

Hobart and William Smith Colleges

Police Funding

More Than a Response to Crime Rates?

Robert Tasker

ECON 304-01: Econometrics

Professor Damar

12/08/2020

Table of Contents

Table of Contents	2
1. Introduction.....	3
1.1 Research Question	3
1.2 Hypothesis and Associated Complications	3
1.3 Model and Associated Data.....	5
1.4 Results Summary	5
2. Literature Review.....	5
2.1 Discredited Bias.....	7
3. Data Description	8
3.1 Initial Overview	8
3.2 Summary Statistics	9
4. Model Results	12
4.1 The Model	12
4.2 Regression Results.....	14
4.3 Extension and Discussion.....	14
4.3.1 Findings.....	18
4.3.2 Heteroscedasticity	19
5. Conclusion	20

1. Introduction

1.1 Research Question

Police funding and policing activity has been a growing, controversial, topic of debate in sociological studies and progressive politics. On one hand, the US police force is said to work in favor of the people to help serve and protect the country from crime (e.g. Medved 2014). On the other hand, some disagree with this statement and argue that there exists ulterior motives and biases within US police funding and enforcement (e.g. Taylor and Okeowo 2020). In addition to this topic of debate, there exists the conflict of what factors cause this bias or perceived bias in police funding and enforcement. Answers to this question are quite vast and range from varying factors of demographics to population density of a respective state or location (e.g. Lum 2009). This paper presents a statistical overview of this topic with the objective of understanding whether police funding can be fully accounted for by crime rates.

1.2 Hypothesis and Associated Complications

Although this topic is quite common in progressive politics, there exists various theoretical and technical complications when justifying conclusions. The model being tested in this paper is one which holds police funding as a dependent variable, with crime rates as the main independent variable of interest. This model supports a hypothesis of police funding with some independent variables as statistically significant determinants, other than crime rates. Through this hypothesis, it is expected that police funding will not be unilaterally correlated with crime rates; the other independent variables are hypothesized to be statistically significant in the model while greatly increasing the coefficient of determination, R^2 , of the regressions being performed.

This hypothesis follows political and sociological theories of biased or problematic police funding and enforcement practices. Although this hypothesis largely agrees with these theories, it is possible that the hypothesis would not hold through theoretical economics. For this reason, some potential complications, that could cause this discrepancy, are to be looked into further in the following discussion.

The first potential complication in this model is reverse causality. Complications could theoretically arise in cases where police funding is actually a determinant of crime rates. This is something that could hold true in economic theory when rational people commit more crimes when policing is lower, due to a decrease in expected opportunity costs. Although this complication seems quite reasonable, it will be discredited to some extent by a strongly positive correlation between police funding and crime rates, along with the literature review presented on this issue in *2.1 Discredited Bias*.

A second, and more broadly applied, complication that could theoretically arise in this model, is the potential for independent variables to hold an unexpectedly small or large marginal effect on police funding. It is likely for some variables to have these limited or exaggerated effects on the dependent variable of police funding, due to theoretical variations between rational, self-interested, behavior of neoclassical economics, and purposeful, cognitive, behavior. In this sense, economic theory could fall short in accurately predicting the effects that variables, such as income and race, actually have on police funding. This paper aims to utilize economic theory on government spending while understanding and accounting for these potential shortcomings of neoclassical economics.

1.3 Model and Associated Data

As a brief introduction to the model and variable data used, most variables of interest will be derived from demographics or other factors similar and related to demographics or government funding. This is done to show that correlation between the dependent variable of police funding, and the independent variables, will be that of importance to the proposed hypothesis. While using these variables that are expected to well-account for police funding, it will thus become more feasible to use collinearity to speak on potential causality. Furthermore, some reliable sources of data are necessary in this analysis, therefore, all data is derived from US government reports by the FBI or US Census Bureau.

1.4 Results Summary

The results found in this paper support the presented hypothesis of the existence of statistically significant independent variables, other than crime rates. In the restricted model developed in *4. Model Results*, statistical significance is found in a model regression of police funding on crime rates, population density, education funding, and race. Although causality cannot be statistically defined by these model results, these results do act as a stepping stone towards a better understanding of the determinants of police funding in the modern day economic and political environment of the United States.

