“Do shall-issues law reduce crime-or not?”

BUAN 6312.003

Applied Econometrics and Time Series Analysis

Fall 2019

- **Kamini Bokefode**
- **Shylaja Vijayaraghavan**
- **Priyanka Savant**
- **Himaja Barla**
- **Varadharajan Hayagreeva Srinivasan**

Table of Contents

1. Introduction

Concealed gun laws have been a very important part of the gun control discussion. Many have challenged and defended the right to carry a firearm concealed on or near the person. The issue of gun control has been and continues to be an important issue in United States history.

states with "shall issue" systems require a license or permit to carry a concealed handgun, and applicants must meet certain well-defined objective criteria. However, unlike "may issue" systems, a "shall issue" state removes all arbitrary bias and discretion, compelling the issuing authority to award the permit. These laws require that the empowered authority “shall issue” a permit to applicants who meet the criteria defined by law.

Generally, the criteria for issuance of a license include proof of residency within the state, a minimum age, fingerprints for a background check, no record of mental illness or adjudication of mental defect by a court, proof or certification from an acceptable handgun safety class (including live-fire range qualification exercises to demonstrate safe and acceptable proficiency), and submitting the required application fee. Ohio is an example of a state with a "shall issue" system of licensing. The details of the requirements differ from state to state.

2. Exploratory Data Analysis:

a. Data Description:

The given dataset consists of balanced panel data for 51 US states (including the District of Columbia), from 1977 to 1999. The following is the detailed description of all the variables:

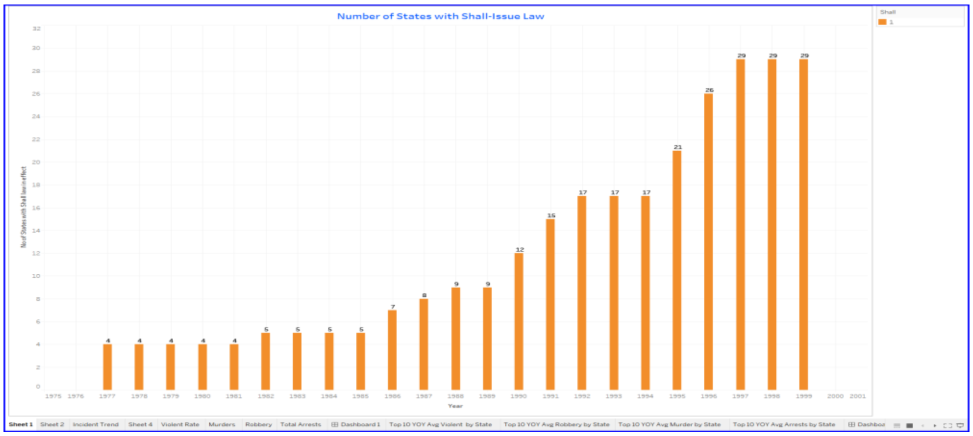
Variable	Definition
vio	violent crime rate (incidents per 100,000 members of the population)
rob	robbery rate (incidents per 100,000)
mur	murder rate (incidents per 100,000)
shall	= 1 if the state has a shall-carry law in effect in that year = 0 otherwise
incarc_rate	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
density	population per square mile of land area, divided by 1000
avginc	real per capita personal income in the state, in thousands of dollars
pop	state population, in millions of people
pm1029	percent of state population that is male, ages 10 to 29
pw1064	percent of state population that is white, ages 10 to 64
pb1064	percent of state population that is black, ages 10 to 64
stateid	ID number of states (Alabama = 1, Alaska = 2, etc.)
year	Year (1977-1999)

State id's with respect to the States are as follows:

Stateid	Abb	State	Stateid	Abb	State	Stateid	Abb	State
1	**AL**	Alabama	18	**LA**	Louisiana	35	**OH**	Ohio
2	**AK**	Alaska	19	**ME**	Maine	36	**OK**	Oklahoma
3	**AZ**	Arizona	20	**MD**	Maryland	37	**OR**	Oregon
4	**AR**	Arkansas	21	**MA**	Massachusetts[E]	38	**PA**	Pennsylvania[E]
5	**CA**	California	22	**MI**	Michigan	39	**RI**	Rhode Island[F]
6	**CO**	Colorado	23	**MN**	Minnesota	40	**SC**	South Carolina
7	**CT**	Connecticut	24	**MS**	Mississippi	41	**SD**	South Dakota
8	**DE**	Delaware	25	**MO**	Missouri	42	**TN**	Tennessee
9	**FL**	Florida	26	**MT**	Montana	43	**TX**	Texas
10	**GA**	Georgia	27	**NE**	Nebraska	44	**UT**	Utah
11	**HI**	Hawaii	28	**NV**	Nevada	45	**VT**	Vermont
12	**ID**	Idaho	29	**NH**	New Hampshire	46	**VA**	Virginia[E]
13	**IL**	Illinois	30	**NJ**	New Jersey	47	**WA**	Washington
14	**IN**	Indiana	31	**NM**	New Mexico	48	**WV**	West Virginia
15	**IA**	Iowa	32	**NY**	New York	49	**WI**	Wisconsin
16	**KS**	Kansas	33	**NC**	North Carolina	50	**WY**	Wyoming
17	**KY**	Kentucky[E]	34	**ND**	North Dakota	51	**DC**	District of Columbia

b. Shall Issue Law:

- Among the 51 states, 29 states had implemented the *shall-law*. 25 states had the law implemented in 1977-1999.
- The number of states with the *shall-law* increased gradually from 4 to 29 in the span of 23 years, i.e. from 1977 – 1999.



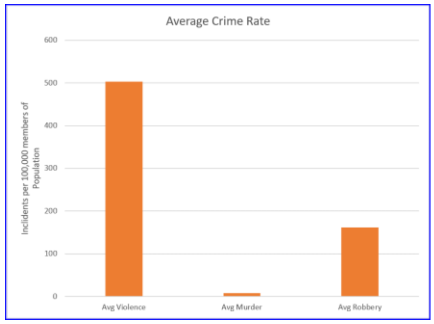
c. Crime Rate:

Crime rate analysis has two parts:

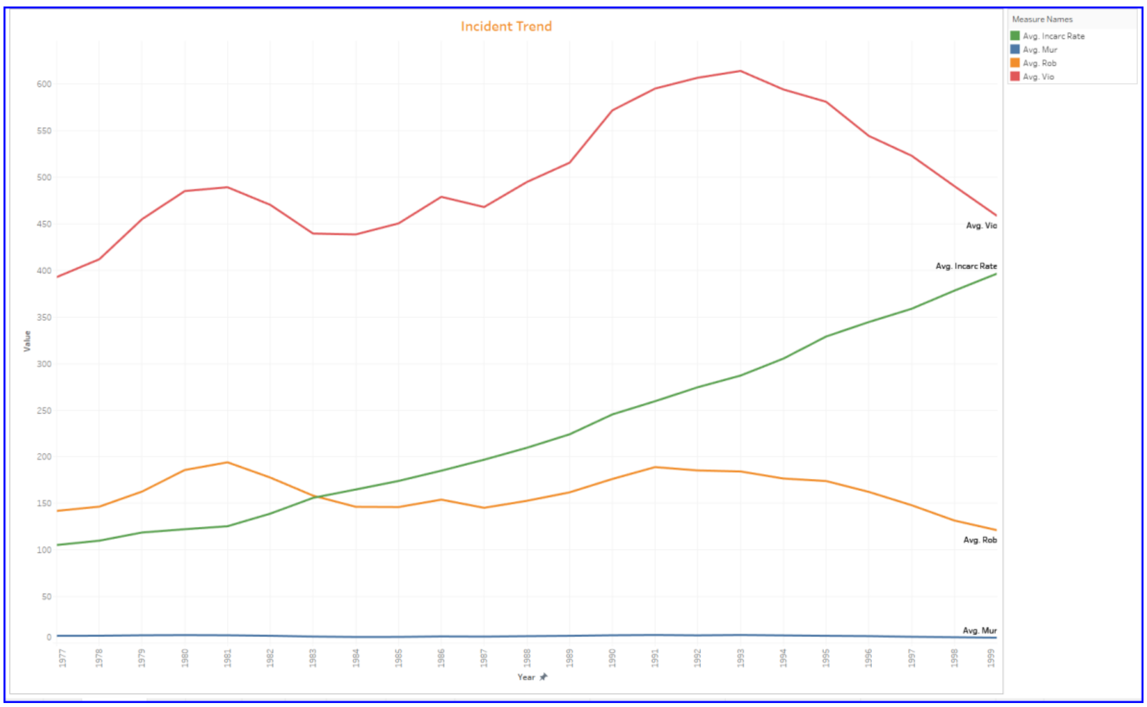
- **Overall Analysis** - which is average across the United States
- **State wise analysis** - which explores the trends and distribution of crime rate across each state and each year.

c.1. Crime Rate – Overall Analysis:

Three different forms of crimes are present in the dataset – Violence, Murder and Robbery. By considering the crime rate factors, it is observed that violence has the highest average rate across the United states, followed by Robbery, and then murder. So, most of the analysis is based on violence rate.



Crime rate is highest in 1993 and there's a decline in the later years.

