Does Crime Breed Inequality?

Igor Barenboim
Harvard University

Dec 24th, 2007

Abstract

Crime and income inequality are positively correlated. Most authors have developed theories that study the causation from income inequality to crime. To our knowledge, no studies have been written on the reverse causality. This paper aims to fill in this gap. Theoretically, we show that if the marginal return to buying protection declines with the amount of protection goods owned, income inequality increases with crime. We also analyze this relationship in an investment and a spatial model. We present some suggestive evidence of the theoretical mechanisms described and quantify it using overcrowding litigations as an instrument.

1 Introduction

Crime and Inequality are positively correlated. Most authors have developed theories that study the causation from income inequality to crime. Becker (1968) explained that in more unequal societies, the return to committing a crime increases, since there is more wealth to be taken away at each crime. Ehrlich (1973) documented this positive relationship using US data. In Figure 1, we plot crime rates per 100,000 inhabitants against the Gini coefficient for a panel of US states from 1969 to 1994 to show this positive correlation. However, Bourgingnon et al. (2003) points out that the evidence of causality from inequality to crime is weak. The same study indicates that these variables are indeed consistently positively correlated in many cross sections studies. However,

---

1 I would like to thank Elias Albagli, Alberto Alesina, David Cutler, Alex Gelber, Edward Glaeser, Alex Kaufman, Lawrence Katz and Andrei Shleifer for their insightful comments and support. Needless to say, all potential mistakes are mine only.
this relationship fails to be significant in a panel data setting in which fixed effects are controlled for. Freeman (1996) has shown this fact for US metropolitan areas and Fajnzylber et al. (2002) for a cross-country panel. Nonetheless, the latter authors did find a positive significant coefficient when they took into account a dynamic relationship between crime and inequality using time series variation\textsuperscript{2}. This evidence suggests that a theory for reverse causality between inequality and crime may be of help to comprehend reality.

There are at least three channels through which crime can breed income inequality. First, consider an environment where individuals can engage in self-protection by buying security goods. Assume that the marginal protection that a security good provides is declining with the amount of security goods already owned. Then, lower income individuals will find optimal to allocate a larger share of their income to acquiring security goods relatively to the rich. As a result, the disposable income distribution, after paying for crime and protection, will be more unequal than the total income one.

The main assumption necessary for the mechanism to go through is that there are negative returns to scale in the protection technology. One would believe that this is a good assumption if one thinks that the amount of security an individual gets from installing a home alarm security system is smaller for individuals who have already a wall around

\textsuperscript{2}using the Arellano-Bover system estimator.
their house than for individuals who do not do so. Di Tella and coauthors (2006) show that after a massive crime increase in Argentina during the 1990s, victimization rates for the poor increased more than for the rich for home robberies. They present evidence that suggests that the rich have invested more in security goods and therefore were less likely to be victimized. They also showed that victimization rates for on-street robberies rose by the same proportion for rich and poor since there is not much for the rich to do besides mimicking the poor while walking on the street. This evidence suggests that there is so much that security goods can do for you and that concavity has to kick in at some point. In sum, this latter study shows that (i) individuals do acquire security goods to self-protect, (ii) the poor do seem to suffer more due to the lack of income to self protect, and (iii) security goods become somewhat redundant once one owns a lot of them.

An alternative assumption that would make the same mechanism work is that there are fixed costs embedded in security goods, or that security goods costs are concave. Think of the cost of building a wall in front of a two bedroom house compared to the cost of building it to protect a four bedroom one. Both endeavors involve transporting the materials to the building site and perhaps the same amount of concrete - the most expensive material for building a wall. The marginal cost would be in terms of bricks and hours worked. A starker example is that to install a home alarm system in both types of houses may cost precisely the same. Or even to buy a German Shepherd and feed it shall entail the same cost for two or four bedroom householders.

Second, consider a two period investment model where individuals can choose in the first period to consume, save, or buy security goods to protect themselves from crime in the second period. Then, lower income individuals will choose to save a smaller share of their income relatively to the rich, because crime creates a wedge in the inter-period marginal rate of substitution. Hence, total income in the second period shall be more unequal than in the first period. In other words, even if the poor allocate a larger share of their income to protection in the first period, they will not be as well protected as the rich, therefore the returns to savings for the poor is lower and hence the total income inequality in the second period shall be higher than in the first.

This investment model has a clear and verifiable empirical implication, that the difference of saving rates of rich and poor is larger when crime rates are higher. Using data from the US Consumption Expenditure Survey we are able to calculate saving rates for above and below median income individuals per state. We then show that the difference

---

3 Levitt (1999) show similar evidence for the US.
of saving rates between rich and poor is positively correlated with per-capita rates of property and violent crime obtained from the Uniform Crime Reports assembled by the FBI.

