American Economic Association

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

Author(s): Lance Lochner and Enrico Moretti

Source: *The American Economic Review*, Vol. 94, No. 1 (Mar., 2004), pp. 155-189
Published by: American Economic Association
Stable URL: http://www.jstor.org/stable/3592774
Accessed: 11-08-2016 18:28 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://about.jstor.org/terms

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

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

By LANCE LOCHNER AND ENRICO MORETTI*

We estimate the effect of education on participation in criminal activity using changes in state compulsory schooling laws over time to account for the endogeneity of schooling decisions. Using Census and FBI data, we find that schooling significantly reduces the probability of incarceration and arrest. NLSY data indicate that our results are caused by changes in criminal behavior and not differences in the probability of arrest or incarceration conditional on crime. We estimate that the social savings from crime reduction associated with high school graduation (for men) is about 14–26 percent of the private return. (*JEL* I2, K42)

Is it possible to reduce crime rates by raising the education of potential criminals? If so, would it be 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.

The motivation for these questions is not limited to the obvious policy implications for crime prevention. Estimating the effect of education on criminal activity may shed some light on the magnitude of the *social return* to education. 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 productivity and earnings of other workers around him. [For example, see James Heckman and Peter Klenow (1999), Daron Acemoglu and Joshua Angrist (2000), and Moretti (2004a, b).] 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.[1]

Crime is a negative externality with enormous social costs. If education reduces crime, then schooling will have social benefits that are not taken into account by individuals. In this case, the *social return* to education may exceed the private return. 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 will affect subsequent crime. First, schooling increases the returns to legitimate work, raising the opportunity costs of illicit behavior.[2] Additionally, punishment for crime

* Lochner: Department of Economics, University of Western Ontario, 1151 Richmond Street, London, Ontario, N6A 5C2, Canada (e-mail: llochner@uwo.ca); Moretti: Department of Economics, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095 (e-mail: moretti@econ.ucla.edu). We are grateful to Daron Acemoglu and Josh Angrist for their data on compulsory attendance laws and useful suggestions. We thank Mark Bils, Elizabeth Caucutt, Janet Currie, Gordon Dahl, Stan Engerman, Jeff Grogger, Jinyong Hahn, Guido Imbens, Shakeeb Khan, David Levine, Jens Ludwig, Darren Lubotsky, Marco Manacorda, Marcelo Moreira, David Mustard, Peter Rupert, Steve Rivkin, Todd Stinebrickner, Edward Vytlacil, Tiemen Woutersen, two referees, and seminar participants at Columbia University, Chicago GSB, NBER Summer Institute, Econometric Society, University of Rochester, UCLA, University of British Columbia, Hoover Institution, and Stanford University for their helpful comments. All the data used in the paper are available at http://www.econ.ucla.edu/moretti/papers.html.

[1] Ann D. Witte (1997) and Lochner (2003) are notable exceptions.

[2] W. K. Viscusi (1986), Richard Freeman (1996), Jeffrey Grogger (1998), Stephen Machin and Costas Meghir (2000), and Eric D. Gould et al. (2002) 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 Freeman, 1983, 1995, or Theodore Chiricos, 1987, for excellent sur-

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

typically entails incarceration. By raising wage rates, schooling makes this "lost time" more costly. Second, education 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 or risk aversion. On net, we expect that most of these channels will lead to a negative relationship between education and typical violent and property crimes.

Despite the many reasons to expect a causal link between education and crime, empirical research is not conclusive.[3] The key difficulty in estimating the effect of education on criminal 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.

To address endogeneity problems, we use changes in state compulsory attendance laws over time to instrument for schooling. Changes in these laws have a significant effect on educational achievement, and we find little evidence that changes in these laws simply reflect preexisting trends toward higher schooling levels. In the years preceding increases in compulsory schooling laws, there is no obvious trend in schooling achievement. Increases in education associated with increased compulsory schooling take place *after* changes in the law. Furthermore, increases in the number of years of compulsory attendance raise high school graduation rates but have no effect on college graduation rates. These two facts indicate that the increases in compulsory schooling raise education, not vice versa. We also examine whether increases in compulsory schooling ages are associated with increases in state resources devoted to fighting crime. They are not.

We use individual-level data on incarceration from the Census and cohort-level data on arrests by state from the FBI 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 education on incarceration. The group quarters type of residence in the Census indicates whether an individual is incarcerated at the Census date. For both blacks and whites, ordinary least-squares (OLS) estimates uncover significant reductions in the probability of incarceration associated with more schooling. Instrumental

veys); however, a number of recent studies that better address problems with endogeneity and unobserved correlates (including Steven Raphael and Rudolf Winter-Ebmer, 2001, and Gould et al., 2002) find a sizeable positive effect of unemployment on crime.

[3] Witte (1997) concludes that "... neither years of schooling completed nor receipt of a high school degree has a significant effect on an individual's level of criminal activity." But, this conclusion is based on only a few available studies, including Helen 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 after controlling for wages. (Of course, increased wages are an important consequence of schooling.) More recently, Lochner (2003) 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 (Michael Gottfredson, 1985; David Farrington et al., 1986; and 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.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

variable estimates reveal a significant relationship between education and incarceration, and they suggest that the impacts are greater for blacks than for whites. One extra year of schooling results in a 0.10-percentage-point reduction in the probability of incarceration for whites, and a 0.37-percentage-point reduction for blacks. To help in interpreting the size of these impacts, we calculate how much of the black-white gap in incarceration rates in 1980 is due to differences in educational attainment. Differences in average education between blacks and whites can explain as much as 23 percent of the black-white gap in incarceration rates.

Because incarceration data do not distinguish between types of offenses, we also examine the impact of education on arrests using data from the UCR. This data allows us to identify the type of crime that arrested individuals have been charged with. Estimates uncover a robust and significant effect of high school graduation 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 graduation 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 schooling on self-reported criminal participation using data from the NLSY. These estimates confirm that education significantly reduces self-reported participation in both violent and property crime among whites. Results for blacks in the NLSY are less supportive, but there is good reason to believe that they are substantially biased due to severe underreporting of crime by high school dropouts. We also use the NLSY to explore the robustness of our findings on imprisonment to the inclusion of rich measures of family background and individual ability. The OLS estimates obtained in the NLSY controlling for the Armed Forces Qualifying Test (AFQT) scores, parental education, family composition, and several other background characteristics are remarkably similar to the estimates obtained using Census data for both blacks and whites.

Given the general consistency in findings across data sets, measures of criminal activity, and identification strategies, we cannot reject that a relationship between education and crime exists. Using our estimates, we calculate the social savings from crime reduction associated with high school completion. Our estimates suggest that a 1-percent increase in male high school graduation rates would save as much as $1.4 billion, or about $2,100 per additional male high school graduate. These social savings represent an important externality of education that has not yet been documented. The estimated externality from education ranges from 14–26 percent 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 I, we briefly discuss the channels through which education may affect subsequent crime, arrests, and incarceration. Section II reports estimates of the impact of schooling on incarceration rates (Census data), and Section III reports estimates of the impact of schooling on arrest rates (UCR data). Section IV uses NLSY data on self-reported crime and on incarceration to check the robustness of UCR and Census-based estimates. In Section V, we calculate the social savings from crime reduction associated with high school graduation. Section VI concludes.

I. The Relationship Between Education, Criminal Activity, Arrests, and Incarceration

Theory suggests several ways that educational attainment may affect subsequent criminal decisions. First, schooling increases individual wage rates, thereby increasing the opportunity costs of crime. Second, punishment is likely to be more costly for the more educated. Incarceration implies time out of the labor market, which is more costly for high earners. Furthermore, previous studies estimate that the stigma of a criminal conviction is larger for white collar workers than blue collar workers (see, e.g., Jeffrey Kling, 2002), which implies that the negative effect of a conviction on earnings extend beyond the time spent in prison for more educated workers.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

Third, schooling may alter individual rates of time preference or risk aversion. That is, schooling may increase the patience exhibited by individuals (as in Gary S. Becker and Casey B. Mulligan, 1997) or their risk aversion. More patient and more risk-averse individuals would place more weight on the possibility of future punishments. Fourth, schooling may also affect individual tastes for crime by directly affecting the psychic costs of breaking the law. (See, e.g., Kenneth Arrow, 1997.)

Fifth, it is possible that criminal behavior is characterized by strong state dependence, so that the probability of committing crime today depends on the amount of crime committed in the past. By keeping youth off the street and occupied during the day, school attendance may have long-lasting effects on criminal participation.[4]

These channels suggest that an increase in an individual's schooling attainment should cause a decrease in his subsequent probability of engaging in crime. But, it is also possible that schooling raises the direct marginal returns to crime. For example, certain white collar crimes are likely to require higher levels of education. Education may also lower the probability of detection and punishment or reduce sentence lengths handed out by judges. David B. Mustard (2001) finds little evidence of the latter.

In this paper, we do not attempt to empirically differentiate between the many channels through which education may affect criminal activity. Instead, we explore a simple reduced-form relationship between adult crime, c_i, and educational attainment, s_i, conditional on other individual characteristics, X_i:

$$c_i = \beta s_i + \gamma X_i + \varepsilon_i. \tag{1}$$

The coefficient β captures the net effect of education on criminal activity. As long as schooling increases the marginal return to work more than crime and schooling does not decrease patience levels or increase risk aversion, we should observe a negative relationship between crime and schooling: $\beta < 0$.

In estimating equation (1), two important difficulties arise. First, schooling is not exogenous. Considering their optimal lifetime work and crime decisions for each potential level of schooling, young individuals will choose the education level that maximizes lifetime earnings. As a result, the same factors that affect decisions to commit crime also affect schooling decisions. (See Lochner, 2003, for a more formal theoretical analysis.) For example, 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, 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.

We deal with the endogeneity of schooling by using variation in state compulsory schooling laws as an instrumental variable for education. The instrument is valid if it induces variation in schooling but is uncorrelated with discount rates and other individual characteristics that affect both imprisonment and schooling. We find no evidence that changes in these laws simply reflect preexisting trends toward higher schooling levels. There are no clear trends in schooling during years preceding changes in compulsory schooling ages. Furthermore, the empirical effects of these laws are focused on high school grades and are unrelated to college completion rates. Both of these findings indicate that the increases in compulsory schooling raise education and not that changes in the law are correlated with underlying changes in education within states. We also test whether increases in compulsory schooling ages are associated with increases in state resources devoted to fighting crime. We find little evidence to support this hypothesis.

A second problem that arises in the estimation of equation (1) is due to data limitations—namely, crime is not observed directly. In this paper, we

[4] Estimates by Brian Jacob and Lars Lefgren (2003) suggest that school attendance reduces contemporaneous juvenile property crime while increasing juvenile violent crime. Their results are consistent with an incapacitation effect of school that limits student capacities for engaging in property crime, but they also may suggest that the increased level of interaction among adolescents facilitated through schools may raise the likelihood of violent conflicts.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

primarily use information on incarceration (from the Census) and arrests (from the FBI Uniform Crime Reports). However, neither of these data sets measures crime directly. 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 at date t, c_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 $\Pr(Arrest_t) = \pi_a c_t(s)$ and incarcerated with probability $\Pr(Inc_t) = \pi_i \pi_a c_t(s)$.

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

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 graduation on arrest and incarceration rates (when measured in logarithms) are

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

and

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

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 graduation on arrest is greater than its effect on crime by $\ln \pi_a(1) - \ln \pi_a(0)$ and its measured effect on imprisonment is larger still by the additional amount $\ln \pi_i(1) - \ln \pi_i(0)$.

