 0 | 0 | 1 |
| IL | 0 | 0 | 1 |
| KS | 0 | 0 | 1 |
| MA | 0 | 0 | 1 |
| MD | 0 | 0 | 1 |
| MI | 0 | 0 | 1 |
| MN | 0 | 0 | 1 |
| MO | 0 | 0 | 1 |
| NE | 0 | 0 | 1 |
| NJ | 0 | 0 | 1 |
| NM | 0 | 0 | 1 |
| NY | 0 | 0 | 1 |
| OH | 0 | 0 | 1 |
| RI | 0 | 0 | 1 |
| WI | 0 | 0 | 1 |

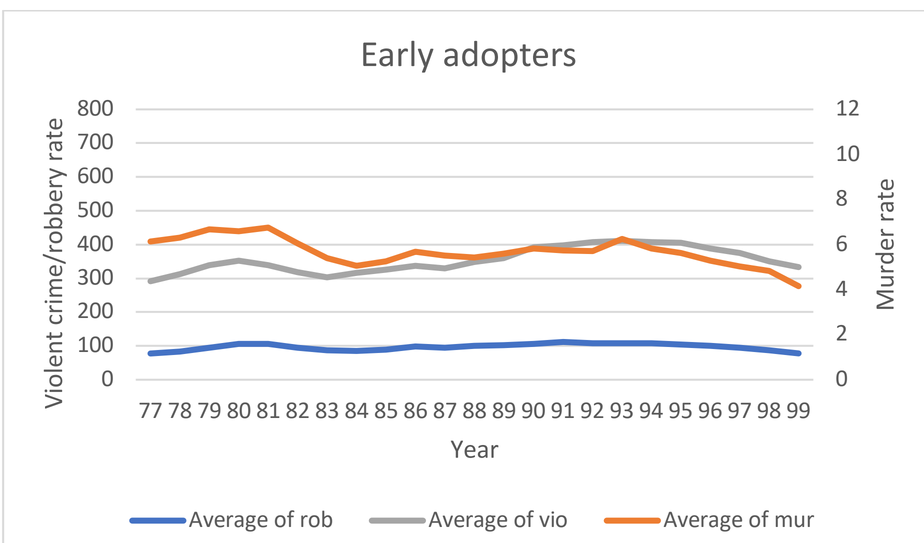
| State | Average of shall | Percentile Rank | Group |
|---|---|---|---|
| KY | 0.130435 | 0.44 | 2 |
| SC | 0.130435 | 0.44 | 2 |
| TX | 0.130435 | 0.44 | 2 |
| AR | 0.173913 | 0.5 | 2 |
| NC | 0.173913 | 0.5 | 2 |
| NV | 0.173913 | 0.5 | 2 |
| OK | 0.173913 | 0.5 | 2 |
| VA | 0.173913 | 0.5 | 2 |
| AK | 0.217391 | 0.6 | 2 |
| AZ | 0.217391 | 0.6 | 2 |
| TN | 0.217391 | 0.6 | 2 |
| WY | 0.217391 | 0.6 | 2 |

| State | Average of shall | Percentile Rank | Group |
|---|---|---|---|
| LA | 0.347826 | 0.68 | 3 |
| MT | 0.347826 | 0.68 | 3 |
| ID | 0.391304 | 0.72 | 3 |
| MS | 0.391304 | 0.72 | 3 |
| OR | 0.391304 | 0.72 | 3 |
| GA | 0.434783 | 0.78 | 3 |
| PA | 0.434783 | 0.78 | 3 |
| WV | 0.434783 | 0.78 | 3 |
| FL | 0.521739 | 0.84 | 3 |
| UT | 0.565217 | 0.86 | 3 |
| ND | 0.608696 | 0.88 | 3 |
| SD | 0.608696 | 0.88 | 3 |
| ME | 0.782609 | 0.92 | 3 |
| IN | 1 | 0.94 | 3 |
| NH | 1 | 0.94 | 3 |
| VT | 1 | 0.94 | 3 |
| WA | 1 | 0.94 | 3 |

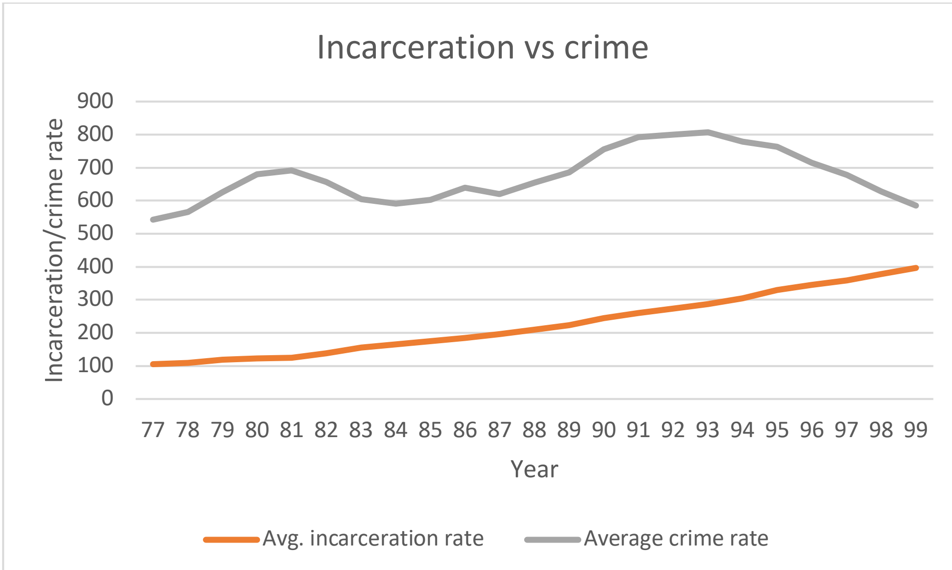
Once the states were split based on their adoption of shall carry law, we observed some difference in trends. The group of states that rejected the shall carry law on an average has higher crime than the states that adopted shall carry. Shall carry law seems to affect murders the most as the increase in murders from 1985 – 1991 in the rejecter states is not present in the adopter states. Also the late adopter states show a sharper drop in the number of murders form 1983 than the other two groups.


Rejecters
800
700
600
500
400
300
200
100
0
Violent crime/robbery rate
12
10
8
6
4
2
0
Murder rate
77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99
Year
Average of rob
Average of vio
Average of mur


Late adopters
800
700
600
500
400
300
200
100
0
Violent crime/robbery rate
12
10
8
6
4
2
0
Murder rate
77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99
Year
Average of rob
Average of vio
Average of mur


Early adopters
800
700
600
500
400
300
200
100
0
Violent crime/robbery rate
12
10
8
6
4
2
0
Murder rate
77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99
Year
Average of rob
Average of vio
Average of mur

### Incarceration vs crime rate:




A bivariate between the previous year's incarceration rate and the current years crime does not show any conclusive trends. Incarceration has been on a steady rise however crime rate has been varying.

## Conclusion from EDA:

1. Shall carry law seems to have a different impact on different types of crimes
2. Incarceration over time does not seem to have a clear impact on crime rate

## Approach:

We will create three different models with the dependent variable being each type of crime:

# Model Building

## Violent Crime Rate:

### Model 1:

We first run a Pooled Effect model with all the independent variables. The results are shown below:

```
. reg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

      Source |       SS           df       MS      Number of obs   =     1,173
-------------+----------------------------------   F(8, 1164)      =    381.10
       Model |  94776406.5         8  11847050.8   Prob > F        =    0.0000
    Residual |  36184330.1     1,164  31086.1942   R-squared       =    0.7237
-------------+----------------------------------   Adj R-squared   =    0.7218
       Total |   130960737     1,172  111741.243   Root MSE        =    176.31

------------------------------------------------------------------------------
         vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .8155948    .044174    18.46   0.000     .7289253    .9022642
      pb1064 |   11.29196   6.864417     1.64   0.100    -2.176053    24.75998
      pw1064 |   2.702383   3.453603     0.78   0.434      -4.0736    9.478366
      pm1029 |   9.316543   4.441374     2.10   0.036      .602549    18.03054
         pop |   18.50567   1.054859    17.54   0.000     16.43603    20.57531
      avginc |   1.234759   3.207344     0.38   0.700    -5.058062    7.527581
     density |   94.66882   5.428439    17.44   0.000      84.0182    105.3194
     1.shall |  -92.72125    13.4257    -6.91   0.000    -119.0625   -66.37997
       _cons |  -178.7774   224.0104    -0.80   0.425    -618.2867    260.7319
------------------------------------------------------------------------------
```

