What drives crime in the U.S. ?

Hiba Bousarsar Samy Mokrani Baptiste Llamas

École Nationale de l'Aviation Civile

May 10, 2023

Presentation Overview

What drives crime in the U.S. ?

- What influences crime rates in the US ?
- Does police presence increase those ?

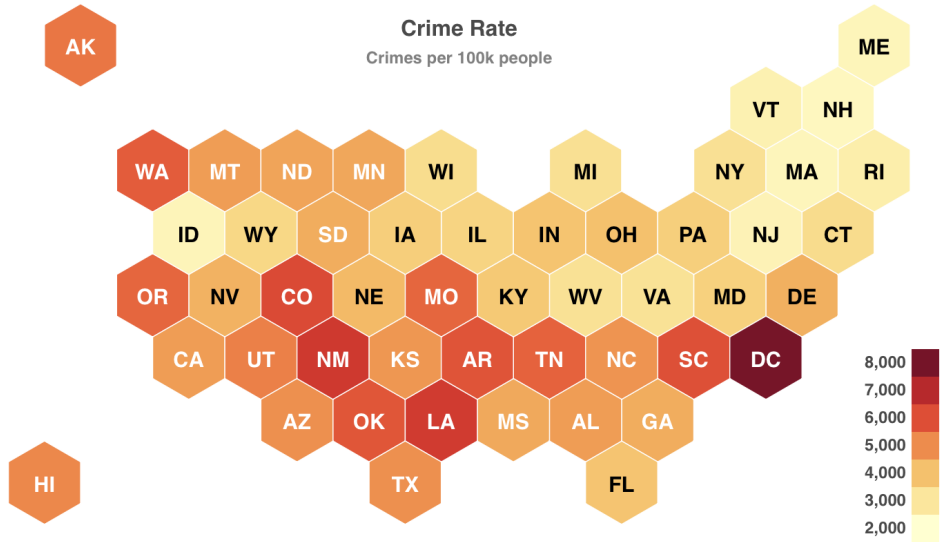


Figure: Crime per 100k people in 2022 in the US, by state

Data Collecting

- US Bank Data Base
- US Government Data Base

Procedure

Cross-sectionnal Model

- 50 states + District of Columbia
- 1 year : 2021

Time-series Model

- 3 states : California, New York, Texas
- 22 years : from 2000 to 2021

Panel Model

- 50 states + District of Columbia
- 11 years : from 2011 to 2021

Cross-sectionnal Data

Variables	Nicknames	Unit	Expected sign
AFRICAN-AMERICAN-POP	afr	% of black	+
BACHELOR-POP	bachelor	% of bachelor	-
HISPANIC-POP	hisp	% of hispanic	+
DEATH-PENALTY	pen	dummy	?
POLICE-OFFICERS	police	per 100k people	-
UNDER-POVERTY-LINE	poverty	% of poor people	+
WEAPON-ALLOWED	weapon	dummy	+
WHITE-POP	white	% of white	-

Table: Cross-sectionnal variables expected signs

First "naive" equation

- Modelization of crime rate per state on a given time (2021)
- We modelled (1):

$$cr_i = \beta_0 + \beta_1 afr_i + \beta_2 hisp_i + \beta_3 white_i + \beta_4 pen_i + \beta_5 police_i + \beta_6 poverty_i + \beta_7 weapon_i + \beta_8 bachelor_i + u_i \quad (1)$$

Equation (1) results

Equation: EQ01 Workfile: CROSS-SECTONNIAL::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CR
Method: Least Squares
Date: 04/13/23 Time: 17:11
Sample: 1 51
Included observations: 51

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3053.974	3030.034	1.007901	0.3193
AFR	-2773.531	2548.810	-1.088167	0.2827
HISP	-1912.617	2160.118	-0.885423	0.3810
WHITE	-4501.350	1685.630	-2.670426	0.0107
PEN	697.5634	370.2608	1.883978	0.0665
POLICE	-3.604547	2.013894	-1.789840	0.0807
POVERTY	25533.51	8662.132	2.947717	0.0052
WEAPON	699.3395	571.4125	1.223879	0.2278
BACHELOR	6530.984	3613.741	1.807264	0.0779

