Impact of Educational Attainment on Crime in the United States: A Cross-Metropolitan Analysis

Georgia Institute of Technology

ECON 3161, Econometric Analysis

Brian Gentry, Rishab Mokkapati & Kiran Rampersad

November 17, 2016

Abstract:

This study seeks to find a relationship between educational attainment of the population aged 18 years and older, and violent and property crime, in 342 metropolitan areas across the United States. While past studies have researched this relation, they have not done so on a metropolitan scale . Regression models were formed using 2015 data obtained from both the US Census Bureau and the Federal Bureau of Investigation (FBI) . The simple linear regression model found a negative relationship between educational attainment and crime. The more educated a metropolitan population is, the lower its crime levels. Multiple linear regression analyses found that this correlation holds even as other variables are added to the regression. Although this negative correlation is weak, it is still statistically significant.

1. Introduction

Throughout recent history, educational attainment has garnered great attention in the American government's agenda. The Bush administration pumped billions of dollars into the *No Child Left Behind* program while the Obama administration invested billions in ventures such as *Race to the Top* and *Education Jobs Fund*. Education is seen as a tool for social improvement, uplifting the poor and changing lives. However, it also has been seen to have an impact upon crime rates. Many of those who turn to crime are stereotyped as uneducated and desperate. Although the stereotype might be false, the underlying principle is thought to be true; increased educational attainment may have a negative correlation upon crime rates. By gaining some level of education, one can find employment and not have to resort to crime.

There have been numerous studies done on the impact of education on crime. Education is generally regarded as a human capital investment that increases work opportunities in the future and thus discourages participation in crime. Moreover, human capital raises the marginal returns from work more than crime which discourage criminal activity (Lochner, 2011). Groot and van den Brink (2010) argue that educational attainment reduces crime levels because it increases the opportunity costs from forgone earnings and expected costs of incarceration. Fella and Gallipoli (2014) in their theoretical study, concluded that a subsidy to high school completion provides large welfare gains and subsequently reduces crime. Thus, by investing in education, the population is more skilled and knowledgeable, and substantial savings on the social costs of crime can be attained.

While there exists an abundance of literature on the association between educational attainment and crime, the relationship between these two variable at the metropolitan level has not been extensively researched. Metropolitan cities are on the rise as they are emerging faster and larger. The goal of this paper is to uncover the empirical relationship between educational attainment and crime levels across 342 metropolitan statistical areas (MSAs) across the United States, using regression analysis. The model developed in this study will seek to find a correlation between educational attainment of individuals who are above the age of 18 and crime levels (violent and property crimes) using metropolitan-specific data obtained for 2015.

2. Literature Review

Other studies linking education with crime show that there is a correlation between the amount of education an individual has and the amount of crime they are predicted to commit. In 2011, Lance Lochner analyzed the relationship between the number of years of schooling individuals had and the number of times they went to and the years they spent in prison. Lochner (2011)

hypothesizes that additional schooling, namely years spent increasing human capital in close proximity to others, lowers the returns on crime relative to the returns on work, and expresses this in terms of dollars per year. He also notes that better schools tend to reduce delinquency rates and later felony conviction rates: students from low income families more disposed to crime who had been placed in high performing schools via lottery showed a 45% reduction in felony convictions comparative to their peers up to seven years afterwards. After the forced desegregation in of schools in the American South, many low income African American families were able to send their children to better funded schools. This led to a 17% drop in homicides among African Americans of high school age. This in turn suggests that better school quality, as Lochner states, is more effective in deterring crime. However, his method measures the conviction rates of crime, not necessarily the actual crime statistics of the students he studies.

Fella and Gallipoli's (2014) comparative study offers significant insight as to the relative magnitude of increased spending on education upon crime. Fella and Gallipoli (2014) do measure the impact of increasing enrollment rates upon the crime rate, and then correlate the enrollment rates to graduation subsidies. The paper definitively shows that increasing the enrollment numbers decreases the amount of crime at a given time, by giving individuals necessary tools to enter the workforce instead of turning to crime. However, their model seeks to keep students who would otherwise drop out of school to work or commit crime by paying them a subsidy, rather than investing in improvements to the schools themselves. The study does not take into account the achievement rates of the individuals in the schools, and therefore shows the impact of merely keeping students in school longer, rather than improving the quality of their education or improving the school systems to motivate students.