2. Literature Review

Many studies done on police funding and crime rates largely involve the study of deterrence. Namely, those along the lines of Vollaard and Koning (2009) and Georgakopoulos (2004) speak on the idea and the main purpose of policing and police funding as being the

deterrence of crime. Deterrence is an interesting point of view on this topic and involves a lot of ideas that are often regarded as being socio-economically problematic. These studies delve into the ideas of race and poverty to better understand the influence that these types of demographic factors have on deterrence-driven police funding and crime rates.

It is stated in such studies as Georgakopoulos (2004) and Ayres (2000) that deterrence is what drives changes in policing efforts and is inevitably what causes the unintentional over-policing of minorities and those who are impoverished. This process of fund allocation is brought about by the usage of ideas such as "optimal allocation of enforcement" and thus "intensity of policing" (Georgakopoulos p.22). When these ideas are applied to the concepts and theories of opportunity costs, it is said that there is an indirect connection between minorities and those who are impoverished, and unintentional race profiling and deterrence efforts.

Although these reviews on police funding and crime rates are quite thorough in their analyses, it can also be said that they disregard the idea of other independent variables besides crime rates, like race and racial profiling. These past studies do elaborate on these other socio-economic and political factors but conclude that they are simply collinear as crime is directly defined by these factors and fully accounts for their effects. According to other research such as Jackson and Carroll (1981), this is not the case.

In the Jackson and Carroll (1981) study of "Race and the War on Crime," it is exhaustively presented that racial composition and civil unrest are largely independent of the effect that crime has on police funding and expenditure. This is a framework that is maintained in this research paper as well. An argument made in Jackson and Carroll (1981) is that most other researchers have overlooked the fact that minority composition and police funding per capita are positively curvilinear. In other words, this relationship holds an increasing marginal

effect as determinants, like racial composition, increase on a per capita, state by state, basis. This is something that was tested for when creating the initial model in this paper, and, unfortunately, the data is unable to support these claims. Although this is the case, it does not discredit Jackson and Carroll's reasoning when accounting for the large time period discrepancy. This paper will utilize a framework similar to Jackson and Carroll's, but opt to do so with linearly correlated variables.

The research evidence being presented does provide limited statically significant support for the hypotheses, and, unfortunately, this is not enough to conclude on 40 years later. With a continuously changing economic, political, and social environment within the US, many factors come into play that must be examined past racial civil unrest. This research paper is tasked with defining this correlation in more recent terms, and further providing evidence to support the claims being made in Jackson and Carroll (1981).

2.1 Discredited Bias

To conclude this literature review, it is helpful to also clarify the stance and hypothesis of this topic and to discredit any issues that could potentially arise or notions that could be used to discredit this research. A large issue that many studies on this topic face is cited as being the potential existence of simultaneity bias or reverse causality. With the presence of studies which describe crime rates as a dependent variable, it can be seen that there is no defined causation of crime rates from changes in policing or police funding. Studies such as Levitt (1997), Levitt (2002), and Samuel Cameron (1988), go as far to say that there is no evidence of the existence of an impact of policing on crime rates. This lack of statistically significant evidence, of reverse causality between police funding and crime rates, helps to discredit the possibility of reverse

causality or simultaneity bias in the model. These studies are thus a justification for this paper's usage of police funding as the dependent variable and crime as the main independent variable.

3. Data Description

3.1 Initial Overview

The main dependent variable of interest for this data set is POLICE. This variable captures police funding on a per capita, state-by-state basis. The data used for the analysis of this variable comes from the US Census Bureau, Annual Survey of State and Local Government Finances. This source was used to find accurate information on police funding in USD per capita, based on state and local government survey responses.

The main independent variable of interest for this data set is CRIME. This independent variable captures crime rates on a per capita, state-by-state basis. The data used in the analysis of this variable comes from the FBI's crime database. Data on crime rates is based on conviction and arrest rates reported by each state in the number of violent crimes convicted per 10,000 population in that respective state.

