The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports*

Lance Lochner
Department of Economics
University of Rochester

Enrico Moretti
Department of Economics
UCLA

July 6, 2001

Abstract

There are many reasons to expect a link between education and crime. However, estimating the causal effect of education on criminal activity is difficult, because unobserved characteristics that affect schooling decisions are likely to be correlated with unobservables influencing the decision to engage in crime. In this paper, we attempt to empirically address this issue using individual-level data from the Census on incarceration and cohort-level data on arrests by state from the FBI.

We begin by analyzing the effect of high school graduation on incarceration using individual-level Census data from 1960, 1970, and 1980. Both OLS and instrumental variable estimates using changes in state compulsory attendance laws as an instrument for high school graduation reveal a significant relationship between education and incarceration among both whites and blacks; though, estimates suggest that the impacts are greater for blacks than whites. About 25% of the difference in incarceration rates between blacks and whites could be eliminated by reducing black dropout rates to the same level as that of whites.

We next examine the impact of high school graduation on arrests using FBI data on state arrest rates by age for the 1960-90 period. Using variation in graduation rates across age groups within a state and year, we uncover a robust negative effect of graduation on arrests for both violent and property crimes. The biggest impacts of graduation are associated with murder, assault, and motor vehicle theft.

We use the National Longitudinal Survey of Youth to examine the effects of drop out on self-reported crime to ensure that our estimates for imprisonment and arrest are caused by changes in criminal behavior and not educational differences in the probability of arrest or incarceration conditional on crime. The results suggest that high school graduation does reduce actual criminal activity.

Using our estimates, we calculate the social savings from crime reduction associated with high school graduation. The social savings are sizeable and represent an important externality of education that has not yet been documented. The externality of education is about 14-26% of the private return to schooling.

*We especially thank Daron Acemoglu and Josh Angrist for their data on compulsory attendance laws and their useful suggestions. We thank Mark Bils, David Card, Elizabeth Caucutt, Janet Currie, Gordon Dahl, Stan Engerman, Jeff Grogger, David Levine, Jens Ludwig, Jeff Kling, Darren Lubotsky, Marco Manacorda, David Mustard, Steve Rivkin, Cecilia Rouse and seminar participants at Columbia University, University of Rochester, UCLA, and University of British Columbia for their helpful comments.

1 Introduction

Is it possible to reduce crime rates by raising the education of potential criminals? If so, is it cost effective with respect to other crime prevention measures? Despite the enormous policy implications, little is known about the relationship between schooling and criminal behavior.

Economists interested in the benefits of schooling have traditionally focused on the private return to education. However, researchers have recently started to investigate whether schooling generates benefits beyond the private returns received by individuals. In particular, a number of studies attempt to determine whether the schooling of one worker raises the earnings of other workers around him (Acemoglu and Angrist 2000, Heckman and Klenow 1999, Moretti 1999, Rausch 1993). Yet, little research has been undertaken to evaluate the importance of other types of external benefits of education, such as its potential effects on crime (Lochner (1999) and Witte (1997) are notable exceptions). Crime is a negative externality with enormous social costs (Miller, Cohen and Wiersema 1996). If education reduces crime, then schooling will have social benefits that are not taken into account by individuals. Given the large social costs of crime, even small reductions in crime associated with education may be economically important.

There are a number of reasons to believe that education can reduce criminal activity. First, schooling increases the returns to legitimate work, raising the opportunity costs of illegal behavior.[1] Additionally, punishment for criminal behavior often entails incarceration. By raising wage rates, schooling makes any time spent out of the labor market more costly. Second, schooling may directly affect the financial or psychic rewards from crime itself. Finally, schooling may alter preferences in indirect ways, which may affect decisions to engage in crime. For example, education may increase one's patience (as in Becker and Mulligan (1997)) or risk aversion. To the extent that schooling reduces individual discount rates, it makes potential punishments more costly, since they often extend long after a crime is committed. Since crime is generally a risky activity, both in its returns and costs, education will reduce crime if it increases individual aversion to risk.

Despite the many reasons to expect a causal link between education and crime, empirical research is not conclusive.[2] The key difficulty in estimating the effect of education on criminal

[1] Freeman (1996), Gould, et al. (2000), Grogger (1998), Machin and Meghir (2000), and Viscusi (1986) empirically establish a negative correlation between earnings levels (or wage rates) and criminal activity. The relationship between crime and unemployment has been more tenuous (see Chiricos (1987) or Freeman (1983, 1995) for excellent surveys); however, a number of recent studies that better address problems with endogeneity and unobserved correlates (including Gould, et al. (2000) and Raphael and Winter-Ebmer (2001)) find a sizeable positive effect of unemployment on crime. Also, see Ehrlich (1975) for an early discussion of the relationship between education and crime.

[2] Witte (1997) concludes that "...neither years of schooling completed nor receipt of a high school degree has a significant affect on an individual's level of criminal activity." But, this conclusion is based on only a few available studies, including Tauchen, et al. (1994) and Witte and Tauchen (1994), which find no significant link between education and crime after controlling for a number of individual characteristics. While Grogger (1998) estimates a significant negative relationship between wage rates and crime, he finds no relationship between education and crime

activity is that unobserved characteristics affecting schooling decisions are likely to be correlated with unobservables influencing the decision to engage in crime. For example, individuals with high criminal returns or discount rates are likely to spend much of their time engaged in crime rather than work regardless of their educational background. To the extent that schooling does not raise criminal returns, there is little reward to finishing high school or attending college for these individuals. As a result, we might expect a negative correlation between crime and education even if there is no causal effect of education on crime. State policies may induce bias with the opposite sign – if increases in state spending for crime prevention and prison construction trade off with spending for public education, a *positive* spurious correlation between education and crime is also possible.

A second difficulty in estimating the effect of schooling on crime arises because actual criminal activity is generally unknown. In this paper, we use individual-level data on incarceration from the Census and cohort-level data on arrests by state from the Uniform Crime Reports (UCR) to analyze the effects of schooling on crime. We then turn to self-report data on criminal activity from the National Longitudinal Survey of Youth (NLSY) to verify that the estimated impacts measure changes in crime and not educational differences in the probability of arrest or incarceration conditional on crime. We employ a number of empirical strategies to account for unobservable individual characteristics and state policies that may introduce spurious correlation.

We start by analyzing the effect of high school graduation on incarceration using Census data from 1960, 1970, and 1980. The group quarters type of residence indicates whether an individual is incarcerated at the Census date. OLS estimates show significant impacts of high school graduation on incarceration for both blacks and whites. Cohorts in a state that have higher than average drop out rates also have higher than average incarceration rates. To address endogeneity problems, we turn to instrumental variable techniques using changes in compulsory attendance laws over time to instrument for high school drop out. Changes in state compulsory attendance laws have a significant effect on high school graduation among whites and blacks, and we reject tests for reverse causality. Instrumental variable estimates reveal a significant relationship between education and incarceration among both whites and blacks. Estimates suggest that the impacts are greater for blacks than for whites. Differences in high school drop out rates between blacks and whites can explain as much as 25% of the 1980 black-white gap in incarceration rates.

after controlling for wages. (Of course, increased wages are an important consequence of schooling.) More recently, Lochner (1999) estimates a significant and important link between high school graduation and crime using data from the National Longitudinal Survey of Youth (NLSY). Other research relevant to the link between education and crime has examined the correlation between crime and time spent in school (Gottfredson 1985, Farrington et al. 1986, Witte and Tauchen 1994). These studies find that time spent in school significantly reduces criminal activity – more so than time spent at work – suggesting a contemporaneous link between school attendance and crime. Previous empirical studies have not controlled for the endogeneity of schooling.

The interpretation of the IV estimator is complicated by the fact our instrument affects schooling progressions at many different grade levels (i.e. it causes some to attend 10 rather than 9 years, 11 rather than 10 years, etc.). When estimating the impact of high school graduation only, OLS and IV estimators estimate different weighted sums of the impact of each schooling progression on the probability of incarceration and not simply the impact of progressing from 11th to 12th grade. Building on results in Angrist and Imbens (1995), we formally develop and empirically implement a formula that clarifies the relationship between OLS and IV estimates. We show that the "weights" placed on the effects of each schooling progression have an intuitive interpretation and are functions of observable quantities. Differences in these "weights" explain nearly all of the difference between our OLS and IV estimates.

Because incarceration data do not distinguish between types of offenses, we also examine the impact of high school graduation on arrests using data on arrest rates from the 1960-90 UCR and high school graduation rates from the 1960-90 Censuses. This data allows us to identify the type of crime that arrested individuals have been charged with. Estimates controlling for many state, age, year, and offense interactions uncover a robust and significant positive effect of drop out on arrests for both violent and property crimes, effects which are consistent with the magnitude of impacts observed for incarceration in the Census data. When arrests are separately analyzed by crime, the greatest impacts of drop out are associated with murder, assault, and motor vehicle theft.

Estimates using arrest and imprisonment measures of crime may confound the effect of education on criminal activity with educational differences in the probability of arrest and sentencing conditional on commission of a crime. To verify that our estimates identify a relationship between education and actual crime, we estimate the effects of high school graduation and college attendance on self-reported criminal participation using data from the NLSY. These estimates suggest that high school graduation significantly reduces self-reported participation in both violent and property crime among young white men. We also use the NLSY to explore the robustness of the estimated effects of drop out on incarceration to the inclusion of rich measures of family background, individual ability, and labor market characteristics. Both self-report and incarceration estimates are similar across sparse and rich empirical specifications. Additionally, the estimated effects on incarceration are consistent with our findings using Census data.

Given the general consistency in findings across data sets, measures of criminal intensity, and empirical methods, we cannot reject that a relationship between high school graduation and crime exists. We use our estimates to calculate the social savings from crime reduction associated with high school completion. Our estimates suggest that a 1% reduction in male high school drop out rates would save as much as \$1.4 billion, or about \$2,100 per additional male high school graduate. Social savings estimates range from 14-26% of the private return to high school graduation, suggesting that a significant part of the social return to education is in the form of externalities from crime reduction.

The remainder of the paper is organized as follows. In Section 2, a simple model of criminal activity is described. Section 3 reports estimates of the impact of schooling on incarceration rates (Census data), and Section 4 reports estimates of the impact of schooling on arrest rates (UCR data). Section 5 uses NLSY data on self-reported crime and on incarceration to check the robustness of UCR and Census-based estimates. In Section 6, we calculate the social savings from crime reduction associated with high school graduation. Section 7 summarizes our findings and concludes.

2 An Economic Framework

To provide some intuition as to why high school graduation might affect criminal behavior, this section discusses a simple economic model of work, school, and crime. The model is by no means a complete description of criminal decisionmaking, but it is a useful reference point for an empirical study of crime and schooling. Individuals are assumed to choose the amount of education they acquire and the amount of time spent on work and crime once they have finished school. We begin by analyzing crime and work decisions conditional on educational choice, then return to the educational choice problem.

Consider the decisions of someone who has completed s years of school and must decide how to allocate his time to work and crime, where k_t is the fraction of time spent committing crime at age t. Let $w_t(s)$ represent his wage rate and $R(k_t, s)$ his total net return from crime at age t if he has s years of schooling.[3] Assume that someone who commits crime in period t has a probability, $\pi(k_t)$, of being punished the next period, $t+1$, where $\pi'(k_t) > 0$. For simplicity, assume that the punishment, P, is constant over time and measured in utility terms. While in school, an individual receives a constant utility of $\bar{u}$. Since we are interested in post-school crime, we ignore crime during school years. At each age after completing school, individuals consume their income from work and crime, receiving utility $u(y_t)$, where $y_t = w_t(s)(1-k_t) + R(k_t, s)$ is total income in period t, $u'(y_t) > 0$ and $u''(y_t) \leq 0$. The individual's maximization problem, conditional on already having chosen s years of school, is

$$V(s) = \max_{\{k_t\}_{t=s+1}^{T}} \left\{ \sum_{t=0}^{s-1} \beta^t \bar{u} + \beta^{s-1} \sum_{t=s+1}^{T} \rho^{t-s}(s)[u(w_t(s)(1-k_t) + R(k_t, s)) - \rho(s)\pi(k_t)P] \right\},$$

where $\beta \in [0,1]$ is an individual's initial discount factor, $\rho(s) \in [0,1]$ is his after-school discount factor (i.e. schooling may affect his rate of time preference), and T is the total number of years

[3]In the analysis that follows, we assume that $w_t'(s) > 0$, $w_t'(s) > \frac{\partial^2 R}{\partial k_t \partial s}$, and $\frac{\partial R}{\partial s} \geq 0$.

he can work or attend school. $V(s)$ represents the lifetime value of choosing s years of school when the individual chooses his criminal behavior optimally.

The interior first order condition for crime, k_t, is given by

$$\frac{\partial R(k_t, s)}{\partial k_t} - w_t(s) = \rho(s)\left(\frac{\pi'(k_t)P}{u'(y_t)}\right) \geq 0. \tag{1}$$

Notice, the gap between current returns from crime and work must be (weakly) positive, since crime involves future punishment. The right hand side of equation (1) represents the compensating differential that must be paid in the criminal sector due to potential punishment. It is increasing in the discount rate $\rho(s)$, since impatient individuals (low $\rho(s)$) heavily discount future punishment costs.

Equation (1) suggests three ways that schooling can affect criminal decisions. First, schooling increases individual wage rates, thereby increasing the opportunity costs of crime. Second, schooling may affect the net marginal returns to crime, $\frac{\partial R}{\partial k_t}$. Finally, schooling may alter individual rates of time preference. That is, schooling may increase the patience exhibited by individuals (i.e. $\rho'(s) > 0$).[4] As long as schooling increases the marginal return to work more than crime ($w_t'(s) > \frac{\partial^2 R}{\partial k_t \partial s}$) and schooling does not decrease patience levels, crime is decreasing in the number of years of schooling. It is also clear that, all else equal, individuals with higher wage rates and lower discount factors (ρ) will commit less crime. High wage rates reduce crime by increasing its opportunity cost, while increased patience makes delayed punishments more costly.

Recall that schooling is not exogenous. After calculating their optimal lifetime work and crime decisions for each potential level of schooling, young individuals will choose the education level that maximizes lifetime earnings, $V(s)$. The same factors that affect decisions to commit crime, and therefore $V(s)$, also affect schooling decisions. For example, it is clear from equation (1) that individuals with lower discount factors will engage in more crime, since more impatient individuals put less weight on future punishments. At the same time, individuals with low discount factors choose to invest less in schooling, since they discount the future benefits of schooling more heavily. Similarly, individuals with a high marginal return from crime are likely to spend much of their time committing crime regardless of their educational attainment. If schooling provides little or no return in the criminal sector (i.e. $\frac{\partial^2 R}{\partial k_t \partial s} \approx 0$), then there is little value to attending school. Both examples suggest that schooling and crime are likely to be negatively correlated, even if schooling has no causal effect on crime.

[4]While this specification does not explicitly allow schooling to affect individual tastes for crime, the relationship between schooling and net criminal returns yields qualitatively similar implications. One could also allow for incarceration (i.e. punishment may include time out of the criminal and labor market), which would make punishments more costly for those with high wage rates or criminal returns. This would provide another channel through which schooling affects crime. These additional channels have been suppressed to maintain simplicity.