Groot and van de Brink (2010) is similar to Lochner's model in terms of the design of their analysis. Focusing on the relative amount of education each individual has and comparing it to their crime statistics later in life, van de Brink maps out the effects of an additional year of education on individuals, and finds that additional years of education decrease tendencies to commit violent crime, breaking and entering, and tax fraud. However this has little to no noticeable effect upon petty crime statistics. By focusing on years of schooling, Groot and van de Brink (2010) try to map out the quantity of schooling a person has comparatively to their predisposition to commit crimes. Their analysis shows that an additional year's schooling tends to result in less perpetration of vandalism, petty theft, and other minor crimes, resulting in massive savings in government expenditure in crime prevention and social welfare. However, their study doesn't address the quality of schooling addressed in Lochner's analysis, a key component in our observations. While learning about how increasing the

amount of time spent in the classroom would certainly be useful, without evaluation on how much personal capital the students gained, the impact of how the schooling actually improved their potential to the point at which crime became unprofitable.

Marlow (2001) writes about the correlational fallacy between education and crime prevention expenditure. By looking at the various factors which might contribute to increased education or crime prevention expenditure (including political party at the head of local government, diversity of local community, and the average education level of the local community) he charts the relative impact an increase in education expenditure might have on expenditures in crime prevention. The study finds that there is little to no correlation between the two forms of expenditure, meaning that increased educational spending doesn't result in a decrease of the amount spent to control crime. This is useful to us because it indicates we don't need to control for crime expenditure in building our multiple regression model; but the study itself doesn't track the effects of the expenditure on education on the crime level at all.

Mark Anderson's (2014) study gives additional credibility to the idea that schooling itself plays a role in keeping juvenile crime rates low. By analyzing the impact of changes in the Minimum Dropout Age (MDA), Anderson found that increased MDA reduced the number of juveniles of the age the MDA now covered by nearly 10 percent. In addition, a movement across the United States to increase the minimum dropout age to 18 would theoretically decrease the juvenile crime rate by 17 percent. This drop could be contributed to the idea that students in school have less incentive or time to go looking fro crimes to commit. Although the study demonstrates a negative correlation between education attendance and crime rates, the study is specifically focused on juvenile crime rates, whereas our study will be covering violent and property crime rates for adults i.e. 18 years and older.

Our research will avoid focusing on the broad picture or the simple amount of schooling received by students, by focusing on such attainment at the metropolitan level. This differs from previous studies in the sense that we will not be looking at the data surrounding individuals- individual grades and individual crime statistics. Instead of only tracking the number of individuals in schools, we will be looking at the level of the education they received, thereby focusing on their educational attainment as opposed to the enrollment rates. Our study hopes to discover whether an increased educational attainment deters crime more later in life- not whether keeping individuals off the street in their youth prevents crime.

3. Data

3.1 Variables

The purpose of this study was to determine the causal relationship between educational attainment and the crime level in 342 randomly selected metropolitan cities across the United States. (Breakdown on region found in table 2) Crime levels were measured with violent and property crime values obtained for each of these metropolitan cities. Violent crimes include rape, murder and manslaughter, robbery, aggravated assault. Property crimes include burglary, larceny-theft, and motor vehicle theft. To measure the level of educational attainment, the number of individuals 18 and older who reported graduating high school was used.

High school graduation was used, as opposed to other degrees for two reasons. The major reason why high school graduation rates were used over any other degrees (or non-degrees) in the explanation of property crime is the considerable increase in employment opportunities between high school graduates and non-graduates. Although these employment opportunities do not provide comparable financial support to those opportunities available to bachelors and masters degree holders, it is sufficient to provide a living. For those who did not graduate high school there are considerably fewer jobs available, and thus they are more likely to be driven to property crime to obtain financial support. In terms of violent crime, the high school diploma represents an individual's completion of four years of interacting with other individuals and the initial determination of their identity. Due to their interactions with other high school students and teachers, and their instruction in the social studies, former high school students are more equipped to control their emotions and find legal, and appropriate means of conveying their emotions.

In order to ensure that the results from this study accounted for population, the overall population of each of these metropolitan per 100,000 people was used to determine the amount of crime, violent and property, and the percentage of the population, 18 years and older, who have graduated from at least high school and received a high school diploma, was used to measure educational attainment.

