Analysis of Violent Crime Rates on the State and County Level

Introduction

Violent crime in the United States is currently at historical lows, with current rates having fallen by 49% since 1993 [1]. However, in recent years there has been a rebound in rates of violent crime across the country, with particularly large surges of crime after the beginning of the COVID-19 pandemic [2]. While crime rates still remain close to historical lows, current trends indicate that increased resources may be needed soon to respond to the rising rates. Understanding what types of violent crime are increasing and their geographic distribution on both the state and county level can help offer insight into why crime rates are rising and can help with decisions relating to allocation of resources needed to fight crime.

Using data collected from Kaggle [3], this analysis will focus on answering these questions by analyzing data on the county level from 2016. This data offers raw crime totals for every US county and county equivalent in the United States for eight types of violent crime, as well as the population of those counties at the time of data collection. This population data further allows us to back calculate in order to find the rate of crime for each county and also allows us to elevate our data to the state level. Once rates are calculated, comparisons can then be made between different states and counties. Additionally, these results can help with assessing the accuracy of the public's conception of violent crime and whether the types of violent crimes, which are increasing, are the same as the types that the public perceives as rising.

It is important to consider the ethical implications of what this research can entail and suggest. Despite crime in the US being close to historical lows, there is still widespread perception among the public that crime rates have increased and are even near all time highs [4]. This misconception has the potential to be taken advantage of in order to promote the targeting of specific groups and raise support for policies which would be unnecessary to fight crime and potentially detrimental to vulnerable communities. This dataset does not contain any demographic information beyond population totals for each county and therefore is limited in the insight it can provide us. Attempts to correlate certain types of crime with other parameters must be done carefully and thoughtfully so as not to perpetuate harmful stereotypes or promote falsehoods. The purpose of this analysis is to determine the distribution of different types of crime within the United States and propose possible reasons for this. The results should not be used to conclude things that lie far outside of that goal.

Finally, these results should be taken as a fluid snapshot of how violent crime is distributed in the United States as opposed to a concrete distribution that is slow changing. Dynamics in the country have changed considerably since 2016 with a plethora of policy changes on the state and local level across the country as well as events contributing to large scale social upheaval such as the COVID-19 pandemic and widespread economic shocks. The combination of these factors as well as others have likely contributed to reverberations in crime rates across the country in just the last few years.

Methods

The data obtained from Kaggle consisted of raw crime totals for eight different types of violent crime (murder, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, and arson) in each county in the United States. Using these raw crime totals and the population of each county, I was able to plug these values into the rate formula in order to find the rate for each of these eight crimes for all counties. This rate then allows us to make comparisons between counties of different population sizes.

$$\frac{Rate}{100{,}000} = \frac{Crime\ Total}{Population}$$

$$Rate = \frac{Crime\ Total}{Population} \bullet 100{,}000$$

Additionally, we can translate this population and crime total data into data on the state level by adding up all counties from the same state together. This was done using python and allowed me to create an entire new dataframe for violent crime at the state level. The rates for each type of crime were likewise calculated in the same way as they were for counties. Following the creation of this rate data, bar charts were created for all states with all eight types of crime analyzed, and for the top 15 counties with the highest rates for each type of crime.

