Non-Production Benefits of Education: Crime, Health, and Good Citizenship

Lance Lochner*
University of Western Ontario

December 15, 2010

1 Introduction

Economists have long recognized and measured the effect of education on an individual's own lifetime earnings.[1] More recently, attention has been paid to the effects of education on other personal and social outcomes, such as criminal behavior, health and mortality, and voting and democratic participation.[2] A growing body of work suggests that education offers a wide-range of benefits that extend beyond increases in labor market productivity. This chapter reviews recent developments on these 'non-production' benefits of education with an emphasis on contributions made by economists.

2 Education's Effects on Crime

In 1997, over two-thirds of all prison inmates in the U.S. were high school dropouts (Harlow 2003). Despite promising evidence that education-based policies and early childhood interventions can play an important role in helping reduce crime, evidence is still limited and sometimes mixed. The link between schooling and crime is more complicated than simple prison statistics suggest. This section reviews evidence in this rapidly growing area and develops a human capital-based theory for interpreting much of this evidence.

We first discuss the relationship between education and crime from an economic perspective, developing a simple lifecycle model that sheds light on key ways in which early childhood programs and policies that encourage schooling may affect both juvenile and adult crime. The model is grounded

---

[1]See, e.g., Card (1999) or Heckman, Lochner, and Todd (2006, 2008).

[2]Other studies attempt to estimate production externalities associated with education (see, e.g., Acemoglu and Angrist, 2001, Heckman and Klenow, 1997, Moretti, 2004a, 2004b, Ciccone and Peri, 2006). Lange and Topel (2006) offer a recent review of these studies.

in human capital theory and paints with a broad brush. It emphasizes the choice individuals face between legitimate work and criminal activity, with its associated punishments. By altering the relative rewards of work and crime, educational investments affect decisions to engage in crime. While the model does not incorporate all avenues through which education may affect crime, it serves as a useful point of reference.

We next discuss evidence on the impacts of educational attainment and school quality/choice on adult crime. The evidence from studies of educational attainment on crime is largely consistent with a human capital-based theory of crime, suggesting that increases in schooling reduce most types of adult crime (e.g. Lochner 2004, Lochner and Moretti 2004). Studies of school choice and increases in school quality paint a more nuanced picture: sizeable improvements in school quality produce minor (at best) improvements in student achievement and educational attainment, while they appear to substantially reduce crime during late adolescence and early adulthood (Cullen, Jacob, and Levitt 2006, Deming 2009a). We next discuss the contemporaneous relationship between school attendance and crime. Using exogenous policy changes and other events that effectively force students to stay in school or take extra days off (e.g. changes in compulsory schooling laws, teacher in-service days and strikes), a few recent studies have shown that school attendance affects crime in rich and complex ways.

We also review a number of recent studies that examine the long-run impacts of early childhood, school-based, and young adult training interventions on juvenile and adult crime. While a few early preschool programs have produced sizeable long-run reductions in crime – most famously, Perry Preschool – other quite similar programs have not. School-based programs focused on improving social development among 'risky' children have been shown to reduce crime through early adulthood. Finally, job training for young adults (e.g. Job Corps) appears to reduce self-reported arrests and convictions during the period of intensive training, but it yields negligible lasting effects on crime. Altogether, the evidence suggests that reductions in crime can be achieved by a wide range of human capital-based intervention strategies.

Finally, we discuss a number of policy issues related to education and its potential role as a crime-fighting strategy.

2.1 The Economics of Education and Crime

Why does education reduce crime and does its impact vary across different types of crimes? How might education and human capital policies help reduce crime? We begin with a simple economic model that formalizes a number of key channels through which education may affect crime. We then briefly discuss other factors that may help determine the relationship between education and crime.

2.1.1 A Dynamic Model of Schooling, Work, and Crime

We develop a dynamic time allocation model of crime, work, and education in which individuals are assumed to maximize lifetime earnings from work and crime. The model is based on Lochner (2004); however, it abstracts from random shocks to the returns from work and crime and simplifies certain lifecycle aspects of the problem to focus on the role of education. It emphasizes the role of education as a human capital investment that increases future legitimate work opportunities, which discourages participation in crime.[3] This is consistent with numerous recent studies that show that higher wages reduce crime (e.g. Grogger 1998, Machin and Meghir 2004, Gould, et al. 2002) and decades of research in labor economics showing that education increases wage rates (see, e.g., Card 1999).

We consider two stages of life: a 'schooling stage' followed by a 'work stage'. During the schooling stage (ages $t = 1, ..., t_s$), individuals decide how much time to allocate to work, crime, and human capital investment at school. During the 'work stage' ($t = t_s + 1, ...$), individuals only decide between work and crime. We assume that the 'work stage' lasts forever.

Time spent at work $L_t \geq 0$ returns income $w H_t L_t$ where w reflects the rental rate on human capital and H_t reflects an individual's legitimate market human capital. Time spent on crime $k_t \geq 0$ offers a reward of $N(k_t, H_t, \theta)$, where $N(\cdot)$ is strictly increasing and concave in k_t, strictly increasing in θ, and weakly increasing in H_t. The parameter θ reflects an individual's criminal 'ability', and it is assumed that $\frac{\partial^2 N}{\partial k \partial \theta} \geq 0$. While we use the term 'criminal ability', θ should be interpreted broadly, including anything which affects the rewards from crime (e.g. local neighborhood conditions, the availability of good 'targets'). Our assumptions allow for the fact that labor market skills may also be rewarded in the criminal sector, an issue we discuss further below. During the 'schooling stage', time and effort spent in school $I_t \geq 0$ offers no current rewards, but it increases future market skill levels according to:

$$H_{t+1} = H_t + A f(I_t, H_t), \tag{1}$$

where $f(\cdot)$ is increasing and concave in each of its arguments and $A > 0$ reflects an individual's 'learning ability'. Individuals are endowed with an initial skill level H_0. While human capital grows during the 'schooling stage' for those who spend some time acquiring skills through education (or training), it is assumed to remain constant during the 'work stage'. We normalize an individual's time endowment to one at all ages, so $L_t + k_t + I_t = 1$ during the 'schooling stage' and $L_t + k_t = 1$ during the 'work stage'.

Individuals who engage in crime at any age may be caught and imprisoned for J years with probability πk_t. We assume that the 'incarceration rate' $\pi \in (0, 1)$, so there is some probability that individuals escape punishment even if they spend all of their time committing crime. During prison, individuals earn nothing and their skills remain unchanged when they return to the labor market.[4]

---

[3]Fella and Gallipoli (2009) embed a similar lifecycle model of crime in a general equilibrium environment.

[4]Most estimates suggest that earnings and employment decline following an arrest or prison term; although, there

For simplicity, we further assume that youth imprisoned during the 'schooling stage' of life leave prison in the 'work stage' of life.

In setting up the model, it is convenient to begin with the 'work stage'. Since this stage lasts forever, we drop the t subscripts and formulate the problem in recursive form. For a time discount rate of $\beta \in (0,1)$, the value function for someone with human capital H who is in the 'work stage' but not currently in prison is

$$W(H) = \max_{k \in (0,1)} \left\{ wH(1-k) + N(k,H,\theta) + \beta \left[\pi k \beta^J + (1 - \pi k) \right] W(H) \right\}, \tag{2}$$

which reflects the fact that the value associated with imprisonment for J years is simply $\beta^J W(H)$, since individuals leave prison in the same state they entered. The interior first order condition for time spent committing crime is

$$wH + \pi \beta (1 - \beta^J) W(H) = \frac{\partial N(k,H,\theta)}{\partial k}, \tag{3}$$

where the left hand side reflects the marginal cost of spending more time on crime while the right hand side reflects the marginal return. Note that crime is costly in terms of current foregone earnings (wH) as well as expected losses due to time spent in prison (note that $\beta^J < 1$).

Before discussing the effects of human capital or criminal ability on criminal choices, it is worth discussing the marginal value of human capital, $W'(H)$. The envelope condition implies that

$$\begin{aligned} W'(H) &= w(1-k) + \frac{\partial N(k,H,\theta)}{\partial H} + \beta \left[\pi k \beta^J + (1 - \pi k) \right] W'(H) \\ &= \frac{w(1-k) + \frac{\partial N(k,H,\theta)}{\partial H}}{1 - \beta + \pi k \beta (1 - \beta^J)} \\ &\geq 0, \end{aligned} \tag{4}$$

where the final inequality is strict if $k < 1$ or $\frac{\partial N}{\partial H} > 0$. Since $\frac{w}{1-\beta}$ reflects the marginal value of human capital for someone who never commits any crime, we observe three ways in which crime affects the the marginal value of human capital during adulthood: (i) by reducing labor market work, it lowers the payoff from human capital by wk; (ii) the rewards from crime may be increasing in human capital ($\frac{\partial N}{\partial H} > 0$) which increases the marginal value of skills in the criminal sector; and (iii) the likelihood of arrest and imprisonment reduces the expected years human capital can be utilized in both the legitimate and criminal sectors. When human capital has negligible effects on criminal earnings (i.e. $\frac{\partial N}{\partial H} \approx 0$), engagement in crime will generally lower the marginal value of human capital due to (i) and (iii).

---

has been some debate about the magnitude and duration of any effects (e.g. see Grogger, 1995, Kling, 2003, Nagin and Waldfogel, 1995, and Waldfogel, 1994). This suggests that losses in earnings power due to stigma effects and/or human capital depreciation outweigh any gains from new skills that may be acquired in prison. We abstract from these features to focus on other issues.

Now, we can discuss the effects of human capital and criminal ability on crime. Equation (3) makes clear that crime is costly for individuals with higher levels of human capital. Totally differentiating this equation with respect to H yields

$$\frac{dk}{dH} = \frac{w - \frac{\partial^2 N}{\partial k \partial H} + \pi\beta(1-\beta^J)W'(H)}{\frac{\partial^2 N}{\partial k^2}}.$$

If human capital raises the marginal returns from work more than crime (i.e. $w \geq \frac{\partial^2 N}{\partial k \partial H}$), it is clear that human capital reduces crime since $W'(H) \geq 0$. Crime will be increasing in human capital (or schooling) only if human capital raises the marginal returns from crime sufficiently more than the returns from legitimate work to compensate for any expected losses from imprisonment associated with crime. From this, it is clear that policies that increase schooling (or the efficiency of schooling) will reduce most types of street crime among adults; although, certain types of white collar crime (e.g. embezzlement, fraud) may increase with education if they sufficiently reward skills learned in school.

One can also differentiate equation (3) with respect to θ to determine the effects of 'criminal ability' on the extent of criminal activity. A higher criminal ability raises the current payoff from crime, but it also raises the value of staying out of prison in the future, which discourages crime. As a result, the effects of θ on crime are generally ambiguous; however, if the probability of punishment is sufficiently low or the length of incarceration sufficiently short, the effect of θ on current criminal payoffs will dominate and more able criminals will commit more crime.[5] Given the relatively low probabilities of arrest and subsequent incarceration for most crimes, this is the more empirically relevant case. In the discussion below, we assume that crime is increasing in θ.

Now, consider the 'schooling stage'. At ages $t \leq t_s$, someone not currently in prison solves the following problem:

$$S_t(H_t) = \max_{k_t, I_t} \left\{ wH_t(1 - I_t - k_t) + N(k_t, H_t, \theta) + \beta\left[\pi k_t \beta^J W(H_{t+1}) + (1 - \pi k_t)S_{t+1}(H_{t+1})\right] \right\} \quad (5)$$

subject to the time constraints and human capital accumulation equation (1).[6] At an interior, the first order conditions for investment and crime are

$$wH_t = \beta[\pi k_t \beta^J W'(H_{t+1}) + (1 - \pi k_t)S'_{t+1}(H_{t+1})]A\frac{\partial f(I_t, H_t)}{\partial I_t} \quad (6)$$

$$wH_t = \frac{\partial N(k_t, H_t, \theta)}{\partial k_t} - \beta\pi[S_{t+1}(H_{t+1}) - \beta^J W(H_{t+1})]. \quad (7)$$

---

[5]Recognizing that $W(\cdot)$ depends on θ as well as H and totally differentiating (3) with respect to θ yields

$$\frac{dk}{d\theta} = \frac{\pi\beta(1-\beta^J)\frac{\partial W}{\partial \theta} - \frac{\partial^2 N}{\partial k \partial \theta}}{\frac{\partial^2 N}{\partial k^2}},$$

where $\frac{\partial W}{\partial \theta} = \frac{\partial N/\partial\theta}{1 - \beta + \pi k\beta(1-\beta^J)}$ and $\frac{\partial^2 N}{\partial k \partial \theta}$ are both positive.

[6]Notice $S_{t_s+1}(H) = W(H)$, since individuals enter the 'work stage' of life at that time.

As long as youth are engaged in all three activities (work, school, and crime), they will equate the marginal values of each.[7] Since the marginal return on work is fixed at any age (wH_t), individuals simply spend their time studying and committing crime up to the point where the net marginal returns on these activities equal their current period wage rate wH_t. Investment offers future returns, while the costs (foregone earnings) are paid up front. Crime is the opposite, providing current returns with costs paid up front (foregone labor market earnings) and in the future (expected imprisonment).[8] The marginal value of human capital during the schooling period is:

$$S'_t(H_t) = w(1-I_t-k_t) + \frac{\partial N(k_t, H_t, \theta)}{\partial H_t} + \beta \left[\pi k_t \beta^J W'(H_{t+1}) + (1-\pi k_t) S'_{t+1}(H_{t+1}) \right] \left[1 + A \frac{\partial f(I_t, H_t)}{\partial H_t} \right].$$

Notice that both current and future criminal activity affect the marginal value of human capital and incentives to invest in skills. Both current and future crime raise the probability of incarceration, which lowers the utilization of skills. Future crime also affects the marginal value of human capital, since crime and work will not generally reward human capital the same. For example, if $\frac{\partial N}{\partial H} = 0$, then future time spent on crime reduces time spent working and utilizing human capital in the same way that time spent in prison does.

On average, street criminals are quite young, have low IQ levels (Kandel, et al., 1988, and White, Moffitt, and Silva, 1989), and acquire little formal education. This suggests that the returns to traditional market skills are substantially lower for common street crimes (e.g. larceny, robbery, auto theft, drug dealing) than in the legitimate labor market. On the other hand, the returns to white collar crimes like forgery, fraud, and embezzlement would seem to depend heavily on market skill levels. It is, therefore, worth distinguishing between *unskilled crimes*, for which we assume $\frac{\partial N}{\partial H} = 0$ and *white collar crimes* for which we expect that $\frac{\partial N}{\partial H} > 0$ and $\frac{\partial^2 N}{\partial H \partial k} > 0$.

How do individual endowments (A, H_0, θ) influence decisions about investment and crime? Learning ability, A, largely influences crime through past investments and their effects on current skill levels. More able youth will generally invest more in their skills and accumulate more skills per unit of investment. As a result, they will commit fewer unskilled crimes at older ages than their less able counterparts. Because incarceration costs are increasing in A (due to the higher productivity of foregone learning opportunities) more able individuals will also commit less crime (than the less able) during their schooling years. For low incarceration probabilities, the latter effect is relatively unimportant and differences in crime by ability A will grow with age as human capital profiles diverge. Because differences in skill levels will tend to persist over time for a cohort, individuals who begin life with more human capital (H_0) will tend to commit fewer unskilled crime at all ages. By contrast, individuals with higher criminal returns (θ) are more likely to engage in crime at any age.

---

[7]Youth who do not work at all will equate the marginal return to crime with the marginal return to schooling investment (i.e. they equate the right hand side of equation (6) with the right hand side of (7)).

[8]It is clear that $S_t(H) \geq W(H)$, since the latter is the value associated with a constrained problem ($I = 0$) that is otherwise the same as the former.

Furthermore, criminal ability (in unskilled crimes) also has a reinforcing indirect effect on crime at older ages through its effect on early human capital investment decisions — by directly encouraging crime over work, θ indirectly reduces the returns on skill investments causing high θ individuals to accumulate less human capital. Altogether, factors that reflect high A and H_0 or low θ should be negatively correlated with unskilled crime. All of these relationships are likely to be muted (or even perverse) for more skill-intensive white collar crimes. A high A and θ are also likely to be associated with greater schooling investments.

Now, consider the relationship between human capital and crime. This relationship is most easily analyzed for unskilled crimes, which offer no reward for human capital. Conditional on ability and other permanent characteristics, adults with high current skill levels face low opportunity costs of crime and should, on average, commit less crime. As such, past investments and education should be negatively correlated with current unskilled crime rates conditional on ability and initial skill levels. Of course investments and education are endogenous. Individuals that expect to commit a lot of unskilled crime in the future regardless of their current investments in skill (e.g. individuals with a high θ) have little incentive to make any investment. Thus, variation in criminal opportunities can also generate a negative correlation between educational attainment and unskilled crime.

Since a higher A and a lower θ cause individuals to invest more and commit less unskilled crime, their joint distribution is important for determining the correlation between educational attainment and unskilled crime in the population. Holding either of these endowments constant in the population, we should expect a negative correlation between schooling and unskilled crime.[9] A positive correlation between education and unskilled crime is, therefore, only likely to arise if A and θ are strongly positively correlated, in which case the best criminals are also the brightest. Given the lower than average IQ levels of most criminals (Kandel, et al., 1988, and White, Moffitt, and Silva, 1989), this seems unlikely.

The correlation between white collar crime and education will typically be less negative (and may even be positive). This is because a high A leads to greater investment, but that investment need not lead to less white collar crime. And, while a high θ encourages white collar crime, it need not cause individuals to invest less in their skills if those skills pay off in the criminal sector. Thus, two important forces generating a negative correlation between unskilled crime and educational attainment are less powerful for white collar crime.[10]

While individual endowments are important determinants of investment, work, and crime decisions, they are not the only factors affecting the education-crime relationship. Although not modeled here (see Lochner (2004)), shocks to wages and criminal returns can influence both current and future

---

[9]While H_0 should be negatively correlated with unskilled crime, it is more neutral with respect to investment since human capital increases both the returns and costs of investing.

[10]Additionally, punishments for white collar crimes tend to be skewed more heavily toward fines relative to incarceration. This generates similar disincentives to engage in white collar crime for all education backgrounds, since fines are largely independent of the offender's labor market skills. In fact, the most educated may be deterred the least by fines due to diminishing marginal utility of wealth. By contrast, incarceration is more costly for the most educated.

decisions. A strong positive wage or crime shock during the teenage years may cause some individuals to drop out of school, which should then affect subsequent decisions about crime. Absent the accumulation of criminal skills or habit persistence, these exogenous shocks only affect subsequent crime through differences in accumulated human capital levels. Variation in the costs of or tastes for schooling may also affect the education-crime relationship among adults through accumulated skill levels. Finally, policies and government interventions may affect schooling decisions. Policies that encourage schooling investments should reduce crime rates among youth as they substitute time from crime to school. Furthermore, increases in human capital acquired through additional schooling should reduce subsequent crime rates as well. The empirical literature has primarily focused on the latter; however, a few studies have attempted to examine the contemporaneous relationship between education policies and crime. We review some of these studies below.

2.1.2 Other Ways in Which Education May Affect Crime

Education may also teach individuals to be more patient (Becker and Mulligan 1997). This would discourage crime, since forward-looking individuals place greater weight on any expected future punishment associated with their criminal activities. To the extent that time preferences are affected by schooling, crimes associated with long prison sentences (or other long-term consequences) should be most affected. Education may also affect preferences toward risk. If schooling makes individuals more risk averse, it should discourage crime with its greatest effects on offenses that entail considerable uncertainty in returns or punishment. Finally, schooling may affect the set of people individuals interact with on a daily basis in school, work, or their neighborhoods. Assuming more educated people interact more with other educated people who are less inclined to engage in crime, this is likely to compound any reductions in crime associated with schooling.[11] In most cases, mechanisms related to changes in preferences or social interactions suggest that educational attainment is likely to reduce most types of crime among adults.

2.1.3 School Attendance and Contemporaneous Crime

It is useful to distinguish between the effects of educational attainment on subsequent criminal activity, and the way in which school attendance itself affects contemporaneous crime. The latter relationship is likely to be driven by three mechanisms (the first two being central to the time allocation problem above). First, school may have an incapacitation effect – youth cannot be in two places at once and many criminal opportunities are more limited in school than on the streets. Of course, school does not last all day, so this effect depends, in part, on the ease with which youth can engage in crime during non-school hours. Second, longer periods of school attendance should increase human capital levels and improve future employment prospects. This, in turn, may make juvenile arrests and long periods

---

[11]See Glaeser, Sacerdote, and Scheinkman (1996) for a model of crime where social interactions are important.

of detention more costly, reducing incentives to engage in crime while enrolled in school. Third, schools bring hundreds of adolescents together for the day and then free them all at the same time. The social interaction effects from doing this are far from obvious, but it is quite possible that this leads to altercations and more general group-based delinquency. The incapacitation and human capital effects are likely to to imply negative effects of school attendance on crime, while the social interaction effect could be positive or negative.

2.2 Evidence on Education and Crime

We now discuss evidence on the effects of educational attainment and school quality and choice on subsequent criminal outcomes. We also review empirical studies that analyze the relationship between school attendance and contemporaneous crime.

2.2.1 Educational Attainment and Crime

We have discussed four primary reasons schooling might affect subsequent crime: (i) education raises wage rates, which raises the opportunity costs of crime; (ii) education may directly affect the financial or 'psychic' rewards from crime; (iii) education may alter preferences for risk-taking or patience; and (iv) schooling may affect the social networks or peers of individuals. For most crimes (except, possibly, white collar crimes), one would expect these forces to induce a negative effect of schooling on adult crime.

Empirically, there is a strong negative correlation between educational attainment and official measures of crime. In 1997, 75% of state and 59% of federal prison inmates in the U.S. did not have a high school diploma (Harlow 2003).[12] After controlling for age, state of birth, state of residence, year of birth and year effects, Lochner and Moretti (2004) still find significant effects of schooling (especially high school completion) on the probability of incarceration in the U.S. as reported in Figure 1.[13] In 2001, more than 75% of convicted persons in Italy had not completed high school (Buonanno and Leonida 2009), while incarceration rates among men ages 21-25 in the United Kingdom were more than 8 times higher for those without an education qualification (i.e. dropouts) relative to those with a qualification Machin, Marie, and Vujic (forthcoming).

Differences by education are also apparent in self-reported survey measures of crime. For example, in the 1980 wave of the National Longitudinal Survey of Youth (NLSY), 34% of American men ages 20-23 with 11 or 12 years of completed schooling self-reported earning some income from crime, compared

---

[12]These figures exclude those who received a General Educational Development (GED) diploma. As shown in Cameron and Heckman (2003) and Heckman and LaFontaine (2006), individuals with a GED perform like high school dropouts rather than graduates in the labor market. Roughly 35% of state inmates and 33% of federal inmates completed their GED with more than two-thirds of these inmates earning their GED while incarcerated. A small percentage of those who did not receive a high school diploma had participated in some vocational or post-secondary courses. See Harlow (2003).

[13]These figures report the coefficients on indicators for different years of completed schooling from the 1960, 1970, and 1980 Censuses for white and black men ages 20-60.

with 24% of those with 12 years of school, and only 17% of those with more than 12 years. The effect of education is magnified if we consider more active criminal engagement: 4.2% of 20-23 year-old NLSY men completing 10 or 11 years of school reported earning more than half their income from crime, compared with 1.4% of those with 12 years and 0.7% of those with at least some college education. Similar patterns are observed for violent crime in the NLSY. See Lochner (2004) for further details.

Western (2006) shows that while U.S. differences in incarceration rates by education existed in the 1970s, they have grown considerably since. Among black men born 1945-49, roughly 4% of those who dropped out of high school would spend some time in prison by age 35, while only 1% of those with a high school degree or GED faced incarceration. By contrast, black men born 1965-69 faced a 59% incarceration rate by age 35 if they had not completed high school compared to an 18% rate among those who had. Trends in the incarceration differential by education among white men are qualitatively similar. As we discuss below, these trends are predictable responses to the rapid growth in the probability of imprisonment and prison sentence lengths, as well as the rising wage returns to schooling over the 1980s and 1990s.

Early studies of the relationship between education and crime focused on their correlation conditional on measured individual and family characteristics using standard regression methods.[14] For example, Witte and Tauchen (1994) find no significant relationship between educational attainment and crime after controlling for a number of individual characteristics. Grogger (1998) estimates a significant negative effect of wages on crime, but he finds no relationship between years of completed schooling and crime after controlling for individual wage rates. Of course, increased wages and earnings are important consequences of schooling. Thus, this study suggests that education may indirectly reduce crime through increased wage rates.

These earlier studies must be interpreted with caution. A negative cross-sectional correlation between education and crime, even after controlling for measured family background and neighborhood characteristics, does not necessarily imply that education reduces crime. Standard regression studies are unlikely to estimate the causal effect of eduction on crime (i.e. the effect of increasing someone's schooling on his criminal activity) for a number of reasons. First, unobserved individual characteristics like patience or risk aversion are likely to directly affect both schooling and criminal decisions. Individuals who choose more schooling (even after conditioning on observable characteristics) might also choose less crime regardless of their education level, in which case regression-based estimates do not identify a causal effect. Second, using variation in crime and education across states or local communities may also produce biased estimates. Governments may face a choice between funding police or good public schools, which would tend to produce a spurious positive correlation between education and crime. Alternatively, unobserved characteristics about communities or their residents

---

[14]Ehrlich (1975) provides an early empirical exploration of predicted effects of education on crime from a human capital perspective. See Witte (1997) for a survey of the early empirical literature on education and crime.

may directly affect the costs or benefits of both education and crime. For example, communities with few job opportunities that reward schooling may also be faced with severe gang problems. While it is often possible to account for permanent unobserved differences across communities by examining the relationship between changes in schooling and crime over time, this approach does not account for the effects of changing unobserved community characteristics. Third, reverse causality is another important concern, for reasons discussed earlier. Individuals who plan to heavily engage in crime (e.g. because they are particularly good at it, enjoy it, or live in areas with plenty of illicit opportunities) are likely to choose to leave school at a young age. Arrests or incarceration associated with juvenile crime may also cause some youth to drop out of school early (Hjalmarsson 2008).

Recently, economists have attempted to address these difficult issues and to estimate the causal effects of schooling on crime using instrumental variable (IV) methods. In the context of estimating the effect of educational attainment on crime, an instrument is valid if it induces variation in schooling but is uncorrelated with other factors that directly affect criminal proclivity (e.g. individual preferences or abilities, local law enforcement). Intuitively, this approach exploits differences in educational attainment across individuals that arise in response to factors that have no direct impact on criminal decisions. An ideal instrument would randomly assign some youth to drop out of high school and others to finish. Then, comparing the differences in crime rates across these groups would identify the causal effect of high school completion on crime. In practice, we typically do not observe such perfect experiments, but researchers can sometimes come close.