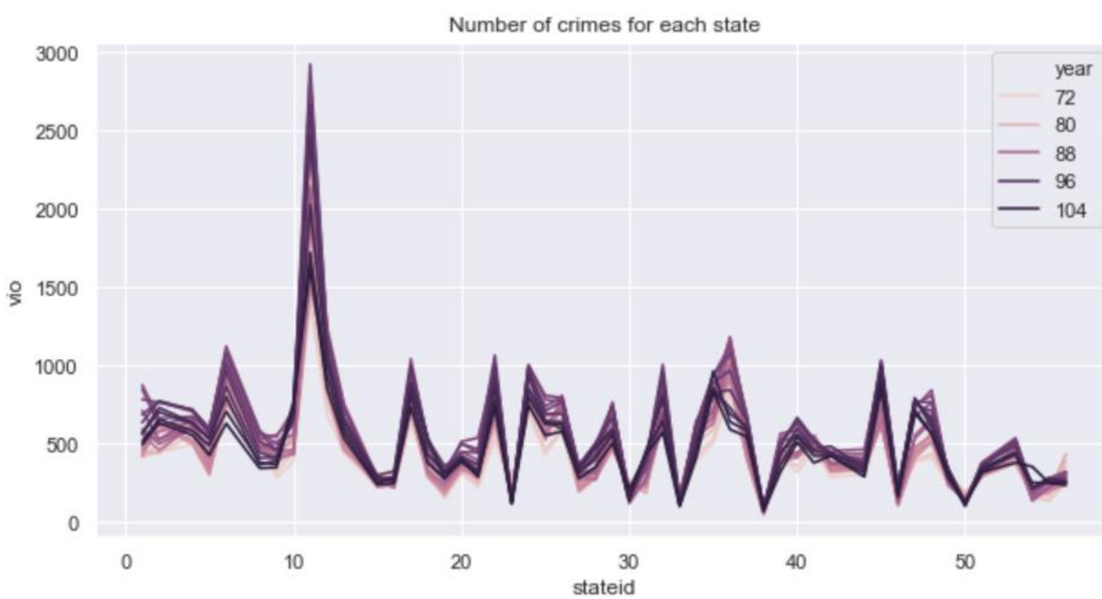


The above trend graph also leads us to the following conclusions:

- The average Incarceration rate has increased over time and it is highest in 1999.
- Also, the Avg violence rate has decreased significantly after 1993 and this could be the effect of heavy Incarceration rate which has steeply increased after 1993.
- Avg robbery rates have also decreased after 1993 which correlates with the higher Incarceration rate.
- No significant trend observed for Avg murder rate, because it has lesser number of observations

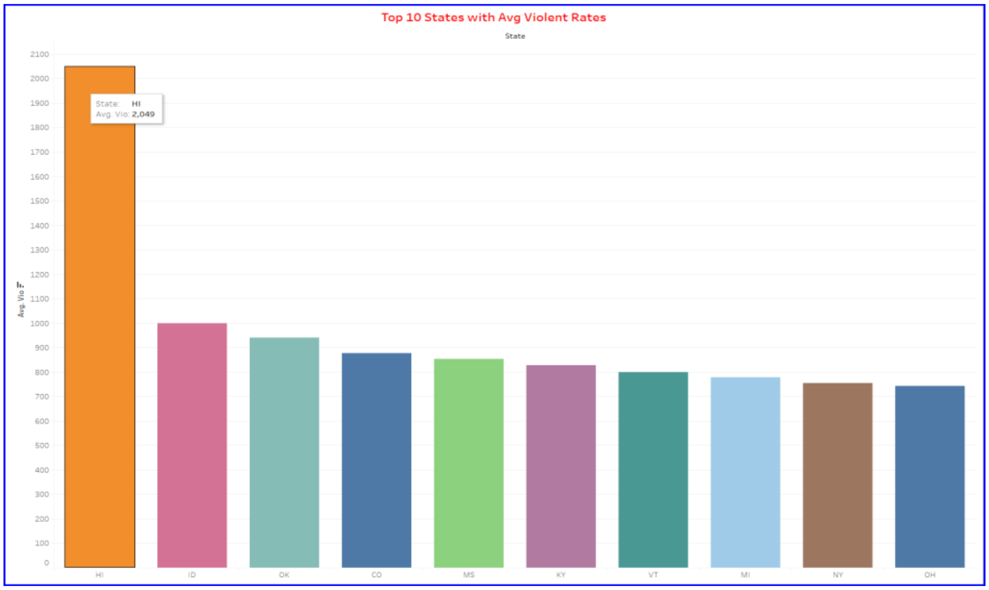
c.2. Crime Rate - State-wise Analysis:

The below dashboard shows various crime rates, and it has been created from the stateid table above (*the order has been retrieved from Wikipedia*). DC is considered as 51st state. The color represents different states and the size of the dots represent crime rates in the corresponding states.

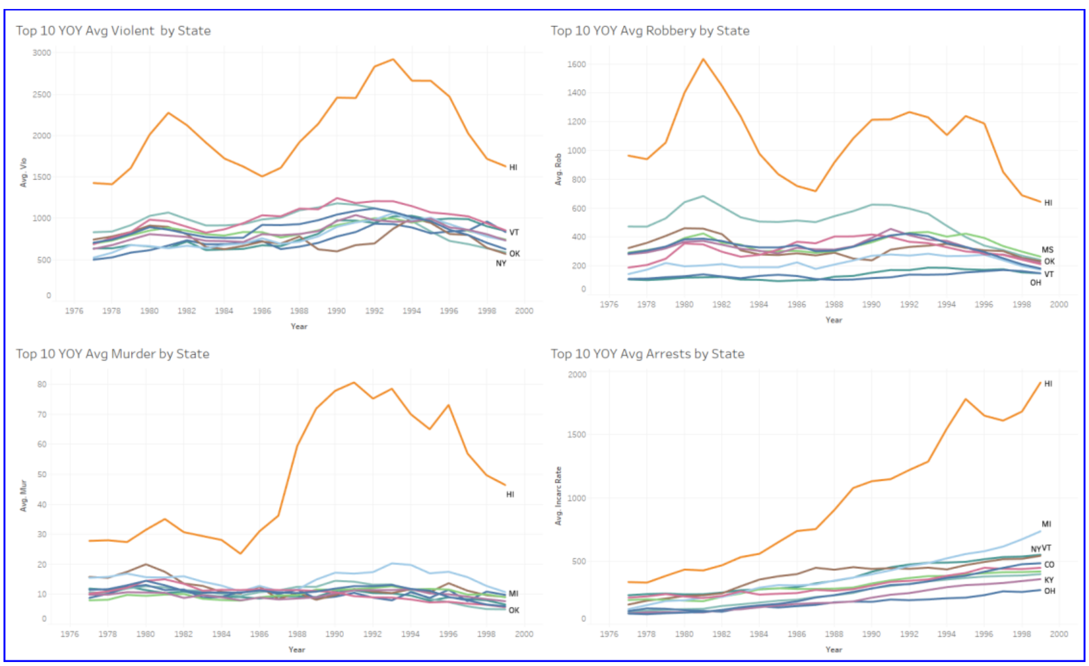


- Hawaii (HI) (state ID – 11) has the highest crime rate and incarceration rate.
- The Pennsylvania (State ID – 38) has the lowest crime rate.

Below are the top 10 states based on average violent crime rate.



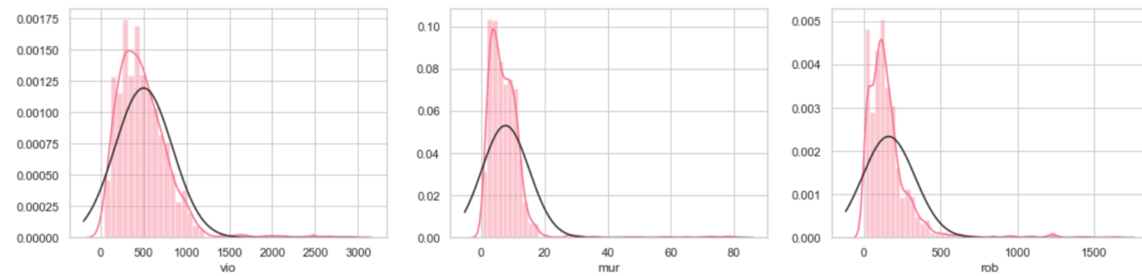
The average crime rates have been compared against incarceration rate across different states, and top 10 states are taken in each category. The below dashboard gives a glimpse of distribution of crime rate for all top 10 states.



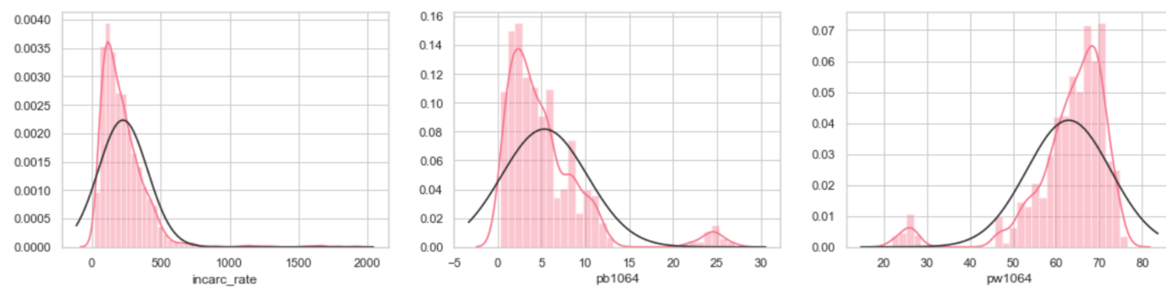
An interesting pattern in observed for the incarceration rates of HI. There is an increasing trend as compared to other forms of crimes.

3. Distribution of variables

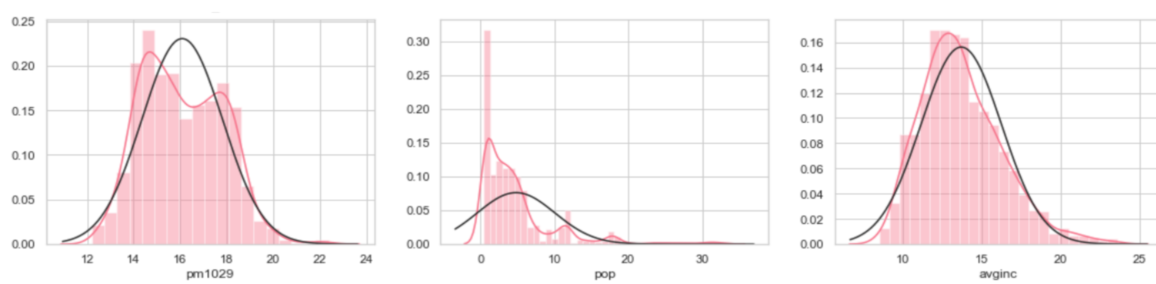
The distributions of vio, mur and rob are positively skewed. So, ln transformations have been taken in order to obtain the approximate normal distributions.