Our next step is to run the same model using the robust standard errors to account for the incorrect standard errors.

```
. reg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

Linear regression                               Number of obs     =      1,173
                                                F(8, 1164)        =     135.12
                                                Prob > F          =     0.0000
                                                R-squared         =     0.7237
                                                Root MSE          =     176.31

------------------------------------------------------------------------------
             |               Robust
         vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .8155948   .0911446     8.95   0.000     .6367687    .9944208
      pb1064 |   11.29196   8.354651     1.35   0.177    -5.099897    27.68382
      pw1064 |   2.702383   4.126139     0.65   0.513    -5.393119    10.79788
      pm1029 |   9.316543   4.778368     1.95   0.051    -.0586355    18.69172
         pop |   18.50567    1.40521    13.17   0.000     15.74865     21.2627
      avginc |   1.234759   3.291204     0.38   0.708    -5.222597    7.692116
     density |   94.66882   8.261852    11.46   0.000     78.45904    110.8786
     1.shall |  -92.72125    14.0256    -6.61   0.000    -120.2395   -65.20297
       _cons |  -178.7774    270.354    -0.66   0.509    -709.2131    351.6584
------------------------------------------------------------------------------
```

Due to the time effect included in panel data, a variable x in time t has an influence over xt+1. Therefore, the assumption that the error terms between different time periods are not serially correlated is violated. Additionally, the variance of the error term may also be different over time.

To account for the problem of serial correlation which causes the coefficient estimates to become inefficient and the standard errors to be misleading, the next step is to run the panel data as a fixed and random effects model.

## Model 2:

A random effects model is run next as it has more degrees of freedom than the fixed effect model. Below are the results from the random effects model:

```
. xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

Random-effects GLS regression                   Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.1665                                         min =         23
     between = 0.7632                                         avg =       23.0
     overall = 0.6967                                         max =         23

                                                Wald chi2(8)      =     399.18
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
         vio |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .3931534   .0397895     9.88   0.000     .3151674    .4711393
      pb1064 |   20.61077   7.196501     2.86   0.004     6.505887    34.71565
      pw1064 |   8.026554   3.093081     2.60   0.009     1.964226    14.08888
      pm1029 |  -9.640604   3.623266    -2.66   0.008    -16.74207   -2.539133
         pop |   13.15851    3.07476     4.28   0.000     7.132095    19.18493
      avginc |  -6.025157   3.514785    -1.71   0.086    -12.91401    .8636953
     density |   104.7166   15.97841     6.55   0.000     73.39954    136.0337
     1.shall |  -37.19344    11.6025    -3.21   0.001    -59.93392   -14.45296
       _cons |  -54.70189   231.1279    -0.24   0.813    -507.7042    398.3004
-------------+----------------------------------------------------------------
     sigma_u |  133.02093
     sigma_e |  98.505026
         rho |  .64583883   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

## Model 3:

The fixed effects model is run next. The results are as follows:

```
. xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.2038                                         min =         23
     between = 0.3446                                         avg =       23.0
     overall = 0.2658                                         max =         23

                                                F(8,1114)         =      35.64
corr(u_i, Xb)  = -0.8388                        Prob > F          =     0.0000

------------------------------------------------------------------------------
         vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0967879   .0573672     1.69   0.092     -.015772    .2093479
      pb1064 |   12.80803   10.88266     1.18   0.239     -8.54479    34.16085
      pw1064 |   10.32325   3.110094     3.32   0.001     4.220951    16.42556
      pm1029 |  -23.86385   3.924761    -6.08   0.000    -31.56461   -16.16309
         pop |   12.24406   5.346751     2.29   0.022     1.753224     22.7349
      avginc |  -4.066645   3.621108    -1.12   0.262    -11.17161    3.038316
     density |  -155.9533   52.11761    -2.99   0.003     -258.213   -53.69351
     1.shall |  -18.57508   11.56324    -1.61   0.108    -41.26327    4.113097
       _cons |   203.0109   235.8217     0.86   0.389    -259.6938    665.7157
-------------+----------------------------------------------------------------
     sigma_u |  500.37955
     sigma_e |  98.505026
         rho |  .96269175   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(50, 1114) = 52.30                   Prob > F = 0.0000
```

## Hausman Test:

Hausman's test is run next to test if the dependent variables of the random effects model are correlated with its error. Below are the results:

```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random        Difference          S.E.
-------------+----------------------------------------------------------------
 incarc_rate |    .0967879     .3931534       -.2963654        .0413254
      pb1064 |    12.80803     20.61077       -7.802738        8.163496
      pw1064 |    10.32325     8.026554          2.2967        .3248615
      pm1029 |   -23.86385    -9.640604       -14.22325        1.508542
         pop |    12.24406     13.15851       -.9144534        4.374196
      avginc |   -4.066645    -6.025157        1.958512        .8710399
     density |   -155.9533     104.7166       -260.6699        49.60782
     1.shall |   -18.57508    -37.19344        18.61836               .
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       64.49
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

The p-value is < 0.00. So, we reject the null hypothesis that the coefficients of the fixed effect and the random effects are the same. Hence, the fixed effects model is selected.

### Model 4:

The fixed effect model is then run, with the Cluster Robust Standard Errors to get more efficient coefficients.

```
. xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce
> (cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.2038                                         min =         23
     between = 0.3446                                         avg =       23.0
     overall = 0.2658                                         max =         23

                                                F(8,50)           =     265.92
corr(u_i, Xb)  = -0.8388                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
         vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0967879   .1567361     0.62   0.540    -.2180259    .4116017
      pb1064 |   12.80803   26.48154     0.48   0.631    -40.38171    65.99777
      pw1064 |   10.32325   6.148925     1.68   0.099    -2.027226    22.67373
      pm1029 |  -23.86385    9.45301    -2.52   0.015    -42.85078   -4.876919
         pop |   12.24406   9.561442     1.28   0.206     -6.96066    31.44878
      avginc |  -4.066645   6.276427    -0.65   0.520    -16.67322    8.539929
     density |  -155.9533   113.4665    -1.37   0.175    -383.8573    71.95084
     1.shall |  -18.57508   19.47905    -0.95   0.345    -57.69992    20.54975
       _cons |   203.0109   368.3606     0.55   0.584    -536.8631    942.8849
-------------+----------------------------------------------------------------
     sigma_u |  500.37955
     sigma_e |  98.505026
         rho |  .96269175   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

None of the variables except for the percent of males in the age group 10 to 29 (pm1029) are significant at a 95% confidence level.

### Model 5:

We then use log variables to account for the skewness observed in the histogram of the variables

```
. xtreg log_vio log_incrate pb1064 pw1064 pm1029 log_pop avginc density i.shall
> , fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.2217                                         min =         23
     between = 0.1131                                         avg =       23.0
     overall = 0.0607                                         max =         23

                                                F(8,50)           =      22.79
corr(u_i, Xb)  = -0.5221                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
     log_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 log_incrate |  -.0633165   .0669849    -0.95   0.349    -.1978596    .0712267
      pb1064 |   .1167926   .0320688     3.64   0.001     .0523805    .1812046
      pw1064 |   .0413543   .0143821     2.88   0.006     .0124671    .0702415
      pm1029 |  -.0643953   .0259147    -2.48   0.016    -.1164465   -.0123441
     log_pop |  -.1258555   .1625252    -0.77   0.442     -.452297     .200586
      avginc |  -.0065709   .0132301    -0.50   0.622    -.0331442    .0200025
     density |  -.1285301   .0788056    -1.63   0.109    -.2868158    .0297557
     1.shall |   -.037671   .0435654    -0.86   0.391    -.1251748    .0498327
       _cons |   4.446236   .8064566     5.51   0.000      2.82642    6.066051
-------------+----------------------------------------------------------------
     sigma_u |  .71725041
     sigma_e |  .16032097
         rho |   .9524155   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

We see that both percent of black people as well as white people in the age group 10 to 64 (pb1064 and pw1064) have become significant at the 95% confidence level.