Because crime itself is difficult to measure, researchers are often forced to use measures of arrest or incarceration rather than actual crimes committed. It is possible that education reduces the probability of arrest and incarceration or the sentence lengths administered by judges, in which case estimates based on measures of arrest or incarceration incorporate these effects in addition to any effects of education on actual crime. While there is little direct evidence on these issues, Mustard (2001) finds negligible effects of defendant education levels on the sentence lengths they receive. Furthermore, results using self-reported measures of crime in the National Longitudinal Survey of Youth (NLSY) support the case that education reduces actual violent and property crime and not just the probability of arrest or incarceration conditional on crime (Lochner 2004, Lochner and Moretti 2004).[15]

Many recent empirical studies analyze crime aggregated at some geographic level, exploring the effects of average educational attainment on crime, arrest, conviction, or incarceration rates. In order to address concerns with endogeneity or unobserved heterogeneity, researchers have typically turned

---

[15] There has been considerable debate among criminologists on the merits of self-reported measures of crime vs. official measures of arrest. Most studies find a reasonably high correlation between the two; however, it is generally agreed that the two measures offer distinct and complementary information about criminal activity. Comparisons of self-reported arrests vs. official arrests tend to find a stronger correlation, with agreement increasing further for self-reported vs. official measures of criminal convictions. A number of studies report greater under-reporting of crimes and arrests by blacks; however, studies vary considerably on this. See the classic Hindelang, Hirschi, and Weis (1981) for comprehensive treatment of the issue or Thornberry and Krohn (2000) for a more recent survey of this literature.

to instrumental variables estimation or a differences-in-differences strategy using changes in state or national rules that affect schooling decisions. An aggregate-level regression is often specified as follows:

$$y_{calt} = E_{alt}\beta + X_{alt}\gamma + d_{lt} + d_{cl} + d_{al} + d_{ct} + d_{at} + d_{ca} + \varepsilon_{calt} \tag{8}$$

where y_{calt} is a measure of the crime, arrest, or incarceration rate for some offense type c, age group a, in location l, in year t. In some cases, only a single measure of crime is used (e.g. incarceration or total arrests), in which case the c subscript is extraneous. E_{alt} is an aggregate measure of educational attainment for age group a in location l at time t (e.g. average schooling attainment or high school completion rates). X_{alt} is a set of observable characteristics that may vary across age, location, and time (e.g. racial composition of an area). The d's represent indicator variables that account for unobserved differences by age/cohort, location, year, and criminal offense types. The term d_{lt} allows for location-specific time effects, which accounts for time varying unobserved location-specific differences that may reflect differences in local public spending, economic conditions, or law enforcement. The inclusion of d_{cl} allows the average distribution of crime or arrest types to differ across areas. For example, some states may focus arrests more heavily on one type of crime, while others focus on other types. Or, some areas may be more amenable to certain crimes while others are not. Similarly, the age distribution of crime or arrests need not be the same across areas – some age groups may be more crime-prone in some areas or arrest policies with respect to age may differ across areas. The term d_{al} absorbs long-run differences in age-arrest patterns across locations. Crime-specific and age-specific time trends in arrest common to all areas are accounted for by d_{ct} and d_{at}, respectively. Finally, d_{ca} accounts for long-term differences in age-crime profiles across different types of criminal offenses. Given these fixed effects, identification of the effect of education on crime is achieved through time variation in cohort educational attainment levels across different locations. The absence of d_{alt} indicator variables in equation (8) is, therefore, central to identification.

Lochner and Moretti (2004) examine state-level male arrest rates by criminal offense and age (five-year age categories beginning at ages 20-24 through 55-59) from the FBI's Uniform Crime Reports (UCR) for the U.S. in 1960, 1970, 1980, and 1990. This data is linked to 1960-90 decennial U.S. Census data on educational attainment and race to estimate equation (8), where y_{calt} represents log arrest rates for a specific offense, age category, state, and Census year. They specifically analyze arrest rates for murder, rape, assault, robbery, burglary, larceny, auto theft, and arson. In using log arrest rates, the effect of education is assumed to be the same in percentage terms for each type of crime included in the regression. They explore the effects of both average years of schooling and high school completion rates at the cohort-level (cohorts are defined by year of birth given year t and age a) in state l as of time t (i.e. E_{alt}). In addition to including all the d fixed effects in equation (8), they also include the percent of males that are black in age group a living in state l at time t.

The main methodological contribution of Lochner and Moretti (2004) is the use of changes in state-

specific compulsory schooling laws over time as instrumental variables for schooling.[16] Intuitively, this strategy measures the extent to which an increase in a state's compulsory schooling age leads to an immediate increase in educational attainment and reduction in subsequent crime rates for affected cohorts. Because the laws only affect schooling at low levels (mainly grades 8-12), their instrumental variable (IV) estimates reflect the impact of an additional year of high school on crime.[17] The impacts of schooling at other margins may differ as discussed below.

Lochner and Moretti (2004) estimate equation (8) using both ordinary least squares (OLS) and IV estimation. Using OLS, they find that a one-year increase in average education levels in a state reduces state-level arrest rates by 11 percent. IV estimates suggest slightly larger effects, although they are not statistically different. These estimated effects are very similar to the predicted effects derived from multiplying the estimated increase in wages associated with an additional year of school by the estimated effects of higher wage rates on crime (from Gould, et al. 2002). This suggests that much of the effect of schooling on crime may come through increased wage rates and opportunity costs as emphasized in the model above. Given the strong relationship between high school completion and incarceration apparent in Figure 1, Lochner and Moretti (2004) also estimate specifications using the high school completion rate as a measure of schooling. OLS estimates suggest that a ten percentage point increase in high school graduation rates would reduce arrest rates by 7%, while IV estimates suggest a slightly larger impact of 9%.

Lochner and Moretti (2004) also estimate separate effects of education for different types of crime using OLS (including interactions of criminal offense type with education in equation (8)). These results suggest similar effects across the broad categories of violent (murder, rape, robbery, and assault) and property (burglary, larceny, motor vehicle theft, and arson) crime — a one year increase in average years of schooling reduces both property and violent crime by about 11-12%. However, 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 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. Additional specifications suggest quantitatively similar effects for a 10-20 percentage point increase in high school graduation rates. Their results for rape are surprising and not easily explained by standard economic models of crime.[18]

---

[16]The relevant compulsory schooling age is based on the state law that applied when a cohort was age 14. Lochner and Moretti's (2004) analysis suggests that changes in compulsory schooling laws are exogenous and not related to prior trends in schooling or state expenditures on law enforcement, so it appears to be a valid instrument. Other studies reach similar conclusions about the exogenous nature of changes in compulsory schooling laws in other contexts (e.g. Acemoglu and Angrist 2001, Lleras-Muney 2002).

[17]More specifically, they reflect the impact of an additional year of high school on crime among those youth that are affected by changes in the schooling age laws. See Angrist and Imbens (1994, 1995) for the interpretation of IV estimators when using discrete instrumental variables.

[18]However, the results are consistent with some specifications in Gould, et al. (2002), which suggests that local wage rates are positively correlated with local crime rates for rape.

Lochner (2004) follows a similar approach using the same UCR data from 1960 to 1980; however, he also examines white collar crime. OLS estimation of equation (8) produces *positive*, though statistically insignificant, effects of schooling on arrest rates for white collar crimes (forgery and counterfeiting, fraud, and embezzlement). Estimates for violent and property crime are negative and similar to those of Lochner and Moretti (2004).

Lochner and Moretti (2004) also use individual-level data on incarceration and schooling from the 1960, 1970, and 1980 U.S. Censuses to estimate the effects of educational attainment on the probability of imprisonment separately for black and white men (ages 20-60). Their estimates control for age of the respondent (three-year age categories), state of birth, state of residence, cohort of birth, and state-specific year effects. Most importantly, controlling for state-specific year effects allows for the possibility that different states may have different time trends for law enforcement policies or may simply exhibit different trends in aggregate criminal activity. Analogous to their analysis of state-level arrest rates, they use state-level changes in compulsory schooling ages as an instrument for educational attainment. Although this analysis uses individual-based measures of incarceration and schooling, variation in schooling laws at the state-year level effectively identifies the effect of education on crime. As with the estimates for aggregate arrest rates, identification comes from the fact that in any given state and year, different age cohorts faced different compulsory schooling laws during their high school years, causing them to acquire different levels of schooling and to commit crime at different rates. Again, both OLS and IV estimates are very similar and suggest that, on average, an extra year of education reduces the probability of imprisonment by slightly more than .1 percentage point for whites and by about .4 percentage points for blacks. In their sample, the probability of incarceration for male whites (blacks) without a high school degree averaged .83% (3.6%), which translates into a 10-15% reduction in incarceration rates for both white and black males associated with an extra year of completed schooling. These estimated effects are comparable to those for arrest rates described earlier. OLS results suggest that completion of the twelfth grade causes the greatest drop in incarceration, while their is little effect of schooling beyond high school (see Figure 1).

Oreopolous and Salvanes (2009) reproduce the Lochner and Moretti (2004) IV results for black males using the same estimation strategy with a slightly different specification and an expanded sample that includes men ages 25-64 from the 1950-80 U.S. Censuses.[19] Their estimate suggests that an additional year of completed schooling reduces incarceration rates among black men by about 20%.

Machin, Marie, and Vujic (forthcoming) exploit a 1972-73 increase in the minimum schooling age (from age 15 to 16) in England and Wales to estimate the effects of schooling on criminal convictions for property and violent crimes over the period 1972-96. Using both IV and regression discontinuity methods, identification effectively comes from cohort-level changes in schooling attainment and crime

---

[19]Most notably, they do not include state and state-specific year effects in their specification. They also remove individuals with greater than twelve years of schooling from their sample. Their measures of compulsory schooling ages differ as well, incorporating the fact that some states allow for exceptions to the dropout age under certain conditions.

for cohorts turning 15 immediately before and after the law change.[20] Among men, they estimate that a one-year increase in average schooling levels reduces conviction rates for property crime by 20-30% and violent crime by roughly one-third to one-half as much.[21] Compared to estimates for the U.S. by Lochner and Moretti (2004), the impacts of education on property crime appear to be greater in the United Kingdom, while the effects on violent crime are weaker.

Buonanno and Leonida (2009) estimate the effects of educational attainment on crime rates in Italy using regional panel data from 1980-95. Their unit of observation is a region-year (they examine 20 Italian regions), and they estimate a restricted form of equation (8) using OLS. Specifically, they control for region and time fixed effects (d_l and d_t), along with region-specific quadratic time trends (assuming $d_{lt} = \delta_{1l}t + \delta_{2l}t^2$), and a rich set of time-varying region-specific covariates.[22] These estimates are identified from the relationship between changes in regional education levels and crime rates (around smooth regional time trends). Their estimates suggest that a ten percentage point increase in high school graduation rates would reduce property crime rates by 4% and total crime rates by about 3%. (Effects on property crime are statistically significant, while effects on total crime are not.) They find no evidence to suggest that university completion reduces crime.[23]

Merlo and Wolpin (2009) take a very different approach to estimating the relationship between schooling and subsequent crime. Using individual-level panel data on black males ages 13-22 from the NLSY, they estimate a discrete choice vector autoregression model in which individuals can choose to engage in crime, attend school, and/or work each year.[24] These decisions are allowed to depend on unobserved individual-specific returns to each activity, as well as crime, schooling, and work choices the previous year. Using estimates for their model, Merlo and Wolpin simulate the effects of changing youth schooling status at age 16 on subsequent outcomes. Their estimates suggest that, on average, attending school at age 16 reduces the probability a black male ever commits a crime over ages 19-22 by 13 percentage points and the probability of an arrest over those ages by 6 percentage points. These represent 42% and 23% reductions in self-reported crime and arrest rates, respectively, for black males not in school at age 16.

A final study worth mentioning examines the effects of an explicit education subsidy on youth

---

[20]They estimate models aggregated to the year-age level for individuals ages 18-40 from 1972-96. To alleviate concerns that other important economic or social factors may have changed at the same time the compulsory schooling age increased, they include a rich set of covariates: year and age indicators, fraction British-born, fraction employed, fraction non-white, and fraction living in London.

[21]Estimated effects on male property crime are statistically significant, while effects on male violent crime are not. Estimated effects for women are, unfortunately, very imprecise.

[22]Covariates include employment rates, GDP per capita, GDP growth rates, average wage rates, the fraction of crimes without an arrest, police per capita, and the length of time in the judicial process.

[23]Buonanno and Leonida (2009) also generalize their econometric specification to allow for an effect of lagged crime rates on current crime rates, estimating this using a generalized method of moments estimator to account for the endogeneity of lagged crime rates. This specification produces similar estimated effects of schooling on crime.

[24]Crime, work, and school are not mutually exclusive activities in this framework – individuals can do any combination of these three activities in each year. An individual is said to have engaged in crime in any year if he self-reported any of the following offenses: theft, other property crime, sold drugs, or assault.

burglary rates in England. Between 1999 and 2002, England piloted Educational Maintenance Allowances (EMA), which provided subsidies of up to £40 per week (plus bonuses for completion of coursework) for low-income 16-18 year old youth to attend school. The program was administered in 15 local areas with low schooling participation rates. During the same time period, the Reducing Burglary Initiative (RBI) funded 63 different local burglary reduction schemes as a separate pilot project. Roughly half of all EMA pilot areas were also selected for the RBI. Sabates and Feinstein (2007) use a differences-in-differences strategy to identify the effects of each pilot program as well as the combination of the two on burglary. Specifically, they compare changes in burglary conviction rates before and after the introduction of RBI, EMA, or both against a set of comparison areas. While baseline burglary conviction rates were much higher in EMA and EMA-RBI combined areas relative to the comparison areas, annual growth rates in burglary conviction rates prior to the programs were quite similar across all three classifications. To reduce concerns about differences between the treated and untreated areas, Sabates and Feinstein control for a number of time-varying area-specific factors likely to affect crime and limit their sample of comparison areas to those that best 'match' the distribution of demographic characteristics in the pilot areas.[25] Their findings suggest that the combination of both the EMA and RBI significantly reduced burglary rates by 1.3 per 1,000 youth (about 5.5%) relative to the 'matched' comparison areas. Effects of the EMA alone were slightly lower but still significant. While there are obvious concerns about the extent to which time-varying determinants of burglary are the same for treated and comparison areas, Sabates and Feinstein (2007) show that estimated effects on burglary rates for 19-21 year olds (who were not offered the education subsidy) were much lower and statistically insignificant.

2.2.2 School Quality and Crime

If human capital acquisition, socialization, or preference modification are important mechanisms determining the impacts of educational attainment on crime, then it is likely that school quality and the type of schools students attend also affect criminal behavior. While there are no studies which directly estimate the effects of measured school quality on crime, three recent studies on school choice and desegregation provide some useful insights.

Cullen, Jacob, and Levitt (2006) and Deming (2009a) examine the importance of school choice in large urban U.S. school districts (Chicago and Charlotte-Mecklenburg, respectively) on a variety of student outcomes, including delinquency and crime. Both studies examine the effects 'winning' a randomized lottery for admission to schools children selectively apply to.[26] By comparing the outcomes

---

[25]Their regressions control for unemployment rates for individuals under 25, proportion of students eligible for free school meals, number of qualified teachers, pupil-teacher ratios, and the number of supplementary staff for ethnic minorities, percent of youth with no schooling qualifications as of age 16 (i.e. dropouts), and the percent of unauthorized half-days missed in secondary school. We discuss results based on the 'matched' sample of comparison areas.

[26]In both cases, students could always choose to attend their neighborhood school. If any additional positions were available in a school, an open enrollment lottery was run based on all other students who applied to that school/program.

for youth who win vs. lose a particular school admission lottery, they estimate the effects of being offered admission to that school relative to the preferred alternative. This reflects the 'intention to treat' (i.e. the effects of being offered the opportunity to attend better schools) and not necessarily the effects of actually attending that school, since many students did not ultimately enroll in schools for which they were admitted by lottery. However, both studies find that 'winning' a lottery does significantly increase enrollment in that school. Since many students applying outside their assigned local school are from disadvantaged backgrounds and neighborhoods, on average, lottery winners end up attending better quality schools, as measured by such things as student achievement scores, value added (i.e. growth in achievement), student behavioral problems, or teacher quality. In this sense, these studies offer an opportunity to examine the effects of school quality, broadly defined, on delinquency and crime.

Cullen, Jacob, and Levitt (2006) find that winning a high school lottery in Chicago significantly raises peer graduation rates by 6% and the share of peers who test above national norms by about 14%; however, lottery winners appear to be placed in lower tracked classes within the better schools. Interestingly, they find no evidence that lottery winners perform better on a wide range of academic measures (e.g. math and reading tests, enrollment, days absent) and some evidence that they are more likely to drop out of high school. The latter may be due to a mismatch between student ability and school demands. Despite the disappointing findings regarding academic outcomes, students who won lotteries to high achievement Chicago public schools reported nearly 60% fewer arrests on a ninth grade student survey. These winners also reported getting into less trouble at school, and school administrative data suggests that they had lower incarceration rates during school ages. Of course, it is possible that schools themselves affect student arrest and incarceration rates through differential disciplinary policies (or criminal opportunities), so it is important to study whether these reductions in arrests/incarceration persist beyond high school.

To this end, Deming (2009a) examines the impacts of open enrollment lotteries (for middle and high schools) on adult criminal outcomes seven years after random assignment.[27] Given his interest in the effects of school choice on crime, he categorizes males based on their likelihood of arrest, which he estimates as a function of demographic characteristics, earlier math and reading test scores, and other school-related behaviors at young ages. For his entire sample of middle and high school lottery participants, 'high-risk' youth (defined as those in the top quintile of predicted arrest probability) have seven times more felony arrests (seven years after random assignment) than the average student from the bottom four quintiles combined.

Like Cullen, Jacob, and Levitt (2006), Deming (2009a) estimates significant effects of winning a school lottery on the quality of school attended, especially among 'high-risk' youth, but no effects on

---

Lotteries were random within population subgroups (e.g. by race or family income).

[27]He merges Charlotte-Mecklenburg school district data with data on adult (ages 16+) arrests and incarceration from Mecklenburg County and the North Carolina Department of Corrections.

achievement tests. There appears to be some effect on student enrollment during high school years, but there is no evidence that 'high-risk' lottery winners are more likely to graduate from high school.[28] Among high school lottery winners in the high-risk category, Deming (2009a) estimates a significant 0.35 (roughly 45%) reduction in the number of adult felony arrests (cumulative as of seven years after the lottery) with an associated savings in victimization costs of $4,600-16,600.[29] Because many crimes do not lead to an arrest, the total benefits to potential victims and society are likely to be much larger. His estimates suggest that winning middle school lotteries also reduces crime among high-risk youth with most effects of a similar order of magnitude.

Court-ordered school desegregation policies enacted since *Brown vs. Board of Education of Topeka* in 1954 dramatically altered the types of schools blacks attended in many American districts. In most cases, the resources and average student achievement of schools attended by blacks would have improved markedly.[30] Guryan (2004) estimates that these desegregation efforts significantly increased high school graduation rates among blacks by 2-3 percentage points but had no effect on white graduation rates. Weiner, Lutz and Ludwig (2009) examine whether these changes affected county-level homicide rates.[31] Their estimates suggest that homicide deaths among blacks ages 15-19 declined by 17% in the first five years after court-ordered desegregation, while homicide deaths among white 15-19 year olds declined by about 23%. Homicide deaths among slightly older whites and blacks also declined. In looking at offenders, they find that arrest rates for homicide declined by one-third for blacks ages 15-19, while there was no decline for young whites. Combining Guryan's (2004) estimated effect on high school graduation rates with the estimated effects of schooling on crime from Lochner and Moretti (2004), they argue that much of the effect may be coming from the increased schooling among blacks associated with desegregation.

For some perspective, it is interesting to compare these findings with those from the Moving-to-Opportunity (MTO) experiment, which provided housing vouchers to low-income families to move out of high poverty neighborhoods. Evaluations of MTO report that families receiving the housing vouchers moved into neighborhoods with about 25% lower poverty rates; however, these moves only led to modest improvements in the quality of schools youth attended and no improvements in their cognitive achievement (Sanbonmatsu, et al. 2006). Kling, Ludwig, and Katz (2005) report that the MTO housing vouchers led to lasting reductions in arrests for both violent and property offenses among young females, short-term reductions in violent crime arrests for males, and delayed *increases* in property crime arrests for males. Overall, reductions in crime were modest at best.

---

[28] There is more evidence of effects on high school graduation and college attendance among the lower risk quintiles.

[29] These victimization costs (in year 2009 dollars) assign costs based on the type of offense using cost estimates from Miller, Cohen and Wiersema (1996). They do not include justice system or enforcement costs. The larger figure uses a cost of $4.3 million for murder while the smaller uses a value of $125,000 (twice the cost of rape).

[30] For example, Reber (2007) shows that integration efforts in Louisiana from 1965-70 were accompanied by large increases in per pupil funding for black students.

[31] They use data on homicide death by year and county over the period 1958-88 from vital statistics and data on homicide victims and arrestees from the Supplemental Homicide Report from 1976-2003.

Taken together, these studies suggest that simply improving the schools attended by disadvantaged youth appears to be much more successful in reducing criminal activity (though not necessarily in improving academic outcomes) than changing neighborhoods. Given the mixed findings on educational attainment levels (with modest positive effects at best), the impacts of better schools on crime appear to be driven largely by school quality and not 'quantity'. Whether it is the quality of teachers and instruction or of student peers is less obvious. The fact that test scores did not improve among lottery winners suggests that the main effects of attending 'better' schools on delinquency and crime are likely to be attributed to better socialization, better peer interactions, improvements in non-cognitive skills, or changes in preferences.[32] It is, therefore, interesting that substantial improvements in 'neighborhood peers' do not yield the same benefits in terms of crime reduction.

2.2.3 Contemporaneous Schooling and Crime

We now consider the relationship between contemporaneous schooling and crime. As noted earlier, there are three main ways in which altering youth's schooling attendance is likely to affect their contemporaneous engagement in crime: (i) incapacitation, (ii) raising the costs of future punishment through human capital accumulation, and (iii) social interactions facilitated by bringing youth together. The incapacitation and human capital effects of schooling on crime are likely to be negative, while the sign of the social interaction effect is theoretically ambiguous.

Three relatively recent studies shed light on these effects by estimating the impacts of different 'interventions' that directly affect youth schooling attendance.[33] Anderson (2009) examines the effects of increasing state compulsory schooling ages on crime among affected youth (i.e. forcing some youth to stay in school), while Jacob and Lefgren (2003) and Luallen (2006) estimate the effects of extra days off from school due to teacher in-service days or teacher strikes (i.e. keeping youth out of school). The policies analyzed by these studies differ in two important respects. First, increases in compulsory schooling ages typically 'require' students to stay in school at least one additional year and sometimes more, whereas teacher in-service days and strikes are of very short duration. Second, while teacher strikes and in-service days release all students from school, changes in compulsory schooling laws typically affect a small set of marginal students. All three potential effects of school attendance on crime are likely to be relevant to changes in compulsory schooling, while the effects of in-service days and teacher strikes are likely to be limited to incapacitation and social interactions. Any social interaction effects are likely to be magnified in the latter cases due to the universal nature of the policy.

Rather than use changes in compulsory schooling laws as instruments for educational attainment,

---

[32]See Heckman, Stixrud, and Urzua (2006) for evidence on the importance of 'cognitive' vs. 'non-cognitive' skills for crime and in the labor market.

[33]Using individual-level data, earlier studies by Gottfredson (1985), Farrington, et al. (1986), and Witte and Tauchen (1994) explore the cross-sectional relationship between time spent in school and contemporaneous crime, concluding that time spent in school significantly 'reduces' criminal activity. Unfortunately, these findings are difficult to interpret given the simultaneous nature of the crime and schooling choices.

Anderson (2009) estimates the direct effect of these laws on contemporaneous county-level arrest rates (from the UCR) from 1980 to 2006 among affected youth ages 16-18. Specifically, his estimates are identified from within-county fluctuations in arrests (around county-specific trends) for 16-18 year-olds (relative to 13-15 year-olds) over time as state compulsory schooling ages change.

Anderson's estimates for total arrest rates imply that a compulsory schooling age of 17 significantly reduces age 17 arrests by about 8% (5.4 arrests per 1,000 youth) compared to a compulsory schooling age of 16 or less. Similarly, an age 18 compulsory schooling age significantly reduces arrests by 9.7-11.5% at ages 16-18. Separating arrests by type of offense, he estimates that compulsory schooling laws significantly reduce both property and violent arrests for 16-18 year olds. Although, estimated effects of schooling age laws on drug-related crimes are sizeable, the effects are typically not statistically significant. Overall, the estimates generally suggest that forcing youth to spend an extra year or two in high school significantly reduces their arrest rates over that period.

Jacob and Lefgren (2003) examine the effects of single day changes in school-wide attendance on juvenile crime and arrest rates in 29 large American cities from 1995 to 1999. Exploiting teacher in-service days across jurisdictions over time as an exogenous source of variation in school days, they essentially compare local juvenile crime rates on days when school is not in session to those when it is.[34] Their findings suggest that an additional day of school reduces serious juvenile property crime by about 14% that day while it *increases* serious juvenile violent crime by 28%. These results are consistent with an 'incapacitation effect' of school that limits participation in property crime. However, the increased level of interaction among adolescents facilitated through schools may raise the likelihood of violent conflicts (and other minor delinquency) after school. Interestingly, they find no evidence to suggest that school days simply shift crime to other days without changing overall crime rates.

Luallen (2006) follows a similar approach, using teacher strikes (typically lasting about 5 days) rather than in-service days as an exogenous source of school days. Using data from the state of Washington for 1980-2001, Luallen (2006) estimates that an extra day of school reduces arrests for property crimes by about 29% while increasing arrests for violent crimes by about 32% in urban areas. The effect on property crime is roughly double the effect estimated in Jacob and Lefgren (2003), while the effect on violent crime is quite similar. In rural and suburban areas, Luallen finds insignificant effects on both violent and property crime arrests. Thus, the incapacitation and social interaction effects appear to be particularly strong in urban areas and negligible elsewhere.

2.3 Evidence on Human Capital-Based Interventions from Birth to Young Adulthood

A growing body of evidence suggests that early childhood and school-age interventions can reduce adult crime rates. Most famously, the High/Scope Perry Preschool Program substantially lowered

---

[34]Their main specification includes controls to account for the possibility that crime may be higher on certain days of the week or that different cities may experience different monthly crime cycles.

arrest rates through age 40 for a sample of low-income minority children in Ypsilanti, Michigan. Several other early childhood interventions have produced similar effects on delinquency; however, others have not. We briefly review studies of early childhood and school-age interventions that have analyzed educational and criminal/delinquency outcomes during late adolescence or adulthood.[35] We then discuss a few programs aimed at improving school participation among adolescents or that directly provide training to adolescents and young adults.

Table 1 summarizes four small-scale early childhood interventions (Abecedarian Project, Chicago Child Parent Center (CPC), High/Scope Perry Preschool, and Infant Health and Development Program (IHDP)), their target populations, study methodology, and estimated effects on educational attainment and crime at ages 18 or older. All of the programs included a preschool component, ranging from full-time full-year care from birth to kindergarten (Abecedarian) to half-day preschool at ages 3 and 4 (Chicago CPC and Perry Preschool). Perry Preschool and IHDP also included regular home visits at preschool ages as part of their curriculums.[36] All of the programs targeted youth facing some form of disadvantage. Abecedarian and Perry Preschool specifically targeted children at-risk of having problems developing normally in school. Children enrolling in the Chicago CPC were all minorities selected from families with low socioeconomic status (SES). IHDP drew from a more heterogeneous population, targeting pre-term children born of low birth-weight (less than 2500g). Overall, these studies cover a reasonably broad range of potential preschool-based interventions and target populations. (We discuss findings for Head Start below.)

Youth from all four of these programs were followed until at least age 18, enough time to determine whether the programs have medium-term effects on the education and criminal behavior of participants. Only the Chicago CPC was not evaluated using randomized trials; however, Reynolds, et al. (2001) use a strong design of matching treated children with other comparison children based on age of kindergarten entry, eligibility for and participation in government funded programs, and neighborhood and family poverty. Children from the matched comparison sample would also have been eligible for the program had they lived in a neighborhood with a center. Sample sizes range from around one hundred children for Perry Preschool to thirteen hundred for Chicago CPC.

Both Chicago CPC and Perry Preschool significantly increased high school completion rates overall; however, the Chicago CPC had more sizeable effects on male graduation rates while Perry Preschool only raised female graduation rates (Reynolds, et al. 2001, Schweinhart, et al. 2005). The IHDP had no effect on high school dropout rates by age 18, while Abecedarian increased college attendance but not high school completion (McCormick, et al. 2006, Campbell, et al. 2002). These programs typically

---

[35]See Karoly, et al. (1998) or Blau and Currie (2006) for more comprehensive surveys of early childhood programs.

[36]All of the programs typically provided other additional services to families and children (e.g. nutritional and health services). While a subsample of the Abecedarian participants received an extended school-age intervention for the first few years of school, we focus on the preschool component of the program. The additional school-age services did not substantially impact the educational attainment or crime outcomes discussed here.

produced short-term gains in achievement scores and sometimes generated lasting gains.