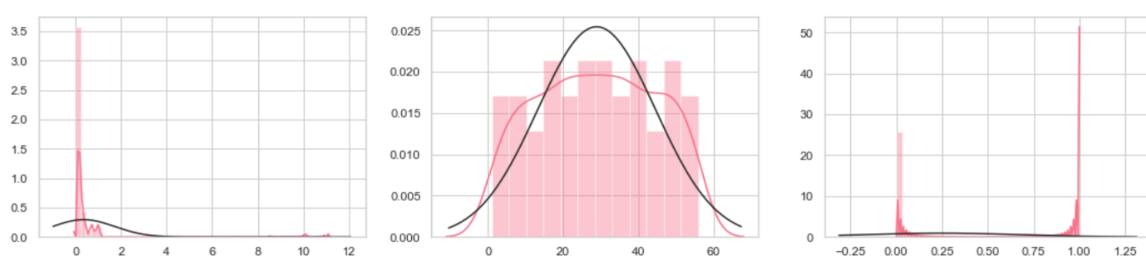
incarc_rate is positively skewed. Hence, ln transformation has been taken for the same. However, pb1064 and pw1064 are approximately normal. Hence the variables can be used as is.



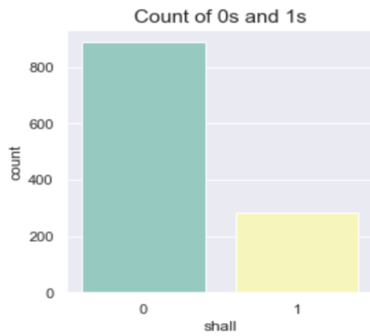
pm1029, pop, avginc are approximately normally distributed.



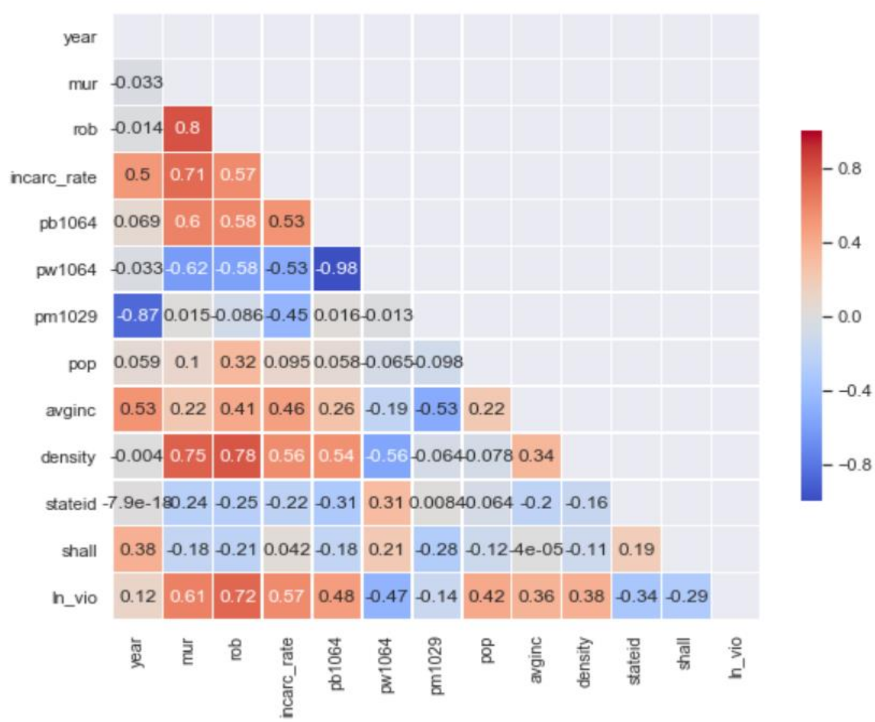
density is positively skewed, so ln transformation has been taken to obtain the approximate normal distribution. stateid is approximately normally distributed. Shall is a binary categorical variable, there's no requirement of ln transformation.



The histogram below explains the number of 0's and 1's in the dataset for Shall variable. More observations have shall = 0, indicating that more states do not have shall- law in effect that year.



Correlation Plot



lnvio is highly correlated with murder, robbery and incarceration rates as expected. pb1064 and pw1064 have almost similar correlation with lnvio because they complement each other. pop, avginc, density have high negative correlation with lnvio . The variable pm1029 has very less correlation with lnvio.

4. Hypothesis Testing:

Based on the EDA performed, the following hypotheses are considered for the Guns dataset to understand the significance of explanatory variables across states with and without the shall law. The results of One-way Anova test to identify the relations between the variables are as follows:

a. Avg Income across state with shall law and without shall-law:

H_0: Average income across state with and without shall law is not significantly different.

Ha: Average income across state with and without shall law is significantly different.

```
anova lnavginc shall

                       Number of obs =      1,173    R-squared     =  0.0005
                       Root MSE      =    .168865    Adj R-squared = -0.0003

                Source | Partial SS         df         MS        F    Prob>F
            -----------+----------------------------------------------------
                 Model |  .01775409          1    .01775409     0.62  0.4302
                       |
                 shall |  .01775409          1    .01775409     0.62  0.4302
                       |
              Residual |  33.391646      1,171     .0285155
            -----------+----------------------------------------------------
                 Total |    33.4094      1,172    .02850631
```

Result: P-value = 0.4302 > 0.05. We conclude that Average income across shall and non-shall state is not significantly different from each other.

b. Average percentage of blacks is different across states with and without shall law:

H_0: Average % of blacks across states with and without shall law is not significantly different

H_a: Average % of blacks across states with and without shall law is significantly different.

```
. anova lnpb1064 shall

                         Number of obs =      1,173    R-squared     =  0.0624
                         Root MSE      =    .635248    Adj R-squared =  0.0616

                  Source | Partial SS         df         MS        F    Prob>F
              -----------+----------------------------------------------------
                   Model |  31.459642          1   31.459642     77.96  0.0000
                         |
                   shall |  31.459642          1   31.459642     77.96  0.0000
                         |
                Residual |  472.54606      1,171   .40354062
              -----------+----------------------------------------------------
                   Total |   504.0057      1,172     .430039
```

Conclusion: Here, the P-value is 0.000, so we reject the null hypothesis. We conclude that the average percentage of blacks across states with and without shall law is significantly different.

c. Average density is different across states with and without shall law:

H_0: Average density across states with and without shall law is not significantly different.

Ha: Average density across states with and without shall law is significantly different

```
. anova lnden shall

                         Number of obs =      1,173    R-squared     =  0.0275
                         Root MSE      =    .346382    Adj R-squared =  0.0267

                  Source | Partial SS         df         MS        F    Prob>F
              -----------+----------------------------------------------------
                   Model |  3.9709599          1   3.9709599     33.10  0.0000
                         |
                   shall |  3.9709599          1   3.9709599     33.10  0.0000
                         |
                Residual |  140.49754      1,171   .11998082
              -----------+----------------------------------------------------
                   Total |   144.4685      1,172   .12326664
```

Conclusion: Here, the P-value is 0.000, and we reject the null hypothesis. We conclude that the average density of population across states with and without shall law is significantly different.

d. Average violent crime is significantly different across states with and without shall law.

H_0: Average violent crime rate across states with and without shall law is not significantly different

Ha: Average violent crime rate across states with and without shall law is significantly different

```
. anova lnvio shall

                         Number of obs =      1,173    R-squared     =  0.0865
                         Root MSE      =    .615238    Adj R-squared =  0.0858

                  Source | Partial SS         df         MS        F    Prob>F
              -----------+----------------------------------------------------
                   Model |  41.992856          1  41.992856   110.94    0.0000
                         |
                   shall |  41.992856          1  41.992856   110.94    0.0000
                         |
                Residual |  443.24385      1,171  .37851738
              -----------+----------------------------------------------------
                   Total |  485.23671      1,172   .4140245
```

Conclusion: Here, the P-value is 0.000, so we reject the null hypothesis. We conclude that the average violent rate across states with and without shall law is significantly different.

General Conclusions:

- Shall-Issue Law has increased from 4 states to 29 states over the period which means there is a possibility that more states trust this law to be effective in bringing down the crime rate.
- Hence, we expect Shall-issue variable to be one of the important variables in determining the violent rate and murder rate, but not for robbery, as robbery has the reachability of gun to many people leading to raise in the robbery rate.
- As we saw in the EDA, incarceration rate has increased over the years and in peak at 1999, but the crime rate is not as high correlated with incarceration rate. So, we could say it is not effectively reducing crime or we can also think there is a causality here as both these variables are related to each other.
- From the hypothesis testing, we could see population density is one of the important parameters for the crime rate changes. There are high chances of crime in densely populated areas rather sparingly populated areas.
- The percentage of whites and black are highly correlated variables as expected and we will retain one of them to see how it impacts the crime rate. But our assumption is this should not affect the crime rate and we do not want to take any racial biasness in our assumption.

5. Regression Analysis:

The following four models are considered to estimate the effect of "Shall" and other variables on the violent crime rate. Though there are three crime factors vio, rob, and mur, from the EDA we have observed that vio has the highest average rate across the United states (highest number of observations in the dataset), therefore most of our analysis is based on this variable as dependent variable in the analysis.