An Empirical Specification

We now clarify the empirical issues involved in estimation. To simplify the discussion further, consider an income maximizing framework ($u(y) = y$) where the probability of facing punishment is simply πk_t and the net return to crime depends on individual characteristics θ according to[5]

$$R(k_t, s) = (\theta + \gamma(s))k_t - \frac{\eta}{2}k_t^2.$$

Solving the first order conditions for k_t yields the following characterization for criminal activity:[6] $k_t(\theta, s) = (1/\eta)(-w_t(s) + \gamma(s) + \theta - \rho(s)\pi P)$. If time preferences are unaffected by schooling and the probability of punishment as well as the punishment itself varies across locations, l, then criminal behavior for person i at age t is described by

$$k_{i,l,t} = -\frac{w_t(s_i) - \gamma(s_i)}{\eta} + \frac{\theta_i}{\eta} - \frac{\rho\pi_l P_l}{\eta}$$

If the difference in returns to schooling for work and crime is linear in schooling, so $(w_t(s) - \gamma(s))/\eta = -c_t + \delta s$, we obtain a simple reduced form specification for criminal behavior for person i as

$$k_{i,l,t} = c_t - \delta s_i + \tilde{\theta}_i - d_l$$

where δ is the parameter of interest, $\tilde{\theta}_i = \theta_i/\eta$ is a measure of unobserved ability, and $d_l = \rho\pi_l P_l/\eta$ is a state-specific dummy representing the deterrent effect of expected punishment. Standard OLS regressions of criminal activity on schooling and current location will be biased if individual criminal abilities or returns (θ_i) are correlated with schooling. Indeed, theory suggests that θ_i and s_i are negatively correlated (i.e. those with high criminal returns will acquire less education) producing a *negative* bias in the estimated impact of schooling on crime. Failure to account for differences in expected punishment levels across location (or state-level effects) will also induce a bias if expected punishments are correlated with schooling outcomes. If states face budget decisions between funding for schools and funding for police and prisons, we might expect a negative correlation between expected punishments and schooling levels across states. In this case, failure to account for state-level variation in expected punishments would lead to *positively* biased estimates. When controlling for state-level differences, only the first bias remains and estimates should be negatively biased due to unobserved heterogeneity in θ_i.

To address the endogeneity of schooling and eliminate the bias induced by correlation between θ_i and s_i, we use instruments that exogenously affect schooling choices. Using valid instruments for schooling and controlling for state-level variation should produce unbiased estimates of δ in the simple framework above. However, the interpretation of IV estimates is more complicated

[5]We assume throughout this analysis that θ is unobserved. Introducing observed characteristics is straightforward.

[6]For an interior solution where $\theta \in (\alpha_t(s), \alpha_t(s) + \eta)$ and $\alpha_t(s) = w_t(s) - \gamma(s) + \rho(s)\pi P$.

when δ varies across schooling levels or across individuals (Angrist and Imbens 1995). In Section 3.4 below, we derive a simple econometric formula that relates IV estimates to OLS estimates when δ is allowed to vary across schooling levels.

A final complication arises due to data limitations. The instruments that we use in this paper are effective when using large data sets on crime like the Census or UCR. However, neither of these data sets measures crime directly. The Census data provide information on incarceration while the UCR provide data on arrests. It is, therefore, important to clarify the relationship between schooling and these alternative measures of crime.

It is reasonable to assume that arrests and incarceration are a function of the amount of crime committed, k_t. Consider first the case where both the probability of arrest conditional on crime (π_a) and the probability of incarceration conditional on arrest (π_i) are constant and age invariant. Then an individual with s years of schooling will be arrested with probability $Pr(Arrest_t) = \pi_a k_t(s)$ and incarcerated with probability $Pr(Inc_t) = \pi_i \pi_a k_t(s)$.

Consider two schooling levels – high school completion (s=1) and drop out (s=0). Then, the effect of drop out on crime is simply $k_t(0) - k_t(1)$, while its effect on arrests is $\pi_a(k_t(0) - k_t(1))$. Its impact on incarceration is $\pi_i \pi_a(k_t(0) - k_t(1))$. The measured effects of drop out on arrest and incarceration rates are less than its effect on crime by factors of π_a and $\pi_i \pi_a$, respectively. However, drop out should have similar effects on crime, arrests, and incarceration when measured in logarithms.

More generally, the probability of arrest conditional on crime, $\pi_a(s)$, and the probability of incarceration conditional on arrest, $\pi_i(s)$, may depend on schooling. This would be the case if, for example, more educated individuals have access to better legal defense resources or are treated more leniently by police officers and judges. In this case, the measured effects of drop out on arrest and incarceration rates (when measured in logarithms) are

$$\ln Pr(Arrest_t|s=0) - \ln Pr(Arrest_t|s=1) = (\ln k_t(0) - \ln k_t(1)) + (\ln \pi_a(0) - \ln \pi_a(1))$$

and

$$\ln Pr(Inc_t|s=0) - \ln Pr(Inc_t|s=1) = (\ln k_t(0) - \ln k_t(1)) + (\ln \pi_a(0) - \ln \pi_a(1)) + (\ln \pi_i(0) - \ln \pi_i(1)),$$

respectively. If the probability of arrest conditional on crime and the probability of incarceration conditional on arrest are larger for less educated individuals, then the measured effect of drop out on arrest is larger than its effect on crime by $\ln \pi_a(0) - \ln \pi_a(1)$ and its measured effect on imprisonment is larger still by the additional amount $\ln \pi_i(0) - \ln \pi_i(1)$. Mustard (2001) provides evidence from U.S. federal court sentencing that high school drop outs are likely to receive a slightly longer sentence than otherwise similar graduates, though the difference is quite small (about 2-3%).

3 The Impact of Schooling on Incarceration Rates

3.1 Data and OLS Estimates

We begin by analyzing the impact of education on the probability of incarceration for men using U.S. Census data. The public versions of the 1960, 1970, and 1980 Censuses report the type of group quarters and, therefore, allow us to identify prison and jail inmates, who respond to the same Census questionnaire as the general population.

We include in our sample males ages 20-60 for whom all the relevant variables are reported. Summary statistics are provided in Table 1. We create a dummy variable equal to 1 if the respondent is in a correctional institution. About 0.6% of men ages 20-60 in the US are interviewed in prison during each of the Census years we examine.[7] The percentage of high school drop outs decreases steadily from 52% in 1960 to 23% in 1980. Table 2 reports incarceration rates by race and age. The probability of imprisonment is substantially larger for blacks than for whites at all ages, and incarceration rates are declining with age for both races. The imprisonment rate for white males ages 20-29 is 0.7%, while the corresponding rate for blacks is more than 6 times larger. The black-white incarceration ratio is fairly stable across age groups, ranging from 5 to 7.5.

This paper focuses on the relationship between crime and education. Figure 1 clearly shows a decline in incarceration rates at the high school graduation stage that is substantially larger than at any other schooling progression. This, combined with our available instruments (compulsory schooling laws) which primarily impact schooling at grades 10-12, motivates our analysis of high school drop out rather than a more general study of the impacts of schooling at all grade levels. Table 3 reports differences in incarceration rates between high school drop outs and high school graduates. Throughout the paper, we define high school graduates to include anyone with 12 or more years of schooling and a high school degree.[8] The top panel reports incarceration rates for all men in our sample, while the bottom panel reports separate figures for blacks and whites. In 1960, incarceration rates for drop outs and graduates were 1.0% and 0.2%, respectively. The stability in aggregate incarceration rates reported earlier in Table 1 masks the underlying trends for both education groups. Table 3 shows increasing incarceration rates within each schooling group over the 1970s. The difference in incarceration rates between high school graduates and drop outs has also increased over time, from 0.8% in 1960 to 1.1% in 1980. The substantial difference in graduate and drop out incarceration rates combined with the more than 25% increase in high school graduation rates explains why aggregate incarceration rates remained relatively stable over

[7] The years under consideration precede the massive prison build-up that began around 1980. Unfortunately, the public version of the 1990 Census does not identify inmates.

[8] We ignore the fact that in some years, high school graduation in South Carolina could be achieved with 11 years of schooling.

time while within education group incarceration rates rose.

The key feature to notice in Table 3 is that the difference between high school drop outs and graduates is substantially larger for backs than for whites in all years.[9] For example, in 1980 this difference is .7% for whites and 2.1% for blacks. Obviously, these mean differences do not necessarily represent the causal effect of high school graduation on the probability of incarceration, since high school drop outs are likely to differ in many respects from individuals with more education. However, the patterns may indicate that the effect of education on imprisonment differs for blacks and whites. In the empirical analysis below, we allow for differential effects by race whenever possible.

To account for other factors in determining incarceration rates, we begin by using OLS to estimate the effect of drop out status on the probability of imprisonment. The top panel in Table 4 reports the coefficients on high school drop out from a linear probability model of imprisonment estimated on the sample of white males.[10] In column 1, the only covariates are year dummies that absorb general trends in incarceration. The point estimate suggests that high school drop outs are 0.6 percentage points more likely to be in prison than individuals with at least a high school degree.[11]

Age, state of birth, state of residence, and cohort of birth are important determinants of incarceration that could induce spurious correlation, resulting in biased estimates. Column 2 accounts for differences in age and state of birth by including 14 dummies for 3-year age groups (20-22, 23-25, 26-28, etc;) and 49 dummies for state of birth (excludes Alaska and Hawaii since our instruments below are unavailable for those states). To account for the many changes that affected Southern born blacks after Brown v. Board of Education, we also include a state of birth specific dummy for black men born in the South who turn age 14 in 1958 or later. The point estimate in column 2 is slightly larger than in column 1 indicating that, on average, state of birth effects and age effects are negatively correlated with drop out. The increase in the coefficient on drop out is mainly driven by the inclusion of age effects, since drop out rates increase with age in our sample – due to the secular trend of increasing education throughout the century – and imprisonment rates decrease with age – since most crimes are committed by younger men. Not surprisingly, the age dummies (not reported) monotonically decrease.

In column 3, we absorb heterogeneity across states in the probability of incarceration by including dummies for state of residence. The point estimate is almost unchanged. We next account for unobserved differences across birth cohorts, allowing for differences in school quality

[9] A second striking feature is the difference in overall incarceration rates between blacks and whites. The probability of incarceration for black high school *graduates* is more than twice as large as the probability of incarceration for white *drop outs*.

[10] Because of sample size considerations, we include only a 90% random sample of white males.

[11] The standard errors in this section are corrected for state of birth - year of birth clustering, since our instrument below varies at the state of birth - year of birth level.

or youth environments, in columns 4 and 5. Column 4 includes dummies for decade of birth (1914-1923, 1924-1933, etc.). Column 5 includes state-specific linear trends in year of birth to allow for the possibility that different states followed different paths for school and cohort quality. The final column also accounts for state of residence×year effects. This absorbs state-specific time-varying shocks or policies that may affect the probability of imprisonment and drop out. For example, an increase in prison spending in any given state may be offset by a decrease in education spending that year. (Notice, however, that since prison inmates may have committed their crime years before they are observed in prison, the state of residence×year effects are an imperfect control.) The estimates are insensitive to these additional controls. Overall, the estimates suggest that high school graduation among whites is associated with a decrease in their probability of imprisonment by about 0.76 percentage points, which is only slightly larger than the unconditional estimate obtained by differencing the average imprisonment rates in Table 3 (0.6 percentage points).[12]

The bottom panel in Table 4 reports analogous estimates for black men. The estimated effect of high school drop out on crime among blacks is about 1.9 percentage points before age and state of birth effects are accounted for (column 1). Adding age and state of births dummies (column 2) raises the coefficient to 3.4. This increase is similar to that observed for whites (the reason for the increase is the same), although it is quantitatively larger for blacks. In columns 2-6 the estimates are stable around 3.4 percentage points. Overall, the effects for blacks are significantly larger than those found for whites, which is consistent with the larger difference in unconditional means reported in Table 3.

To help in interpreting the size of the drop out impacts on incarceration, one can use these estimates to calculate how much of the black-white gap in incarceration rates is due to differences in drop out rates. In 1980, the difference in incarceration rates for whites and blacks is about 2.4%. Using the OLS estimates for blacks, we conclude that 25% of the difference in incarceration rates between blacks and whites could be wiped out by reducing black drop out rates to the same level as that of whites.

Alternatively, consider that the fall in white and black drop out rates between 1970 and 1980 are 13 and 21 percentage points, respectively.[13] In this case, the fall in drop out rates among whites should have caused a decrease of 0.1 percentage points in their incarceration rates (compared to a base of 0.6% among white drop outs in 1970). The corresponding figure for blacks is 0.6 percentage points (compared to a base of 2.9% among black drop outs in 1970). Later, we show that these effects are consistent with predicted declines in crime associated with the wage increases that accompany high school completion, using the elasticities of criminal participation

[12]If some inmates graduate from high school while in prison, these estimates will be biased toward finding no effect of graduation on crime.

[13]These drop out rates refer to the proportion of 20-60 year old men who did not finish high school and not the drop out rate of cohorts from those years.

and arrest with respect to wages that have been estimated in the literature.

To probe the robustness of our results, we explored a number of alternative specifications. First, probit models yield similar estimated effects of drop out. Second, models that are linear in years of schooling produce estimates that lead to similar conclusions. For example, coefficients on years of schooling in a specification similar to the those in columns 4-6 are -.0009 (.00002) for whites and -.0036 (.0001) for blacks.[14] It is not surprising that these estimates are smaller given the patterns observed in Figure 1 and the fact that a linear specification averages the impacts for all schooling transitions.

3.2 The Effect of Compulsory Attendance Laws on Schooling Achievement

The OLS estimates just presented are consistent with the hypothesis that high school graduation reduces the probability of imprisonment. If so, the effect appears to be statistically significant for both whites and blacks, and quantitatively larger for blacks. However, these estimates may reflect the effects of unobserved individual characteristics that influence the probability of committing crime and dropping out of high school. For example, the theoretical model in Section 2 suggests that individuals with a high discount rate or taste for crime, presumably from more disadvantaged backgrounds, are likely to commit more crime and attend less schooling. To the extent that variation in unobserved discount rates and criminal proclivity across cohorts is important, OLS estimates could overestimate the effect of schooling on imprisonment.

It is also possible that juveniles who are arrested or confined to youth authorities while in high school may may face limited educational opportunities. Even though we examine men ages 20 and older, some are likely to have been incarcerated for a few years, and others may be repeat offenders. If their arrests are responsible for their drop out status, this should generate a negative correlation between education and crime. Fortunately, this does not appear to be an important empirical problem.[15]

The ideal instrumental variable induces exogenous variation in high school drop out status but is uncorrelated with discount rates and other individual characteristics that affect both imprisonment and schooling. We use changes over time in the number of years of compulsory education that states mandate as an instrument for high school drop out. Years of compulsory attendance are defined as the maximum between (i) the minimum number of years that a child is required to stay in school and (ii) the difference between the earliest age that he is required to be in

[14]IV estimates using the instrument described below are statistically similar to these OLS estimates.

[15]A simple calculation using NLSY data suggests that the bias introduced by this type of reverse causality is small. The incarceration gap between high school graduates and drop outs among those who were not in jail at ages 17 or 18 is 0.044, while the gap for the full sample is only slightly larger (0.049). Since the first gap is not affected by reverse causality, at most 10% of the measured gap can be explained away by early incarceration resulting in drop out. If some of those who were incarcerated would have dropped out anyway (not an unlikely scenario), far less than 10% of the gap is eliminated.

school and the latest age he is required to enroll.[16] Figure 2 plots the evolution of compulsory attendance laws over time for 49 states (all but Alaska and Hawaii). In the years relevant for our sample, 1914 to 1974, states changed compulsory attendance levels several times, and not always upward.