Estimates using arrest and imprisonment measures of crime may, therefore, 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 also estimate the effects of schooling on self-reported criminal participation using data from the NLSY. Unless education substantially alters either the probability of arrest, the probability of incarceration, or sentence lengths, we should expect similar percentage changes in crime associated with schooling whether we measure crime by self-reports, arrests, or incarceration rates.[5]

II. The Impact of Schooling on Incarceration Rates

A. *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 create a dummy variable equal to 1 if the respondent is in a correctional institution.[6] We include in our sample males ages 20–60 for whom all the relevant variables are reported. Summary statistics are provided in Table 1. Roughly 0.5–0.7 percent of the respondents are in prison during each of the Census years we examine. Average years of schooling

[5] Mustard (2001) provides evidence from U.S. federal court sentencing that high school graduates are likely to receive a slightly shorter sentence than otherwise similar graduates, though the difference is quite small (about 2–3 percent).

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

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 1—CENSUS DESCRIPTIVE STATISTICS: MEAN (STANDARD DEVIATION) BY YEAR

Variable	1960	1970	1980
In prison	0.0067	0.0051	0.0068
	(0.0815)	(0.0711)	(0.0820)
Years of schooling	10.54	11.58	12.55
	(3.56)	(3.39)	(3.07)
High school graduate +	0.48	0.63	0.77
	(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 EDUCATION (IN PERCENTAGE TERMS)

	All years	1960	1970	1980
White men				
High school dropout	0.83	0.76	0.69	0.93
High school graduate	0.34	0.21	0.22	0.39
Some college	0.24	0.21	0.13	0.27
College +	0.07	0.03	0.02	0.08
Black men				
Dropout	3.64	2.94	2.94	4.11
High school graduate	2.18	1.80	1.52	2.35
Some college	1.97	0.81	0.89	2.15
College +	0.66	0.00	0.26	0.75

Notes: High school dropouts are individuals with less than 12 years of schooling or 12 years but no degree; high school graduates have exactly 12 years of schooling and a high school degree. Individuals with some college have 13–15 years of schooling, and college graduates have at least 16 years of schooling and a college degree.

increase steadily from 10.5 in 1960 to 12.5 in 1980.[7]

Table 2 reports incarceration rates by race and educational attainment. The probability of imprisonment is substantially larger for blacks than for whites, and this is the case for all years and education categories. Incarceration rates for white men with less than 12 years of schooling are around 0.8 percent while they average about 3.6 percent for blacks over the three decades. Incarceration rates are monotonically declining with education for all years and for both blacks and whites.

An important feature to notice in Table 2 is that the reduction in the probability of imprisonment associated with higher schooling is substantially larger for blacks than for whites. For example, in 1980 the difference between high school dropouts and college graduates is 0.9 percent for whites and 3.4 percent for blacks. Because high school dropouts are likely to differ in many respects from individuals with more education, these differences do not necessarily represent the causal effect of education on the probability of incarceration. However, the patterns indicate that the effect may differ for blacks and whites. In the empirical analysis below, we allow for differential effects by race whenever possible.[8]

To account for other factors in determining incarceration rates, we begin by using OLS to examine the impacts of education. Figure 1 shows how education affects the probability of imprisonment at all schooling levels after controlling for age, state of birth, state of residence, cohort of birth, and year effects (i.e., the graphs display the coefficient estimates on the complete set of schooling dummies). The figure clearly shows a decline in incarceration rates with schooling beyond eighth grade, with a larger decline at the high school graduation stage than at any other schooling progression.

Ideally, we would like to estimate a general model where the effect of education on imprisonment varies across years of schooling. Because the instruments we use are limited in the range of schooling years affected and in the amount of actual variation, this is not empirically feasible. In fact, we cannot even use two-

[7] The data used in this paper are available at www.econ.ucla.edu/moretti.

[8] The stability in aggregate incarceration rates reported in Table 1 masks the underlying trends within each education group, which show substantial increases over the 1970's. The substantial difference in high school graduate and dropout incarceration rates combined with the more than 25-percent increase in high school graduation rates over this time period explains why aggregate incarceration rates remained relatively stable over time while within-education-group incarceration rates rose.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

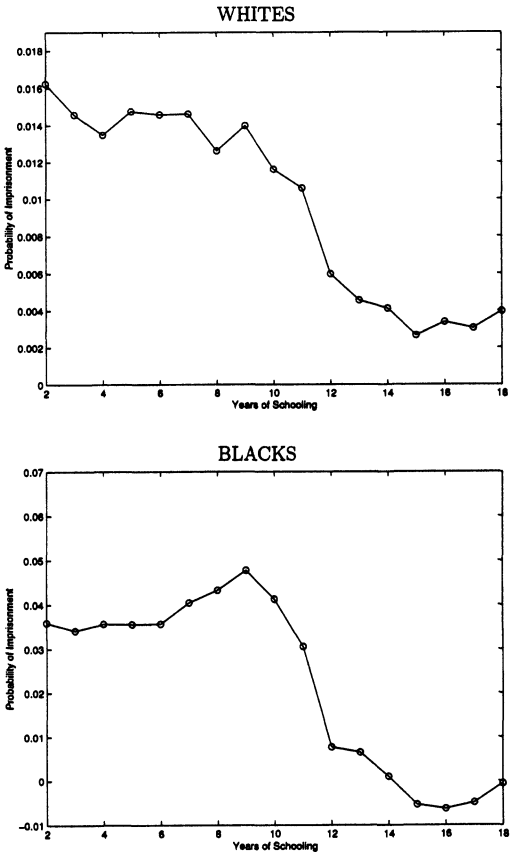

FIGURE 1. REGRESSION-ADJUSTED PROBABILITY OF INCARCERATION, BY YEARS OF SCHOOLING

Note: Regression-adjusted probability of incarceration is obtained by conditioning on age, state of birth, state of residence, cohort of birth, and year effects.

stage least squares (2SLS) to estimate a model of incarceration that is linear in school with a separate "sheepskin" effect of high school completion. Throughout the paper we present results both for models where the main independent variable is years of schooling and models where the main independent variable is a dummy for high school graduation.

Table 3 reports the estimated effects of years of schooling on the probability of incarceration using a linear probability model. Estimates for whites are presented in the top row with estimates for blacks in the bottom. In column (1), covariates include year dummies, age (14 dummies for three-year age groups, including 20–22, 23–25, 26–28, etc.), state of birth, and state

TABLE 3—OLS ESTIMATES OF THE EFFECT OF YEARS OF SCHOOLING ON IMPRISONMENT (IN PERCENTAGE TERMS)

	(1)	(2)	(3)
WHITES	−0.10	−0.10	−0.10
	(0.00)	(0.00)	(0.00)
BLACKS	−0.37	−0.37	−0.37
	(0.01)	(0.01)	(0.01)
Additional controls:			
Cohort of birth effects		y	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 (all coefficient estimates are multiplied by 100). All specifications control for age, year, state of birth, and state of residence. Sample in the top panel includes white males ages 20–60 in 1960, 1970, and 1980 Censuses; N = 3,209,138. Sample in the bottom panel includes black males ages 20–60 in 1960, 1970, and 1980 Censuses; N = 410,529. Age effects include 14 dummies (20–22, 23–25, etc.). State of birth effects are 49 dummies for state of birth (Alaska and Hawaii are excluded) and the District of Columbia. Year effects are three dummies for 1960, 1970, and 1980. State of residence effects are 51 dummies for state of residence and the District of Columbia. Cohort of birth effects are dummies for decade of birth (1914–1923, 1924–1933, 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.*

of current residence, which are all likely to be important determinants of criminal behavior and incarceration.[9] 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.[10] These estimates suggest that an additional year of schooling reduces the probability of incarceration by 0.1 percentage points for whites and by 0.37 percentage points for

[9] All specifications exclude Alaska and Hawaii as a place of birth, since our instruments below are unavailable for those states.

[10] Although the landmark *Brown v. Board of Education* was decided in 1954, there was little immediate response by states. We allow for a break in 1958, since at least two southern states made dramatic changes in their schooling policy that year in response to forced integration—both South Carolina and Mississippi repealed their compulsory schooling statutes to avoid requiring white children to attend school with black children.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

blacks.[11] The larger effect for blacks is consistent with the larger differences in unconditional means displayed in Table 2.

Column (2) accounts for unobserved differences across birth cohorts, allowing for differences in school quality or youth environments by including dummies for decade of birth (1914–1923, 1924–1933, etc.). Column (3) further controls for state of residence × year effects. This absorbs state-specific time-varying shocks or policies that may affect the probability of imprisonment and graduation. For example, an increase in prison spending in any given state may be offset by a decrease in education spending that year.[12] Both sets of estimates are insensitive to these additional controls.[13]

To gauge the size of these 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 educational attainment. In 1980, the difference in incarceration rates for whites and blacks is about 2.4 percent. Using the estimates for blacks, we conclude that 23 percent of this difference could be eliminated by raising the average education levels of blacks to the same level as that of whites.

B. *The Effect of Compulsory Attendance Laws on Schooling Achievement*

The OLS estimates just presented are consistent with the hypothesis that education 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 school. For example, 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 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.[14]

The ideal instrumental variable induces exogenous variation in schooling 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 education. Compulsory schooling laws have different forms. The laws typically determine the earliest age that a child is required to be in school and/or the latest age he is required to enroll and/or a minimum number of years that he is required to stay in school. We follow Acemoglu and Angrist (2000) and define years of compulsory attendance 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 school and the latest age he is required to enroll. Figure 2 plots the evolution of compulsory attendance laws over time for 48 states (all

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

[12] 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.

[13] Models that include AFQT scores, parents' education, whether or not the individual lived with both of his natural parents at age 14 and whether his mother was a teenager at his birth estimated using NLSY data yield results that are remarkably similar to those based on Census data. (See Section IV.) Probit models also yield similar estimated effects.

[14] 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 dropouts 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 percent 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), less than 10 percent of the gap is eliminated.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

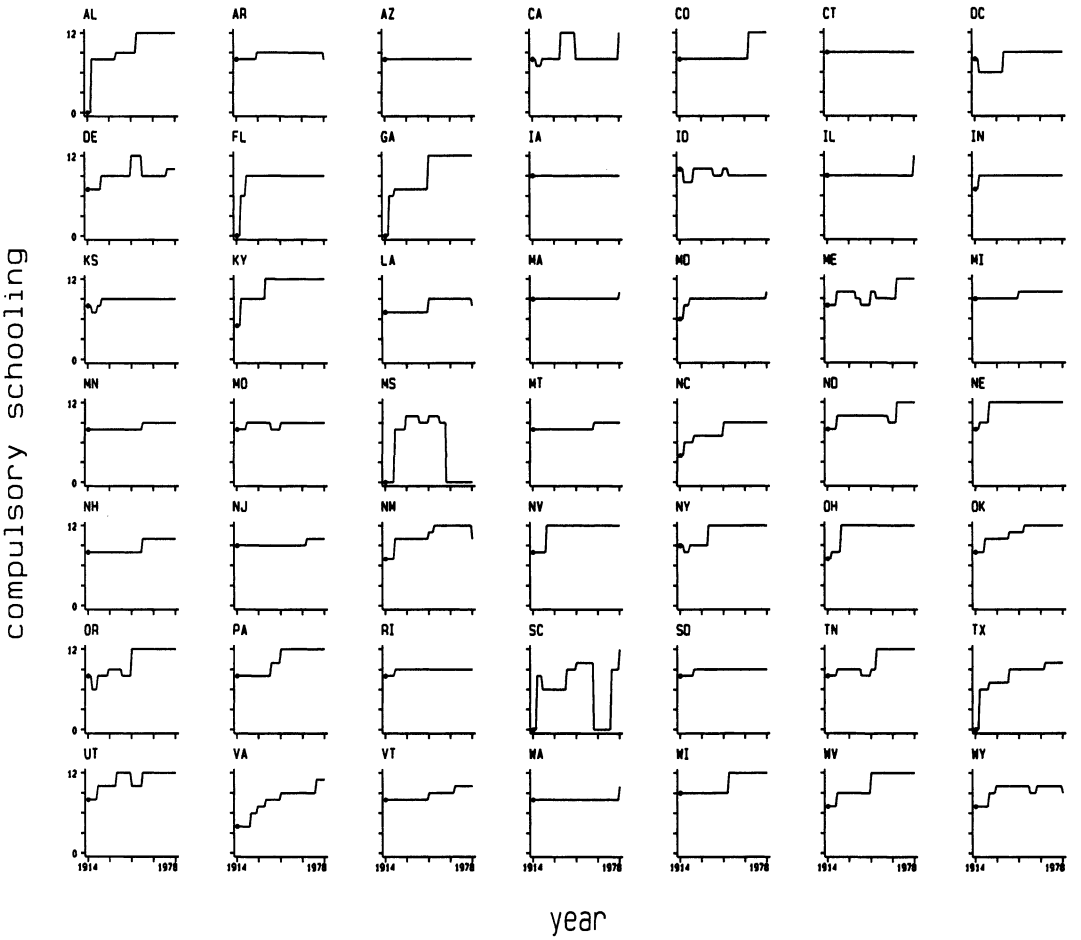

FIGURE 2. CHANGES IN COMPULSORY ATTENDANCE LAWS BY STATE 1914–1978

but Alaska and Hawaii) and the District of Columbia. In the years relevant for our sample, 1914 to 1974, states changed compulsory attendance levels several times, and not always upward.[15]

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.[16] The fraction of individuals belonging to each compulsory attendance group are reported in Table 1.