1. Pooled OLS
2. Entity Fixed Effects
3. Entity and Time Fixed Effects
4. Random Effects

a. Considering lnvio as the dependent variable the following set of models are run:
   - Pooled OLS
   - Pooled OLS with Cluster Robust Standard Errors
   - Entity Fixed Effects
   - Entity Fixed Effects with Cluster Robust Standard Errors
   - Entity and Time Fixed Effects
   - Entity and Time Fixed Effects with Cluster Robust Standard Errors
   - Random Effects
   - Random Effects with Cluster Robust Standard Errors

b. White Test is performed to check for Heteroskedasticity

c. Hausman Test is performed to check for Endogeneity

a. Pooled OLS Estimates

In Pooled OLS model the data on different individuals are simply pooled together.

Model:

$$\ln(vio_{it}) = \beta_1 + \beta_2 shall_{it} + \beta_3 \ln\left(incarc_{rate_{it}}\right) + \beta_4 pb1064_{it} + \beta_5 pm1029_{it} + \beta_6 pop_{it} + \beta_7 avginc_{it} + \beta_8 ln(density_{it}) + \varepsilon_{it}$$

Regression Output:

```
. reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall

      Source |       SS           df       MS      Number of obs   =     1,173
-------------+----------------------------------   F(7, 1165)      =    339.76
       Model |  327.976983         7  46.8538548   Prob > F        =    0.0000
    Residual |  160.654575     1,165  .137900923   R-squared       =    0.6712
-------------+----------------------------------   Adj R-squared   =    0.6692
       Total |  488.631558     1,172  .416921125   Root MSE        =    .37135

------------------------------------------------------------------------------
       lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |   .6942355   .0251826    27.57   0.000     .6448272    .7436439
      pb1064 |  -.0033903   .0031686    -1.07   0.285    -.0096071    .0028266
      pm1029 |     .11824   .0097339    12.15   0.000     .0991421    .1373378
         pop |   .0240022   .0022951    10.46   0.000     .0194992    .0285052
      avginc |   .0248689   .0054591     4.56   0.000     .0141581    .0355796
   lndensity |   .0921688    .008831    10.44   0.000     .0748423    .1094953
       shall |  -.2794279   .0274716   -10.17   0.000    -.3333272   -.2255287
       _cons |   .3777348   .2678552     1.41   0.159    -.1477978    .9032674
------------------------------------------------------------------------------
```

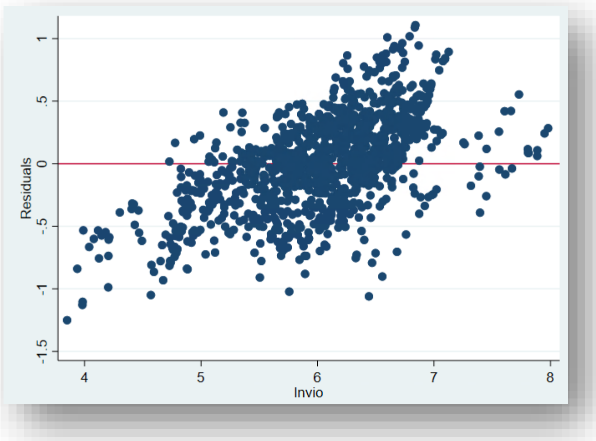
Model based on regression output:

$$\ln(vio_{it}) = 0.3777 - 0.2794\, shall_{it} + 0.6942\, \ln(incarc\_rate_{it}) - 0.0034 pb1064_{it} + 0.1182 pm1029_{it} + 0.024 pop_{it} + 0.02487 avginc_{it} + 0.0922 ln(density_{it}) + \varepsilon_{it}$$

Based on the output it can be said that, the presence of shall issue law reduces the violence rate by 27.94% and this value is highly significant as per the model. Except pb1064 all the other variables impact the violence rate positively and these variables are significant.

White Test

To further check heteroskedasticity in the model, white test is performed. The results are not very surprising as chi^2 value is very high. The null is clearly rejected in this case and it can be concluded that lnvio is heteroskedastic. Below are the results of white test performed.




```
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedastic

         chi2(34)      =    225.46
         Prob > chi2   =    0.0000

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	225.46	34	0.000
Skewness	34.02	7	0.000
Kurtosis	0.10	1	0.755
Total	259.57	42	0.000

Solution to Heteroskedasticity problem:

- It is more appropriate to use the cluster robust standard errors.
- The advantage of using the Cluster Robust model is though the estimates will be inefficient, the cluster robust standard errors will be correct and the SE, T-value, P-value and the confidence interval will be calculated better when compared to Pooled OLS model

a. Pooled OLS with Cluster Robust Standard Errors

It is possible to compute correct standard errors for the least squares estimator using the Clustered Robust Standard Errors:

```
. reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, vce(cluster stateid)

Linear regression                               Number of obs     =      1,173
                                                F(7, 50)          =      27.17
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6712
                                                Root MSE          =     .37135

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

lnvio	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	.6942355	.092407	7.51	0.000	.5086306	.8798405
pb1064	-.0033903	.0119538	-0.28	0.778	-.0274002	.0206197
pm1029	.11824	.0292443	4.04	0.000	.0595011	.1769789
pop	.0240022	.0075225	3.19	0.002	.0088928	.0391117
avginc	.0248689	.016665	1.49	0.142	-.0086038	.0583415
lndensity	.0921688	.0336732	2.74	0.009	.0245343	.1598033
shall	-.2794279	.0789335	-3.54	0.001	-.4379705	-.1208854
_cons	.3777348	.8932571	0.42	0.674	-1.416425	2.171895

The standard errors are now corrected, and so are inflated. Large differences between standard errors imply that there are individual characteristics that are not completely captured by the included explanatory variables that are correlated over time. Ignoring the within-individual correlation means that the reliability of the pooled OLS estimates is overstated. This makes the variables more insignificant.

Major problems with Pooled OLS are as follows:

- Serial Correlation (The correlation between errors of the same entity)
- Heteroskedasticity (increase in variance of error)
- Endogeneity (correlation between error term and the dependent variable)

Impact of these problems are as follows:

- The first two problems related to serial correlation and heteroskedasticity leads to inefficient estimates even though they are unbiased and consistent.
- The last problem of endogeneity results in an invalid model with biased and inconsistent estimates.

The Pooled OLS with cluster robust standard errors is still not efficient estimator since it does not take into consideration the serial correlation between the error terms. The violent crime rates in 1993 affects violent crime rates in 1994. Pooled OLS does not take this into account, and hence provides incorrect results.

In order to control the unobserved heterogeneity in the model, it appropriate to use the Fixed Effects Estimator.

b. Entity Fixed Effects

The coefficients of explanatory variables might be different for different states and can cause endogeneity problem because of both observed and unobserved heterogeneity, which accounts for the Least Squares estimates being biased and inconsistent. Running Fixed effect model allows us to control for unobserved heterogeneity among entities (states), making estimates unbiased and consistent; however, time effects will still not be accounted.

Model:

$$\ln(vio_{it}) = \beta_{1i} + \beta_2 shall_{it} + \beta_3 \ln(incarc\_rate_{it}) + \beta_4 pb1064_{it} + \beta_5 pm1029_{it} + \beta_6 pop_{it} + \beta_7 avginc_{it} + \beta_8 ln(density_{it}) + \varepsilon_{it}$$

Regression Output:

```
. xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe

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

R-sq:                                           Obs per group:
     within  = 0.1765                                         min =         23
     between = 0.0899                                         avg =       23.0
     overall = 0.0667                                         max =         23

                                                F(7,1115)         =      34.15
corr(u_i, Xb)  = -0.6575                        Prob > F          =     0.0000

------------------------------------------------------------------------------
       lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |  -.0090545   .0281105    -0.32   0.747    -.0642099    .0461009
      pb1064 |   .0241139   .0126498     1.91   0.057    -.0007061     .048934
      pm1029 |  -.0512861   .0082642    -6.21   0.000    -.0675012   -.0350711
         pop |   .0114121   .0094163     1.21   0.226    -.0070636    .0298878
      avginc |  -.0014284    .005886    -0.24   0.808    -.0129772    .0101204
   lndensity |  -.2671709   .0884615    -3.02   0.003    -.4407407   -.0936011
       shall |   .0204939   .0181277     1.13   0.258    -.0150743    .0560621
       _cons |   6.049705   .3313636    18.26   0.000     5.399538    6.699871
-------------+----------------------------------------------------------------
     sigma_u |  .80755232
     sigma_e |  .16483395
         rho |  .96000325   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(50, 1115) = 95.96                   Prob > F = 0.0000
```

Model based on regression output:

$$\ln(vio_{it}) = 6.0497 + 0.0205 shall_{it} - 0.0091 \ln(incarc\_rate_{it}) + 0.0241 pb1064_{it} - 0.0513 pm1029_{it} + 0.0114 pop_{it} - 0.0014 avginc_{it} - 0.2672 ln(density_{it}) + \varepsilon_{it}$$

c. Entity Fixed Effects – Cluster Robust Standard Errors

Though the Fixed effects coefficients are inefficient, we can account for unobserved heterogeneity by using the cluster robust standard errors.

```
. xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe cluster(stateid)

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

R-sq:                                           Obs per group:
     within  = 0.1765                                         min =         23
     between = 0.0899                                         avg =       23.0
     overall = 0.0667                                         max =         23

                                                F(7,50)           =       5.51
corr(u_i, Xb)  = -0.6575                        Prob > F          =     0.0001

                                (Std. Err. adjusted for 51 clusters in stateid)
-------------------------------------------------------------------------------
              |               Robust
        lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
lnincarc_rate |  -.0090545   .0595141    -0.15   0.880    -.1285921    .1104831
       pb1064 |   .0241139   .0259551     0.93   0.357    -.0280184    .0762463
       pm1029 |  -.0512861   .0224432    -2.29   0.027    -.0963647   -.0062076
          pop |   .0114121   .0141647     0.81   0.424    -.0170386    .0398628
       avginc |  -.0014284   .0131434    -0.11   0.914    -.0278276    .0249709
    lndensity |  -.2671709   .1769004    -1.51   0.137    -.6224859     .088144
        shall |   .0204939    .039491     0.52   0.606    -.0588262     .099814
        _cons |   6.049705   .6698403     9.03   0.000     4.704291    7.395119
--------------+----------------------------------------------------------------
      sigma_u |  .80755232
      sigma_e |  .16483395
          rho |  .96000325   (fraction of variance due to u_i)
-------------------------------------------------------------------------------
```

From the output we can see that the standard errors have been improved, but the estimates remain inefficient. Thus, the estimates are unbiased and consistent but inefficient.