We assign compulsory attendance laws to individuals on the basis of state of birth and the year when the individual was 14 years old. To the extent that individuals migrate across states between birth and age 14, the instrument precision is diminished, though IV estimates will still be consistent. We create four indicator variables, depending on whether years of compulsory attendance are 8 or less, 9, 10, and 11 or 12.[17] The fractions of individuals belonging to each compulsory attendance group are reported in Table 1.

In examining the effect of compulsory attendance laws on educational achievement, our specifications include controls for age, year, state of birth, state of residence, and cohort of birth effects. To account for the impact of Brown v. Board of Education on the schooling achievement of Southern born blacks, we also include an additional state of birth dummy for black cohorts born in the South turning age 14 in 1958 or later. Identification of the estimates comes from *changes* over time in the number of years of compulsory education in any given state. The identifying assumption is that conditional on state of birth, cohort of birth, state of residence and year, the timing of the changes in compulsory attendance laws within each state is orthogonal to characteristics of individuals that affect criminal behavior like family background, ability, risk aversion, or discount rates.

The top panel in Table 5 reports the relevant estimates for whites. Three points are worth making. First, the more stringent the compulsory attendance legislation, the lower is the percentage of high school drop outs. In states/years requiring 11 or more years of compulsory attendance, the percentage of high school drop outs is 5.5% lower than in states/years requiring 8 years or less (the excluded case). These effects have been documented by Acemoglu and Angrist (2000).[18] Second, the coefficients in columns 1 and 2 are roughly equal, but with opposite sign. For example, in states/years requiring 9 years of schooling, the share of high school drop outs is 3.3 percentage points lower than in states/years requiring 8 years or less of schooling; the share of high school graduates is 3.3 percentage points higher. This suggests that compulsory attendance legislation does reduce the number of high school drop outs by 'forcing' them to stay in school. Third, the effect of compulsory attendance is smaller, and in most cases, not significantly different

[16]The data sources for compulsory attendance laws are given in Appendix B of Acemoglu and Angrist (2000).

[17]We use the same cut off points as Acemoglu and Angrist (2000). We experimented with an instrument linear in the number of years of compulsory attendance, and found qualitatively similar results. We also experimented with a matching based on the year the individual is age 16 or 17, and found qualitatively similar results.

[18]Having a compulsory attendance law equal to 9 or 10 years has a significant effect on high school graduation. Possible explanations include "lumpiness" of schooling decisions (Acemoglu and Angrist, 2000), educational sorting (Lang and Kropp 1986) or peer effects.

from zero in columns 3 and 4. Finding a *positive* effect on higher levels of schooling may indicate that the laws are correlated with an underlying trend of increasing education, which would cast doubt on their exogeneity. This does not appear to be a problem in the data. Although very small in magnitude, the coefficient on compulsory attendance ≥ 11 is significantly different from 0 for individuals with some college. Surprisingly, it is negative, suggesting that states imposing the most stringent compulsory attendance laws experience small declines in the number of individuals attending college. This result may indicate a shift in state resources from local colleges to high schools following the decision to raise compulsory attendance laws.

The bottom panel in Table 5 reports the estimated effect of compulsory attendance laws on the educational achievement of blacks. These estimates are also generally consistent with the hypothesis that higher compulsory schooling levels reduce high school drop outs rates, although the coefficients in column 1 are not monotonic as they are for whites. The coefficients in column 3 are negative, suggesting that increases in compulsory attendance are associated with decreases in the percentage of black men attending local colleges. The magnitudes are smaller than the effect on high school graduation rates, but they are statistically significant and larger than the corresponding coefficients for whites. This may reflect a shift in resources from local black colleges to white high schools, and to a lesser extent, to black high schools.[19] As expected, compulsory attendance laws have little effect on college graduation.

Are compulsory schooling laws valid instruments? We start to address this question by examining whether increases in compulsory schooling ages are associated with increases in state resources devoted to fighting crime. If increases in mandatory schooling correspond with increases in the number of policemen or police expenditures, IV estimates might be too large. However, we do not expect this to be a serious problem. First, legislators enacting the laws did not appear to be acting in response to problems with juvenile delinquency, youth unemployment, or other factors related to crime. (See, for example, Kotin and Aikman (1980) for a history of compulsory schooling legislation.) Second, in contrast to most studies using state policy changes as an instrument, simultaneous changes in compulsory schooling laws and increased enforcement policies are not necessarily problematic for the instrument in this study, since we examine incarceration among individuals many years after schooling laws are changed and drop out decisions are made. Recall that we assign compulsory attendance based on the year an individual is age 14, and our sample only includes individuals ages 20 and older. For the instrument to be invalid, state policy changes that take place when an individual is age 14 must directly affect his crime years later (in his twenties and thirties). In general, this does not appear to be a likely scenario. However, as an additional precaution, we absorb time-varying state policies in our regressions by

[19] To the extent that compulsory attendance laws reduce college attendance, IV estimates will be biased toward finding no effect (or even a negative effect) of drop out on crime.

including state of residence$\times$year effects.[20]

Finally, we directly test for whether increases in compulsory attendance laws are associated with increases in the amount of police employed in the state. We find little evidence that higher compulsory attendance laws are associated with greater police enforcement. Column 1 in Table 6 reports the correlation between the instruments and the per capita number of policemen in the state. Data on policemen are from the 1920 to 1980 Censuses.[21] Columns 2 and 3 report the correlation between the instruments and state police expenditures and per capita police expenditures, respectively, using annual data on police expenditures from 1946 to 1978. Data on police expenditures are from ICPSR Study 8706. No clear pattern emerges from columns 1 and 2, while there appears to be a negative correlation in column 3. Overall, we strongly reject the hypothesis that higher compulsory attendance laws are associated with an increase in police resources. If anything, per capita police expenditures may have *decreased* slightly in years when compulsory attendance laws increased (consistent with trade-offs associated with strict state budget constraints). Unfortunately, data on sentencing laws over a long enough time span proved to be elusive, so we could not examine the correlation between compulsory schooling laws and criminal sentencing laws.

Another important concern with using compulsory attendance laws as an instrument is that the cost of adopting more stringent versions of the laws may be lower for states that expect faster increases in high school graduation rates. It is, therefore, possible that changes in compulsory attendance laws simply reflect underlying state-specific trends in graduation rates. We address this issue in two ways. In the IV regressions below, we include state-specific linear time trends, so identification comes from deviations from the trend in schooling induced by changes in compulsory attendance laws. Here, we examine the relationship between future compulsory attendance laws and current graduation rates. If causality runs from compulsory attendance laws to high school graduation, we should observe that future laws do not affect current drop out rates conditional on current compulsory attendance laws. Results of this test are reported in Table 7. The coefficients in the first row, for example, represent the effect of compulsory attendance laws that are in place 4 years after individuals are age 14. All models condition on compulsory attendance laws in place when the individual is age 14, 15, 16, and 17 (these coefficients are not reported but are generally significant). To minimize problems with multicollinearity, we run separate regressions for each future year (i.e. each row is a separate regression), although results are similar when we run a single regression of compulsory attendance on all future years. Negative coefficient estimates on future schooling laws would be consistent with causality running from schooling

[20]We note, however, that state-time dummies at the time of incarceration may be noisy measures of period effects at the time the crime was committed.

[21]The number of policemen in 1920, 1930, and 1940 are taken from Census reports on occupations and the labor force for the entire U.S. population. Data from 1950, 1960, 1970, and 1980 are from the IPUMS 1% Census samples.

levels to compulsory attendance, and would cast doubt on our identifying assumption. Overall, the results in Table 7 suggest that states with faster expected declines in drop out rates are not more likely to change their compulsory attendance laws.[22] This result is consistent with the findings of Lleras-Muney (2000) who examines these laws covering the first half of the twentieth century.

3.3 Instrumental Variable Estimates

We use an instrumental variable strategy to estimate specifications like those in Table 4. The top panel of Table 8 reports IV estimates for whites. IV estimates in columns 1 and 2 are very similar to the corresponding OLS estimates, although they are less precise. Estimates in columns 3 and 4 control for unobserved heterogeneity across cohorts that may introduce spurious correlation, and column 5 controls for state-specific time effects.

In general, the estimated effect of drop out on the probability of imprisonment is quite stable around .8-.9% across all specifications except column 4. This specification replaces nationwide cohort effects (column 3) with state-specific trends in cohort of birth. Identification derives from discontinuous changes (or deviations from the state trend) in drop out rates that occur when schooling attendance laws change. This accounts for the possibility that states in which drop out rates are falling faster may be more likely to increase compulsory attendance levels, absorbing state-specific long run trends in imprisonment and high school graduation. However, it clearly demands a lot from the data. The first stage estimated coefficients are smaller than those in other columns, and the second stage estimated effect is more than twice as large as the estimates in other specifications. The standard error more than doubles. Such a decrease in precision is not surprising given the loss in identifying variation in the first stage.

The bottom panel of Table 8 reports IV estimates for blacks. The estimated effects range from 0.06 to 1.0, and all are statistically significant except that of column 4. As with whites, the inclusion of the state-specific cohort trend in column 4 absorbs most of the variation in the first stage, so that the second stage estimate is not well-identified and the standard error is large. The most robust estimates suggest an impact of around 0.07-0.08, roughly twice the OLS estimate for all blacks. However, a Hausman test fails to reject the hypothesis that OLS estimates are unbiased and equal to the IV estimates, which is not surprising given the size of the standard errors for the IV estimates. (We also report the test statistic for an F-test of whether the compulsory schooling attendance laws all have zero coefficients.)

[22] Only one estimated coefficient for whites is significantly negative (t=+18). The only significant negative coefficients for blacks refer to laws 15 or more years in the future, too far ahead to be comfortably interpreted as causal. Furthermore, for those years where the coefficients are negative, there is no relationship between stringency of the law and high school drop out, making it difficult to interpret this finding.

3.4 Interpreting IV Estimates

A potentially important source of bias in OLS estimation is unobserved individual heterogeneity in criminal ability and/or discount rates (see Section 2), which would suggest that OLS estimates are biased towards finding too large an effect. Therefore, finding a larger IV estimate for blacks is surprising. Despite the fact that standard statistical tests do not reject that the OLS and IV estimates are identical, it is worth exploring potential reasons for a larger IV estimate.

The interpretation of IV and the comparison with OLS estimates requires care when the instruments affect schooling progressions at many different levels. Figure 1 suggests that the effect of schooling on imprisonment is not constant but varies widely across schooling levels. For example, the reduction in probability of imprisonment that occurs moving from 11 to 12 years of schooling is more than three times larger than the reduction that occurs moving from 10 to 11 years of schooling. This motivates our emphasis on high school drop out. However, when estimating the impact of high school drop out on the probability of incarceration, IV estimates reflect a weighted sum of causal responses to each single-year change in schooling (i.e. moving from 9 to 10 years, 10 to 11 years, 11 to 12 years, etc.), with weights that depend on the fraction of individuals induced to make each change in schooling by the compulsory attendance laws (Angrist and Imbens 1995). On the other hand, OLS estimates reflect the average difference in incarceration rates between high school graduates and drop outs, which represents a different weighted sum of causal responses.

Here, we build on the Angrist and Imbens result to show under what conditions OLS and IV strategies estimate different parameters. Both OLS and IV estimators can be written as a weighted sum of causal responses to a unit change in schooling, with different weights for the two estimators. As a result, the expected difference between the estimators depends on the difference in the weights as well as the impacts of schooling on incarceration at each grade level. Since these weights are observable quantities, we can determine how much of the difference between OLS and IV estimates for blacks and whites can be explained by the different weighting schemes and how much reflects other factors.

Consider an outcome $y_i(s) = \sum_{j=0}^{s} \beta_j + \varepsilon_i$, which depends on the level of schooling $s = 1, 2, ..., S$ and a mean zero iid error term ε_i. If D is an indicator equal to one when $s \geq k$ (in our case, $k = 12$, so D represents high school graduation) and zero otherwise (for purposes of this analysis, assume D is independent of ε), then the OLS regression

$$y_i = \alpha + \beta D_i + \varepsilon_i \tag{2}$$

yields an estimate for β with the expectation

$$E(\hat{\beta}_{OLS}) \quad = \quad E(y|s \geq k) - E(y|s < k)$$

$$= \beta_k + \sum_{j=1}^{k-1} \phi_j \beta_j + \sum_{j=k+1}^{S} \theta_j \beta_j,$$

where $\phi_j = Pr(s < j | s < k)$ and $\theta_j = Pr(s \geq j | s \geq k)$. The β's are pictured in Figure 1.

If an instrumental variable, $Z \in \{0, 1\}$, exists and satisfies the monotonicity and independence assumptions of Angrist and Imbens (1995), then the expected value of the IV estimate for β in equation (2) is given by

$$\begin{aligned} E(\hat{\beta}_{IV}) &= \frac{E(y|Z=1) - E(y|Z=0)}{E(D|Z=1) - E(D|Z=0)} \\ &= \beta_k + \sum_{j \neq k} \omega_j \beta_j, \end{aligned}$$

where $\omega_j = \frac{Pr(s_1 \geq j > s_0)}{Pr(s_1 \geq k > s_0)}$ and s_z is the schooling choice for someone when $Z = z$.

More generally, the instrument may take on more than two values: $Z \in \{0, 1, 2, ..., I\}$. In our case, Z=0 if compulsory schooling is 8 years or less; Z=1 if compulsory schooling is 9 years; Z=2 if compulsory schooling is 10 years; Z=3 if compulsory schooling is 11 or 12 years. The two stage least squares estimate for β using I indicator variables ($z_i = 1$ if $Z = i$ and zero otherwise) as instruments for D in equation (2) has expectation

$$\begin{aligned} E(\hat{\beta}_{2SLS}) &= \frac{E\{y[E(D|Z) - E(D)]\}}{E\{E(D|Z)[E(D|Z) - E(D)]\}} \\ &= \beta_k + \sum_{j \neq k} \lambda_j \beta_j, \end{aligned}$$

where

$$\lambda_j = \frac{\sum_{i=1}^{I} \Pr(Z=i)[E(D|Z=i) - E(D)]Pr(s_i \geq j > s_0)}{\sum_{i=1}^{I} \Pr(Z=i)[E(D|Z=i) - E(D)]Pr(s_i \geq k > s_0)}. \tag{3}$$

The 2SLS estimator is a weighted sum of the causal responses to each unit change in schooling. The λ-weights depend on the number of individuals whose education level is affected by each of the instruments ($Pr(s_i \geq j > s_0)$) and the distance between $E(D|Z = i)$ and E(D). See Appendix A for the derivation of the λ-weights.

In general, the OLS, IV, and 2SLS estimates need not coincide. All three estimators consistently estimate different functions of the underlying β's. Comparing the 2SLS and OLS estimates produces the following relationship:

$$E(\hat{\beta}_{2SLS}) - E(\hat{\beta}_{OLS}) = \sum_{j=1}^{k-1} (\lambda_j - \phi_j)\beta_j + \sum_{j=k+1}^{S} (\lambda_j - \theta_j)\beta_j. \tag{4}$$

Equation 4 makes clear that two special cases will yield OLS and 2SLS estimates with the same expectation. First, schooling may have no effect on the outcome y at any level except moving from $s = k-1$ to $s = k$. In this case, $\beta_j = 0$ for all $j \neq k$, and the two estimators consistently estimate the effect on y of moving from $k-1$ to k years of school. Second, $\lambda_j = \phi_j$ for all $j < k$ and $\lambda_j = \theta_j$ for all $j \geq k$. In this unlikely event, both estimators will yield the same weighted sums of all causal effects. Only if all weights are identically zero (so the instrument has no effect on any grade transition other than moving from $k-1$ to k) will both estimates consistently estimate β_k.