The sample period being used in this data set is 2017. The US population in 2017 was 325,719,178 people and this is the value being used for the overarching number of observations in the sample. For data analysis purposes, 51 observations are used to represent the 50 US states and the District of Columbia. 2017 is used as the period being analyzed in this paper due to a lack of more recent data reported by the US Census Bureau. This time period was also held consistent wherever possible during data acquisition.

3.2 Summary Statistics

Table 3.1: POLICE Summary Statistics.

Percentiles		Summary	
1%	186.2024	Observations	51
5%	217.1665		
10%	239.4804		
25%	261.4439	Mean	333.2574
		Std. Dev.	112.435
50%	311.41	Variance	12641.64
75%	376.6649	Skewness	2.80165
90%	429.6744	Kurtosis	14.76516
95%	493.7982		
99%	910.3208		

Table 3.2: CRIME Summary Statistics.

Percentiles		Summary	
1%	121	Observations	51
5%	198.7		
10%	226.4		
25%	250.6	Mean	389.4275
		Std. Dev.	171.5953
50%	358	Variance	29444.96
75%	456.2	Skewness	1.349132
90%	555.9	Kurtosis	5.398225
95%	783.5		
99%	1004.9		

Table 3.3: POLICE when CRIME is more than or equal to its mean of 389.4275.

Percentiles		Summary	
1%	200.1456	Observations	22
5%	223.5696		
10%	240.7138		
25%	261.4439		
50%	316.7734	Mean	359.3313
75%	413.1419	Std. Dev.	149.5868
90%	486.5561	Variance	22376.21
95%	493.7982	Skewness	2.327216
99%	910.3208	Kurtosis	9.412471

Table 3.4: POLICE when CRIME is less than or equal to its mean of 389.4275.

Percentiles		Summary	
1%	186.2024	Observations	29
5%	217.1665		
10%	232.4864		
25%	262.1051		
50%	311.41	Mean	313.4772
75%	338.3626	Std. Dev.	69.66207
90%	400.7735	Variance	4852.804
95%	429.6744	Skewness	0.910287
99%	529.9101	Kurtosis	4.632944

Table 3.5: Regression of POLICE and CRIME

police	Coefficient	Std. Error	p-value
CRIME	0.3320451	0.0806956	0.000
Constant	203.9499	34.28579	0.000
N	51		
R ²	0.2568		

Tables 3.3 and 3.4 shown above, present potential data patterns from the Stata analysis that was executed. It can be seen that the mean and median values of POLICE generally increase when the mean of CRIME increases. Although this can be seen, the change in the median and the change in the mean are also largely different. This effect shows that the data is becoming more and more skewed as CRIME increases. One common cause of this is the existence of few very large values in the data set. Although these large values may act as outliers, it would be unjustified to attempt to drop an entire state from the data set.

Along with this analysis and view of the correlation and pattern between the two variables, the p-value shown in Table 3.5 holds a value of 0.000 which also shows that the null hypothesis, of CRIME having no effect on POLICE ($\beta = 0$), is rejected. This helps to solidify the existence of a correlation between police funding and crime, but likely does not paint the whole picture. The regression is statistically significant but has an R² of 0.2568, thus, further data is necessary for this evaluation. For further statistical analysis of this data; race, income (poverty rates), and population are some examples of independent variables that will be helpful to analyze. As hypothesized, these additional independent variables should be statistically significant and largely increase the value of R² of the regression.

4. Model Results

4.1 The Model

$$POLICE = \beta_1 + \beta_2 CRIME + \beta_3 DEM + \beta_4 POP_DENS + \beta_5 INCOME + \beta_6 EDUCATION + \beta_7 WHITE + e$$

As mentioned in the data description section of this paper, the main dependent variable of interest for this data set is POLICE. This variable captures police funding in USD per capita, on a state-by-state basis. The independent variables used in this model are CRIME, DEM, POP_DENS, INCOME, EDUCATION, and WHITE. With the usage of these variables, the model is estimated via OLS (ordinary least squares).