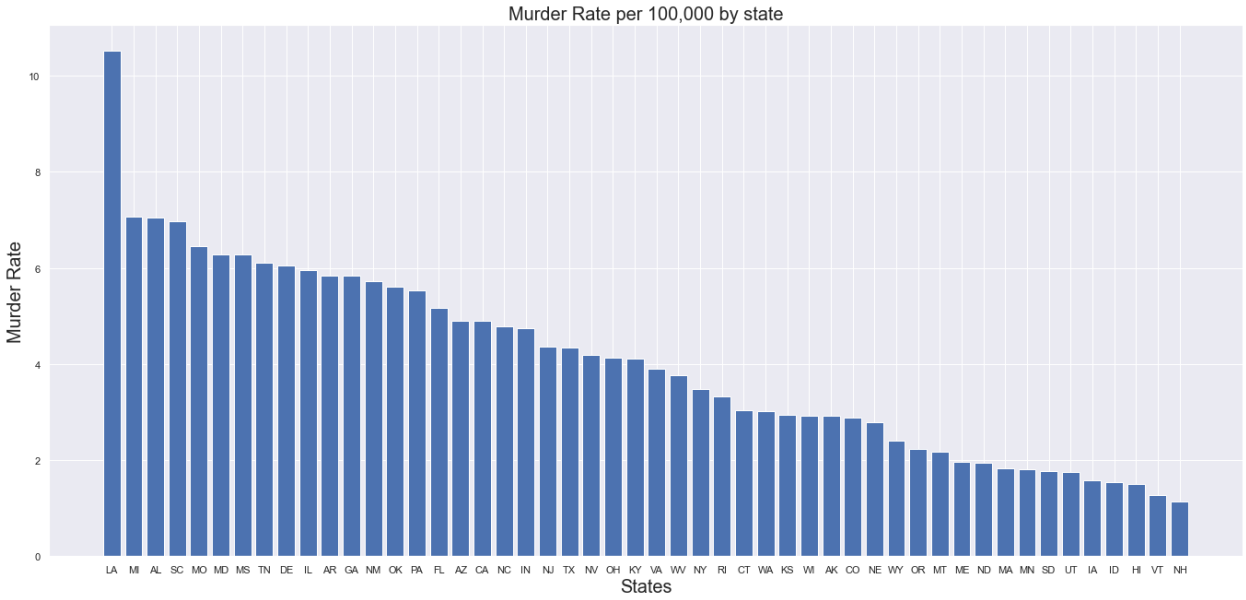
Lastly, a breakdown of counties by population was done in order to detect differences in how crime is distributed throughout the country based on population. Separations were made between counties which have populations ranging from over 500,000, from 100,001 to 500,000, from 50,001 to 100,000, and from 50,000 and less. This breakdown can be used as a proxy for trends between urbanization of a county and the types of crime occuring there.

Results

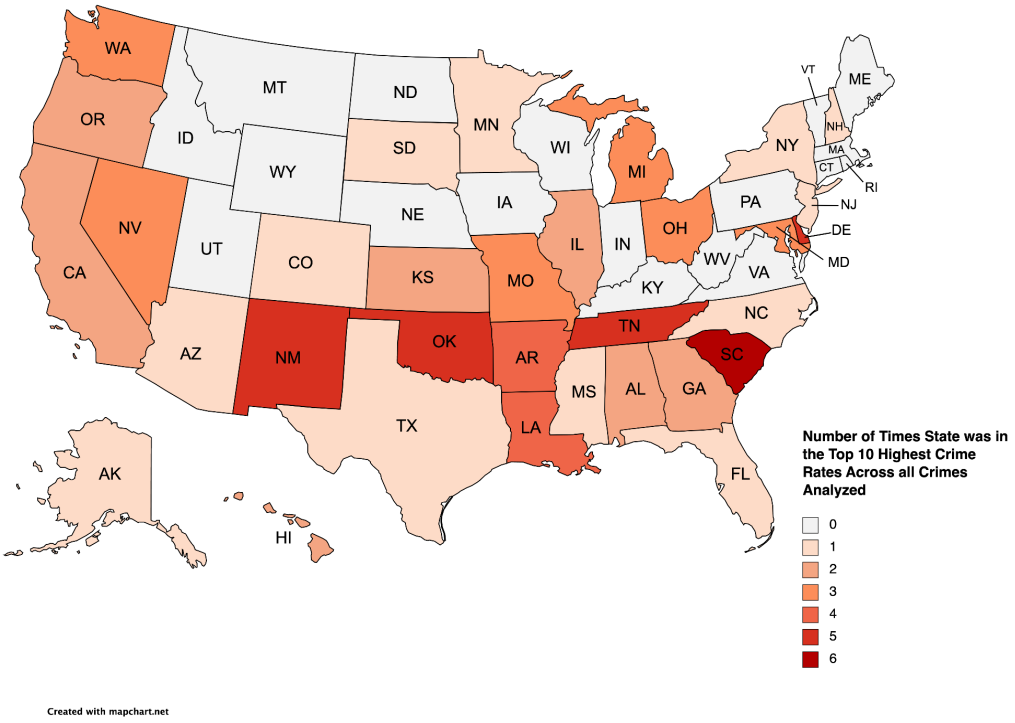
Crime Rates per 100,000 for Entire United States (2016)

Murder	Rape	Robbery	Aggravated Assault	Burglary	Larceny	Motor Theft	Arson
4.55	25.55	109.14	231.63	641.23	1866.84	220.85	16.27

Crime rates for the entire United States were calculated by summing the raw county data and plugging into the aforementioned formula. Larceny stands out as the crime with by far the highest rate. Murder, despite getting the most news coverage, stands out as being the lowest. The breakdown of these rates state by state and county by county has a wide distribution across the entire country, but some patterns do stand out. The South by far had the highest concentration of the country's crime after adjusting for population. Seven of the states with the highest rates of murder were all located in the South, with Louisiana leading the country with a murder rate of 10.53, nearly 3.5 more than Michigan in second place.