Since the λ's, ϕ's and θ's are functions of observable quantities, equation (4) can be estimated. Figure 3 plots the 2SLS and OLS weights as well as the difference between the two sets of weights for whites and blacks.

Under the assumption that the OLS estimates of β_j reflect the causal responses to a unit change in schooling (i.e. there is no correlation between the error term and D),

$$\begin{aligned} E(\hat{\beta}_{2SLS}) - E(\hat{\beta}_{OLS}) &= -0.0001 \quad \text{for whites} \\ &= 0.0159 \quad \text{for blacks.} \end{aligned} \tag{5}$$

In other words, when D_i and ε_i are uncorrelated, we should expect 2SLS estimates to be larger than OLS estimates for blacks and roughly equal to OLS estimates for whites. In fact, the expected differences between 2SLS and OLS estimates in equation (5) are remarkably consistent with the actual differences reported in Tables 4 and 8 for both whites and blacks. We, therefore, conclude that the larger 2SLS estimates for blacks are not an anomaly, but they reflect differences in the weights that OLS and 2SLS give to the effect of different schooling levels on crime. This is also suggestive that the OLS estimates reflect causal impacts and are not biased by endogeneity of drop out status. Some of the discrepancy may also be explained by the lack of precision in our 2SLS estimates. After all, we cannot statistically reject that the 2SLS and OLS estimates are the same.

We also explore two other potential explanations for the difference between OLS and 2SLS estimates for blacks, but we find little empirical support for either of them. First, we study whether heterogeneous rates of return to schooling can explain the discrepancy. In most cases, IV estimates of the Mincerian rate of return to schooling are greater than OLS estimates despite the presumption that OLS estimates should be biased upward (Card 1995). IV estimates reflect a local average effect for those individuals *induced to graduate from high school by compulsory attendance laws* (Imbens and Angrist 1994) and not the average effect of high school graduation in the entire population, which is the parameter estimated by OLS. This is relevant to our analysis because of the link between wages and crime implied by economic theory. If the return to schooling is higher for individuals induced to graduate from high school by compulsory attendance laws, then IV estimates of the effect of education on crime may also predict a larger effect than OLS estimates imply. We find that IV estimates are indeed larger than OLS estimates, but they are larger for both blacks and whites.[23] IV estimates of the effect of drop out on crime are only larger for blacks, so this does not appear to be a decisive factor in explaining the discrepancy in IV and OLS estimates for blacks but not for whites.

Another potential reason for differences between OLS and IV may be the existence of spillover or contagion effects, which have been suggested by Glaeser, et. al (1996) and Gaviria and Raphael (2001). If individual decisions to commit crime depend on average education levels or crime rates for other individuals in their cohort, IV (using state-cohort level instruments as we do) will estimate the combined effect of own drop out on crime as well as the effect of average cohort drop out rates on crime. That is, IV will estimate the sum of the individual effect and spillover effect. If cross-state and cohort variation in average drop out rates are small relative to overall variation in drop out rates, then OLS will only estimate the individual effect of drop out on crime. See Appendix B for a formal discussion. Empirically, this does not appear to be an important factor.[24]

4 The Impact of Schooling on Arrest Rates

Evidence in the previous section is consistent with the hypothesis that high school graduation reduces criminal activity. One limitation of Census data is that they do not differentiate among different types of criminal offenses. In this section, we investigate the impact of education on specific crime rates by using data on arrests by offense. Because individual-level data that contain education of the arrested do not exist, we use arrest data collected by the FBI Uniform Crime Reports (UCR) by state, criminal offense, and age for 1960, 1970, 1980, and 1990. For each year and reporting agency, arrests are reported by age group, gender, and offense type. (Unfortunately, arrest rates are not reported by race in addition to state, age, and year.) We only study males ages 20-59 in our analysis. See Chilton and Weber (1998) for further details about the data.

To relate arrest rates to schooling and racial composition, we augment the arrest data with drop out rates by age and the percentage black in each state from the 1960-1990 Censuses. We estimate the following model:

$$\ln A_{cast} = \beta D_{ast} + \gamma B_{ast} + d_{st} + d_{sc} + d_{sa} + d_{ct} + d_{at} + d_{ac} \tag{6}$$

[23]OLS estimates (using a specification that controls for year, age, state of birth, state of current residence, and cohort of birth and Census data) are 0.061 (0.0004) for whites and 0.056 (0.0006) for blacks. IV estimates are 0.139 (0.016) and 0.095 (0.015), respectively.

[24]OLS regressions using cohort-year-state aggregate incarceration rates and drop out rates yield estimates similar to those produced by our individual-level regressions. The coefficient on drop out rates in a regression of cohort-year-state incarceration rates on cohort-year-state drop out rates that controls for cohort, state and year effects is .008 (.001) for whites and 0.020 (0.006) for blacks. Since regressions based on aggregate measures should produce estimates of the combined individual and spillover effects, the similarity in estimates suggests that spillover effects at the state-cohort level are not important in determining incarceration rates. Other attempts to estimate spillover effects at the state-cohort level yielded similar conclusions.

where $\ln A_{cast}$ is the logarithm of the male arrest rate for crime c, age group a, in state s in year t (from UCR); D_{ast} is the male drop out rate for age group a in state s at time t (from Census); B_{ast} is the percent of males that are black in age group a in state s at time t (from Census). In using log arrest rates, the effect of drop out rates on arrest rates is assumed to be the same *in percentage terms* for all crimes.[25] In a few specifications, we allow the effect of drop out rates to vary by type of crime (β_c).

The d's represent indicator variables that account for unobserved heterogeneity across states, years, cohorts, and criminal offense types. In particular, d_{st} is a state$\times$year effect that absorbs time varying, state-specific shocks that may induce spurious correlation. The level of arrests reflects both the level of criminal activity and police resources devoted to making arrests. If a state decides to reduce spending for public education and increase spending for police or prisons, a spurious positive correlation between arrests and drop out rates may arise. Including state-year effects is more robust than including observable state-level variables reflecting differences in spending or punishment. Since for each state-year combination there are many age groups in our data, we can control for unrestricted state-specific time-varying shocks without fully saturating the model.

In estimating equation (6), the distribution of crimes across states does not need to be uniform. Some states may focus arrests more heavily on some types of crimes than others, either because more of those crimes are committed or because that state is simply harsher on those crimes. Also, the age of arrestees need not be the same across states – some age groups may be more prone to commit crimes in some states or the arrest policy with respect to age may differ across states. The terms d_{sc} and d_{sa} absorb permanent state$\times$crime and state$\times$age heterogeneity in arrest rates.

Crime-specific and age-specific trends in arrest common to all states are accounted for by crime$\times$year dummies, d_{ct}, and age$\times$year dummies, d_{at}, respectively. Finally, age$\times$crime effects, d_{ac}, account for the fact that some age groups might always be more likely to commit certain types of crimes and to be arrested for those crimes. In the data, we have 8 age groups (20-24, 25-30, etc.), 9 crimes (murder, rape, assault, robbery, burglary, larceny, auto theft, and arson), and 51 states.

Most crimes do not result in an arrest. We are interested in arrests, however, because there is presumably a link between the amount of crime that takes place and the number of arrests that are made. To establish that link, we first compare our arrest data with crime reported to the police in the FBI's Uniform Crime Reports. The crime reported to the police in the UCR is used by the FBI to calculate official crime rates. The average arrest-crime ratio across all years and states is 0.6 for murder, and declines substantially as we move toward less serious

[25]This assumption is consistent with that made by Levitt (1998). We also estimated specifications in arrest rates (rather than log arrest rates) and arrived at similar conclusions.

crimes. Although this fact suggests that very few arrests are made for each crime committed, the correlation between arrests and crimes committed is remarkably high: 0.97 for burglary, 0.96 for rape and robbery, 0.94 for murder, assault and burglary, and 0.93 for motor vehicle theft. This suggests that variation in arrest rates closely tracks variation in actual crimes committed.[26]

The estimated impacts of high school drop out on arrest rates are reported in the top panel of Table 9. All models are weighted by cell size. Since variation in arrest rates occurs across offense type, age, state, and year, and variation in drop out rates occurs across age, state, and year, standard errors are corrected for state-year-age clustering. The estimates in column 1 control for age, year, state, and offense effects, and columns 2-4 progressively control for more interaction terms. The estimated impacts of high school drop out rates on log arrest rates range from 0.4-0.7 depending on the specification (all are statistically significant). Richer specifications suggest a sizeable impact of around 0.7, suggesting that a 10 percentage point decrease in drop out rates would reduce arrest rates by about 7%.

The bottom panel of Table 9 allows for differential effects of high school drop out rates on violent and property crimes. Again, estimates vary across specifications (with the impacts on violent crime typically greater than on property crime), but the richer specifications of the final two columns suggest that a ten percentage point decrease in drop out rates is associated with a 6% reduction in property crime rates and an 8% reduction in violent crime rates. The difference is not statistically significant. Note that when an individual is arrested for committing more than one crime, only the most serious is recorded. For example, if a murder is committed during a burglary, the arrest is recorded as murder. This may blur the distinction between violent and property crime. We discuss other potential reasons for the slightly larger effects on violent crime below.

Because arrest rates are not reported by race in addition to state, age, and year, it is difficult to determine whether drop out has differential effects on arrest by race. We attempt to examine this issue by controlling for both the share of black and white drop outs in the state. To do this, we interact black (and white) drop out rates by age and state with the fraction of men who are black (and white) in that same age and state category. If total arrests are the sum of arrests for blacks and for whites, then coefficients on these variables will give us the impacts of drop out on arrests for each race. In a specification analogous to that of column (3) in Table 9, we find evidence that black drop out rates have greater effects on crime than do white drop out rates. The coefficient estimate for the interaction of black drop out rates with percent black and violent

[26]Levitt (1998) transforms arrest rates into implied crime rates using the following algorithm: $Crime_{ast} = Arrest_{ast} \times (Crime_{st}/Arrest_{st})$ under the assumption that the number of crimes committed by a cohort in a given state and year is proportional to that cohort's share of total arrests in that state and year. Since we use the *logarithm* of arrests, and we control for state$\times$year effects, our specification is similar to Levitt's (1998). (They would be identical if we studied only one type of crime.)

crime is 2.49 (0.49), while it is 1.50 (0.49) for property crime. The corresponding estimates for whites are only 0.38 (0.24) and 0.31 (0.25).[27]

In Table 10, we allow for differential effects of drop out rates across more detailed offense categories. The table suggests that the effects vary considerably within the broad categories of violent and property crime.[28] For example, the estimates imply that a ten percentage point decrease in drop out rates would reduce murder and assault arrest rates by about 20%, motor vehicle theft by about 13%, arson by 8%, and burglary and larceny by slightly less than 3% (which is insignificant from zero). Estimated effects on robbery are not statistically different from zero, while those for rape are significantly negative. This final result is surprising and not easily explained by standard economic models of crime.[29]

As a whole, these OLS results suggest that high school completion is negatively correlated with many types of crime even after controlling for a rich set of covariates that absorb heterogeneity at the state, year, crime, and age level. Furthermore, high school drop out rates appear to have a slightly larger effect on violent crimes (especially murder and assault) than property crimes. This may be surprising since one channel through which schooling can affect crime is through raising wage rates and, therefore, the opportunity costs of crime. But, it is consistent with the fact that punishments for violent crimes typically involve substantially longer prison sentences, which are more costly when wages and schooling are high. And, to the extent that schooling increases patience levels or risk aversion, the long prison sentences associated with violent crimes become more costly. Non-economic factors may also play an important role in determining criminal activity. For example, finishing high school may cause individuals to change their lifestyles, residential location, or peer groups, reducing the amount of criminal opportunities they come into contact with and choose to engage in. Finally, the large coefficients on murder and assault may, in part, reflect the fact that only the most serious crime gets reported by the FBI when multiple crimes are committed.

For completeness, we also report instrumental variable estimates based on changes in compulsory attendance legislation. Unfortunately, we cannot assign compulsory attendance directly to individuals as we could with the Census data. Nor can we assign compulsory attendance based on the *state of birth*, since it is not available in the FBI aggregate data. Instead, we assign the

[27] When we also control for state-specific year effects as in column (4) of Table 9, the lack of race-specific arrest rates makes precise estimation of race-specific drop out impacts difficult.

[28] In general, estimating the effects of drop out separately for each crime (rather than a grouped specification as in Table 10) produces qualitatively similar results, with two exceptions: separate estimates for rape are much closer to zero and insignificant (even positive for a specification similar to column 2) while separate estimates for assault become small and insignificant.

[29] We originally thought that it may be explained by differential reporting rates by education, with more educated women more likely to report a rape. To test this hypothesis we examined reporting rates from the National Criminal Victimization Survey (U.S. Dept. of Justice 2000), but we failed to find evidence of such differential reporting. It is still possible that less educated women tend to be more restrictive in their definition of rape.

compulsory attendance laws based on the state where the arrest took place and the year the arrestees were age 14. Because of these data limitations, we expect a substantial decrease in precision.

Instrumental variable estimates are reported in Table 11. First stage coefficients are consistent with a negative effect of compulsory attendance laws on drop out, as was found in the Census data. More importantly, second stage estimation controlling for numerous state, age, offense, and year interactions suggests a sizeable impact of drop out rates on arrest rates that is consistent with the OLS estimates, though the estimates are slightly larger and considerably less precise here: a ten percentage point decrease in drop out rates causes an 8.7-9.5% reduction in arrest rates.

Are these estimates consistent with the Census-based incarceration estimates of the previous section? As discussed in Section 2, if sentence lengths or the probability of incarceration given arrest are greater for less educated individuals, the log difference in incarceration rates by drop out status should exceed the log difference in arrest rate by the log difference in the probability of incarceration given arrest. Since Mustard (2001) finds differences of only 2-3% in sentencing by drop out status, we should expect comparable effects of drop out on log arrest rates and log incarceration rates. From Table 3, the log difference in incarceration rates for all men in the Census is about 1.4 (OLS and IV estimates produce similar effects for whites and larger impacts for blacks). The IV estimates in Table 11, obtained using data on all offenses, suggest that drop out increases arrest rates among all men by about 1 log point. OLS estimates suggest an overall effect of about 0.7 log points, while crime-specific estimates suggest effects as large as 2.2 log points for violent crimes (carrying a long prison sentence) such as assault and murder. These simple comparisons suggest that the estimated effects on arrest and incarceration rates are roughly consistent.

One might also expect effects of this magnitude based on the estimated impact of increased wage rates on crime and arrest rates. For example, Grogger (1998) estimates an elasticity of criminal participation with respect to wages of around 1-1.2 using self-report data from the NLSY. Gould, et. al (2000) estimate the elasticity of arrest rates to the local wage rates of unskilled workers to be in the neighborhood of 1-2. When using March CPS data from 1964-90, a standard log wage regression controlling for race, experience, experience-squared, year effects, and college attendance yields an estimated coefficient on high school graduation of 0.49. Combining this estimate of the effect of schooling on wages with the elasticity of arrests with respect to wages estimated by Gould, et. al (2000) produces an impact of 0.5-1.0. That is, a 10% increase in high school drop out rates should increase arrest rates by 5-10% through increased wages alone. This covers the range of estimates in Tables 9 and 11 and confirms that an important explanation for the effect of high school drop out on crime resides in the lower wage rates associated with dropping out of high school.