We used seven different variables to explain crime rates in each metropolitan area. Table 1 shows the description of the variables used in the simple linear regression model. Education, if used to try to solely explain crime rates, faces a serious omitted variable bias; a high school education alone is unlikely to deter someone from crime. The addition of other explanatory variables was necessary to cover other factors towards crime.

Unemployment (unemploy) represents the percent of the population which reported itself as unemployed in 2015 for each metropolitan area. We expected a positive correlation between Unemployment and crime- as more individuals in an area are employed, they have less reason to commit violent or property crimes as the opportunity costs from incarceration rise. Median age was also accounted for in our analysis since the age plays a factor in determining if an individual commits property or violent crime. We expect a negative correlation between crime levels and median age in the metropolitan area due to the tendency of younger individuals to be the perpetrators of violent and property crime. We expect a positive relationship between Gini levels and crime- the higher the income disparity, the more temptation for those in poverty to engage in criminal activity. Finally, poverty was included and expected to have a positive correlation on crime since those in poverty have a greater incentive to try to improve their lives through crime.

Table 1: Variable Descriptions

Variable	**Description**
crime	Crime level per 100,000 of population (18 years and older)
educ	Percentage of population who graduated from high school (18 years and older)
age	Median age of entire population
unemploy	Unemployment rate of workforce
gini	Gini coefficient - Income distribution of population (measure of inequality)
poverty	Percentage of people who earned an income under poverty level

Table 2: Geographical Distribution of Metropolitan Statistical Areas

US Region	**Number of Metropolitan Statistical Areas**
Northeast	51
Midwest	71
Southeast	105
Southwest	38
West	77
Total	**342**

3.2 Sources

All of the education data was obtained through the American Fact Finder based on US Census Bureau data through the American Community Survey (ACS). The survey provides vital information about the United States and its people. The survey yielded information regarding the number of people who received a high school degree or higher among an 18 or older population in metropolitan cities around the nation. The additional variables that were added for use in the multiple regression analysis were obtained through the US Census Bureau's American FactFinder database. The crime data and the population information for metropolitan cities was obtained through the Federal Bureau of Investigation's Uniform Crime Reporting (UCR) database. The UCR provides crime data for 18,000 cities, university/college, county, state, tribal, and federal law enforcement agencies voluntarily participating in the program. All data collected was from 2015.

3.3 Summary Statistics

Table 3 shows the summary statistics of the six variables used in our regression models.

Table 3: Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
crime	342	3808.89	1299.05	211.46	7692.20
educ	342	87.67	5.05	64.96	98.78
age	342	38.07	4.98	24.60	66.50
unemploy	342	6.31	2.05	2.3	16.5
gini	342	0.46	0.03	0.39	0.54
poverty	342	15.65	4.34	6.62	32.43

3.4 Gauss Markov Assumptions

The first assumption states that the model should be linear in parameters. Since the models used in our regression analyses are written in the form of $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k + u$, assumption 1 is satisfied. For the second assumption, there should be random sampling of regressors. The 342 metropolitan statistical areas used in our regression analysis were randomly selected from a

total of 374. The third assumption states that there should exist no perfect collinearity between any of the independent variables. Table 4 shows that no two independent variables are perfectly positively or negatively correlated. Since there are no exact linear relationships between the regressors, assumption 3 is satisfied.

Table 4: Correlation Between Variables

	crime	*educ*	*age*	*unemploy*	*gini*	*poverty*
crime	1	-	-	-	-	-
educ	-0.35	1	-	-	-	-
age	-0.23	0.13	1	-	-	-
unemploy	0.30	-0.40	0.04	1	-	-
gini	0.20	-0.17	-0.05	0.17	1	-
poverty	0.42	-0.52	-0.28	0.49	0.45	1

