What drives crime in the U.S.?

Hiba Bousarsar Samy Mokrani Baptiste Llamas

École Nationale de l'Aviation Civile

May 10, 2023

<日 > < 部 > < 注 > < 注 > ご = の < @</p>

ENAC

Group 1 PREV Empirical Project ime-series Results

Panel Result

Conclusion

Presentation Overview

1 Introduction

- **2** Cross-sectionnal Results
- **3** Time-series Results
- 4 Panel Results

5 Conclusion

ENAC

Group 1 PREV Empirical Project ime-series Results

Panel Results

Conclusion

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

ENAC

Group 1

Time-series Results

Panel Result

Conclusion

Data Collecting

- US Bank Data Base
- US Government Data Base

ENAC

Group 1 PREV Empirical Project

< 口 > < 同 >

A B > A B >

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):

 $cr_{i} = \beta_{0} + \beta_{1}afr_{i} + \beta_{2}hisp_{i} + \beta_{3}white_{i} + \beta_{4}pen_{i} + \beta_{5}police_{i} + \beta_{6}poverty_{i} + \beta_{7}weapon_{i} + \beta_{8}bachelor_{i} + u_{i}$ (1)

ENAC

Group 1

lime-series Results

Panel Result

Conclusion

Equation (1) results

🗖 Equatio	on: EQ01	L Wo	rkfile:	CROSS-	SECTONN	IIAL::Un	titled\		_ 🗆 ×
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									
۷	ariable		Coe	efficient	Std. Ei	rror t	Statis	tic	Prob.
V PC W BAC	C AFR HISP VHITE PEN OLICE VERTY EAPON CHELOR		30 -27 -19 -45 69 -3.6 25 69 65	53.974 73.531 12.617 01.350 7.5634 304547 533.51 9.3395 30.984	3030.0 2548.8 2160.1 1685.6 370.26 2.0138 8662.1 571.41 3613.7	034 1 310 -1 118 -0 530 -2 508 1 394 -1 132 2 125 1 741 1	.0079 .0881 .8854 .6704 .8839 .7898 .9477 .2238 .8072	01 67 23 26 78 40 17 79 64	0.3193 0.2827 0.3810 0.0107 0.0665 0.0807 0.0052 0.2278 0.0779
R-square Adjusted S.E. of reg Sum squ Log likelil F-statistic Prob(F-st	d R-squar gression ared res hood : atistic)	ed id	0.4 0.3 10 433 -42 4.2 0.0	48744 343743 16.236 374925 0.5318 273708 00803	Mean de S.D. dep Akaike ir Schwarz Hannan- Durbin-V	pendent endent v fo criterio criterion Quinn cr Vatson s	var ar on iter. tat	43 12 16 17 16 2.4	59.824 54.463 84438 18530 97466 487297

▲ロト▲御ト▲臣ト▲臣ト 臣 の父(で

ENAC

Group 1

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

ENAC

Group 1

ime-series Results

Panel Results

Conclusion

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

ENAC

Group 1

Introduction



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)	0.3286
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

Introduction

Cross-sectionnal Results

Time-series Results

Panel Results

Conclusion

Tests



Figure: Residuals

ENAC

Group 1

Second equation

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

 $cr_{i} = \beta_{0} + \beta_{1} weapon_{i} + \beta_{2} pen_{i} + \beta_{3} police_{i} + \beta_{4} police_{i}^{2} + \beta_{5} poverty_{i} + \beta_{6} afr * police_{i} + \beta_{7} afr * poverty_{i} + \beta_{8} hisp * police_{i} + \beta_{9} hisp * poverty_{i} + u_{i}$

(2)

lime-series Results

Panel Result

Equation (2) results

Equation:	EQ02	Wo	rkfile:	CROSS-	SECTONN	IAL::Un	titled\		_ 🗆 🗙
View Proc Ob	ject	Print	Name	Freeze	Estimate	Forecast	Stats	Resids	
Dependent V Method: Leas Date: 04/20/2 Sample: 1 51 Included obs	′ariab st Squ 23 Ti l ervati	le: CF Jares me: 1 ons: !	₹ 0:38 51						
Varia	ble		Coe	fficient	Std. E	ror t	Statis	tic	Prob.
C WEAF POLI POLIC POVE AFR*PO HISP*PO HISP*PO	PON CE RTY DLICE VERT OLICI VERT	Y T	93 51 -24 -24 -14 -26 10 -30 12	05.063 9.4750 6.2472 .33697 I31888 894.33 .36409 4588.9 .94582 2518.3	3397.3 578.00 339.11 15.533 0.0169 12670 12.399 38956 20.420 63659	320 2 089 0 68 1 351 -1 361 1 363 -1 364 -1 365 -1 361 1 363 -1 364 -1 37 -2 39 2 39 2 39 -2 39 -1 .46 1	.7389 .8987 .6697 .5667 .8800 .1755 .1262 .6847 .5153 .9245	42 32 71 40 48 06 24 69 97 90	0.0091 0.3740 0.1026 0.1249 0.0672 0.2466 0.0396 0.0104 0.1373 0.0612
R-squared Adjusted R-s S.E. of regres Sum square Log likelihoo F-statistic Prob(F-statis	quan ssion d resi d	ed d	0.4 0.3 10 436 -42 3.6 0.0	45410 23670 31.661 37281 0.6856 58718 01975	Mean de S.D. dep Akaike ir Schwarz Hannan- Durbin-V	pendent endent v criterion Quinn cr Vatson s	var ar on iter. tat	43 12 16 17 17 2.3	59.824 54.463 .88963 .26842 .03438 254864

▲ロト▲聞ト▲目ト▲目ト 目 のへで

Group 1 PREV Empirical Project Cross-sectionnal Results

Time-series Results

Panel Results

< □ > < 同 > < 三 > .