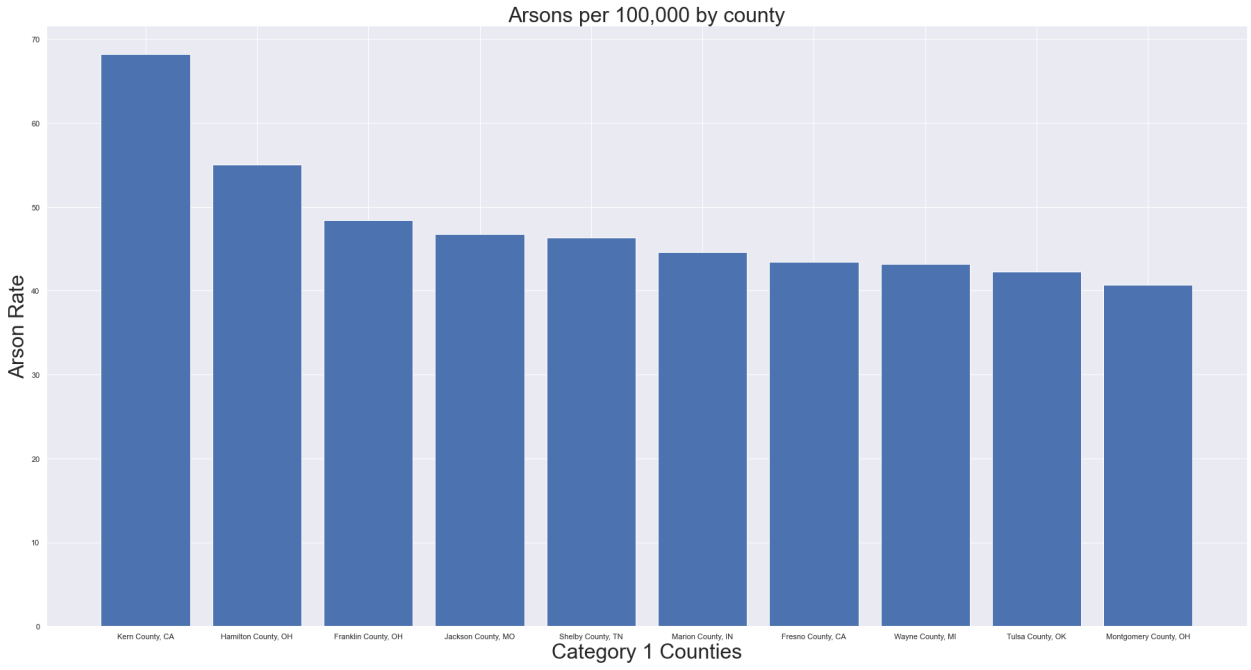


Again, this distribution was not limited to just murder, as the South appears overrepresented in the top ten spots for each crime. South Carolina appears in the top ten spots for six out of eight crimes looked at. Four states appeared in the top ten spots five times, of which only one of those states (New Mexico) was not located within the South.