The independent variable of CRIME captures crime rates on a per capita, state-by-state basis. This variable is the most relevant independent variable to the research question as talked about in the data description section of this paper. Thus, the main coefficient of interest is β_2 and given my model, it is expected that the coefficient estimate, b_2 , is positive. In other words, an increase in crime rates is hypothesized to increase police funding. This being said, the research topic expands on this idea by rejecting the claim that CRIME exclusively explains POLICE and thus complicating this definition of a main variable of interest.

To begin describing the independent variables used in this model, the independent variable DEM captures democratic voters per capita, on a state-by-state basis. This is a political affiliation, indicator variable that represents the population percentage of people affiliated with each party. Political affiliation values are taken from the recent 2020 presidential election results. Although this is an inconsistency from the discussion in 3.1 Initial Overview, it should be noted

that this was intentional. It is expected for 2020 political affiliation data to more accurately describe 2017 voter sentiment than that of 2016.

The independent variable POP_DENS captures the population density on a state-by-state basis. POP_DENS is calculated by using a state's population divided by that respective state's total land area in square miles. Total land area is the state's total surface area minus the area of any bodies of water within that state.

The next independent variable, INCOME, captures household income on a per capita, state-by-state basis. This variable is expected to largely determine the tax base and is also helpful in understanding demographics. It is expected for higher income levels to increase police funding due to this tax base effect.

Similarly, the independent variable EDUCATION captures that respective state's education expenditure on a per capita, state-by-state basis. This variable is used to provide insight into supplementary government funding, and acts as a proxy for the education or intelligence level of those who live in that respective state. This is important to include in the model to help provide comparable government funding data and include any marginal effect that education or intelligence has on crime rates and police funding. Without this variable, there would likely also be omitted variable bias on CRIME because of this duality.

The final independent variable in this model is WHITE. WHITE is an indicator variable which accounts for the overall demographic of each state. This value is a relative percentage to the reference group of all other races. The reference group is used here to represent population demographics of racial minorities.

4.2 Regression Results

Table 4.1: Initial regression results.

police	Coefficient	Std. Error	p-value
CRIME	0.1616101	.0567064	0.007
POP_DENS	26.61924	6.488685	0.000
DEM	91.82879	101.0388	0.368
INCOME	1.366348	1.04262	0.197
EDUCATION	10.2579	2.776628	0.001
WHITE	-119.5457	60.04911	0.053
Constant	89.09041	93.50014	0.346
N	51		
R ²	0.8293		

The regression results in Table 4.1 above show that crime rates do in fact help explain police funding, but not entirely. This result is as expected of POP_DENS and EDUCATION both agreeing with the overarching hypothesis by holding a positive correlation with POLICE. This being said, a surprising result is the rejection of the null of hypothesis of $\beta_i = 0$ (at $\alpha = 5\%$) for DEM, INCOME, and WHITE. This is a clear indication of correlation among these variables and is something that must be looked into further in Tables 4.2-4.5 of *4.3 Extension and Discussion*.

4.3 Extension and Discussion

Table 4.2: Pairwise correlation results.

	DEM	INCOME	WHITE
DEM	1.0000		
INCOME	0.6703	1.0000	
WHITE	-0.4821	-0.3000	1.0000

As can be seen in Table 4.2 above, there is correlation among these variables. It can be said that the largest correlation is between INCOME, DEM and WHITE. This is likely due to these variables being key factors of demographics. To remedy this correlation, one solution is to create a more restricted model by removing DEM. Through various manipulations of the model, it is evident that two of the three correlated variables must be removed to remedy the collinearity in the independent variables. These results can be seen in Tables 4.3-4.5 below.