< 3 D

Conclusion

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):

 $RATE_{t} = \beta_{0} + \beta_{1}NBOFFICERS_{t} + \beta_{2}PERAFRICAN_{t} + \beta_{3}PERBACHELOR_{t} + \beta_{4}PERHISPANIC_{t} + \beta_{5}PERPOVERTY_{t} + \beta_{6}PERURBAN_{t} + \beta_{7}PER18MORE_{t} + u_{t}$ (3)

• We then modelled (4): $D(RATE_t) = \gamma_0 + \gamma_1 D(NBOFFICERS_t) + ... + \gamma_7 D(PER18MORE_t) + v_t$ (4)

Equation (3) results

Variables	California	New York	Texas
NB-OFFICERS	- (90%)	- (99%)	+ (99%)
PER-AFRICAN	+ (99%)	/	/
PER-BACHELOR	/	/	- (95%)
PER-HISPANIC	/	<mark>-</mark> (95%)	/
PER-POVERTY	<mark>+ (99%)</mark>	<mark>+ (95%)</mark>	+ (90%)
PER-URBAN	/	/	/
PER-18MORE	/	/	/
Observations	22	22	22
R^2	0.965	0.983	0.942
Adjusted <i>R</i> 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 <i>R</i> 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
<i>R</i> ² Normality in the error ? Homoskedasticity Autocorrelation	Very High Yes Yes Yes (90%)	Very High Yes Yes Yes (90%)	Very High Yes No (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):

 $cr_{it} = \beta_0 + \beta_1 AFRICANAMERICANPOP_{it} + \beta_2 BACHELORPOP_{it} + \beta_3 DEATHPENALTY_{it} + \beta_4 POLICEOFFICERS_{it} + \beta_5 UNDERPOVERTYLINE_{it} + \beta_6 WEAPONALLOWED_{it} + u_{it}$ (5)