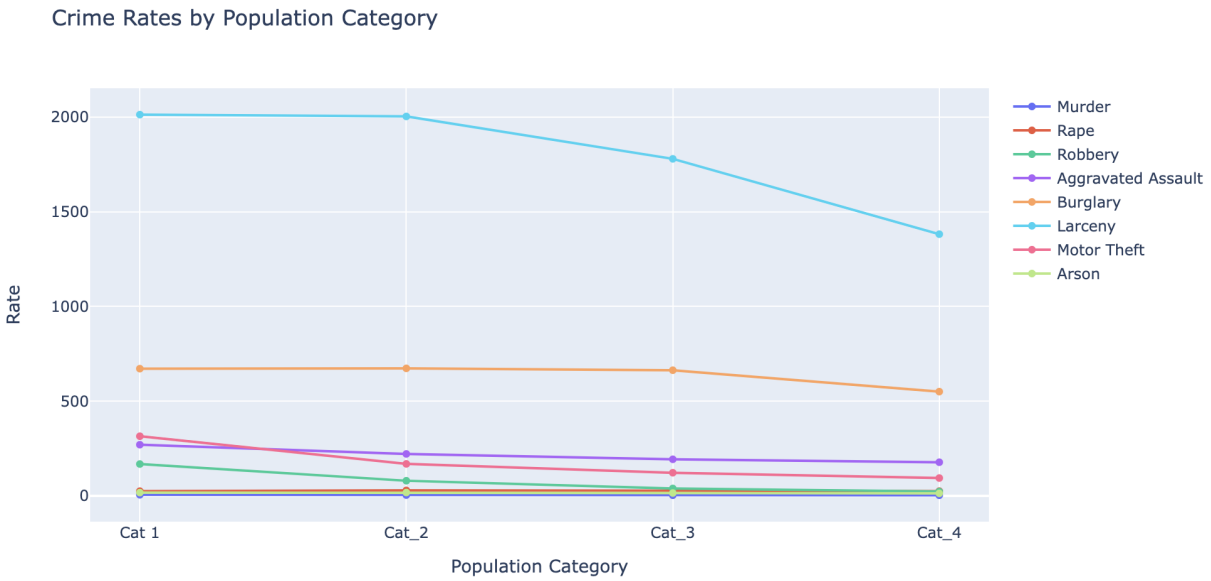


Outside of the South, the Midwest and West Coast stand out as the next highest regions of the country. There appears to be correlation between higher crime rates and poorer economic conditions, as impoverished regions of the country appear to be where the over-representations lie. Rust Belt states (Ohio, Michigan, Illinois, and Missouri) stand out as abnormally high. Another particularly interesting point about the midwest is that rates of arson appear to be high compared with the rest of the country. Ohio ranks first with a rate of 30.91, with Michigan and Missouri joining them in the top ten. A county breakdown of arson is also very telling. Among counties grouped into category 1 (populations greater than 500,000), the midwest is highly overrepresented, with six of the top 10 counties being in the Midwest. The cities of Detroit, Cincinnati, Indianapolis, Columbus, Dayton, Kansas City, and Dayton are located among those counties.

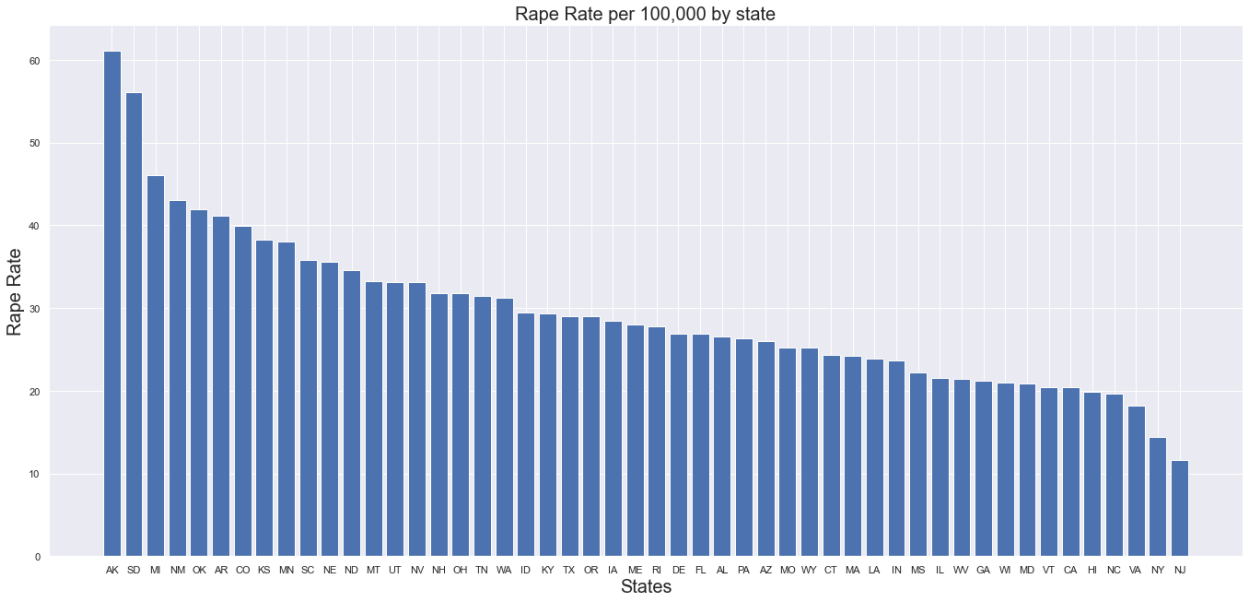
As previously alluded to, poor economic conditions can explain some of the abnormally high crime rates across the Midwest, but another factor may be contributing to the region's abnormally high arson rate. The loss of industrial and manufacturing jobs across the United States has hit the Rust Belt particularly hard, leaving thousands of abandoned factories and warehouses behind. These locations provide would-be arsonists with plenty of potential sites to target. This hypothesis is likely not too far fetched, as others have pointed out similar trends between regions with high manufacturing losses and arsons [5].