5 The Impact of Schooling on Criminal Participation and Incarceration in the NLSY

Since crime is not directly observed, we have used data on arrests and incarceration to estimate the impacts of drop out on crime. Those results suggest that high school drop out is associated with a higher probability of arrest and imprisonment. Because those estimates can confound the effects of drop out on actual crime with any educational differences in the probability of arrest or incarceration conditional on commission of a crime (Section 2), we turn to the National Longitudinal Survey of Youth to study the relationship between education and self-reported crime. Although self-reported crime may suffer from under-reporting, it is the most *direct* measure of criminal participation available.

The NLSY also offers an abundance of individual-level variables that may determine crime but which are not available in the Census or arrest data we have used thus far. Therefore, a second important advantage of the NLSY is that it can be used to determine the robustness of our earlier results to the inclusion of more control variables likely to be related to crime. In particular, the survey records scores on the Armed Forces Qualifying Test (AFQT) that can be used as a measure of cognitive ability. Parents' age and education are available, as is family income. The NLSY also indicates whether or not individuals lived with both of their natural parents at age 14 and whether the mother was a teenager when she gave birth. Because the NLSY follows respondents who become incarcerated, we are able to verify our Census-based findings in Section 3.

We create three self-reported crime categories corresponding to more serious offenses: (i) property crimes consist of thefts greater than or equal to $50 as well as shop-lifting; (ii) violent crimes consist of using force to get something or attacking with intent to injure or kill (i.e. robbery and assault); and (iii) drug crimes consist of selling marijuana or hard drugs. Individuals are considered to be incarcerated if (i) they were surveyed in prison or (ii) they reported incarceration as a reason they were not looking for work when they were unemployed during the survey year (post-1988 only).

While it is virtually impossible to verify self-reported crime, most studies agree that young black men are more likely to under-report their criminal behavior than young white men. (See for example the exhaustive study by Hindelang, Hirsch, and Weis (1981).) Our calculations based on NLSY data suggest that black drop outs may be substantially under-reporting criminal activity, while there is less reason to believe that black high school graduates and whites are under-reporting to the same degree.[30] Because a correlation between under-reporting and education

[30] Among black drop outs, the self-reported crime rate at ages 18-23 is 0.22, but the incarceration rate over ages 22-28 is 0.32. While self-reported criminal activity may suffer from under-reporting, the incarceration data

would bias any estimates of the impact of drop out on crime, we focus attention on white men in the NLSY.

Columns 1 and 2 in Table 12 show average criminal participation and incarceration rates among young white men in the NLSY based on their high school graduation status. Self-reported crime measures are for men ages 18-23 in 1980, while incarceration measures represent the annual rate of incarceration over ages 22-28. Column 3 shows that self-reported criminal participation in all types of crime and incarceration rates are significantly lower for high school graduates. To account for individual differences that may be correlated with schooling, we estimate probit models of self-reported crime and incarceration. Two goals are pursued. First, we examine the impacts of schooling on self-reported crime to compare with the results for arrests and incarceration. Second, to determine the robustness of our findings, we explore much richer specifications that control for family background, individual ability, and local labor markets.

We begin with sparse specifications analogous to those used in the previous sections, controlling for age and region of residence. High school graduation significantly reduces participation in violent, property, and drug crimes. The average estimated effect of high school graduation on participation is reported in column 4 of Table 12.[31] Graduation reduces participation rates in violent crime by 0.09, drug sales by 0.04, property crime by 0.07, and overall criminal participation by 0.11. Column 5 controls for age, family background,[32] ability (as measured by AFQT percentile), race and ethnicity, geographic location (region of residence and SMSA status), local unemployment rates, and statewide incarceration rates (the ratio of prisoners to crimes committed in the state as taken from Levitt (1998)).[33] The striking result is that these estimates obtained by conditioning on a rich set of individual and family background characteristics are nearly identical to the raw differences. In other words, ignoring cognitive ability and family background does not introduce an upward bias in estimating the effect of high school graduation on criminal participation.

How do these effects compare with our findings for arrest rates? We compare arrest results

are reliable, since they are primarily based on whether the respondent is interviewed in prison. Given that crime typically declines with age among adults and 32% of the black high school drop outs in the sample were incarcerated over ages 22-28, it seems highly unlikely that only 22% of young black drop outs participated in crime just a few years earlier. In the absence of gross incarceration of innocent black men, it is likely that black drop outs substantially under-reported their criminal involvement in the NLSY. Among whites and black graduates, self-reported crime rates are more consistent with subsequent incarceration rates. As a result, differential reporting by drop out status is likely to be less of a problem among whites. More accurate reporting among whites accords with previous studies (Hindelang, Hirsch and Weis 1981).

[31] Average effects correspond to the discrete version of an average derivative, measuring the average difference in the probability of participation associated with a change in high school graduation status. Coefficient estimates are available from the authors upon request.

[32] Family background measures include: current enrollment in school, parents highest grade completed, whether or not the individual lived with both of his natural parents at age 14, whether his mother was a teenager at his birth, and family income.

[33] OLS estimates of the effect of schooling on self-reported crime among blacks are quite similar to the raw differences reported in Table 12.

from Table 9 with the log difference in self-reported crime by high school drop out status reported in Table 12. The difference in self reported log violent crime rates is 0.92, slightly larger than the measured effect on violent arrests, 0.79. The difference in self reported log property crime rates is 0.43, slightly less than the estimated effect on property arrests, 0.62. These findings suggest that the estimated impacts of drop out on arrests and incarceration are not simply the result of differential treatment by police and judges. Completing high school has a real effect on crime that is measurably similar to its effects on both arrest and incarceration.[34] This reconciles with the finding of Mustard (2001) that average prison sentences are quite similar across high school graduates and drop outs.

We next examine the impact of education on incarceration in the NLSY to verify our earlier results using Census data. The estimated effects of high school graduation on incarceration during early adulthood are shown in the bottom panel of Table 12. As in Section 3, high school graduation significantly lowers the probability a young man will be incarcerated. Both sets of probit estimates are close to the raw difference. The final row suggests that graduation reduces the annual probability of incarceration by about 0.02-0.03 among white men ages 22-28.[35] While these estimated effects are larger than those found with the Census data, the discrepancy is explained by the fact that the Census estimates report incarceration rates over ages 20-60, while the NLSY-based estimates refer to men ages 22-28. When annual incarceration rates are compared for 22-28 year-old men, both data sources yield remarkably similar predictions.[36]

Two points are evident from the NLSY data. First, high school graduation significantly reduces self-reported crime, and the estimated effects are consistent with the impacts estimated for arrests and incarceration in Sections 3 and 4. This implies that the impacts estimated for arrests and incarceration in Sections 3 and 4 reflect a true effect on crime, and not simply educational differences in the probability of arrest or incarceration conditional on commission of a crime. Second, controlling for individual ability, family background, and local labor markets does not change the estimated effect.

[34]It should be noted that self-report estimates measure the effects on criminal participation at the extensive margin, so they need not correspond perfectly to arrest rates, which include changes at the intensive and extensive margin.

[35]These estimates adjust the impact of graduation on the probability of incarceration over the entire age span of 22-28 to an annual impact using the ratio of incarceration rates over the seven year period to the annual incarceration rate (a factor of 3). Though not shown, results for blacks are also consistent with Census results. A specification analogous to the one in column 5 of Table 12 suggests that high school graduation reduces the annual probability of imprisonment by about 0.06 (compared with a raw difference of 0.09) among 22-28 year-old black men. Results available from authors upon request.

[36]Annual incarceration rates among males ages 22-28 in the Census are 0.025 and 0.005 for white drop outs and graduates, respectively.

6 Social Savings from Crime Reduction

Given the estimated impact of high school drop out on crime, it is possible to determine the social savings associated with reducing drop out rates. Because the social costs of crime differ substantially across crimes, we use estimates based on the impact of drop out on arrests by offense type to determine the social benefits of increased education.

These estimates are subject to two important caveats. First, they assume that estimates in Table 10 produces a consistent estimate of the effect of graduation on arrest.[37] Second, consistent with most other studies of crime, these estimates do not account for general equilibrium effects on wages of increasing the supply of graduates. However, in Appendix C, we present a simple general equilibrium model to assess how sensitive our estimates of social savings might be to the inclusion of general equilibrium effects. The intuition of the model is very simple. An increase in the supply of high school graduates reduces their wage levels which should increase their crime rate. This would suggest that our social benefit calculations overestimate the true social savings. At the same time, however, a reduction in the supply of drop outs increases their wage rates which should decrease their crime rate causing us to understate the true social savings. A back of the envelope calculation reported in Appendix C suggests that the net effect of changing wages on crime is trivial. If anything, when we move 1% of the population from dropouts to graduates, the reduction in wages among graduates is more than offset by the increase in wages among drop outs, so that the net effect on crime when general equilibrium effects are included is no smaller than what we report.

Recognizing the limitations of the exercise, we nonetheless provide a rough estimate of the social savings from crime reduction resulting from a 1% increase in high school graduation rates. Columns 1 to 4 of Table 13 report the costs per crime associated with murder, rape, robbery, assault, burglary, larceny/theft, motor vehicle theft, and arson. Victim costs and property losses are taken from Miller, et al. (1996). Victim costs reflect an estimate of productivity and wage losses, medical costs, and quality of life reductions based on jury awards in civil suits. Incarceration costs per crime equal the incarceration cost per inmate multiplied by the incarceration rate for that crime (approximately $17,000).[38] Total costs are computed by summing incarceration costs and victim costs less 80% of property losses, which are already included in victim costs and may be considered a partial transfer to the criminal.[39] The table reveals substantial variation

[37]To the extent that the effects of drop out vary in the population, these estimates may be too large or too small depending on which individuals change graduation status. Our IV estimates suggest that those on the margin may show a slightly *larger* response.

[38]Incarceration rates by offense type are calculated as the total number of individuals in jail or prison (from Lynch, et al. (1994)) divided by the total number of offenses that year (where the number of offenses are adjusted for non-reporting to the police). Incarceration costs per inmate are taken from Stephan (1999). Offenses known to the police and reporting rates are given by the Uniform Crime Reports and National Criminal Victimization Survey.

[39]For the crime of arson, total costs equal victim costs plus incarceration costs, since it is assumed that none of

in costs across crimes: violent crimes like murder and rape impose enormous costs on victims and their family members, while property crimes like burglary and larceny serve more to transfer resources from the victim to the criminal.

It is important to recognize that many costs of crime are not included in this table. For example, the steps individuals take each day to avoid becoming victimized – from their choice of neighborhood to leaving the lights on when they are away from home – are extremely difficult to estimate. More obvious costs such as private security measures are also not included in Table 13. Even law enforcement (other than costs directly incurred when pursuing/solving a particular crime) and judicial costs are absent here, mostly because they are difficult to attribute to any particular crime. Finally, the costs of other crimes not in the table may be sizeable. Nearly 25% of all prisoners in 1991 were incarcerated for drug offenses, costing more than $5 billion in jail and prison costs alone (Lynch et al. 1994). Given the NLSY findings for the effects of high school graduation on drug offenses, there is good reason to believe these costs of crime are also relevant for this analysis.

Column 5 reports the predicted change in total arrests in the U.S. based on the arrest estimates reported in column 2 of Table 10 and the total number of arrests in the Uniform Crime Reports. Our estimates imply that nearly 400 fewer murders and 8,000 fewer assaults would have taken place in 1990 if high school drop out rates had been one percentage point lower. Column 6 adjusts the arrest effect in column 5 by the number of crimes per arrest. In total, nearly 100,000 fewer crimes would take place. The implied social savings from reduced crime are obtained by multiplying column 4 by column 6 and are shown in column 7. Savings from murder alone are as high as $1.1 billion. Savings from reduced assaults amount to nearly $370,000. Because our estimates suggest that drop out increases rape and robbery offenses, they partially offset the benefits from reductions in other crimes. The final row reports the total savings from reductions in all eight types of crime. These estimates suggest that the social benefits of a one percent reduction in male U.S. high school drop out rates (from reduced crime alone) would have amounted to $1.4 billion. And, these calculations leave out many of the costs associated with crime and only include a partial list of all crimes. Given these omissions, $1.4 billion should be viewed as an under-estimate of the true social benefit.

One might worry that our large estimated effects for murder combined with the high social costs of murder account for most of the benefits. When we, instead, use the estimated effects for violent and property crime in Table 9, the resulting total social benefits from crime reduce to $782 million. (An overly conservative estimate that only considered savings from reductions in incarceration costs would yield a savings of around $50 million.)

The social benefit *per additional male graduate* amounts to around $1,170-$2,100, depending

---

the property loss is transferred to the criminal.

on whether estimates in Table 9 or 10 are used. To put these amounts into perspective, it is useful to compare the private and social benefits of completing high school. Completing high school would raise average annual earnings by about $8,040.[40] Therefore, the positive externality in crime reduction generated by an extra male high school graduate is between 14% and 26% of the private return to high school graduation. The externalities from increasing high school graduation rates among black males are likely to be even larger given the larger estimated impacts of drop out on incarceration and arrest rates among blacks.

For another interesting comparison, consider what a one percent reduction in male drop out rates entails. The direct costs of one year of secondary school were about $6,000 per student in 1990. Comparing this initial cost with $1,170-$2,100 in social benefits per year thereafter reveals the tremendous upside of reducing drop out.[41]

How do these figures compare with the deterrent effects of hiring additional police? Levitt (1997) argues that an additional sworn police officer in large U.S. cities would reduce annual costs associated with crime by about $200,000 at a public cost of roughly $80,000 per year. To generate an equivalent social savings from crime reduction would require graduating 100 additional high school students for a one-time public expense of around $600,000 in schooling expenditures (and a private expense of nearly three times that amount in terms of foregone earnings). Of course, such a policy would also raise human capital and annual productivity levels of the new graduates by more than 40% or $800,000 based on our estimates using standard log wage regressions. So, while increasing police forces is a cost effective policy proposal for reducing crime, increasing high school graduation rates offers far greater benefits when both crime reductions and productivity increases are considered.

7 Conclusions

There are many theoretical reasons to expect that education reduces crime. By raising earnings, education raises the opportunity cost of crime and the cost of time spent in prison. Education may also make individuals less impatient or more risk adverse, further reducing the propensity to commit crimes. To empirically explore the importance of the relationship between schooling and criminal participation, this paper uses individual-level data from the Census on incarceration, state-level data on arrests from the Uniform Crime Reports, and self-report data on crime and

[40]This is based on a regression of log earnings on dummies for high school completion, college attendance, and other standard controls using males in the 1990 Census. The coefficient on the high school dummy, 0.42, was multiplied by $19,146, the average earnings for male workers with 10 or 11 years of schooling in the 1990 Census.

[41]Because the arrest estimates reflect the average difference between all high school graduates and all drop outs (rather than comparing those with 12 versus 11 years of schooling), the estimated benefits are likely to be greater than the benefits that result from simply increasing the schooling of those with eleven years by one additional year (recall the discussion in Section 3.4). However, as Figure 1 reveals, most of the reductions seem to be associated with finishing the final year of high school.

incarceration from the National Longitudinal Survey of Youth. All three of these data sources produce similar conclusions: high school graduation significantly reduces criminal activity. Our estimated effects of education on crime cannot be explained away by unobserved characteristics of criminals or unobserved state policies that affect both crime and schooling. Evidence from other studies regarding the elasticity of crime with respect to wage rates suggests that a significant part of the measured effect of education on crime can be attributed to the increase in wages associated with schooling.