The final column of Table 1 reports estimated effects of these programs on late juvenile and adult crime. As alluded to above, Perry Preschool had significant effects on lifetime crime measured as of age 40 (Schweinhart, et al. 2005). Reductions in the fraction arrested five or more times were substantial for both males and females. Both showed reductions of about one-third; however, the size of the effect in absolute terms is much larger for males given their higher baseline crime rate. Reductions in crime for Perry Preschool students were observed across a broad range of crimes (e.g. drug, property, and violent crimes) and were apparent even at younger ages. The Chicago CPC also reduced arrest rates (by age 18) by about one-third (Reynolds, et al. 2001). Another widely cited family support and preschool program, the Syracuse University Family Development Research Program, showed significant reductions in juvenile delinquency measured at a slightly earlier age: 6% of preschool participants had been placed under probation services by age 15 compared to 22% of controls (Lally, et al. 1988).[37]

The estimated savings in reduced criminal justice expenditures and victimization costs resulting from the crime reductions of Perry Preschool and Chicago CPC are sizeable. Using a 3% discount rate, Belfield, et al. (2006) estimate that the Perry Preschool produced a social benefit of over $150,000 (year 2000 dollars) per child from crime reduction alone.[38] Reynolds, et al. (2002) estimate that reductions in juvenile crime through age 18 associated with the Chicago CPC saved society roughly $8,000. Findings like these, especially those for Perry Preschool, led Donohue and Siegelman (1998) to conclude that small, rigorous early intervention programs may pay for themselves through reduced crime rates alone, if they can be targeted to high-crime groups.

Not all early childhood programs in Table 1 yield reductions in crime. While modest reductions in self-reported convictions and incarceration through age 21 are observed for Abecedarian, none of these effects are statistically significant (Campbell, et al. 2002). Based on administrative records of adult criminal charges in North Carolina, Clarke and Campbell (1998) report nearly identical rates of arrests and criminal charges (as of age 21, on average) for treatment and control children in the Abecedarian study. Similarly, IHDP produced no significant effects on crime through age 18 (McCormick, et al. 2006).

What is different about Abecedarian and IHDP that these programs did not produce the same reductions in crime? It is difficult to point to any particular curriculum difference; although, not all preschools are alike. Abecedarian began preschool at infancy and continued through kindergarten –

---

[37]The Elmira Nurse Home Visitation Program provided home visits by nurses to first-time mothers who were young, unmarried, or of low SES. Nurses visited homes for randomly assigned mothers during pregnancy and for the first two years of the child's life. Olds, et al. (1998) report mixed but encouraging effects of the program on delinquency at age 15: treated youth were more likely to self-report being stopped by the police but had fewer incidences of arrests and convictions.

[38]This figure is for benefits through age 40. Using a 7% discount rate produces a social benefit from crime of about $67,000 (Belfield, et al. 2006). Heckman, et al. (2009) report that savings from crime reduction account for about 40-65% of the benefit-cost ratio for Perry Preschool, depending on assumptions about discount rates (0-7%) and the cost of murder.

the longest of any program. It was also full-day year-round, unlike Perry Preschool or Chicago CPC. Like Perry Preschool, it showed sizeable gains in achievement and IQ, so it is difficult to attribute its lack of effects on crime to inadequate intervention. The only obvious program difference between Abecedarian and Perry Preschool or Chicago CPC that might explain the absence of any impact on crime is its lack of a 'home visit' component, but IHDP included home visits by nurses from birth through three years of age. IHDP began early but also ended when Perry Preschool and Chicago CPC began (age 3), so it is possible that the early 'home visit' combined with later preschool care is a key combination of services necessary for long-term impacts on delinquency and crime.

An alternative hypothesis is that the environments more than the specifics of the programs were important in determining impacts on crime. Chapel Hill is a mid-sized mostly white and relatively affluent university city in the South, while Ypsilanti is a smaller industrial city with a sizeable minority population. Chicago CPC sites were in low-income neighborhoods in a large urban midwestern city. (IHDP had sites throughout the U.S.) It seems quite possible that the same program might have different effects in each city. As noted by Barnett and Masse (2007), crime rates were 70% higher in Ypsilanti than Chapel Hill when the respective program participants would have been age 15. They speculate that there may have been little crime to prevent among the Abecedarian sample; however, Clarke and Cambpell (1998) report that the two control samples (Perry and Abecedarian) had very similar arrest rates (around 40%) by their early 20s. McCormick, et al. (2006) report that juvenile arrest rates among controls were similar for the IHDP and Chicago CPC as well. So, among the target populations for these programs, crime rates were fairly similar even if local crime rates were quite different. Of course, it is possible that the long-term effects of early childhood programs depend as much on the environment in which participants grow up as on individual and family characteristics of the participants themselves. If so, it is important to exercise caution in extrapolating benefits from any single program or community to the wider population.

Despite the fact that children targeted by all programs were disadvantaged, there is a sizeable difference in baseline educational attainment levels between Abecedarian and IHDP on the one hand and Chicago CPC and Perry Preschool on the other. High school graduation rates were 70% among Abecedarian controls; dropout rates (as of age 18) were only 10% among the IHDP controls. These both compare quite favorably with Chicago CPC and Perry Preschool controls who had high school completion rates ranging from 30-50%. Neither IHDP nor Abecedarian increased high school graduation rates. While Abecedarian improved college attendance rates, this does not appear to be an important margin for crime (see Figure 1). Given the tight link between high school dropout and crime discussed earlier, it may not be particularly surprising that Abecedarian and IHDP did not reduce crime given their negligible effects on high school completion. Yet, Perry Preschool substantially reduced male crime rates without raising educational attainment among males. Clearly, early interventions may reduce delinquency and criminal behavior without significantly improving final schooling

outcomes.

In the end, there is no easy explanation for the different findings across studies. While the results from these studies are individually powerful given their research designs (most are based on random assignment), it is difficult to draw strong conclusions overall about the efficacy of early childhood interventions as a national crime-fighting strategy. The fact that sample sizes are quite modest and that program populations are not necessarily representative of the U.S. raises additional questions. This itself may explain some of the variation in findings across studies. It is natural to ask how these programs would affect other populations. Questions about scalability have also been raised: can these programs and their effects be reproduced at a larger scale? These issues have led a number of researchers to analyze the largest early childhood program in the U.S.: Head Start. This program targets children from low-income families usually living in low-income communities and has served hundreds of thousands of children throughout the U.S. since its inception in 1967.

Because no large-scale long-term random assignment studies of Head Start are available, researchers have employed non-experimental methods. These studies generally examine impacts on national samples of individuals served by Head Start, using data from the Panel Survey of Income and Dynamics (PSID) or Children of the National Longitudinal Survey of Youth (CNLSY). We next discuss those studies that examine the impacts of Head Start on behavioral problems, delinquency, or measures of adult crime.

Garces, Thomas, and Currie (2002) and Deming (2009b) use a family fixed effects approach to estimate the effects of Head Start on a variety of long-term outcomes. By comparing siblings who did and did not attend a Head Start program at ages 3-5, they address important concerns about permanent or long-run differences across families that may affect decisions about preschool or Head Start enrollment.[39] Garces, Thomas, and Currie (2002) use data from the PSID, examining adult outcomes for individuals born between 1964 and 1977, while Deming (2009b) uses data from the CNLSY and examines outcomes for individuals born in the late 1970s and early 1980s. Despite using the same empirical approach, the two studies find quite different patterns for Head Start impacts on educational attainment and criminal behavior. Garces, Thomas, and Currie (2002) estimate significant increases in high school completion (by 20 percentage points) and college attendance (by 28 percentage points) for whites only, while Deming (2009b) estimates an 11 percentage point increase in high school completion rates and a 14 percentage point increase in college attendance for blacks only. Excluding GED recipients, Deming (2009b) estimates a smaller (7 percentage points) and statistically insignificant effect on high school completion for blacks, suggesting that much of their apparent improvement in high school completion is due to increases in the GED.[40] Regarding crime, estimates by Garces, Thomas,

---

[39]Of course, they leave unanswered the question as to why some siblings enroll in Head Start while others from the same family do not and, more importantly, whether different enrollment decisions are related to underlying differences in child abilities or other factors that may affect outcomes later in life.

[40]The substitution between high school degrees and GED receipt is less relevant for the earlier cohort studied by

24

and Currie (2002) suggest that Head Start reduces the probability of being booked or charged with a crime by about 12 percentage points among blacks, with no effect on whites. Deming (2009b) finds no significant effects of Head Start on crime for blacks or whites.[41]

Carneiro and Ginja (2008) use a regression discontinuity design to estimate the effects of Head Start on adolescent outcomes, including the probability that someone is sentenced for a crime. Their approach exploits the fact that Head Start imposes strict eligibility criteria related to family income and structure: children ages 3-5 are eligible if family income is below the federal poverty guidelines or if the family is eligible for public assistance. Since these criteria vary across states and time, the income thresholds vary across these dimensions as well. They exploit this exogenous variation in eligibility, assuming the effects of family income (when children are ages 3-5) on subsequent outcomes are continuous. Using data from the CNLSY on youth who would have enrolled in Head Start during the 1980s and 1990s, they estimate that participation in Head Start at ages 3-5 significantly reduces the probability (by 31 percentage points) a 16-17 year-old male is sentenced for a crime (based on self-reports). They estimate similar effects for a sample of blacks only. These estimates measure the effect of Head Start on children who were at the margin of eligibility for the program and, therefore, represent the effects we might expect with modest expansions of the program.

Altogether, the non-experimental evidence on Head Start appears to suggest some long-term effects on education and crime, but findings vary in important ways across studies.[42] The strongest effects on crime appear to exist for blacks; although, Deming (2009) finds no effect on crime for either blacks or whites. Combined with the evidence from smaller scale programs evaluated by randomized trials, there is limited but important evidence that early childhood interventions can reduce crime later in life for youth from disadvantaged backgrounds.

A recent program, Fast Track, introduced in four sites around the U.S., provides group- and individual-based services to children from grades one through ten. The program specifically targets children from high crime and poverty neighborhoods who exhibit conduct problems in kindergarten, with the primary aim of preventing antisocial behavior and psychiatric disorders. The program focuses on three elements of development: social and cognitive skills, peer relationships, and parenting.

Garces, Thomas, and Currie (2002), since the GED was much less common in the 1980s relative to more recent years.

[41] His measure of crime is an indicator equal to one if the respondent reports having been convicted of a crime, been on probation, sentenced by a judge, or is in prison at the time of the interview.

[42] While Head Start may affect juvenile and adult crime even if it has no effect on educational attainment (as with males in the Perry Preschool program), one might speculate that any increases in schooling (especially high school years) associated with Head Start should lead to reductions in crime as estimated by Lochner and Moretti (2004). Under this assumption, estimates from Ludwig and Miller (2007), which suggest that roughly doubling Head Start spending (per capita) increases high school completion rates by as much as four percentage points, imply that this policy should also reduce arrest rates by up to 3-4%. Of course, multiplying the Garces, Thomas, and Currie (2002) estimated effects of Head Start on schooling attainment among whites by Lochner and Moretti's (2004) estimated effects of education on crime suggests that Head Start attendance should significantly reduce incarceration rates among whites, while analogous estimates from Deming (2009b) suggest that Head Start should reduce crime among blacks. Yet, these studies estimated no effect of Head Start attendance on self-reported measures of arrest, conviction, or incarceration rates for these populations.

During early grades, parents were offered training and home visits to help improve parenting skills, while children were engaged in group activities to foster friendships and tutoring sessions in reading. As children aged, more individualized services were provided, along with group sessions aimed at dealing with the transition to middle school, resistance to drugs, etc. The program also incorporated a classroom intervention during grades 1-5 at schools with program children. Teachers implemented 2-3 sessions per week designed to promote social and emotional competence and to reduce aggression. Experimental estimates based on random assignment suggest that the program produced sustained improvements in conduct disorders and anti-social behavior over grades 3-9 (Conduct Problems Prevention Research Group, CPPRG, 2007). As of grade 9, high risk youth (those from the top 3% of conduct problems in kindergarten) receiving the Fast Track program showed significant reductions in self-reported delinquency and criminal behavior; however, no significant effects on anti-social behavior were found for other youth.[43] Two recent follow-up studies (CPPRG 2010a, 2010b) suggest that the reductions in crime and conduct problems extend at least two years beyond the conclusion of the program (last measured at grade 12/age 19) and continue to be focused on youth that were initially 'high risk'. Effects on juvenile conduct disorders did not appear to decline after the program, while effects on crime showed some fade-out.

Experimental evaluations of two earlier, more limited elementary school-age interventions are worth commenting on, since they also focused largely on social development among 'high-risk' children. The Montreal Longitudinal Experimental Study provided social skills training to first and second grade children, along with teacher and parent training over those same years. Boisjoli, et al. (2007) report that by age 24, children receiving the intervention (compared to control children) were twice as likely to have completed high school and only half as likely to have a criminal record. The Seattle Social Development Project intervened over a longer period (grades 1-6); however, it only provided teacher and parent training (aimed at improving child social and emotional skill development). As of age 21, Hawkins, et al. (2005) estimate that the six-year intervention had increased high school graduation rates from 81 to 91 percent and significantly reduced self-reported crime and official lifetime court charges (from 53 to 42 percent).

Altogether, the evidence from Fast Track, the Montreal Longitudinal Experimental Study, and the Seattle Social Development Project suggests that comprehensive school-age programs designed to improve social development can produce lasting impacts on educational attainment, conduct disorders, and criminal behavior. In many ways, these programs emphasized social over cognitive development relative to the preschool programs summarized in Table 1. Of course, both sets of programs were broad-based and yielded improvements in both domains.[44]

---

[43]Results for anti-social behavior are based on an index created from self-reports of serious delinquent/criminal actions like stealing something worth more than $100, assault, selling heroin or LSD, and sexual assault.

[44]This is largely consistent with recent estimates of skill production functions for both cognitive and 'non-cognitive' skills (e.g., see Cunha and Heckman 2008).

Programs targeted to older adolescents and young adults have shown mixed results. The Quantum Opportunity Program provided entering high school students with a mentor/tutor that aided them in schoolwork and community activities for four years. Financial incentives designed to encourage high school graduation and college enrollment were provided for educational, service, and developmental activities. A recent random assignment evaluation of the program reported no significant improvements in schooling or reductions in crime six years after scheduled high school graduation (Schirm, Stuart, and McKie 2006). In part, this may be due to the relatively low participation by youth in program activities.[45]

The Job Corps provides intensive basic educational and vocational training for economically disadvantaged youth and young adults ages 16-24 throughout the U.S. The program also offers a wide range of other services (e.g. counseling, social skills training, health education, job placement services). The average participant is enrolled for about 8 months, with most living in residence at training sites. The program's primary goal is to improve employment and earnings prospects. Based on a recent random assignment evaluation, Schochet, Burghardt and Glazerman (2001) conclude that the program produced modest positive impacts on post-program employment and earnings. The program also reduced self-reported arrest rates by about 30% during the first year after random assignment, when most youth would have been enrolled. Reductions in subsequent years were smaller and statistically insignificant. The program also significantly reduced conviction rates by about 17% during the four years following random assignment.[46] Conclusions from the less-expensive and non-residential JOBSTART program are largely consistent with these findings (Cave, et al. 1993).[47]

Collectively, these studies indicate that human capital-based interventions from early childhood to early adulthood can reduce juvenile and adult crime, at least for some populations. To understand why, it is useful to return briefly to the model developed earlier. The model suggests that effective interventions may reduce juvenile and adult crime by improving child learning productivity, A, increasing adolescent human capital levels, H_1, or by socializing children (i.e. lowering θ). While preschool programs highlighted in Table 1 may raise learning abilities, achievement gains are generally short-lived and limited to primary school ages. Evidence of reduced criminal activity among adolescents

---

[45] On average, youth spent only 177 hours per year on educational, community, and developmental activities. Roughly one-in-four spent no time at all in these activities by the fourth year of the program.

[46] An earlier study by Long, et al. (1981) estimated that the social benefits from reduced criminal activity among Job Corps participants amounted to over $7,000 (in 2008 dollars) per participant – almost 30% of the total social benefit of the program.

[47] JOBSTART offered many of the same basic components of the Job Corps to a similar population. Cave, et al. (1993) find modest (and statistically insignificant) positive effects on earnings 3-4 years after random assignment for the full sample; however, earnings increased roughly 25% (in years 3 and 4) for male participants with a prior arrest (i.e. had an arrest since age 16 but prior to random assignment). Among male participants with no prior arrest, the program significantly reduced self-reports of an arrest (6.4 percentage points or 36%) during the first year after random assignment (i.e. the training year) but did not reduce the fraction arrested in subsequent years. Among males with a prior arrest, the program (insignificantly) reduced the fraction reporting an arrest over the first four years after random assignment by about 8% and had negligible effects on arrests during the first year. There were no significant effects on arrests for female participants.

attributed to early intervention programs, suggests that these programs raise adolescent market skills (H_1) and/or reduce criminal returns (θ) through socialization. School-based programs for high-risk youth like Fast Track emphasized social development (i.e. lowering θ) over cognitive achievement; yet, they also likely improved adolescent human capital levels H_1. Despite the difference in emphasis between the two types of programs, both have shown the ability to significantly reduce juvenile and adult crime. Job training programs for adolescents and young adults directly operate on the incentives to invest in human capital (analogous to an increase in the subsidy rate s in our model) and have led to modest reductions in crime during periods of heavy training. These programs have produced only modest increases in earnings and negligible long-run effects on crime, however, suggesting that simply training low-skilled adolescents does not provide the same promise as earlier interventions that act on individual endowments.

2.4 Policy Lessons

We next discuss a number of important policy lessons regarding human capital policies and crime. First, we summarize evidence on the social savings from crime reduction that we might expect from policies that increase educational attainment or enrollment, improve school choice and quality, or expand access to early childhood interventions. Second, we highlight a few sub-populations and schooling margins that are likely to yield the greatest social gain from crime reduction. Finally, we discuss a few other lessons based on the evidence.

2.4.1 Valuing the Social Benefits from Crime Reduction

Lochner and Moretti (2004) estimate that increasing educational attainment levels in the population yields sizeable social benefits. Specifically, they calculate the social savings from crime reduction that would result from a one percentage point increase in high school graduation rates in the U.S. Table 2 summarizes their exercise, translating all dollar values into 2008 dollars using the Consumer Price Index for All Urban Consumers (CPI-U). Column 1 reports total costs per crime associated with murder, rape, robbery, assault, burglary, larceny/theft, motor vehicle theft, and arson.[48] Column 2 reports the predicted change in total U.S. arrests based on the Lochner and Moretti (2004) offense-specific arrest estimates discussed earlier and the total number of arrests in the 1990 Uniform Crime Reports. Column 3 adjusts the arrest effect in column 2 by the number of crimes per arrest. In total, nearly 100,000 fewer crimes would have taken place in 1990 if high school graduation rates had been one percentage point higher. The implied social savings from reduced crime are shown in column 4. Savings from murder alone are as high as $1.7 billion. Savings from reduced assaults amount to nearly $550 million. Because the estimates suggest that schooling increases rape and robbery offenses, increased costs associated with these crimes partially offset the benefits from reductions in other

---

[48]These costs include incarceration and victim costs. See notes to Table 2 or Lochner and Moretti (2004) for details.

crimes.

The final row reports the total savings from reductions in all eight types of crime. Because these figures only include a partial list of crimes (e.g. nearly 25% of all prisoners in 1991 were incarcerated for drug offenses according to the U.S. Dept. of Justice (1994)) and do not include all costs associated with each crime (e.g. private security measures are omitted), these amounts are likely to under-estimate the true social benefit associated with increasing high school graduation rates. Still, the savings are substantial: the social benefits of a one percentage point increase in male U.S. high school graduation rates (from reduced crime alone) in 1990 would have amounted to more than $2 billion.[49] This represents more than $3,000 in annual savings *per additional male graduate.*

Machin, Marie and Vujic (forthcoming) also estimate the social savings from crime reduction associated with increasing the population of individuals with an education qualification (similar to high school completion in the U.S.). Accounting only for benefits from property crime reduction (estimated effects on violent crime are statistically insignificant), their estimates suggest a savings of over £10,000 per additional student qualification.

Open school enrollment lotteries and desegregation efforts appear to reduce crime rates by improving school quality. Deming (2009a) estimates that reductions in arrests associated with offering better quality school options to a high-risk youth produces a roughly $16,000 social savings to victims over the next seven years. Because better schools are also likely to have reduced crimes that never led to an arrest, total victimization savings are likely to be substantially higher. Total social savings should be still larger once savings on prisons and other crime prevention costs are factored in.

The effects of school attendance on contemporaneous juvenile crime rates are more complicated. Studies estimating the effects of day-to-day changes in attendance suggest that in urban communities additional school days reduce property crime while increasing violent crime (Jacob and Lefgren 2003, Luallen 2006). Overall, the social costs associated with increased violence are likely to dominate the benefits from reduced property crime. On the other hand, Anderson (2009) estimates reductions in both violent and property juvenile crime associated with increases in compulsory schooling ages. Thus, his findings suggest an overall social savings from juvenile crime reduction, although he does not attempt to put a dollar value on the effects.

Evidence on the effects of early childhood and school-age interventions are mixed. Long-run impacts on juvenile delinquency and adult crime can be substantial for disadvantaged youth. For example, estimates suggest that Perry Preschool produced a social benefit from crime reduction of roughly $150,000 per child (through age 40). On the other hand, Abecedarian produced no significant impacts on crime.

---

[49]These calculations are partial equilibrium in nature, implicitly assuming that the skill prices of workers do not change in response to a policy that increases national high school completion rates. Fella and Gallipoli (2009) show that the social benefits of such a policy would be even larger if general equilibrium effects on the wages of high school dropouts and graduates were incorporated.

2.4.2 Where are the Big Returns?

Given the most sizeable reductions in crime appear to result from the final years of high school, policies that encourage high school completion would seem to be most promising in terms of their impacts on crime.[50] Because crime rates are already quite low among high school graduates, policies that encourage college attendance or completion are likely to yield much smaller social benefits from crime reduction; although, they may be desirable on other grounds.[51]

In general, policies designed to encourage schooling among more crime-prone groups are likely to produce the greatest benefits from crime reduction. Consistent with this, the school-age Fast Track program appears to have reduced juvenile crime only among very high-risk children, showing little impact on even moderately high-risk children (CPPRG 2007, 2010b). Similarly, Deming (2009a) estimates that improved school choice for middle and high school students leads to significant reductions in arrests for high-risk youth but not for others. As Donohue and Siegelman (1998) conclude, the overall efficiency of early childhood programs as a crime-fighting strategy is likely to depend heavily on the ability to target high-risk children at very young ages. The same is likely to be true for school-age interventions.

Social benefits from crime reduction also vary across gender and race. Men commit much less crime than women, on average. Thus, it is not surprising that crime-related benefits from education policies and interventions are typically much smaller for females than males (e.g. Perry Preschool, Job Corps). This is true even though programs sometimes reduce female and male crime rates by similar amounts in percentage terms. Among men, Lochner and Moretti (2004) estimate much larger effects of additional schooling on incarceration rates among blacks relative to whites. Garces, Currie and Thomas (2002) estimate that Head Start significantly reduces crime for blacks but not whites; however, Deming (2009b) estimates no effect on crime for either group while Carneiro and Ginja (2008) estimate similar large effects on both. Because crime rates are much higher among blacks than whites, on average, policies would generally need to produce much larger proportional reductions in white crime rates to achieve similar absolute reductions in crime. None of the evidence surveyed here suggests that this is the case.

2.4.3 Additional Policy Lessons

A few other useful lessons can be drawn from the studies surveyed here.

First, education policies can reduce property crime as well as violent crime. In the U.S., the estimated effects of educational attainment or school enrollment on property and violent offenses appear

---

[50]See Hanushek and Lindseth (2009), Jacob and Ludwig (2008), or Murnane (2008) for recent discussions of policies to improve schooling outcomes in the U.S.

[51]The fact that crime declines substantially with high school completion but not college attendance suggests that net expected returns from crime for most individuals lie somewhere between the wages of high school dropouts and graduates. See Freeman (1999) for a summary of evidence regarding criminal wages and earnings.

to be quite similar in percentage terms (Lochner and Moretti 2004, Anderson 2009).[52] Even murder appears to be quite responsive to changes in educational attainment and school quality (Lochner and Moretti 2004, Weiner, Lutz, and Ludwig 2009).

Second, higher wages increase the opportunity costs (including work foregone while incarcerated) of both property and violent crime. Lochner and Moretti (2004) show that the estimated effects of educational attainment on crime can largely be accounted for by the effects of schooling on wages and the effects of wages on crime. This is important, since it suggests that policymakers can reduce crime simply by increasing labor market skills; they need not alter individual preferences or otherwise socialize youth.[53] Of course, as evidence from the Job Corps and other training programs suggests, this is not necessarily an easy task. Training programs targeted at low-skill adolescents and young adults have modest (at best) effects on earnings and crime. On the other hand, encouraging youth to finish high school (e.g. through compulsory schooling laws) appears to substantially increase earnings and reduce crime. Preventing early school dropout is likely to be more successful than trying to compensate for dropout a few years later.

Third, education-based policies need not increase educational attainment to reduce crime. Studies on school choice lotteries (Cullen, Jacob, and Levitt 2006, Deming 2009a) suggest that providing disadvantaged urban youth with better schools can substantially reduce juvenile and adult crime, even if it has small effects on traditional education outcomes. Perry Preschool had no effect on male schooling levels but substantially reduced male crime rates through age 40 (Schweinhart, et al. 2005).

Fourth, evidence that violent crime is higher on school days than non-school days in urban districts suggests that social interaction effects are particularly important for juvenile violent crime (Jacob and Lefgren 2003, Luallen 2006). Smart policing efforts may be able to help address some of the problems associated with schools releasing lots of adolescents at the same time. For example, an increased police presence immediately after school or other major adolescent congregations let out may be warranted. Or, on non-school days, it may be wise for police to focus more on targets or areas of juvenile property crime, worrying less about violent crime. The 'hot spot' or 'problem-oriented policing' literature in criminology suggests that informed targeting of police efforts to high crime areas (and, by extension, times) can be effective at reducing overall crime rates.[54] Alternatively, it may be useful to consider ways of designing after-school youth programs or other weekend activities to minimize violent behavior afterwards.

---

[52]Estimates from Machin, Marie, and Vujic (forthcoming) suggest that education reduces property crime more than violent crime in the U.K.

[53]As Heckman, Stixrud and Urzua (2006) show, both 'cognitive' and 'non-cognitive' skills are acquired in school, are rewarded in the labor market, and affect crime.

[54]For a recent survey of this literature, see Braga (2005).

2.5 Interpreting the Evidence on Crime

Much of the empirical evidence is broadly consistent with a human capital-based model of crime and work premised on the notion that schooling raises legitimate wage rates much more than the returns to most types of crime. Indeed, Lochner and Moretti (2004) argue that the reductions in violent and property crime associated with increased schooling in the U.S. is roughly equivalent to the effect of education on wages multiplied by the effect of increased wages on crime.[55] Thus, most of the effect of education on violent and property crime may come from increased wages. By contrast, education may increase the returns to white collar crime more than the returns to work. Consistent with this, Lochner (2004) finds that arrest rates for white collar crime increase when education levels rise.

The human capital-based model of crime is further supported by recent trends in incarceration differentials by education in the U.S. as documented by Western (2006). Since the early 1980s, the U.S. has witnessed a dramatic increase in the wage returns to schooling, the probability of incarceration, and prison sentence lengths. All of these trends have served to increase the opportunity costs of crime relatively more for the most educated. As predicted by the model, incarceration rates have become more strongly correlated with education over this period.

Education-based programs may also socialize youth, reducing personal or psychic rewards from crime. Heckman and Rubinstein (2001) argue that 'non-cognitive' skills may be more important determinants of many life outcomes than 'cognitive' skills among those at the bottom of the education distribution. Emphasizing social and emotional development, school-age programs like Fast Track have shown the ability to significantly reduce later conduct disorders and crime (among high-risk children). These programs also improved educational outcomes, which may explain some of their impacts on crime. Perry Preschool reduced male (and female) crime rates without affecting male schooling outcomes. Thus, the program appears to have improved social development or increased early skill levels (without noticeably affecting subsequent schooling investments). Evidence from school choice lotteries suggests that improvements in school and peer quality can lead to reductions in crime without raising student achievement or educational attainment. The most likely explanation for the reduction in crime is that higher quality schools better socialize youth or provide them with a better set of peers. Yet, evidence from the MTO experiment suggests that moving families to lower poverty neighborhoods does not produce the same reductions in crime, complicating any explanation related to peer effects or social networks.