R-squared	0.448744	Mean dependent var	4359.824
Adjusted R-squared	0.343743	S.D. dependent var	1254.463
S.E. of regression	1016.236	Akaike info criterion	16.84438
Sum squared resid	43374925	Schwarz criterion	17.18530
Log likelihood	-420.5318	Hannan-Quinn criter.	16.97466
F-statistic	4.273708	Durbin-Watson stat	2.487297
Prob(F-statistic)	0.000803		

Sign comparaison

Nicknames	Unit	Expected sign	Sign
afr	% of black	+	/
hisp	% of hispanic	+	/
white	% of white	-	-
pen	dummy	?	+
police	per 100k people	-	-
poverty	% of poor people	+	+
weapon	dummy	+	/
bachelor	% of bachelor	-	+

Table: Equation (1) coefficients signs

Correlation

	AFR	BACHELOR	CR	HISP	PEN	POLICE	POVERTY	WEAPON	WHITE
AFR	1.000000								
BACHE...	0.159367	1.000000							
CR	0.314029	-0.006751	1.000000						
HISP	-0.123214	0.101319	0.180792	1.000000					
PEN	0.111560	-0.517098	0.293805	0.035045	1.000000				
POLICE	0.291927	-0.089813	-0.074667	0.117962	-0.136942	1.000000			
POVERTY	0.516963	-0.449708	0.448285	0.078586	0.373266	0.221536	1.000000		
WEAPON	-0.247504	-0.356258	-0.082755	-0.335408	0.085480	-0.075819	-0.119520	1.000000	
WHITE	-0.484421	-0.225448	-0.404889	-0.572008	0.047098	-0.397524	-0.293137	0.390666	1.000000

Figure: Correlation Matrix

Tests

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 2 lags

F-statistic	2.594979	Prob. F(2,40)	0.0872
Obs*R-squared	5.857227	Prob. Chi-Square(2)	0.0535

Figure: Breusch-Godfrey Correlation LM Test

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	1.189186	Prob. F(8,42)	0.3285
Obs*R-squared	9.418658	Prob. Chi-Square(8)	0.3082
Scaled explained SS	4.565661	Prob. Chi-Square(8)	0.8028