We further argue that the impact of high school graduation on crime suggests that there are benefits to education not taken into account by individuals themselves. The social return to schooling is, therefore, larger than the private return. The estimated social externalities from reduced crime are sizeable. A 1% increase in the high school completion rate of all men ages 20-60 would save the United States as much as $1.4 billion per year in reduced costs from crime incurred by victims and society at large. The externality from education amounts to $1,170-2,100 per additional high school graduate or 14-26% of the private return to schooling. It is difficult to imagine a better reason to develop policies that prevent high school drop out.

References

Acemoglu, D. and Angrist, J. (2000), How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws. Working Paper.

Angrist, J. and Imbens, G. (1995), 'Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity', *JASA* **90**(430), 431–442.

Becker, G. and Mulligan, C. (1997), 'The Endogenous Determination of Time Preference', *Quarterly Journal of Economics* **112**(3), 729–758.

Card, D. (1995), 'Earnings, Schooling, and Ability Revisited', *Research in Labor Economics* **14**, 23–48.

Chilton, R. and Weber, D. (1998), *Uniform Crime Reporting Program (United States): Arrests by Age, Sex, and Race for Police Agencies in Metropolitan Statistical Areas, 1960-1995 [Computer file]*, ICPSR version.

Chiricos, T. (1987), 'Rates of Crime and Unemployment: An Analysis of Aggregate Research', *Social Problems* **34**(2), 187–211.

Ehrlich, I. (1975), On the Relation Between Education and Crime, *in* F. T. Juster, ed., 'Education, Income, and Human Behavior', McGraw-Hill Book Co., New York, chapter 12.

Farrington, D. et al. (1986), 'Unemployment, School Leaving and Crime', *British Journal of Criminology* **26**, 335–56.

Federal Bureau of Investigation, United States Department of Justice (1990), *Uniform Crime Reports for the United States*, USGPO, Washington, DC.

Freeman, R. (1983), Crime and Unemployment, *in* J. Q. Wilson, ed., 'Crime and Public Policy', ICS Press, San Francisco, chapter 6.

Freeman, R. (1995), The Labor Market, *in* J. Q. Wilson and J. Petersilia, eds, 'Crime', ICS Press, San Francisco, chapter 8.

Freeman, R. (1996), 'Why Do So Many Young American Men Commit Crimes and What Might We Do About It?', *Journal of Economic Perspectives* **10**(1), 25–42.

Gaviria, A. and Raphael, S. (2001), 'School-Based Peer Effects and Juvenile Behavior', *Review of Economics and Statistics* .

Glaeser, E., Sacerdote, B. and Scheinkman, J. (1996), 'Crime and Social Interactions', *Quarterly Journal of Economics* **111**(2), 507–48.

Gottfredson, D. (1985), 'Youth Employment, Crime, and Schooling', *Developmental Psychology* **21**, 419–32.

Gould, E., Mustard, D. and Weinberg, B. (2000), Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997. Working Paper.

Grogger, J. (1998), 'Market Wages and Youth Crime', *Journal of Labor Economics* **16**(4), 756–91.

Heckman, J. and Klenow, P. (1999), Human Capital Policy. Working Paper.

Hindelang, M., Hirsch, T. and Weis, J. (1981), *Measuring Delinquency*, Sage, Beverly Hills, CA.

Imbens, G. and Angrist, J. (1994), 'Identification and Estimation of Local Average Treatment Effects', *Econometrica* **62**(2), 467–75.

Kotin, L. and Aikman, W. (1980), *Legal Foundations of Compulsory School Attendance*, National University Publications: Port Washington.

Lang and Kropp (1986), 'Human Capital versus Sorting: Evidence from compulsory schooling laws', *Quarterly Journal of Economics* **101**, 609–624.

Levitt, S. (1997), 'Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime', *American Economic Review* **87**(3), 270–90.

Levitt, S. (1998), 'Juvenile Crime and Punishment', *Journal of Political Economy* **106**(6), 1156–85.

Lleras-Muney, A. (2000), Were Compulsory Attenance and Child Labor Laws Effective? An Analysis from 1915 to 1939. Working Paper.

Lochner, L. (1999), Education, Work, and Crime: Theory and Evidence. Rochester Center for Economic Research Working Paper No. 465.

Lynch, J. et al. (1994), *Profile of Inmates in the United States and in England and Wales, 1991*, U.S. Department of Justice, Washington, DC.

Machin, S. and Meghir, C. (2000), Crime and Economic Incentives. Institute for Fiscal Studies, Working Paper.

Miller, T., Cohen, M. and Wiersema, B. (1996), Victim Costs and Consequences: A New Look. Final Summary Report to the National Institute of Justice.

Moretti, E. (1999), Estimating the Social Return to Education: Evidence from Longitudinal and Cross-Sectional Data. Center for Labor Economics, University of California, Berkeley, Working Paper No. 22.

Mustard, D. (2001), 'Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the US Federal Courts', *Journal of Law and Economics* **44**(1).

Raphael, S. and Winter-Ebmer, R. (2001), 'Identifying the Effect of Unemployment on Crime', *Journal of Law and Economics* **44**(1).

Rausch, J. (1993), 'Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities', *Journal of Urban Economics* **34**, 380–400.

Stephan, J. (1999), *State Prison Expenditures, 1996*, U.S. Department of Justice, Washington, DC.

Tauchen, H., Witte, A. D. and Griesinger, H. (1994), 'Criminal Deterrence: Revisiting the Issue with a Birth Cohort', *Review of Economics and Statistics* **76**(3), 399–412.

U.S. Dept. of Justice (2000), *National Crime Surveys: National Sample of Rape Victims, 1973-1982 [Computer file]*, ICPSR version, 3rd ed.

Viscusi, K. (1986), Market Incentives for Criminal Behavior, *in* R. Freeman and H. Holzer, eds, 'The Black Youth Employment Crisis', University of Chicago Press, Chicago, chapter 8.

Witte, A. D. (1997), Crime, *in* J. Behrman and N. Stacey, eds, 'The Social Benefits of Education', University of Michigan Press, Ann Arbor, chapter 7.

Witte, A. D. and Tauchen, H. (1994), Work and Crime: An Exploration Using Panel Data. NBER Working Paper 4794.

Figure 1: Probability of Incarceration, by Years of Schooling - All Men 1980

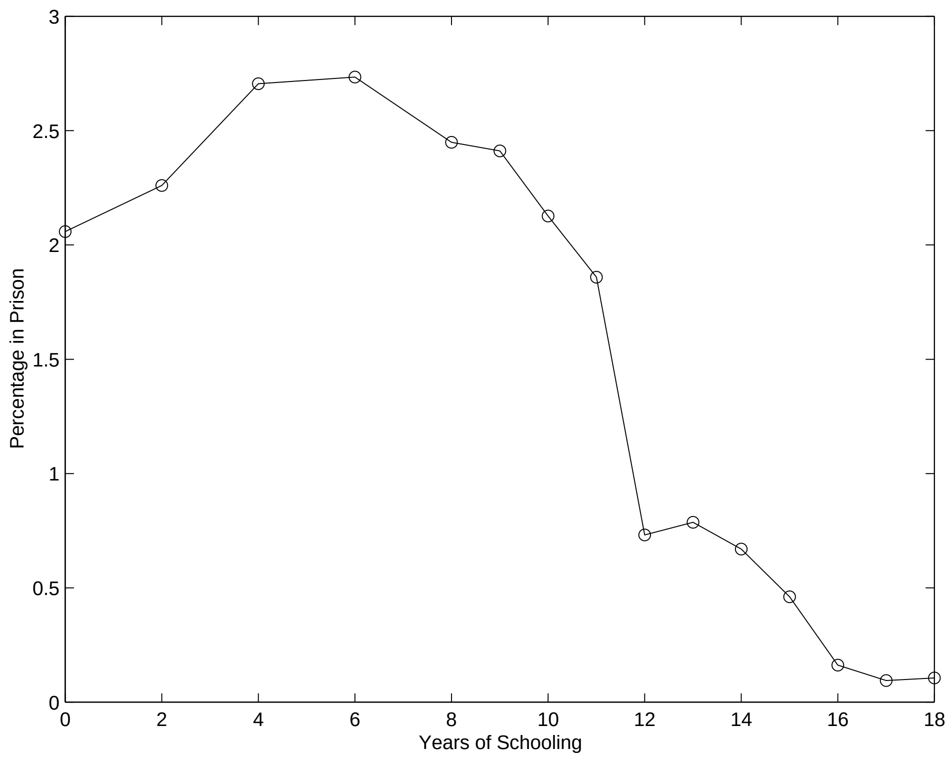


Figure 2: Changes in Compulsory Attendance Laws by State 1914-1978

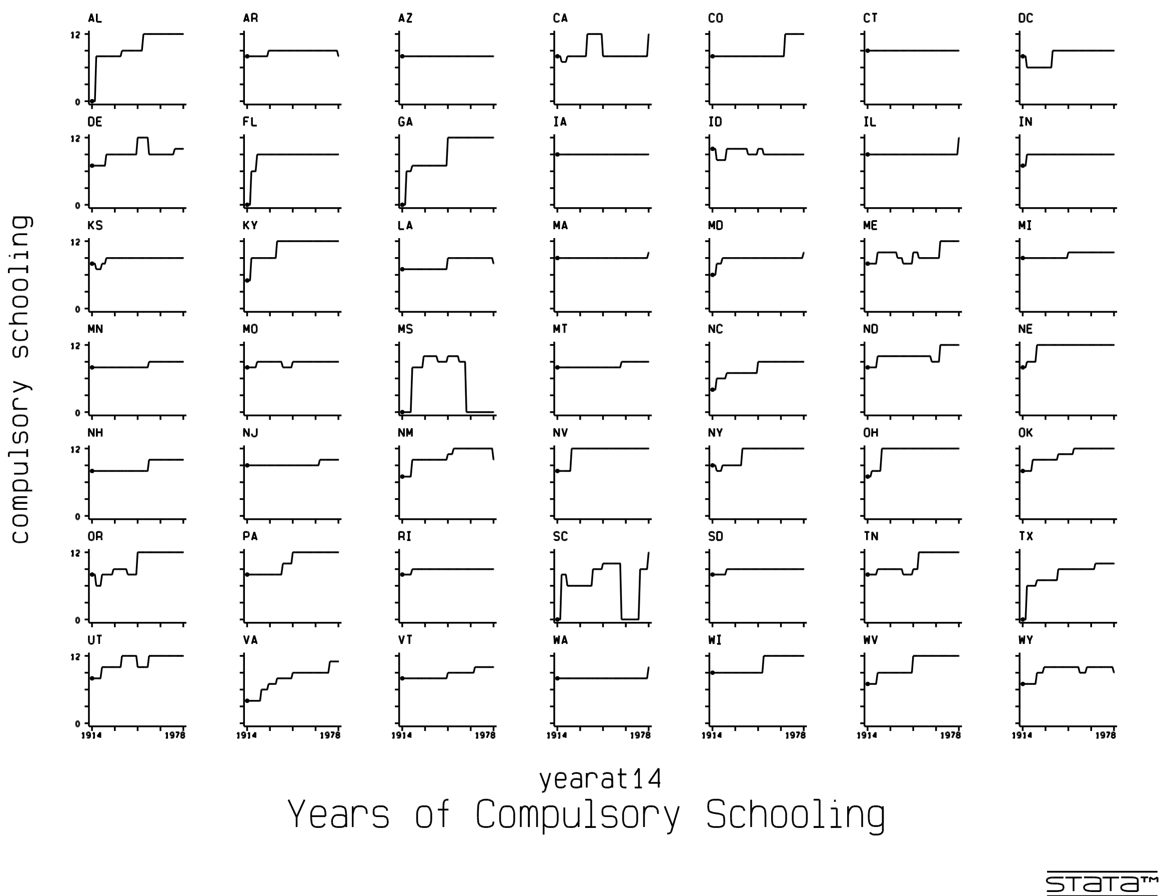
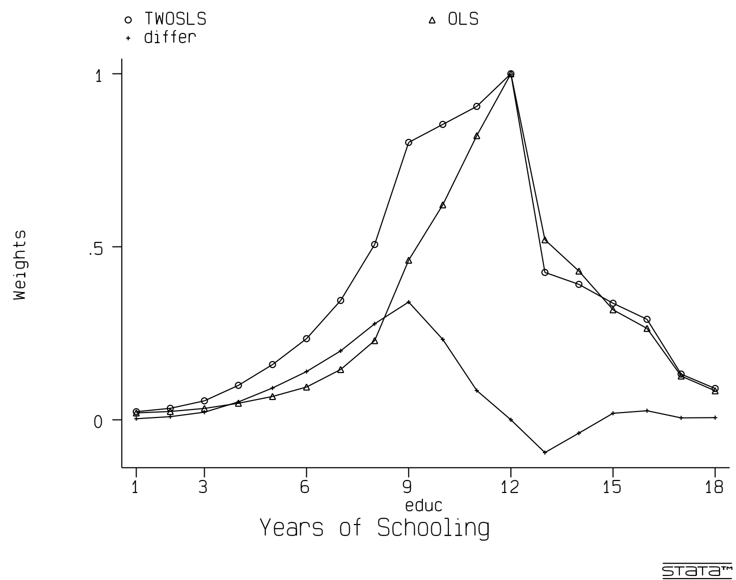
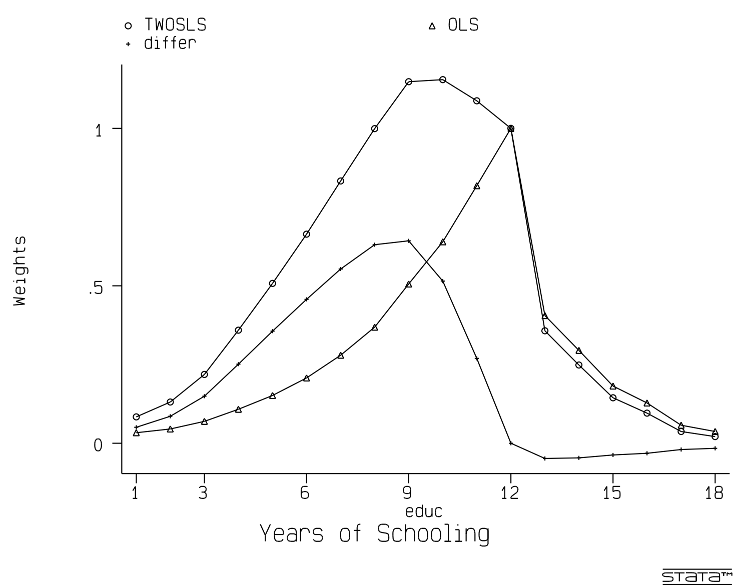


Figure 3: Comparing OLS and 2SLS Estimates: 2SLS Weights and OLS Weights for Whites (top) and Blacks (bootom)





The line with circles plots 2SLS weights for each year of schooling (λ in the text); The line with triangles plots OLS weights (θ and ϕ in the text). The line with dots plots the difference. The top panel is for whites; the bottom panel is for blacks.