Third, consider a spatial model of a metropolitan area in which individuals can choose to live in the city center with high density and hence high crime rates, or in the suburbs with lower crime rates. The supply of houses in the city is fixed and the supply of housing in the suburbs is given by the marginal cost of construction. Using the absence of spatial arbitrage as our equilibrium concept\(^4\), and assuming decreasing returns to scale in the protection technology, we can show that, in equilibrium, the rich and the poor will choose to live in the city, while the middle class will flee to the suburbs. This is the case because house prices in the city center drop because of crime, which works as an attraction for the poor. The rich are able to engage in protection and are also therefore better off living in the city center. The middle class is not rich enough to afford enough protection and hence moves to the suburbs\(^5\). We present some suggestive evidence that corroborates with our theoretical finding. We show that in the early 90s, when crime rates were still on the rise, migration to the suburbs was lower for rich and poor and higher for the middle class. We also show that the reverse is true for migration to the city centers.

We finalize this article with an empirical exercise, which consists of finding an instrument that is correlated with crime and not with inequality to obtain an estimate of the impact of crime on inequality. We use overcrowding litigation of prisons in the US as an instrument for property crime - an instrument formerly used by Levitt (1996). The data shows that crime has a positive and significant effect on inequality after three or more years\(^6\). Our point estimates tell us that when crime per-capita rates increase by 10\% the Gini coefficient rises by 1\%.

In the next section we proceed by building up a formal analysis of the theoretical mechanisms we have described here, in Section 3 we provide a discussion and some suggestive empirical evidence on critical assumptions and consequences of our models. In Section 4, we describe the data and the methodology for the instrumental variable estimation of the impact of crime on inequality. In Section 5, we show our estimates of the impact of crime and inequality. The last section concludes the study.

\(^4\) i.e. that individuals cannot be better off by moving to the other area of town
\(^5\) I am thankful to Edward Glaeser for this insight.
\(^6\) up to five years.
2 Formal Analysis

In this section we build the theoretical foundations of the mechanisms through which crime can breed income inequality. We start by looking at a static model, we then look at a two period model to incorporate investment. Next, we analyze a spatial model in which choice of location to live is endogenous.

2.1 Static Model

Consider an environment with a continuum of individuals with an exogenous income distribution \( F(y) \) with support \([0, y_{\text{max}}]\). This economy has an exogenous crime rate \( k \in [0, 1] \), which works as a toll on \( k \) share of individual income. The proceeds from this toll are thrown away. Individuals can buy security goods \( s \) at price 1 to protect themselves from crime. Security goods provide protection \( \pi(s) \in [0, 1] \) from the crime toll, where \( \pi'(s) > 0 \) and \( \pi''(s) < 0 \). Define the disposable income after crime and protection as

\[
\hat{y} = y[1 - k(1 - \pi(s))] - s
\]

and suppose for simplicity that individuals maximize disposable income\(^7\) in choosing \( s \). Define the distribution of disposable income yielded by this maximization is \( G(\hat{y}) \).

**Proposition 1** Crime convexifies the distribution of disposable income. In other words, any individual \( i \) will be richer relatively to the population below her and poorer relatively to the population above her in terms of disposable income than in terms of total income. (Proof in the Appendix)

This result entails that for any distribution of income \( F(y) \), crime would yield a distribution of disposable income with fatter tails. This is the case because the poor in the original distribution find it optimal to allocate a larger share of their income to security goods, since the marginal return of acquiring these goods is much higher for them. While the rich would allocate a smaller share of their income to reach their most desired and relatively high level of protection, shifting them to the right of the disposable income distribution. Figure 2 illustrates the mechanics of the model well showing how an increase in the crime rate shifts the distribution of disposable income. The solid 45 degree line shows that when crime is zero income and disposable income are precisely the same.
However when crime increases disposable income shifts down unevenly for different income levels breeding disposable income inequality.

As we have mentioned previously, an alternative assumption that would make the same mechanism work is that there are fixed costs embedded in security goods, or that security goods costs are concave. To see how that would work formally, just replace $\pi(s)$ by $\pi(\phi(s))$, where $\phi(.)$ is a cost function such that $\phi'(.) > 0$ and $\phi''(.) < 0$.