Figure 3 shows how the increases in compulsory schooling affect educational attainment

[15] The most dramatic examples of downward changes are South Carolina and Mississippi, which repealed their compulsory attendance statutes in 1958 in order to avoid requiring white children to attend racially mixed schools (Lawrence Kotin and William Aikman, 1980).

[16] The data sources for compulsory attendance laws are given in Appendix B of Acemoglu and Angrist (2000). We use the same cut-off points as Acemoglu and Angrist (2000). We experimented with a matching based on the year the individual is age 16 or 17, and found qualitatively similar results.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

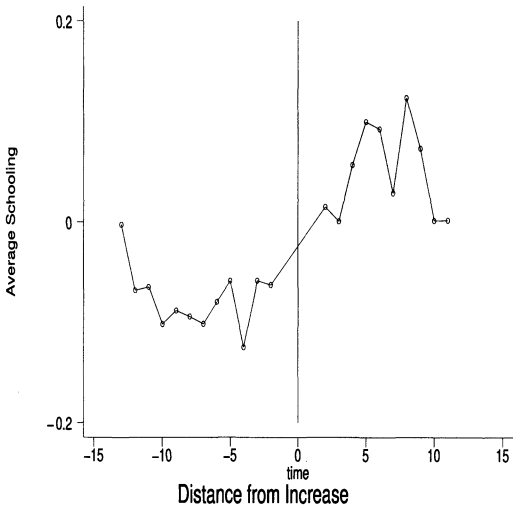

FIGURE 3. THE EFFECT OF INCREASES IN COMPULSORY ATTENDANCE LAWS ON AVERAGE YEARS OF SCHOOLING

over time, controlling for state and year of birth.[17] In the 12 years before the increase, there is no obvious trend in schooling achievement. All of the increase in schooling associated with stricter compulsory schooling laws takes place *after* changes in the law. This figure is important because it suggests that changes in compulsory schooling laws appear to raise education levels and not that they simply respond to underlying trends in schooling. More formal tests are provided below.

Table 4 quantifies the effect of compulsory attendance laws on different levels of educational achievement. These 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, they 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.

Consider the estimates for whites presented in the top panel. Three points are worth making. First, the more stringent the compulsory attendance legislation, the lower is the percentage of high school dropouts. In states/years requiring 11 or more years of compulsory attendance, the number of high school dropouts is 5.5 percent lower than in states/years requiring eight years or less (the excluded case). These effects have been documented by Acemoglu and Angrist (2000) and Adriana Lleras-Muney (2002).[18] Second, the coefficients in columns (1) and (2) are roughly equal, but with opposite sign. For example, in states/years requiring nine years of schooling, the share of high school dropouts is 3.3 percentage points lower than in states/years requiring eight 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 dropouts by "forcing" them to stay in school. Third, the effect of compulsory attendance is smaller, and in most cases, not significantly different from zero in columns (3) and (4). Finding a *positive* effect on higher levels of schooling may indicate that the laws are correlated with underlying trends of increasing education, which would cast doubt on their exogeneity. This does not appear to be a problem in the data. The coefficient on compulsory attendance ≥ 11 for individuals with some college is negative, although small in magnitude, suggesting that states imposing the most stringent compulsory attendance laws experience small declines in the number of individuals attending community college. This result may indicate a shift in state re-

[17] The figure shows the estimated coefficients on leads and lags of an indicator for whether compulsory schooling increases in an individual-level regression that also controls for state of birth and year of birth effects. The dependent variable is years of schooling. Lags include years -12 to -3. Leads include years $+3$ to $+12$. Time $= 0$ represents the year the respondent is age 14.

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

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 4—THE EFFECT OF COMPULSORY ATTENDANCE LAWS ON SCHOOLING ACHIEVEMENT (IN PERCENTAGE TERMS)

	Dropout (1)	High school (2)	Some college (3)	College+ (4)
WHITES				
Compulsory attendance = 9	−3.25	3.27	−0.04	0.03
	(0.34)	(0.37)	(0.17)	(0.20)
Compulsory attendance = 10	−3.31	4.01	−0.30	−0.39
	(0.45)	(0.51)	(0.30)	(0.33)
Compulsory attendance ≥ 11	−5.51	5.82	−0.68	0.36
	(0.47)	(0.52)	(0.26)	(0.32)
F-test [p-value]	47.91	45.47	3.05	1.67
	[0.000]	[0.000]	[0.027]	[0.171]
R^2	0.12	0.02	0.04	0.05
BLACKS				
Compulsory attendance = 9	−2.36	3.09	−0.69	−0.03
	(0.46)	(0.41)	(0.23)	(0.16)
Compulsory attendance = 10	−1.76	4.06	−1.82	−0.47
	(0.65)	(0.64)	(0.39)	(0.23)
Compulsory attendance ≥ 11	−2.96	5.02	−1.89	0.16
	(0.69)	(0.62)	(0.34)	(0.25)
F-test [p-value]	10.09	27.13	12.76	1.85
	[0.000]	[0.000]	[0.000]	[0.136]
R^2	0.19	0.07	0.06	0.02

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 dropout. Coefficient estimates are multiplied by 100. The dependent variables in columns (2)–(4) are dummies for high school, some college, and college, respectively. All specifications control for age, year, state of birth, state of residence, and cohort of birth. Sample in the top panel includes white males ages 20–60 in 1960, 1970, and 1980 Censuses; N = 3,209,138. Sample in the bottom panel includes 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) and the District of Columbia. Year effects are three dummies for 1960, 1970, and 1980. State of residence effects are 51 dummies for state of residence and the District of Columbia. Cohort of birth effects are dummies for decade of birth (1914–1923, 1924–1933, 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*. F-tests are for whether the coefficients on the excluded instruments are jointly equal to zero, conditional on all the controls (3 degrees of freedom).

sources from local community colleges to high schools following the decision to raise compulsory attendance laws.

The bottom panel in Table 4 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-out 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 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

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

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 5—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 expenditures (3)
Compulsory attendance = 9	0.002	0.103	−0.002
	(0.008)	(0.186)	(0.002)
Compulsory attendance = 10	−0.003	−0.430	−0.015
	(0.010)	(0.209)	(0.003)
Compulsory attendance = 11	−0.008	−0.340	−0.011
	(0.010)	(0.180)	(0.003)
R^2	0.81	0.89	0.85
N	343	1,500	1,500

Notes: Standard errors are in parentheses. All specifications control for year and state effects. The dependent variable in column (1) is the percentage policemen in the state. Sample in column (1) includes observations from 48 states and the District of Columbia in years 1920, 1930, 1940, 1950, 1960, 1970, and 1980. The number of policemen in 1920–1940 are taken from Census reports on occupations and the labor force for the entire U.S. population. Data from 1950–1980 are from the IPUMS 1 percent 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: "City Police Expenditures, 1946–1985." See text for details.

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, 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 including state of residence × year effects.

Second, 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 5 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. 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.[20] No clear pattern emerges from columns (1) and (2), while there appears to be a negative correlation in column (3). Overall, we 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).

[20] Data on police expenditures are taken from ICPSR Study 8706: "City Police Expenditures, 1946–1985." To obtain state-level expenditures, we added the expenditures of all available cities in a state.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 6—THE EFFECT OF FUTURE COMPULSORY ATTENDANCE LAWS ON CURRENT GRADUATION STATUS (IN PERCENTAGE TERMS)

	WHITES			BLACKS		
	Compulsory attendance = 9 (1)	Compulsory attendance = 10 (2)	Compulsory attendance ≥ 11 (3)	Compulsory attendance = 9 (4)	Compulsory attendance = 10 (5)	Compulsory attendance ≥ 11 (6)
$t = +4$	−0.32	0.25	−1.41	0.54	−1.53	−1.64
	(1.22)	(1.82)	(2.14)	(0.67)	(1.10)	(1.44)
$t = +5$	0.04	0.85	−0.07	−0.04	−0.98	−0.68
	(0.78)	(1.13)	(1.41)	(0.46)	(0.81)	(1.01)
$t = +6$	0.06	1.00	0.27	−0.43	−1.32	−1.60
	(0.69)	(0.93)	(1.21)	(0.45)	(0.73)	(0.95)
$t = +7$	0.01	1.07	0.27	−0.72	−1.36	−0.24
	(0.57)	(0.78)	(1.21)	(0.43)	(0.79)	(0.90)
$t = +8$	0.13	1.06	0.91	−0.99	−1.06	−0.47
	(0.54)	(0.71)	(0.86)	(0.42)	(0.79)	(0.83)
$t = +9$	0.16	0.92	−0.94	−1.26	−1.04	−0.60
	(0.51)	(0.67)	(0.80)	(0.41)	(0.79)	(0.70)
$t = +10$	0.11	0.95	1.23	−1.40	−0.84	−0.41
	(0.46)	(0.63)	(0.71)	(0.45)	(0.78)	(0.75)
$t = +11$	−0.13	0.63	1.31	−1.56	−0.71	−0.20
	(0.43)	(0.55)	(0.69)	(0.49)	(0.75)	(0.78)
$t = +12$	−0.61	0.16	0.80	−1.58	−0.17	−0.42
	(0.47)	(0.54)	(0.72)	(0.50)	(0.70)	(0.75)
$t = +15$	−0.92	−0.18	0.078	−0.97	1.22	−0.44
	(0.46)	(0.54)	(0.66)	(0.52)	(0.63)	(0.79)
$t = +18$	−0.67	0.19	1.31	−0.20	2.71	−0.61
	(0.46)	(0.55)	(0.56)	(0.55)	(0.61)	(0.85)
$t = +20$	−0.65	0.40	0.76	0.13	3.49	0.40
	(0.50)	(0.60)	(0.59)	(0.64)	(0.71)	(0.83)

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 graduate. Coefficient estimates are multiplied by 100. Each row is a separate regression. All models control for compulsory attendance laws at $t = 0$, $t = 1$, $t = 2$, and $t = 3$, as well as year, age, state of birth, state of residence, and cohort of birth. 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) and the District of Columbia. Year effects are three dummies for 1960, 1970, and 1980. State of residence effects are 51 dummies for state of residence and the District of Columbia. Cohort of birth effects are dummies for decade of birth (1914–1923, 1924–1933, 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.* In column (1), (2), and (3) sample includes white males ages 20–60 in 1960, 1970, and 1980 Censuses. In column (4), (5), and (6) sample includes black males ages 20–60 in 1960, 1970, and 1980 Censuses. N = 3,209,138 for whites; N = 410,529 for blacks.