Figure: Breusch-Pagan-Godfrey Heteroskedasticity Test

Tests

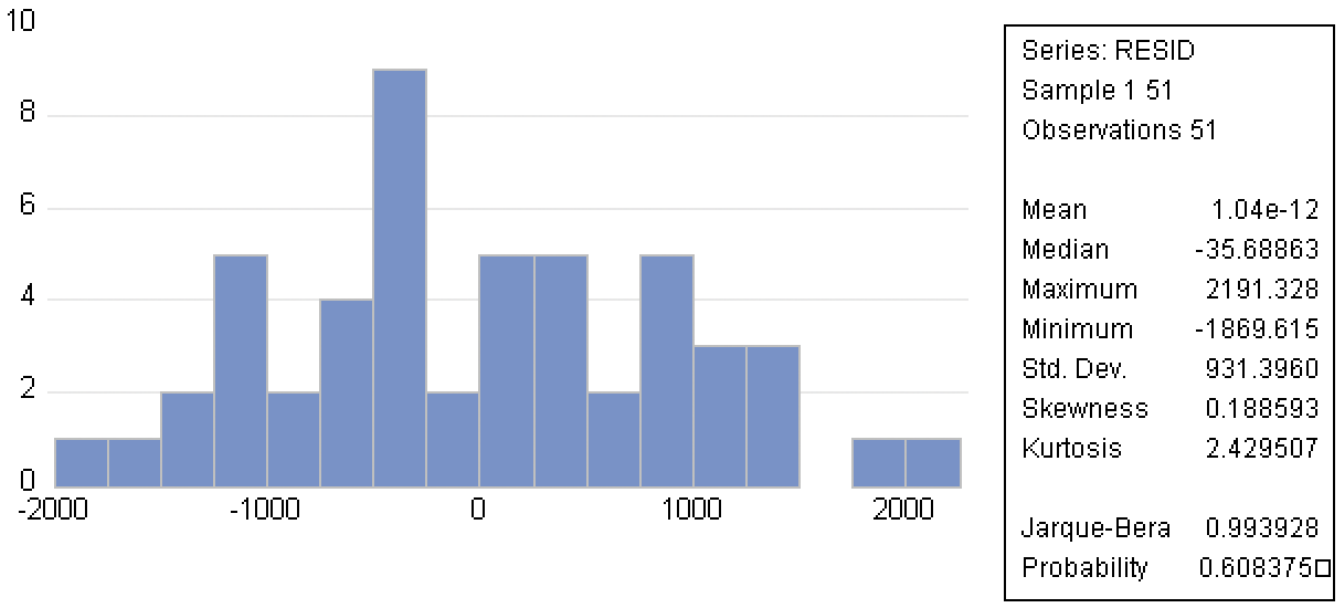


Figure: Residuals

Second equation

- Modelization of crime rate per state on a given time (2021), but improved
- We modelled (2):

$$cr_i = \beta_0 + \beta_1 weapon_i + \beta_2 pen_i + \beta_3 police_i + \beta_4 police_i^2 + \beta_5 poverty_i + \beta_6 afr * police_i + \beta_7 afr * poverty_i + \beta_8 hisp * police_i + \beta_9 hisp * poverty_i + u_i \quad (2)$$

Equation (2) results

Equation: EQ02 Workfile: CROSS-SECTONNIAL::Untitled\

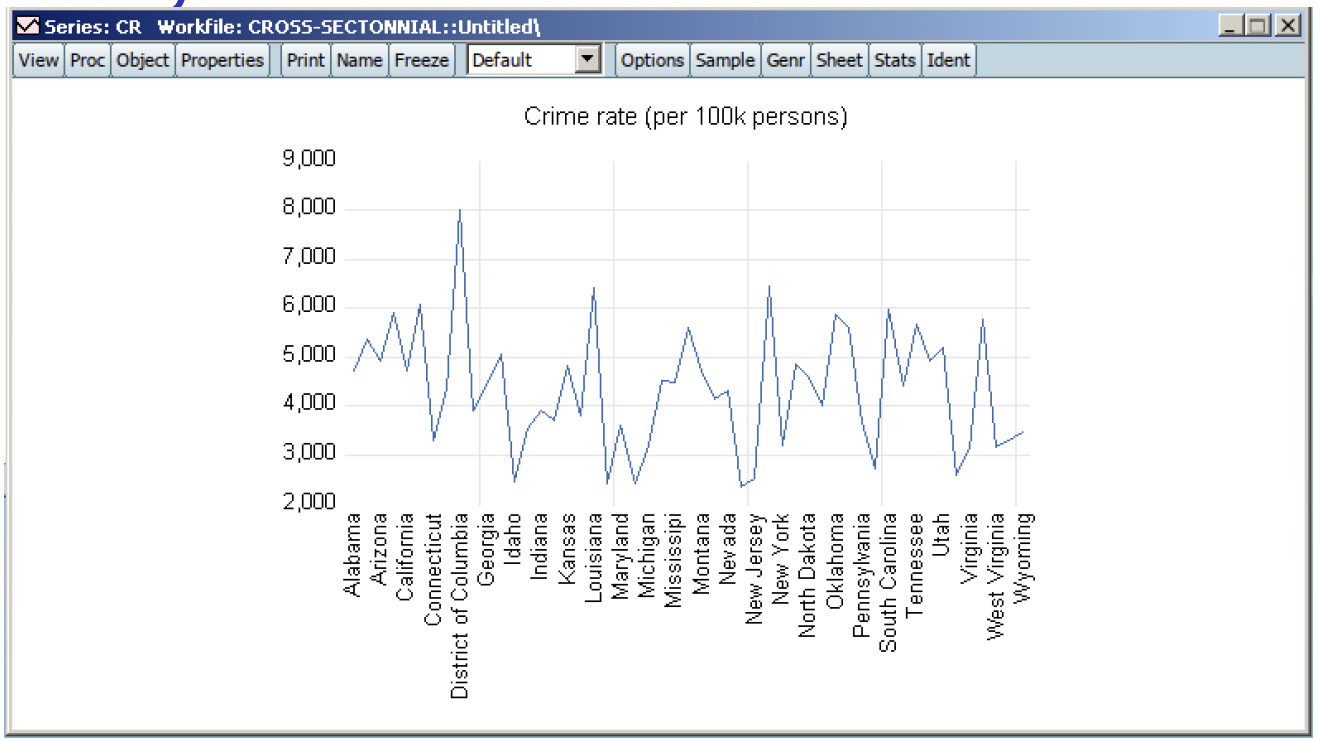
View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CR
Method: Least Squares
Date: 04/20/23 Time: 10:38
Sample: 1 51
Included observations: 51