2.2 Investment Model

Consider a two period investment model where individuals can choose in the first period to consume right away $c_1$, invest to consume in the next period $I$ with rate of return $R$, but subject to a crime toll $k$, which can be mitigated by acquiring security goods $s$ at price 1 in the first period. Protection is again an increasing and concave function $\pi(s)$. Individuals have exogenous income distributed by $F(y)$ and maximize their time additive concave utility function.

$$\max u(c_1) + \beta u(c_2)$$

__Notes__

7 Utility can also be defined as concave, but this adds no value to the results here. So we proceed with the simplest formulation.
where $\beta$ is the discount rate. For simplicity, assume $\beta R = 1$. Individuals maximize utility subject to the following budget constraints

$$BC_1: y = c_1 + I + s$$
$$BC_2: RI[1 - k(1 - \pi(s))] = c_2$$

Taking the first order condition with respect to investment, we can verify that crime creates a wedge between the marginal utility of period 1 and that of period 2. More specifically crime creates a disincentive to save and consume in the second period and hence in equilibrium the marginal utility in period 2 is larger than in period 1. This implies that in this setting it is optimal to consume less in the second period than in the first one.

$$\frac{w'(c_1)}{w'(c_2)} = 1 - k(1 - \pi(s)) < 1$$

Note that the wedge between the marginal utilities is not only a function of crime, but also a function of security goods which depends on the level of income. This observation leads us to the next proposition.

**Proposition 2** In the presence of crime, consumption inequality rises over time.

The proof of this proposition is straighforward. First note that if the crime rate is zero, there is no wedge between the marginal utilities. Next, note that the wedge between marginal utilities diminishes when security goods consumption rises, since the ratio of marginal utilities becomes closer to one. But security goods are normal goods\(^8\) and hence consumption is smoother over time for higher levels of income. The intuition for this result is that lower income individuals will choose to save a smaller share of their income relatively to the rich, because crime creates a wedge in the inter-period marginal rate of substitution. Hence, total income in the second period shall be more unequal than in the first period. In other words, even if the poor allocate a larger share of their income to protection in the first period, they will not be as well protected as the rich, therefore the returns to savings for the poor is lower and thus total income inequality in the second period shall be higher than in the first.

In Figure 3, we illustrate this result. The solid thick line shows that in the absence of crime individuals smooth consumption over time, i.e,
consumption is the same in period 1 and 2. However as crime rises, it is optimal for poorer individuals to consume more in the first period relative to the second, however this difference in inter-temporal consumption choice declines with income for all levels of crime.

Note that for the previous result we did not have to assume concavity of the protection technology function. Another result that we can obtain from this model which is connected to the results obtained for our static model is that lower income individuals would decide to save a smaller share of their income relatively to higher income individuals.

**Proposition 3** *Crime convexifies savings, leading to higher total income inequality over time. (Proof in the appendix)*

This result shows that in a two-period dynamic model the same intuition that we got from the static model holds. There, crime provoked higher disposable income inequality because the returns to acquiring security goods was higher for lower income individuals. Here crime breeds income inequality by inducing a lower saving rate to poor individuals.
2.3 Spatial Model

We go back to our static setting in which individuals maximize their disposable income \( y \) after crime, security expenditures and now housing outlays \( h \). The focus now is on spatial income inequality and for that matter the variable of interest is location choice of individuals. The total population in this economy is normalized to 1. Think of a metropolitan area in which individuals can choose to live in the city center with limited amount of housing \( \frac{1}{2} \), higher density and therefore a higher level of crime \( k \). The price of housing in the city is hence endogenous depending on the demand. The other option is to live in the suburbs where the crime rate is lower \( k \) which we normalize to zero. Here again individuals can engage in self protection by acquiring security goods and we also assume that the protection technology is concave. We also assume an exogenous distribution of income \( F(y) \). So the individual maximization problem can be written as follows.

\[
\max u(\hat{y}) = \hat{y} = y[1 - k(1 - \pi(s))] - s - h
\]

Where the housing costs \( h \) are equal to the marginal cost of construction \( q \) if individuals choose to live in the suburbs, or equal to \( p^* \) the equilibrium price of housing in the city center. Now maximize disposable income by choosing security goods expenditures to obtain the optimal \( s \).

\[
s^* = \pi^{-1} \left( \frac{1}{ky} \right)
\]

To solve this model we use an equilibrium concept from the urban economics literature, which is the absence of spatial arbitrage. This consists on individuals not being able to improve upon moving to another area of town, in other words prices adjust so that the utility of individuals are equalized across suburbs and city center. Mathematically this condition can be written as

\[
y - q = y[1 - \bar{k}(1 - \pi(\pi^{-1} \left( \frac{1}{ky} \right)))] - \pi^{-1} \left( \frac{1}{ky} \right) - p
\]