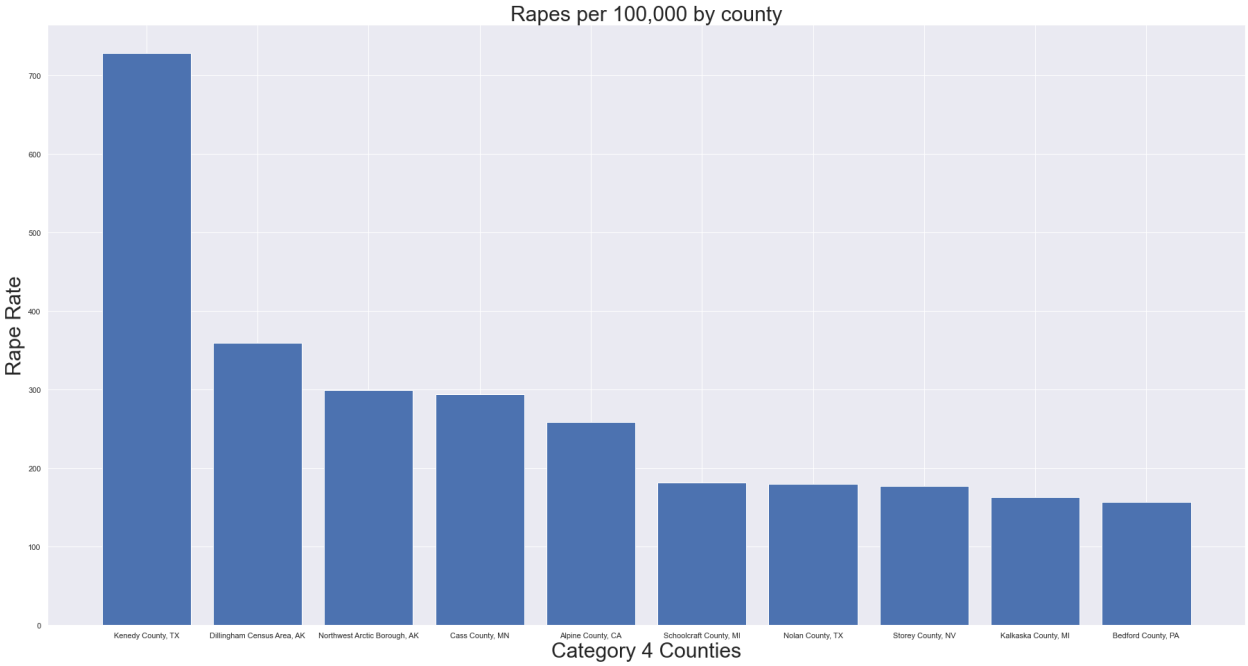


Counties outside of the South in other population categories have higher rates of arson, but their populations being comparatively low mean that very little raw crime has actually occurred in those communities. When analyzing counties by their different population categories, we find a generally positive correlation between population and crime rates, as population categories of higher ranges typically correspond with relatively high crime rates. This decline becomes muddled for only a handful of types of crime.