Model Significance:

- pm1029 (percentage of male population aged 10-29) for the given state is significant at 5% level. A 1% increase in the percentage of males decreases the violent crime rate by 5%. In the real world scenario, young men are bound to commit more violent crimes than the old men.

- Density (population per square mile of land area divided by 1000) has a negative effect on violent crime. This is as expected since areas with a sparse population density rate are locations that are more prone to crime. The coefficient is insignificant as compared to the model with normal standard errors.

- Shall-carry law, incarceration rate, pb1064, avg income, population variables are statistically insignificant in this model.

The interpretation of shall variable is as follows:

Holding other variables fixed, states having shall-carry law in effect in a particular year have violent crime rate 2% higher than the states not having shall-carry law in effect.

d. Entity and Time Fixed Effects

The Entity fixed model gave us a satisfactory result. But we are not limited with that model. We have considered only the fixed effects between the entities and we have not considered the factors that varies over time. For example, it can be a policy change across the country, socio-economic conditions of the country say a severe recession or a war or economic breakdown. These are not changing between the states, but they might change over the time. So, we need to understand how these time varying factors effects our dependent variable.

Model:

$$
\begin{aligned}
\ln(vio_{it}) = \alpha_i &+ \delta_2 year1978_{it} + \delta_3 year1979_{it} + \delta_4 year1980_{it} + \delta_5 year1981_{it} + \delta_6 year1982_{it} \\
&+ \delta_7 year1983_{it} + \delta_8 year1984_{it} + \delta_9 year1985_{it} + \delta_{10} year1986_{it} + \delta_{11} year1984_{it} \\
&+ \delta_{12} year1985_{it} + \delta_{13} year1986_{it} + \delta_{14} year1987_{it} + \delta_{15} year1988_{it} + \delta_{16} year1989_{it} \\
&+ \delta_{17} year1990_{it} + \delta_{18} year1991_{it} + \delta_{19} year1992_{it} + \delta_{20} year1993_{it} + \delta_{21} year1994_{it} \\
&+ \delta_{22} year1995_{it} + \delta_{23} year1996_{it} + \delta_{24} year1997_{it} + \delta_{25} year1998_{it} + \delta_{26} year1999_{it} \\
&+ \beta_2 shall_{it} + \beta_3 \ln(incarc\_rate_{it}) + \beta_4 pb1064_{it} + \beta_5 pm1029_{it} + \beta_6 pop_{it} + \beta_7 avginc_{it} \\
&+ \beta_8 ln(density_{it}) + \mu_{it}
\end{aligned}
$$

Regression-output:

```
37 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall i.year, fe

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

R-sq:                                           Obs per group:
     within  = 0.4256                                         min =         23
     between = 0.2499                                         avg =       23.0
     overall = 0.1768                                         max =         23

                                                F(29,1093)        =      27.92
corr(u_i, Xb)  = -0.7890                        Prob > F          =     0.0000

------------------------------------------------------------------------------
       lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |   -.103518   .0278617    -3.72   0.000    -.1581865   -.0488494
      pb1064 |     -.0089   .0111455    -0.80   0.425     -.030769    .0129691
      pm1029 |   .0772543   .0111447     6.93   0.000     .0553869    .0991218
         pop |   .0064205   .0079546     0.81   0.420    -.0091876    .0220286
      avginc |   .0021565   .0060196     0.36   0.720    -.0096547    .0139677
   lndensity |   -.252022    .075973    -3.32   0.001    -.4010914   -.1029527
       shall |  -.0282769   .0172283    -1.64   0.101    -.0620811    .0055273
             |
        year |
         78  |   .0671185   .0277998     2.41   0.016     .0125715    .1216656
         79  |   .1856235   .0281526     6.59   0.000     .1303842    .2408627
         80  |   .2474713   .0284554     8.70   0.000     .1916379    .3033048
         81  |   .2553967   .0290742     8.78   0.000     .1983492    .3124442
         82  |   .2485782    .030686     8.10   0.000     .1883681    .3087883
         83  |   .2268473   .0329679     6.88   0.000     .1621598    .2915347
         84  |   .2685999    .035651     7.53   0.000     .1986477     .338552
         85  |   .3267886   .0383976     8.51   0.000     .2514473      .40213
         86  |   .4145254   .0418175     9.91   0.000     .3324738    .4965769
         87  |   .4230712   .0451694     9.37   0.000     .3344426    .5116998
         88  |   .4943054   .0487162    10.15   0.000     .3987176    .5898933
         89  |     .55905   .0521011    10.73   0.000     .4568205    .6612794
         90  |      .6928    .055672    12.44   0.000     .5835639    .8020361
         91  |   .7569464   .0584498    12.95   0.000     .6422599     .871633
         92  |    .799334   .0616223    12.97   0.000     .6784226    .9202455
         93  |   .8311132   .0638293    13.02   0.000     .7058714    .9563549
         94  |   .8268423   .0663677    12.46   0.000     .6966196    .9570649
         95  |   .8323325   .0691252    12.04   0.000     .6966994    .9679657
         96  |   .7876731     .07185    10.96   0.000     .6466936    .9286526
         97  |    .776527   .0743257    10.45   0.000     .6306898    .9223642
         98  |   .7308377   .0770233     9.49   0.000     .5797075     .881968
         99  |   .6808563   .0790625     8.61   0.000     .5257248    .8359877
             |
       _cons |   4.178464   .2985587    14.00   0.000     3.592651    4.764277
-------------+----------------------------------------------------------------
     sigma_u |  .93692641
     sigma_e |  .13904812
         rho |  .97844948   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(50, 1093) = 123.89                  Prob > F = 0.0000
```

e. Entity and Time Fixed Effects – Cluster Robust Standard Errors

Cluster robust model has been utilized for correcting the standard errors and make the model to have better standard errors and confidence interval. Thus, the estimates are consistent, unbiased but inefficient.

Regression Output:

```
. xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall i.year, fe cluster(stateid

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

R-sq:                                           Obs per group:
     within  = 0.4256                                         min =         23
     between = 0.2499                                         avg =       23.0
     overall = 0.1768                                         max =         23

                                                F(29,50)          =      52.25
corr(u_i, Xb)  = -0.7890                        Prob > F          =     0.0000

                                 (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
       lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |   -.103518     .07226    -1.43   0.158    -.2486564    .0416205
      pb1064 |     -.0089   .0248878    -0.36   0.722    -.0588886    .0410887
      pm1029 |   .0772543   .0308334     2.51   0.016     .0153237     .139185
         pop |   .0064205   .0131308     0.49   0.627    -.0199534    .0327945
      avginc |   .0021565   .0158065     0.14   0.892    -.0295917    .0339048
   lndensity |   -.252022   .1816408    -1.39   0.171    -.6168583    .1128142
       shall |  -.0282769   .0402029    -0.70   0.485    -.1090268     .052473
             |
        year |
         78  |   .0671185   .0113206     5.93   0.000     .0443804    .0898567
         79  |   .1856235   .0202953     9.15   0.000     .1448591    .2263879
         80  |   .2474713   .0312928     7.91   0.000     .1846179    .3103248
         81  |   .2553967   .0335756     7.61   0.000     .1879581    .3228352
         82  |   .2485782   .0429835     5.78   0.000     .1622433    .3349131
         83  |   .2268473   .0520702     4.36   0.000     .1222613    .3314332
         84  |   .2685999   .0619992     4.33   0.000     .1440708     .393129
         85  |   .3267886   .0728826     4.48   0.000     .1803995    .4731777
         86  |   .4145254   .0882331     4.70   0.000     .2373039    .5917468
         87  |   .4230712   .1011085     4.18   0.000     .2199888    .6261536
         88  |   .4943054   .1108335     4.46   0.000     .2716899     .716921
         89  |     .55905   .1231386     4.54   0.000     .3117189     .806381
         90  |      .6928   .1341262     5.17   0.000     .4233995    .9622005
         91  |   .7569464   .1423801     5.32   0.000     .4709675    1.042925
         92  |    .799334   .1504489     5.31   0.000     .4971485     1.10152
         93  |   .8311132   .1572605     5.28   0.000     .5152462     1.14698
         94  |   .8268423   .1628332     5.08   0.000     .4997821    1.153902
         95  |   .8323325   .1671919     4.98   0.000     .4965176    1.168147
         96  |   .7876731   .1736542     4.54   0.000     .4388783    1.136468
         97  |    .776527   .1795981     4.32   0.000     .4157936     1.13726
         98  |   .7308377   .1853886     3.94   0.000     .3584737    1.103202
         99  |   .6808563   .1889471     3.60   0.001     .3013449    1.060368
             |
       _cons |   4.178464   .7984377     5.23   0.000     2.574754    5.782173
-------------+----------------------------------------------------------------
     sigma_u |  .93692641
     sigma_e |  .13904812
```

Model Significance:

- All the year coefficients are significant and have positive effect on crime. As compared to the base year 1977, on average crime rate increases in the subsequent years. This is similar to the results obtained in EDA where we noticed the increase in crime rate over the years across the country.

- pm1029 (percentage of male population aged 10-29) is significant at the 5% level. This again correlates with our expectation that more crime by younger people than the older people.

- lnincarc_rate and lndensity are insignificant even at the 10% significance level which makes sense. The incarceration rate though not highly correlated as expected but has a positive effect on the crime reduction as observed in the EDA.