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 experience faster increases in high school graduation rates. As discussed earlier, Figure 2 shows that increases in average education levels follow increases in compulsory schooling ages. We now quantify the relationship between future compulsory attendance laws and current graduation rates, since that is an important education margin affected by the laws. If causality runs from compulsory attendance laws to schooling, we should observe that future laws do not affect current graduation rates conditional on current compulsory attendance laws. Results of this test are reported in Table 6. The coefficients in the first row, for example, represent the effect of compulsory attendance laws that are in place four 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. Overall, the results in Table 6 suggest that states with faster expected increases in graduation rates are not more likely to change their

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

compulsory attendance laws.[21] This result is consistent with the findings of Lleras-Muney (2002), who examines these laws from 1925–1939.

C. *Instrumental Variable Estimates*

We now present 2SLS estimates of the impact of schooling on the probability of incarceration using models identical to our earlier OLS specifications. The 2SLS estimates in Table 7 suggest that one extra year of schooling reduces the probability of imprisonment by about 0.1 percentage points for whites and 0.3–0.5 percentage points for blacks. These estimates are stable across specifications and nearly identical to the corresponding OLS estimates shown in Table 3. (We cannot reject that they are the same using a standard Hausman test.) This indicates that the endogeneity bias is not quantitatively important after controlling for age, time, state of residence, and state of birth.

An important concern with an IV approach is the possible use of weak instruments, which tends to bias 2SLS estimates towards OLS estimates and may weaken standard tests for endogeneity. The existing econometric literature defines weak instruments based on the strength of the first-stage equation (e.g., Paul Bekker, 1994; Douglas Staiger and James H. Stock, 1997; Stock and Motohiro Yogo, 2003). Are our instruments weak by this standard? F-statistics based on the test of whether compulsory schooling attendance laws all have zero coefficients (conditioning on all other controls) range between 36.2 and 52.5 for whites and between 41.5 and 88.1 for blacks. These test statistics are well above the critical values for weak instruments as reported by Stock and Yogo (2003). This is true for both the critical values based on 2SLS bias and the ones based on 2SLS size. (These critical values are obtained using weak instruments asymptotic distributions.) This implies that, according to traditional tests for weak instruments, our first stage has good power and our instruments are not weak.

Still, estimates suggest that "reduced-form" models that directly regress incarceration on the compulsory schooling laws produce a fairly weak relationship. The estimated reduced-form effects of compulsory schooling laws on the probability of incarceration [corresponding to the specification reported in column (3)] for whites are −0.14 (0.09), −0.08 (0.14), and −0.31 (0.13) for compulsory schooling ages equal to 9, 10, and 11 or 12 years, respectively. Corresponding estimates for blacks are −0.005 (0.01), −0.014 (0.02), and −0.056 (0.02).[22] The reduced-form F-tests are 3.18 (with a p-value of 0.023) and 2.07 (with a p-value of 0.10) for whites and blacks, respectively.

Given the weak reduced-form effects, it is, perhaps, surprising that our 2SLS estimates of the effect of crime on schooling are statistically significant. At first glance, this would appear to be a contradiction. Upon closer look, it is not. In general, there need not be any relationship between significance in the reduced form and significance for 2SLS estimates. This is because the reduced-form residual is the sum of the first-stage equation residual and the outcome equation residual. If these two residuals are negatively correlated, we should expect larger standard errors for reduced-form estimates than 2SLS estimates. We show this point formally for the case with a single endogenous regressor and a single instrument in Appendix A.[23]

[21] Only one estimated coefficient for whites is significantly positive ($t = +18$). The only significant positive 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 positive, there is no relationship between stringency of the law and high school dropout, making it difficult to interpret this finding.

[22] These estimates are reported in percentage terms and are comparable to those in related tables.

[23] The weak instruments literature has focused on the strength of the first-stage regression rather than the reduced-form equation. Intuitively, this focus is motivated by the fact that a weak first stage leads to invertability problems for the 2SLS estimator while a weak reduced form does not [i.e., the standard IV estimator, $(d'x)^{-1}d'y$, with dependent variable y, regressor x, and instrument d, breaks down when $d'x$ is near zero while it does not when $d'y$ approaches zero]. More generally, there is not a one-to-one correspondence between the power of the first stage and the power of the reduced form in overidentified 2SLS models. See Jinyong Hahn and Jerry Hausman (2002) and their discussion of both the "forward" and "reverse" model. In our context, the instruments are strong for the "forward" model (regression of incarceration on schooling) but they are weak instruments for the "reverse" model (regression of schooling on incarceration). This suggests that we should obtain consistent estimates for our model but would obtain biased estimates of the reverse model. Because Limited Information Maximum Likelihood (LIML) can be understood as a combination of the "forward" and "reverse" 2SLS estimators (see Hahn and Hausman, 2002) and in our case one of them is problematic, we cannot use LIML despite its potential advantages.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 7—IV AND CONTROL FUNCTION ESTIMATES OF THE EFFECT OF YEARS OF SCHOOLING ON IMPRISONMENT (IN PERCENTAGE TERMS)

	IV estimates			Control function
	(1)	(2)	(3)	(4)
WHITES				
Second stage				
Years of schooling	−0.11	−0.09	−0.14	−0.09
	(0.02)	(0.05)	(0.06)	(0.05)
First stage				
Compulsory attendance = 9	0.278	0.222	0.202	
	(0.026)	(0.024)	(0.024)	
Compulsory attendance = 10	0.213	0.199	0.176	
	(0.035)	(0.034)	(0.033)	
Compulsory attendance ≥ 11	0.422	0.340	0.329	
	(0.037)	(0.033)	(0.033)	
First stage F-test (d.o.f. = 3)	52.5	38.6	36.2	
Hausman test (p-value)	0.35	0.90	0.73	
Control function				
$\hat{v}$				−0.04
				(0.05)
$\hat{v}$ × years of schooling				0.00
				(0.00)
BLACKS				
Second stage				
Years of schooling	−0.47	−0.33	−0.41	−0.35
	(0.12)	(0.18)	(0.19)	(0.18)
First stage				
Compulsory attendance = 9	0.672	0.454	0.421	
	(0.043)	(0.040)	(0.039)	
Compulsory attendance = 10	0.664	0.476	0.434	
	(0.079)	(0.071)	(0.070)	
Compulsory attendance ≥ 11	0.794	0.528	0.509	
	(0.068)	(0.063)	(0.062)	
First stage F-test (d.o.f. = 3)	88.1	45.9	41.5	
Hausman test (p-value)	0.87	0.85	0.83	
Control function				
$\hat{v}$				0.20
				(0.18)
$\hat{v}$ × years of schooling				−0.02
				(0.00)
Additional controls:				
Cohort of birth effects		y	y	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. Second stage and control function estimates are multiplied by 100. All specifications control for age, year, state of birth, and state of residence. See Table 4 for a description of the sample and regressors. The F-test tests whether the coefficients on the excluded instruments are jointly equal to zero. Degrees of freedom for the Hausman tests is 1.

If the effect of schooling on imprisonment varies across individuals, then OLS and 2SLS may not estimate the "average treatment effect" of schooling [i.e., $E(\beta)$]. Under conditions specified by John Garen (1984) and David Card (1999), a linear control function approach can be used to estimate the "average treatment effect" when β varies in the population.[24] We specify these assumptions and the resulting estimating equation in Appendix

[24] See Jeffrey M. Wooldridge (1997) or Heckman and Edward Vytlacil (1998), for a discussion of the conditions needed for OLS and 2SLS to identify the average treatment effect. See Heckman and Richard J. Robb (1985) for a general treatment of control function methods.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 8—THE EFFECT OF YEARS OF SCHOOLING ON INCARCERATION (IN PERCENTAGE TERMS)—ROBUSTNESS CHECKS

		WHITES		BLACKS	
		OLS (1)	2SLS (2)	OLS (3)	2SLS (4)
(A)	Base case	−0.10	−0.14	−0.37	−0.41
		(0.00)	(0.06)	(0.01)	(0.19)
	First-stage F-test (d.o.f. = 3)		36.2		41.5
(B)	Region of birth × cohort trend	−0.10	−0.19	−0.37	−0.73
		(0.00)	(0.10)	(0.01)	(0.26)
	First-stage F-test (d.o.f. = 3)		12.79		28.39
(C)	Region of birth × cohort effects	−0.10	−0.22	−0.37	−0.34
		(0.00)	(0.17)	(0.01)	(0.35)
	First-stage F-test (d.o.f. = 3)		5.74		22.41
(D)	State of birth × cohort trend	−0.10	−0.34	−0.37	−0.67
		(0.00)	(0.21)	(0.01)	(0.32)
	First-stage F-test (d.o.f. = 3)		5.83		19.15
(E)	Age effects × cohort effects	−0.10	−0.17	−0.37	−0.33
		(0.00)	(0.07)	(0.01)	(0.23)
	First-stage F-test (d.o.f. = 3)		37.90		35.68
(F)	Education	−0.38	−0.65	−1.41	−0.72
		(0.01)	(0.14)	(0.04)	(0.63)
	Education × age	0.01	0.01	0.02	0.00
		(0.00)	(0.00)	(0.00)	(0.01)
	Effect at age 20	−0.24	−0.55	−1.09	−0.67
		(0.01)	(0.14)	(0.04)	(0.65)
	Effect at age 40	−0.17	−0.38	−0.68	−0.54
		(0.01)	(0.15)	(0.05)	(0.71)
	First-stage F-test (d.o.f. = 6)		19.2–28.8		24.3–34.8

Notes: Standard errors corrected for state of birth–year of birth clustering are in parentheses. All specifications control for age, state of birth, cohort of birth, and state of residence × year. See Table 4 for a description of the sample and regressors. The F-test is for whether the coefficients on the excluded instruments are jointly equal to zero, conditional on all the controls.

B. In column (4) of Table 7, we report control function estimates of a model that includes dummies for age, year, state of residence, state of birth, and cohort of birth. These estimates are very similar to the corresponding OLS and 2SLS estimates in column (2), suggesting that heterogeneity across individuals does not appear to be important in estimation of the "average treatment effect" of schooling on incarceration.[25]

In Table 8, we probe the robustness of our OLS and 2SLS estimates to different specifications. All specifications control for age, year × state of residence, state of birth, and cohort of birth. Specification A reports the base case results from Table 7 [column (3)] for ease of comparison. The following three models aim at absorbing trends that are specific to the region or the state of birth to account for geographic differences in school quality over time, as well as differences in other time-varying factors that

[25] We also employed Amemiya's Generalized Least-Squares estimator (Whitney K. Newey, 1987), the probit analog with endogenous regressors. The estimated effects of schooling on the probability of incarceration were generally negative but smaller in magnitude and more sensitive to the specification.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

are specific to the state of birth and correlated with schooling. Specification B includes region of birth-specific linear trends in year of birth. Specification C includes the interaction of region of birth effects and cohort of birth effects. Specification D further relaxes the model by allowing for different trends in cohort quality at the state level.

These three specifications come close to fully saturating the model. For example, in specification D the 2SLS estimator is identified only by deviations of compulsory attendance laws from a linear trend. The loss of identifying variation in the first stage is indicated by the drop in reported first-stage F-test statistics. OLS estimates are unchanged. While the 2SLS estimates show greater effects, they are much less precise and statistically indistinguishable from the base case estimates.