Table 1: Census Descriptive Statistics: Mean (Standard Deviation) by Year

Variable	1960	1970	1980
Prison	0.0067	0.0051	0.0068
	(0.0815)	(0.0711)	(0.0820)
Drop out	0.52	0.37	0.23
	(0.50)	(0.48)	(0.42)
Age	38.79	38.54	37.00
	(11.21)	(11.95)	(11.94)
Compulsory Attendance ≤ 8	0.32	0.20	0.14
	(0.46)	(0.40)	(0.35)
Compulsory Attendance $= 9$	0.43	0.45	0.40
	(0.49)	(0.49)	(0.49)
Compulsory Attendance $= 10$	0.06	0.07	0.09
	(0.24)	(0.26)	(0.29)
Compulsory Attendance ≥ 11	0.17	0.26	0.34
	(0.37)	(0.44)	(0.47)
Black	0.096	0.090	0.106
	(0.295)	(0.287)	(0.307)
Sample Size	392,103	880,404	2,694,731

Table 2: Census Incarceration Rates for Men by Age and Race

Age	White	Black
20-29	0.0070	0.0446
30-39	0.0040	0.0298
40-49	0.0023	0.0126
50-60	0.0012	0.0060

Table 3: Census Incarceration Rates for Men Ages 20-60 by Drop Out Status

	All Years	1960	1970	1980
All Men				
Drop Out	.012	.010	.010	.015
HS Graduate +	.003	.002	.002	.004
Difference	.009	.008	.008	.011
White Men				
Drop Out	.008	.007	.006	.009
HS Graduate +	.002	.001	.001	.002
Difference	.006	.006	.005	.007
Black Men				
Drop Out	.036	.029	.029	.041
HS Graduate +	.019	.013	.012	.020
Difference	.017	.016	.017	.021

Table 4: OLS Estimates of the Effect of Drop Out on Imprisonment

	(1)	(2)	(3)	(4)	(5)	(6)
WHITES						
drop out	0.0062	0.0075	0.0076	0.0077	0.0076	0.0077
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
R-squared	0.001	0.004	0.004	0.004	0.004	0.004
BLACKS						
drop out	0.0193	0.0337	0.0339	0.0339	0.0338	0.0339
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
R-squared	0.0035	0.0203	0.0215	0.0218	0.0221	0.0219
Year Effects	y	y	y	y	y	
Age Effects		y	y	y	y	y
State of Birth Effects		y	y	y	y	y
State of Residence Effects			y	y	y	
Cohort of Birth Effects				y		y
Trend in Year of Birth × State of Birth					y	
State of Residence × Year Effects						y

Notes: Standard errors corrected for State of Birth - Year of Birth clustering are in parentheses. The dependent variable is a dummy equal to 1 if the respondent is in prison. Sample in the top panel includes 90% random sample of all white males ages 20-60 in 1960, 1970, and 1980 Censuses; N = 3,209,138. Sample in the bottom panel includes all black males ages 20-60 in 1960, 1970, and 1980 Censuses. N = 410,529. Age effects are 14 dummies (20-22, 23-25, etc.). State of Birth effects are 49 dummies for state of birth (Alaska and Hawaii are excluded). Year effects are 3 dummies for 1960, 1970, and 1980. State of Residence effects are 51 dummies for state of residence. Cohort of Birth Effects are dummies for decade of birth (1914-23, 1924-33, etc.). Models for blacks in columns 2 to 6 also include an additional state of birth dummy for cohorts born in the South turning age 14 in 1958 or later to account for the impact of Brown v. Board of Education.

Table 5: The Effect of Compulsory Attendance Laws on Schooling Achievement

	(1) drop out	(2) high school	(3) some college	(4) college+
WHITES				
Compulsory Attendance = 9	-.0325	.0327	-.0004	.0003
	(.0034)	(.0037)	(.0017)	(.0020)
Compulsory Attendance = 10	-.0331	.0401	-.0030	-.0039
	(.0045)	(.0051)	(.0030)	(.0033)
Compulsory Attendance $\geq$ 11	-.0551	.0582	-.0068	.0036
	(.0047)	(.0052)	(.0026)	(.0032)
F-test (p-value)	0.0000	0.0000	0.027	0.171
R-squared	0.12	0.02	0.04	0.05
BLACKS				
Compulsory Attendance = 9	-.0236	.0309	-.0069	-.0003
	(.0046)	(.0041)	(.0023)	(.0016)
Compulsory Attendance = 10	-.0176	.0406	-.0182	-.0047
	(.0065)	(.0064)	(.0039)	(.0023)
Compulsory Attendance $\geq$ 11	-.0296	.0502	-.0189	.0016
	(.0069)	(.0062)	(.0034)	(.0025)
F-test (p-value)	0.0000	0.0000	0.0000	0.136
R-squared	0.19	0.07	0.06	0.02
Year Effects	y	y	y	y
Age Effects	y	y	y	y
State of Birth Effects	y	y	y	y
State of Residence Effects	y	y	y	y
Cohort of Birth Effects	y	y	y	y

Notes: Standard errors corrected for State of Birth - Year of Birth clustering are in parentheses. The dependent variable in column 1 is a dummy equal to 1 if the respondent is a high school drop out. The dependent variables in columns 2-4 are dummies for high school, some college, and college, respectively. Sample in the top panel includes 90% random sample of all white males ages 20-60 in 1960, 1970, and 1980 Censuses; N = 3,209,138. Sample in the bottom panel includes all black males ages 20-60 in 1960, 1970, and 1980 Censuses; N = 410,529. Age effects are 14 dummies (20-22, 23-25, etc.). State of Birth effects are 49 dummies for state of birth (Alaska and Hawaii are excluded). Year effects are 3 dummies for 1960, 1970 and 1980. State of Residence effects are 51 dummies for state of residence. Cohort of Birth Effects are dummies for decade of birth (1914-23, 1924-33, etc.). Models for blacks also include an additional state of birth dummy for cohorts born in the South turning age 14 in 1958 or later to account for the impact of Brown v. Board of Education.

Table 6: Are changes in compulsory attendance laws correlated with the number of policemen or state police expenditures?

	Number of Policemen (1)	Police Expenditures (2)	Per Capita Police Expend. (3)
Compulsory Attendance =9	.0024	.103	-.002
	(.0080)	(.186)	(.002)
Compulsory Attendance =10	-.0031	-.430	-.015
	(0.0104)	(.209)	(.003)
Compulsory Attendance =11	-.0080	.-340	-.011
	(.0102)	(.180)	(.003)
R square	0.81	0.89	0.85
N	343	1500	1500
Year Effects	y	y	y
State Effects	y	y	y

Notes: Standard errors in parentheses. The dependent variable in column 1 is the percentage policemen in the state. Sample in column 1 includes observations from 49 states in years 1920, 1930, 1940, 1950, 1960, 1970, and 1980. The number of policemen in 1920-40 are taken from Census reports on occupations and the labor force for the entire U.S. population. Data from 1950-80 are from the IPUMS 1% Census samples. The dependent variable in column 2 is state police expenditures/$100 billions in constant dollars; Sample in column 2 includes observations from 49 states in all years from 1946 to 1978. The dependent variable in column 3 is state per capita police expenditures in constant dollars; Sample in column 3 includes observations from 49 states in years all years from 1946 to 1978. Data on police expenditures are from ICPSR 8706. See text for details.

Table 7: The Effect of Future Compulsory Attendance Laws on Current Drop Out Status

	WHITES			BLACKS		
	Compuls. Att. = 9	Compuls. Att. = 10	Compuls. Att. ≥11	Compuls. Att. = 9	Compuls. Att. = 10	Compuls. Att.≥ 11
	(1)	(2)	(3)	(4)	(5)	(6)
t = +4	.0032	-.0025	.0141	-.0054	.0153	.0164
	(.0122)	(.0182)	(.0214)	(.0067)	(.0110)	(.0144)
t = +5	-.0004	-.0085	.0007	.0004	.0098	.0068
	(.0078)	(.0113)	(.0141)	(.0046)	(.0081)	(.0101)
t = +6	-.0006	-.0100	-.0027	.0043	.0132	.0160
	(.0069)	(.0093)	(.0121)	(.0045)	(.0073)	(.0095)
t = +7	-.0001	-.0107	-.0027	.0072	.0136	.0024
	(.0057)	(.0078)	(.0121)	(.0043)	(.0079)	(.0090)
t = +8	-.0013	-.0106	-.0091	.0099	.0106	.0047
	(.0054)	(.0071)	(.0086)	(.0042)	(.0079)	(.0083)
t = +9	-.0016	-.0092	.0094	.0126	.0104	.0060
	(.0051)	(.0067)	(.0080)	(.0041)	(.0079)	(.0070)
t = +10	-.0011	-.0095	-.0123	.0140	.0084	.0041
	(.0046)	(.0063)	(.0071)	(.0045)	(.0078)	(.0075)
t = +11	.0013	-.0063	-.0131	.0156	.0071	.0020
	(.0043)	(.0055)	(.0069)	(.0049)	(.0075)	(.0078)
t = +12	.0061	-.0016	-.0080	.0158	.0017	.0042
	(.0047)	(.0054)	(.0072)	(.0050)	(.0070)	(.0075)
t = +15	.0092	.0018	-.00078	.0097	-.0122	.0044
	(.0046)	(.0054)	(.0066)	(.0052)	(.0063)	(.0079)
t = +18	.0067	-.0019	-.0131	.0020	-.0271	.0061
	(.0046)	(.0055)	(.0056)	(.0055)	(.0061)	(.0085)
t = +20	.0065	-.0040	-.0076	-.0013	-.0349	-.0040
	(.0050)	(.0060)	(.0059)	(.0064)	(.0071)	(.0083)
Year Effects	y	y	y	y	y	y
Age Effects	y	y	y	y	y	y
State of Birth Effects	y	y	y	y	y	y
State of Resid. Effects	y	y	y	y	y	y
Cohort of Birth Effects	y	y	y	y	y	y

Notes: Standard errors corrected for State of Birth - Year of Birth clustering are in parentheses. The dependent variable is a dummy equal to 1 if the respondent is a high school drop out. In column 1, 2 and 3, sample includes a 90% random sample of all white males ages 20-60 in 1960, 1970, and 1980 Censuses. In column 4, 5, and 6 sample includes all black males ages 20-60 in 1960, 1970, and 1980 Censuses. Age effects are 14 dummies (20-22, 23-25, etc.). State of Birth effects are 49 dummies for state of birth (Alaska and Hawaii are excluded). Year effects are 3 dummies for 1960, 1970 and 1980. State of Residence effects are 51 dummies for state of residence. Cohort of Birth Effects are dummies for decade of birth (1914-23, 1924-33, etc.). Columns 4, 5, and 6 also include an additional state of birth dummy for cohorts born in the South turning age 14 in 1958 or later to account for the impact of Brown v. Board of Education. N = 3,209,138 for whites; N = 410,529 for blacks.

Table 8: Instrumental Variable Estimates of the Effect of Drop Out on Imprisonment

	(1)	(2)	(3)	(4)	(5)
WHITES					
Second-Stage					
drop out	0.0076	0.0077	0.0061	0.0212	0.0089
	(0.0028)	(0.0029)	(0.0035)	(0.0079)	(0.0037)
Hausman Test: IV=OLS	-0.00	0.00	0.01	1.44	0.03
[$Prob > \chi^2$]	-	[0.99]	[0.99]	[0.99]	[0.99]
First Stage					
Compulsory Attendance = 9	-0.0393	-0.0398	-0.0325	-0.0160	-0.0327
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)
Compulsory Attendance = 10	-0.0362	-0.0358	-0.0331	-0.0076	-0.0313
	(0.0046)	(0.0046)	(0.0046)	(0.0051)	(0.0045)
Compulsory Attendance $\geq$ 11	-0.0676	-0.0661	-0.0551	-0.0297	-0.0539
	(0.0048)	(0.0048)	(0.0048)	(0.0050)	(0.0047)
First Stage R-squared	0.12	0.12	0.12	0.12	0.13
F-test for Instruments	64.65	63.48	47.91	15.61	48.05
BLACKS					
Second-Stage					
drop out	0.0929	0.0967	0.0723	0.0613	0.0800
	(0.0235)	(0.0260)	(0.0366)	(0.0396)	(0.0378)
Hausman Test: IV=OLS	49.81	35.62	6.14	0.07	1.88
[$Prob > \chi^2$]	[0.99]	[0.99]	[0.99]	[0.99]	[0.99]
First Stage					
Compulsory Attendance = 9	-0.0316	-0.0307	-0.0236	-0.0131	-0.0233
	(0.0051)	(0.0051)	(0.0046)	(0.0059)	(0.0046)
Compulsory Attendance = 10	-0.0225	-0.0221	-0.0176	0.0036	-0.0163
	(0.0076)	(0.0076)	(0.0065)	(0.0090)	(0.0064)
Compulsory Attendance $\geq$ 11	-0.0484	-0.0481	-0.0296	-0.0373	-0.0288
	(0.0070)	(0.0068)	(0.0069)	(0.0081)	(0.0068)
First Stage R-squared	0.17	0.18	0.19	0.19	0.19
F-test for Instruments	20.54	20.50	10.09	8.70	10.01
Year Effects	y	y	y	y	
Age Effects	y	y	y	y	y
State of Birth Effects	y	y	y	y	y
State of Residence Effects		y	y	y	
Cohort of Birth Effects			y		y
Trend in Year of Birth $\times$ State of Birth				y	
State of Residence $\times$ Year Effects					y

Notes: Standard errors corrected for State of Birth - Year of Birth clustering are in parentheses. See notes in Table 4

Table 9: OLS Estimates of the Effect of Drop Out Rates on Arrest Rates

	(1)	(2)	(3)	(4)
ALL CRIMES				
drop out rate	0.398	0.618	0.674	0.710
	(0.170)	(0.183)	(0.181)	(0.283)
percent. black	2.083	3.007	2.948	2.684
	(0.539)	(0.491)	(0.503)	(0.403)
R-squared	0.89	0.93	0.95	0.96
BY TYPE OF CRIME				
drop out × violent crime	0.597	0.966	0.751	0.793
	(0.172)	(0.194)	(0.198)	(0.291)
drop out × property crime	0.195	0.244	0.593	0.621
	(0.175)	(0.197)	(0.208)	(0.304)
percent. black	2.086	2.998	2.947	2.684
	(0.540)	(0.491)	(0.504)	(0.403)
R-squared	0.89	0.93	0.96	0.96
age effects	y			
year effects	y			
state effects	y			
offense effects	y			
age × offense effects		y	y	y
offense × year effects		y	y	y
age × year effects		y	y	y
state × age effects		y	y	y
state × offense effects			y	y
state × year				y

Notes: Standard errors corrected for state-year-age clustering are in parentheses. The dependent variable is the logarithm of the arrest rate by age, type of offense, state, and year. Drop out rate and percentage black are by age group, state, and year (see text). Violent crimes include murder, rape, robbery, and assault. Property crimes include burglary, larceny, vehicle theft, and arson. There are 8 age groups, 8 offenses, 50 states, and 4 years. All models are weighted by cell size.