- Shall-law – our primary variable to be considered for the analysis, is still not statistically significant at 5% but it will be eventually become significant at slightly above 10%. The estimate has a negative effect on the crime rate which is as expected.

Significance of Time Effects:

- F-Test can be used to check the significance of Time Effect in the Fixed effect model:
- From the model, we can say that the time effects are significant. In order to make sure that the results are consistent an F-test for joint significance is conducted.

H_o: The year dummy variables coefficients are all equal to zero.
H_a: At least one or more-year dummy coefficients are not equal to zero.

The results of the joint hypotheses to test the significance of the coefficients of time are given below

```
         rho |  .97844948   (fraction of variance due to u_i)
------------------------------------------------------------------------------

.
. 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) =   39.57
            Prob > F =    0.0000
```

Conclusion:

The p-value is zero and this clearly states that we can reject the null hypothesis and conclude that at least one of the variables is not zero and that they have significant effect in the model.

f. Random Effects

So far, of all the models that were executed and analyzed it has been observed that the Time Entity Fixed effect model gives a better explanation of our dependent variable vio. It is known that Random effect model is much more efficient than that fixed effect model when there is no endogeneity problem with the error term and any of the explanatory variables or omitted variables.

Model:

$$\ln(vio_{it}) = \beta_{1i} + \beta_2 shall_{it} + \beta_3 \ln(incarc\_rate_{it}) + \beta_4 pb1064_{it} + \beta_5 pm1029_{it} + \beta_6 pop_{it} + \beta_7 avginc_{it} + \beta_8 ln(density_{it}) + \varepsilon_{it}$$

Regression output:

```
. xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall

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

R-sq:                                           Obs per group:
     within  = 0.1547                                         min =         23
     between = 0.4705                                         avg =       23.0
     overall = 0.4411                                         max =         23

                                                Wald chi2(7)      =     272.24
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
        lnvio |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
lnincarc_rate |   .0540951   .0282982     1.91   0.056    -.0013684    .1095586
       pb1064 |   .0351525   .0080151     4.39   0.000     .0194432    .0508619
       pm1029 |  -.0234883   .0079295    -2.96   0.003    -.0390299   -.0079466
          pop |   .0168241   .0063451     2.65   0.008      .004388    .0292603
       avginc |  -.0020607   .0059365    -0.35   0.729     -.013696    .0095746
    lndensity |   .0585901     .02985     1.96   0.050      .000085    .1170951
        shall |  -.0160969   .0183351    -0.88   0.380    -.0520331    .0198393
        _cons |   6.036365   .2735228    22.07   0.000      5.50027     6.57246
--------------+----------------------------------------------------------------
      sigma_u |  .29473274
      sigma_e |  .16483395
          rho |  .76174327   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

$$\ln(vio_{it}) = 6.0364 - 0.0161 shall_{it} + 0.0541 \ln(incarc\_rate_{it}) + 0.0352 pb1064_{it} - 0.0235 pm1029_{it} + 0.0168 pop_{it} - 0.0021 avginc_{it} - 0.0586 ln(density_{it}) + \varepsilon_{it}$$

g. Random Effects – Cluster Robust Standard Errors

To further check for any unobserved heteroskedasticity in the model we ran random effects model with Cluster Robust Standard Errors. The results were as follows:

Regression Output:

```
. xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, cluster(stateid)

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

R-sq:                                           Obs per group:
     within  = 0.1547                                         min =         23
     between = 0.4705                                         avg =       23.0
     overall = 0.4411                                         max =         23

                                                Wald chi2(7)      =      50.54
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

                                 (Std. Err. adjusted for 51 clusters in stateid)
--------------------------------------------------------------------------------
               |               Robust
         lnvio |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
lnincarc_rate  |   .0540951   .0564448     0.96   0.338    -.0565346    .1647248
       pb1064  |   .0351525   .0160786     2.19   0.029     .0036391    .0666659
       pm1029  |  -.0234883   .0193941    -1.21   0.226       -.0615    .0145235
          pop  |   .0168241    .011733     1.43   0.152    -.0061722    .0398205
       avginc  |  -.0020607   .0122593    -0.17   0.867    -.0260884     .021967
    lndensity  |   .0585901    .059781     0.98   0.327    -.0585785    .1757586
        shall  |  -.0160969   .0376442    -0.43   0.669    -.0898782    .0576845
        _cons  |   6.036365   .5857812    10.30   0.000     4.888255    7.184475
---------------+----------------------------------------------------------------
      sigma_u  |  .29473274
      sigma_e  |  .16483395
          rho  |  .76174327   (fraction of variance due to u_i)
--------------------------------------------------------------------------------
```

The below thoughts explain that random effects model is not suitable for our dataset.

- The given dataset is not a random sample. This is a dataset with complete states list and hence there is no much randomness we expect. Hence, random effects model will not give a better result.
- The data has lot of unobserved variables like state/national policies, the percentage of police to people ratio, social and cultural behaviors which varies from state to state. These are not considered in our model and hence this can cause omitted variables bias and more endogeneity problem to our model.

h. Hausman Test

Although it is known to use the Fixed Effects model when not working with random data. Hausman test has been performed to check for endogeneity. If endogeneity exist then we will go for fixed effect model.

To check for Endogeneity, Hausman Test is conducted with following hypotheses:

H_0 : No Endogeneity

H_1 : Endogeneity exists

The results of Hausman test are as follows:

```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
lnincarc_r~e |   -.0090545     .0540951      -.0631496               .
      pb1064 |    .0241139     .0351525      -.0110386        .0097864
      pm1029 |   -.0512861    -.0234883      -.0277979        .0023279
         pop |    .0114121     .0168241      -.0054121        .0069575
      avginc |   -.0014284    -.0020607       .0006323               .
   lndensity |   -.2671709     .0585901       -.325761        .0832731
       shall |    .0204939    -.0160969       .0365908               .
------------------------------------------------------------------------------
                           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(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                         =  -195.36    chi2<0 ==> model fitted on these
                                       data fails to meet the asymptotic
                                       assumptions of the Hausman test;
                                       see suest for a generalized test
```

The results shows a significant difference in the fixed and random effects standard errors proving that there is Endogeneity leading to choosing Fixed Effects model over Random Effects

6. Conclusion:

Out of all the models, "Entity and Time Fixed effects" model gives the better results because it takes care of observed and unobserved heterogeneity and endogeneity. Following are the conclusions drawn from the model:

- Shall-carry is insignificant variable in the model. However, it had a negative coefficient in the model meaning the crime rate appears to be reduced in the states with shall-carry law in effect in a year when compared to states with no shall-carry law in effect in a year.
- Increase in percentage of males aged 10-29 increases the crime rate.

We do not have a conclusive evidence that shall-carry law has any real impact on the reduction in the crime rate. Overall in the given data, whether or not shall-law is implemented there is an increase in the crime rate over the years. States implementing shall-laws have lower growth rate than the states not implementing shall-laws over the years. Hence, we conclude that **"More guns do not bring down the crime rate".**

7. Recommendation:

After a thorough analysis, we bring this recommendation. We feel that we can explain the effect on crime rates more precisely if information regarding variables such as National/State policies, number of police units in the state, percentage of intoxicated people, education status, social or cultural attitudes, rich to poor ratio etc. are known besides what is provided.

Below are our recommendations:

- Create awareness and improve the education status of the people which might decrease the crime rate
- Deploy enough police force in the densely populated areas by which we can reduce the crime rate at densely populated areas
- Improve the socio-economic status of the people by improving more employment opportunities, provide more vocational trainings for people who are skilled but uneducated.
- Improve the technological surveillance and introduce state of art technologies like closed circuit TV monitoring or drones in crowded areas and deserted areas to reduce the crime rate.

```
clear *
use "C:\Users\HXB180009\Desktop\Project\guns.dta"
describe

gen lnvio = ln(vio)
gen lnincarc_rate = ln(incarc_rate)
gen lndensity = ln(density)


*Hypothesis Testing:

*1.  Avg Income across shall law state and non-shall law state:

anova lnavginc shall

*2.  Avg % of blacks across shall law state and non-shall law state:

anova lnpb1064 shall

*3.  Avg Density Vs Shall law states

anova lnden shall

*4.  Avg Violent rate across shall law state and non-shall law state

anova lnvio shall



***************Pooled OLS ********************
 *Model 1:
 * All the results will be inefficient. It does not consider serial correlation and hetersskedasticity. It does not
 * take into account the unobserved heterogenity. The OLS are inefficient. They will still be unbiased and consistent.
 *Impact : incorrect standard errors.

 *dropped year,stateid, mur , rob, pw1064

 *1.1 Pooled OLs
reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall
  *  shall = -.2794279   ,  .0274716
predict ehat1, res
graph twoway scatter ehat1 lnvio, yline(0)

estat imtest, white

 *1.2 Pooled OLS with clusteres Robust standard errors
reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, vce(cluster stateid)
  *  shall = -.2794279  , .0789335

  *** We can see the standard errors have increased in clustered robust model  which means..

 ********** Fixed Effects *********************


xtset stateid year
xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe
xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe cluster(stateid)

*** Entity and time fixed effect*******

xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall i.year, fe
xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall i.year, fe cluster(stateid)

testparm i.year


******* Random Effects **************
```

```
xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall
xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, cluster(stateid)


 ********* Hausmann Test -- FE and RE ************

xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe
estimates store fixed

xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall
estimates store random

hausman fixed random
***** difference = 0 --> no endogenity. Null is rejected. we should go with Fixed Effect.
** In fixed effect Time & entity fixed is best as the testparam clearly ***
```

```
                                                  ___  ____  ____  ____  ____(R)
                                                 /__    /   ____/   /   ____/
                                                ___/   /   /___/   /   /___/
                                                  Statistics/Data Analysis

                                                    User: Output
```

```
  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   15.1   Copyright 1985-2017 StataCorp LLC
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
     Special Edition                  College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

65-user Stata network license expires  4 Oct 2020:
       Serial number:  401609247590
         Licensed to:  Jindal School of Management
                       University of Texas at Dallas

Notes:
      1.  Unicode is supported; see help unicode_advice.
      2.  Maximum number of variables is set to 5000; see help set_maxvar.