Education may also increase patience or alter preferences for risk; however, neither seems to be

---

[55]Machin and Meghir (2004) estimate the elasticity of property crime to low-skill wages to be a little over one in the U.K. Grenet (2010) estimates that an extra year of school (induced by the U.K.'s 1973 increase in compulsory schooling) increases wages by about 7%. Combining these figures, the effect of an extra year of school through increased wages should be to lower property crime by 7-10% in the U.K. Estimates from Machin, Marie, and Vujic (forthcoming) suggest that an extra year of school reduces overall property crime by 10-15% (20-30% among men and very little among women). Altogether, these calculations suggest that increased wages explain anywhere from one-half to all of the effect of education on property crime in the U.K.

central to the estimated impacts on crime. Property crimes are generally associated with less than one month of expected time in jail or prison conditional on being sentenced (see Table 3), hardly enough time for modest changes in patience to play much of a role. Property crimes also have very low expected probabilities of arrest (typically less than 10% chance) and even lower probabilities of incarceration (typically less than 1%), so there is little actual uncertainty in outcomes associated with these crimes (see Table 3). Yet, estimated impacts of schooling on property crime are similar to those for violent crime, which entails much longer and more uncertain prison sentences.

Altogether, the evidence suggests that while efforts to socialize youth can be effective, simply providing them with valuable market skills can discourage them from choosing a life of crime. In terms of crime reduction, human capital-based policies that target the most disadvantaged (and crime-prone) are likely to be the most efficient, while also promoting a more equitable society. To that end, increasing high school graduation rates and improving our nation's worst inner city schools are likely to yield the greatest social return.

Although policies that increase school attendance for a year or more (e.g. increased compulsory schooling ages) appear to reduce both violent and property crime (Anderson 2009), a few extra days off from school may actually lead to reductions in violent crime, especially in urban areas (Jacob and Lefgren 2003, Luallen 2006). From a human capital perspective, the increased opportunities that open up for youth attending an additional year of schooling should raise the future costs of incarceration associated with juvenile crime. This may serve as an important additional criminal deterrent that does not exist for day-to-day changes in the school calendar. In general, the effects of longer periods of attendance on contemporaneous juvenile crime are consistent with the subsequent effects of additional schooling on adult crime.

There are many ways by which early childhood interventions may affect juvenile and adult crime. The human capital approach favored in this chapter highlights the potential effects of these programs on learning abilities, adolescent skill levels, and socialization or tastes for crime. These programs may also affect childhood preferences, including risk aversion, patience, or self-control. While a few early childhood programs have produced sizeable reductions in both juvenile and adult crime – most famously, Perry Preschool – other quite similar programs have not. School-age interventions focused on developing social and emotional skills have proven successful at reducing later conduct disorders and crime, especially among very high-risk children. The benefits from reduced crime associated with successful programs certainly warrant the attention they have received; yet, we still need to know much more about why other programs have not produced the same effects. Two things are clear. First, preschool and school-age programs have substantially reduced crime for some disadvantaged high-risk populations. Even if these gains cannot be expected in all cases, they are large enough to warrant careful consideration on a broader scale. Second, successful programs did not always increase educational attainment, even when they significantly reduced juvenile and adult crime rates. Thus,

disappointing achievement or educational outcomes need not imply the absence of benefits from crime reduction.

Given current evidence, it is difficult to draw strong conclusions about the relative benefits of trying to target and 'treat' children at very young ages vs. intervening at later ages to keep adolescents from dropping out of high school. Of course, we need not choose one or the other. Indeed, both are likely to be important components of a broad-based national crime-fighting agenda. Calculations by Lochner and Moretti (2004) and Donohue and Seigelman (1998) suggest that both human capital-oriented policies are competitive with more traditional law enforcement and incarceration efforts when all benefits are considered.

3   Education's Effects on Health and Mortality

Health and mortality gaps by education are large and have been growing for decades (see, e.g., Pappas, et al. 1993, Goesling 2007, Meara, et al. 2008). In 2000, white males with at least some college education could expect to live 6.2 years more than their less educated counterparts (Meara, et al. 2008). As highlighted by recent surveys of the literature on education and health (e.g. Grossman and Kaestner 1997, Grossman 2000, 2006), education is more strongly correlated with health than is income or occupation. As discussed below, the literature has identified many reasons education may improve health and reduce mortality.

We address three questions in this section: (i) Does education actually improve health and mortality outcomes (i.e. is there a causal effect of education)? (ii) If so, why? (iii) What does this mean for education or health policy?

We first develop a simple model of education and mortality to shed light on a few key mechanisms at play. In particular, the model incorporates three often-discussed channels: First, education may directly increase health production by raising the marginal productivity of health inputs or behaviors, sometimes referred to as 'productive efficiency' (Grossman 1972a). Second, education may enhance one's ability to acquire and process health information or to follow more complicated treatments. As a result, education may improve the 'allocative efficiency' of health inputs (Rosenzweig and Schultz 1982). Third, education generally increases earnings, which makes costly health care and insurance purchases more affordable. An increase in income also raises the demand for health and longevity by increasing consumption opportunities (Grossman 1972a). Overall, the model clarifies how education affects mortality directly and indirectly through the choice of health inputs.

We next discuss recent evidence on whether education has a causal effect on health and mortality.[56] In particular, we focus on the growing number of studies that attempt to address well-known concerns about the endogeneity of schooling and unobserved heterogeneity in preferences using instru-

---

[56]For surveys of earlier work, see Grossman and Kaestner (1997) or Grossman (2000, 2006).

mental variable (IV) or regression discontinuity (RD) techniques. We discuss both the credibility of these studies as well as their findings, concluding that the most convincing of these studies find that additional schooling leads to modest improvements in health and small reductions in mortality.

Much of the most recent literature on education and health has attempted to explain why schooling may improve health and reduce mortality. Most of these studies take one of two approaches. Some studies attempt to empirically decompose the education – health gradient by regressing various health outcomes or behaviors on education and different subsets of potential mediators (e.g. income, ability, preferences, information about treatments or diseases) to see how much of the education – health gradient can be 'explained by' these mediators. Other studies focus on a single potential mechanism, theory, or explanation. Most studies (by economists) in this vein have attempted to determine whether differences in health/mortality by education are driven by differences in specific health knowledge or behavior. We briefly review both types of studies and discuss the strengths and weaknesses of each approach.

Finally, we discuss a few important policy issues. First, we argue that is important to determine whether individuals implicitly or explicitly pay for any health benefits associated with greater education (e.g. better health insurance or more expensive treatments). This is crucial for determining the net return on education and the value of education as a health-improving policy goal, since additional costs should be netted out. Second, it is important to determine whether health benefits accrue exclusively to the individual that becomes educated, to other family members, or to broader social networks and society at large. Evidence of important health externalities may justify public expenditures on schooling. Third, it is reasonable to ask whether individuals are aware of future health benefits associated with additional schooling. If health benefits are effectively free and individuals are unaware of those benefits (or fail to incorporate them) when making their education decisions, governments may wish to subsidize education even in the absence of externalities.

3.1   A Simple Model of Education and Mortality

Grossman (1972a, 1972b, 1975) develops a dynamic model of health investment related to the traditional human capital investment models of Becker (1967) and Ben-Porath (1967).[57] In his framework, health affects the time endowment for workers, while education may affect health as well as returns in the labor market. Grossman (2006) discusses key aspects of the model in a simpler static environment.

In this section, we develop a simple model in which education and 'health inputs' affect mortality. The model helps formalize and clarify three mechanisms: productive efficiency, allocative efficiency, and differential demand for health by income. It also facilitates a 'back of the envelope' calculation offered below on the potential benefits of education through improved mortality, an important and objective measure of health. In many ways, the model developed here is similar to that of Becker

---

[57]Also, see Ehrlich and Chuma (1990).

and Mulligan (1997), which allows individuals to affect their rate of time preference through costly up-front expenditures.[58]

Individuals wish to maximize discounted lifetime utility from consumption c, but they face some probability of death each period. Schooling $s \in \{\underline{s}, \bar{s}\}$ may directly improve the survival probability ('productive efficiency') or individuals may purchase costly inputs x at price p to increase this probability. While we model these inputs as goods with price p, one could easily introduce other costs more directly related to foregone opportunities (e.g. giving up smoking or drinking). We assume that the survival probability at any age, $\pi(x, s)$, is strictly increasing and concave in inputs $x \geq 0$ and increasing in schooling s. We also assume complementarity between health inputs and schooling: $\frac{\partial^2 \pi}{\partial x \partial s} \geq 0$. Furthermore, $\pi(0, \underline{s}) \geq 0$ and $\lim_{x \to \infty} \pi(x, \bar{s}) = \bar{\pi} < 1$, ensuring that people have some probability of surviving each period but do not live forever. Since we are not concerned with lifecycle issues here, we assume that $\pi(x, s)$ is a time invariant function and that individuals must choose the same x at all ages. Naturally, one would want to relax these assumptions in order to study how health or mortality investments vary over the lifecycle.

Schooling also increases individual earnings $y(s)$ so $y'(s) \geq 0$. However, we ignore any lifecycle effects on income in assuming that post-school earnings are constant over time. We also abstract from savings/borrowing decisions and assume consumption $c = y(s) - px$. While alive, individuals receive utility $u(c)$ from consumption at each date and discount utility from future consumption at rate $\tilde{\beta} \in (0, 1)$. It is also useful to define the 'effective discount rate' $\beta(x, s) \equiv \tilde{\beta}\pi(x, s)$. Without loss of generality, we effectively normalize the utility of death at zero.[59]

We begin by considering optimal health input decisions given schooling attainment s. Given our assumptions, individuals who have completed s years of school solve the following problem:

$$
\begin{aligned}
V(s) &= \max_{x:x \geq 0} \left\{ \sum_{t=0}^{\infty} [\tilde{\beta}\pi(x, s)]^t u[y(s) - px] \right\} \\
&= \max_{x:x \geq 0} \left\{ \frac{u[y(s) - px]}{1 - \beta(x, s)} \right\}.
\end{aligned}
$$

Optimal health input x decisions depend on the marginal benefits less the marginal costs. These net marginal benefits are reflected in the following function:

$$
N(x, s) \equiv \frac{\partial \beta(x, s)}{\partial x} u(c) - [1 - \beta(x, s)]pu'(c), \tag{9}
$$

where the first term reflects the marginal benefits of extending life from an increase in health expenditures while the second term reflects the loss in utility from foregone consumption. If inputs x are not

---

[58]There are some important differences. Becker and Mulligan (1997) assume a fixed resource endowment, the cost of increasing patience (survival probability in our case) is paid up front with a one-time expenditure, and individuals can borrow and save to smooth consumption optimally over time. Given the emphasis on mortality here, our model assumes that individuals receive income each period they remain alive and must pay each period to reduce their mortality rate. For simplicity, we assume that individuals cannot borrow or save.

[59]The 'bad' state in our model could be something other than death. The key assumption is that utility in this state must be independent of schooling and health expenditures.

sufficiently beneficial, some individuals may choose not to purchase them at all (i.e. individuals with $N(0, s) < 0$). Given our assumption that $\frac{\partial^2 \pi}{\partial x \partial s} \geq 0$, there exists a threshold schooling level $\hat{s}$, above which individuals invest $x = x^*(s) > 0$ (satisfying the interior first order condition $N(x^*(s), s) = 0$), and below which individuals invest $x = 0$.[60] (In general, we would not expect all health inputs to be strictly zero for any individuals. Allowing for multiple inputs would generally yield separate thresholds for each input, such that more educated individuals would utilize a wider range of inputs.)

Among individuals with $s > \hat{s}$ who purchase positive health inputs, it is straightforward to show that those inputs are increasing in completed schooling:

$$\frac{dx^*(s)}{ds} = \frac{\left[\frac{\partial^2 \beta}{\partial x \partial s} + \left(\frac{1}{1-\beta}\right) \frac{\partial \beta}{\partial s} \frac{\partial \beta}{\partial x}\right] u(c) + \left[\frac{\partial \beta}{\partial x} u'(c) - (1-\beta) p u''(c)\right] y'(s)}{-\frac{\partial^2 \beta}{\partial x^2} u(c) - p^2 (1-\beta) u''(c)} > 0.$$

This reveals two main effects of past schooling on health inputs. The first term in the numerator reflects the fact that schooling and health inputs are complementary, both directly and indirectly. The indirect complementarity arises because the marginal value of extending life today is greater if future survival rates are also high. Thus, by increasing future survival rates, schooling raises the marginal value of investing in health today. The second term in the numerator reflects simple income effects due to the market returns from schooling. By increasing earnings, schooling increases the demand for longevity and, therefore, health inputs. Overall, health inputs may be zero for a low range of schooling levels ($s \leq \hat{s}$), and will be strictly increasing in schooling among those with $s > \hat{s}$.

Since health inputs are increasing in schooling, it is clear that

$$\frac{d\pi}{ds} = \frac{\partial \pi}{\partial s} + \frac{\partial \pi}{\partial x} \frac{dx}{ds} > \frac{\partial \pi}{\partial s} \geq 0$$

and that schooling increases survival rates. Among individuals with low schooling levels ($s \leq \hat{s}$), the total effect of schooling equals the partial effect (i.e. $\frac{d\pi}{ds} = \frac{\partial \pi}{\partial s} \geq 0$). Among more educated individuals with $s > \hat{s}$, the total effect of schooling $\frac{d\pi}{ds}$ on mortality is greater than its partial effect $\frac{\partial \pi}{\partial s}$ due to the responsiveness of health inputs. Indirect effects through increased health expenditures will be large when (i) schooling and health inputs exhibit strong direct complementary, (ii) schooling raises income a lot, and (iii) mortality is very responsive to health inputs and/or schooling. This suggests that productive efficiency may lead to a stronger education – health/mortality gradient at the top end of the education distribution (i.e. $s > \hat{s}$) where health input decisions play a more important role.

It interesting to consider how a change in the price of health inputs affects health decisions. It is straightforward to show that input purchases are decreasing in price ($dx^*/dp < 0$), while $\hat{s}$ is increasing in prices. Thus, innovations that lower the price of health inputs will cause highly educated individuals to increase their health investments and some less-educated individuals to begin purchasing the cheaper

---

[60]The complementarity assumption $\frac{\partial^2 \pi}{\partial x \partial s} \geq 0$ ensures that $N(0, s)$ is increasing in s, yielding the threshold $\hat{s}$ as the solution to $N(0, \hat{s}) = 0$. Second order conditions for a maximum are satisfied by the assumptions on $\pi(x, s)$ and $u(c)$.

input. The latter effect will tend to shrink health and mortality differences by education; however, it is more difficult to sign the effect of a price change on $\frac{dx^*}{ds}$ among those already using the input. While a decline in the price of an input should unambiguously reduce education differences in the fraction of individuals using that input, it could actually widen education differences in the quantity of inputs used (among initial users).

What happens if new information arises suggesting that some particular health input (or behavior) is good for one's health? For example, suppose $\frac{\partial \beta}{\partial x}$ was previously thought to be zero but is now thought to improve health, i.e. $\frac{\partial \beta}{\partial x} > 0$. This new information should cause individuals with $s > \hat{s}$ to begin using the new input, with the most educated increasing their use the most. By contrast, lower-educated individuals may not respond at all to the new information. The higher the price of the new input, the narrower the response will be among the most educated. In general, new information that some input or behavior is beneficial for health is likely to increase education – health/mortality gradients even if that information is distributed and interpreted equally by everyone.[61] We would expect a similar response to the introduction of new health inputs.

3.1.1 Valuing Mortality Benefits from Education

Using the envelope condition, it is straightforward to derive the marginal lifetime benefit of schooling:

$$
\begin{aligned}
V'(s) &= \left[\frac{u'(c)}{1-\beta}\right]\left[\left(\frac{\tilde{\beta}}{1-\beta}\frac{u(c)}{u'(c)}\right)\frac{\partial \pi}{\partial s} + y'(s)\right] \\
&= \left[\frac{u'(c)}{1-\beta}\right]\left[VSL \times \frac{\partial \pi}{\partial s} + y'(s)\right],
\end{aligned} \tag{10}
$$

The marginal value of schooling can be decomposed into a 'mortality effect' and a standard lifetime earnings effect. Interestingly, the mortality effect can be calculated from standard estimates of the value of a statistical life (VSL) and the estimated (partial) effects of schooling on survival rates.[62]

Below, we calculate the 'mortality effect' using plausible values for the effect of schooling on mortality and measures of the value of a statistical life. Empirically, it is difficult to distinguish between the total effect of schooling on mortality $\frac{d\pi}{ds}$ and the partial effect $\frac{\partial \pi}{\partial s}$ observed in equation (10). Most recent studies that attempt to estimate the causal effects of schooling on health estimate the

---

[61]As with price changes, it is more difficult to generally sign the response of dx^*/dp to marginal increases in the productivity of inputs when some of the population is already using x.

[62]The VSL is typically estimated as the amount of money someone would need to be compensated at date t to raise their mortality risk by some amount $d\pi_t$. Based on our model, we can write the discounted utility for someone as of date t as $V_t = u(y_t - x) + \tilde{\beta}\pi_t[u(c) + \tilde{\beta}\pi u(c) + (\tilde{\beta}\pi)^2 u(c) + ...]$, where we specifically allow y_t and π_t to differ from all future values since we want to consider changing both while holding lifetime utility constant. Totally differentiating V_t with respect to π_t, we obtain the VSL as the amount income would have to adjust in response to keep lifetime utility constant:

$$
VSL \equiv -\frac{dy_t}{d\pi_t}\mid_{V_t = \bar{V}_t} = \frac{\tilde{\beta}}{1-\beta}\frac{u(c)}{u'(c)}.
$$

total effect rather than the partial effect; however, as this analysis and our discussion below suggests, the partial effect is crucial for education and health policy.

3.1.2 Allocative Efficiency

The notion that education may affect the allocation of health inputs is often discussed by economists. See Rosenzweig and Schultz (1982) for an early treatment of the issue. Health economists have discussed a number of reasons more educated individuals may make different health-related choices. For example, more educated individuals may be better at acquiring or processing health information, or they may better understand and follow complicated directions. Education may also raise the demand more for some health inputs than others. We briefly discuss some of these ideas within the context of a generalized version of the model above.

To incorporate the idea of 'allocative efficiency', consider two potential inputs x_1 and x_2 with prices p_1 and p_2, respectively. To keep things very simple, assume that these inputs are perfect substitutes in production (e.g. two different forms of medicine for the same disease), so we can write the survival probability as $\pi(a_1 x_1 + a_2 x_2, s)$.

Except when $\frac{a_1}{a_2} = \frac{p_1}{p_2}$, perfect substitutability implies that individuals will never purchase both inputs. They will purchase input x_1 if $\frac{a_1}{a_2} > \frac{p_1}{p_2}$; otherwise, they will choose x_2. The least-educated may choose to purchase neither input. The characterization of input decisions above carries over here to the desired input. As such, the introduction of a new input x_2 that is equally productive as x_1 (i.e. $a_1 = a_2$) but which has a lower price should cause individuals to switch to the new input, choosing a higher input level. The effects on health are equivalent to those observed for a price reduction in the more simple single-input model above.

If the relative productivity of investments a_1/a_2 or prices p_1/p_2 depend on educational attainment (e.g. more educated persons may be better at implementing more complicated treatments), then the choice of inputs will naturally depend on education. If education raises the productivity of or lowers the price of x_2 relative to x_1, then more-educated individuals will choose to use x_2 while less-educated individuals will favor x_1. For example, a new treatment that only lowered costs for the most educated should cause only them to switch treatments, increasing the education – health gradient. New treatments that greatly simplify health regimens may reduce costs for the least educated relatively more, having the opposite effect (Goldman and Lakdawalla 2005).

Knowledge about more health options/inputs should generally improve health, since more knowledgeable individuals may be able to take advantage of more productive or less costly inputs. This is the presumption in much of the health economics literature. However, it is also quite possible that individuals are aware of their health options but are simply mis-informed about the productivity of some inputs. In this case, it is not obvious that they should under-estimate health productivity. It seems equally plausible that some individuals may over-estimate the value of certain health behaviors

or treatments (e.g. the value of exercise or not drinking). This should lead to excess investments in health. In this case, better information may actually cause individuals to invest less in their health and become less healthy (albeit better off in utility terms). This suggests that the nature of information deficiencies may be as important as the extent of those deficiencies.

3.2   Evidence on the effects of education on health and mortality

We first review recent empirical studies that estimate the causal effects of education on health and mortality, focusing attention on studies that most convincingly address concerns about the endogeneity of schooling and unobserved heterogeneity. We then discuss evidence aimed at distinguishing between various mechanisms that might explain the relationship between education and health.

3.2.1   Does education affect health and mortality?

It is important to recognize that education may be correlated with health and mortality even if schooling has no causal effect on either. For example, healthier individuals are likely to be more efficient producers of human capital through schooling (Grossman 1972a). These effects may be compounded over time if previous health levels affect present health status. More generally, unobservable individual characteristics (e.g. genetic background, family environment, patience) may jointly affect both education choices and health-related behaviors (Fuchs 1982). An important goal of recent empirical studies has been to identify the causal effects of education on health or mortality outcomes. These studies generally rely on IV or RD estimation strategies designed to address concerns about the endogeneity of schooling and bias from unobserved heterogeneity. We discuss this more recent literature here, referring the reader to Grossman and Kaestner (1997) or Grossman (2000, 2006) for comprehensive surveys of earlier research.

Many of the most convincing and recent IV and RD studies use changes in compulsory schooling laws as instruments for educational attainment. Estimates from these analyses reflect the effects of education on health for youth who respond to changes in schooling laws (i.e. lesser educated youth constrained from dropping out earlier).[63] As such, they measure improvements in health associated with lower levels of education (typically secondary school). Few studies employ instruments that would help identify the effects of post-secondary schooling on health (de Walque (2007a) and Grimard and Parent (2007) are notable exceptions).

Most studies estimate health/mortality equations of the following form:

$$H_{iclt} = \beta E_{icl} + X_{icl}\gamma + d_c + d_l + d_t + \varepsilon_{iclt}, \tag{11}$$

[63]Angrist and Imbens (1994, 1995) show that, under certain conditions, IV estimation using discrete instrumental variables identifies the average effect of the endogenous regressor on the outcome for the population that changes its behavior in response to changes in the instrument. They refer to this as the 'local average treatment effect'.

where H_{iclt} reflects a health outcome for person i born in year c and location l measured as of year t. E reflects educational attainment, while X reflects other characteristics that may affect health. Some studies (especially those analyzing mortality) aggregate across individuals by gender, cohort and place of birth, but otherwise estimate similar specifications.

Since health and survival rates have generally improved over time and across cohorts, it is important to account for health trends across cohorts. This is especially true for studies using school reforms as instrumental variables, since most reforms increased educational attainment for affected cohorts. Failure to account for secular improvements in health may incorrectly attribute those changes to school reforms, biasing estimates toward finding health benefits of schooling. The extent to which researchers account for general trends in mortality (or other health outcomes) across cohorts is crucial for the credibility of most IV estimates. Some of the best studies, implicitly or explicitly, exploit an RD design by (i) focusing on changes in schooling and health only around the time of a change in school reforms or (ii) allowing for general cohort trends.

While the importance of accounting for general health trends cannot be overstated, RD strategies are likely to under-estimate the long-run effects of education on health if there are important cross-cohort spillovers. These spillovers may arise for a number of reasons. For example, firms typically offer a limited set of health insurance options to all employees, reflecting the average demand across all cohorts employed by the firm. Thus, health insurance policies may be relatively homogeneous across cohorts despite differences in demand for insurance. Alternatively, social interactions (including between spouses) may exist across cohorts. As a result, demand for smoking, drinking, or exercise may depend on what friends and family members do. RD estimates will fail to incorporate any impact due to these cross-cohort interactions.[64] Therefore, when cross-cohort spillovers are thought to be important, there is a tension between accounting for secular trends in health via RD estimation and identifying the full long-term effects of education.

Tables 4-7 summarize IV studies from the past five years that estimate the impact of education on health and mortality. All of the studies in these tables employ credible identification strategies and provide estimates with reasonable levels of precision. Tables 4-6 document the following characteristics of these studies: data sources, important sample characteristics, details on the health outcome of

---

[64]Consider a simple example of social interactions, in which the utility associated with health input x for individual i from cohort t is given by $U(x_{it}) = \alpha_t x_{it} - \frac{\beta}{2}\left[\theta\left(x_{it}^2 + \gamma[x_{it} - \bar{x}]^2\right) + (1-\theta)\sum_{j=1}^{J}\omega_j(x_{it} - \bar{x}_{t-j})^2\right]$, where $\bar{x}_t$ reflects the cohort t average for x and $\sum_{j=1}^{J}\omega_j = 1$. This specification implies that individuals consider their own personal benefits and costs associated with x, but they would also like to choose x to mimic the choices of others in their cohort as well as J previous cohorts. θ reflects the importance of one's own private cost and the desire to conform with others in the same cohort relative to the desire to conform with older cohorts. If we consider a policy that increases the marginal returns to x by Δ at time $t = 1$, so $\alpha_t = \alpha_0 \; \forall t \leq 0$ and $\alpha_t = \alpha_0 + \Delta \; \forall t > 0$, then in equilibrium, the immediate change in x is given by $\bar{x}_1 - \bar{x}_0 = \frac{\Delta}{\beta}$, while the long-run change is given by $\bar{x}_\infty - \bar{x}_0 = \frac{\Delta}{\beta\theta}$. This implies that an RD estimator comparing cohorts 0 and 1 will under-estimate the long-run effect of the policy change on x by a factor of θ.

interest, instruments for schooling, and the measure of schooling used in the analysis (e.g. years of schooling). Two sets of estimated effects of schooling on health/mortality are reported: (i) regression-based estimates (including probits) and (ii) IV or RD estimates that attempt to address concerns about endogeneity or unobserved heterogeneity. Table 4 describes studies that estimate the effect of schooling on mortality, while Tables 5 and 6 describe studies focused on other health outcomes in the U.S. (Table 5) and Europe (Table 6).

Our discussion is organized largely around Table 7, which categorizes selected estimates by general health outcome or behavior across all countries. This table further reports the magnitude of estimated effects of schooling (using IV or RD strategies) as a percent of base outcome levels (e.g. percent reductions in mortality or smoking).

Lleras-Muney (2005, 2006) provides the first serious attempt to estimate the causal effect of schooling on mortality. Using U.S. Census data from 1960, 1970, and 1980, she analyzes ten-year mortality rates among native-born whites born 1901-1925. She first calculates mortality rates by gender, year of birth and state of birth using consecutive Censuses. Then, she combines these mortality rates with average years of completed schooling and compulsory schooling ages (both by gender, year of birth and state of birth) to estimate the effect of an additional year of education on mortality using standard two sample IV strategies.[65] As reported in Table 4, her IV specification suggests that an additional year of schooling reduces ten-year mortality rates by about 6 percentage points (an effect that is statistically significant and nearly twice the effect estimated by OLS). This estimate is almost certainly greater than the true effect, since it implies a nearly 60% reduction per year of school given baseline ten-year mortality rates of 0.11 in her sample.

Because Lleras-Muney (2005, 2006) controls for region-specific cohort trends, her estimates are identified from state-level deviations from regional trends in schooling and mortality associated with changes in state compulsory schooling ages. However, Mazumder (2008) notes that if states with larger increases in compulsory schooling ages also had greater secular improvements in life expectancy (relative to the regional average), then failing to control for state-specific trends will cause her to overestimate the effect of schooling on mortality. Using an expanded sample to include 1960-2000 U.S. Censuses (while continuing to study native whites born 1901-25), Mazumder (2008, 2010) shows that controlling for state-specific cohort trends (rather than just region-specific trends) produces negligible estimated effects of education on mortality.[66]

Using U.S. data on disease-specific mortality, Glied and Lleras-Muney (2008) estimate that in-

---

[65]Here main specification controls for gender, cohort, state of birth, and Census year fixed effects. She also controls for region-specific cohort trends and observed cohort-specific state-of-birth characteristics (e.g. percent urban, percent black, percent employed in manufacturing, average manufacturing wage, etc.). Like Lochner and Moretti (2004), she links cohorts to compulsory schooling ages in their state of birth when they were age 14.