Table 10: OLS Estimates for Arrest Rates by Detailed Type of Crime

	(1)	(2)
drop out rate × murder	2.062	2.133
	(0.403)	(0.403)
drop out rate × rape	-1.094	-1.049
	(0.307)	(0.353)
drop out rate × robbery	-0.184	-0.113
	(0.253)	(0.333)
drop out rate × assault	2.136	2.179
	(0.226)	(0.326)
drop out rate × burglary	0.202	0.250
	(0.268)	(0.347)
drop out rate × larceny	0.235	0.277
	(0.209)	(0.311)
drop out rate × vehicle	1.227	1.271
	(0.251)	(0.346)
drop out rate × arson	0.745	0.784
	(0.358)	(0.408)
percent black	2.908	2.620
	(0.333)	(0.404)
R-squared	0.96	0.96
age × offense effects	y	y
offense × year effects	y	y
age × year effects	y	y
state × age effects	y	y
state × offense effects	y	y
state × year		y

Notes: Standard errors corrected for state-year-age clustering are in parentheses. The dependent variable is the logarithm of the arrest rate by age, type of offense, state, and year. Drop out rate and percentage black are by age group, state, and year (see text). There are 8 age groups, 8 offenses, 50 states, and 4 years. All models are weighted by cell size.

Table 11: IV Estimates of the Effect of Drop Out Rates on Arrest Rates: All crimes

	(1)	(2)	(3)	(4)
Second Stage				
drop out rate	0.944	0.946	0.941	0.873
	(0.397)	(0.491)	(0.522)	(0.669)
percent. black	1.646	3.271	3.121	1.172
	(0.589)	(0.553)	(0.561)	(0.628)
First Stage				
Compulsory Attendance =9	-0.050	-0.035	-0.036	-0.030
	(0.008)	(0.004)	(0.004)	(0.005)
Compulsory Attendance=10	-0.035	-0.038	-0.038	-0.028
	(0.012)	(0.006)	(0.006)	(0.006)
Compulsory Attendance$\geq$ 11	-0.079	-0.053	-0.054	-0.049
	(0.011)	(0.006)	(0.006)	(0.008)
First stage R-squared	0.89	0.97	0.97	0.97
age effects	y			
year effects	y			
state effects	y			
offense effects	y			
age $\times$ offense effects		y	y	y
offense $\times$ year effects		y	y	y
age $\times$ year effects		y	y	y
state $\times$ age effects		y	y	
state $\times$ offense effects			y	y
state $\times$ year				y

Notes: Standard errors corrected for state-year-age clustering are in parentheses. The dependent variable is the logarithm of the arrest rate by age, type of offense, state, and year. Drop out rate and percentage black are by age group, state, and year (see text). There are 8 age groups, 8 offenses, 50 states, and 4 years. All models are weighted by cell size.

Table 12: Effect of High School Graduation on Self-Reported Crime and Incarceration Rates for Whites (NLSY)

	Mean for drop-out (1)	Mean for graduates (2)	Raw Diff. (1)-(2) (3)	Probit Spec. 1 (4)	Probit Spec. 2 (5)
Self-Reported Crime					
Violent Crime	0.15	0.06	-0.09	-0.09	-0.08
	(0.01)	(0.01)	(0.01)	(0.02)	*
Drug Sales	0.06	0.02	-0.04	-0.04	-0.05
	(0.01)	(0.00)	(0.01)	(0.00)	(0.03)
Property Crime	0.23	0.15	-0.08	-0.07	-0.10
	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
Any Crime	0.31	0.19	-0.11	-0.11	-0.15
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Incarcerated	0.04	0.01	-0.03	-0.03	-0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Controls:					
Age/Cohort				y	y
Region of Residence				y	y
Family Background					y
Ability					y
SMSA Status					y
Local Unemployment Rate					y
State Incarceration Rate					y

Notes: Self-reported crimes are based on men ages 18-23 in 1980. Violent crimes correspond to robbery and assault, while property crimes include shoplifting and all other thefts of over $50. Estimates in columns 4 and 5 are the average differences in predicted criminal participation or incarceration associated with a change in high school graduation status (derived from probit estimates). Each cell in column 4 and 5 represents the result of a separate probit regression. The dependent variable in the final row is a dummy equal 1 if the individual was incarcerated over ages 22-28. The coefficients reported are obtained by adjusting the ages 22-28 incarceration rates by the ratio of the probability of incarceration for the seven year span to the annual incarceration probability (over those same ages). Family background measures include: current enrollment in school, parents highest grade completed, whether or not the individual lived with both of his natural parents at age 14, whether his mother was a teenager at his birth, and family income. Standard errors were calculated using the delta method (* denotes a covariance matrix for the coefficient estimates that was not positive definite).

Table 13: Social Costs per Crime and Social Benefits of Increasing High School Completion Rates by 1%

	Victim Costs per crime (1)	Property Loss per crime (2)	Incarc. Cost per crime (3)	Total Cost per crime (4)	Est. Change in Arrests (5)	Est. Change in Crimes (6)	Social Benefit $(4)\times(6)$ (7)
Violent Crimes							
Murder	2,940,000	120	845,455	3,024,359	-373	-373	$1,129,596,562
Rape	87,000	100	2,301	89,221	347	1,559	-$139,109,278
Robbery	8,000	750	1,985	9,385	134	918	-$8,617,191
Assault	9,400	26	538	9,917	-7,798	-37,135	$368,252,227
Property Crimes							
Burglary	1,400	970	363	987	-653	-9,467	$9,342,643
Larceny/Theft	370	270	44	198	-1,983	-35,105	$6,944,932
Motor Vehicle Theft	3,700	3,300	185	1,245	-1,355	-14,238	$17,728,056
Arson	37,500	15,500	1,542	39,042	-69	-469	$18,323,748
Total					11,750	94,310	$1,402,461,698

Notes: Victim costs and property losses taken from Table 2 of Miller, Cohen, and Wiersema (1996). Incarceration costs per crime equal the incarceration cost per inmate, $17,027 (Stephan 1999), multiplied by the incarceration rate (Lynch et al. 1994). Total costs are calculated as the sum of victim costs and incarceration costs less 80% of the property loss (already included in victim costs) for all crimes except arson. Total costs for arson are the sum of victim costs and incarceration costs. See text for details. Estimated change in arrests calculated from column 2 of Table 10 and the total number of arrests in 1990 Uniform Crime Reports (1990). Estimated changes in crimes adjusts the arrest effect by the number of crimes per arrest. The social benefit is the estimated change in crimes in column 6 times the total cost per crime in column 4. All dollar figures are in 1993 dollars. See text for details.

Appendix A Derivation of λ-Weights for 2SLS Estimation

This appendix derives the λ-weights for the 2SLS estimator as described in equation (3) in the text.

Let $\alpha_i = Pr(Z=i)[E(D|Z=i) - E(D)]$. Then,

$$\frac{E\{y[E(D|Z) - E(D)]\}}{E\{E(D|Z)[E(D|Z) - E(D)]\}} = \frac{\sum_{i=0}^{I} E(y|Z=i)\alpha_i}{\sum_{i=0}^{I} Pr(s \geq k|Z=i)\alpha_i} \tag{7}$$

$$= \beta_k + \sum_{j \neq k} \beta_j \left(\frac{\sum_{i=0}^{I} \alpha_i Pr(s \geq j|Z=i)}{\sum_{i=0}^{I} \alpha_i Pr(s \geq k|Z=i)} \right), \tag{8}$$

where the first inequality follows from the definition of α_i and the fact that $E(D|Z=i) = Pr(s \geq k|Z=i)$. The second inequality follows from $E(y|Z=i) = \sum_{j=0}^{S} Pr(s \geq j|Z=i)\beta_j$.

To finish the proof, it is necessary to verify that the expression in parentheses in equation (8) is equivalent to λ_j. Re-write $Pr(s \geq j|Z=i) = Pr(s_i \geq j)$ and consider the term

$$\begin{aligned}\sum_{i=0}^{I} \alpha_i Pr(s_i \geq j) &= \sum_{i=0}^{I} \alpha_i Pr(s_0 \geq j) + \sum_{i=1}^{I} \alpha_i [Pr(s_i \geq j) - Pr(s_0 \geq j)] \\ &= \sum_{i=1}^{I} \alpha_i Pr(s_i \geq j > s_0),\end{aligned}$$

where the first equality follows from adding and subtracting the term $\sum_{i=1}^{I} \alpha_i Pr(s_0 \geq j)$. The second inequality follows from the fact that $\sum_{i=1}^{I} \alpha_i = 0$ since $E[E(D|Z)] = E(D)$. Substituting these terms in for the expression in equation (8), it is clear that the expression equals λ_j.

Appendix B IV and OLS in a Model with Spillovers

We show that OLS and IV estimates will differ in a simple econometric model of crime when spillovers are introduced. Consider the following (constant coefficient) model of criminal behavior:

$$c_{is} = \alpha + \beta d_{is} + \gamma d_s + \varepsilon_{is},$$

where c_{is} represents crime committed by individual i in state s, d_{is} represents drop out status for individual i living in state s, d_s represents average drop out rates in state s, and ε_{is} is a mean zero random error term assumed to be independent of D_{is} and D_s. To simplify the analysis, assume we have a balanced panel with n individuals in each of S states. We also define $d = \frac{1}{nS}\sum_s \sum_i d_{is} = \frac{1}{S}\sum_s d_s$, which represents average drop out rate in the entire economy.

OLS estimates of a regression of c_{is} on d_{is} and a constant term produce the following estimate of β:

$$\hat{\beta}_{ols} = \beta + \gamma \left(\frac{\sum_s \sum_i (d_{is} - d)(d_s - d)}{\sum_s \sum_i (d_{is} - d)^2} \right) + \frac{\sum_s \sum_i \varepsilon_{is}^2}{\sum_s \sum_i (d_{is} - d)^2}$$

$$
\begin{aligned}
&= \beta+\gamma\left(\frac{\sum_s (d_s-d)(d_s-d)}{\sum_s\sum_i (d_{is}-d)^2}\right)+\frac{\sum_s\sum_i \varepsilon_{is}^2}{\sum_s\sum_i (d_{is}-d)^2}\\
&\to \beta+\gamma\left(\frac{V(d_s)}{V(d_{is})}\right).
\end{aligned}
$$

IV estimates of the same estimating equation, using a valid state-level instrument z_s (with sample mean z) to instrument for drop out rates, produces the following estimate for β:

$$
\begin{aligned}
\hat{\beta}_{IV} &= \beta+\gamma\left(\frac{\sum_s\sum_i (z_s-z)(d_s-d)}{\sum_s\sum_i (z_s-z)(d_{is}-d)}\right)+\frac{\sum_s\sum_i (z_s-z)\varepsilon_{is}}{\sum_s\sum_i (z_s-z)(d_{is}-d)}\\
&= \beta+\gamma\left(\frac{\sum_s (z_s-z)(d_s-d)}{\sum_s (z_s-z)(d_s-d)}\right)+\frac{\sum_s\sum_i (z_s-z)\varepsilon_{is}}{\sum_s\sum_i (z_s-z)(d_{is}-d)}\\
&\to \beta+\gamma.
\end{aligned}
$$

For $\gamma > 0$, we will observe $\beta_{IV} \geq \beta_{ols}$ since $V(d_s) \leq V(d_{is})$. For small cross-state variation in drop out rates (i.e. $V(d_s) \approx 0$), OLS will estimate the own-effect of drop out on criminal participation while IV estimates using state-level instruments will estimate the combined own-effect and spillover effect of average drop out rates.

Defining $c_s = \frac{1}{n}\sum_i c_{is}$, estimates from a regression of average crime rates on average drop out rates in a state will produce estimates of the combined own and spillover effects, since $c_s = \alpha + (\beta+\gamma)d_s + \varepsilon_s$.

A similar analysis can be performed when average crime rates (rather than average drop out rates) affect individual criminal behavior. Now, suppose

$$
c_{is} = \alpha + \beta d_{is} + \gamma c_s + \varepsilon_{is}.
$$

Assuming $\frac{1}{n}\sum_i \varepsilon_{is} = 0$, taking the average of c_{is} yields

$$
c_s = \frac{\alpha}{1-\gamma} + \left(\frac{\beta}{1-\gamma}\right) d_s.
$$

As above, we can calculate OLS and IV estimates of β using a regression of c_{is} on d_{is} and a constant term:

$$
\beta_{ols} \to \beta + \gamma\left(\frac{V(c_s)}{V(d_{is})}\right) = \beta+\gamma\left(\frac{\beta}{1-\gamma}\right)^2\left(\frac{V(d_s)}{V(d_{is})}\right)
$$

and

$$
\beta_{IV} \to \beta + \frac{\gamma\beta}{1-\gamma} = \frac{\beta}{1-\gamma}.
$$

Provided $\gamma > 0$ and $\beta + \gamma < 1$, we will observe $\beta_{IV} \geq \beta_{ols}$. A regression of average crime rates on average drop out rates should produce an estimate that converges to the IV estimate.

Appendix C Wage Changes in General Equilibrium

Consider a CRS aggregate production technology, $F(N_d, N_g, K)$, that uses high school dropouts, N_d, high school graduates, N_g, and physical capital, K:

$$
F(N_d, N_g, K) = aK^{\beta}(bN_d^{\rho} + (1-b)N_g^{\rho})^{(1-\beta)/\rho}.
$$

In a competitive market, inputs are paid their marginal products, so

$$w_d = (1-\beta)abK^{\beta}Q^{\frac{1-\beta-\rho}{\rho}}N_d^{\rho-1} \tag{9}$$
$$w_g = (1-\beta)a(1-b)K^{\beta}Q^{\frac{1-\beta-\rho}{\rho}}N_g^{\rho-1}, \tag{10}$$

where $Q = bN_d^{\rho} + (1-b)N_g^{\rho}$. Changes in log wages are given by

$$dlog(w_d) = \beta\, dlog(K) + \left(\frac{1-\beta-\rho}{\rho}\right)\left(\frac{\partial log(Q)}{\partial N_d}\, dN_d + \frac{\partial log(Q)}{\partial N_g}\, dN_g\right) + (\rho-1)\, dlog(N_d)$$
$$dlog(w_g) = \beta\, dlog(K) + \left(\frac{1-\beta-\rho}{\rho}\right)\left(\frac{\partial log(Q)}{\partial N_d}\, dN_d + \frac{\partial log(Q)}{\partial N_g}\, dN_g\right) + (\rho-1)\, dlog(N_g)$$

To determine the wage responses of any increase in the graduation rate (holding physical capital, K, fixed), it is necessary to know initial dropout rates as well as the technology parameters b, β and ρ. The parameter ρ can be determined from the elasticity of substitution between high school dropouts and graduates, σ, according to $\rho = (\sigma-1)/\sigma$, while b can be determined from the high school graduate - dropout wage ratio, $R = w_g/w_d$, as well as the supply of graduates and dropouts. Taking the ratio of graduate to dropout wages using equations (9) and (10) yields $b = \left[1 + R(N_g/N_d)^{1-\rho}\right]^{-1}$. The parameter β is given by capital's share of output.

A high school graduation rate of $N_g = 0.85$ ($N_d = 0.15$), graduate - dropout wage ratio of $R = 1.4$, elasticity of substitution between graduates and dropouts of $\sigma = 2$, and capital share of $\beta = 0.33$ yields a 0.0350 increase in the log wages of high school drop outs and a 0.0054 decrease in the log wages of graduates in response to a 1% increase in the high school graduation rate (ignoring changes in physical capital).

If the elasticity of crime with respect to wage rates, η, is identical for all workers, graduates and drop outs alike, then the aggregate effect on log crime rates is given by

$$dlog(crime) = \eta(N_g\, dlog(w_g) + N_d\, dlog(w_d)).$$

Using an elasticity of crime with respect to wages of 2 (on the high end of the estimates from Gould, et al., 2000) yields a net *decrease* in log crime rates of 0.0014. An elasticity of 1.5 yields a decrease of 0.0024.