### Model 6:

Dummy variables for each year is then added to check for time effects on the fixed effects model. The results are given below:

```
. xtreg log_vio log_incrate pb1064 pw1064 pm1029 log_pop avginc density i.shall
>  i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.4269                                         min =         23
     between = 0.4615                                         avg =       23.0
     overall = 0.2735                                         max =         23

                                                F(30,50)          =      69.24
corr(u_i, Xb)  = -0.7584                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 51 clusters in stateid)

------------------------------------------------------------------------------
             |               Robust
     log_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 log_incrate |  -.1042194   .0691802    -1.51   0.138     -.243172    .0347332
      pb1064 |   .0073456    .053356     0.14   0.891     -.099823    .1145142
      pw1064 |   .0000793   .0249287     0.00   0.997    -.0499914      .05015
      pm1029 |   .0804462   .0526424     1.53   0.133    -.0252892    .1861817
     log_pop |  -.1936537   .1932852    -1.00   0.321    -.5818785    .1945711
      avginc |  -.0005546   .0162047    -0.03   0.973    -.0331028    .0319935
     density |  -.1041343   .0796751    -1.31   0.197    -.2641666    .0558979
     1.shall |    -.02888   .0393632    -0.73   0.467    -.1079432    .0501833
             |
        year |
          78 |   .0674092   .0160657     4.20   0.000     .0351404     .099678
          79 |   .1846194   .0275775     6.69   0.000     .1292283    .2400104
          80 |   .2430197   .0399504     6.08   0.000      .162777    .3232624
          81 |   .2509862   .0433771     5.79   0.000     .1638608    .3381116
          82 |   .2438153   .0545544     4.47   0.000     .1342396     .353391
          83 |   .2229004   .0673075     3.31   0.002     .0877094    .3580915
          84 |   .2667187   .0822677     3.24   0.002     .1014791    .4319583
          85 |    .326206   .0976441     3.34   0.002     .1300821      .52233
          86 |   .4154293   .1142612     3.64   0.001     .1859289    .6449297
          87 |   .4249587   .1312944     3.24   0.002     .1612461    .6886713
          88 |   .4970081    .145989     3.40   0.001     .2037805    .7902356
          89 |   .5623485   .1604123     3.51   0.001     .2401509    .8845462
          90 |   .6952929   .2055371     3.38   0.001     .2824595    1.108126
          91 |   .7585728   .2150783     3.53   0.001     .3265754     1.19057
          92 |   .8009791   .2275746     3.52   0.001     .3438821    1.258076
          93 |   .8321949   .2375269     3.50   0.001     .3551081    1.309282
          94 |   .8278492   .2461974     3.36   0.001     .3333472    1.322351
          95 |   .8328327   .2576959     3.23   0.002     .3152352     1.35043
          96 |   .7876821   .2695923     2.92   0.005       .24619    1.329174
          97 |   .7763838   .2774862     2.80   0.007     .2190363    1.333731
          98 |   .7311231   .2891897     2.53   0.015     .1502684    1.311978
          99 |   .6810194   .2995348     2.27   0.027     .0793861    1.282653
             |
       _cons |   4.994805   1.318697     3.79   0.000     2.346123    7.643486
-------------+----------------------------------------------------------------
     sigma_u |   .8266663
     sigma_e |  .13895235
         rho |  .97252288   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

## F-Test:

An F-test is performed on the year dummy-variables, to assess if the fixed effect model is affected by the time component of the panel data.

```
. testparm i.year

 ( 1)  78.year = 0
 ( 2)  79.year = 0
 ( 3)  80.year = 0
 ( 4)  81.year = 0
 ( 5)  82.year = 0
 ( 6)  83.year = 0
 ( 7)  84.year = 0
 ( 8)  85.year = 0
 ( 9)  86.year = 0
 (10)  87.year = 0
 (11)  88.year = 0
 (12)  89.year = 0
 (13)  90.year = 0
 (14)  91.year = 0
 (15)  92.year = 0
 (16)  93.year = 0
 (17)  94.year = 0
 (18)  95.year = 0
 (19)  96.year = 0
 (20)  97.year = 0
 (21)  98.year = 0
 (22)  99.year = 0

       F( 22,    50) =   21.71
            Prob > F =    0.0000
```

The results show that the time effects do indeed affect the fixed effect model.

## Model 7:

To understand the effect of non-linear relationship of the independent variables on violent crime rates, interaction variables were added. The only significant interaction was shall*percentage of black population. The results are as follows:

```
. xtreg log_vio pw1064 avginc pm1029 log_pop c.log_incrate density i.shall##c.p
> b1064 i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.4727                                         min =         23
     between = 0.4730                                         avg =       23.0
     overall = 0.2895                                         max =         23

                                                F(31,50)          =      74.84
corr(u_i, Xb)  = -0.7825                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
```

| log_vio | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| pw1064 | .0023161 | .0247531 | 0.09 | 0.926 | -.047402 | .0520342 |
| avginc | -.0040728 | .0142901 | -0.29 | 0.777 | -.0327754 | .0246298 |
| pm1029 | .0982148 | .0517161 | 1.90 | 0.063 | -.00566 | .2020896 |
| log_pop | -.2238996 | .1834917 | -1.22 | 0.228 | -.5924535 | .1446543 |
| log_incrate | -.1014184 | .0658296 | -1.54 | 0.130 | -.233641 | .0308042 |
| density | -.0651606 | .0724121 | -0.90 | 0.373 | -.2106046 | .0802835 |
| 1.shall | -.2185981 | .0630807 | -3.47 | 0.001 | -.3452995 | -.0918967 |
| pb1064 | -.0123516 | .0545962 | -0.23 | 0.822 | -.1220113 | .0973081 |
| shall# | | | | | | |
| c.pb1064 | | | | | | |
| 1 | .0357479 | .008361 | 4.28 | 0.000 | .0189544 | .0525415 |
| year | | | | | | |
| 78 | .0733303 | .0168294 | 4.36 | 0.000 | .0395276 | .1071331 |
| 79 | .1945539 | .0289443 | 6.72 | 0.000 | .1364176 | .2526903 |
| 80 | .2567862 | .0413274 | 6.21 | 0.000 | .1737777 | .3397946 |
| 81 | .2682949 | .0445629 | 6.02 | 0.000 | .1787876 | .3578022 |
| 82 | .2700751 | .055902 | 4.83 | 0.000 | .1577927 | .3823575 |
| 83 | .2563013 | .0689346 | 3.72 | 0.001 | .1178421 | .3947605 |
| 84 | .3091804 | .0848483 | 3.64 | 0.001 | .1387576 | .4796032 |
| 85 | .376629 | .1008441 | 3.73 | 0.000 | .1740777 | .5791804 |
| 86 | .4784283 | .1167863 | 4.10 | 0.000 | .243856 | .7130006 |
| 87 | .4984314 | .1336427 | 3.73 | 0.000 | .230002 | .7668607 |
| 88 | .5783546 | .1486463 | 3.89 | 0.000 | .2797897 | .8769196 |
| 89 | .6510437 | .1630797 | 3.99 | 0.000 | .3234884 | .978599 |
| 90 | .7864282 | .2106893 | 3.73 | 0.000 | .3632462 | 1.20961 |
| 91 | .8550086 | .2203393 | 3.88 | 0.000 | .4124441 | 1.297573 |
| 92 | .9009866 | .2328843 | 3.87 | 0.000 | .4332248 | 1.368748 |
| 93 | .9362999 | .2428609 | 3.86 | 0.000 | .4484994 | 1.4241 |
| 94 | .9364301 | .2520807 | 3.71 | 0.001 | .4301112 | 1.442749 |
| 95 | .9452227 | .2658301 | 3.56 | 0.001 | .4112871 | 1.479158 |
| 96 | .8971862 | .2780709 | 3.23 | 0.002 | .3386643 | 1.455708 |
| 97 | .885901 | .2865608 | 3.09 | 0.003 | .3103267 | 1.461475 |
| 98 | .843688 | .2990441 | 2.82 | 0.007 | .2430402 | 1.444336 |
| 99 | .7955219 | .3104489 | 2.56 | 0.013 | .1719669 | 1.419077 |
| _cons | 4.671467 | 1.296198 | 3.60 | 0.001 | 2.067976 | 7.274958 |
| sigma_u | .8573313 | | | | | |
| sigma_e | .13334069 | | | | | |
| rho | .97638175 | (fraction of variance due to u_i) | | | | |

The adjusted R-squared value was calculated at .4584

The linear variable percentage of males is significant at the 10% significance level, while the variable shall and the interaction effect (shall*percentage of black population) along with the time periods from 1978 through 1999 were found to be significant at the 95% confidence level.

The interpretations of the coefficients are given below:

- As the years increase the violent crime rate increases.
- When a state has the shall law issued, there is an associated substantial decrease in violent crime rate.
- The percentage of males in the population has an increasing effect on the violent crime rate.
- When the shall law is implemented in a state, the percentage of black population on violent crime rate increases.

## Murder Rate

### Model 1:

We first run a Pooled Effect model with all the independent variables. The results are shown below:

```
. reg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

      Source |       SS           df       MS      Number of obs   =     1,173
-------------+----------------------------------   F(8, 1164)      =    523.14
       Model |  51892.2351         8  6486.52939   Prob > F        =    0.0000
    Residual |  14432.6089     1,164  12.3991486   R-squared       =    0.7824
-------------+----------------------------------   Adj R-squared   =    0.7809
       Total |   66324.844     1,172  56.5911638   Root MSE        =    3.5212

------------------------------------------------------------------------------
         mur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0241046   .0008822    27.32   0.000     .0223736    .0258355
      pb1064 |    .030323   .1370932     0.22   0.825    -.2386544    .2993004
      pw1064 |  -.0326635   .0689739    -0.47   0.636    -.1679905    .1026635
      pm1029 |   1.011478   .0887012    11.40   0.000     .8374463    1.185511
         pop |   .1661031   .0210672     7.88   0.000     .1247691     .207437
      avginc |  -.3346584   .0640557    -5.22   0.000    -.4603359   -.2089809
     density |   2.496043   .1084144    23.02   0.000     2.283334    2.708753
     1.shall |  -1.077353   .2681323    -4.02   0.000     -1.60343   -.5512763
       _cons |   -8.99181   4.473839    -2.01   0.045     -17.7695     -.21412
------------------------------------------------------------------------------
```

Our next step is to run the same model using the robust standard errors in order to account for the incorrect standard errors.