Specification E allows the cohort effects to vary with age, capturing the possibility that age-crime patterns have varied over time. Estimates are similar to the base case.

Finally, specification F allows the impact of education on the probability of incarceration to vary with age. Ideally, one would like to split the sample into two or three age groups, running separate regressions for each group. However, there is not enough variation in the data to obtain precise IV estimates separately for each age group. The estimates of model F suggest that the effects are larger for younger men, declining with age. In addition to the coefficient estimates, we report the implied effects at ages 20 and 40. Among white men, the 2SLS estimates suggest that an additional year of schooling reduces the probability of incarceration by about 0.55 percentage points at age 20 and by 0.38 percentage points at age 40. The corresponding estimates for blacks imply an effect of 0.67 and 0.54 percentage points at ages 20 and 40, respectively. These estimates suggest that racial differences in the estimated effect of education on the probability of incarceration are partially due to differences in age levels among blacks and whites in the population. Garen estimates corresponding to specification F suggest even larger effects at all ages.[26]

We also explore the robustness of our findings to aggregation within age–state of birth–year cells to shed light on any "aggregation bias" that may arise in our estimation of the effects of education on aggregate arrest rates. Specifically, we aggregate our sample to compute incarceration rates and average schooling levels by age (eight age categories), state of birth, and year. We then use these aggregate observations to estimate specifications analogous to those in Tables 4 and 7.[27] The results of this procedure are quite similar to those using individual-level regressions and are reported in Appendix Tables C1 and C2.

Overall, our findings indicate that endogeneity bias is not likely to be empirically important for OLS estimation after controlling for age, time, state of residence, and state of birth. Our estimates suggest an economically important and statistically significant effect of schooling on the probability of incarceration with larger effects for blacks than for whites. Based on our base case specification which controls for age, year, state of birth, state of residence, cohort of birth, and year-specific state of residence effects, an additional year of schooling reduces the probability of incarceration by about 0.1 percentage point for whites and 0.4 percentage points for blacks.

D. *The Effect of High School Graduation on Imprisonment*

While we have estimated the effects of schooling on the probability of incarceration assuming a linear relationship between the two, Figure 1 suggests that the effects of education on crime may be nonlinear. In this case, our OLS and 2SLS linear-in-schooling estimators identify weighted averages of all grade transition effects on the probability of incarceration.[28] Because of the limited variation in our

[26] Garen estimates for whites suggest 1.1-percentage-point reduction at age 20 and a 0.7-percentage-point reduction at age 40. For blacks, the estimated reductions are 2.5 and 1.5 percentage points at ages 20 and 40, respectively.

[27] Rather than using state of residence as a dummy regressor as was done in the individual-level regressions, we use the fraction of men in a particular age–state of birth–year cell residing in each state as regressors.

[28] OLS weights depend on the distribution of schooling in the population (Shlomo Yitzhaki, 1996), while 2SLS weights depend on the fraction of individuals switching

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 9—ESTIMATES OF THE EFFECT OF HIGH SCHOOL GRADUATION ON IMPRISONMENT (IN PERCENTAGE TERMS)

	OLS estimates		IV estimates		Control function
	(1)	(2)	(3)	(4)	(5)
WHITES					
Second stage					
High school	−0.77	−0.77	−0.61	−0.89	−0.97
	(0.02)	(0.02)	(0.35)	(0.37)	(0.32)
First-stage F-test (d.o.f. = 3)			47.91	48.05	
Hausman test (p-value)			0.99	0.78	
$\hat{v}$					−2.16
					(0.36)
$\hat{v} \times$ high school					2.02
					(0.12)
BLACKS					
Second stage					
High school	−3.39	−3.39	−7.23	−8.00	−11.40
	(0.01)	(0.01)	(3.66)	(3.78)	(3.68)
First-stage F-test (d.o.f. = 3)			10.09	10.01	
Hausman test (p-value)			0.27	0.20	
$\hat{v}$					−7.02
					(3.65)
$\hat{v} \times$ high school					3.96
					(0.90)
Additional controls:					
State of residence × year effects		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 in prison. All coefficient estimates are multiplied by 100. All specifications control for age, year, state of birth, cohort of birth, and state of residence. Sample in the top panel includes white males ages 20–60 in 1960, 1970, and 1980 Censuses; N = 3,209,138. Sample in the bottom panel includes black males ages 20–60 in 1960, 1970, and 1980 Censuses. N = 410,529. Age effects include 14 dummies (20–22, 23–25, etc.). State of birth effects are 49 dummies for state of birth (Alaska and Hawaii are excluded) and the District of Columbia. Year effects are three dummies for 1960, 1970, and 1980. State of residence effects are 51 dummies for state of residence and the District of Columbia. Cohort of birth effects are dummies for decade of birth (1914–1923, 1924–1933, 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.* The F-test is for whether the coefficients on the excluded instruments are jointly equal to zero, conditional on all the controls. The degree of freedom for the Hausman test is 1.

schooling instruments, it is impossible to estimate the effects of each grade transition on the probability of incarceration (as estimated by OLS and represented in Figure 1) using 2SLS.

But given the importance placed on high school graduation by policy makers and the large apparent effect of high school graduation on crime reflected in Figure 1, we estimate a specification which includes an indicator for high school completion rather than total years of completed schooling.

OLS and 2SLS estimates of the impact of high school completion are reported in Table 9.[29] The OLS estimates indicate that white

from one schooling level to another in response to the introduction of compulsory schooling laws (Guido W. Imbens and Angrist, 1994). Because these weights are based on observable information, they can be estimated. See Appendix A and Figures 3 and 4 of Lochner and Moretti (2001) for a detailed discussion and empirical representation of these weights.

[29] We ignore the fact that in some years, high school graduation in South Carolina could be achieved with 11 years of schooling. We also ignore the fact that some inmates graduate in prison, which is uncommon in the years

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

high school graduates have a 0.76-percentage-points lower probability of incarceration than do dropouts. 2SLS estimates are quite similar. Incarceration rates among black graduates are 3.4 percentage points lower than among dropouts according to the OLS estimates. 2SLS estimates are larger, ranging from -7 to -8 percentage points. We cannot reject that OLS and 2SLS estimates are equal for either blacks or whites using standard Hausman specification tests. In Lochner and Moretti (2001), we discuss in detail how nonlinearities in the schooling-crime relationship and differences between OLS and 2SLS "weights" on each grade-specific effect can generate the observed differences in these OLS and 2SLS estimates.

As with the linear-in-schooling case, OLS and 2SLS might differ if the effect of high school graduation on crime varies across individuals. The final column of Table 9 reports estimates using the linear control function approach of Garen (1984). These estimates are larger than both the OLS and 2SLS estimates for both whites and blacks. If there is any bias in our OLS estimates due to unobserved heterogeneity or self-selection, these estimates suggest that it is toward finding no effect of education on crime.

How do these results compare with models based on years of schooling in Tables 3 and 7? For whites, the mean gap in education between high school dropouts and those with at least high school is 5.34. If we multiply this gap by the OLS estimate in Table 3, we get $-0.10 \times 5.34 = -0.53$. This is less than the corresponding estimate in Table 9 [columns (1)–(2)]: -0.77. The discrepancy is slightly smaller for 2SLS estimates. If we multiply the gap by the 2SLS estimate in Table 7 [column (3)], we get $-0.14 \times 5.34 = -0.75$. The corresponding estimate in Table 9 is -0.89. For blacks, the education gap is quite similar: 5.33. For OLS estimates, the comparison is $-0.37 \times 5.33 = -2.0$ vs. -3.39. For IV estimates the comparison is $-0.41 \times 5.33 = -2.2$ vs. -8.0.

Nonlinearities in the relationship between crime and schooling may explain why the effect of high school graduation estimated in Table 9 is larger than the effect of an additional year of school estimated in Tables 3 and 7 multiplied by the average difference in schooling. Figure 1 suggests a large drop in the probability of incarceration when moving from 11 to 12 years of schooling. At the same time, OLS estimates based on the high school graduation dummy specification tend to more heavily weight the effect of finishing grade 12 (relative to finishing other grades) than does the linear-in-schooling specification (see Lochner and Moretti, 2001, for empirical estimates of these weights). So, if finishing high school has a larger effect than other transitions (as suggested by Figure 1), the estimated graduate-dropout difference should be greater when using the graduation dummy specification than the linear-in-schooling specification. The comparison of 2SLS estimates is more complicated, but they also differ when nonlinearities in the crime-schooling relationship are present (Lochner and Moretti, 2001).

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

III. The Impact of Schooling on Arrest Rates

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.

To relate arrest rates to schooling and racial composition, we augment the arrest data with average education levels and high school graduation rates by age and state as well as the percentage black by age in each state from the 1960–1990 Censuses. We estimate the following model:

$$(2) \quad \ln A_{cast} = \beta E_{ast} + \gamma B_{ast} + d_{st} + d_{sc} + d_{sa} + d_{ct} + d_{at} + d_{ac} + e_{cast}$$

where $\ln A_{cast}$ is the logarithm of the male arrest rate for crime c, age group a, in state s in year

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

t (from UCR); E_{ast} is either average education or the high school graduation rate for males in 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 education on arrest rates is assumed to be the same *in percentage terms* for all crimes.[30] In a few specifications, we allow the effect of schooling 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 × 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 schooling 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. For example, average schooling and arrest rates of men ages 20–24 are different from average schooling and arrest rates of men ages 25–29 in the same state and year.

In estimating equation (2), 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 × crime and state × age heterogeneity in arrest rates. Crime-specific and age-specific trends in arrest common to all states are accounted for by crime × year dummies, d_{ct}, and age × year dummies, d_{at}, respectively. Finally, age × 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 eight age groups (20–24, 25–30, etc.), nine crimes (murder, rape, assault, robbery, burglary, larceny, auto theft, and arson), and 50 states plus the District of Columbia.

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 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.[31]

The estimated impacts of education on arrest rates are reported in Table 10. The top half reports the effects of average education levels and the bottom half reports the effects of high school graduation rates. Columns (1)–(3) report OLS estimates, and columns (4)–(6) report 2SLS estimates using compulsory schooling laws as instruments. We assign the compulsory attendance laws based on the state where the arrest took place and the year the arrestees were age 14.[32] All models are weighted by cell size.

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

[31] 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 × year effects, our specification is similar to Levitt's (1998). (They would be identical if we studied only one type of crime.)

[32] Unfortunately, we cannot assign compulsory attendance directly to individuals as we could with the Census

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 10—OLS AND IV ESTIMATES OF THE EFFECT OF SCHOOLING ON ARREST RATES

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
(A) AVERAGE EDUCATION						
Average years of education	−0.114	−0.116	−0.111	−0.176	−0.182	−0.162
	(0.024)	(0.023)	(0.042)	(0.080)	(0.080)	(0.105)
R^2	0.89	0.93	0.95			
(B) HIGH SCHOOL GRADUATION						
High school graduation rate	−0.618	−0.674	−0.710	−0.946	−0.941	−0.873
	(0.183)	(0.181)	(0.283)	(0.491)	(0.522)	(0.669)
R^2	0.93	0.95	0.96			
Controls:						
age × offense effects	y	y	y	y	y	y
offense × year effects	y	y	y	y	y	y
age × year effects	y	y	y	y	y	y
state × age effects	y	y	y	y	y	y
state × offense effects		y	y		y	y
state × year			y			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. Average schooling and high school graduation rate is by age group, state, and year (see text). All models control for percentage black. There are eight age groups, eight offenses, 50 states plus the District of Columbia, and four years. All models are weighted by cell size.