Checking for updates...
(contacting http://www.stata.com)

Update status
    Last check for updates:  12 Dec 2019
    New update available:    none         (as of 12 Dec 2019)
    Current update level:    26 Aug 2019  (what's new)

Possible actions

    Do nothing; all files are up to date.

Click to edit automatic update checking preferences

1 . doedit "C:\Users\HXB180009\Desktop\Project\Project.do"

2 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

3 . clear *

4 . use "C:\Users\HXB180009\Desktop\Project\guns.dta"

5 . describe

Contains data from C:\Users\HXB180009\Desktop\Project\guns.dta
  obs:         1,173
 vars:            13                          5 Sep 2014 22:29
 size:        48,093
------------------------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------------------------------------------
year            byte    %9.0g
vio             float   %9.0g                 Violent Crime Rate per 100,000 population (BJS)
mur             float   %9.0g                 Murder Crime Rate per 100,000 population (BJS)
rob             float   %9.0g                 Robbery Crime Rate per 100,000 population (BJS)
incarc_rate     int     %8.0g                 72-99 ONLY - Lagged Rate per 100,000 resident pop of
pb1064          float   %9.0g
pw1064          float   %9.0g
pm1029          float   %9.0g
pop             float   %9.0g
avginc          float   %9.0g
density         float   %9.0g
stateid         byte    %9.0g
shall           byte    %9.0g
------------------------------------------------------------------------------------------------
Sorted by:
```

```
 6 .
   end of do-file

 7 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

 8 . gen lnvio = ln(vio)

 9 . *gen lnmur = ln(mur)
10 . *gen lnrob = ln(rob)
11 . gen lnincarc_rate = ln(incarc_rate)

12 . gen lndensity = ln(density)

13 .
   end of do-file

14 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

15 . reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall
```

Source	SS	df	MS			
				Number of obs	=	1,173
				F(7, 1165)	=	339.76
Model	327.976983	7	46.8538548	Prob > F	=	0.0000
Residual	160.654575	1,165	.137900923	R-squared	=	0.6712
				Adj R-squared	=	0.6692
Total	488.631558	1,172	.416921125	Root MSE	=	.37135

lnvio	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	.6942355	.0251826	27.57	0.000	.6448272	.7436439
pb1064	-.0033903	.0031686	-1.07	0.285	-.0096071	.0028266
pm1029	.11824	.0097339	12.15	0.000	.0991421	.1373378
pop	.0240022	.0022951	10.46	0.000	.0194992	.0285052
avginc	.0248689	.0054591	4.56	0.000	.0141581	.0355796
lndensity	.0921688	.008831	10.44	0.000	.0748423	.1094953
shall	-.2794279	.0274716	-10.17	0.000	-.3333272	-.2255287
_cons	.3777348	.2678552	1.41	0.159	-.1477978	.9032674

```
16 .
   end of do-file

17 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

18 . predict ehat1, res

19 .
   end of do-file

20 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

21 . graph twoway scatter ehat1 lnvio, yline(0)
```

```
22 .
   end of do-file

23 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

24 . estat imtest, white

   White's test for Ho: homoskedasticity
            against Ha: unrestricted heteroskedasticity

            chi2(34)      =    225.46
            Prob > chi2   =    0.0000

   Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	225.46	34	0.0000
Skewness	34.02	7	0.0000
Kurtosis	0.10	1	0.7550
Total	259.57	42	0.0000

```
25 .
   end of do-file

26 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

27 . reg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, vce(cluster stateid)

   Linear regression                               Number of obs     =      1,173
                                                   F(7, 50)          =      27.17
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.6712
                                                   Root MSE          =     .37135

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

lnvio	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	.6942355	.092407	7.51	0.000	.5086306	.8798405
pb1064	-.0033903	.0119538	-0.28	0.778	-.0274002	.0206197
pm1029	.11824	.0292443	4.04	0.000	.0595011	.1769789
pop	.0240022	.0075225	3.19	0.002	.0088928	.0391117
avginc	.0248689	.016665	1.49	0.142	-.0086038	.0583415
lndensity	.0921688	.0336732	2.74	0.009	.0245343	.1598033
shall	-.2794279	.0789335	-3.54	0.001	-.4379705	-.1208854
_cons	.3777348	.8932571	0.42	0.674	-1.416425	2.171895

```
28 .
   end of do-file
```

```
29 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

30 . xtset stateid year
          panel variable:  stateid (strongly balanced)
           time variable:  year, 77 to 99
                   delta:  1 unit

31 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe

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

   R-sq:                                           Obs per group:
        within  = 0.1765                                         min =         23
        between = 0.0899                                         avg =       23.0
        overall = 0.0667                                         max =         23

                                                   F(7,1115)         =      34.15
   corr(u_i, Xb)  = -0.6575                        Prob > F          =     0.0000

   ------------------------------------------------------------------------------
           lnvio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
   --------------+----------------------------------------------------------------
   lnincarc_rate |  -.0090545   .0281105    -0.32   0.747    -.0642099    .0461009
          pb1064 |   .0241139   .0126498     1.91   0.057    -.0007061     .048934
          pm1029 |  -.0512861   .0082642    -6.21   0.000    -.0675012   -.0350711
             pop |   .0114121   .0094163     1.21   0.226    -.0070636    .0298878
          avginc |  -.0014284    .005886    -0.24   0.808    -.0129772    .0101204
       lndensity |  -.2671709   .0884615    -3.02   0.003    -.4407407   -.0936011
           shall |   .0204939   .0181277     1.13   0.258    -.0150743    .0560621
           _cons |   6.049705   .3313636    18.26   0.000     5.399538    6.699871
   --------------+----------------------------------------------------------------
         sigma_u |  .80755232
         sigma_e |  .16483395
             rho |  .96000325   (fraction of variance due to u_i)
   ------------------------------------------------------------------------------
   F test that all u_i=0: F(50, 1115) = 95.96                   Prob > F = 0.0000

32 .
   end of do-file

33 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

34 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe cluster(stateid)

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

   R-sq:                                           Obs per group:
        within  = 0.1765                                         min =         23
        between = 0.0899                                         avg =       23.0
        overall = 0.0667                                         max =         23

                                                   F(7,50)           =       5.51
   corr(u_i, Xb)  = -0.6575                        Prob > F          =     0.0001
```

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

lnvio	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	**-.0090545**	**.0595141**	**-0.15**	**0.880**	**-.1285921**	**.1104831**
pb1064	**.0241139**	**.0259551**	**0.93**	**0.357**	**-.0280184**	**.0762463**
pm1029	**-.0512861**	**.0224432**	**-2.29**	**0.027**	**-.0963647**	**-.0062076**
pop	**.0114121**	**.0141647**	**0.81**	**0.424**	**-.0170386**	**.0398628**
avginc	**-.0014284**	**.0131434**	**-0.11**	**0.914**	**-.0278276**	**.0249709**
lndensity	**-.2671709**	**.1769004**	**-1.51**	**0.137**	**-.6224859**	**.088144**
shall	**.0204939**	**.039491**	**0.52**	**0.606**	**-.0588262**	**.099814**
_cons	**6.049705**	**.6698403**	**9.03**	**0.000**	**4.704291**	**7.395119**
sigma_u	**.80755232**					
sigma_e	**.16483395**					
rho	**.96000325**	(fraction of variance due to u_i)				

35 .
end of do-file

36 . do "C:\Users\HXB180~1\AppData\Local\Temp\24\STD3dfc_000000.tmp"

37 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity shall i.year, fe

Fixed-effects (within) regression		Number of obs	=	**1,173**
Group variable: **stateid**		Number of groups	=	**51**
R-sq:		Obs per group:		
within	= **0.4256**	min	=	**23**
between	= **0.2499**	avg	=	**23.0**
overall	= **0.1768**	max	=	**23**
		F(**29,1093**)	=	**27.92**
corr(u_i, Xb)	= **-0.7890**	Prob > F	=	**0.0000**

lnvio	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	**-.103518**	**.0278617**	**-3.72**	**0.000**	**-.1581865**	**-.0488494**
pb1064	**-.0089**	**.0111455**	**-0.80**	**0.425**	**-.030769**	**.0129691**
pm1029	**.0772543**	**.0111447**	**6.93**	**0.000**	**.0553869**	**.0991218**
pop	**.0064205**	**.0079546**	**0.81**	**0.420**	**-.0091876**	**.0220286**
avginc	**.0021565**	**.0060196**	**0.36**	**0.720**	**-.0096547**	**.0139677**
lndensity	**-.252022**	**.075973**	**-3.32**	**0.001**	**-.4010914**	**-.1029527**
shall	**-.0282769**	**.0172283**	**-1.64**	**0.101**	**-.0620811**	**.0055273**
year						
78	**.0671185**	**.0277998**	**2.41**	**0.016**	**.0125715**	**.1216656**
79	**.1856235**	**.0281526**	**6.59**	**0.000**	**.1303842**	**.2408627**
80	**.2474713**	**.0284554**	**8.70**	**0.000**	**.1916379**	**.3033048**
81	**.2553967**	**.0290742**	**8.78**	**0.000**	**.1983492**	**.3124442**
82	**.2485782**	**.030686**	**8.10**	**0.000**	**.1883681**	**.3087883**
83	**.2268473**	**.0329679**	**6.88**	**0.000**	**.1621598**	**.2915347**
84	**.2685999**	**.035651**	**7.53**	**0.000**	**.1986477**	**.338552**
85	**.3267886**	**.0383976**	**8.51**	**0.000**	**.2514473**	**.40213**
86	**.4145254**	**.0418175**	**9.91**	**0.000**	**.3324738**	**.4965769**
87	**.4230712**	**.0451694**	**9.37**	**0.000**	**.3344426**	**.5116998**
88	**.4943054**	**.0487162**	**10.15**	**0.000**	**.3987176**	**.5898933**
89	**.55905**	**.0521011**	**10.73**	**0.000**	**.4568205**	**.6612794**
90	**.6928**	**.055672**	**12.44**	**0.000**	**.5835639**	**.8020361**
91	**.7569464**	**.0584498**	**12.95**	**0.000**	**.6422599**	**.871633**
92	**.799334**	**.0616223**	**12.97**	**0.000**	**.6784226**	**.9202455**
93	**.8311132**	**.0638293**	**13.02**	**0.000**	**.7058714**	**.9563549**
94	**.8268423**	**.0663677**	**12.46**	**0.000**	**.6966196**	**.9570649**
95	**.8323325**	**.0691252**	**12.04**	**0.000**	**.6966994**	**.9679657**

96	.7876731	.07185	10.96	0.000	.6466936	.9286526
97	.776527	.0743257	10.45	0.000	.6306898	.9223642
98	.7308377	.0770233	9.49	0.000	.5797075	.881968
99	.6808563	.0790625	8.61	0.000	.5257248	.8359877
_cons	4.178464	.2985587	14.00	0.000	3.592651	4.764277
sigma_u	.93692641					
sigma_e	.13904812					
rho	.97844948	(fraction of variance due to u_i)				

F test that all u_i=0: F(50, 1093) = 123.89 Prob > F = 0.0000

38 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity shall i.year, fe cluster(stateid)

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

R-sq:                                           Obs per group:
     within  = 0.4256                                         min =         23
     between = 0.2499                                         avg =       23.0
     overall = 0.1768                                         max =         23

                                                F(29,50)          =      52.25
corr(u_i, Xb)  = -0.7890                        Prob > F          =     0.0000
```