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9305.063	3397.320	2.738942	0.0091
WEAPON	519.4750	578.0089	0.898732	0.3740
PEN	566.2472	339.1168	1.669771	0.1026
POLICE	-24.33697	15.53351	-1.566740	0.1249
POLICE^2	0.031888	0.016961	1.880048	0.0672
POVERTY	-14894.33	12670.58	-1.175506	0.2466
AFR*POLICE	-26.36409	12.39949	-2.126224	0.0396
AFR*POVERTY	104588.9	38956.39	2.684769	0.0104
HISP*POLICE	-30.94582	20.42093	-1.515397	0.1373
HISP*POVERTY	122518.3	63659.46	1.924590	0.0612

R-squared	0.445410	Mean dependent var	4359.824
Adjusted R-squared	0.323670	S.D. dependent var	1254.463
S.E. of regression	1031.661	Akaike info criterion	16.88963
Sum squared resid	43637281	Schwarz criterion	17.26842
Log likelihood	-420.6856	Hannan-Quinn criter.	17.03438
F-statistic	3.658718	Durbin-Watson stat	2.254864
Prob(F-statistic)	0.001975		

Crime rate by state in 2021



Conclusion

- States very different (population, police presence)
- Strong effects from Poverty, Police, African, Hispanic

Time-series Data

Variables	Unit	Expected sign
NB-OFFICERS	number of police officers	?
PER-AFRICAN	% of african	+
PER-BACHELOR	% of bachelor	-
PER-HISPANIC	% of hispanic	+
PER-POVERTY	% of poor people	+
PER-URBAN	% of urban-living people	-
PER-18MORE	% of 18 y.o. people	+

Table: Time-series variables expected signs

Characteristics of the model

- Modelization of crime rate per year on given states (California, New York, Texas)
- We modelled (3):

$$RATE_t = \beta_0 + \beta_1 NBOFFICERS_t + \beta_2 PERAFRICAN_t + \beta_3 PERBACHELOR_t + \beta_4 PERHISPANIC_t + \beta_5 PERPOVERTY_t + \beta_6 PERURBAN_t + \beta_7 PER18MORE_t + u_t \quad (3)$$

- We then modelled (4):

$$D(RATE_t) = \gamma_0 + \gamma_1 D(NBOFFICERS_t) + \ldots + \gamma_7 D(PER18MORE_t) + v_t \quad (4)$$

Equation (3) results

Variables	California	New York	Texas
NB-OFFICERS	- (90%)	- (99%)	+ (99%)
PER-AFRICAN	+ (99%)	/	/
PER-BACHELOR	/	/	- (95%)
PER-HISPANIC	/	- (95%)	/
PER-POVERTY	+ (99%)	+ (95%)	+ (90%)
PER-URBAN	/	/	/
PER-18MORE	/	/	/
Observations	22	22	22
R^2	0.965	0.983	0.942
Adjusted R^2	0.947	0.974	0.912
F-statistic	55.08 (99%)	114.25 (99%)	32.26 (99%)