[66]Mazumder (2008, 2010) further shows that the key difference in estimated effects arises due to the inclusion of state-specific cohort trends and not the addition of 1990 and 2000 Census data. Baseline mortality rates are considerably higher in Mazumder (2008, 2010) than in Lleras-Muney (2005, 2006), because the 1901-25 birth cohorts are older at the time of the more recent Censuses.

creases in compulsory schooling ages led to greater reductions in death rates for diseases that experienced greater technological improvements in recent years, whether the latter is measured by the number of new active drug ingredients or disease-specific declines in mortality in recent years. Given the absence of state- or region-specific cohort trends in their specification, it seems likely that they over-estimate the average effect of schooling on mortality. However, if state-specific cohort trends in mortality do not vary systematically across diseases, then differences in mortality reductions across different diseases should be consistently estimated. In this case, their finding that mortality reductions associated with schooling are greatest for diseases experiencing faster technological improvements is more credible. We return to this issue below.

Two recent international studies, one in the U.K. (Clark and Royer 2010) and the other in France (Albouy and Lequien 2009), use national increases in compulsory schooling ages and RD designs to estimate the effects of schooling on mortality. They effectively estimate changes in educational attainment and mortality for the first cohorts affected by increases in compulsory schooling ages relative to the last cohorts before the reform. Both studies estimate statistically insignificant effects: estimates by Albouy and Lequien (2009) are modest in size, while those of Clark and Royer (2010) suggest that education actually *increases* mortality.[67]

A growing literature studies the effects of education on self-reported health status, disability, mobility, and activity. Mazumder (2008) uses data from the Survey of Income and Program Participation (SIPP) from 1984-96 to estimate the effects of schooling on self-reported health for native-born white Americans born 1901-25. As with mortality, he instruments for schooling using state-level compulsory schooling laws, controlling for state-specific cohort trends. He estimates sizeable and statistically significant effects of an additional year of school on the probability someone reports being in fair or poor health (23% reduction) as well as on the probability someone reports a health limitation (17.5% reduction). Estimates focusing on specific limitations suggest that schooling primarily reduces problems with seeing, hearing, and speaking; estimated effects on walking, climbing stairs, or lifting are small and statistically insignificant. Using data from 1950-2000 U.S. Censuses and changes in state-level compulsory schooling laws as instruments for educational attainment, Oreopoulos (2006) estimates that an additional year of school reduces disabilities limiting personal care or mobility by about 30% among 25-84 year-old Americans.[68]

A number of recent studies examine the impacts of education on self-reported health and limitations

---

[67]Clark and Royer (2010) do not actually estimate the effect of schooling on mortality. Instead, they estimate the reduced form effect of changes in the compulsory schooling age on mortality and (using a slightly different specification) the effect of compulsory schooling age on years of completed schooling. Their estimates suggest that the increase in compulsory schooling increased both education and mortality rates. For comparison purposes, Tables 4 and 7 report the results from dividing the effect of the 1947 reform on mortality by the reform's effect on educational attainment to arrive at an 'IV estimate' (in the spirit of a Wald estimator).

[68]His sample includes Americans born 1901-56. His specifications control for general national cohort and time trends, but they do not account for differences in health trends across states. Like Lleras-Muney (2005), he does control for time-varying state-level demographic and labor market characteristics.

43

in Europe using changes in compulsory schooling laws. Silles (2009) uses the General Household Survey for England, Scotland and Wales to estimate the effect of education on self-reported health in the U.K. Using an RD design around national increases in compulsory schooling ages in 1947 and 1973, she estimates that an extra year of secondary school increases self-reported good health by about 10%, a modest effect about half as large as Mazumder's estimate for self-reported bad health in the U.S. Clark and Royer (2010) exploit the same increases in the compulsory schooling age (also with an RD design) using data from the Health Survey of England (1991-2004). They estimate small and statistically insignificant effects of education on self-reported health, long-term illness, and physical activity. Using data from the General Household Survey for England and its Northern Ireland counterpart, Oreopoulos (2006, 2008) reaches a similar conclusion about the effect of education on self-reported health in the U.K. His analysis focuses on the 1947 increase in the compulsory schooling age in England along with a similar increase in Northern Ireland that took place ten years later.[69] Kemptner, Jurges and Reinhold (2010) estimate the effect of increasing minimum schooling levels in (formerly West) Germany from 8 to 9 years using German Microcensus data. Interestingly, they find that education significantly reduces reported long-term illness (by about 20%) among men but not women. Point estimates for women are positive and sizeable, but less precisely measured.

Based on the studies summarized in Table 7, education appears to have a weaker effect on mortality, self-reported health, and physical activity in Europe than in the United States. This is not because education and health are unrelated in Europe – indeed, some of the OLS results in Tables 5 and 6 imply a stronger relationship in Europe. It is tempting to speculate that the effects of education depend on access to health care, general social welfare and unemployment policies, and the level of overall inequality.

While most studies examine self-reported measures of health, Powdthavee (2010) exploits the Health Survey of England (1991-2007) and changes in compulsory schooling in the U.K. to estimate the effects of education on hypertension as determined from blood pressure measurements (see Table 6). IV-probit estimates exploiting the 1947 increase in the compulsory schooling age (from 14 to 15 years), suggest that an extra year of schooling reduces hypertension by slightly more than 10% for both men and women. Unfortunately, RD estimates using the 1973 increase in the compulsory schooling age (from 15 to 16 years) are much more imprecise, yielding positive but statistically insignificant point estimates. Aside from this study, there is little evidence on the effects of education on objective measures of health.

The final panel of Table 7 reports estimates of the effects of education on smoking and obesity, two important contributors to a number of serious and chronic health problems. For these outcomes, studies mostly agree: education significantly reduces smoking but has negligible effects on obesity.[70]

---

[69]His specifications allow for general cohort trends; however, he assumes that those trends are the same for England and Northern Ireland.

[70]Many of these studies also find that education increases the probability of quitting smoking conditional on having

Clark and Royer (2010) estimate that an additional year of schooling reduces the probability someone is currently smoking by about 13% in England, very similar to estimates for men in the U.S. by de Walque (2007a). Unlike all other studies discussed thus far, de Walque (2007a) identifies the effect of post-secondary schooling rather than primary or secondary schooling. This is because he uses the risk of being drafted for the Vietnam War to instrument for schooling attainment.[71] Grimard and Parent (2007) also exploit the effect of the Vietnam draft on male college attendance decisions for cohorts born between 1945 and 1950. Their estimates suggest that additional post-secondary schooling has even greater effects on rates of frequent smoking, reducing the likelihood that someone smokes every day by roughly 40%.[72] Kenkel, et al. (2006) estimate the effects of receiving a GED or finishing high school on smoking for NLSY79 respondents in their late 30s and early 40s using state-specific high school and GED requirements and policies as instruments (e.g. the number of courses required to complete high school, an indicator for whether local districts set graduation requirements, the fraction of youth taking the GED test, an index of GED policies related to score requirements and fees, and the log of per capital education spending averaged over primary and secondary school ages). They estimate very large effects of high school graduation on smoking, with effects for women especially large and statistically significant (88% reduction). Estimated effects of GED receipt are much weaker (not reported here).[73] In contrast to previous studies, Kemptner, Jurges, and Reinhold (2010) estimate relatively small and insignificant effects of schooling on smoking in Germany. This difference may be due to the cohorts analyzed in their study. In exploiting increases in compulsory schooling (from 8 to 9 years) from 1949-69, they largely identify the effects of schooling on cohorts who would have made their initial smoking decisions before the harms of smoking were well-known. Among these early cohorts, smoking was only weakly correlated with education (e.g. see the relatively small OLS estimates for this study relative to others in Tables 5 and 6).[74] By this reasoning, the relatively large estimated effects of schooling on smoking by Kenkel, et al. (2006) may be explained by the fact that these individuals would all have made their smoking decisions long after the 1964 U.S. Surgeon General widely publicized the dangers of smoking. Consistent with the model above, the effects of education on smoking seem to depend on the perceived impacts of smoking on health.

The estimated effects of education on weight problems in Europe and the United States are uni-

---

smoked; although, this is not as easily interpreted since education affects the likelihood of starting smoking.

[71]In some specifications, he also uses the risk of induction multiplied by the probability of being killed to instrument for schooling. When he controls for the endogeneity of veteran status, de Walque (2007a) estimates even greater effects of education on smoking than reported in Tables 5 and 7.

[72]Grimard and Parent (2007) and de Walque (2007) effectively identify the effects of college for men who chose to avoid the draft. Since these men may be more risk averse (or may value their survival more) than the average male, the effect of college-going on smoking for them may differ from that of a more representative population.

[73]One concern with this study is that differences in education policies or expenditures across states and over time may affect skill levels conditional on educational attainment. That is, it may mean something different to finish high school in a state that spends a lot on secondary schools and has high graduation requirements relative to a state with low expenditures and weaker standards. This would, in general, violate their exogeneity assumptions; however, Kenkel, et al. (2006) test and do not reject their exclusion restrictions in their context, so this may not be a serious problem.

[74]See de Walque (2010) for a detailed analysis of how smoking differences by education evolved over time in the U.S.

formly small. None of the studies we survey report statistically significant effects on obesity (body mass index, BMI, of 30 or more); however, a few find small but statistically significant effects on the probability that someone is overweight (BMI of 25 or more). See Table 7.

Tables 5 and 6 do not include a few earlier IV studies of the effects of education on health; however, there are good reasons to question the exogeneity of at least some of their instruments.[75] Both Auld and Sidhu (2005) and Adams (2002) use parental background as instruments; however, these instruments have generally been dismissed as endogenous in the literature on schooling and earnings (e.g., see Carneiro and Heckman (2002)). Furthermore, Kenkel, et al. (2006) reject the validity of parental education levels as instruments in estimating the effects of education on smoking and obesity. Adams (2002) also uses quarter of birth as an instrument for schooling; however, a large literature on weak instruments suggests that this may produce IV estimates biased towards OLS estimates (see, e.g., Bound, Jaeger, and Baker (1995)). Arkes (2003) and Auld and Sidhu (2005) use local unemployment rates when individuals are in their teenage years to instrument for educational attainment; however, the strong serial correlation in local labor market conditions over time may lead to bias as mobility rates are often low and contemporaneous local labor market conditions may affect health through a variety of channels. Cameron and Taber (2004) show that, at least in the estimation of earnings equations, it is important to control for current local labor market conditions when using past conditions as an instrument for schooling. Finally, the Danish study by Arendt (2005) uses two national school reforms, which both raised compulsory schooling ages and affected the nature of schooling (e.g. removal of student tracking based on tests, convergence in curricula between village and city schools). The latter raises similar concerns to those with Kenkel, et al. (2006). Even ignoring any concerns about endogeneity, the IV estimates in Arendt (2005) are extremely imprecise.

Conti, Heckman and Urzua (2010a, 2010b) depart from the standard IV/RD approach in studying the relationship between education and health. They estimate a multi-factor model of schooling, earnings, and health outcomes using data from the British Cohort Study, which has followed all babies born in the U.K. during one week in 1970 periodically through age 30. Their model assumes that education, post-school earnings, and health behaviors/outcomes depend on family background characteristics, age 10 observable health endowments, and three unobserved latent factors (cognitive ability, non-cognitive ability, and an unobserved health endowment). Because their data contain multiple measurements of the three (correlated) unobserved factors, they are able to estimate the joint distribution of these factors as well as their effects on health behaviors and outcomes. They focus attention on the impacts of completing mandatory schooling (through age 16) on smoking, obesity, and self-reported health (fair or bad). Their estimates suggest that the causal effect of education explains

---

[75]These tables also exclude a few recent IV studies exploiting schooling reforms (e.g. Grabner (2009), Jurges, Kruk and Reinhold (2009), and Reinhold and Jurges (2010)). While the identification strategies in these analyses are reasonably credible, the reported IV estimates are very imprecise and unable to rule out a large range of effects (including the estimates reported in Tables 5 and 6).

60-70% of the raw differences in smoking by educational attainment, 35-55% of the raw differences in self-reported health, and one-third of the differences in obesity for men (none of the difference for women). Thus, education has important causal effects on smoking and self-reported health for men and women, but it only reduces obesity rates (by a fairly small amount) for men. These findings are largely consistent with the evidence presented in Table 7.

The approach of Conti, Heckman and Urzua (2010a, 2010b) enables them to draw a number of other interesting conclusions. First, they show that the effects of education are greatest for individuals with relatively high cognitive skills and/or low non-cognitive skills. Second, they show that cognitive ability is not a very important determinant of smoking decisions or obesity. There is a modest effect of cognitive ability on self-reported health for women, but little effect for men. Finally, they estimate that non-cognitive skills are generally more important for smoking, obesity, and self-reported health. In fact, non-cognitive skills are roughly as important as early health endowments in explaining these outcomes at age 30.

Parental Education and Child Health

There is a growing literature examining the effects of parental education on child (often infant) health and mortality. We briefly discuss a few recent studies that exploit IV and RD techniques to address concerns about unobserved heterogeneity. Currie and Moretti (2003) provide the first evidence that maternal education causally affects maternal pre-natal behavior and birth outcomes. They exploit openings of two- and four-year colleges by county in the U.S. from 1940-96 as instruments for maternal education. These college openings increased average years spent in college among their sample of 25-45 year-old women considerably (roughly one year for four-year colleges and 0.15-0.20 years for two-year colleges). Relative to a high school graduate, their IV estimates suggest that an extra year of college reduces the likelihood of a low birthweight child by about 20% and a pre-term birth by about 15%. Examining maternal behaviors, they estimate that an extra year of college reduces smoking during pregnancy by roughly one-third but only increases the incidence of pre-natal care by 3%. The impact on maternal smoking is a likely channel for the improvements in birth outcomes. McCrary and Royer (2009) exploit a very different source of exogenous variation in maternal schooling: school entry age laws. State laws often specify that children must be age five by a certain date before entering kindergarten. Using data from Texas (1989-2001) and California (1989-2002), they show that these laws induce a discontinuity in schooling attainment for women in their sample (age 23 or less) based on date of birth. In contrast to Currie and Moretti (2003), their RD estimates suggest negligible effects of schooling on low-birthweight, pre-maturity, and infant mortality. In most cases, their estimates are precise enough to rule out impacts as large as those in Currie and Moretti (2003). They also find little evidence to suggest that schooling impacts maternal smoking decisions; however, they find some evidence that it improves pre-natal care in Texas.

Lindeboom, et al. (2009) use the U.K.'s 1947 increase in the compulsory schooling age (from 14 to 15) discussed above to estimate the effect of maternal and paternal education on child health outcomes in the U.K. Using National Child Development Study data on all births from March 3-9, 1958, they estimate small and statistically insignificant effects of parental education on birth outcomes (low birthweight, illness at birth) as well as child health outcomes (chronic conditions, overweight) at ages 7-16; however, standard errors for their estimates are sizeable and cannot generally rule out impacts of 20% or more. Finally, Chou, et al. (2010) exploit the rapid construction of junior high schools in Taiwan from 1968 to 1973. Using variation in openings across counties as instruments for parental education, they estimate very large effects on birth outcomes. Their estimates suggest that, on average, the expansion of junior high schools increased schooling levels by 0.11 to 0.16 years and reduced the incidence of low birthweight children by 5% and infant mortality by 8-19%. The implied effects of an additional year of schooling are enormous – much larger than those of Currie and Moretti (2003).

Given the substantial differences in impacts estimated by these studies, it is difficult to draw any strong conclusions about the effects of parental education on child health. Looking only at the U.S. and U.K. studies, it is tempting to conclude that, in developed countries at least, additional college education improves child outcomes, while additional years of high school do not. Furthermore, the findings of Chou, et al. (2010) suggest that educational attainment may be more important in a developing economy like Taiwan, even at lower levels of education. While these tentative conclusions seem reasonable, there are also important differences in the estimation strategies employed by these studies that complicates any comparison of results. First, the exogenous sources of variation that affect schooling also tend to affect other factors that may influence child health. For example, the variation in school entry ages used by McCrary and Royer (2009) also leads to differences in age-for-grade. Students who start a year later are a year older than their counterparts during every grade at school. This may affect success in school beyond the number of years attended; it may also affect the timing of marriage and fertility or the choice of spouse. In this case, estimated effects of the policy on children will reflect the influence of all of these factors in addition to changes in parental education. Second, additional schooling is likely to affect the timing of fertility decisions. The aforementioned studies typically control for maternal age, but this may be an important channel through which schooling affects infant health and mortality. One final issue that complicates the interpretation of most of these studies is the role of marriage markets and both parents' education. On the one hand, both parents' education may be important, yet it is often difficult to account for both due to data limitations and the strong correlation between the two (Lindeboom, et al. (2009) is unusual in that it does separately estimate the effects of both parents' education). Many of the policies used as instruments are likely to affect both parents' education to different degrees, so examining the effects of only one parent's education likely over-states that parent's role. These policies are also likely to affect marriage markets,

with effects spilling over to other cohorts/groups assumed to be unaffected by the policies. All of these issues make estimation of the effects of parental education on child health much more complicated than the analysis of education's impacts on one's own health. Future research on this topic will need to grapple with these in terms of estimation and the interpretation of results.

3.2.2 Why does education affect health and mortality?

To the extent that education does have small to modest effects on health and mortality, what are the key operating channels or mediating factors? Education may improve decisionmaking abilities, which may lead to better health decisions and more efficient use of health inputs (productive efficiency). More educated individuals may also be better at acquiring or interpreting information from doctors, other individuals, or the press/internet. This, too, may lead to better decisions about health, including more appropriate treatments (allocative efficiency). Economists have devoted considerable attention to the distinction between productive efficiency (better health production given a set of inputs) and allocative efficiency (better allocation of health inputs) as it pertains to the relationship between education and health.

There are many other channels through which education may affect health and mortality as shown in Table 8, which also lists any personal costs associated with these channels and whether associated benefits are private (accruing to the individual who receives more schooling) or public (accruing to others). In addition to improving the production of health via productive or allocative efficiency, education may (costlessly) improve health by alleviating stress, as emphasized by some social epidemiologists. This may be due to improvements in social standing or socioeconomic well-being that result in fewer stressful situations or less long-term stress (e.g. Marmot 2004, Marmot and Wilkinson 2006).

Education may also lead to better health outcomes, because it raises income levels. As the model above shows, this increases the demand for better health and may cause individuals to spend more to remain healthy. More educated individuals are more likely to purchase better health insurance or to spend more on costly treatments.[76] More educated individuals may also choose healthier lifestyles. This may include taking greater safety precautions (e.g. wearing seat belts, installing smoke alarms), eating a healthier diet, exercising more, drinking alcohol in moderation, or smoking less. In fact, Leigh (1983) estimates that much of the positive effect of schooling on health can be explained by exercise, smoking, and occupational choices; however, more recent studies have reached a different conclusion as discussed below. While it is possible that education directly modifies preferences towards these activities, it seems more likely that causal effects of education on these behaviors are due to costly behavioral changes or to changes in peers, friends, or neighbors.[77] If, in the absence of health benefits,

---

[76]Recent quasi-experimental studies estimate significant health benefits from health insurance across a wide range of outcomes. See McWilliams (2009) for a review of this literature.

[77]Education may be correlated with these behaviors due to underlying heterogeneity in preferences as well (e.g. indi-

people would prefer to eat fatty foods, drink, smoke, and drive without a seat belt, then foregoing these activities is costly. Similarly, more educated individuals may choose healthier or safer occupations and employers. As the literature on compensating differentials (Rosen 1986) reminds us, these 'benefits' come at a cost, usually lower wages. More educated individuals may also choose to pay higher housing prices to live in neighborhoods that offer a healthier environment (e.g. less pollution, traffic, etc.).[78] In all of these cases, health benefits from increased education are effectively paid for, at least on the margin.

Finally, education may affect co-workers and the friends one interacts with on a frequent basis. To the extent that more educated persons interact more with other educated persons, effects of education on health may be increased through the desire to conform or the spread of health-related information. These 'social multiplier' effects are essentially costless and provide public benefits not taken into account when schooling decision are made.

Empirically, economists have largely relied on two broad approaches for identifying the channels through which education may reduce mortality or improve health. One approach examines whether controlling for a number of potential mediating factors (e.g. income, occupation, health knowledge) affects estimates of the education – health gradient. Put another way, these studies explore the extent to which the correlation between education and health can be explained by different factors. A second approach focuses more on testing or measuring the importance of specific factors. For example, Glied and Lleras-Muney (2008) examine which diseases are most affected by schooling to see whether differences in technological innovation are important for the education – mortality gradient. Other studies examine changes in knowledge or behavior in response to new public information (e.g. information that smoking is unhealthy). We briefly discuss the strengths and limitations of these approaches and review the findings from more recent or influential studies.

Numerous studies from a broad range of social scientists have explored the extent to which the relationship between education and mortality or health can be explained by the well-known differences in income (or economic well-being more generally) associated with schooling. This approach typically begins by regressing some measure of health (including mortality) on education and a few standard demographic characteristics (e.g. race, gender, age). Then, a measure of income or economic well-being is included in the regression, and the extent to which the coefficient on education declines (towards zero) is taken to be the effect of education 'explained' or 'mediated' by income. We discuss two of these

---

viduals with more patience or self-control may prefer a healthier lifestyle and to acquire more schooling even if education itself does not alter these behaviors). Tables 5-7 suggest a significant causal effect of schooling on smoking but not obesity.

[78]See Chay and Greenstone (2005) for evidence on the effects of air quality on housing prices. Chay and Greenstone (2003) and Currie and Neidell (2005) provide evidence that air pollution significantly increases child mortality; however, effects on adult mortality appear to be weaker (Chay, Dobkin and Greenstone 2003). Moretti and Neidell (2009) find evidence of important short-run effects of pollution on health. Lleras-Muney (2010) estimates that ozone pollution significantly increases hospitalizations among military children ages 2-5. Neidell (2004) estimates important adverse effects of pollution on child health, with larger effects on children of lower socioeconomic status.

studies, which extend this approach to study the importance of a broad range of mediating factors.[79]

Cutler and Lleras-Muney (2010) explore the importance of the following mediating factors on health behavior: (1) income and economic resources, (2) personality, self-esteem, and sense of control, (3) social integration, (4) health-specific knowledge, (5) cognitive ability/skill, (6) preferences for risk and time discounting. Their analysis sheds light on the role of several mechanisms in affecting health and mortality through important health and lifestyle choices. Because they look at such a wide array of mediating factors, they exploit numerous data sources in the U.S. (the National Health Interview Survey, NHIS, National Survey of Midlife Development in the United States, MIDUS, the Health and Retirement Survey, HRS, National Longitudinal Survey of Youth, NLSY, and the Survey on Smoking SOS) as well as the National Childhood Development Study (NCDS) in the U.K.

While Cutler and Lleras-Muney (2010) estimate the relationship between education and numerous health behaviors (smoking, diet/exercise, alcohol consumption, and preventive care), it is useful to summarize these relationships based on the importance of different behaviors for actual health outcomes. They, therefore, calculate a 'mortality weighted average effect' of the relationship between education and smoking, drinking, and obesity (three of the more important and oft-measured behaviors). This summary measure weights education – behavior relationships in proportion to their likely impacts on mortality. We briefly summarize these 'mortality weighted average effects'. Cutler and Lleras-Muney (2010) first estimate the extent to which differences in household income (and economic resources more generally) can account for the correlation between education and health behaviors. Their estimates suggest that economic resources can account for 11% to 32% of this correlation in the U.S. and U.K. Next, Cutler and Lleras-Muney (2010) control for other potential mediators along with household income to explore their additional influence. Of the other types of mediators, only cognitive ability/skills and social integration appear to be important. Using the NLSY, they find that cognitive ability, as measured by a battery of aptitude tests taken during late teenage years, explains about 15% of the correlation between education and health behaviors.[80] Economic resources and cognitive ability/skills combine to explain roughly 30% of this correlation in the NLSY. In the MIDUS data, social integration explains 7% of the education – health behavior correlation after accounting for income; together income and social integration account for nearly 20% of the correlation. The strongest effects of cognitive ability/skills and social integration comes from the NCDS in the U.K., where ability accounts for 44% and social integration 15% of the correlation.[81] Simultaneously controlling for economic resources, cognitive ability/skills, social integration, time preference, and sense of control account for 72% of the correlation between education and health behaviors in the NCDS;

---

[79]See Cutler and Lleras-Muney (2006) for a recent review of this literature.

[80]All respondents in the NLSY were given a battery of ten tests measuring such skills as math, science, reading comprehension, word knowledge, and electronics. These tests make up the Armed Services Vocational Aptitude Battery, four of which go into the well-known Armed Forces Qualifying Test (AFQT).

[81]Ability in the NCDS is based on drawing, reading and math tests administered at ages 7, 11, and 16; although, Cutler and Lleras-Muney (2010) show that scores at ages 7 are not very important.

however, nearly all of this effect can be accounted for by income and cognitive ability/skills alone. In all specifications that explore the role of preferences for risk, time discounting, self-efficacy, or sense of control, these factors are found to explain little if any of the relationship between education and health behaviors. Lastly, Cutler and Lleras-Muney (2010) explore whether individuals who are better informed about the health impacts of smoking and drinking are less likely to smoke and drink, and whether this can explain any of the correlation between education and these unhealthy habits. Their findings are mixed. While knowledge about the harms of smoking appears to explain about 15% of the relationship between education and smoking, knowledge about the harms of drinking do not explain differences in drinking by education.

The main conclusion from Cutler and Lleras-Muney (2010) is that both income and cognitive ability/skills play an important role in explaining the relationship between education and health behaviors. These two factors combine to explain roughly 30-50% of the education – health behavior correlation in the U.S. and nearly 70% of the correlation in the U.K.[82] Social integration plays a modest role, and all other potential mediating factors appear to be relatively unimportant.

While commonly measured health behaviors are important determinants of health and mortality, previous studies report that smoking, drinking, exercise, and other health behaviors can only explain about 30% of the relationship between education and health (e.g. Marmot 1994, Cutler and Lleras-Muney 2006). Furthermore, examining the effects of mediating factors on health behaviors does not necessarily inform us about the full importance of these mediating factors for health status or mortality, since some of these mediators may affect health through other channels. For example, increases in income may improve health by reducing stress or by allowing individuals to move to safer and healthier neighborhoods.

Ross and Wu (1995) directly examine the importance of the following mediating factors for the relationship between education and health: (1) work and economic resources, (2) social/psychological resources, and (3) health lifestyle. The first two are also analyzed in Cutler and Lleras-Muney (2010); however, the third factor is treated very differently in the two studies. Unlike Cutler and Lleras-Muney (2010), who examine the effects of other factors on health behaviors, Ross and Wu (1995) consider the mediating effect of health-related behaviors on health itself. Using data from the 1990 Work, Family, and Well-Being (WFW) Survey and the 1979-80 National Survey of Personal Health Practices and Consequences (HP), Ross and Wu (1995) examine the extent to which these factors explain the relationship between education and both self-reported health and physical functioning.[83] In both data

---

[82]It is not clear whether the difference between the U.S. and the U.K. is real or due to relatively more accurate measures of economic resources and ability in the U.K.'s NCDS. Furthermore, Smith (1999, 2007) shows that correlations between income and health are at least partially due to the adverse effects of disease onset on income and labor supply. The extent of this 'reverse causality' may differ across countries.

[83]Measures for work and economic conditions include household income, employment status, economic hardship, and work fulfillment (WFW only). Social-psychological resource measures include sense of control and social support. Health lifestyle measures include self-reported exercise, smoking, drinking (HP only), and health checkups (HP only). Controls for sex, race, age, and marital status are included in all specifications.

sets, education is significantly associated with better self-reported health and physical functioning. Assuming work and economic resources may affect social/psychological resources, which may in turn, influence health-related habits, Ross and Wu (1995) first add controls for work and economic resources, then controls for social-psychological resources, and finally controls for health behaviors. Consistent with earlier studies examining the relationship between education, income, and mortality (e.g. Elo and Preston (1996), Kitagawa and Hauser (1973)) they find that economic resources explain 42-46% of the correlation between education and self-reported health and 34-54% of education's association with physical functioning. By comparison, the roles played by both social-psychological factors and health habits are much weaker, each explaining less than 10% of the association between education and either health measure. Ross and Wu (1995) estimate that all three broad classes of factors combine to explain 46-59% (66-71%) of the association between education and self-reported health (physical functioning).