```
. reg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

Linear regression                               Number of obs     =      1,173
                                                F(8, 1164)        =     138.76
                                                Prob > F          =     0.0000
                                                R-squared         =     0.7824
                                                Root MSE          =     3.5212

------------------------------------------------------------------------------
             |               Robust
         mur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0241046   .0020404    11.81   0.000     .0201012    .0281079
      pb1064 |    .030323   .1491101     0.20   0.839    -.2622317    .3228777
      pw1064 |  -.0326635   .0695919    -0.47   0.639     -.169203    .1038761
      pm1029 |   1.011478   .1151811     8.78   0.000     .7854927    1.237464
         pop |   .1661031   .0263041     6.31   0.000     .1144942    .2177119
      avginc |  -.3346584    .070592    -4.74   0.000    -.4731603   -.1961565
     density |   2.496043   .3271928     7.63   0.000      1.85409    3.137997
     1.shall |  -1.077353   .2202453    -4.89   0.000    -1.509475   -.6452307
       _cons |   -8.99181   4.232699    -2.12   0.034    -17.29638    -.687236
------------------------------------------------------------------------------
```

Due to the time effect included in a panel data, a variable x in time t has an influence over xt+1. Therefore, the assumption that the error terms between different time periods are not correlated is violated. Additionally, the variance of the error term may also be different over time.

To account for the problem of serial correlation which causes the coefficient estimates to become inefficient and the standard errors to be misleading, the next step is to run the panel data as a fixed and random effects model.

Model 2:

A random effects model is run next as it has more degrees of freedom than the fixed effect model. Below are the results from the random effects model:

```
. xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

Random-effects GLS regression                   Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.1641                                         min =         23
     between = 0.8928                                         avg =       23.0
     overall = 0.7575                                         max =         23

                                                Wald chi2(8)      =     753.31
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
         mur |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0212448   .0010772    19.72   0.000     .0191335     .023356
      pb1064 |   -.382155   .1770604    -2.16   0.031     -.729187    -.035123
      pw1064 |  -.1868612   .0859997    -2.17   0.030    -.3554174   -.0183049
      pm1029 |   1.237827   .1009954    12.26   0.000     1.039879    1.435774
         pop |  -.0084263   .0508474    -0.17   0.868    -.1080853    .0912327
      avginc |   .1282011   .0921206     1.39   0.164     -.052352    .3087542
     density |   2.264481   .2402161     9.43   0.000     1.793666    2.735296
     1.shall |    -1.3641   .3260685    -4.18   0.000    -2.003183   -.7250177
       _cons |  -5.437536   6.152176    -0.88   0.377    -17.49558    6.620508
-------------+----------------------------------------------------------------
     sigma_u |  1.6370471
     sigma_e |  2.7074164
         rho |  .26772388   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

### Model 3:

The fixed effects model is run next. The results are as follows:

```
. xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.3204                                         min =         23
     between = 0.7745                                         avg =       23.0
     overall = 0.5752                                         max =         23

                                                F(8,1114)         =      65.66
corr(u_i, Xb)  = -0.9787                        Prob > F          =     0.0000

------------------------------------------------------------------------------
         mur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0079266   .0015767     5.03   0.000     .0048329    .0110203
      pb1064 |  -1.783213   .2991105    -5.96   0.000    -2.370096   -1.196329
      pw1064 |  -.0710444   .0854811    -0.83   0.406    -.2387666    .0966777
      pm1029 |   .5870231   .1078723     5.44   0.000     .3753673    .7986789
         pop |  -.5212885   .1469558    -3.55   0.000    -.8096298   -.2329472
      avginc |   .5839008   .0995264     5.87   0.000     .3886205     .779181
     density |  -6.825431   1.432456    -4.76   0.000    -9.636045   -4.014816
     1.shall |  -.6550583   .3178163    -2.06   0.040    -1.278644   -.0314723
       _cons |    7.47542   6.481573     1.15   0.249    -5.242047    20.19289
-------------+----------------------------------------------------------------
     sigma_u |  20.321614
     sigma_e |  2.7074164
         rho |  .98255975   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(50, 1114) = 17.10                   Prob > F = 0.0000
```

### Hausman Test:

Hausman's test is run next to test if the dependent variables of the random effects model are correlated with its error. Below are the results:

```
. hausman fixed2 random2

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed2       random2       Difference          S.E.
-------------+----------------------------------------------------------------
 incarc_rate |    .0079266     .0212448       -.0133182         .0011514
      pb1064 |   -1.783213     -.382155       -1.401058         .2410741
      pw1064 |   -.0710444    -.1868612        .1158167                .
      pm1029 |    .5870231     1.237827       -.6508035         .0378992
         pop |   -.5212885    -.0084263       -.5128622         .1378787
      avginc |    .5839008     .1282011        .4556996         .0376735
     density |   -6.825431     2.264481       -9.089911          1.41217
     1.shall |   -.6550583      -1.3641        .7090419                .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      313.25
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

The p-value is < 0.00. So, we reject the null hypothesis that the coefficients of the fixed effect and the random effects are the same. Hence, the fixed effects model is selected.

### Model 4:

The fixed effect model is then run, with the Cluster Robust Standard Errors to get more efficient coefficients.

```
. xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce
> (cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.3204                                         min =         23
     between = 0.7745                                         avg =       23.0
     overall = 0.5752                                         max =         23

                                                F(8,50)           =    2942.43
corr(u_i, Xb)  = -0.9787                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
         mur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0079266   .0112422     0.71   0.484    -.0146541    .0305072
      pb1064 |  -1.783213   2.063557    -0.86   0.392    -5.927989    2.361563
      pw1064 |  -.0710444   .2061573    -0.34   0.732    -.4851234    .3430346
      pm1029 |   .5870231   .4259051     1.38   0.174    -.2684324    1.442479
         pop |  -.5212885   .3999463    -1.30   0.198    -1.324604    .2820274
      avginc |   .5839008   .3060605     1.91   0.062    -.0308398    1.198641
     density |  -6.825431   10.92914    -0.62   0.535    -28.77726     15.1264
     1.shall |  -.6550583   .3940816    -1.66   0.103    -1.446595     .136478
       _cons |    7.47542   11.74442     0.64   0.527    -16.11394    31.06478
-------------+----------------------------------------------------------------
     sigma_u |  20.321614
     sigma_e |  2.7074164
         rho |  .98255975   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

None of the variables are significant at a 95% confidence level.