One of the types of crime which does not see an expected drop off in rate as population declines is Rape. This is particularly interesting, because as was the case with arson, the distribution of sexual assault throughout the country follows a pattern. Upon looking at rates of sexual assault across the United States, Alaska ranks first in the nation, with New Mexico, South Dakota, and Oklahoma joining them in the top ten. These collections of states together are relevant because despite being geographically distinct (New Mexico in the South west, South Dakota in the Great Plains, Oklahoma in the South, and Alaska non-continuous), these states are united in having some of the largest portions of their state be Native American. In fact, these four states just so happen to have the four highest relative native populations in the entire country [6]. Looking outside of just the top ten, we see that Montana and North Dakota sit at 12th and 13th respectively, notable for likewise having very large native populations. The apparent correlation between native populations and rape is present still on the county level. As mentioned, Category 4 counties with populations less than 50,000 actually have a slightly higher rate than Category 1 counties. Notable considering that large segments of the US native population reside in reservations in counties with populations under 50,000. When looking at the county data, there is a remarkable correlation between counties with large native populations and those within close proximity to a largely native counties and there being heightened levels of sexual assault. There is an established trend of Native Americans experiencing increased levels of violence as compared to white Americans, and proximity is notable in this case because most violent crimes committed Native Americans are by non-natives [7]. Native women in particular experience disproportionately high rates of sexual violence and related crimes such as stalking, and even experience these crimes more often than other groups.





The correlations that types of crime have with various economic and demographic data suggests that crime and what causes crime is far more dynamic than how media often portrays it. Rises and falls in crime rates are often assumed to be the direct results of policy changes made by governments, allowing partisans to argue for and against politicians as either weak or strong on crime. However, the results of this analysis appear to indicate that underlying complexities exist for different types of crime which can't be fixed with something like a mere increase in penalties.

Discussion

Crime in the United States is a more complicated subject than popular opinion implies. This is evident in how crime is distributed based on type and the demographics and economics of the regions those crimes are in. Policies relating to things outside of the realm of criminal justice can have even stronger effects on crime rates than actual criminal justice policy. This is evident by the apparent trend in job outsourcing within the rust belt and arson rates. The long legacy of discrimination against Native Americans added with difficulties in prosecuting non-natives has likewise significantly affected the rate of sexual assault across the country, especially in some of the nation's small counties.

All of these observations are important to make note of and understand given the misconceptions that exist about crime. Discussions of crime are closely tied with discussions of crime penalties and whether those penalties are strong enough. Additionally, they're held with the assumption that big cities are where the bulk of this crime occurs. While there are elements of truth in this, since larger counties by population generally have higher rates of violent crime, this assumption is too broad to mindfully cover all the nuances that exist across these different types of crime, and it erases the issue of crime and its effects in smaller and rural communities, which still have to cope with rates which aren't always substantially lower than big cities.

More research and analysis into this type of data would be incredibly helpful in better understanding the sources of different types of crime and how to better curtail crime in meaningful ways. Poverty is already an established risk factor for higher rates of crime [8], but there are clearly other parameters to search for when looking at ways to reduce specific types of violent crime. Those arguing for stricter penalties in response to rising crime rates are likely making knee-jerk reactions to the problem, as sources of the rise in crime appear correlated to other factors. An appropriate evidence based approach to combating crime would pay crucial attention to the types of trends found in this analysis and find case by case solutions in order to target crime rises at the source.

Reference

[1] J. Gramlich, "What the data says (and doesn't say) about crime in the United States," *Pew Research Center*, November 20, 2020.

[2] N. James, "Recent Violent Crime Trends in the United States," *Congressional Research Service*, June 20, 2018.

[3] M. Johnson Jr., "United States crime rates by county," *Kaggle*.

[4] C. April K. and C. Michael, "Mind the Gap: Public Attitudes on Crime Miss the Mark," *International Research Journal of Social Sciences*, January 27th, 2013.

[5] G. Ridley, "Arson: Blue-collar America is burning," *MLive*, April 8th, 2013.

[6] T. Norris, P. L. Vines, and E. M. Hoeffel, "The American Indian and Alaska Native Population: 2010," *2010 Census Briefs*, January 2012.

[7] A. B. Rosay, "Violence Against American Indian and Alaska Native Women and Men," *National Institute of Justice Journal*, June 1st, 2016.

[8] L. Lieberman and A. B. Smith, "Crime Rates and Poverty - A Reexamination," *Crime and Social Justice*, 1986.