The finding of Ross and Wu (1995) that health behaviors explain little of the overall education – health relationship suggests that the decomposition of the education – health behavior gradient by Cutler and Lleras-Muney (2010) is limited in its implications for health outcomes. Consistent with this finding, Cutler, et al. (2010) show that despite sizeable cross-sectional differences in mortality risk factors (e.g. smoking, obesity, high blood pressure, high cholesterol) by education, trends in these risk factors explain little of the observed widening in mortality gaps (generally and specifically for cancer and cardiovascular diseases) by educational attainment.

Altogether, these 'decomposition studies' offer a few lessons but are not without caveats. Taken at face value, these studies suggest that income and ability combined are important for explaining differences in health behaviors or outcomes by education, while differences in preferences, specific health knowledge and psycho-social factors are relatively unimportant. It should be noted, however, that evidence from Cutler and Lleras-Muney (2010) on the importance of ability relative to income is more difficult to interpret. This is because they control for income when examining the importance of ability, but they do not control for ability when estimating the importance of income. As shown in the Appendix, this ordering should be reversed since post-school income is strongly influenced by adolescent ability conditional on education (e.g., see Cawley, et al. 2000).[84] Despite the suggestion by Cutler and Lleras-Muney (2010) that income is more important, the reverse is also quite possible. The fact that other studies (e.g. Ross and Wu 1995, Elo and Preston 1996, Kitagawa and Hauser 1973) also find that income plays an important mediating role does not alleviate this concern, since they too neglect measures of ability in their analysis. Conti, Heckman and Urzua (2010a, 2010b) argue that non-cognitive skills may be more important factors in explaining education – health gradients than

---

[84]In general, if one is interested in the importance of some mediating factor x in explaining the relationship between education and health, then it is important to control for any additional mediating factors that may affect both x and health directly. By contrast, one would not want to control for additional mediating factors affected by x if they do not also affect x themselves. See the Appendix for details.

cognitive skills, further complicating matters. Altogether, it appears that education has important effects on health behaviors, health outcomes, and mortality via its combined effects on income and ability (both cognitive and non-cognitive).[85] A second lesson from Ross and Wu (1995) and Cutler, et al. (2010) is that differences in health behaviors (e.g. smoking, drinking, obesity) by education explain little of the education – health gradient or its evolution in recent decades. An important caveat cannot be overlooked. None of these studies address issues of endogeneity, nor do they simultaneously control for all factors that may directly affect health and which may be correlated with potential mediators. These studies aim to decompose the correlation between education and health, but the relative importance of different factors in terms of correlation may not reflect the relative importance in terms of causal effects.

Rather than decompose education – health gradients into a variety of potential mediating factors or explanations, a number of other recent studies have focused on estimating and testing the implications of specific channels, including the role of productive efficiency, allocative efficiency, and peer effects/social networks. In drawing conclusions from these types of studies, it is important not to focus too much on any one mechanism or disease in trying to understand the link between education and health. Link and Phelan (1995) forcefully and convincingly argue that education and socioeconomic differences are 'fundamental causes' of health disparities. That is, these factors affect the financial resources, knowledge, and access to power, which determine the ability and desire of individuals to avoid risks and address health problems. The mechanisms and diseases that lead to health disparities by socioeconomic status continually evolve over time, arguing against a narrow research or policy focus on any one mechanism or disease. Instead, it is important to interpret the body of evidence in light of more basic theories about the role of financial resources, differential costs, and information. We organize our discussion of this evidence along these lines.

If education improves productive efficiency, more educated individuals should (i) be healthier conditional on all health inputs and (ii) use fewer inputs conditional on input prices and health outcomes (Grossman 1972a, 1972b). Earlier studies by health economists (e.g. Grossman 1972b, Wagstaff 1986, Erbsland, Ried and Ulrich 1995, and Gilleskie and Harrison 1998) largely focused on measuring the role of productive efficiency based on these two implications. Findings from these studies are generally supportive of productive efficiency. Unfortunately, these methods require good measures of all health inputs or input prices. If some inputs or prices are missing, it is extremely difficult to estimate the role of productive efficiency, since unobserved inputs or prices are likely to be

---

[85]Smith (1999, 2007) argues that changes in income have relatively small effects on the onset of major health diseases; instead, the onset of disease tends to adversely affect subsequent income. This might suggest that health plays a 'mediating' role in determining the effects of education on income and wealth. In decompositions of the relationship between education on health, this would incorrectly be attributed to a mediating role of income. Of course, the correlation between income and health may be largely due to differences in treatment and management of disease by income rather than differences in the likelihood of disease onset as emphasized by Smith (1999, 2007). In this case, education – health decompositions are likely to measure the mediating effect of income.

correlated with education. In practice, these approaches also typically rely on strong functional form assumptions about the health production function — assumptions that are difficult to verify.

Kenkel (1991) is the first to directly examine the importance of allocative efficiency by estimating whether specific health knowledge affects behavior. Using the same decomposition approach as described earlier, he estimates whether controlling for knowledge about the harms of smoking, drinking, and exercise reduces the correlation between education and those behaviors. His estimates suggest that both knowledge and education affect smoking, heavy drinking, and exercise; however, differences in health knowledge explain less than 20% of the correlation between education and those behaviors (controlling for income, employment and standard demographic factors). Thus, health-specific knowledge matters, but it does not explain much of the difference in health behavior. This finding is echoed in the more recent study by Cutler and Lleras-Muney (2010).

A few recent studies have taken a more indirect approach to evaluating the role of health-specific knowledge and allocative efficiency. These studies estimate the extent to which new public health information affects the relationship between education and specific health behaviors. De Walque (2010) and Aizer and Stroud (2010) show that education – smoking gradients increased substantially in the U.S. after the 1964 Surgeon General's Report announced the dangers of smoking. De Walque (2007b) finds that the education – HIV prevalence gradient evolves in a similar way following a major HIV/AIDS prevention campaign in Uganda during the late 1980s and early 1990s. Price and Simon (2009) examine the impact of a medical research study published in 2001 which showed that the risks of complications from attempting a vaginal birth after a caesarean (VBAC) were higher than previously thought. Their analysis shows that VBACs declined sharply after publication of the article, with significantly larger declines among more educated women. In a variety of contexts (and countries), these studies show that new publicly provided information about health procedures or behaviors tends to widen gaps in those procedures/behaviors by education.

As Lange (forthcoming) notes, this type of evidence is consistent with two potential explanations: (1) more educated individuals may gather and process information more easily and/or (2) more educated individuals may respond more to new information because it is of greater value to them (e.g., see section 3.1).[86] Assuming information is exogenously acquired, Lange (forthcoming) shows that examining actual beliefs about the likelihood of a health outcome (or someone's risk level) may help identify whether explanation (1) is important. If more educated individuals are better informed, then their beliefs should be more strongly correlated with objective risks. However, the fact that more educated individuals may have a greater demand for health does not, by itself, make that information any better. If information is *exogenously* acquired, then there is no reason to expect more educated

---

[86]In some cases, it may be useful to distinguish between the potential for education to improve the ability to acquire information and the ability to process that information and make good decisions. In the former case, attempts to better distribute information to lesser educated individuals should be helpful, while the latter suggests efforts to simplify and explain the implications of new information may be needed.

individuals to have more accurate beliefs due to their greater demand for health. Based on this insight, Lange (forthcoming) estimates the correlation between individuals' beliefs about their risk of cancer and objective risk levels (based on demographics and family history). His estimates suggest that this correlation is stronger for more educated individuals. Consistent with this finding, Aizer and Stroud (2010) show that knowledge about the harms of smoking increased more quickly among the most educated following the 1964 Surgeon General's warning. Both Lange (forthcoming) and Aizer and Stroud (2010) conclude that those with more education are better informed and that this affects their health decisions (i.e. cancer screenings and smoking decisions). An obvious complication is the likely possibility that information about risks of cancer or smoking-related health problems is not exogenously acquired. Because more educated individuals have a greater demand for health generally, they are also likely to have a greater demand for accurate health information about their objective risks, causing them to invest more in acquiring that information. Thus, differences in endogenously acquired information should reflect differences in demand for health in the same way differences in health behaviors do.

Overall, these studies convincingly show that education affects the acquisition of health-related information. Whether or not improved information comes at a cost is more difficult to determine. Furthermore, does the information translate into important differences in health behaviors and, ultimately, health outcomes? Aizer and Stroud (2010) show that the widening of the smoking differential corresponds to the emergence of a knowledge gap about the effects of smoking on heart disease. Lange (forthcoming) also shows that more educated individuals are both better informed about their cancer risks and more likely to undergo screening for cancer. Despite these findings, evidence by Kenkel (1991) and Cutler and Lleras-Muney (2010) suggests that differences in health-specific knowledge by education account for only small differences in behavior. Furthermore, differences in behavior explain only a small fraction of differences in health (Ross and Wu 1995).

A number of recent studies explore the extent to which new technologies affect the relationship between education and health. Goldman and Smith (2002) estimate that more educated patients who suffer from diabetes or from HIV better self-manage their diseases by adhering to complex treatments. Most notably, they find that less-educated diabetes patients benefitted more (than more-educated patients) from an experimental treatment regimen with intensive patient monitoring. Because treatments were provided freely as part of the experiment, differences in income are unlikely to explain the differential responses by education. This strongly suggests that factors other than income (e.g. cognitive ability) may play an important role for difficult-to-manage diseases and that differences in these factors by education may help explain the education – health gradient. Unfortunately, it is difficult to know whether these other factors are as important outside the 'laboratory' where more intensive treatments also tend to be more costly.[87]

---

[87]Maitra (2010) estimates that the impacts of education on diabetes in the Health and Retirement Survey are largely

As noted earlier, Glied and Lleras-Muney (2008) find that more educated individuals experienced greater survival advantages for diseases that showed the fastest technological progress. They argue that more educated individuals may be more likely to adopt new treatments and medicines. However, Goldman and Smith (2005) find no evidence of differential diffusion of hypertension drugs by education over the 1980s and 1990s. Goldman and Lakdawalla (2005) show that the impacts of new treatments on education – health relationships are likely to depend on whether those treatments are more complicated or simpler than previous treatments. Consistent with Goldman and Smith (2002), they show that antiretroviral therapies for HIV introduced in the 1990s led to greater health improvements among more educated HIV patients. By contrast, the emergence of beta-blockers as a form of treatment for hypertension reduced education disparities in hypertensive cardiovascular disease. More generally, they point out that innovations in health practices, technologies, or pharmaceuticals that raise optimal treatment costs are likely to exacerbate health and mortality differences by education, while innovations that lower treatment costs should do the opposite. Of course, most new technologies tend to be more expensive than their predecessors, at least initially, suggesting that periods of rapid medical innovation are likely to be associated with widening education – health gradients.

It is difficult to know whether differences in disease management and medical treatments are due to differences in the ability to utilize best practices/treatments or to differences in demand for health. Do more educated individuals better self-manage their diseases, because it is easier for them to do so or because they have more to gain from putting forth the effort? Do less educated individuals fail to take the newest and most effective drugs, because they find the regimens difficult or because they are too expensive? Despite serious efforts taken in these studies to narrow the focus on a single mechanism, these open questions reveal the difficulty of the task at hand.

A final issue worth discussing is the spillover effects of education for health. There is some evidence regarding the peer effects associated with smoking and other health outcomes that speaks to the potential for education to have external effects on the health of others. A number of recent studies (e.g. Norton, Lindrooth and Ennett 1998, Gaviria and Raphael 2001, and Powell, Tauras, and Ross 2005, Fletcher 2010) use IV techniques to estimate the influence of peers on smoking and drinking among school-age adolescents. Most of these studies assume that school peers' family background or neighborhood characteristics (e.g. parental education, family composition, race) have no direct affect on whether someone engages in risky behaviors conditional on the individual's own family background and school characteristics. (They also implicitly assume that these neighborhood or parental characteristics are uncorrelated with unobserved factors that may affect youth's decisions.) These studies typically find statistically significant peer effects; however, the magnitudes vary considerably. Using class-level

---

due to differences in economic resources and health insurance coverage rather than differences in disease management or cognitive ability as emphasized by Goldman and Smith (2002). Maitra (2010) does not dispute the experimental evidence by Goldman and Smith (2001); however, his results may suggest that income and health insurance differences by education are more important for diabetes management once the actual costs of different treatments are factored in.

data, Fletcher (2010) shows that incorporating school fixed effects and allowing for some contextual effects of peer background characteristics on smoking reduces the estimated peer effects on smoking by about 30% relative to the approach taken in most previous studies. Krauth (2005) eschews the strong exogeneity assumptions employed by these studies and instead takes a more structural approach to bound the effects of peers on smoking with different assumptions about the correlation in unobserved tastes across individuals. His approach suggests that the effects of peers on smoking are likely to be much smaller than IV estimates suggest. Altogether, these studies suggest that peer effects for risky behaviors like smoking and drinking may amplify the effects of policies on youth; however, Cohen-Cole and Fletcher (2008) estimate insignificant effects of peer obesity levels on one's own obesity among middle and high school students in the U.S. Little is known about the extent to which peer effects and social networks are important for adults and their health behaviors. Given the evidence that differences in habits and health behaviors by education only explain a small share of the overall education – health gradient, it seems unlikely that these behavior spillover effects are economically very important. The most important spillover is likely to be from parental education to child health as discussed earlier.

In summary, evidence on the extent of productive efficiency is suggestive but comes with many caveats. Growing evidence more clearly suggests that education affects the allocation of health inputs and behaviors; however, it is still not clear whether these allocations are more efficient for more educated individuals or simply different due to differential demand for health. The fact that differences in income explain one-quarter to one-third of the correlation between education and health/mortality is consistent with a differential demand-based story. The fact that differences in ability (cognitive and/or non-cognitive) seem to matter suggests that productive and allocative efficiency may also be important. Peer effects appear to exist among adolescents for some risky behaviors like smoking and drinking, but there is no evidence to suggest that this type of externality is important among adults or for broader measures of health.

3.3   Implications for education and health policy

For public policy purposes, it is constructive to categorize potential channels according to the following: (i) Does the potential benefit to health come at some direct cost to the individual (other than through costs associated with schooling)? (ii) Does the health benefit come primarily from one's own schooling or does it result from the schooling of others?

If acquiring more schooling makes someone healthier or live longer, there is every reason to think individuals will take that into account when making their education decisions – if they are aware of the benefit and value it. If the benefits are private and known, then there is little case to be made (at least on efficiency grounds) for subsidizing schooling with the aim of improving health outcomes. On the other hand, if the benefits are unknown or are public in nature (e.g. peer effects), then

education subsidies may be efficient. Thus, it is clearly important to distinguish between channels implying public vs. private benefits. Most mechanisms that link education and health suggest that the benefits are private, or at least contained within the family. Empirical evidence on peer effects in health behaviors are limited to a few risky behaviors and adolescents. While these studies tend to find some role for peers, it is unlikely that the peer effects of education on health more generally are very important. There is some evidence that parental education affects child health outcomes. To the extent that individuals internalize benefits to their future offspring and recognize benefits to their own health when making education decisions, they have a strong incentive to optimally weigh those benefits against the costs of schooling.

If individuals are unaware of the health benefits of education (for themselves or their future children) at the time schooling decisions are made, it is important to distinguish between benefits that must be paid for and those that are freely acquired through schooling. If individuals are unaware of benefits that come without cost (e.g. lower stress, better at gathering/interpreting information, better decisionmaking), youth may acquire too little schooling. However, failure to fully consider future benefits that come at a personal cost (e.g. better health care or healthier lifestyle choices) is not likely to distort education decisions in the same way it is not important for youth to know how much they will spend on golf during retirement.

Distinguishing between costly and free health benefits of education may also be of interest to researchers interested in quantifying the total net benefits from schooling. To this end, health benefits (net of their associated costs) should be incorporated. These benefits may be important for accurate cost-benefit analyses of specific education-related programs or more generally for evaluating whether individuals appear to appropriately balance the total costs and benefits of schooling. If the total private benefits substantially outweigh the private costs, it may indicate that distortions exist in the education sector (e.g. borrowing constraints, lack of information about the costs or benefits of schooling).

What do we know about the extent to which the health benefits of education are paid for? On the one hand, evidence that health behaviors and habits play little role in explaining the education – health gradient would seem to suggest that differences in these types of costly inputs are not particularly important (even though education does appear to affect smoking behavior). On the other hand, the fact that differences in income by education explain one-fourth to one-third of the education – health gradient is consistent with some role for costly inputs. It is not clear why income should be such an important mediator otherwise. Of course, as noted earlier, none of these studies necessarily identifies the channels through which education causally affects health. Furthermore, much of the education – health relationship is left unexplained. Unfortunately, existing studies are not particularly informative about the extent to which the health benefits of education are effectively paid for through costly changes in behavior or in health inputs.

In calculating the total returns to schooling, it would be useful to put a dollar value on all estimated

health benefits. Most benefits are difficult to value (e.g. improvements in self-reported health) even if we could accurately estimate their magnitudes. Fortunately, equation (10) shows how we can value the estimated reductions in mortality associated with schooling using measures of the value of a statistical life (VSL). Typical VSL estimates of $3-5 million suggest that a 0.01 reduction in ten-year mortality rates should be valued at roughly $3-5 thousand in annual gains.[88] Since the IV estimates in Table 4 generally measure the total effect of education on mortality, $d\pi/ds$, rather than the desired partial effect, $\partial\pi/\partial s$, a high-end estimate of the total benefit from mortality reduction is probably around $9-15 thousand dollars (assuming ten-year mortality rates are reduced by 0.03, roughly half the IV estimates of Lleras-Muney (2005, 2006)). This is quite large relative to estimated increases in annual earnings for an additional year of school; however, it is probably much higher than the actual benefit given the small and statistically insignificant estimates on mortality found in three of the four studies reported in Table 4. If education reduces ten-year mortality rates by 0.01 — roughly 10% of baseline mortality rates for 45-80 year-olds in Lleras-Muney (2005) (a figure more in line with the larger set of estimates on self-reported health, health limitations, and activity) — and if roughly half of that reduction is 'paid for' in the form of costly health investments and behavior changes, then a ballpark figure for the mortality benefits of an extra year of school is probably on the order of $1,500-2,500 per year, a non-trivial sum comparable in size to the gains from crime reduction discussed earlier. The value of more general health improvements may also be sizeable.

Of course, even if education does produce health-related externalities or if some of the health benefits of education are not incorporated by youth making their education decisions, it is not clear that education-related policies are the best way to address these problems. Other, more targeted information or public health campaigns may be more efficient policies worth considering. General education policies may also produce different health effects in different environments. Table 7 suggests that the health and mortality benefits from education (and education policies) may be weaker in Europe (compared to the U.S.) where there is greater access to health care, more generous social support, and lower income inequality. We might expect the opposite in developing countries characterized by high inequality and very limited access to health care.

3.4   Suggestions for Future Research on Education and Health

Research on the effects of education on health has accelerated in recent years. Much of that research has justifiably focused on (i) identifying causal effects using IV or RD methods or (ii) on identifying the channels by which education influences health. Most studies of the former type exploit changes in compulsory schooling laws in a few developed countries. Looking forward, we are likely to learn the most from exploiting other sources of exogenous variation in schooling at different margins or by looking at a wider range of countries (or by focusing on different subpopulations). Does schooling raise

---

[88]See Viscusi (1993) and Viscusi and Aldy (2003) for summaries of VSL estimates.

health in developed economies? Do the effects of education differ across countries with very different underlying levels of inequality or access to medical care? Answers to these questions are not only interesting in their own right, but they may also shed additional light on key mechanisms.

Thanks to a number of recent studies, we know much about factors that 'explain' the correlation between education and health, but we know little about how these factors actually determine the *causal* relationship. Attempts to identify different channels through which education causally impacts health and mortality should, therefore, be high on the list of priorities. In particular, it is important to better determine the extent to which health and mortality benefits of education are driven by differences in demand, in which case they are likely to be paid for via costly inputs and foregone opportunities. This is of central concern for policy analysis.

In addition to searching for exogenous sources of variation in schooling induced by policies and other natural experiments, efforts to put some structure on the education and health decision problems, combined with additional data on individual abilities and skills as in Conti, Heckman and Urzua (2010a, 2010b), is also likely to yield new insights about which subpopulations benefit most from education and why.

4 Education's Effect on Citizenship, Political Participation, and Democracy

The hypothesis that education encourages and strengthens democracy has a long intellectual history, with Lipset (1959) crediting the basic idea to Aristotle. Early views by Lipset (1959) and Aristotle emphasize the role of education in informing citizens and increasing their capacity to make 'good' electoral decisions while resisting demagoguery. Education may also affect both the benefits and costs of voting and other forms of political engagement. For example, education may instill civic and democratic values, either through the explicit design of education systems (especially in democratic countries) or indirectly by improving analytical skills and an awareness of history and a diversity of opinions.[89] Education may also indirectly affect political participation by altering social networks and peers. By raising wage rates, education may affect the time costs associated with active political participation and voting. As noted by Verba, et al. (1995), individuals report that lack of time is the most important reason for political inactivity.

Despite the plethora of hypotheses linking education and democracy, formal economic models of this link are scarce. A notable exception is Glaeser, Ponzetto and Shleifer (2007), who emphasize the social nature of political action and education's role in facilitating social interaction. They develop a model in which education endogenously affects political participation, assuming more educated

---

[89]Spilimbergo (2009) finds that countries which send students abroad for higher education experience greater democratization if the host countries are democratic. This suggests that direct efforts by education systems to teach democratic ideals may be important.

individuals are better at persuading others to become politically active (or, alternatively are more persuadable by other active participants). To the extent that these social influences are more valuable to larger political constituencies with less active participation (i.e. broader political regimes), the authors show that a more educated population is likely to be more inclusive and democratic.

While democracy and political freedoms are, no doubt, intrinsically valuable, economists have largely been interested in the link between education and democracy based on the potential for democratic institutions to facilitate economic growth.[90] This is evident in two other political economy models that consider the inter-relationship between the distribution of education, democracy, and economic growth. These models focus more on the conflict between different political constituencies as defined by education and income rather than potential reasons educated citizens may behave more democratically. Bourguignon and Verdier (2000) assume that power is initially concentrated among a small group of educated elites and that the larger uneducated populace is excluded from the political process. By assumption, only the educated vote in their framework. Their main objective is to understand conditions under which the educated elite will selfishly choose policies to increase access to education and, therefore, broaden the political base. Their results suggest that a more equal initial distribution of education leads to faster democratization and more economic growth. Rajan (2009) instead emphasizes the fact that less powerful constituencies have competing education and policy preferences. He argues that in a plutocracy in which political power is determined by wealth, a growth in wealth and power among educated lower-middle class constituents may threaten the economic rents earned by wealthier elites more than the prospect of a fully democratic government in which everyone has equal power regardless of wealth. In this case, the distribution of education may affect decisions to democratize; however, democratization, if it occurs, is likely to be characterized by policy gridlock and weak economic growth.

Beginning with Barro (1999), economists have sought to empirically determine whether increases in a nation's education strengthen democratic institutions. Most studies are macro in nature, using aggregate measures of democracy and education across countries and over time to study the issue. A few microeconometric studies exploit individual-level data within the U.S., U.K., and Germany to estimate the effects of an individual's own education on their likelihood to participate in the political process in one way or another. We review both types of studies below.

---

[90]Proponents of the idea that democracy encourages growth are diverse in their reasoning (e.g., see Hayek 1960, Wittman 1989, North 1990, Olson 1993, and Sen 1999). Others are more skeptical (e.g., see de Tocqueville 1835, Huntington 1968, Buchanan and Tullock 1962, Alesina and Rodrik 1994, Persson and Tabellini 1994, Besley and Coate 1998). Evidence on the effects of democracy on growth is mixed (e.g., see Przeworski and Limongi 1993, Barro 1997, Minier 1998, La Porta, et al. 1999, Tavares and Wacziarg 2001, Doucouliagos and Ulubasoglu 2008, Papaioannou and Siourounis 2008).

4.1 Macro Evidence on Education and Democracy

Empirically, education and democracy are highly correlated across countries. For example, Barro (1999) shows that countries with higher average years of primary schooling also have greater electoral rights and civil liberties. Economists have begun to examine this relationship more closely to determine whether changes in education actually lead to changes in democracy and political freedom, attempting to assign causality. These studies typically estimate an equation of the form:

$$D_{c,t} = \rho D_{c,t-1} + E_{c,t-1}\beta + X_{c,t-1}\gamma + \delta_t + \mu_c + \varepsilon_{c,t}, \tag{12}$$

where $D_{c,t}$ reflects some measure of democracy or political openness in country c and period t, $E_{c,t-1}$ reflects some measure of a country's education level in period $t-1$, $X_{c,t-1}$ reflects any number of potential time-varying determinants of democracy, δ_t reflects unrestricted time effects, and μ_c reflects country fixed effects.

Glaeser, et al. (2004) estimate a version of equation (12) using standard fixed effects estimation methods (and five-year periods). Their estimates suggest that lagged education improves a variety of measures of democracy.[91] However, Acemoglu, et al. (2005) point out that Glaeser, et al. (2004) do not account for time effects in their specification, so identification largely comes from widespread international increases in education and democracy over time. Accounting for time effects (using the same specifications and data), Acemoglu, et al. (2005) find little remaining effect of education.[92]

More importantly, it is well-known that simple fixed effects estimates of equation (12) are likely to be biased with short panels. This is because $D_{c,t-1}$ is correlated with all past values of $\varepsilon_{c,t-j}$ for $j \geq 1$. (Note that $E_{c,t-1}$ is also likely to be correlated with past ε's if there are any feedback effects of democracy on schooling.) Arellano and Bond (1991) show that if the ε's are serially uncorrelated, then one can obtain consistent estimates by first differencing equation (12) and using lagged values of democracy (and schooling) to instrument for their differences, based on the following moment condition: $E[(\varepsilon_{c,t} - \varepsilon_{c,t-1})D_{c,t-j}] = 0$ for all $j \geq 2$. Using this approach and five-year differences from 1965 to 1995, Acemoglu, et al. (2005) estimate small and statistically insignificant effects of average educational attainment on political freedom. They, therefore, conclude that cross-country correlations between education and democracy are due to omitted country-specific factors (e.g. political, social, and economic institutions).[93]

Most recently, Bobba and Coviello (2007) and Castello-Climent (2008) argue that the Arellano and Bond estimator is likely to suffer from weak instrument problems, since democracy is highly persistent over time (in this case, past values of democracy are only weakly correlated with changes in

---

[91]Specifically, they find significant effects on a measure of 'executive constraints' and a general measure of democracy (Jaggers and Marshall 2000) as well as autocracy as measured by Alvarez, et al. (2000).

[92]Papaioannou and Siourounis (2008b) use a sample of countries that are non-democratic in 1960 to estimate the effects of education on the probability of making a democratic transition. Their estimates suggest that education significantly increases the probability, intensity, and speed of democratization.

[93]Also, see Acemoglu, Johnson and Robinson (2005) on the importance of institutions.

democracy). A similar problem arises with respect to measures of education if one instruments for its changes using past values. The presence of weak instruments implies that estimates will be imprecise and biased towards the standard first difference estimator (that does not use any instruments). These authors, therefore, suggest that the 'system GMM' estimator of Blundell and Bond (1998) be used instead. This estimator includes the additional moment equation,

$$E[(D_{c,t-1} - D_{c,t-2})(\mu_c + \varepsilon_{c,t})] = 0, \tag{13}$$

which implies that changes in democracy must be orthogonal to country fixed effects. Both Bobba and Coviello (2007) and Castello-Climent (2008) estimate much larger and statistically significant effects of educational attainment on democracy using this estimator. More interestingly, Castello-Climent (2008) finds that the distribution of education in a country is more important than the average level. Her estimates suggest that greater equity in educational outcomes is associated with stronger measures of democracy.[94]

The system GMM estimator has the advantage of using additional information/instruments; however, the cost is an additional stationarity assumption (equation 13) that is quite strong, since it requires that changes in democracy be unrelated to steady state levels of democracy (i.e. $\mu_c/(1-\rho)$). Given the widespread increases in democracy over the past few decades, it seems likely that many countries are still converging towards their long-run steady states. In this case, we would expect μ_c to be positively correlated with changes in democracy, violating the additional assumption of the system GMM estimator. In this case, there is no clear preference between the Arellano-Bond estimator and the Blundell-Bond system estimator – the former likely suffers from bias due to weak instruments while the latter likely suffers from bias due to an invalid moment restriction. Given these concerns, we turn next to microeconometric evidence on the causal effects of education on voting and citizenship.