Equation (5) results

C Object Print Name Freeze Estimate Forecast Stats Resids Dependent Variable: CRIME_RATE_PER_100K_PERSONS_ Method: Less Variable Cast Stats Resids Date: 04/22/23 Time: 22.36 Sample: 1561 Included observations: 561 Variable Coefficient Std. Error 1-Statistic Prob. AFRICAN_MERICAN_POP 9145953 1.451692 6.300200 0.0000 BACHELOR_POP 9.145953 1.451692 6.300200 0.0000 UNDEETH_PENALTY POLICEOFFICERS_PER 100K_PER_0 0.19837 0.066137 2.976200 0.0300 UNDER_POVERTY_UNE 2.177029 2.27183 Mean dependent var 3.84.9576 Adjusted R-squared 0.275128 Mean dependent var 3.84.9576 Adjusted R-squared 0.275128 Mean dependent var 1.73140 S.E. of regression 148.8569 Akaike info criterion 1.2.87541 Usin squared resid 1.2277742	Equation: UNTITLED Workfile: PANEL::	Panel\			- 🗆 💌
Dependent Variable: CRIME_RATE_PER_100K_PERSONS_ Method. Least Squares Method. Least Squares Date. 04/22/32 Time: 22.36 Sample: 1561 Included observations: 561 Variable Coefficient Std. Error t-Statistic Prob. ARICAN_MERICAN_POP 2.029394 0.732812 2.769323 0.0090 ARICAN_MERICAN_POP 2.029394 0.732812 2.769323 0.0058 BACHELOR_POP 9.145953 1.451692 6.00200 0.0000 DEATH_PERNALTY 10.66337 0.976200 0.030 UNDER_POVERTY_LINE 21.77029 2.521838 8.632708 0.0091 WEAPON_ALLOWED -20.05978 2.377081 -0.84883 0.3991 R-squared 0.267277 S.D. dependent var 384.9576 Adjusted R-squared 0.267277 S.D. dependent var 12.87641 Sum squared resid 1.2277742 S.Marar criterion 12.87641 Sum squared resid 1.2277742 S.Howarz criterion 12.87641 Sum squared resid 1.54549 2.46541 12.83766	/iew Proc Object Print Name Freeze Estima	ate Forecast St	ats Resids		
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 0.0050 BACHELOR_POP 9.145953 1.451692 6.00200 0.0000 DEATH_PENALTY 10.62312 14.90854 0.712552 0.4764 DULCE_OFFLCERS_PER_100K_PER. 0.196837 0.06137 2.976200 0.0030 UNDER_POVERTY_LINE 2.177029 2.521838 8.632708 0.0000 WEAPON_ALLOWED -20.05978 23.77081 0.84383 3.3991 R-squared 0.275128 Mean dependent var 384.9576 SL: of regression 148.9600 Akaike info criterion 12.81750 Sum squared resid 1.2277742 Schwarz criterion 12.81750 Log likelihood -3599.223 Haman-Quinn criterion 12.81750 - statistic 0.34549 Durburn Valson stat 1.933676	Dependent Variable: CRIME_RATE_PER Method: Least Squares Date: 04/22/23 Time: 22:36 Sample: 1 561 Included observations: 561	_100K_PERS	ONS_		
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 6.00200 0.0000 DEATH_PENALTY 10.62312 14.90854 0.712552 0.4764 DULCE_OFICERS_PER_100K_PER. 10.96837 0.976200 0.0030 WEAPON_ALLOWED -20.05978 2.3.77081 0.984383 0.3991 R-squared 0.275128 Mean dependent var 73.8149 3.9971 SE of regression 148.9660 Akaike info criterion 12.81750 Sum squared resid 1.2277742 Schwarz criterion 12.81750 Sum squared resid 1.227742 349479 12.81750 Sum squared resid 1.227742 340400 nor riterion 12.81750 Sum squared resid 1.227742 3404000 nor stat 1.238766	Variable	Coefficient	Std. Error	t-Statistic	Prob.
R-squared 0.275128 Mean dependent var 384.9576 Adjusted R-squared 0.267277 S.D. dependent var 172.9140 S.E. of regression 148.0690 Akaike info criterion 12.85641 Sum squared resid 12277742 Schwarz criterion 12.91044 Log likelihood -3599.223 Hannan-Quinn criter. 12.87750 - statistic 35.04549 Durbin-Vvalson stat 1.933676	C AFRICAL AMERICAN_POP BACHELOR, POP DEATH, PENALTY POLICE_OFFICERS_PER_100K_PER UNDER_POVERTY_LINE WEAPON_ALLOWED	-195.9629 2.029394 9.145953 10.62312 0.196837 21.77029 -20.05978	74.77677 0.732812 1.451692 14.90854 0.066137 2.521838 23.77081	-2.620638 2.769323 6.300200 0.712552 2.976200 8.632708 -0.843883	0.0090 0.0058 0.0000 0.4764 0.0030 0.0000 0.3991
Prob(F-statistic) 0.000000	R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood F-statistic Prob(F-statistic)	0.275128 0.267277 148.8690 12277742 -3599.223 35.04549 0.000000	Mean dependent var S.D. dependent var Akaike info criterion Schwarz criterion Hannan-Quinn criter. Durbin-Watson stat		384.9576 173.9140 12.85641 12.91044 12.87750 1.933676

▲ロト▲御ト▲臣ト▲臣ト 臣 の父(で

ENAC

Group 1

Time-series Results

Panel Results

Conclusion

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

▲ロト▲御▶▲臣▶▲臣▶ 臣 のへで

ENAC

Group 1

ime-series Results

Panel Results

Tests

Equation: UNTITLED Workfile: PANEL::Pa	nel\			×
View Proc Object Print Name Freeze Estimate	Forecast St	ats Resids		
Heteroskedasticity Test: Breusch-Pagan-Go Null hypothesis: Homoskedasticity	dfrey			^
F-statistic Obs*R-squared Scaled explained SS	11.72523 63.21299 102.1556	Prob. F(6,554) Prob. Chi-Square(6) Prob. Chi-Square(6)	0.0000 0.0000 0.0000	
Equation: UNTITLED Workfile: PANEL::Pa	inel\	ate Deside		×
Equation: UNTITLED Workfile: PANEL:-Pa View Proc Object Print Name Freeze Estimat Breusch-Godfrey Serial Correlation LM Test Null hypothesis: No serial correlation at up to	rnel\ e Forecast St 2 lags	ats Resids		×

Interpretation

- Heteroskedasticity
- No autocorrelation

ime-series Results

Panel Results

Residuals



Interpretation

No normality in the errors

・ロト・母・・ヨト・ヨー ろくぐ

Group 1

Cross-sectionnal Result

lime-series Results

Panel Results

Conclusion

Crime rate for each state by time

CRIME_RATE by YEAR, STATE



- Alabama	- Alaska	- Arizona	- Arkenses
California		Connecticut	Delaware
- District of Columbia		Georgia	Hawaii
Idatio	- Illinois	Indiana	lowa
	- Kernucky	- Louisiana	Maine
Maryland	Massachusetts	Michigan	Minnesota
- Mississippi	- Missouri	- Montana	Nebraska
	New Hampshire	New Jersey	New Mexico
New York	North Carolina	North Dakota	Ohio
Oklahoma	- Oregon	Pennsylvania	Rhode Island
	South Dakota	Tennessee	Texas



ENAC

Group 1

< 口 > < 同

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?

<ロ> (四) (四) (三) (三) (三) (三)