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

lnvio	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	-.103518	.07226	-1.43	0.158	-.2486564	.0416205
pb1064	-.0089	.0248878	-0.36	0.722	-.0588886	.0410887
pm1029	.0772543	.0308334	2.51	0.016	.0153237	.139185
pop	.0064205	.0131308	0.49	0.627	-.0199534	.0327945
avginc	.0021565	.0158065	0.14	0.892	-.0295917	.0339048
lndensity	-.252022	.1816408	-1.39	0.171	-.6168583	.1128142
shall	-.0282769	.0402029	-0.70	0.485	-.1090268	.052473
year						
78	.0671185	.0113206	5.93	0.000	.0443804	.0898567
79	.1856235	.0202953	9.15	0.000	.1448591	.2263879
80	.2474713	.0312928	7.91	0.000	.1846179	.3103248
81	.2553967	.0335756	7.61	0.000	.1879581	.3228352
82	.2485782	.0429835	5.78	0.000	.1622433	.3349131
83	.2268473	.0520702	4.36	0.000	.1222613	.3314332
84	.2685999	.0619992	4.33	0.000	.1440708	.393129
85	.3267886	.0728826	4.48	0.000	.1803995	.4731777
86	.4145254	.0882331	4.70	0.000	.2373039	.5917468
87	.4230712	.1011085	4.18	0.000	.2199888	.6261536
88	.4943054	.1108335	4.46	0.000	.2716899	.716921
89	.55905	.1231386	4.54	0.000	.3117189	.806381
90	.6928	.1341262	5.17	0.000	.4233995	.9622005
91	.7569464	.1423801	5.32	0.000	.4709675	1.042925
92	.799334	.1504489	5.31	0.000	.4971485	1.10152
93	.8311132	.1572605	5.28	0.000	.5152462	1.14698
94	.8268423	.1628332	5.08	0.000	.4997821	1.153902
95	.8323325	.1671919	4.98	0.000	.4965176	1.168147
96	.7876731	.1736542	4.54	0.000	.4388783	1.136468
97	.776527	.1795981	4.32	0.000	.4157936	1.13726
98	.7308377	.1853886	3.94	0.000	.3584737	1.103202
99	.6808563	.1889471	3.60	0.001	.3013449	1.060368
_cons	4.178464	.7984377	5.23	0.000	2.574754	5.782173
sigma_u	.93692641					
sigma_e	.13904812					

```
         rho |  .97844948   (fraction of variance due to u_i)
------------------------------------------------------------------------------

39 .
40 . 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) =   39.57
            Prob > F =    0.0000

41 .
42 .
43 . ******* Random Effects **************
44 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall

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

R-sq:                                           Obs per group:
     within  = 0.1547                                         min =         23
     between = 0.4705                                         avg =       23.0
     overall = 0.4411                                         max =         23

                                                Wald chi2(7)      =     272.24
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       lnvio |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |   .0540951   .0282982     1.91   0.056    -.0013684    .1095586
      pb1064 |   .0351525   .0080151     4.39   0.000     .0194432    .0508619
      pm1029 |  -.0234883   .0079295    -2.96   0.003    -.0390299   -.0079466
         pop |   .0168241   .0063451     2.65   0.008      .004388    .0292603
      avginc |  -.0020607   .0059365    -0.35   0.729     -.013696    .0095746
   lndensity |   .0585901     .02985     1.96   0.050      .000085    .1170951
       shall |  -.0160969   .0183351    -0.88   0.380    -.0520331    .0198393
       _cons |   6.036365   .2735228    22.07   0.000      5.50027     6.57246
-------------+----------------------------------------------------------------
     sigma_u |  .29473274
     sigma_e |  .16483395
         rho |  .76174327   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
45 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, cluster(stateid)

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

R-sq:                                           Obs per group:
     within  = 0.1547                                         min =         23
     between = 0.4705                                         avg =       23.0
     overall = 0.4411                                         max =         23

                                                Wald chi2(7)      =      50.54
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

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

lnvio	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnincarc_rate	.0540951	.0564448	0.96	0.338	-.0565346	.1647248
pb1064	.0351525	.0160786	2.19	0.029	.0036391	.0666659
pm1029	-.0234883	.0193941	-1.21	0.226	-.0615	.0145235
pop	.0168241	.011733	1.43	0.152	-.0061722	.0398205
avginc	-.0020607	.0122593	-0.17	0.867	-.0260884	.021967
lndensity	.0585901	.059781	0.98	0.327	-.0585785	.1757586
shall	-.0160969	.0376442	-0.43	0.669	-.0898782	.0576845
_cons	6.036365	.5857812	10.30	0.000	4.888255	7.184475
sigma_u	.29473274					
sigma_e	.16483395					
rho	.76174327	(fraction of variance due to u_i)				

```
46 .
47 .
48 .  ********* Hausmann Test -- FE and RE ************
49 .
50 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall, fe

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

R-sq:                                           Obs per group:
     within  = 0.1765                                         min =         23
     between = 0.0899                                         avg =       23.0
     overall = 0.0667                                         max =         23

                                                F(7,1115)         =      34.15
corr(u_i, Xb)  = -0.6575                        Prob > F          =     0.0000
```

lnvio	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnincarc_rate	-.0090545	.0281105	-0.32	0.747	-.0642099	.0461009
pb1064	.0241139	.0126498	1.91	0.057	-.0007061	.048934
pm1029	-.0512861	.0082642	-6.21	0.000	-.0675012	-.0350711
pop	.0114121	.0094163	1.21	0.226	-.0070636	.0298878
avginc	-.0014284	.005886	-0.24	0.808	-.0129772	.0101204
lndensity	-.2671709	.0884615	-3.02	0.003	-.4407407	-.0936011
shall	.0204939	.0181277	1.13	0.258	-.0150743	.0560621
_cons	6.049705	.3313636	18.26	0.000	5.399538	6.699871
sigma_u	.80755232					
sigma_e	.16483395					
rho	.96000325	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(50, 1115) = 95.96                   Prob > F = 0.0000
```

```
51 . estimates store fixed

52 .
53 . xtreg lnvio lnincarc_rate pb1064 pm1029 pop avginc lndensity  shall

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

R-sq:                                           Obs per group:
     within  = 0.1547                                         min =         23
     between = 0.4705                                         avg =       23.0
     overall = 0.4411                                         max =         23

                                                Wald chi2(7)      =     272.24
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       lnvio |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnincarc_rate |   .0540951   .0282982     1.91   0.056    -.0013684    .1095586
      pb1064 |   .0351525   .0080151     4.39   0.000     .0194432    .0508619
      pm1029 |  -.0234883   .0079295    -2.96   0.003    -.0390299   -.0079466
         pop |   .0168241   .0063451     2.65   0.008      .004388    .0292603
      avginc |  -.0020607   .0059365    -0.35   0.729     -.013696    .0095746
   lndensity |   .0585901     .02985     1.96   0.050      .000085    .1170951
       shall |  -.0160969   .0183351    -0.88   0.380    -.0520331    .0198393
       _cons |   6.036365   .2735228    22.07   0.000      5.50027     6.57246
-------------+----------------------------------------------------------------
     sigma_u |  .29473274
     sigma_e |  .16483395
         rho |  .76174327   (fraction of variance due to u_i)
------------------------------------------------------------------------------

54 . estimates store random

55 .
56 . hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
lnincarc_r~e |   -.0090545     .0540951       -.0631496               .
      pb1064 |    .0241139     .0351525       -.0110386        .0097864
      pm1029 |   -.0512861    -.0234883       -.0277979        .0023279
         pop |    .0114121     .0168241       -.0054121        .0069575
      avginc |   -.0014284    -.0020607        .0006323               .
   lndensity |   -.2671709     .0585901        -.325761        .0832731
       shall |    .0204939    -.0160969        .0365908               .
------------------------------------------------------------------------------
                           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(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =  -195.36    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test
```

```
57 . ***** difference = 0 --> no endogenity. Null is rejected. we should go with Fixed Effect.
58 . ** In fixed effect Time & entity fixed is best as the testparam clearly ***
59 .
60 .
61 .
   end of do-file

62 .
```