Note that if crime is zero \( p = q \). Now, we can rewrite this expression to obtain the demand price for city center housing for each income level \( y \).

\[
p = q - y\bar{k} \left[1 - \pi(\pi^{-1} \left( \frac{1}{ky} \right)) \right] - \pi^{-1} \left( \frac{1}{ky} \right)
\]

The city center housing demand expression is intuitive. The city center housing price is equal to the housing price in the suburbs minus the share of income that is taken away by crime and protection. We are now ready to state our last proposition.
Proposition 4 In equilibrium the poor and the rich decide to locate in the city center and the middle class flees to the suburbs, breeding thereby spatial inequality in the city.

To understand the economics behind this proposition we will go over the brief proof. We start by seeing how the demand price for city center housing varies with income and to do that we take its derivative with respect to income.

\[
\frac{\partial p}{\partial y} = -k[1 - \pi(s^*)] + \frac{k}{\pi''(s^*)} - y\frac{k^2\pi'(s^*)}{\pi''(s^*)}
\]

The first two terms of the expression above are negative and they correspond respectively to the loss of income to crime as income rises and the higher expenditures in security goods that richer people choose to do if they locate in the city. The last term is however positive and it corresponds to the protection from crime that individuals get from the security goods acquired. Note that the last term depends linearly and positively on income. So optimal demand prices for housing in the city center vary with income in the following manner. For zero income individuals the demand price for housing in the city is the same as the price in the suburbs \( q \). Then as income rises the demand price starts declining until income reaches \( \tilde{y} = \frac{1}{k\pi'(s^*)}\{1 - \pi''(s^*)[1 - \pi(s^*)]\} > 0 \) and demand prices reach their lower bound at \( \tilde{p} = q - \frac{1 - \pi(s^*)}{\pi'(s^*)}\{1 - \pi''(s^*)[1 - \pi(s^*)]\} - s^* \). After this level of income, the demand price starts rising as the effect of having protection dominates. Since there is not enough housing in the city for all the population in the city center the equilibrium price \( p^* \) will be higher than \( \tilde{p} \), so the middle class which is willing to pay a price close to \( \tilde{p} \) will be driven out of the city center by the poor and the rich. We depict the economics of this proof in Figure 4.

3 Empirical Evidence

In this section, we provide empirical support for the most critical modelling assumptions. We also provide suggestive evidence that supports some of our main theoretical findings. We start by asking if it is right to assume that individuals engage in self-protection and if better-off individuals do more of that. Next, we evaluate the plausibility of the concave protection technology assumption. Third, we look at correlations of crime and the saving rates for different income levels to check if our result that crime curbs more the savings of the poor goes through in the data. Finally, we use metropolitan area migration data to look for some suggestive evidence that supports the prediction of our spatial model.
3.1 Evidence on the Static Model

There are three key assumptions that we make in the static model (i) individuals acquire security goods to self-protect in the presence of crime, (ii) security good is a normal good, i.e. its consumption rises with income and (iii) security goods become somewhat redundant once one owns a lot of them. All of these assumptions are common sense, but in any case we look for evidence of them to get a better grasp if they are true. One consequence of the model is that the poor are more likely to be victimized so we look for evidence on that as well.

Di Tella et al (2006) analyzed victimization survey data in Argentina through the 1990s, when crime rates more than doubled. Looking at the data, we can see that in the early 1990s - low crime era- 10% of the rich (population above the median income) and 2 % of the poor had alarms at their homes. In 2001 these numbers increased respectively to 25% and 8% for the rich and the poor. A similar pattern is observed for the percentage of population that hired security guards. These data suggests that (i) people do engage in costly self-protection activities to avoid crime and that (ii) rich people do more of that.
The remaining assumption necessary for the mechanism to go through is that there are negative returns to scale in the protection technology. One would believe that this is a good assumption if one realizes that for some type of crimes it is really hard to avoid victimization despite investment in security goods. Di Tella and coauthors (2006) show that after a massive crime increase in Argentina during the 1990s, victimization rates for the poor increased more than for the rich for home robberies. More specifically, for the period 1990-1994, high-income households suffered a home victimization rate that was more than double than that observed by low-income families (11 percent versus 5 percent). After that period, low-income households suffered a significant increase in victimization, while high-income families showed a non-significant decline. The cross-sectional difference becomes insignificant in those subsequent periods. Thus, the victimization rate of the low-income households caught up to the high-income rate during the decade. However, they also showed that victimization rates for on-street robberies rose by the same proportion for rich and poor since there is not much the rich can do besides mimicking the poor while walking on the street. This evidence suggests that there is so much that security goods can do for you and that concavity has to kick in at some point.