Since variation in arrest rates occurs across offense type, age, state, and year, and variation in graduation rates occurs across age, state, and year, standard errors are corrected for state-year-age clustering.

The OLS estimates suggest that a one-year increase in average education levels is estimated to reduce arrest rates by 11 percent. 2SLS estimates suggest slightly larger effects, although they are not statistically different. While the standard errors more than double when using 2SLS, the estimates are still generally statistically significant. Given the importance of high school completion in determining incarceration rates, we also explore the relationship between high school graduation rates and arrest rates in the bottom half of the table. The OLS estimated impacts of high school graduation rates range from 0.6–0.7, while 2SLS estimates suggest a larger effect (though they are less precisely estimated).[33]

Table 11 allows for differential effects of schooling across different types of crime. The top half distinguishes between violent and property crimes, while the bottom half examines arrests for more detailed types of crimes. In interpreting these results, recall 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. Estimates for years of schooling are in columns (1) and (2). The upper panel shows similar effects across the broad categories of violent and property crime; however, the bottom panel suggests that the effects vary considerably within these categories. A one-year increase in average years of schooling reduces murder and assault by almost 30 percent, motor

data. Nor can we assign compulsory attendance based on the *state of birth,* since it is not available in the FBI aggregate data. Because of these data limitations, we expect a decrease in precision. Still the first-stage estimated effects of compulsory schooling laws on education are significant.

[33] Note that OLS and IV estimates do not necessarily estimate the average treatment effect when the effect of schooling varies in the population. See the earlier discussion in Section II, subsections C and D.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 11—OLS ESTIMATES FOR ARREST RATES BY TYPE OF CRIME

	Average education		High school graduation rate	
	(1)	(2)	(3)	(4)
(A) VIOLENT vs. PROPERTY CRIME				
Violent crime	−0.121	−0.116	−0.751	−0.793
	(0.025)	(0.044)	(0.198)	(0.291)
Property crime	−0.111	−0.105	−0.593	−0.621
	(0.026)	(0.044)	(0.208)	(0.304)
(B) BY DETAILED TYPE OF CRIME				
Murder	−0.276	−0.274	−2.062	−2.133
	(0.041)	(0.058)	(0.403)	(0.403)
Rape	0.113	0.118	1.094	1.049
	(0.037)	(0.048)	(0.307)	(0.353)
Robbery	−0.007	−0.005	0.184	0.113
	(0.031)	(0.047)	(0.253)	(0.333)
Assault	−0.297	−0.292	−2.136	−2.179
	(0.028)	(0.048)	(0.226)	(0.326)
Burglary	−0.057	−0.052	−0.202	−0.250
	(0.032)	(0.048)	(0.268)	(0.347)
Larceny	−0.058	−0.052	−0.235	−0.277
	(0.027)	(0.045)	(0.209)	(0.311)
Vehicle theft	−0.201	−0.197	−1.227	−1.271
	(0.030)	(0.048)	(0.251)	(0.346)
Arson	−0.133	−0.127	−0.745	−0.784
	(0.044)	(0.053)	(0.358)	(0.408)
Additional controls: state × year		y		y

Notes: Standard errors corrected for state-year-age clustering are in parentheses. Violent crimes include murder, rape, robbery, and assault. Property crimes include burglary, larceny, vehicle theft, and arson. Average schooling and high school graduation rate are by age group, state, and year (see text). All specifications control for percentage black, age × offense effects, offense × year effects, age × year effects, state × age effects, and state × offense effects. There are eight age groups, eight offenses, 50 states plus the District of Columbia, and four years. All models are weighted by cell size.

vehicle theft by 20 percent, arson by 13 percent, and burglary and larceny by about 6 percent. Estimated effects on robbery are negligible, while those for rape are significantly positive. This final result is surprising and not easily explained by standard economic models of crime.[34]

We find very similar patterns when looking at the relationship between high school graduation rates and arrest rates, reported in columns (3) and (4). The estimates for detailed arrests imply that a 10-percentage-point increase in graduation rates would reduce murder and assault arrest rates by about 20 percent, motor vehicle theft by about 13 percent, and arson by 8 percent.[35]

[34] 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, 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.

[35] High school graduation 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

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

Because arrest rates are not reported by race in addition to state, age, and year, it is difficult to determine whether schooling has differential effects on arrest by race. We attempt to examine this issue by controlling for both the schooling levels of blacks and whites in each state. To do this, we interact black (and white) educational attainment 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 education on arrests for each race. We find some evidence that the impact is greater for blacks.[36]

As a whole, these results suggest that schooling 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. Both IV and OLS estimates are similar, again suggesting that endogeneity problems are empirically unimportant.

Are these estimates consistent with the Census-based incarceration estimates of the previous section? As discussed in Section I, if sentence lengths or the probability of incarceration given arrest are greater for less educated individuals, the log difference in incarceration rates by education 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 percent in sentencing by graduation status, we should expect comparable effects of education on log arrest rates and log incarceration rates. The log difference in incarceration rates between high school dropouts and graduates for all men in the Census is about 1.4 (IV estimates produce larger impacts for blacks). The IV estimates in Table 10, obtained using data on all offenses, suggest that graduation reduces arrest rates among all men by nearly 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. (2002) 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 1964–1990, 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. (2002) produces an impact of 0.5–1.0. That is, a 10-percent increase in high school graduation rates should reduce arrest rates by 5–10 percent through increased wages alone. This covers the range of estimates in Tables 10 and 11 and confirms that an important explanation for the effect of high school graduation on crime resides in the higher wage rates associated with finishing high school.

IV. 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 education on crime. Those results suggest that schooling is associated with

costly. Noneconomic 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.

[36] For example, in a specification analogous to that of column (2) in the bottom panel of Table 10, the coefficient estimate for the interaction of black graduation rates with percent black and violent 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). When we also control for state-specific year effects as in column (3) of Table 10, the lack of race-specific arrest rates makes precise estimation of race-specific graduation impacts difficult.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

a lower probability of arrest and imprisonment. Because those estimates may confound the effects of schooling on actual crime with any educational differences in the probability of arrest or incarceration conditional on commission of a crime (see Section I), 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 underreporting, 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 AFQT that can be used as a measure of cognitive ability. Parents' age and education are available. 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 II.

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 shoplifting; (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 underreport their criminal behavior than young white men. (See for example the exhaustive study by Michael Hindelang et al., 1981.) Our calculations based on NLSY data suggest that black dropouts may be substantially underreporting criminal activity, while there is less reason to believe that black high school graduates and whites are underreporting to the same degree.[37] Because a correlation between underreporting and education would bias any estimates of the impact of schooling on crime, results for black self-reported crime should be treated with suspicion. Still, we present them along with results for whites for completion.

Table 12 reports the estimated effects of schooling on self-reported criminal participation and incarceration among young men in the NLSY using OLS. 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. 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 state of residence. Because the sample is so young and many of the men are still in school, we also control for school enrollment. As indicated by columns (1) and (3), both years of schooling and high school graduation significantly reduce participation in violent, property, and drug crimes among whites but not blacks. Due to the suspected underreporting of crime by black dropouts, the negligible effects of education are

[37] Among black dropouts, 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 underreporting, the incarceration data 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 percent of the black high school dropouts in the sample were incarcerated over ages 22–28, it seems highly unlikely that only 22 percent of young black dropouts participated in crime just a few years earlier. In the absence of gross incarceration of innocent black men, it is likely that black dropouts substantially underreported 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 educational attainment is likely to be less of a problem among whites. More accurate reporting among whites accords with previous studies (Hindelang et al., 1981).

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 12—THE EFFECT OF EDUCATION ON SELF-REPORTED CRIME AND INCARCERATION IN THE NLSY (IN PERCENTAGE TERMS)

	Years of school		High school graduate	
	(1)	(2)	(3)	(4)
WHITES				
Self-reported crime				
Violent crime	−1.87	−1.29	−8.89	−9.06
	(0.69)	(0.76)	(2.02)	(2.10)
Drug sales	−1.15	−0.99	−5.11	−5.02
	(0.44)	(0.48)	(1.28)	(1.33)
Property crime	−1.84	−1.38	−10.15	−11.21
	(0.98)	(1.07)	(2.86)	(2.94)
Any crime	−2.78	−2.21	−13.62	−14.71
	(1.08)	(1.18)	(3.14)	(3.25)
Incarcerated	−0.59	−0.62	−3.69	−3.47
	(0.06)	(0.08)	(0.30)	(0.34)
BLACKS				
Self-reported crime				
Violent crime	1.92	0.85	−0.40	−1.71
	(1.24)	(1.38)	(3.57)	(3.75)
Drug sales	−0.27	−0.58	−0.63	−0.79
	(0.56)	(0.60)	(1.57)	(1.62)
Property crime	−1.35	−2.91	−2.61	−4.43
	(1.30)	(1.43)	(3.70)	(3.90)
Any crime	2.02	0.46	2.38	0.07
	(1.52)	(1.68)	(4.33)	(4.53)
Incarcerated	−2.00	−1.74	−9.23	−7.94
	(0.23)	(0.28)	(0.98)	(1.04)
Controls:				
Age/cohort	y	y	y	y
Area of residence	y	y	y	y
Enrolled in school	y	y	y	y
Family background		y		y
Ability		y		y
SMSA status		y		y
Local unemployment rate		y		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. Each row represents a separate OLS regression. The dependent variables for the self-reported crimes are dummy variables equal to 1 if the person participated in that type of crime; for incarceration, it is a dummy equal to 1 if the individual was incarcerated at any time over ages 22–28. All coefficient estimates are multiplied by 100. The reported coefficients for incarceration are obtained by adjusting the ages 22–28 incarceration rates by the ratio of annual incarceration rates (over those ages) to incarceration rates over the full seven-year period (a factor of 0.3692 for whites and 0.4171 for blacks). 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, and whether his mother was a teenager at his birth. Area of residence refers to state dummies in columns (1) and (3) for self-reported crimes and to region-level dummies for all other specifications. Incarceration specifications do not control for current enrollment.

not surprising for black males. For white males, the estimates suggest that an additional year of school reduces participation in each type of crime by around 1–3 percentage points. High school graduation reduces white participation rates in violent crime by 9 percentage points, drug sales by 5 percentage points, property crime by 10 percentage points, and overall criminal participation by 14 percentage points.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

Columns (2) and (4) control for age, family background,[38] ability (as measured by AFQT percentile), race and ethnicity, geographic location [region of residence and an indicator for residence in a Standard Metropolitan Statistical Area (SMSA)], and local unemployment rates. The striking result is that these estimates obtained by conditioning on a rich set of individual and family background characteristics are quite similar to the parsimonious specifications used throughout the paper. In other words, ignoring cognitive ability and family background does not introduce a systematic 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 from Table 11 with the log difference in self-reported crime by high school graduation status in the NLSY. 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 graduation on arrests and incarceration are not simply the result of differential treatment by police and judges. Education has a real effect on crime that is measurably similar to its effects on both arrest and incarceration.[39] This reconciles with the finding of Mustard (2001) that average prison sentences are quite similar across high school graduates and dropouts.

We next examine the impact of education on incarceration in the NLSY to verify our earlier results using Census data. The estimated effects of schooling on incarceration during early adulthood are shown in the bottom row of Table 12. As in Section II, education significantly reduces the probability that a young man will be incarcerated. Estimates for both years of schooling and high school graduation are similar across the parsimonious and rich specifications, suggesting that an additional year of schooling reduces the annual probability of incarceration by about 0.6 percentage points for whites and 2 percentage points for blacks. High school graduation reduces the probability by 3–4 percentage points among white men ages 22–28 and 8–9 percentage points among black men over those ages.[40] While these estimated effects are larger than the average effects estimated with the Census data, the discrepancy is explained by the fact that the Census estimates report average incarceration effects over ages 20–60, while the NLSY-based estimates refer to men ages 22–28. Comparing the effects for 20-year-old men in the Census (see specification F of Table 8) with the NLSY results yields a remarkable consistency.