4.2 Estimating 'Causal' Effects of Education on Citizenship

A large literature in political science demonstrates a strong correlation between educational attainment and political participation, voting, and civic awareness at the individual level. See Nie, et al. (1996) for an extensive review. Of course, as noted repeatedly above, correlation need not imply causation. Three recent studies by economists (Dee 2004, Milligan, Moretti, and Oreopoulos 2004, Siedler 2010) use individual-level data and instrumental variables strategies to estimate the causal effects of educational attainment on voter registration, voting, support for free speech, and other measures of civic engagement (e.g. newspaper readership, group membership). Table 9 summarizes evidence from these studies on voting and voter registration, while Table 10 summarizes evidence on other measures of political/social involvement and views.

---

[94]Spilimbergo (2009) uses the system GMM estimator to estimate the effects of 'sending' post-secondary students abroad to democratic countries on the originating country. His estimates suggest that foreign-educated individuals promote democracy back home if the education is acquired in democratic countries.

Dee (2004) examines the effects of education on voting and other forms of civic engagement in the U.S. using data from the High School and Beyond (HS&B) longitudinal study and the General Social Surveys (GSS). HS&B followed a cohort of high school sophomores in 1980 through 1992, while Dee's analysis of the 1972-2000 GSS considers individuals born 1900-1964. Milligan, Moretti, and Oreopoulos (2004) examine political outcome measures in the U.S. using the biannual National Election Surveys (NES) from 1948 to 2000 and the 1978-2000 waves of the November Voting Supplement to the Current Population Survey (CPS). The former offers more detailed and consistent measures of political activity for a long time period, while the latter offers much larger sample sizes. Milligan, Moretti, and Oreopoulos (2004) also analyze political behavior in the U.K. using the British Election Surveys (BES) for 1964, 1974, 1979, 1983, 1987, 1992, and 1997, as well as 50 surveys from the Eurobarometer (EB) U.K. surveys covering 1973-1998. Siedler (2010) studies political outcomes in (formerly Western) Germany using data from even numbered years of the ALLBUS (German Social Survey) from 1980 to 2004 (plus the 1991 survey) as well as the large ForsaBus survey on political attitudes from 1991 to 2006.

All three studies exploit variation in compulsory schooling or minimum work age laws as discussed in previous sections. Cross-state variation in the these laws over time is exploited in the U.S. and Germany, while increases in minimum schooling ages in U.K. occurred nationally in 1947 and 1973. Specifications using variation in these laws identify the effects of secondary schooling on political behaviors and attitudes. By contrast, Dee's (2004) analysis using the HS&B data exploits cross-sectional variation in (i) distance two a 2-year college and (ii) number of 2-year colleges in the county of residence in 1983. He, therefore, estimates the causal effect of college enrollment (rather than secondary schooling) on voting and volunteering using these data.[95]

Table 9 reveals sizeable effects of schooling on voting behavior in the U.S., but much weaker and statistically insignificant effects in the U.K. and Germany. In the U.S., the effects of high school graduation on voting are comparable to the effects of college enrollment: the latter increases rates of voting by 30-50% (Dee 2004), while the former increases voting by 40-70% (Milligan, Moretti and Oreopoulos 2004).[96] Education not only affects voting in the U.S., but it also affects voter registration. Milligan, Moretti and Oreopoulos (2004) argue that differences in registration may drive the impacts of education on voting in the U.S. The fact that voter registration is voluntary and left to the individual in the U.S., while it is mandatory and the responsibility of local governments in the U.K., may explain the larger impacts of education on voting in the U.S.

Table 10 summarizes the estimated effects of education on political interest, efforts to acquire

---

[95]We refer to Dee's (2004) bivariate probit specifications (with instruments excluded from the voting choice equation) as IV-based estimates. We report his estimates of the average marginal effect of enrollment.

[96]Dee (2004) notes that the estimated effect of a year of completed schooling in the GSS, 0.07, is comparable in magnitude to the effect of college enrolment, 0.18-0.20, after accounting for the fact that college enrollment is associated with roughly 2.5 years of additional schooling, on average.

information about politics and elections, political views about civil rights and freedoms, and general political and social involvement. Dee (2004) and Milligan, Moretti, and Oreopoulos (2004) estimate sizeable effects of education on a broad range of political domains in the U.S. In particular, they find that additional years of high school significantly increase interest in politics, efforts to acquire information about political issues/campaigns, and beliefs in freedom of speech. By contrast, evidence regarding volunteering and group membership, as emphasized in the theory of Glaeser, Ponzetto and Shleifer (2007), receives little support in U.S. data. There is, however, modest evidence from the U.K. that additional secondary schooling increases individuals' efforts to discuss politics with others and to persuade others to share their views (Milligan, Moretti, and Oreopoulos 2004). Evidence from Germany is generally inconsistent with the view that education encourages democratic ideals (Siedler 2010).

Overall, the micro-empirical evidence suggests important effects of education on a wide range of political behaviors and views in the U.S.; however, this is not the case in the U.K. or Germany. Evidence from the U.S. is most consistent with the older views of Aristotle and Lipset that education leads to a more informed and engaged citizenry. The evidence is less supportive of the idea that schooling influences political outcomes by substantially altering the nature of social interaction.

5   Conclusions

A fast growing literature has established that education and human capital impact a wide range of personal decisions and activities. Education has been shown to reduce crime, improve health, lower mortality, and increase political participation. The social benefits from these impacts can be sizeable. For example, Lochner and Moretti (2004) estimate that high school completion may lower the annual social costs of crime by roughly $3,000 per male graduate. Increasing high school completion rates in the U.S. by one percentage point would reap a savings of more than $2 billion. Annual benefits from reductions in mortality are likely to be in the neighborhood of $1,500-2,500 per additional graduate; however, there is considerable uncertainty about this value.

Much of the evidence on the causal effects of education on crime, health, and citizenship has come from changes in compulsory schooling laws, which primarily affect secondary schooling choices. School lotteries have also been used to estimate the effects of improvements in middle and high school quality on delinquency and crime. As a result, we know much about the impact of additional years of high school on crime, health, and citizenship, but evidence on the effects of higher education is more scarce. Given the rise in college attendance throughout the world, additional efforts to study this margin are needed. There is good reason to believe that increases in college-going are not likely to yield dramatic benefits from crime reduction (at least in the near future), since studies have shown that education-based interventions and policies appear to reduce crime and delinquency most among the least able, most disadvantaged. A few studies have estimated significant reductions in smoking

and improvements in political participation in response to additional years of college, but studies that measure the impacts of higher education on health or citizenship are the exception. There is growing evidence that preschool and school interventions at early ages can reduce delinquency and crime years later. While some programs (e.g. Perry Preschool, CCPC) have been credited with remarkable reductions in adult crime, other similar programs (e.g. Abecedarian, IHDP) have not. Unfortunately, we do not yet understand these differential impacts.

Much of the evidence is U.S. based. A growing number of very recent studies have begun to analyze the non-production benefits of education in the U.K., Germany, and Europe more generally; however, very few studies exploit data from developing countries where education levels are much lower. One might expect substantial differences in the impact of education on crime, health, and citizenship across countries with very different criminal justice, health care, and political systems. Indeed, comparisons across estimates from the U.S. and Europe seem to suggest that education may improve health and mortality less in Europe, where health care tends to be universal and economic inequality is generally lower. Education also appears to impact voting and political participation less in Europe, where voter registration is required and governments are more active in registering voters. While it is tempting to speculate about factors that might explain observed differences in estimates across countries, we are far from understanding them. Much can be learned from simply extending successful empirical strategies to a broader range of countries.

In addition to documenting the impacts of education on non-production activities, this chapter has also discussed what is known about the underlying mechanisms that produce those impacts. We have argued that education may largely affect crime, because it raises human capital and legitimate wage rates. It seems less likely that education reduces crime by altering preferences for risk or impatience. Decompositions of the impacts of education on health suggest that improvements in earnings may explain as much as one-third of the education – health gradient. The fact that education increases knowledge and ability also appears to be important; however, there is little evidence that changes in preferences or psychological traits play important roles. Evidence from the U.S. on the impacts of education on citizenship and democracy appear to support the views of Aristotle and Lipset that schooling leads to a more informed and engaged citizenry.

What role should government play given the findings of this chapter? Crime reduction is an obvious externality that may justify expenditures on policies that improve the skills of the most disadvantaged (e.g. targeted preschool programs, improvements in school quality in low-income areas, or policies that encourage high school completion). Current evidence suggests that well-targeted education-based programs can be more cost-effective than traditional law enforcement policies once all costs and benefits are accounted for. Education policies targeted to the most disadvantaged have the added benefit of reducing economic inequality. There is little evidence of important education externalities in the health domain (most gains are private or, at least, contained within the family),

so arguments for education interventions based on health gains are likely to be based on equity and social justice or on the argument that individuals are unaware of some health benefits. It is important to acknowledge that equity-based proposals that focus on educating the most disadvantaged are likely to be inefficient if individuals are aware of the full costs and benefits of education. Indeed, some individuals may benefit more (in terms of health and in other ways) from transfers of income than from additional schooling.[97] However, if individuals are unaware of some health benefits and unknown benefits are achieved without cost, then education-based policies may be justified. Unfortunately, we still do not know the extent to which most health benefits are effectively paid for through higher health care expenditures or costly changes in behavior. These costs must be factored into policy decisions. Finally, it is clear that increases in political participation will affect the democratic process; however, it is difficult to know exactly how and even more difficult to put a value on this.

Appendix: Estimating Mediating Effects

Researchers are often interested in the effect of education on health (or other outcomes) that operates through or is mediated by some particular variable. This appendix considers this problem in a simple setting. See Alwin and Hauser (1975) for an early treatment of this type of problem. Also, see Hauser and Goldberger (1971) and Goldberger (1972) for discussions of identification and estimation in more general environments.

Consider the following model for health outcome y as a function of education E, and mediating variables x_1 and x_2:

$$y = E\alpha + x_1\beta_1 + x_2\beta_2 + \varepsilon \tag{14}$$

$$x_1 = E\gamma_E + x_2\gamma_2 + \eta_1 \tag{15}$$

$$x_2 = E\delta_E + x_1\delta_1 + \eta_2. \tag{16}$$

To focus on the appropriate way to attribute effects of education through a mediating variable, we abstract from concerns about endogeneity, assuming $\varepsilon \perp\!\!\!\perp (E, x_1, x_2)$. The direct, or partial, effect of E on y holding x_1 and x_2 constant is given by α in equation (14). We assume $0 \leq \gamma_2\delta_1 < 1$ so the system is well-behaved.

By solving for x_1 and x_2 as functions of E, η_1 and η_2, then substituting these into equation (14), we can write y as a function of E (and error terms) only:

$$y = E\left[\alpha + \frac{(\gamma_E + \gamma_2\delta_E)\beta_1 + (\delta_E + \delta_1\gamma_E)\beta_2}{1 - \delta_1\gamma_2}\right] + \left[\frac{\beta_1 + \beta_2\delta_1}{1 - \delta_1\gamma_2}\right]\eta_1 + \left[\frac{\beta_2 + \beta_1\gamma_2}{1 - \delta_1\gamma_2}\right]\eta_2 + \varepsilon. \tag{17}$$

This defines the 'total effect' of E on y: $a = \alpha + \frac{(\gamma_E + \gamma_2\delta_E)\beta_1 + (\delta_E + \delta_1\gamma_E)\beta_2}{1 - \delta_1\gamma_2}$.

---

[97]Of course, income transfers typically create disincentives for work, so education interventions may provide a less inefficient way of improving equity regardless of its health benefits.

Suppose we are interested in the effect of education on health that operates through or is mediated by x_1. In the model defined by equations (14)-(16), this is given by

$$
\begin{aligned}
ME_1 &= \gamma_E[\beta_1 + \delta_1\beta_2 + \delta_1\gamma_2\beta_1 + \delta_1^2\gamma_2\beta_2 + \delta_1^2\gamma_2^2\beta_1 + ...] \\
&= \frac{\gamma_E(\beta_1 + \delta_1\beta_2)}{1 - \delta_1\gamma_2}.
\end{aligned}
\tag{18}
$$

A common decomposition approach for estimating ME_1 is to regress y on E and x_2; then regress y on E, x_1, and x_2 (i.e. equation (14)), subtracting the latter estimated coefficient on E from the former. Estimation of equation (14) yields $\hat{\alpha} \overset{p}{\to} \alpha$ given the assumption $\varepsilon \perp\!\!\!\perp (E, x_1, x_2)$. Further assuming $E(\eta_1|E, x_2) = 0$ and estimating $y = Ea_2 + x_2b_2 + \nu_2$ yields $\hat{a}_2 \overset{p}{\to} \alpha + \gamma_E\beta_1$. The estimate of the mediating effect of x_1 is then:

$$
\widetilde{ME}_1 = \hat{a}_2 - \hat{\alpha} \overset{p}{\to} \gamma_E\beta_1 = ME_1 - \frac{\gamma_E\delta_1(\gamma_2\beta_1 + \beta_2)}{1 - \delta_1\gamma_2}.
$$

This strategy obtains consistent estimates for the effect mediated through x_1, ME_1, if any of three special cases hold: (a) no effect of x_1 on x_2 ($\delta_1 = 0$); (b) no direct effect of E on x_1 ($\gamma_E = 0$); or (c) no effect of either x_1 or x_2 on outcome y ($\beta_1 = \beta_2 = 0$). In general, this strategy leads to bias, because including x_2 in the specifications eliminates the mediating role x_1 plays through its impact on x_2. If all parameters are positive, this produces a downward biased estimate of the mediating effect of x_1.[98] Also, notice that even if $\beta_2 = 0$ (so x_2 does not directly affect outcome y), this estimator may be biased if x_1 and x_2 affect each other.

A second common decomposition used to estimate ME_1 regresses y on E; then regresses y on E and x_1 and subtracts the latter estimated coefficient on E from the former. That is, estimate $y = Ea + \nu$ and $y = Ea_1 + x_1b_1 + \nu_1$, then subtract the estimate of a_1 from the estimate of a. This strategy is often used (implicitly or explicitly) when x_2 is unavailable in the data, in which case the first estimation strategy is infeasible. As long as $E(\eta_1|E) = E(\eta_2|E) = 0$ (and $\varepsilon \perp\!\!\!\perp E$ as assumed throughout), it is clear that $\hat{a} \overset{p}{\to} a$. If $E(\eta_2|E, x_1) = 0$, then $\hat{a}_1 \overset{p}{\to} \alpha + \beta_2\delta_E$. This approach yields the following estimator for the mediating effect of x_1:

$$
\widehat{ME}_1 = \hat{a} - \hat{a}_1 \overset{p}{\to} \frac{(\gamma_E + \gamma_2\delta_E)(\beta_1 + \delta_1\beta_2)}{1 - \delta_1\gamma_2} = ME_1 + \frac{\gamma_2\delta_E(\beta_1 + \delta_1\beta_2)}{1 - \delta_1\gamma_2}.
$$

This strategy obtains consistent estimates for the effect mediated through x_1, ME_1, if any of three special cases hold: (a) no effect of x_2 on x_1 ($\gamma_2 = 0$); (b) no direct effect of E on x_2 ($\delta_E = 0$); or (c) no effect of either x_1 or x_2 on outcome y ($\beta_1 = \beta_2 = 0$). It is not enough for x_2 to have no direct effect on y to obtain consistent estimates of ME_1. In this case, the bias $\frac{\gamma_2\delta_E\beta_1}{1 - \delta_1\gamma_2}$ depends on the effect of x_1 on x_2, x_2 on x_1, and the direct effect of education on x_2.

The decision regarding which estimator to use depends on the relationship between the x variables, the direct effects of each x on the outcome y, and the effects of education on each x. Notice that

---

[98]This strategy does, however, identify the effect of increasing E through x_1 if any effects on x_2 could somehow be offset.

$\widehat{ME}_1 = \widetilde{ME}_1 = ME_1$ if $\beta_1 = \beta_2 = 0$, so both estimation strategies will identify the mediating effect of x_1 in this case. Of course, this is neither surprising nor interesting. If we assume that both x variables affect the outcome of interest, the decision of which estimator to use depends on the relationship between education and both x's and on the interdependence of the two x's on each other. Assuming both x variables affect each other, then the first estimator, $\widehat{ME}_1$, yields a consistent estimate of the mediating effect of x_1 if education does not directly affect x_2 while the second estimator, $\widetilde{ME}_1$, yields a consistent estimate if education does not directly affect x_1. If we instead assume that education directly affects both x variables, then the choice of estimators depends on whether we can rule out an effect of one x variable on the other. If x_1 does not affect x_2, then the first estimator will yield a consistent estimate of ME_1, while the second estimator is consistent if x_2 does not affect x_1.

These findings imply the following general lessons when estimating the mediating effect of some variable using the standard de-composition approach: On the one hand, failure to control for other variables can bias estimates of the mediating effect when (a) education directly affects those variables and (b) those variables directly impact the 'mediating variable' of interest. On the other hand, controlling for other variables can also bias estimates of the mediating effect when (a) the mediating variable directly affects these other variables and (b) education directly affects those variables.

If x_2 is available in the data, it is possible under some conditions to use the two estimators to bound the mediating effect of x_1. For example, suppose x_1 and x_2 are normalized such that they both have positive effects on y (i.e. β_1 and β_2 are positive).[99] If x_1 and x_2 have positive cross-effects (i.e. $\gamma_2 \geq 0$ and $\delta_1 \geq 0$) and if education increases both x_1 and x_2 (i.e. $\gamma_E \geq 0$ and $\delta_E \geq 0$), then the two estimation strategies bound the mediating effect of x_1: $\widetilde{ME}_1 \leq ME_1 \leq \widehat{ME}_1$. Alternatively, under the exogeneity assumptions assumed throughout this Appendix, one could simply construct a consistent estimate of the mediating effect of x_1 using equation (18) and parameter estimates from equations (14)-(16).

References

[1] Daron Acemoglu and Joshua Angrist. How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual 2000*, 15:9–74, 2001.

[2] Daron Acemoglu, Simon Johnson, and James A. Robinson. Institutions as a fundamental cause of long-run growth. volume 1, Part 1 of *Handbook of Economic Growth*, chapter 6, pages 385–472. Elsevier, 2005.

[3] Daron Acemoglu, Simon Johnson, James A. Robinson, and Pierre Yared. From education to democracy? *The American Economic Review*, 95(2):44–49, 2005.

---

[99]Note that this can be verified from estimation of equation (14).

[4] Scott J. Adams. Educational attainment and health: Evidence from a sample of older adults. *Education Economics*, 10(1):97–109, 2002.

[5] Anna Aizer and Laura Stroud. Education, knowledge and the evolution of disparities in health. NBER Working Paper 15840, 2010.

[6] Valerie Albouy and Laurent Lequien. Does compulsory education lower mortality? *Journal of Health Economics*, 28(1):155–68, 2009.

[7] Alberto Alesina and Dani Rodrik. Distributive politics and economic growth. *The Quarterly Journal of Economics*, 109(2):465–90, 1994.

[8] M. Alvarez, J. A. Cheibub, F. Limongi, and A. Przeworski. *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990.* Cambridge University Press, Cambridge, 2000.

[9] Duane F. Alwin and Robert M. Hauser. The decomposition of effects in path analysis. *American Sociological Review*, 40(1):37–47, 1975.

[10] Mark Anderson. In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime. Working Paper, 2009.

[11] Joshua D. Angrist and Guido W. Imbens. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

[12] Joshua D. Angrist and Guido W. Imbens. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, 90(430):431–42, 1995.

[13] Manuel Arellano and Stephen Bond. Some tests of specificatino for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–97, 1991.

[14] Jacob Nielsen Arendt. Does education cause better health? a panel data analysis using school reforms for identification. *Economics of Education Review*, 24(2):149–60, 2005.

[15] Jeremy Arkes. Does schooling improve adult health? RAND DRU-3051, 2003.

[16] M. Christopher Auld and Nirmal Sidhu. Schooling, cognitive ability and health. *Health Economics*, 14(10):1019–34, 2005.

[17] W. Steven Barnett and Leondard N. Masse. Comparative benefit-cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review*, 26:113–25, 2007.

[18] Robert J. Barro. *Determinants of economic growth: A cross-country empirical study.* The MIT Press, Cambridge, Mass., 1997.

[19] Robert J. Barro. Determinants of democracy. *The Journal of Political Economy*, 107(6):S158–83, 1999.

[20] Gary S. Becker. Human Capital and the Personal Distribution of Income: An Analitical Approach. Woytinsky Lecture 1, University of Michigan, Institute of Public Administration, Ann Arbor, 1967.

[21] Gary S. Becker and Casey B. Mulligan. The endogenous determination of time preference. *The Quarterly Journal of Economics*, 112(3):729–58, 1997.

[22] Clive R. Belfield, Milagros Nores, W. Steven Barnett, and Lawrence Schweinhart. The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup. *Journal of Human Resources*, 41:162–190, 2006.

[23] Yoram Ben-Porath. The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, 75(4):352–65, 1967.

[24] Timothy Besley and Stephen Coate. Sources of inefficiency in a representative democracy: A dynamic analysis. *The American Economic Review*, 88(1):139–56, 1998.

[25] David Blau and Janet Currie. Pre-School, Day Care, and After-School Care: Who's Minding The Kids. In E. Hanushek and F. Welch, editors, *Handbook of the Economics of Education*, volume 2, chapter 20. Elsevier, Science B.V., Amsterdam, 2006.

[26] Richard Blundell and Stephen Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–43, 1998.

[27] Matteo Bobba and Decio Coviello. Weak instruments and weak identification, in estimating the effects of education, on democracy. *Economics Letters*, 96(3):301–06, 2007.

[28] Rachel Boisjoli, Frank Vitaro, Eric Lacourse, Edward D. Barker, and Richard E. Tremblay. Impact and Clinical Significance of a Preventive Intervention For Disruptive Boys. *British Journal of Psychiatry*, 191:415–419, 2007.

[29] Franois Bourguignon and Thierry Verdier. Oligarchy, democracy, inequality and growth. *Journal of Development Economics*, 62(2):285–313, 2000.

[30] Anthony Braga. Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials. *Journal of Experimental Criminology*, 1:317–342, 2005.

[31] Giorgio Brunello, Daniele Fabbri, and Margherita Fort. Years of schooling, human capital and the body mass index of European females. IZA Discussion Papers 4667, Institute for the Study of Labor (IZA), 2009.

[32] James M. Buchanan and Gordon Tullock. *The calculus of consent: logical foundations of constitutional democracy.* University of Michigan Press, Ann Arbor, MI, 1962.

[33] Paolo Buonanno and Leone Leonida. Education and Crime: Evidence from Italian Regions. *Applied Economics Letters*, 13:709–13, 2006.

[34] Stephen V. Cameron and James J. Heckman. The Nonequivalence of High School Equivalents. *Journal of Labor Economics*, 11(1, part 1):1–47, 1993.

[35] Stephen V. Cameron and Christopher Taber. Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1):132–82, 2004.

[36] Frances A. Campbell, Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science*, 6:42–57, 2002.

[37] David Card. The Economics of Crime. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 3A, chapter 30. Elsevier, Science B.V., Amsterdam, 1999.

[38] Pedro Carneiro and Rita Ginja. Preventing Behavior Problems in Childhood and Adolescence: Evidence from Head Start. Working Paper, 2008.

[39] Pedro Carneiro and James J. Heckman. The Evidence on Credit Constraints in Post-secondary Schooling. *The Economic Journal*, 112(482):705–34, 2002.

[40] Amparo Castelló-Climent. On the distribution of education and democracy. *Journal of Development Economics*, 87(2):179–90, 2008.

[41] John Cawley, James Heckman, Lance Lochner, and Edward Vytlacil. Understanding the role of cognitive ability in accounting for the recent rise in the economic return to education. In *Meritocracy and Economic Inequality*, pages 230–65. Princeton University Press, 2000.

[42] Kenneth Chay, Carlos Dobkin, and Michael Greenstone. The clean air act of 1970 and adult mortality. *Journal of Risk and Uncertainty*, 27(3):279–300, 2003.

[43] Kenneth Y. Chay and Michael Greenstone. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics*, 118(3):1121–67, 2003.

[44] Kenneth Y. Chay and Michael Greenstone. Does air quality matter? evidence from the housing market. *Journal of Political Economy*, 113(2):376–424, 2005.

[45] Shin-Yi Chou, Jin-Tan Liu, Michael Grossman, and Ted Joyce. Parental education and child health: Evidence from a natural experiment in Taiwan. *American Economic Journal: Applied Economics*, 2(1):33–61, 2010.

[46] Antonio Ciccone and Giovanni Peri. Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies*, 73(2):381–412, 2006.

[47] Damon Clark and Heather Royer. The Effect of Education on Adult Health and Mortality: Evidence from Britain. Working Paper, 2010.

[48] Stevens H. Clarke and Frances A. Campbell. Can Intervention Early Prevent Crime Later? The Abecedarian Progect Compared with Other Programs. *Early Childhood Research Quarterly*, 13:319–43, 1998.

[49] Ethan Cohen-Cole and Jason M. Fletcher. Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics*, 27(5):1382–87, 2008.

[50] Conduct Problems Prevention Research Group. Fast Track Randomized Controlled Trial to Prevent Externalizing Psychiatric Disorders: Findings From Grades 3 to 9. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46:1250–62, 2007.

[51] Conduct Problems Prevention Research Group. The Effects of the Fast Track Preventiv Intervention on the Development of Conduct Disorder across Childhood. *Child Development*, 2010a.

[52] Conduct Problems Prevention Research Group. Fast Track Intervention Effects on Youth Arrests and Delinquency. *Journal of Experimental Criminology*, 2010b.

[53] Gabriella Conti, James Heckman, and Sergio Urzua. The Education-Health Gradient. *American Economic Review*, 100(2):234–38, 2010a.

[54] Gabriella Conti, James J. Heckman, and Sergio Urzúa. Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions. Working Paper, June 2010b.

[55] Julie Cullen, Brian Jacob, and Steven Levitt. The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74:1191–1230, 2006.

[56] Flavio Cunha and James J. Heckman. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *The Journal of Human Resources*, 43:738–782, 2008.

[57] Janet Currie and Enrico Moretti. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532, 2003.

[58] Janet Currie and Matthew Neidell. Air pollution and infant health: What can we learn from california's recent experience? *The Quarterly Journal of Economics*, 120(3):1003–30, 2005.

[59] David M. Cutler and Adriana Lleras-Muney. Education and health: Evaluating theories and evidence. Working Paper 12352, National Bureau of Economic Research, 2006.

[60] David M. Cutler and Adriana Lleras-Muney. Understanding differences in health behaviors by education. *Journal of Health Economics*, 29(1):1–28, 2010.

[61] Alexis de Tocqueville. *Democracy in America*. Wordsworth, Hertfordshire, 1998. (originally published in 1835).

[62] Damien de Walque. Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics*, 26:877–95, 2007a.

[63] Damien de Walque. How does the impact of an hiv/aids information campaign vary with educational attainment? evidence from rural Uganda. *Journal of Development Economics*, 84(2):686–714, 2007b.

[64] Damien de Walque. Education, information and smoking decisions: Evidence from smoking histories in the United States, 1940-2000. *Journal of Human Resources*, 45(3):682–716, 2010.

[65] Thomas S. Dee. Are there civic returns to education? *Journal of Public Economics*, 88(9-10):1697–1720, 2004.

[66] David Deming. Better Schools, Less Crime? Working Paper, 2009a.

[67] David Deming. Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3):111–34, 2009b.

[68] John J. Donohue III and Peter Siegelman. Allocating Resources Among Prisons and Social Programs in the Battle Against Crime. *The Journal of Legal Studies*, 27:1–44, 1998.

[69] Hristos Doucouliagos and Mehmet Ali Ulubasoglu. Democracy and economic growth: A meta-analysis. *American Journal of Political Science*, 52(1):61–83, 2008.