An alternative assumption that would make the same mechanism work is that there are fixed costs embedded in security goods, or that security goods costs are concave. Think of the cost of building a wall in front of a two bedroom house compared to the cost of building it to protect a four bedroom one. Both endeavors involve transporting the materials to the building site and perhaps the same amount of concrete - the most expensive material for building a wall. The marginal cost would be in terms of bricks and hours worked. A starker example is that to install a home alarm system in both types of houses may cost precisely the same. Or even to buy a German Shepard and feed it shall entail the same cost for two or four bedroom householders.

Finally, we ought to address one of the consequences of our model, which is that despite the higher returns to investment in security goods for the poor, they still become more likely to be victimized than the rich. Di Tella et al. (2006) also addresses this issue. Their paper has two passages describing their evidence for Argentina and evidence by Levitt on the US that support our theoretical finding.

"The rich start the decade with double the victimization rate than the poor (22 percent versus 11 percent, a difference that is significant at the 1 percent level). By the year 2001, the rates had risen to approximately 40 percent and were statistically indistinguishable."
"The evidence suggests that the poor have been the recipients of most of the increase in crime. The increase in crime for the poor has been approximately 1.5 times that suffered by the rich. The difference-in-differences change of the victimization rates between the first and last period of our study is significant at the 5 percent level. As a comparison note that, for the US, Levitt (1999) finds that property crime has become more concentrated on the poor over time. The magnitude of our finding is in line with his estimates. He reports that while in the 1970’s high-income households were slightly more likely to be burglarized than low-income households, by the 1990’s low income households were 60 percent more likely to be the victims of crime."

3.2 Evidence on the Investment Model

The investment model has a clear and verifiable empirical implication, that the difference of saving rates of rich and poor is larger when crime rates are higher. Using data from the US Consumption Expenditure Survey of 1993 we are able to calculate saving rates for above and below median income individuals per state. We than look at how the saving rates of the rich and the poor and their difference is correlated with state per-capita rates of property and violent crime obtained from the Uniform Crime Reports assembled by the FBI. These correlations provide suggestive evidence that this central result of our investment model takes place in reality.

The Consumer Expenditure Survey provides data for more than 5000 households on their total yearly income and on their total consumption expenditure per quarter. To calculate the saving rates we simply calculate total yearly income minus the sum of total quarterly consumption expenditure and divide this number by total income. We also filter the data for outliers which we define as households who save more than 100% of their income and that consume more than 150% of their income. There are many reasons why we would like to exclude these observations for our sample such as the well know problems in self-reports for individuals that are in the extremes of the income distribution. Poor people tend to get cash transfers from family members and rich people tend to under-report income. We remain with 2832 observations after filtering the data.

In Figure 4, we provide a simple correlation matrix using cross sectional variation for the US states for saving rates of the population above and below the median income for each state and two types of per-capita crime rates: property and violent. We look at past crime rates (3, 5 and 10 year averages) to help our inference of causality, since one could
possibly think of a story for reverse causality. The picture tells us that saving rates of the poor are much more correlated with crime rates than that of the rich. More specifically, the correlation between savings of the poor and of the rich with 10 year violent crime average is respectively -21.5% and -9.9%. While the correlation with past 10 year averages of property crime is smaller for both types of households, we still observe the same pattern. The correlation with savings of the poor is -9.9% and of the rich is approximately zero. Also, in the last row of Figure 5, we can see that the difference between saving rates of the poor and of the rich is positively correlated with violent crime (16.5%) and with property crime (9.4%). This evidence is consistent with our theoretical findings, which predicts that in high crime areas, poor people will allocate a larger share of their income to acquiring security goods and forgo savings, whereas this effect will be smaller for the rich.

We proceed with our analyzes by studying the data more carefully. There are two observations that stand out, Washington DC for the very high crime rates and Oklahoma, for which we have very few household observations, this state displays a difference in saving rates between rich and poor of more than four standard deviations above the mean. We now proceed by excluding these two outliers and looking at scatter plots of crime measures and difference of saving rates for rich and poor. In Figure 5 Panel A, we display a plot of violent crime rates per 100 thousand inhabitants and the difference and in Panel B we show the same variable on the horizontal axis and property crime rates on the vertical axis.