Two points are evident from the NLSY data. First, education significantly reduces self-reported crime among young white men, and the estimated effects are consistent with the impacts estimated for arrests and incarceration in Sections II and III. This implies that the impacts estimated for arrests and incarceration reflect a true effect on crime, and not simply educational differences in the probability of arrest or incarceration conditional on commission of a crime. (Due to suspected underreporting among black dropouts, it is impossible to say whether the same is true for black males.) Second, controlling for individual ability, family background, and local labor markets has little impact on the estimated effects.

V. Social Savings from Crime Reduction

Given the estimated impact of education on crime, it is possible to determine the social savings associated with increasing education levels. Because the social costs of crime differ substantially across crimes, we use estimates based on the impact of schooling on arrests by offense type to determine the social benefits of

[38] 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, and whether his mother was a teenager at his birth.

[39] 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.

[40] 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 annual incarceration rates (over those ages) to incarceration rates over the full seven-year period (a factor of 0.3692 for whites and 0.4171 for blacks).

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE 13—SOCIAL COSTS PER CRIME AND SOCIAL BENEFITS OF INCREASING HIGH SCHOOL COMPLETION RATES BY 1 PERCENT

	Victim costs per crime (1)	Property loss per crime (2)	Incarceration cost per crime (3)	Total cost per crime (4)	Estimated change in arrests (5)	Estimated change in crimes (6)	Social benefit (4) × (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 et al. (1996). Incarceration costs per crime equal the incarceration cost per inmate, $17,027 (U.S. Department of Justice, 1999), multiplied by the incarceration rate (U.S. Department of Justice, 1994). Total costs are calculated as the sum of victim costs and incarceration costs less 80 percent 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 panel B, column (4) of Table 11 and the total number of arrests in the 1990 Uniform Crime Reports. 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.

increased education. Recognizing that the effects of schooling tend to be more important during the high school years (particularly at the 12th-grade level) and due to the substantial policy interest in high school completion, we estimate the social benefits through reduced crime of increasing the high school graduation rate by 1 percent.

These estimates are subject to two important caveats. First, they assume that estimates in Table 11 produce a consistent estimate of the effect of graduation on arrest. Second, consistent with most other studies of crime, these estimates do not account for general-equilibrium effects on wages resulting from an increase in the supply of graduates. However, in Lochner and Moretti (Appendix B, 2001), 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 dropouts 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 Lochner and Moretti (Appendix B, 2001) suggests that the net effect of changing wages on crime is trivial. If anything, when 1 percent of the population is moved from dropout to graduate status, the reduction in wages among graduates is more than offset by the increase in wages among dropouts, so that the net effect on crime when general-equilibrium effects are included is no smaller than what is reported here.

Recognizing the limitations of the exercise, we nonetheless provide a rough estimate of the social savings from crime reduction resulting from a 1-percent 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 Ted 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

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

the incarceration rate for that crime (approximately $17,000).[41] Total costs are computed by summing incarceration costs and victim costs less 80 percent of property losses, which are already included in victim costs and may be considered a partial transfer to the criminal.[42] The table reveals substantial variation 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 percent of all prisoners in 1991 were incarcerated for drug offenses, costing more than $5 billion in jail and prison costs alone (U.S. Department of Justice, 1994). Given the NLSY findings for the effects of high school graduation on drug offenses, there is good reason to believe that these costs of crime are also relevant for this analysis.

Column (5) reports the predicted change in total arrests in the United States based on the arrest estimates reported in panel B, column (4) of Table 11 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 graduation rates had been 1 percentage point higher. 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 graduation 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 1-percent increase in male U.S. high school graduation 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 underestimate 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 the top panel of Table 11, 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 on whether estimates in the top or bottom panel of Table 11 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.[43] Therefore, the positive externality in crime reduction generated by an extra male high school graduate is between 14 percent and 26

[41] Incarceration rates by offense type are calculated as the total number of individuals in jail or prison (from U.S. Department of Justice, 1994) divided by the total number of offenses that year (where the number of offenses are adjusted for nonreporting to the police). Incarceration costs per inmate are taken from U.S. Department of Justice (1999). Offenses known to the police and reporting rates are given by the Uniform Crime Reports and National Criminal Victimization Survey.

[42] For the crime of arson, total costs equal victim costs plus incarceration costs, since it is assumed that none of the property loss is transferred to the criminal.

[43] 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.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

percent of the private return to high school graduation. The externalities from increasing high school graduation rates among black males are likely to be even greater given the larger estimated impacts on incarceration and arrest rates among blacks. On the other hand, the fact that women commit much less crime than men, on average, suggests that the education externality stemming from reduced crime is likely to be substantially smaller for them.

For another interesting comparison, consider what a 1-percent increase in male graduation 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 completing high school.[44]

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 forgone earnings). Of course, such a policy would also raise human capital and annual productivity levels of the new graduates by more than 40 percent 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.

VI. 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 averse, further reducing the propensity to commit crimes. To empirically explore the importance of the relationship between schooling and criminal participation, this paper uses three data sources: 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 incarceration from the National Longitudinal Survey of Youth.

All three of these data sources produce similar conclusions: schooling significantly reduces criminal activity. This finding is robust to different identification strategies and measures of criminal activity. The estimated effect of schooling on imprisonment is consistent with its estimated effect on both arrests and self-reported crime. Both OLS and IV estimates produce similar conclusions about the quantitative impact of schooling on incarceration and arrest. The estimated impacts on incarceration and self-reports are unchanged even when rich measures of individual ability and family background are controlled for using NLSY data. Finally, we draw similar conclusions using aggregated state-level UCR data as we do using individual-level data on incarceration and self-reported crime in the Census or NLSY.

Given the consistency of our findings, we conclude that the estimated effects of education on crime cannot be easily explained away by unobserved characteristics of criminals, unobserved state policies that affect both crime and schooling, or educational differences in the conditional probability of arrest and imprisonment given crime. 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 education on crime implies that there are benefits to education not taken into account by individuals themselves, so the *social return* to schooling is larger than the private return. The estimated social externalities from reduced crime are sizeable. A 1-percent increase in the high school completion rate of all men ages 20–60 would

[44] Because the arrest estimates reflect the average difference between all high school graduates and all dropouts (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 11 years by one additional year. However, as Figure 1 reveals, 70 percent of the reductions seem to be associated with finishing the final year of high school.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

save the United States as much as $1.4 billion per year in reduced costs from crime incurred by victims and society at large. Such externalities from education amount to $1,170–2,100 per additional high school graduate or 14–26 percent of the private return to schooling. It is difficult to imagine a better reason to develop policies that prevent high school drop out.

Appendix A: Comparison of Instrumental Variable and Reduced-Form Strength

Under fairly general conditions, our IV estimates of the effect of schooling on crime are likely to be more significant than are reduced-form estimates of the effect of compulsory schooling laws on crime. To see this, consider the following model:

$$y = x\beta + \varepsilon$$

$$x = d\gamma + u$$

where ε and u are independently and identically distributed (i.i.d.) errors which may be correlated. Let σ_ε^2 and σ_u^2 represent the variances of ε and u, respectively, and $\sigma_{u\varepsilon}$ their covariance. To keep things simple, consider the case with a single regressor, x, and a single instrument, d. Also, consider the reduced-form estimating equation:

$$y = (d\gamma + u)\beta + \varepsilon = d\alpha + v$$

where $\alpha = \gamma\beta$ and $v = u\beta + \varepsilon$.

The just-identified IV estimator is

$$\hat{\beta}_{IV} = (d'x)^{-1}d'y$$

and its estimated variance is

$$\hat{V}(\hat{\beta}_{IV}) = (d'x)^{-1}d'd(d'x)^{-1}\hat{\sigma}_\varepsilon^2,$$

where $\hat{\sigma}_\varepsilon^2 = (y - x\hat{\beta}_{IV})'(y - x\hat{\beta}_{IV})/N$. The t-statistic is given by

$$t_\beta = \frac{\hat{\beta}_{IV}}{\sqrt{\hat{V}(\hat{\beta}_{IV})}} = \frac{d'y}{(d'd)^{1/2}\hat{\sigma}_\varepsilon}.$$

Now, consider the reduced-form OLS estimator for α:

$$\hat{\alpha} = (d'd)^{-1}d'y$$

and its estimated variance,

$$\hat{V}(\hat{\alpha}) = (d'd)^{-1}\hat{\sigma}_v^2,$$

where $\hat{\sigma}_v^2 = (y - d\hat{\alpha})'(y - d\hat{\alpha})/N$. The corresponding t-statistic is given by

$$t_\alpha = \frac{\hat{\alpha}}{\sqrt{\hat{V}(\hat{\alpha})}} = \frac{d'y}{(d'd)^{1/2}\hat{\sigma}_v}.$$

Taking the ratios of t-statistics, we obtain

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

$$\frac{t_\beta}{t_\alpha} = \frac{\hat{\sigma}_v}{\hat{\sigma}_\varepsilon} \xrightarrow{p} \frac{\sigma_v}{\sigma_\varepsilon} = \frac{\sqrt{\beta^2\sigma_u^2 + 2\beta\sigma_{u\varepsilon} + \sigma_\varepsilon^2}}{\sigma_\varepsilon}.$$

So, as long as $\beta\sigma_{u\varepsilon} \geq 0$, we should generally expect a smaller t-statistic for the reduced-form estimate of α than the IV estimate of β.

Appendix B: Control Function Estimators

In order to discuss the linear control function estimator described in Garen (1984) and Card (1999), consider the following simplified version of our model:

$$y = \alpha + \beta s + u \tag{3}$$

$$s = \Pi Z + v, \tag{4}$$

where y represents incarceration, s represents schooling, and Z are instruments.

Assume $E(u|s, Z) = 0$ and $E(v|Z) = 0$. Garen (1984) further assumes that α and β may vary in the population such that

$$E(\alpha - \bar{\alpha}|Z) = 0,$$

$$E(\alpha - \bar{\alpha}|s, Z) = \theta_s s + \theta_z Z,$$

$$E(\beta - \bar{\beta}|Z) = 0,$$

$$E(\beta - \bar{\beta}|s, Z) = \phi_s s + \phi_z Z.$$

Together, these assumptions imply that $\theta_s\Pi = -\theta_z$ and $\phi_s\Pi = -\phi_z$.

Taking expectations of equation (3) conditional on (s, Z) we obtain

$$\begin{aligned} E(y|s, Z) &= \bar{\alpha} + \bar{\beta}s + [\theta_s s + \theta_z Z] + [\phi_s s + \phi_z Z]s, \\ &= \bar{\alpha} + \bar{\beta}s + [\theta_s(\Pi Z + v) + \theta_z Z] + [\phi_s(\Pi Z + v) + \phi_z Z]s, \\ &= \bar{\alpha} + \bar{\beta}s + \theta_s v + \phi_s vs. \end{aligned}$$

Estimating this equation using a consistently estimated $\hat{v}$ in place of v from a first-stage regression of equation (4) yields an estimate of the "average treatment effect" of s, or $\bar{\beta}$.

Since this method only requires mean independence of u conditional on (s, Z) rather than full statistical independence, it is not incompatible with a linear probability model or binary s.