Table: Comparaison between 3 states in (3)

Equation (4) results

Variables	California	New York	Texas
D(NB-OFFICERS)	/	- (90%)	/
D(PER-AFRICAN)	+ (95%)	/	/
D(PER-BACHELOR)	/	/	- (99%)
D(PER-HISPANIC)	/	/	/
D(PER-POVERTY)	- (95%)	/	/
D(PER-URBAN)	/	/	/
D(PER-18MORE)	/	/	/
Observations	21	21	21
R^2	0.633	0.380	0.702
Adjusted R^2	0.476	0.046	0.0.541
F-statistic	4.03 (95%)	1.14	4.37 (99%)

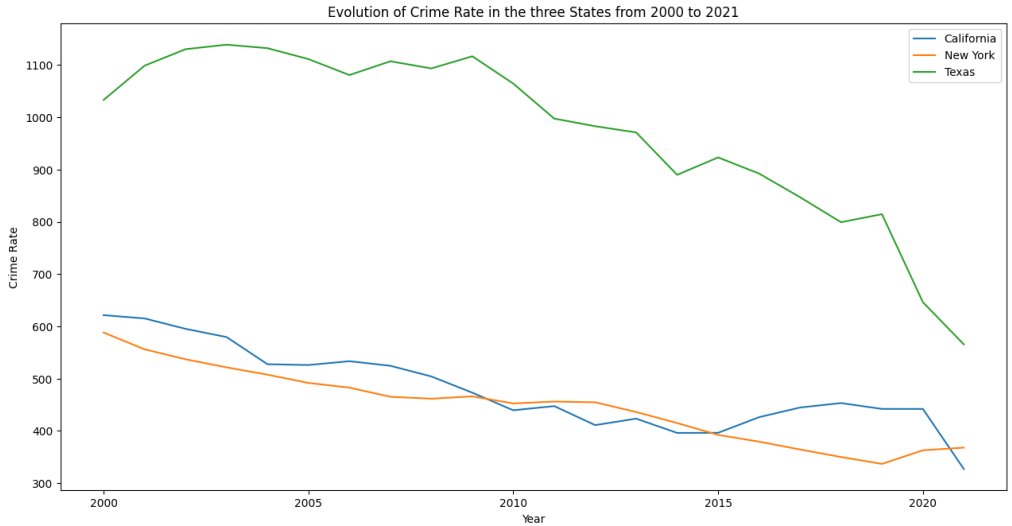
Table: Comparaison between 3 states in (4)

Tests on (3)

	California	**New York**	**Texas**
R^2	Very High	Very High	Very High
Normality in the error ?	Yes	Yes	Yes
Homoskedasticity	Yes	Yes	No (90%)
Autocorrelation	Yes (90%)	Yes (90%)	No

Table: Comparaison tests between 3 states in (3)

Crime rate for 3 states by time



Conclusion

- The three states have different crime rates through time
- Strong effect from percentage of population living under the line of poverty

Panel Data

Variables	Unit	Expected sign
AFRICAN-AMERICAN-POP	% of african	+
BACHELOR-POP	% of bachelor	-
DEATH-PENALTY	dummy	?
POLICE-OFFICERS	number per 100k people	?
UNDER-POVERTY-LINE	% of poor people	+
WEAPON-ALLOWED	dummy	+

Table: Panel variables expected signs

Panel equation

- We modelled (5):

$$\begin{aligned} cr_{it} = \beta_0 + \beta_1 AFRICANAMERICANPOP_{it} \\ + \beta_2 BACHELORPOP_{it} + \beta_3 DEATHPENALTY_{it} \\ + \beta_4 POLICEOFFICERS_{it} + \beta_5 UNDERPOVERTYLINE_{it} \\ + \beta_6 WEAPONALLOWED_{it} + u_{it} \end{aligned} \tag{5}$$