Without these two outliers the correlations do drop around 3 percentage points, for example the correlation with 10 year average of property crime drops from 9.4% to 6.3% and for violent crime the correlation drops from 16.7% to 13.9%. Nonetheless they remain positive and significant. Running a simple regression of the difference in saving rates on violent crime yields a t-statistic of 11, dropping the two outliers the t-statistic drops to 9. A similar pattern is observed for property crime. The regression without excluding the outliers yields a t-statistic of 6, while after dropping the outliers the t-statistic drops to 4.
3.3 Evidence on the Spatial Model

In this section, we present some evidence to provide some support for our theoretical finding that crime leads to spatial inequality in the city center. To do that we look at migration data for the US metropolitan areas and check if the middle class is indeed more likely to move to the suburbs when crime is rising and if the poor and the rich are more likely to stay put. We also look at the reverse migration flow, we look to see if the poor and the rich are more likely to move to the city center than the middle class.

In Figure 7, one can see the percentage of the population living in metropolitan areas in the US that moved from central cities to suburbs and from suburbs to central cities sorted by income level. We use data from the Current Population Survey from 1990 to 1995, a period when crime rates were still on the rise and for which the CPS has migration data sorted by income and top coded at a high enough level $100,000\textsuperscript{9}.

In the table we can see that the households earning between $30,000 to $70,000 are more likely to move to the suburbs and the households with earnings below $30,000 are more likely to move to the central city relatively to the other income brackets.

Although, we do not provide information on the population already living across different areas of the metropolitan areas, we believe that looking at changes in the population during an era of rising crime has a flavor of a difference in difference estimation and hence entails suggestive evidence to support our model.

\textsuperscript{9}1994 US dollars.
4 Data and Methodology for Estimating the Impact of Crime on Inequality

The theoretical mechanisms we have described and the evidence that support them discussed above give no quantitative measure of the effect we are studying. To try to address this concern and quantify the impact of crime on inequality, we need some source of exogenous variation that affects crime only, with no other impact on inequality other than through crime. For this purpose we use a variable called overcrowding litigation. In the US, when prison conditions deteriorated, human right groups would file judicial cases against the state responsible for that prison. In some cases, more precisely twelve states, the entire state prison system was under court order concerning overcrowding. According to Levitt (1996) in the three years prior to the initial filling of litigation in these twelve states, prison population growth was higher than the national average by 2.3%. In the three years after the filling of litigation, prisoner growth rates were 2.5 percentage points lower than the national average. In the three years after a final court order, growth rates lagged the national average by 4.8%.

Therefore, keeping crime constant across the nation, six years after the petition was filed, states that had their prison system under court control would have incapacitated 24% less criminals than the national average. This measure gives us an idea of how big the effect of our instrument is. Levitt (1996) uses this instrument to estimate that every non-incapacitated criminal produces on average 15 crimes per year. To get a better feel of the numbers, it is useful to look at the following back of the envelope calculation. During the period that we estimate our regressions 1970 - 1994, the average prison population in the US was around 750,000 persons. Twenty four percent less incapacitation in 12 out of 50 states in the nation imply an increase of 700,000 crimes in a period of six years, in these twelve states. Hence our instrument does
provide a meaningful variation for crime.

We believe it is plausible to assume, that overcrowded prisons and less incapacitation has no other first order impact on inequality rather than through crime\textsuperscript{10}. One caveat to that statement is that prisoners are not taken into account in Gini index calculations. Therefore we have to worry about the mechanical effect. If imprisonment is biased towards very low income people, then letting criminals free should increase the Gini mechanically. The average six years growth of the prison population in our sample is of 400,000 prisoners which account for 0.2\% of the US population on average in the period. Therefore if all prisoners had zero income, this mechanical effect accounts for 0.002 point in the Gini. Hence our estimates should be read taking into account this small negative bias. However, our estimates are of one order of magnitude larger than this mechanical effect, so the purpose of this comment is to note its lack of importance for our main results.

Data on state crime rates are based on the number reported to the police over the course of the year, as compiled annually by the Uniform Crime Reports. Although victimization data would be preferable to reported crimes theoretically, such data is not available \textsuperscript{11}. Reported crime data is available for the seven crime categories: murder and non-negligent homicide, forcible rape, aggravated assault, robbery, burglary, larceny and motor vehicle theft. The first three are considered violent crime and the latter four property crime. The use of reported crime data instead of victimization can lead to a bias, which we try to correct to by regressing time changes of gini on changes of crime, we also believe that after controlling for state and time fixed effects in addition to time differencing the data, it is unlikely that systematic measurement error is driving the results.