## Model 5:

We then use the log form of density to account for the skewness observed in the histogram.

```
. xtreg mur incarc_rate pb1064 pw1064 pm1029 log_pop avginc log_density i.shall
> , fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.3483                                         min =         23
     between = 0.1564                                         avg =       23.0
     overall = 0.1293                                         max =         23

                                                F(8,50)           =       9.55
corr(u_i, Xb)  = -0.9997                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
         mur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0087949   .0043522     2.02   0.049     .0000532    .0175367
      pb1064 |  -2.524701   1.136673    -2.22   0.031    -4.807776   -.2416258
      pw1064 |  -.0104977   .1470838    -0.07   0.943    -.3059243    .2849289
      pm1029 |   .4505844     .30753     1.47   0.149    -.1671078    1.068277
     log_pop |  -190.3064   46.89719    -4.06   0.000    -284.5022   -96.11065
      avginc |   .5805541   .2367686     2.45   0.018     .1049903    1.056118
 log_density |   184.6991   47.20772     3.91   0.000     89.87961    279.5186
     1.shall |  -.6888283   .3563621    -1.93   0.059    -1.404603    .0269461
       _cons |   680.3142   174.8007     3.89   0.000     329.2166    1031.412
-------------+----------------------------------------------------------------
     sigma_u |  267.28259
     sigma_e |  2.6513293
         rho |  .99990161   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The variables incarceration rate, percentage of black population, log of population, average income and log of density are significant at the 95% confidence level.

### Model 6:

Dummy variable for each year is then added to check for time effects on the fixed effect model. The results are given below:

```
. xtreg mur incarc_rate pb1064 pw1064 pm1029 log_pop avginc log_density i.shall
>  i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.4374                                         min =         23
     between = 0.1515                                         avg =       23.0
     overall = 0.1263                                         max =         23

                                                F(30,50)          =      54.30
corr(u_i, Xb)  = -0.9994                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
```

| mur | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| incarc_rate | .0105974 | .0043489 | 2.44 | 0.018 | .0018623 | .0193325 |
| pb1064 | -2.917727 | 1.876761 | -1.55 | 0.126 | -6.687311 | .8518581 |
| pw1064 | -.2606319 | .4428649 | -0.59 | 0.559 | -1.150152 | .6288884 |
| pm1029 | 1.18437 | 1.076745 | 1.10 | 0.277 | -.9783349 | 3.347075 |
| log_pop | -136.8291 | 44.31992 | -3.09 | 0.003 | -225.8483 | -47.8099 |
| avginc | .9581055 | .2575099 | 3.72 | 0.001 | .4408816 | 1.475329 |
| log_density | 131.2615 | 44.31893 | 2.96 | 0.005 | 42.24434 | 220.2787 |
| 1.shall | -.1771192 | .4208353 | -0.42 | 0.676 | -1.022392 | .6681535 |
| year | | | | | | |
| 78 | .0048033 | .2254295 | 0.02 | 0.983 | -.4479851 | .4575917 |
| 79 | .8002806 | .4146133 | 1.93 | 0.059 | -.0324947 | 1.633056 |
| 80 | 1.701343 | .7723262 | 2.20 | 0.032 | .15008 | 3.252606 |
| 81 | 1.716899 | 1.053163 | 1.63 | 0.109 | -.3984403 | 3.832239 |
| 82 | 1.533093 | 1.200636 | 1.28 | 0.208 | -.878456 | 3.944642 |
| 83 | .7634724 | 1.335948 | 0.57 | 0.570 | -1.919857 | 3.446802 |
| 84 | .0544027 | 1.488729 | 0.04 | 0.971 | -2.935797 | 3.044602 |
| 85 | .171753 | 1.605452 | 0.11 | 0.915 | -3.052893 | 3.396399 |
| 86 | .7795916 | 1.906351 | 0.41 | 0.684 | -3.049426 | 4.60861 |
| 87 | .6941002 | 2.224872 | 0.31 | 0.756 | -3.774686 | 5.162886 |
| 88 | 1.209253 | 2.834292 | 0.43 | 0.671 | -4.48359 | 6.902096 |
| 89 | 1.614317 | 3.210232 | 0.50 | 0.617 | -4.833623 | 8.062258 |
| 90 | 2.623762 | 4.173563 | 0.63 | 0.532 | -5.759085 | 11.00661 |
| 91 | 3.386994 | 4.507692 | 0.75 | 0.456 | -5.666972 | 12.44096 |
| 92 | 2.990373 | 4.567859 | 0.65 | 0.516 | -6.184443 | 12.16519 |
| 93 | 3.565938 | 4.787037 | 0.74 | 0.460 | -6.049109 | 13.18099 |
| 94 | 3.090143 | 4.759568 | 0.65 | 0.519 | -6.469732 | 12.65002 |
| 95 | 2.591256 | 4.762234 | 0.54 | 0.589 | -6.973972 | 12.15648 |
| 96 | 2.201133 | 5.058536 | 0.44 | 0.665 | -7.959235 | 12.3615 |
| 97 | 1.282473 | 4.88219 | 0.26 | 0.794 | -8.523695 | 11.08864 |
| 98 | .3123943 | 4.86412 | 0.06 | 0.949 | -9.457477 | 10.08227 |
| 99 | -.3849733 | 4.887195 | -0.08 | 0.938 | -10.20119 | 9.431247 |
| _cons | 485.4728 | 168.0552 | 2.89 | 0.006 | 147.9239 | 823.0217 |
| sigma_u | 188.98368 | | | | | |
| sigma_e | 2.4882256 | | | | | |
| rho | .99982668 | (fraction of variance due to u_i) | | | | |

The percentage of black population became insignificant after including the time effect.

## F-Test:

An F-test is performed on the year dummy-variables, to assess if the fixed effect model is affected by the time component of the panel data.

```
. testparm i.year

 ( 1)  78.year = 0
 ( 2)  79.year = 0
 ( 3)  80.year = 0
 ( 4)  81.year = 0
 ( 5)  82.year = 0
 ( 6)  83.year = 0
 ( 7)  84.year = 0
 ( 8)  85.year = 0
 ( 9)  86.year = 0
 (10)  87.year = 0
 (11)  88.year = 0
 (12)  89.year = 0
 (13)  90.year = 0
 (14)  91.year = 0
 (15)  92.year = 0
 (16)  93.year = 0
 (17)  94.year = 0
 (18)  95.year = 0
 (19)  96.year = 0
 (20)  97.year = 0
 (21)  98.year = 0
 (22)  99.year = 0

       F( 22,    50) =    8.49
            Prob > F =    0.0000
```

The results show that the time effects do indeed affect the fixed effect model.

### Model 7:

To understand the effect of non-linear relationship of the independent variables on murder rate, many combinations of interactions variables were included. The most meaningful interactions were found to be log of population*incarceration rate and log of density*percentage of black population. The results are as follows:

```
. xtreg mur pw1064 pm1029 c.log_pop##c.incarc_rate avginc c.log_density##c.pb10
> 64 i.shall i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.5224                                         min =         23
     between = 0.1432                                         avg =       23.0
     overall = 0.1200                                         max =         23

                                                F(32,50)          =     374.15
corr(u_i, Xb)  = -0.9993                        Prob > F          =     0.0000