Equation (5) results

Equation: UNTITLED Workfile: PANEL::Panel\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CRIME_RATE__PER_100K_PERSONS_
Method: Least Squares
Date: 04/22/23 Time: 22:36
Sample: 1 561
Included observations: 561

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-195.9629	74.77677	-2.620638	0.0090
AFRICAN_AMERICAN_POP	2.029394	0.732812	2.769323	0.0058
BACHELOR_POP	9.145953	1.451692	6.300200	0.0000
DEATH_PENALTY	10.62312	14.90854	0.712552	0.4764
POLICE_OFFICERS_PER_100K_PER...	0.196837	0.066137	2.976200	0.0030
UNDER_POVERTY_LINE	21.77029	2.521838	8.632708	0.0000
WEAPON_ALLOWED	-20.05978	23.77081	-0.843883	0.3991

R-squared	0.275128	Mean dependent var	384.9576
Adjusted R-squared	0.267277	S.D. dependent var	173.9140
S.E. of regression	148.8690	Akaike info criterion	12.85641
Sum squared resid	12277742	Schwarz criterion	12.91044
Log likelihood	-3599.223	Hannan-Quinn criter.	12.87750
F-statistic	35.04549	Durbin-Watson stat	1.933676
Prob(F-statistic)	0.000000		

Sign comparaison

Variables	Unit	Expected sign	Sign
AFRICAN-AMERICAN-POP	% of african	+	+
BACHELOR-POP	% of bachelor	-	+
DEATH-PENALTY	dummy	?	/
POLICE-OFFICERS	number per 100k people	?	+
UNDER-POVERTY-LINE	% of poor people	+	+
WEAPON-ALLOWED	dummy	+	/

Table: Equation (5) coefficients signs

Tests

Equation: UNTITLED Workfile: PANEL::Panel\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	11.72523	Prob. F(6,554)	0.0000
Obs*R-squared	63.21299	Prob. Chi-Square(6)	0.0000
Scaled explained SS	102.1556	Prob. Chi-Square(6)	0.0000

Equation: UNTITLED Workfile: PANEL::Panel\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

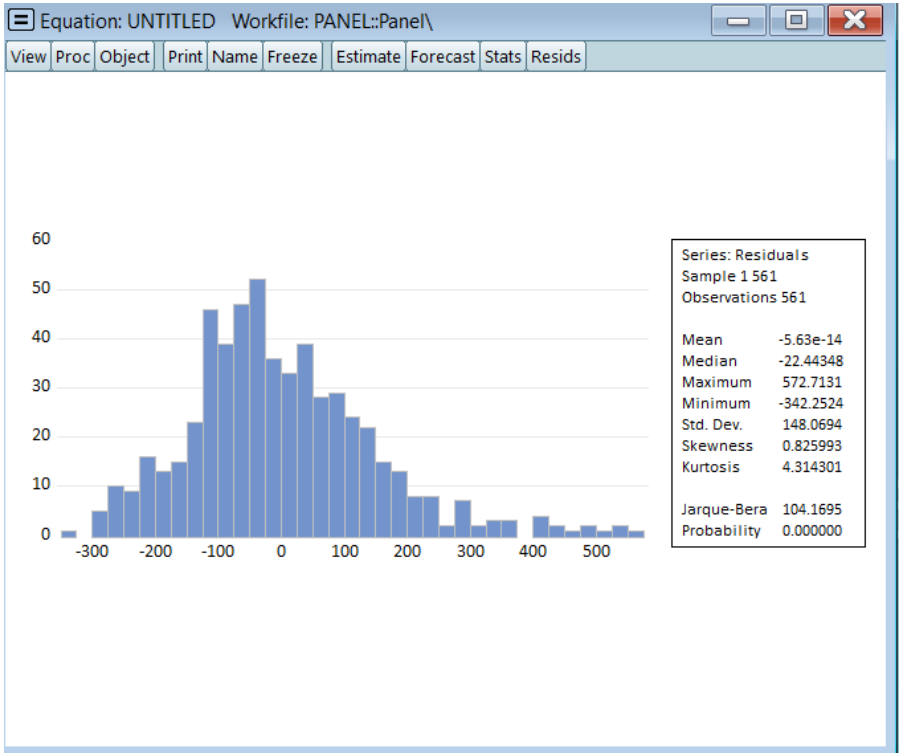
Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 2 lags