To measure income inequality we use the Gini Coefficient constructed by Galbraith and Hale (2006) for the U.S. states. The authors construct the index as follows\textsuperscript{12}: at 10-year intervals, the Census Bureau (2005)\textsuperscript{11}

\textsuperscript{10}It is important to make the caveat, that instrumental variable estimates are many times imperfect, since the exogeneity of the instrument is hardly ever full with certainty. Alberto Alesina has suggested that if in more unequal states, more crimes are likely to be committed and if no more prisons are built to address this concern, then perhaps there could be some correlation between inequality and overcrowding litigation, rather than through crime. Although, it is usually possible to cook up a not so implausible theory about the non-exogeneity of an instrument, we believe that the impact of these theories are of second order and hence we proceed with the estimation.

\textsuperscript{11}See O’Brien (1985) and Gove, Hughes and Geerken (1985) for different views on the validity of the use of reported crime data

\textsuperscript{12}We give a brief description of how the authors estimate the gini coefficients here, but for more details see original paper.
produces a measure of income inequality at the state level for 1969, 1979, 1989, and 1999. To move from decennial to annual data, the authors find an annual dataset that measures wages or incomes for a large proportion of the population of each state, they then create a panel of inequality measures using this underlying data, and use the decennial Census values to transform these yearly inequality measures into estimates of the appropriate Gini coefficient. The ideal dataset for constructing state inequality measures would contain individual level income data for every American—by state—in every year. Such data do not exist, however, the Bureau of Economic Analysis (BEA) in the U.S. Department of Commerce collects data necessary to create internally consistent measures of state pay inequality for the last three decades. For every year since 1969, the BEA has compiled data on wages and employment across dozens of industrial classifications for every state.

We now describe in detail the instrument we chose to disentangle the causal relationship between crime and inequality. The first case on overcrowding litigation was filed in 1965 on the grounds of cruel and unusual punishment. Similar lawsuits took place in 47 states and in DC. Of the approximately 70 cases brought to court, all have achieved at least partial victory but 6. Court orders on overcrowding took form typically by an imposition of population caps, leaving to the administrators to determine the means to comply with the court order (early release programs, construction of new facilities, fewer offenders sent to prison). Only in extreme cases, judges mandated the release of prisoners. The court frequently judged compliance to be inadequate leading to the further step of contempt orders, or court appointed monitors. In twelve states the entire prison system fell under the court order concerning overcrowding. We, as in Levitt (1996), restrict our instrument solely to the states where the entire state prison system fell under the control of the courts, since this states will not be able to comply with court orders on overcrowding simply by rearranging prisoners across prisons within the state. Levitt captured the prison litigation status by six indicator variables and we proceed similarly here. The categories are as follows: (1) Prefilling: no prison overcrowding litigation filed in the state. (2) Filed: litigation filed, but no court decision. (3) Preliminary decision: a court decision is available, but is under appeal. (4) Final decision: no further appeals. (5) Further action: subsequent court intervention on the issue of overcrowding, including appointment of special monitors, contempt orders. (6) Released by court: dismissal of case or relinquishing of court’s oversight of prisons. In Figure 8 the categories that originate the indicator variables that we use as instruments for crime are fully described.
There is wide variation in the timing of prison overcrowding litigation status across the different states. Final court decisions were taken as early as 1971 and as late as 1991. The state prison systems that fell under court order are predominantly Southern though not exclusively so. To avoid major bias from the use of the cross-state variation we regress changes in addition to use state fixed effects.

5 Estimating the Impact of Crime on Inequality

Having described the data in the previous section we can now proceed to estimate the model, using an instrumental variable technique. We first concentrate on checking if our instrument is well correlated with crime. Next, we proceed to the instrumental variable estimation.

To check for the correlation between our instrument and crime, we look at the following first stage regression\[13\].

\[
\Delta \log \text{crime}_{s,t} = \eta + \sum_{i=1}^{5} \varphi_i \Delta \text{Dummy}_{i,s,t-1} + \sum_{i=1}^{5} \kappa_i \Delta \text{Dummy}_{i,s,t-2}
\]

where the dependent variable is the percentage change in per-capita crime rates, which is regressed on a constant \( \eta \), the first lag and the second lag of the change in the five overcrowding litigation status dummies described in the previous section. This is a panel data regression and hence the index \( s \) represent the state, the index \( t \) represent the year and \( i \)

\[13\]Our approach to how to use the instrument follows Levitt (1996).
indexes the overcrowding dummies. In Figure 9 we display the results of this first stage regression. The t-statistics of the overcrowding litigation significant coefficients range from 2 to as high as 10. More importantly the F-statistic for the regression is 14.25 and hence according to Staiger and Stock (1997) we shall not worry about the weak instrument problem.