                             (Std. Err. adjusted for 51 clusters in stateid)
```

| mur | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| pw1064 | -.7906742 | .3973302 | -1.99 | 0.052 | -1.588735 | .007387 |
| pm1029 | 1.580729 | 1.071888 | 1.47 | 0.147 | -.5722208 | 3.733678 |
| log_pop | -131.4554 | 37.50356 | -3.51 | 0.001 | -206.7836 | -56.12732 |
| incarc_rate | .0038021 | .0021354 | 1.78 | 0.081 | -.0004869 | .0080912 |
| c.log_pop# c. incarc_rate | -.003262 | .0012722 | -2.56 | 0.013 | -.0058172 | -.0007067 |
| avginc | 1.045786 | .1746025 | 5.99 | 0.000 | .6950868 | 1.396486 |
| log_density | 128.6868 | 37.93322 | 3.39 | 0.001 | 52.49571 | 204.8779 |
| pb1064 | -3.976124 | 1.339472 | -2.97 | 0.005 | -6.666532 | -1.285715 |
| c. log_density# c.pb1064 | -.8507762 | .1370476 | -6.21 | 0.000 | -1.126044 | -.5755079 |
| 1.shall | -.2059363 | .281548 | -0.73 | 0.468 | -.7714422 | .3595696 |

| | | | | | | |
|---|---|---|---|---|---|---|
| year | | | | | | |
| 78 | .0009581 | .1846999 | 0.01 | 0.996 | -.3700225 | .3719388 |
| 79 | .8534713 | .3612955 | 2.36 | 0.022 | .127788 | 1.579155 |
| 80 | 1.682538 | .6939248 | 2.42 | 0.019 | .2887491 | 3.076327 |
| 81 | 1.770138 | .9067815 | 1.95 | 0.057 | -.0511859 | 3.591463 |
| 82 | 1.694621 | 1.0194 | 1.66 | 0.103 | -.3529034 | 3.742146 |
| 83 | 1.021242 | 1.206436 | 0.85 | 0.401 | -1.401955 | 3.444439 |
| 84 | .3052377 | 1.368936 | 0.22 | 0.824 | -2.44435 | 3.054826 |
| 85 | .4306723 | 1.503104 | 0.29 | 0.776 | -2.588401 | 3.449745 |
| 86 | 1.048517 | 1.851297 | 0.57 | 0.574 | -2.669923 | 4.766957 |
| 87 | .9959423 | 2.180739 | 0.46 | 0.650 | -3.3842 | 5.376085 |
| 88 | 1.537691 | 2.850245 | 0.54 | 0.592 | -4.187194 | 7.262575 |
| 89 | 1.970244 | 3.260389 | 0.60 | 0.548 | -4.57844 | 8.518928 |
| 90 | 4.13956 | 4.154494 | 1.00 | 0.324 | -4.204987 | 12.48411 |
| 91 | 5.093141 | 4.462665 | 1.14 | 0.259 | -3.870385 | 14.05667 |
| 92 | 4.811734 | 4.525828 | 1.06 | 0.293 | -4.27866 | 13.90213 |
| 93 | 5.481632 | 4.727054 | 1.16 | 0.252 | -4.012935 | 14.9762 |
| 94 | 5.149108 | 4.714888 | 1.09 | 0.280 | -4.321023 | 14.61924 |
| 95 | 4.906196 | 4.726352 | 1.04 | 0.304 | -4.586961 | 14.39935 |
| 96 | 4.650774 | 4.991216 | 0.93 | 0.356 | -5.374378 | 14.67593 |
| 97 | 3.830479 | 4.798286 | 0.80 | 0.428 | -5.807161 | 13.46812 |
| 98 | 2.996614 | 4.769789 | 0.63 | 0.533 | -6.583789 | 12.57702 |
| 99 | 2.429816 | 4.800178 | 0.51 | 0.615 | -7.211624 | 12.07126 |
| _cons | 497.0006 | 141.8224 | 3.50 | 0.001 | 212.1419 | 781.8593 |
| sigma_u | 177.07854 | | | | | |
| sigma_e | 2.2945713 | | | | | |
| rho | .99983212 | (fraction of variance due to u_i) | | | | |

The adjusted R-squared value was found to be .509.

It was found that the independent variables incarceration rate and percentage of white population was significant at the 10% significance level, while log of population, average income, log of density and percentage of black population were significant at the 95% confidence level. Additionally, the interaction effects log of population*incarceration rate and log of density*percentage of black population were also significant. The time dummy variables were only significant for the year 1979 and 1980 at the 5% significant level and 1981 at the 10% significance level.

The interpretations of the coefficients are below:

- Increase in percentage of white population was inversely related with murder rate.
- Increase in incarceration rate had a very negligible negative effect on murder rate.
- When population increased by 1%, there was an associated 1.31 unit decrease in murder rate.
- Increase in average income leads to increase in murder rate by 1.04 units.

- More densely populated areas lead to an increase in murder rate. In specific, 1% increase in log of density is associated with a 1.28 unit increase in murder rate.
- Increase in percentage of black population by 1 unit leads to decrease in murder rate by 3.97 units.

# Robbery

## Model 1

We first run a Pooled Effect model with all the independent variables. The results are shown below:

```
. reg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

      Source |       SS           df       MS      Number of obs   =     1,173
-------------+----------------------------------   F(8, 1164)      =    532.09
       Model |  26757434.5         8  3344679.31   Prob > F        =    0.0000
    Residual |  7316879.52     1,164  6285.97897   R-squared       =    0.7853
-------------+----------------------------------   Adj R-squared   =    0.7838
       Total |    34074314     1,172  29073.6468   Root MSE        =    79.284

------------------------------------------------------------------------------
         rob |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0862563   .0198641     4.34   0.000     .0472829    .1252298
      pb1064 |   7.243095   3.086785     2.35   0.019      1.18681    13.29938
      pw1064 |   1.502237   1.553013     0.97   0.334    -1.544781    4.549255
      pm1029 |   3.501497   1.997193     1.75   0.080    -.4170035    7.419998
         pop |   10.98124   .4743482    23.15   0.000     10.05056    11.91191
      avginc |   3.414262   1.442275     2.37   0.018     .5845116    6.244012
     density |   84.22337   2.441055    34.50   0.000       79.434    89.01273
     1.shall |   -27.4472   6.037257    -4.55   0.000    -39.29232   -15.60207
       _cons |   -169.972   100.7328    -1.69   0.092    -367.6101    27.66618
------------------------------------------------------------------------------
```

Our next step is to run the same model using the robust standard errors in order to account for the incorrect standard errors.

```
. reg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

Linear regression                               Number of obs     =      1,173
                                                F(8, 1164)        =     122.35
                                                Prob > F          =     0.0000
                                                R-squared         =     0.7853
                                                Root MSE          =     79.284

------------------------------------------------------------------------------
             |               Robust
         rob |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |   .0862563   .0434361     1.99   0.047     .0010344    .1714782
      pb1064 |   7.243095   3.584797     2.02   0.044     .2097078    14.27648
      pw1064 |   1.502237   1.742477     0.86   0.389    -1.916509    4.920983
      pm1029 |   3.501497   2.024115     1.73   0.084    -.4698241    7.472819
         pop |   10.98124   .7943054    13.82   0.000     9.422807    12.53967
      avginc |   3.414262   1.486466     2.30   0.022     .4978086    6.330715
     density |   84.22337   6.556728    12.85   0.000     71.35904    97.08769
     1.shall |   -27.4472   4.909143    -5.59   0.000    -37.07896   -17.81544
       _cons |   -169.972   114.7728    -1.48   0.139    -395.1567     55.2128
------------------------------------------------------------------------------
```

Due to the time effect included in a panel data, a variable x in time t has an influence over xt+1. Therefore, the assumption that the error terms between different time periods are not correlated is violated. Additionally, the variance of the error term may also be different over time.

To account for the problem of serial correlation which causes the coefficient estimates to become inefficient and the standard errors to be misleading, the next step is to run the panel data as a fixed and random effects model.

## Model 2:

A random effects model is run next as it has more degrees of freedom than the fixed effect model. Below are the results from the random effects model:

```
. xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

Random-effects GLS regression                   Number of obs      =     1,173
Group variable: stateid                         Number of groups   =        51

R-sq:                                           Obs per group:
     within  = 0.0377                                         min =        23
     between = 0.8239                                         avg =      23.0
     overall = 0.7575                                         max =        23

                                                Wald chi2(8)       =    400.64
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
         rob |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |  -.0137862   .0190538    -0.72   0.469    -.0511311    .0235586
      pb1064 |   15.96248   3.326353     4.80   0.000     9.442944    22.48201
      pw1064 |   5.346849   1.502161     3.56   0.000     2.402668     8.29103
      pm1029 |  -1.905452   1.754347    -1.09   0.277    -5.343908    1.533005
         pop |   9.033511   1.288485     7.01   0.000     6.508126     11.5589
      avginc |  -5.937499   1.686247    -3.52   0.000    -9.242482   -2.632515
     density |   87.40912   6.398136    13.66   0.000       74.869    99.94924
     1.shall |  -3.250151   5.646313    -0.58   0.565    -14.31672     7.81642
       _cons |  -218.1524   111.0545    -1.96   0.049    -435.8153   -.489552
-------------+----------------------------------------------------------------
     sigma_u |  51.305528
     sigma_e |  48.244039
         rho |  .53072435   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

## Model 3:

The fixed effects model is run next. The results are as follows:

```
. xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.0739                                         min =         23
     between = 0.3313                                         avg =       23.0
     overall = 0.2849                                         max =         23

                                                F(8,1114)         =      11.12
corr(u_i, Xb)  = -0.7981                        Prob > F          =     0.0000

------------------------------------------------------------------------------
         rob |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |  -.1415018   .0280963    -5.04   0.000    -.1966294   -.0863742
      pb1064 |   23.48345   5.329915     4.41   0.000     13.02564    33.94125
      pw1064 |   5.610663   1.523207     3.68   0.000     2.621986     8.59934
      pm1029 |  -7.975287     1.9222    -4.15   0.000    -11.74683   -4.203747
         pop |   3.223055   2.618636     1.23   0.219    -1.914961     8.36107
      avginc |  -7.471394   1.773482    -4.21   0.000    -10.95113   -3.991652
     density |   -62.3304   25.52524    -2.44   0.015    -112.4134   -12.24745
     1.shall |   7.899847   5.663237     1.39   0.163    -3.211967    19.01166
       _cons |   -49.3018   115.4966    -0.43   0.670    -275.9171    177.3135
-------------+----------------------------------------------------------------
     sigma_u |  228.38104
     sigma_e |  48.244039
         rho |  .95728236   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(50, 1114) = 40.59                   Prob > F = 0.0000
```

### Hausman Test:

Hausman's test is run next to test if the dependent variables of the random effects model are correlated with its error. Below are the results:

```
. hausman fixed3 random3

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed3      random3       Difference          S.E.
-------------+----------------------------------------------------------------
 incarc_rate |   -.1415018    -.0137862       -.1277156        .0206483
      pb1064 |    23.48345     15.96248         7.52097        4.164537
      pw1064 |    5.610663     5.346849        .2638138        .2523314
      pm1029 |   -7.975287    -1.905452       -6.069836        .7855694
         pop |    3.223055     9.033511       -5.810456        2.279707
      avginc |   -7.471394    -5.937499       -1.533895        .5493716
     density |    -62.3304     87.40912       -149.7395        24.71035
     1.shall |    7.899847    -3.250151           11.15        .4374973
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       65.95
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

The p-value is < 0.00. So, we reject the null hypothesis that the coefficients of the fixed effect and the random effects are the same. Hence, the fixed effects model is selected.

### Model 4:

The fixed effect model is then run, with the Cluster Robust Standard Errors to get more efficient coefficients.