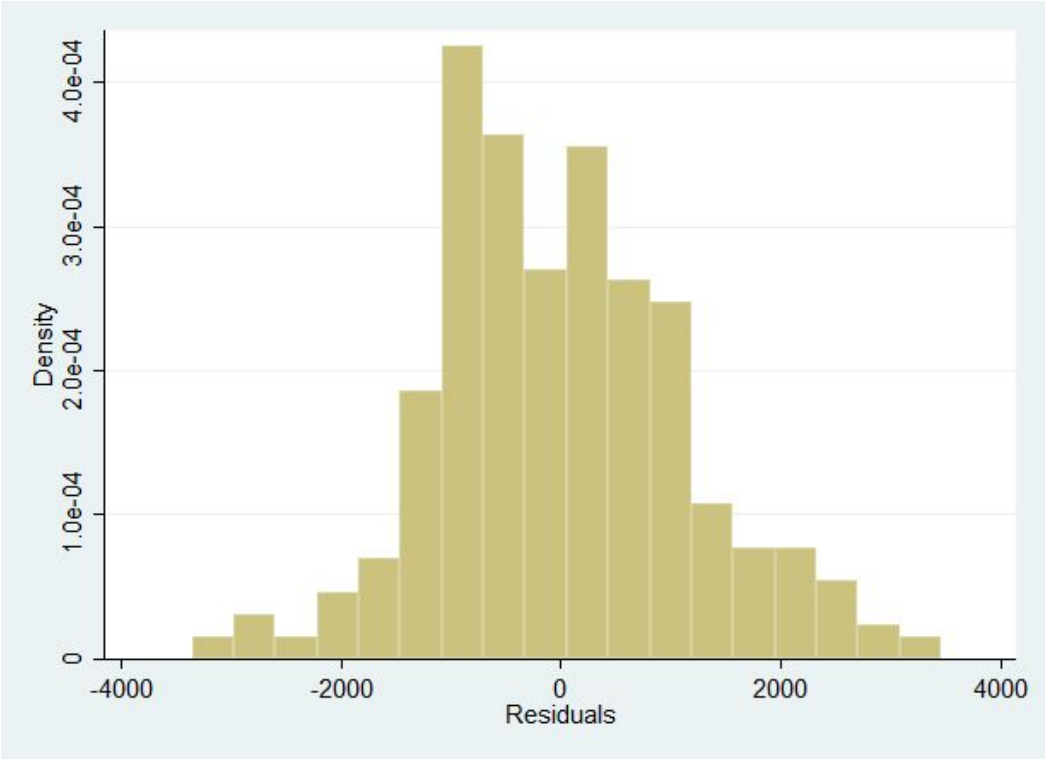
The fourth assumption states that the error term, u, should have an expected value of zero given any value of the independent variables. The summary statistics of the residuals from MLR2 is shown in Table 5 below. Since the mean of the residuals was found to be 4.35e-07, which is approximately zero, assumption 4 is met.

Table 5: Summary Statistics for Residuals in MLR2

	Observations	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
resid	342	4.35e-07	1143.39	-3352.91	3460.77

The fifth and final assumption requires the error term, u, to have a constant variance given any value of the explanatory variables. The residual distribution in Figure 1 demonstrates a normal curve. Thus, assumption 5 is met.

Figure 1: PDF of Residuals



Note: This is the pdf of the residual from regressing *crime* on *educ*, *age*, *unemploy*, *gini* and *poverty*.