F-statistic	1.357983	Prob. F(2,552)	0.2580
Obs*R-squared	2.746734	Prob. Chi-Square(2)	0.2533

Interpretation

- Heteroskedasticity
- No autocorrelation

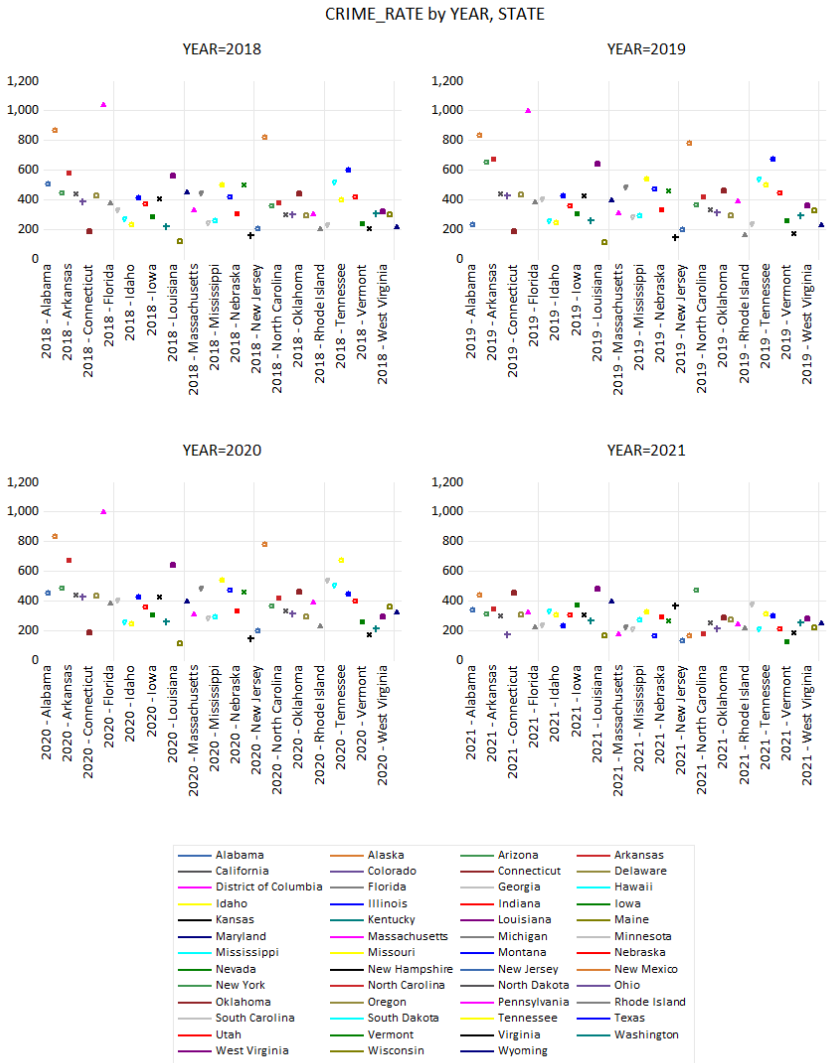
Residuals



Interpretation

- No normality in the errors

Crime rate for each state by time



Conclusion

- Different models for different interpretation
- Strong effect from Poverty, Minorities and Police
- Take into account the U.S. particularity : death penalty, Second Amendement

References

Bruce Western (2004)
Crime, punishment, and American inequality
Social inequality, 771–796.

Joe R. Feagin (2007)
Handbook of the sociology of racial and ethnic relations
Springer

Samuel Walker, Cassia Spohn and Miriam DeLone (1996)
The Color of Justice, Race, Ethnicity and Crime in America

Questions? Comments?