Since the impact of crime on inequality may take some time to go through as individuals could take a while to realize that there is a permanent change in place and therefore a change in the allocation of income may take a while to materialize, we use lags in our main estimation. This also contributes to mitigate endogeneity worries. We now proceed to estimate the main regression:

\[
\log(Gini_{s,t} - Gini_{s,t-j}) = \alpha + \beta \Delta \log crime_{s,t-j} + \gamma X_t + \delta + \zeta + \varepsilon_t
\]

Where \( \Delta \log crime \), stands for the yearly percentage change of the per-capita crime rate, \( s \) for state, \( t \) for time and \( j \) for the lag. \( X_t \) are the socioeconomic controls\(^{14}\) used, \( \delta \) denotes the time effects.

\(^{14}\)We control for the standard deviation of the percentage of the population that is white, black, American Indian, Asian and Pacific Islander, which is a positive non-linear function of the percentage non-white population. The advantage of this
and $\zeta$ the state fixed effects, which given our specification act as a state specific time trend because our dependent variable is in log changes. In Table 8 we present estimates for $\beta$, the elasticity of inequality to crime, for three different specifications. We first run an OLS regression using time and fixed effects, in which we find a positive significant coefficient of one order of magnitude smaller than our IV estimates. We then run the IV specification with time effects TE. The last specification we run IV TE CSE also controls for cross section effects which here work as state specific time trends. We do this for lags $j = 3, 4, 5$. Our estimates in Figure 10 suggest that a doubling of property crime, should increase the gini coefficient by about 10% from 3 to 5 years later. The coefficient of interest is significant across all specifications with t-statistics close to or above 3.

Perhaps the best robustness check to perform to check the validity of our estimates is to look the same regressions using as dependent variables different types of crime. In doing that, we find that the results hold perfectly for property crime. For violent crime the standard errors grow
more and we find positive coefficients of the same order of magnitude signifcat at 20%.

6 Conclusion

In this study, we show theoretically that crime distorts the optimal allocation of income differently across the income spectrum and thereby breeds inequality. We have also provided suggestive evidence of the mechanisms we describe theoretically. Finally, we have quantified the impact of crime on inequality using overcrowding litigation as an instrument and noted its quantitative importance. In light of this work, the way we think about the relationship between crime and inequality should probably change. Causality could be present in both directions and the channel from crime to inequality appears to be of first order.

7 Appendix

Proof. (Proposition 1) The maximization yields the following FOC:

$$\frac{\partial \hat{y}}{\partial s} = y k \pi'(s) - 1 = 0.$$  Applying the implicit funtion theorem we obtain

$$\frac{\partial y}{\partial s} = -\frac{x'(s)}{y^\pi'(s)} > 0.$$  Note that \( \hat{y} \) is increasing on \( y, \frac{\partial y}{\partial y} = [1 - k(1 - \pi(s))] \).

Now we have all ingredients to see that crime convexifies disposable income.

$$\frac{\partial^2 \hat{y}}{\partial y^2} = k \pi'(s) \frac{\partial y}{\partial s} > 0$$

Proof. (Proposition 3) Take the first order conditions. \( FOC_I: -u(c_1) + \beta R[1 - k(1 - \pi(s))] u'(c_2) = 0 \) and \( FOC_s: -u(c_1) + \beta R(k\pi'(s))u'(c_2) = 0 \).

The FOCs an be combined to \( I - \frac{1 - k(1 - \pi(s))}{k\pi'(s)} = 0 \). Now apply the implicit function theorem to this expression to obtain

$$\frac{\partial I}{\partial y} = \frac{\{[k\pi'(s)]^2 - k \pi''(s)[1 - k(1 - \pi(s))}\} \frac{\partial s}{\partial y}}{\{[k\pi'(s)]^2 - \{[k\pi'(s)]^2 - k \pi''(s)[1 - k(1 - \pi(s))]\} \frac{\partial s}{\partial I}} > 0$$

since \( \frac{\partial s}{\partial I} = -\frac{k \pi'(s)^2}{k \pi'(s)^2 - k \pi''(s)[1 - k(1 - \pi(s))] < 0 \).

Now that we have shwoned that savings increase with income we can show that in the presence of crime it increases at an increasing rate.

$$\frac{\partial^2 I}{\partial y^2} = \frac{2k^2 \pi''(s) \frac{\partial s}{\partial y} k \pi''(s)[1 - k(1 - \pi(s)]}{\{[k\pi'(s)]^2 - \{ [k\pi'(s)]^2 - k \pi''(s)[1 - k(1 - \pi(s))]\} \frac{\partial s}{\partial I}}^2 > 0$$
References