Table 4.3: Regression results while preserving DEM*.

police	Coefficient	Std. Error	p-value
CRIME	.2091285	.0510032	0.000
POP_DENS	23.81319	6.660625	0.001
DEM*	253.7979	84.12377	0.004
EDUCATION	10.66283	2.575368	0.000
Constant	-10.35177	47.85173	0.830
N	51		
R ²	0.8025		

Table 4.4: Regression results while preserving INCOME*.

police	Coefficient	Std. Error	p-value
CRIME	.2187536	.0518369	0.000
POP_DENS	28.8682	6.141658	0.000
INCOME*	2.719787	.9557215	0.007
EDUCATION	9.44704	2.821963	0.002
Constant	-43.19855	57.0328	0.453
N	51		
R ²	0.7989		

Table 4.5: Regression results while preserving WHITE*.

police	Coefficient	Std. Error	p-value
CRIME	.1219043	.053806	0.028
POP_DENS	31.6212	5.733976	0.000
WHITE*	-180.2337	50.5757	0.001
EDUCATION	13.40291	2.295895	0.000
Constant	233.1595	57.64821	0.000
N	51		
R ²	0.8146		

Deciding which of these restricted models is best, is an important topic to this discussion. For this reason, all three tables are provided above. It can be argued that all three of these models are helpful in understanding the results, but for this research topic, at first glance, it can be identified that Table 5 best responds to the hypothesis. By removing DEM and INCOME, the estimated coefficient of CRIME, b_2 , is minimized and the p-value of CRIME is maximized. This helps to discredit CRIME from being the main variable of interest and furthers the hypothesis of the necessity for other statistically significant variables in explaining POLICE.

As an even larger, and potentially more important, reason for keeping WHITE and removing DEM and INCOME, this resulting regression helps limit omitted variable bias on CRIME. If all three variables were removed from the model, there would be an omitted variable in the error term that would be correlated with both, CRIME and POLICE. This omitted variable can be accounted for by using a proxy that accurately describes the DEMOGRAPHIC of each respective state. Using DEMOGRAPHIC as an independent variable is not quantitatively feasible, thus WHITE is instead used as a proxy to account for this omitted variable bias. Although it was expected that the combination of DEM, INCOME, and WHITE would be

optimal, this is not the case as can be seen in the statistically insignificant results from DEM and INCOME.

To add some insight on the economic causes of this omitted variable bias, and to clarify the collinearity, it must be understood why DEMOGRAPHIC plays the role that it does. Intuitively, it is seen in everyday life, and across varying academic disciplines, that factors such as age, race, sex, employment, income, political affiliation, religious affiliation, etc. have large effects on personal opportunities, values, and beliefs. The intertwining channels of these cause and effect relationships is something that largely affects crime rates yet the culmination is not easily measureable with numerical values.

DEMOGRAPHIC is also largely collinear with POLICE as varying populations have varying effects on local and state government spending. For example, factors such as race, income, age, and employment all have large influences on the tax base which thus determines funding of local government institutions, including the police force. What makes this such a difficult variable to quantify is the multi-collinearity that these variables have with each other and DEMOGRAPHIC itself. As stated, the best solution found to account for this is using WHITE as a proxy for DEMOGRAPHIC because many factors such as race, income, employment, political affiliation, and religious affiliation vary largely between WHITE and minority races and ethnicities such as BLACK, ASIAN, INDIAN, and HISPANIC. WHITE therefore acts as a proxy for all demographic variables, as well as income.

Furthermore, it can be seen in Table 6, below, that out of DEM, INCOME, and WHITE, WHITE is the only variable that has statistically significant collinearity with CRIME. Since WHITE is correlated with, both, CRIME and POLICE, the restricted model below is the best option as WHITE best accounts for the biased effect that DEMOGRAPHIC has on both, CRIME

and POLICE. This new, restricted, model has an R-squared value of 0.8146, as seen in Table 5.