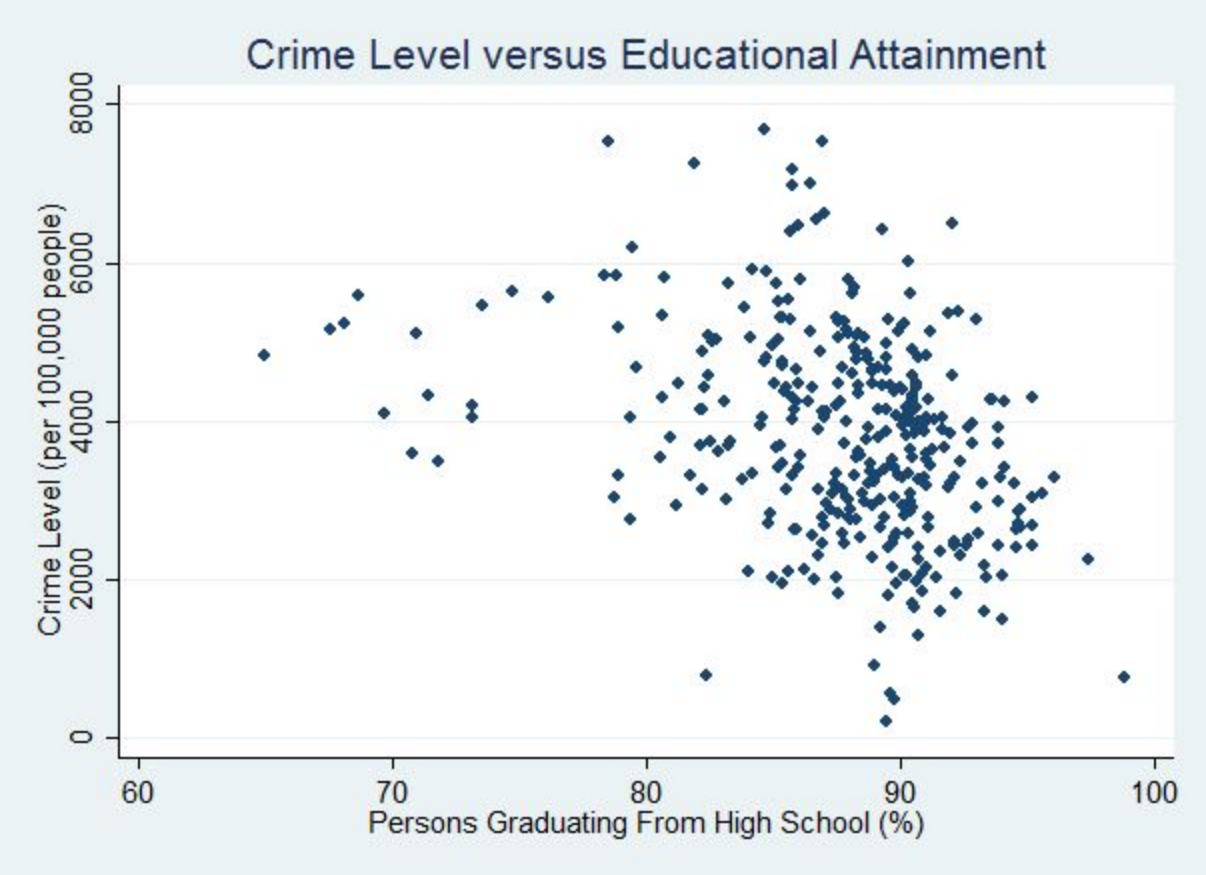
4. Results

4.1 Simple Linear Regression

In the simple regression model, crime was the dependent variable and educational attainment was the independent variable, as seen in Equation 1 below. Our model gave a negative relationship between crime and educational attainment (*educ*) as shown in the scatter plot in Figure 2. For each additional one percent of the population who graduates from high school, the linear regression shows that there is a decrease of 89.1 violent and property crimes per 100 thousand people within each metropolitan area. Although this negative correlation supports our hypothesis, it only does so very loosely- the r-squared value of the simple linear regression is only 11.98 percent. This is due to the omitted variable bias- although education has a negative relationship with crime, lack of education is not enough to explain crime statistics (see table 6).

SLR Model: $crime = \beta_0 + \beta_1 educ + u$

Figure 2: Scatter Plot of Crime Level (per 100,000 people) versus Persons Graduating From High School (%)



Note: Both crime level and persons graduating from high school were measured for the population, 18 years and older. Two outliers were removed from the original data set.

4.2 Multiple Linear Regression

As discussed earlier, simply regressing education and crime will not show the true effects education has on crime due to the omitted variable bias. In order to rectify this discrepancy, several other variables were added in order to create a better relationship, and multiple regression analysis was performed.

The first multiple linear regression model includes the explanatory variables age (age) and unemployment rates (unemploy). These were selected because an aging median population should reduce the amount of individuals fit enough to engage in violent or property crimes. With the

reduction of the omitted variable bias placed upon educ, the number of major crimes reduced by an increase in educated population decreases to 60.42, but the other variables show increased impact upon the crime rates: for each year older the median population is in each metropolitan area, crime rates fall by 52.88 per 100k, and for each additional percent unemploy increases, the amount of crime increases by 132.63. The r-squared value has increased significantly, to 18.88 percent. The addition of the new explanatory variables still does not control for all of the reasons behind crime, but each of the individual variables have added additional validity to the model. (See table 6)

MLR1 Model: $crime = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 unemploy + u$

The second and final multiple linear regression model adds the gini coefficient (gini) and the poverty rate (poverty) of each metropolitan area to the model. This in turn increases the r-squared value to 22.53 percent, but lead to much decreased impact of each of the original variables. Educ still reduces violent crimes per 100k by 40.14 per percent increase, each additional year added to the population median still decreases crime rates by 38, and each additional percent of the population unemployed increases the number of crimes by 77.82. Gini has an extremely strong positive correlation; as the gini coefficient increases from 0 to 1, we see an increase in property and violent crimes by 2434.6. Poverty likewise shows a positive correlation; for each percent the poverty rate increases, the number of crimes increases by 65.25. However, the significance of the variables can be called into question, as will be discussed later on (see table 6).