Appendix C: Aggregating Census Data

This Appendix discusses estimation of the effects of average education on average incarceration rates using aggregated Census data. Specifically, we aggregate our Census sample to compute incarceration rates and average schooling levels by age, state of birth, and year. In aggregating by age, we use eight age groups (ages 20–24, 25–29, etc.), which correspond to those used in our arrest specifications. Using these aggregate observations, we estimate specifications analogous to those in Tables 4 and 7. Rather than using state of residence as a dummy regressor as in the individual-level specifications, we use the fraction of men in a particular age–state of birth–year cell residing in each state as regressors. The results are reported in Tables C1 and C2.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE C1—THE EFFECT OF COMPULSORY ATTENDANCE LAWS ON SCHOOLING (IN PERCENTAGE TERMS)—AGGREGATE SAMPLE

	Dropout (1)	High school (2)	Some college (3)	College+ (4)
WHITES				
Compulsory attendance = 9	−3.2	3.1	−0.0	−0.2
	(0.4)	(0.4)	(0.2)	(0.2)
Compulsory attendance = 10	−3.4	3.8	−0.0	−0.3
	(0.5)	(0.6)	(0.4)	(0.3)
Compulsory attendance ≥ 11	−4.9	5.6	−0.7	0.02
	(0.5)	(0.5)	(0.6)	(0.3)
F-test [*p*-value]	31.6	34.7	2.9	0.81
	[0.00]	[0.00]	[0.03]	[0.49]
BLACKS				
Compulsory attendance = 9	−2.2	2.9	−0.6	−0.1
	(0.5)	(0.4)	(0.2)	(0.2)
Compulsory attendance = 10	−1.6	3.6	−1.6	−0.4
	(0.6)	(0.6)	(0.4)	(0.2)
Compulsory attendance ≥ 11	−2.5	4.6	−1.8	0.3
	(0.6)	(0.6)	(0.3)	(0.3)
F-test [*p*-value]	8.85	22.3	9.9	1.4
	[0.00]	[0.00]	[0.00]	[0.21]

Notes: This table replicates the IV results of Table 4, except that models are estimated on aggregate data. The data have been aggregated at the state of birth, year, and age level. All coefficients are multiplied by 100. There are eight age groups (ages 20–24, 25–29, etc.). State of residence represents the fraction of men in a state of birth-year-age cell living in each state. Sample sizes are 6,273 for whites and 5,259 for blacks. All models are weighted by cell size. Standard errors corrected for state of birth–year of birth clustering are in parentheses. The *F*-test is for whether the coefficients on the excluded instruments are jointly equal to zero, conditional on all the controls.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

TABLE C2—IV ESTIMATES OF THE EFFECT OF YEARS OF SCHOOLING ON IMPRISONMENT (IN PERCENTAGE TERMS)—AGGREGATE SAMPLE

	(1)	(2)	(3)
WHITES			
Second stage			
Years of schooling	−0.08	−0.05	−0.10
	(0.06)	(0.07)	(0.09)
First stage			
Compulsory attendance = 9	0.297	0.230	0.191
	(0.032)	(0.029)	(0.031)
Compulsory attendance = 10	0.223	0.222	0.164
	(0.045)	(0.045)	(0.041)
Compulsory attendance ≥ 11	0.379	0.306	0.270
	(0.046)	(0.049)	(0.040)
F-test [*p*-value]	32.0	25.3	17.1
	[0.000]	[0.000]	[0.000]
BLACKS			
Second stage			
Years of schooling	−0.40	−0.23	−0.59
	(0.13)	(0.20)	(0.23)
First stage			
Compulsory attendance = 9	0.692	0.427	0.389
	(0.045)	(0.040)	(0.041)
Compulsory attendance = 10	0.595	0.437	0.388
	(0.085)	(0.077)	(0.079)
Compulsory attendance ≥ 11	0.694	0.437	0.404
	(0.072)	(0.063)	(0.064)
F-test [*p*-value]	65.9	39.7	31.1
	[0.000]	[0.000]	[0.000]
Additional controls:			
Cohort of birth effects		y	y
State of residence × year effects			y

Notes: This table replicates the IV results of Table 7, except that models are estimated on aggregate data. Second-stage coefficient estimates are multiplied by 100. The data have been aggregated at the state of birth, year, and age level. There are eight age groups (ages 20–24, 25–29, etc.). State of residence represents the fraction of men in a state of birth-year-age cell living in each state. Sample sizes are 6,273 for whites and 5,259 for blacks. All models are weighted by cell size. Standard errors corrected for state of birth–year of birth clustering are in parentheses. The *F*-test is for whether the coefficients on the excluded instruments are jointly equal to zero, conditional on all the controls.

REFERENCES

Acemoglu, Daron and Angrist, Joshua. "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws." in B. Bernanke and K. Rogoff, eds., *NBER macroeconomics annual*, Vol. 15. Cambridge, MA: MIT Press, 2000, pp. 9–59.

Arrow, Kenneth. "The Benefits of Education and the Formation of Preferences," in Jere Behrman and Nevzer Stacey, eds., *The social benefits of education*. Ann Arbor, MI: University of Michigan Press, 1997, pp. 11–16.

Becker, Gary S. and Mulligan, Casey B. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics*, August 1997, *112*(3), pp. 729–58.

Bekker, Paul. "Alternative Approximations to the Distributions of Instrumental Variable Estimators." *Econometrica*, May 1994, *62*(3), pp. 657–81.

Card, David. "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of labor economics*, Vol. 3A. Amsterdam: Elsevier Science, 1999, pp. 1801–63.

Chiricos, Theodore. "Rates of Crime and

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

Unemployment: An Analysis of Aggregate Research." *Social Problems*, May 1987, *34*(2), pp. 187–211.

Farrington, David; Gallagher, Bernard; Morley, Lynda; St. Ledger, Raymond and West, Donald. "Unemployment, School Leaving and Crime." *British Journal of Criminology*, Autumn 1986, *26*(4), pp. 335–56.

Freeman, Richard. "Crime and Unemployment," in J. Q. Wilson, ed., *Crime and public policy*. San Francisco, CA: ICS Press, 1983, pp. 89–106.

_____. "The Labor Market," in J. Q. Wilson and J. Petersilia, eds., *Crime*. San Francisco, CA: ICS Press, 1995, pp. 171–91.

_____. "Why Do So Many Young American Men Commit Crimes and What Might We Do about It?" *Journal of Economic Perspectives*, Winter 1996, *10*(1), pp. 25–42.

Garen, John. "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." *Econometrica*, September 1984, *52*(5), pp. 1199–218.

Gottfredson, Michael. "Youth Employment, Crime, and Schooling." *Developmental Psychology*, May 1985, *21*(3), pp. 419–32.

Gould, Eric D.; Weinberg, Bruce A. and Mustard, David B. "Crime Rates and Local Labor Market Opportunities in the United States: 1977–1997." *Review of Economics and Statistics*, February 2002, *84*(1), pp. 45–61.

Grogger, Jeffrey. "Market Wages and Youth Crime." *Journal of Labor Economics*, October 1998, *16*(4), pp. 756–91.

Hahn, Jinyong and Hausman, Jerry. "A New Specification Test for the Validity of Instrumental Variables." *Econometrica*, January 2002, *70*(1), pp. 163–89.

Heckman, James and Klenow, Peter. "Human Capital Policy." Working paper, University of Chicago, 1999.

Heckman, James J. and Robb, Richard J. "Alternative Methods for Evaluating the Impact of Interventions," in J. Heckman and B. Singer, eds., *Longitudinal analysis of labor market data*. Cambridge: Cambridge University Press, 1985, pp. 156–246.

Heckman, James and Vytlacil, Edward. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return Is Correlated with Schooling." *Journal of Human Resources*, Fall 1998, *33*(4), pp. 974–87.

Hindelang, Michael; Hirsch, Travis and Weis, Joseph. *Measuring delinquency*. Beverly Hills, CA: Sage, 1981.

Imbens, Guido W. and Angrist, Joshua D. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, March 1994, *62*(2), pp. 467–75.

Jacob, Brian A. and Lefgren, Lars. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime." *American Economic Review*, December 2003, *93*(5), pp. 1560–1577.

Kling, Jeffrey. "The Effect of Prison Sentence Length on the Subsequent Employment and Earnings of Criminal Defendants." Working paper, Princeton University, 2002.

Kotin, Lawrence and Aikman, William. *Legal foundations of compulsory school attendance*. Port Washington, NY: National University Publications, 1980.

Lang, Kevin and Kropp, David. "Human Capital versus Sorting: Evidence from Compulsory Attendance Laws." *Quarterly Journal of Economics*, August 1986, *101*(3), pp. 609–24.

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

_____. "Juvenile Crime and Punishment." *Journal of Political Economy*, December 1998, *106*(6), pp. 1156–85.

Lleras-Muney, Adriana. "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939." *Journal of Law and Economics*, October 2002, *45*(2), pp. 401–36.

Lochner, Lance. "Education, Work, and Crime: A Human Capital Approach." Working paper, 2003.

Lochner, Lance and Moretti, Enrico. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8605, 2001.

Machin, Stephen and Meghir, Costas. "Crime and Economic Incentives." Institute for Fiscal Studies Working Paper No. 00/17, 2000.

Miller, Ted; Cohen, Mark and Wiersema,

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms

Brian. "Victim Costs and Consequences: A New Look." Final Summary Report to the National Institute of Justice, February 1996.

Moretti, Enrico. "Estimating the Social Return to Higher Education: Evidence from Cross-Sectional and Longitudinal Data." *Journal of Econometrics*, 2004a (forthcoming).

_____. "Human Capital Externalities in Cities," in V. Henderson and J. F. Thisse, eds., *Handbook of regional and urban economics*, Vol. 4. Amsterdam: North-Holland, 2004b (forthcoming).

Mustard, David B. "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts." *Journal of Law and Economics*, April 2001, *44*(1), pp. 285–314.

Newey, Whitney K. "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables." *Journal of Econometrics*, November 1987, *36*(2), pp. 231–50.

Raphael, Steven and Winter-Ebmer, Rudolf. "Identifying the Effect of Unemployment on Crime." *Journal of Law and Economics*, April 2001, *44*(1), pp. 259–83.

Staiger, Douglas and Stock, James H. "Instrumental Variables Regressions with Weak Instruments." *Econometrica*, May 1997, *65*(3), pp. 557–86.

Stock, James and Yogo, Motohiro. "Testing for Weak Instruments in Linear IV Regression." Working paper, Harvard University, 2003.

Tauchen, Helen; Witte, Ann Dryden and Griesinger, Harriet. "Criminal Deterrence: Revisiting the Issue with a Birth Cohort." *Review of Economics and Statistics*, August 1994, *76*(3), pp. 399–412.

U.S. Department of Justice. *Profile of inmates in the United States and in England and Wales, 1991*. Washington, DC: U.S. Government Printing Office, 1994.

_____. *State prison expenditures, 1996*. Washington, DC: U.S. Government Printing Office, 1999.

Viscusi, W. K. "Market Incentives for Criminal Behavior," in R. Freeman and H. Holzer, eds., *The black youth employment crisis*. Chicago: University of Chicago Press, 1986, pp. 301–46.

Witte, Ann D. "Crime," in J. Behrman and N. Stacey, eds., *The social benefits of education*. Ann Arbor, MI: University of Michigan Press, 1997, pp. 219–46.

Witte, Ann D. and Tauchen, Helen. "Work and Crime: An Exploration Using Panel Data." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4794, 1994.

Wooldridge, Jeffrey M. "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model." *Economics Letters*, October 1997, *56*(2), pp. 129–33.

Yitzhaki, Shlomo. "On Using Linear Regressions in Welfare Economics." *Journal of Business and Economic Statistics*, October 1996, *14*(4), pp. 478–86.

This content downloaded from 130.91.144.125 on Thu, 11 Aug 2016 18:28:16 UTC
All use subject to http://about.jstor.org/terms