This value is quite large for only having four variables in the model. The R-squared value shows that the model explains roughly 81.46% of the variation in POLICE.

$$POLICE = \beta_1 + \beta_2 CRIME + \beta_3 POP_DENS + \beta_6 EDUCATION + \beta_7 WHITE + e$$

Table 4.6: Regression of DEM, INCOME, and WHITE on CRIME.

crime	Coefficient	Std. Error	p-value
DEM	80.30058	261.5154	0.760
INCOME	-3.821116	2.792578	0.178
WHITE*	-558.2966	149.8506	0.001
Constant	962.3502	194.0675	0.000
N	51		
R ²	0.2740		

4.3.1 Findings

With the restricted model in hand, and all necessary regressions run, it is significantly less demanding to discuss these findings. The regression run in Table 4.5 shows the final regression results of this restricted model. Through these results it can be seen that CRIME, POP_DENS, WHITE (demographics), and EDUCATION are the statistically significant independent variables of the model. As expected, CRIME, POP_DENS and EDUCATION have a positive correlation with POLICE.

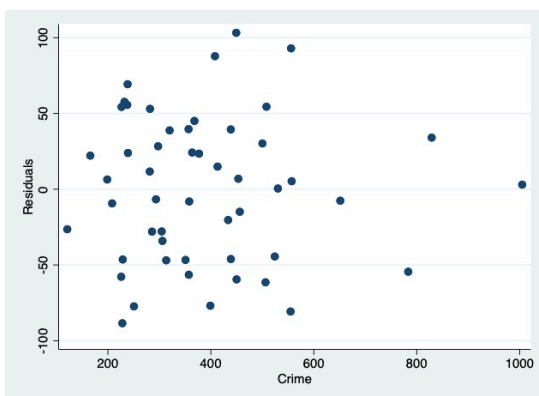
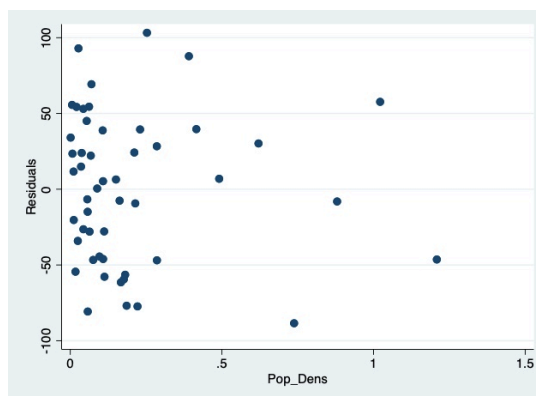
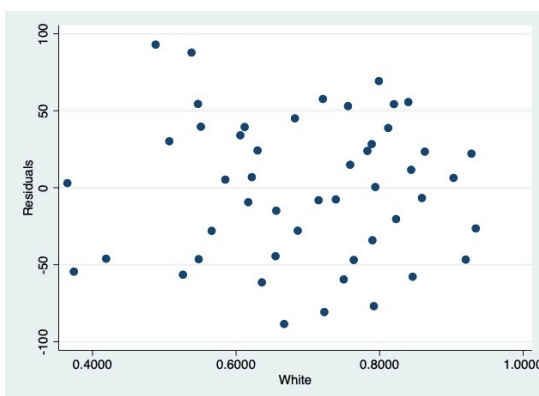
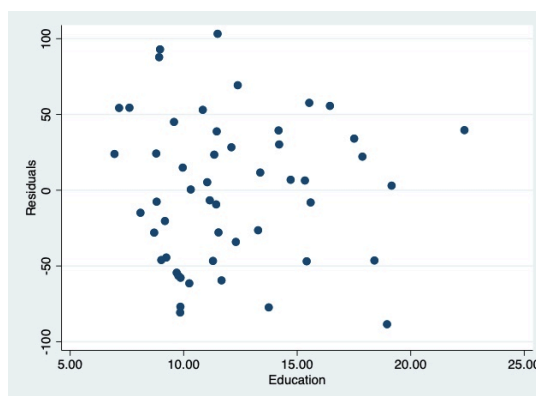
WHITE also has the expected negative correlation with POLICE as this variable describes a percent of the population in comparison to the reference group of minorities. These minorities therefore also hold the expected positive correlation with POLICE, in comparison to

WHITE. Each coefficient value represents the estimated marginal effect that the respective independent variable has on the dependent variable, POLICE. Each independent variable's marginal effect must also be interpreted through the respective units listed in *4.1 The Model*. For example, a one unit increase in crime will have a marginal effect on POLICE based on crime rates per 10,000 population in that respective state.