One of the aspects which we find most significant about all our data sets is the relatively high level of our intercept. Although it drops significantly between our second and third model, the intercept starts at 11621 in our SLR, drops to 10282 in our first MLR, and drops further to 6152 in our second and final MLR. This is indicative, perhaps, of a baseline level of crime that occurs in any economy.

MLR2 Model: $crime = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 unemploy + \beta_4 gini + \beta_5 poverty + u$

Table 6: OLS Regression Estimates

Dependent Variable: *crime*			
Independent Variables	**SLR Model**	**MLR1 Model**	**MLR2 Model**
educ	-89.111 (0.000)***	-60.417 (0.000)***	-40.139 (0.007)***
age	-	-52.886 (0.000)***	-37.996 (0.005)***
unemploy	-	132.631 (0.000)***	77.823 (0.035)**
gini	-	-	2434.609 (0.362)
poverty	-	-	65.251 (0.002)***
Intercept	11621.07 (0.000)***	10282.64 (0.000)***	6152.061 (0.001)***
No. of Observations	342	342	342
R-Squared	0.1198	0.1888	0.2253
Significance Level Key: *10%, **5%, ***1% level			

Note: Values in parenthesis are the respective p-values for each variable.

4.3 Statistical Inferences

The first two models are straightforward when it comes to significance; all of the variables have t-values which places them as significance at all of the three standard levels of significance. In the first model, this data is very misleading; the high t-values mask the fact that there is significant omitted variable bias, leading to overdependence on education, which contributes to the low r-squared value of the model. The second model is much better; with additional variables, the pressure on educ to explain crime in its entirety is reduced, while each of the variables is still significant at every significance level. The low overall r-squared value is likely because no one variable can account for every crime, and we can barely scratch the surface of the number of reasons to commit one.

The final model, which included gini and poverty, has several disconcerting changes. Although in previous models all of the variables were significant, gini and unemployment are no longer significant at the 1%, and gini is not significant at any of the standard significance levels, with a p-value of only 0.91. Gini is an interesting statistic, as it is graded on a scale of 0 to 1, and wit its relatively low variance of 0.15, small changes can seem to have massive impacts, as we can see from its correlation, However, its lack of significance level is telling- even at the 10% level, it fails to be significant. This shows that income inequality may be correlated with higher crime rates, but is unlikely to explain it. Unemployment, however, has only seen a slight decrease in its effectiveness, and is now only significant at the 10% and 5% values. This may be because of moderate levels of colinearity between poverty and unemploy. As unemploy increased, poverty does as well, as we see in table 4 above. Poverty itself, interestingly enough, is significant at all three primary significance levels. One of the promising impacts of the final model is that all of the variables save gini have significant t values- but none have excessive t-values. We surmise that this indicates that the final model manages to avoid completely the omitted variable bias, which plagued educ in the earlier models. Thus, the third model is the most representative of the impact of each variable on crime.

4.4 Robustness

We conducted several F-tests and partial F-tests to determine the usefulness of our models. Table 7 below contains all the results from these robustness tests. First, the SLR model was found to be useful as a whole. Next, we tested our MLR1 model, where we added the variables *age* and *unemploy*. The MLR1 model was also useful as a whole. From the partial F-test on MLR1, *age* and *unemploy* proved to be jointly significant indicating that the entire unrestricted model (MLR1) should be chosen. Lastly, we tested our MLR2 model, where we added the variables *gini* and *poverty*. This model was useful as a whole. However, *gini* is insignificant at all levels, and *unemploy* loses significance at the 1% level. The partial F-test on MLR2, proves that although *gini* and *unemploy* loses some significance individually, they are jointly significant with the other additional independent variables *educ*, *age* and *poverty*. Thus, the unrestricted model (MLR2) should be chosen.