```
. xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce
> (cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.0739                                         min =         23
     between = 0.3313                                         avg =       23.0
     overall = 0.2849                                         max =         23

                                                F(8,50)           =      94.12
corr(u_i, Xb)  = -0.7981                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
         rob |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 incarc_rate |  -.1415018   .0478384    -2.96   0.005     -.237588   -.0454156
      pb1064 |   23.48345   11.62043     2.02   0.049     .1431261    46.82377
      pw1064 |   5.610663   2.346839     2.39   0.021     .8968973    10.32443
      pm1029 |  -7.975287   2.951379    -2.70   0.009    -13.90331   -2.047268
         pop |   3.223055   4.901349     0.66   0.514    -6.621595     13.0677
      avginc |  -7.471394   3.036872    -2.46   0.017    -13.57113   -1.371656
     density |   -62.3304    47.5462    -1.31   0.196    -157.8298    33.16895
     1.shall |   7.899847    8.25673     0.96   0.343    -8.684282    24.48398
       _cons |   -49.3018    147.181    -0.33   0.739    -344.9236      246.32
-------------+----------------------------------------------------------------
     sigma_u |  228.38104
     sigma_e |  48.244039
         rho |  .95728236   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The variables incarceration rate, percentage of black, white and male populations and average income are significant at the 95% confidence level.

### Model 5:

Dummy variable for each year is then added to check for time effects on the fixed effect model. The results are given below:

```
. xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall i.year,
>  fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.2017                                         min =         23
     between = 0.3673                                         avg =       23.0
     overall = 0.2358                                         max =         23

                                                F(30,50)          =     187.06
corr(u_i, Xb)  = -0.6995                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
```

| rob | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| incarc_rate | -.1155623 | .0709902 | -1.63 | 0.110 | -.2581504 | .0270257 |
| pb1064 | 4.915188 | 10.18847 | 0.48 | 0.632 | -15.54897 | 25.37934 |
| pw1064 | -1.7858 | 7.179589 | -0.25 | 0.805 | -16.20643 | 12.63483 |
| pm1029 | 11.91268 | 11.56907 | 1.03 | 0.308 | -11.32449 | 35.14985 |
| pop | -.0267939 | 5.157911 | -0.01 | 0.996 | -10.38676 | 10.33318 |
| avginc | -2.278894 | 3.263392 | -0.70 | 0.488 | -8.83361 | 4.275823 |
| density | -34.25171 | 64.47687 | -0.53 | 0.598 | -163.7573 | 95.25391 |
| 1.shall | 14.47587 | 8.592869 | 1.68 | 0.098 | -2.783415 | 31.73515 |
| year | | | | | | |
| 78 | 7.657708 | 3.037794 | 2.52 | 0.015 | 1.556119 | 13.7593 |
| 79 | 26.01776 | 6.300022 | 4.13 | 0.000 | 13.36379 | 38.67173 |
| 80 | 49.30214 | 12.6902 | 3.89 | 0.000 | 23.81313 | 74.79115 |
| 81 | 59.80562 | 18.64722 | 3.21 | 0.002 | 22.35158 | 97.25966 |
| 82 | 46.94134 | 17.04638 | 2.75 | 0.008 | 12.70267 | 81.18001 |
| 83 | 32.83726 | 16.70774 | 1.97 | 0.055 | -.7212213 | 66.39574 |
| 84 | 26.47968 | 16.73897 | 1.58 | 0.120 | -7.141527 | 60.10089 |
| 85 | 31.14552 | 18.58551 | 1.68 | 0.100 | -6.184572 | 68.47561 |
| 86 | 43.73714 | 21.28271 | 2.06 | 0.045 | .9895713 | 86.48471 |
| 87 | 39.66934 | 24.01913 | 1.65 | 0.105 | -8.574506 | 87.91318 |
| 88 | 51.9991 | 30.09511 | 1.73 | 0.090 | -8.448699 | 112.4469 |
| 89 | 65.97257 | 36.3295 | 1.82 | 0.075 | -6.997368 | 138.9425 |
| 90 | 87.54741 | 52.17075 | 1.68 | 0.100 | -17.24063 | 192.3355 |
| 91 | 103.083 | 54.09403 | 1.91 | 0.062 | -5.568071 | 211.734 |
| 92 | 103.0294 | 57.92358 | 1.78 | 0.081 | -13.31355 | 219.3723 |
| 93 | 105.0671 | 59.52927 | 1.76 | 0.084 | -14.50095 | 224.6352 |
| 94 | 101.728 | 60.86035 | 1.67 | 0.101 | -20.51366 | 223.9696 |
| 95 | 102.0039 | 66.22315 | 1.54 | 0.130 | -31.00922 | 235.017 |
| 96 | 91.80394 | 67.41183 | 1.36 | 0.179 | -43.5967 | 227.2046 |
| 97 | 79.40119 | 64.08699 | 1.24 | 0.221 | -49.32132 | 208.1237 |
| 98 | 66.65684 | 65.16497 | 1.02 | 0.311 | -64.23085 | 197.5445 |
| 99 | 59.42903 | 68.02213 | 0.87 | 0.386 | -77.19744 | 196.0555 |
| _cons | 62.50443 | 359.6663 | 0.17 | 0.863 | -659.9066 | 784.9154 |
| sigma_u | 201.34205 | | | | | |
| sigma_e | 45.242494 | | | | | |
| rho | .95193474 | (fraction of variance due to u_i) | | | | |

## F-Test:

An F-test is performed on the year dummy-variables, to assess if the fixed effect model is affected by the time component of the panel data.

```
. testparm i.year

 ( 1)  78.year = 0
 ( 2)  79.year = 0
 ( 3)  80.year = 0
 ( 4)  81.year = 0
 ( 5)  82.year = 0
 ( 6)  83.year = 0
 ( 7)  84.year = 0
 ( 8)  85.year = 0
 ( 9)  86.year = 0
 (10)  87.year = 0
 (11)  88.year = 0
 (12)  89.year = 0
 (13)  90.year = 0
 (14)  91.year = 0
 (15)  92.year = 0
 (16)  93.year = 0
 (17)  94.year = 0
 (18)  95.year = 0
 (19)  96.year = 0
 (20)  97.year = 0
 (21)  98.year = 0
 (22)  99.year = 0

       F( 22,    50) =    4.53
            Prob > F =    0.0000
```

The results show that the time effects do indeed affect the fixed effect model.

### Model 6:

To understand the effect of non-linear relationship of the independent variables on violent crime rates, various interaction variables were added. While most of the interaction effects were found to be insignificant, the best model was found to include the interaction effects population*percentage of black population, average income*density and the density*density. The output is as follows:

```
. xtreg rob pw1064 pm1029 c.pop##c.pb1064 c.avginc##c.density c.density#c.densi
> ty i.shall incarc_rate c.incarc_rate#c.pb1064 i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-sq:                                           Obs per group:
     within  = 0.3288                                         min =         23
     between = 0.7158                                         avg =       23.0
     overall = 0.6664                                         max =         23

                                                F(34,50)          =  130109.72
corr(u_i, Xb)  = -0.9881                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 51 clusters in stateid)
```