From these findings, it can be concluded that crime is not the only independent variable necessary in largely describing the variation in police funding. With the addition of POP_DENS, WHITE, and EDUCATION, the R^2 value of the regression increases by 0.5578, from 0.2568 as can be seen in Table 3.5, to 0.8146 as can be seen in Table 4.5. This shows a significant change in the explained variation of police funding and is quite helpful in concluding on these findings.

4.3.2 Heteroscedasticity

As a final note in this discussion, it should be recognized that heteroscedasticity was tested for during the performed regressions. Creating robust regressions is a form of accounting for potential patterned variability in the error residuals. As can be seen below, Graphs:1-4 depict the least square residuals vs. each variable used in the restricted model. It can be concluded, through visual inspection, that there is no evidential heteroscedasticity in these variables, and the robust model is therefore not necessary. This can be interpreted by recognizing the cloud-like formations, and the lack of patterned residual data, in these four scatter plots.

Graph 1: Residuals vs. CRIME**Graph 2: Residuals vs. POP_DENS****Graph 3: Residuals vs. WHITE****Graph 4: Residuals vs. EDUCATION**

5. Conclusion

In conclusion, police funding is likely more than a response to crime. It should be noted although, that this statement is simply stated as a likelihood as the time period and data used in this paper is too limited to conclude on causation between the independent variables and POLICE. Although this is the case, the level of significance, that the hypothesized variables show, is in support of the overarching hypothesis. Namely, this proposed hypothesis is that POLICE holds some independent variables as statistically significant determinants, other than CRIME.

The significance levels found on POP_DENS, EDUCATION, and WHITE, are also largely in support of the presented discussion by Jackson and Carroll (1981). The literature written by Jackson and Carroll delves into the connection between police funding, racial composition, and population density in presenting the direct connection between these variables. This paper aimed to maintain a similar framework, thus these results help to add consequence to the results.

Since crime was only able to account for 25.68% of the variation in POLICE, the conclusions of Georgakopoulos (2004) and Ayres (2000) are also discredited to some extent. These conclusions held that variables such as demographic, population density, and education, only held indirect effects on police funding, through their effects on crime rates. As presented in the preformed regressions and following discussions, this is not the case; other variables are significantly necessary in explaining variations in police funding.

Works Cited

- Taylor, Keeanga-Yamahtta, and Alexis Okeowo. *We Should Still Defund the Police*. 14 Aug. 2020, www.newyorker.com/news/our-columnists/defund-the-police.
- Medved, Michael. "Slow down, Police Are the Good Guys: Column." *USA Today*, Gannett Satellite Information Network, 21 Aug. 2014.
- Lum, Cynthia. "Does the 'Race of Places' Influence Police Officer Decision Making?" *Administration of Justice Department Center for Evidence-Based Crime Policy*, Dec. 2009.
- Georgakopoulos, N. L. Self-Fulfilling Impressions of Criminality: Unintentional Race Profiling. *SSRN Electronic Journal*, 2004.
- Vollaard, Ben, and Pierre Koning. "The Effect of Police on Crime, Disorder and Victim Precaution. Evidence from a Dutch Victimization Survey." *International Review of Law and Economics*, vol. 29, no. 4, 2009, pp. 336–348.
- Ayres, Ian, and Steven Levitt. "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack." 1997.
- Levitt, Steven D. "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply." *American Economic Review*, 2002.
- Levitt, Steven D. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review*, 1995.

Jackson, Pamela Irving, and Leo Carroll. "Race and the War on Crime: The Sociopolitical Determinants of Municipal Police Expenditures in 90 Non-Southern U.S. Cities." *American Sociological Review*, vol. 46, no. 3, 1981, p. 290.

Cameron, Samuel. "The Economics of Crime Deterrence: A Survey of Theory and Evidence." *Kyklos*, vol. 41, no. 2, 1988, pp. 301–323.