Table 7: Results from F-Tests

SLR Model	**Equation:** $crime = \beta_0 + \beta_1 educ + u$		
Whole Model F-Test	$F_{model} = 46.27$	$F_{0.05,1,340} = 3.84$	$F_{model} > F_{0.05,1,340}$
MLR1 Model	**Unrestricted Equation:** $crime = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 unemploy + u$ **Restricted Equation:** $crime = \beta_0 + \beta_1 educ + u$		
Whole Model F-Test	$F_{model} = 26.22$	$F_{0.05,3,338} = 2.60$	$F_{model} > F_{0.05,3,338}$
Partial F-Test	$F_{partial} = 14.37$	$F_{0.05,2,338} = 3.00$	$F_{partial} > F_{0.05,2,338}$
MLR2 Model	**Unrestricted Equation:** $crime = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 unemploy + \beta_4 gini + \beta_5 poverty + u$ **Restricted Equation:** $crime = \beta_0 + \beta_1 educ + u$		
Whole Model F-Test	$F_{model} = 19.54$	$F_{0.05,5,336} = 2.21$	$F_{model} > F_{0.05,5,336}$
Partial F-Test	$F_{partial} = 11.44$	$F_{0.05,4,336} = 2.37$	$F_{partial} > F_{0.05,4,336}$

5. Conclusion

The results show that our hypothesis was correct- an increase in the percent of individuals educated at the high school level in each metropolitan area leads to a decrease in crime. Although the impact of education on crime is indisputable within our model, the model itself shows that there is a low level of correlation between high school attainment and crime rates overall.

The large intercept indicates that even with all factors optimized to reduced crime rates, some crime would continue to occur. This is almost certainly due to the most basic economic conflict between unlimited wants and limited resources; as long as individuals want what is scarce, they will try to take from others (leading to property crime) and as long as we are stressed at all (whether from lack of resources or other conflicts) we will lash out at others (leading to violent crimes). This can occur no matter how educated or affluent a person is. This leads to a high intercept, and in turn, to a looser fit to our model.

Our model is unable to encapsulate all the reasons people turn to crime. Even people who are not desperate or who have resources or education can be influenced by passion or want more. A

simple model is unable to fully plot crime rates. This in turn leads to a low r-squared; there may be no deciding factor which correlated with crime rates.

Another major factor may be the fact that our research only covers metropolitan areas. Whereas in metropolitan areas population density leads to greater competition for jobs, rural areas have less such competition. A high school education might be sufficient to get a job in a rural area when there's less competition for each individual open position, but in comparison the same individual may not be able to find work in the city. Further research might want to cover rural counties as well as metropolitan areas.

The low level of correlation does not detract from the results, however. There is a negative correlation between the percent of the population which has a high school diploma and the amount of crime committed within each metropolitan area. This could be because of one of three reasons. First, the higher income possible due to increased education might make crime a less attractive means to make a living. Second, crime rates might be lower because continued socialization and indoctrination into normal society throughout high school might make people less likely to commit violent or property crimes. Third, delaying a person's possible entry into the criminal lifestyle until age eighteen might restrict their evolution to property and violent crime.

The propensity to commit crime is influenced by a variety of factors. The results of this study show that educational attainment, regardless of what function it plays in deterring crime, is a factor in predicting the level of crime in a metropolitan city. As a result, to reduce violent and property crime levels in cities around the nation and to shift government funding away from prisons, policy makers should look towards educating the youth and increasing educational attainment.

6. Works Cited

Anderson, D. Mark. "In School And Out Of Trouble? The Minimum Dropout Age And Juvenile Crime." *Review Of Economics And Statistics* 96.2 (2014): 318-331. *EconLit*. Web. 5 Oct. 2016.

FBI:Uniform Crime Reporting Database. "2015 Crime in the United States." *Table 6: Crime in the United States by Metropolitan Statistical Area* (2015). Web. 24 Sept. 2016.

Fella, G., and G. Gallipoli. "Education and Crime over the Life Cycle." *The Review of Economic Studies* 81.4 (2014): 1484-517. Web. 26 Sept. 2016.

Groot, W., and H. M. Van Den Brink. "The Effects of Education on Crime." *Applied Economics* 42.3 (2010): 279-89. Web. 26 Sept. 2016.