| rob | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| pw1064 | .6210708 | 5.95493 | 0.10 | 0.917 | -11.33976 | 12.5819 |
| pm1029 | 18.31693 | 11.17369 | 1.64 | 0.107 | -4.126098 | 40.75995 |
| pop | 27.85932 | 8.364489 | 3.33 | 0.002 | 11.05875 | 44.65989 |
| pb1064 | 20.60148 | 14.79236 | 1.39 | 0.170 | -9.109858 | 50.31282 |
| c.pop# c.pb1064 | -3.368779 | .832841 | -4.04 | 0.000 | -5.04159 | -1.695969 |
| avginc | -1.223457 | 2.983269 | -0.41 | 0.683 | -7.21553 | 4.768615 |
| density | 1134.294 | 390.71 | 2.90 | 0.005 | 349.5302 | 1919.058 |
| c.avginc# c.density | -7.888837 | 1.043358 | -7.56 | 0.000 | -9.984483 | -5.793192 |
| c.density# c.density | -54.54341 | 19.85128 | -2.75 | 0.008 | -94.41589 | -14.67093 |
| 1.shall | 3.055642 | 6.625512 | 0.46 | 0.647 | -10.25209 | 16.36338 |
| incarc_rate | -.2364494 | .0729753 | -3.24 | 0.002 | -.3830247 | -.0898742 |
| c. incarc_rate# c.pb1064 | .0245751 | .0048652 | 5.05 | 0.000 | .0148031 | .0343472 |

| | | | | | | |
|---|---|---|---|---|---|---|
| year | | | | | | |
| 78 | 6.78982 | 2.685558 | 2.53 | 0.015 | 1.395719 | 12.18392 |
| 79 | 22.82477 | 4.87604 | 4.68 | 0.000 | 13.03096 | 32.61859 |
| 80 | 43.66056 | 9.468083 | 4.61 | 0.000 | 24.64335 | 62.67776 |
| 81 | 54.42022 | 14.3207 | 3.80 | 0.000 | 25.65624 | 83.18419 |
| 82 | 43.35435 | 12.89936 | 3.36 | 0.001 | 17.44523 | 69.26347 |
| 83 | 30.60543 | 12.97822 | 2.36 | 0.022 | 4.537915 | 56.67295 |
| 84 | 26.96907 | 14.54208 | 1.85 | 0.070 | -2.239549 | 56.1777 |
| 85 | 32.99647 | 16.96872 | 1.94 | 0.057 | -1.086206 | 67.07914 |
| 86 | 47.90789 | 20.01846 | 2.39 | 0.020 | 7.699634 | 88.11614 |
| 87 | 45.95395 | 22.59815 | 2.03 | 0.047 | .564242 | 91.34367 |
| 88 | 59.90605 | 27.24752 | 2.20 | 0.033 | 5.177796 | 114.6343 |
| 89 | 74.14719 | 32.23669 | 2.30 | 0.026 | 9.397897 | 138.8965 |
| 90 | 92.1776 | 44.8253 | 2.06 | 0.045 | 2.143344 | 182.2119 |
| 91 | 108.2644 | 47.07038 | 2.30 | 0.026 | 13.72079 | 202.8081 |
| 92 | 108.9246 | 50.3186 | 2.16 | 0.035 | 7.856742 | 209.9925 |
| 93 | 110.8022 | 51.98897 | 2.13 | 0.038 | 6.379286 | 215.2251 |
| 94 | 104.5985 | 51.56055 | 2.03 | 0.048 | 1.036096 | 208.1609 |
| 95 | 102.8208 | 53.74936 | 1.91 | 0.061 | -5.138012 | 210.7795 |
| 96 | 94.08673 | 55.82316 | 1.69 | 0.098 | -18.03739 | 206.2108 |
| 97 | 82.50341 | 54.21519 | 1.52 | 0.134 | -26.39101 | 191.3978 |
| 98 | 68.25523 | 55.03791 | 1.24 | 0.221 | -42.29166 | 178.8021 |
| 99 | 58.39949 | 56.45517 | 1.03 | 0.306 | -54.99405 | 171.793 |
| _cons | -602.5247 | 302.2424 | -1.99 | 0.052 | -1209.596 | 4.547086 |
| sigma_u | 590.70661 | | | | | |
| sigma_e | 41.560938 | | | | | |
| rho | .99507414 | (fraction of variance due to u_i) | | | | |

The adjusted R-squared value was found to be .3087.

We find that the independent variables density, population and incarceration rate are significant at the 95% confidence level along with the interaction effects population*percentage of black population, average income*density and density*density. The time dummy variables for the years 1978 to 1996 are significant at the 90% confidence level.

The interpretations of the coefficients are below:

- Increase in density is associated with a substantial increase in robbery rate.
- Increase in population is also associated with an increase in robbery rate.
- When more people are incarcerated, there is a decline in the robbery rate.
- In a state with more population, if the percentage of black population increases, the rate at which robbery rate increases is diminishing. For a state with lower population, when the percentage of black people increases, increase in robberies is more compared to a state with a high population.
- When the average income of a densely populated area is high, robbery rate declines.

- As the population density of a state increases, the marginal rate at which it affects the increase in robbery rate is diminishing.

## Conclusion:

The pooled OLS model led to inflated estimators due to endogeneity, which was caused by the presence of individual state and time effects. These limitations were controlled for by using the state fixed and time fixed model. The state fixed aspect took into account each of their unique heterogeneity, whereas the time fixed aspect took into account effects that vary across time but not across states, like federal policies that affected all states.

There was also the limitation of limited data. With more data, even more variance could have been accounted for leading to more confident interpretations of the coefficients of the independent variables.

Since many of the features could not strongly explain the different types of crime rates, there exists a possibility of missing variables. Adding additional variables like unemployment rates, or a proxy for cultural attitude towards usage of guns may help engender a better model.

Based on the findings from the models created for the different crime rates, we find that the shall-issue law and the incarceration rate have no significant effect on murder rates and violent crime rates. We do find that shall-issue law significantly affects robbery rate at a 95% confidence level, by reducing it by 29.3 units when the law is adopted.

# Appendix

## Stata Code

```
**Basic Analysis

summarize vio mur rob incarc_rate pb1064 pw1064 pm1029 pop avginc density shall

xtset stateid year

xtsum vio mur rob incarc_rate pb1064 pw1064 pm1029 pop avginc density shall

corr vio mur rob incarc_rate pb1064 pw1064 pm1029 pop avginc density shall


**Dummy variables for time

summarize i.year

testparm i.year


*Log variables

histogram vio

gen log_vio = log(vio)

histogram log_vio

histogram mur

gen log_mur = log(mur)

histogram log_mur

histogram rob

gen log_rob = log(rob)

histogram log_rob

histogram pop

gen log_pop = log(pop)

histogram log_pop

histogram density

gen log_density = log(density)

histogram log_density
```

```
histogram incarc_rate

gen log_incrate = log(incarc_rate)

histogram log_incrate


**Violent Crime Rate

*Pooled OLS

reg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

reg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

*Random Effects Model

xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

estimates store random

*Fixed-Effects Model

xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

estimates store fixed

*Hausman Test

hausman fixed random

*Robust Standard Errors

xtreg vio incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce(cluster stateid)

xtreg log_vio log_incrate pb1064 pw1064 pm1029 log_pop avginc density i.shall, fe vce(cluster stateid)

*Dummy variables for time

xtreg log_vio log_incrate pb1064 pw1064 pm1029 log_pop avginc density i.shall i.year, fe vce(cluster
stateid)

*Interaction Variables

xtreg log_vio pw1064 avginc pm1029 log_pop c.log_incrate density i.shall##c.pb1064 i.year, fe
vce(cluster stateid)

ereturn list


**Murder Rate

*Pooled OLS
```

```
reg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

reg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

*Random Effects Model

xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

estimates store random2

*Fixed-Effects Model

xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

estimates store fixed2

*Hausman Test

hausman fixed2 random2

*Robust Standard Errors

xtreg mur incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce(cluster stateid)

xtreg mur incarc_rate pb1064 pw1064 pm1029 log_pop avginc log_density i.shall, fe vce(cluster stateid)

*Dummy variables for time

xtreg mur incarc_rate pb1064 pw1064 pm1029 log_pop avginc log_density i.shall i.year, fe vce(cluster
stateid)

*Interaction Variables

xtreg mur pw1064 pm1029 c.log_pop##c.incarc_rate avginc c.log_density##c.pb1064 i.shall i.year, fe
vce(cluster stateid)

ereturn list


**Robbery Rate

*Pooled OLS

reg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall

reg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, robust

*Random Effects Model

xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, re

estimates store random3

*Fixed-Effects Model
```

```
xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe

estimates store fixed3

*Hausman Test

hausman fixed3 random3

*Robust Standard Errors

xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall, fe vce(cluster stateid)

*Dummy variables for time

xtreg rob incarc_rate pb1064 pw1064 pm1029 pop avginc density i.shall i.year, fe vce(cluster stateid)

*Interaction Variables

xtreg rob pw1064 pm1029 c.pop##c.pb1064 c.avginc##c.density c.density#c.density i.shall incarc_rate
c.incarc_rate#c.pb1064 i.year, fe vce(cluster stateid)

ereturn list
```