Lochner, Lance. "Non-Production Benefits of Education: Crime, Health, and Good Citizenship." (2011). Working Paper 16722: n. pag. *National Bureau of Economic Research*. Web. 26 Sept. 2016.

Marlow, M. L., and A. F. Shiers. "Do Crime-Related Expenditures Crowd out Higher Education Expenditures?" *Public Finance Review* 29.5 (2001): 369-93. Web. 26 Sept. 2016.

OECD. *Education at a Glance 2013: OECD Indicators*. OECD Publishing. (2013). Web. 24 Sept. 2016

United States Census Bureau. "Educational Attainment" *American Community Survey 1-Year Estimates* (2015). Web. 1 Oct. 2016.

United States Census Bureau. "Employment Status in the Past 12 Months" *American Community Survey 1-Year Estimates* (2015). Web. 5 Oct. 2016.

United States Census Bureau. "Gini Index of Income Inequality" *American Community Survey 1-Year Estimates* (2015). Web. 1 Oct. 2016.

United States Census Bureau. "Median Age By Sex" *American Community Survey 1-Year Estimates* (2015). Web. 5 Oct. 2016.

United States Census Bureau. "Poverty Status in the Past 12 Months By Sex By Age" *American Community Survey 1-Year Estimates* (2015). Web. 1 Oct. 2016.

Appendix

A1: STATA Output for Correlation Between Variables

```
. corr crime educ age unemploy gini poverty
(obs=342)
```

	crime	educ	age	unemploy	gini	poverty
crime	1.0000					
educ	-0.3461	1.0000				
age	-0.2257	0.1346	1.0000			
unemploy	0.2953	-0.4022	0.0409	1.0000		
gini	0.2025	-0.1720	-0.0467	0.1684	1.0000	
poverty	0.4228	-0.5171	-0.2837	0.4931	0.4521	1.0000

A2: STATA Output for Residuals Summary Statistics

```
. predict resid, residuals

. summ resid
```

Variable	Obs	Mean	Std. Dev.	Min	Max
resid	342	4.35e-07	1143.392	-3352.913	3460.771

A3: STATA Output for Simple Linear Regression Model

```
. regress crime educ
```

Source	SS	df	MS		
Model	68933147.1	1	68933147.1	Number of obs	= 342
Residual	506517047	340	1489756.02	F(1, 340)	= 46.27
				Prob > F	= 0.0000
Total	575450194	341	1687537.22	R-squared	= 0.1198
				Adj R-squared	= 0.1172
				Root MSE	= 1220.6

crime	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	-89.1108	13.10008	-6.80	0.000	-114.8782	-63.34339
_cons	11621.07	1150.354	10.10	0.000	9358.36	13883.77

A4: STATA Output for Multiple Linear Regression Models

MLR1:

```
. regress crime educ age unemploy
```

Source	SS	df	MS	Number of obs	=	342
				F(3, 338)	=	26.22
Model	108625421	3	36208473.6	Prob > F	=	0.0000
Residual	466824773	338	1381138.38	R-squared	=	0.1888
				Adj R-squared	=	0.1816
Total	575450194	341	1687537.22	Root MSE	=	1175.2

crime	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	-60.41684	13.96882	-4.33	0.000	-87.8936	-32.94007
age	-52.88631	12.97213	-4.08	0.000	-78.40259	-27.37004
unemploy	132.631	34.10579	3.89	0.000	65.54464	199.7173
_cons	10282.64	1337.47	7.69	0.000	7651.827	12913.45

MLR2:

```
. regress crime educ age unemploy gini poverty
```

Source	SS	df	MS	Number of obs	=	342
				F(5, 336)	=	19.54
Model	129645456	5	25929091.1	Prob > F	=	0.0000
Residual	445804738	336	1326799.82	R-squared	=	0.2253
				Adj R-squared	=	0.2138
Total	575450194	341	1687537.22	Root MSE	=	1151.9

crime	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	-40.13897	14.77265	-2.72	0.007	-69.1975	-11.08044
age	-37.99638	13.47435	-2.82	0.005	-64.5011	-11.49166
unemploy	77.82322	36.68442	2.12	0.035	5.663162	149.9833
gini	2434.609	2666.026	0.91	0.362	-2809.595	7678.813
poverty	65.25128	21.06797	3.10	0.002	23.80954	106.693
_cons	6152.061	1824.655	3.37	0.001	2562.874	9741.248