[70] Isaac Ehrlich and Hiroyuki Chuma. A model of the demand for longevity and the value of life extension. *Journal of Political Economy*, 98(4):761–82, 1990.

[71] Irma T. Elo and Samuel H. Preston. Educational differentials in mortality: United States, 1979-85. *Social Science and Medicine*, 42(1):47–57, 1996.

[72] Manfred Erbsland, Walter Ried, and Volker Ulrich. Health, health care, and the environment. econometric evidence from german micro data. *Health Economics*, 4(3):169–82, 1995.

[73] David P. Farrington, Bernard Gallagher, Lynda Morley, Raymond J. St. Ledger, and Donald J. West. Unemployment, School Leaving and Crime. *British Journal of Criminology*, 26:335–56, 1986.

[74] Guiulio Fella and Giovanni Gallipoli. Education and Crime over the Life Cycle. Working Paper, University of British Columbia, 2009.

[75] Jason M. Fletcher. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. *Health Economics*, 19(4):466–84, 2010.

[76] Victor R. Fuchs. Time preference and health: an exploratory study. In Victor R. Fuchs, editor, *Economic aspects of health*, chapter 3, pages 93–120. University of Chicago Press, Chicago, 1982.

[77] Eliana Garces, Duncan Thomas, and Janet Currie. Longer-Term Effects of Head Start. *American Economic Review*, 92:999–1012, 2002.

[78] Alejandro Gaviria and Steven Raphael. School-based peer effects and juvenile behavior. *The Review of Economics and Statistics*, 83(2):257–68, 2001.

[79] Donna B. Gilleskie and Amy L. Harrison. The effect of endogenous health inputs on the relationship between health and education. *Economics of Education Review*, 17(3):279–95, 1998.

[80] Edward Glaeser, Giacomo Ponzetto, and Andrei Shleifer. Why does democracy need education? *Journal of Economic Growth*, 12:77–99, 2007.

[81] Edward Glaeser, Bruce Sacerdote, and Jose Scheinkman. Crime and Social Interactions. *Quarterly Journal of Economics*, 111(2):507–48, 1996.

[82] Edward L. Glaeser, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. Do institutions cause growth? *Journal of Economic Growth*, 9:271–303, 2004.

[83] Sherry Glied and Adriana Lleras-Muney. Technological innovation and inequality in health. *Demography*, 45(3):741–61, 2008.

[84] Brian Goesling. The Rising Significance of Education for Health? *Social Forces*, 85(4):1621–44, 2007.

[85] Arthur S Goldberger. Structural equation methods in the social sciences. *Econometrica*, 40(6):979–1001, 1972.

[86] Dana Goldman and James P. Smith. Socioeconomic differences in the adoption of new medical technologies. *American Economic Review*, 95(2):234–37, 2005.

[87] Dana P. Goldman and Darius N. Lakdawalla. A theory of health disparities and medical technology. *Contributions to Economic Analysis & Policy*, 4(1):1–32, 2005.

[88] Dana P. Goldman and James P. Smith. Can patient self-management help explain the ses health gradient? *Proceedings of the National Academy of Sciences of the United States of America*, 99(16):10929–34, 2002.

[89] Denise Gottfredson. Youth Employment, Crime, and Schooling. *Developmental Psychology*, 21:419–32, 1985.

[90] Eric Gould, David Mustard, and Bruce Weinberg. Crime Rates and Local Labor Market Opportunities in the United States: 1977-1997. *Review of Economics and Statistics*, 84:45–61, 2002.

[91] Michael J. Grabner. The Impact of Education on Health Using Compulsory Schooling Laws. Working Paper, 2009.

[92] Julien Grenet. Is it Enough to Increase Compulsory Education to Raise Earnings? Evidence from French and Brisish Compulsory Schooling Laws. Working Paper, 2010.

[93] Franque Grimard and Daniel Parent. Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking? *Journal of Health Economics*, 26:896–926, 2007.

[94] Jeffrey Grogger. The Effect of Arrests on the Employment and Earnings of Young Men. *Quarterly Journal of Economics*, 110:51–71, 1995.

[95] Jeffrey Grogger. Market Wages and Youth Crime. *Journal of Labor Economics*, 16:756–91, 1998.

[96] Michael Grossman. On the concept of health capital and the demand for health. *The Journal of Political Economy*, 80(2):223–55, 1972a.

[97] Michael Grossman. *The Demand for Health: A Theoretical and Empirical Investigation.* Columbia University Press, New York, 1972b.

[98] Michael Grossman. The human capital model. In Anthony J. Culyer and Joseph P. Newhouse, editors, *Handbook of Health Economics*, volume 1, Part 1, chapter 7, pages 347–408. Elsevier B.V., 2000.

[99] Michael Grossman. Education and nonmarket outcomes. In E. Hanushek and F. Welch, editors, *Handbook of the Economics of Education*, volume 1, chapter 10, pages 577–633. Elsevier B.V., 2006.

[100] Michael Grossman and Robert Kaestner. Effects of education on health. In Jere R. Behrman and Nevzer Stacey, editors, *The social benefits of education*, chapter 4, pages 69–123. University of Michigan Press, Ann Arbor, 1997.

[101] Jonathan Guryan. Desegregation and Black Dropout Rates. *American Economic Review*, 94(4):919–43, 2004.

[102] Caroline W. Harlow. *Education and Correctional Populations*. U.S. Dept. of Justice, Bureau of Justice Statistics, 2003.

[103] Robert M. Hauser and Arthur S. Goldberger. The treatment of unobservable variables in path analysis. *Sociological Methodology*, 3:81–117, 1971.

[104] J. David Hawkins, Richard F. Catalano, Rick Kosterman, Robert Abbott, and Karl G. Hill. Promoting Positive Adult Functioning Through Social Development Intervention in Childhood. Long Term Effects From the Seattle Social Development Project. *Archives of Pediatrics and Adolescent Medicine*, 159:25–31, 2005.

[105] Friedrich A. von Hayek. *The constitution of liberty*. University of Chicago Press, Chicago, 1960.

[106] James J. Heckman and Peter Klenow. Human capital policy. Working Paper, 1997.

[107] James J. Heckman and Paul LaFontaine. Bias-Corrected Estimates of GED Returns. *Journal of Labor Economics*, 24(3):661–700, 2006.

[108] James J. Heckman, Lance Lochner, and Petra Todd. Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond. In E. Hanushek and F. Welch, editors, *Handbook of the Economics of Education*, volume 1, chapter 7, pages 307–458. Elsevier B.V., 2006.

[109] James J. Heckman, Lance Lochner, and Petra Todd. Earnings Functions and Rates of Return. *Journal of Human Capital*, 2(1):1–31, 2008.

[110] James J. Heckman, Seong H. Moon, Rodrigo Pinto, Peter Savelyev, and Adam Yavitz. Ther Rate of Return to the High/Scope Perry Preschool Program. NBER Working Paper No. 15471, 2009.

[111] James J. Heckman and Yona Rubinstein. The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *American Economic Review*, 91(2):145–59, 2001.

[112] James J. Heckman, Jora Stixrud, and Sergio Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–82, 2006.

[113] Michael J. Hingelang, Travis Hirschi, and Joseph G. Weis. *Measuring Delinquency*. Sage Publications, Inc., 1981.

[114] Randi Hjalmarsson. Criminal Justice Involvement and High School Completion. *Journal of Urban Economics*, 63(2):613–30, 2008.

[115] Samuel P. Huntington. *Political order in changing societies*. Yale University Press, New Haven, 1968.

[116] Brian Jacob and Lars Lefgren. Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime. *American Economic Review*, 93:1560–77, 2003.

[117] K. Jaggers and M. G. Marshall. Polity iv project. Technical report, Center for International Development and conflict Management, University of Maryland, 2000.

[118] Hendrik Jürges, Eberhard Kruk, and S. Reinhold. The Effect of Compulsory Schooling on Health - Evidence from Biomarkers. Working paper, Mannheim Research Institute for the Economics of Aging (MEA), University of Mannheim, 2009.

[119] E. Kandel et al. IQ as a Protective Factor for Subjects at High Risk for Antisocial Behavior. *Journal of Consulting and Clinical Psychology*, 56:224–6, 1988.

[120] Lynn A. Karoly et al. *Investing in Our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Interventions*. RAND, 1998.

[121] Daniel Kemptner, Hendrik Jürges, and Steffen Reinhold. Changes in compulsory schooling and the causal effect of education on health: Evidence from germany. Working paper 10200, Mannheim Research Institute for the Economics of Aging (MEA), University of Mannheim, 2010.

[122] Donald Kenkel, Dean Lillard, and Alan Mathios. The roles of high school completion and ged receipt in smoking and obesity. *Journal of Labor Economics*, 24(3):635–60, 2006.

[123] Donald S. Kenkel. Health behavior, health knowledge, and schooling. *The Journal of Political Economy*, 99(2):287–305, 1991.

[124] Evelyn M. Kitagawa and Philip M. Hauser. *Differential Mortality in the United States*. Harvard University Press, Cambridge, 1973.

[125] Jeffrey Kling. The Effect of Prison Sentence Length on the Subsequent Employment and Earnings of Criminal Defendants. Working Paper, Princeton University, 2002.

[126] Jeffrey Kling, Jens Ludwig, and Lawrence Katz. Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *Quarterly Journal of Economics*, 120(1):87–130, 2005.

[127] Brian Krauth. Structural Estimation of Peer Effects in Youth Smoking. *Advances in Health Economics and Health Services Research*, 16:201–11, 2005.

[128] Rafael La Porta, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert Vishny. The quality of government. *Journal of Law, Economics, and Organization*, 15(1):222–79, 1999.

[129] J. Ronald Lally, Peter Mangione, and Alice Honig. The Syracuse University Family Development Research Program: Long-Range Impact of an Early Intervention with Low-Income Children and Their Families. In D. Powell, editor, *Parent Education as Early Childhood Intervention: Emerging Directions in Theory, Research, and Practice*, pages 79–104. Ablex, Norwood, 1988.

[130] Fabian Lange. The Role of Education in Complex Health Decisions: Evidence from Cancer Screening. *Journal of Health Economics*, forthcoming.

[131] Fabian Lange and Robert Topel. The social value of education and human capital. In E. Hanushek and F. Welch, editors, *Handbook of the Economics of Education*, volume 1, chapter 8, pages 459–509. Elsevier, 2006.

[132] J. Paul Leigh. Direct and indirect effects of education on health. *Social Science and Medicine*, 17(4):227–34, 1983.

[133] Maarten Lindeboom, Ana Llena-Nozal, and Bas van der Klaauw. Parental education and child health: Evidence from a schooling reform. *Journal of Health Economics*, 28(1):109–31, 2009.

[134] Seymour Lipset. Some social requisites for democracy: Economic development and political legitimacy. *American Political Science Review*, 53(1):69–105, 1959.

[135] Adriana Lleras-Muney. Were Compulsory Attenance and Child Labor Laws Effective? An Analysis from 1915 to 1939. *Journal of Law and Economics*, 45(2), 2002.

[136] Adriana Lleras-Muney. The relationship between education and adult mortality in the United States. *The Review of Economic Studies*, 72(1):189–221, 2005.

[137] Adriana Lleras-Muney. Erratum: The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, 73(3):847, 2006.

[138] Adriana Lleras-Muney. The Needs of the Army: Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Childrens Health. *Journal of Human Resources*, 45:549–90, 2010.

[139] David Long, Charles Mallar, and Craig Thornton. Evaluating the Benefits and Costs of the Job Corps. *Journal of Policy Analysis and Management*, 1:55–76, 1981.

[140] Jeremy Luallen. School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes. *Journal of Urban Economics*, 59:75–103, 2006.

[141] Jens Ludwig and Douglas L. Miller. Does Head Start Improve Children's Life Chances? Evidence From a Regression Discontinuity Design. *Quarterly Journal of Economics*, 122:159–208, 2007.

[142] Stephen Machin, Olivier Marie, and Suncica Vujic. The Crime Reducing Effect of Education. *Economic Journal*, forthcoming.

[143] Stephen Machin and Costas Meghir. Crime and Economic Incentives. *Journal of Human Resources*, 39:958–79, 2004.

[144] Sudeshna Maitra. Can patient self-management explain the health gradient? Goldman and Smith's "patient self-management help explain the SES health gradient?" (2002) revisited. *Social Science & Medicine*, 70(6):802–12, 2010.

[145] Michael Marmot. *Status syndrome : how your social standing directly affects your health and life expectancy.* Bloomsbury Pub., London, 2004.

[146] Michael Marmot and Richard G. Wilkinson. *Social determinants of health.* Oxford University Press, Oxford ; New York, 2006.

[147] Michael G. Marmot. Social differentials in health within and between populations. *Daedalus*, 123(4):197–216, 1994.

[148] Bhashkar Mazumder. Does education improve health? a reexamination of the evidence from compulsory schooling laws. *Economic Perspectives*, pages 2–16, 2008.

[149] Bhashkar Mazumder. Erratum: Does education improve health? A reexamination of the evidence from compulsory schooling laws. November 16 2010.

[150] Marie C. McCormick et al. Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age fir the Infant Health and Development Program. *Pediatrics*, 117:771–80, 2006.

[151] Justin McCrary and Heather Royer. The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. NBER Working Paper No. 12329, 2009.

[152] Ellen R. Meara, Seth Richards, and David M. Cutler. The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 19812000. *Health Affairs*, 27(2):350–60, 2008.

[153] Antonio Merlo and Kenneth Wolpin. The Transition from School to Jail: Youth Crime and High School Completion among Black Males. Penn Institute for Economic Research Working Paper 09-002, 2009.

[154] Ted Miller, Mark Cohen, and Brian Wiersema. Victim Costs and Consequences: A New Look. Final Summary Report to the National Institute of Justice, Feb. 1996.

[155] Kevin Milligan, Enrico Moretti, and Philip Oreopoulos. Does education improve citizenship? evidence from the United states and the United Kingdom. *Journal of Public Economics*, 88(9-10):1667–95, 2004.

[156] Jenny A. Minier. Democracy and growth: Alternative approaches. *Journal of Economic Growth*, 3:241–66, 1998.

[157] Enrico Moretti. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1-2):175–212, 2004. Higher education (Annals issue).

[158] Enrico Moretti. Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *The American Economic Review*, 94(3):656–90, 2004.

[159] Enrico Moretti and Matthew Neidell. Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. Working Paper 14939, NBER, 2009.

[160] D. Nagin and J. Waldfogel. The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders. *International Review of Law and Economics*, 15:109–26, 1995.

[161] Matthew Neidell. Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, 23(6):1209–36, 2004.

[162] Douglass C. North. *Institutions, institutional change, and economic performance.* Cambridge University Press, Cambridge and London, 1990.

[163] Edward C. Norton, Richard C. Lindrooth, and Susan T. Ennett. Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use. *Health Economics*, 7:439–53, 1998.

[164] David L. Olds et al. Long-Term Effects of nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15-Year Follow-up of a Randomized Controlled Trial. *Journal of the American Medical Association*, 280:1238–44, 1998.

[165] Mancur Olson. Dictatorship, democracy, and development. *The American Political Science Review*, 87(3):567–76, 1993.

[166] Philip Oreopoulos. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *The American Economic Review*, 96(1):152–75, 2006.

[167] Philip Oreopoulos. Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum. University of British Columbia Working Paper, 2008.

[168] Philip Oreopoulos and Kjell G. Salvanes. How Large are Returns to Schooling? Hint: Money Isn't Everything. NBER Working Paper No. 15339, 2009.

[169] Elias Papaioannou and Gregorios Siourounis. Democratisation and growth. *The Economic Journal*, 118(532):1520–51, 2008a.

[170] Elias Papaioannou and Gregorios Siourounis. Economic and social factors driving the third wave of democratization. *Journal of Comparative Economics*, 36(3):365–87, 2008b.

[171] Gregory Pappas, Susan Queen, Wilbur Hadden, and Gail Fisher. The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. *The New England Journal of Medicine*, 329(2):103–109, July 8 1993.

[172] Torsten Persson and Guido Tabellini. Is inequality harmful for growth? *The American Economic Review*, 84(3):600–21, 1994.

[173] Nattavudh Powdthavee. Does Education Reduce the Risk of Hypertension? Estimating the Biomarker Effect of Compulsory Schooling in England. *Journal of Human Capital*, 4(2):173–202, 2010.

[174] Lisa M. Powell, John A. Tauras, and Hana Ross. The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. *Journal of Health Economics*, 24(5):950–68, 2005.

[175] Joseph Price and Kosali Simon. Patient education and the impact of new medical research. *Journal of Health Economics*, 28(6):1166–74, 2009.

[176] Adam Przeworski and Fernando Limongi. Political regimes and economic growth. *The Journal of Economic Perspectives*, 7(3):51–69, 1993.

[177] Raghuram G. Rajan. Rent preservation and the persistence of underdevelopment. *American Economic Journal: Macroeconomics*, 1(1):178–218, 2009.

[178] Sarah Reber. School Desegregation and Educational Attainment for Blacks. NBER Working Paper Paper No. 13193, 2007.

[179] Steffen Reinhold and Hendrik Jürges. Secondary school fees and the causal effect of schooling on health behavior. *Health Economics*, 19(8):994–1001, 2010.

[180] Arthur J. Reynolds, Judy A. Temple, Dylan L. Robertson, and Emily A. Mann. Long-term Effects of an Early Childhood Intervention on Education Achievement and Juvenile Arrest: A 15-Year Follow-up of Low-Income Children in Public Schools. *Journal of the American Medical Association*, 285:2339–46, 2001.

[181] Sherwin Rosen. The theory of equalizing differences. In Orley C. Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics*, volume 1, chapter 12, pages 641–92. Elsevier B.V., 1986.

[182] Mark R. Rosenzweig and T. Paul Schultz. The behavior of mothers as inputs to child health: The determinants of birth weight, gestation, and rate of fetal growth. In *Economic Aspects of Health*, NBER Chapters, pages 53–92. National Bureau of Economic Research, Inc, 1982.

[183] Catherine E. Ross and Chia-ling Wu. The links between education and health. *American Sociological Review*, 60(5):719–45, 1995.

[184] Ricardo Sabates and Leon Feinstein. Effects of Government Initiatives on Youth Crime. *Oxford Economic Papers*, 60(3):462–83, 2008.

[185] Allen Schirm, Elizabeth Stuart, and Allison McKie. *The Quantum Opportunity Program Demonstration: Final Impacts.* Mathematica Policy Research, Inc., 2006.

[186] Peter Schochet, John Burghardt, and Steven Glazerman. National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes. Mathematica Policy Research, June 2001.

[187] Lawrence J. Schweinhart, Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores. *Lifetime Effects: The High/Scope Perry Preschool Study through Age 40.* High/Scope Press, 2005.

[188] Amartya Sen. *Development as freedom.* Anchor Brooks, New York, 1999.

[189] Thomas Siedler. Schooling and citizenship: Evidence from compulsory schooling reforms. Discussion Papers of DIW Berlin 665, DIW Berlin, German Institute for Economic Research, 2007.

[190] Thomas Siedler. Schooling and citizenship in a young democracy: Evidence from postwar Germany. *Scandinavian Journal of Economics*, 112(2):315–38, 2010.

[191] Mary A. Silles. The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education Review*, 28:122–28, 2009.

[192] James P. Smith. Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives*, 13(2):145–66, 1999.

[193] James P. Smith. The impact of socioeconomic status on health over the life-course. *Journal of Human Resources*, XLII(4):739–64, 2007.

[194] Antonio Spilimbergo. Democracy and foreign education. *American Economic Review*, 99(1):528–43, 2009.

[195] Jos Tavares and Romain Wacziarg. How democracy affects growth. *European Economic Review*, 45(8):1341–78, 2001.

[196] Terence Thornberry and Marvin Krohn. The Self-Report Method for Measuring Delinquency and Crime. In D. Duffee et al., editors, *Criminal Justice 2000: Innovations in Measurement and Analysis.* National Institute of Justice, 2000.

[197] U.S. Department of Justice. *Profile of Inmates in the United States and in England and Wales, 1991.* Washington, DC, 1994.

[198] U.S. Department of Justice. *State Prison Expenditures, 1996.* Washington, DC, 1999.

[199] S. Verba, K. L. Scholzman, and H. E. Brady. *Voice and Equality.* Harvard University Press, Cambridge, 1995.

[200] W. Kip Viscusi. The value of risks to life and health. *Journal of Economic Literature*, 31(4):1912–46, 1993.

[201] W Kip Viscusi and Joseph E Aldy. The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76, 2003.

[202] Adam Wagstaff. The demand for health : Some new empirical evidence. *Journal of Health Economics*, 5(3):195–233, 1986.

[203] Joel Waldfogel. The Effect of Criminal Conviction on Income and the Trust 'Reposed in the Workmen'. *The Journal of Human Resources*, 29:62–81, 1994.

[204] David Weiner, Byron Lutz, and Jens Ludwig. The Effects of School Desegregation on Crime. NBER Working Paper Paper No. 15380, 2009.

[205] J. White, T. Moffitt, and P. Silva. A Prospective Replication of the Protective Effects of IQ in Subjects at High Risk for Juvenile Delinquency. *Journal of Consulting and Clinical Psychology*, 57:719–24, 1989.

[206] Ann D. Witte. Crime. In J. Behrman and N. Stacey, editors, *The Social Benefits of Education*, chapter 7. University of Michigan Press, Ann Arbor, 1997.

[207] Ann D. Witte and Helen Tauchen. Work and Crime: An Exploration Using Panel Data. NBER Working Paper 4794, 1994.

[208] Donald Wittman. Why democracies produce efficient results. *The Journal of Political Economy*, 97(6):1395–1424, 1989.

Figure 1: Regression-Adjusted Probability of Incarceration by Education (Men Ages 20-60)

(a) Whites
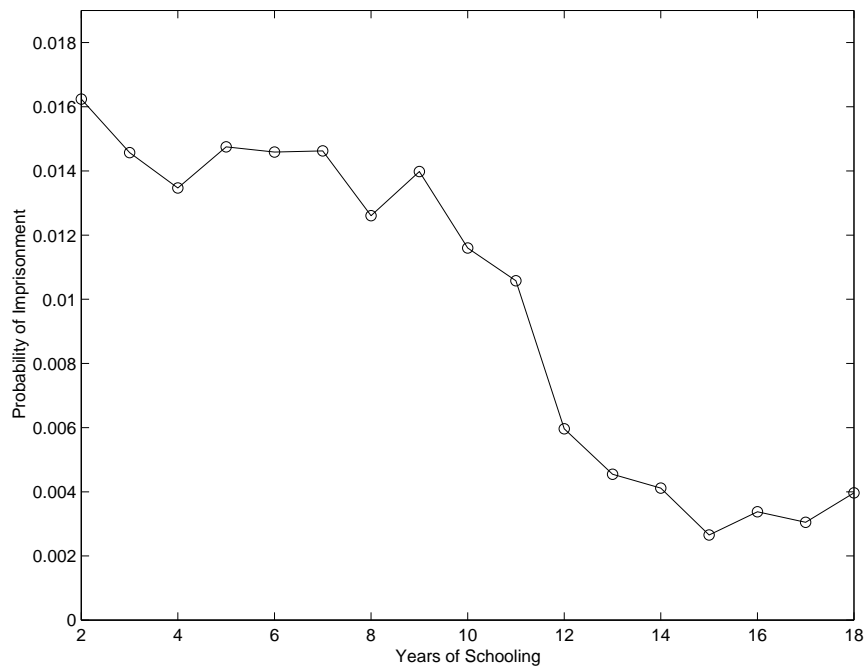
(b) Blacks
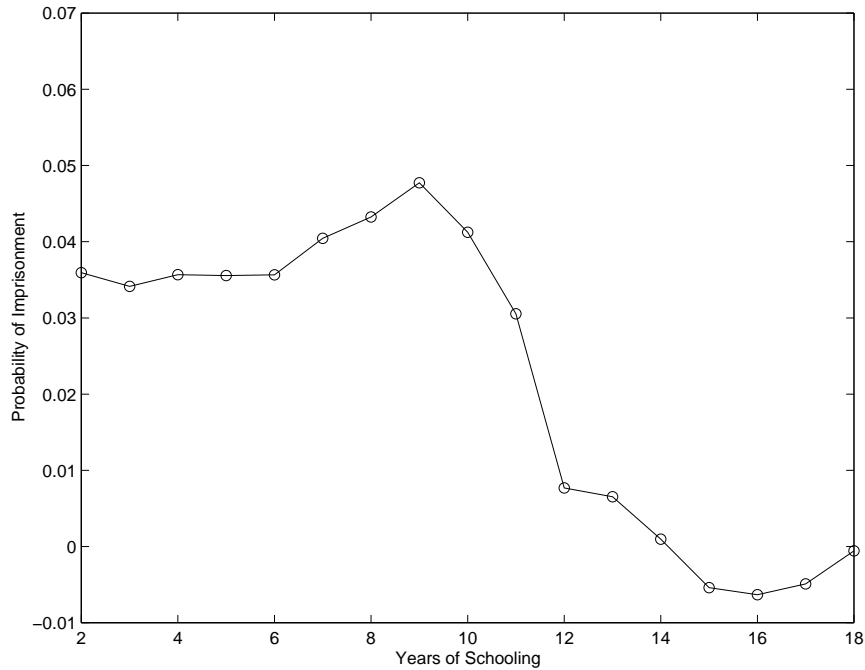
Notes: From 1960, 1970, and 1980 U.S. Censuses. Regressions control for age, state of birth, state of residence, cohort of birth, state, and year effects. Source: Lochner and Moretti (2004).

Table 1: Effects of Selected Early Childhood Programs on Educational Attainment and Adult Crime

Program	Program Description	Program Population	Methodology	Education Effects	Crime Effects
Abecedarian Project	full-time full-year pre-school from infancy to kindergarten	developmentally at-risk children, Chapel Hill, NC	Random Assign.	increased high school graduation rate by 0.03 (0.70 vs. 0.67) and enrollment in 4-yr college by 0.22[*] (0.36 vs. 0.14)	no sig. effects by age 21
Chicago Child Parent Center	half-day preschool (school year) ages 3 and 4	low-income minority children, Chicago, IL	Matched Sample	increased high school completion rate by 0.09 (0.57 vs. 0.48) for females and 0.14[*] (0.43 vs. 0.29) for males	by age 18, reduced fraction arrested by 0.08[*] (0.17 vs. 0.25)
High/Scope Perry Preschool	half-day preschool (school year) ages 3 and 4, bi-weekly home visits	low-income black children at risk of school failure, Ypislanti, MI	Random Assign.	increased high school graduation rates by 0.52[*] (0.84 vs. 0.32) for females and reduced grad. rates by 0.04 (0.50 vs. 0.54) for males	by age 40, reduced fraction arrested 5 or more times by 0.10[*] (0.24 vs. 0.34) for females and 0.24[*] (0.45 vs. 0.69) for males
Infant Health & Development Program (IHDP)	weekly/bi-weekly home visits from 0-36 months, full-time full-year pre-school 12-36 months	low birth-weight pre-term infants, 8 sites	Random Assign.	no sig. effect on high school dropout (approx. 10% dropout rate)	no sig. effects on arrests by age 18

Notes: Effects for Abecedarian Project taken from Campbell, et al. (2002) and Clarke and Campbell (1998). Effects for Chicago Child Parent Center taken from Reynolds, et al. (2001). Effects for Perry Preschool taken from Schweinhart, et al. (2005). Effects for IHDP taken from McCormick, et al. (2006). [*]denotes difference is statistically significant at 0.05 level.

Table 2: Social Benefits of Increasing High School Completion Rates by 1 Percent

	Total Cost per crime	Est. Change in Arrests	Est. Change in Crimes	Social Benefit
Violent Crimes				
Murder	4,506,253	-373	-373	$1,683,083,243
Rape	132,938	347	1,559	-$207,270,899
Robbery	13,984	134	918	-$12,839,495
Assault	14,776	-7,798	-37,135	$548,690,721
Property Crimes				
Burglary	1,471	-653	-9,467	$13,920,409
Larceny/Theft	295	-1,983	-35,105	$10,347,853
Motor Vehicle Theft	1,855	-1,355	-14,238	$26,414,558
Arson	58,171	-69	-469	$27,302,131
Total		-11,750	-94,310	$2,089,648,519

Notes: These costs reflect incarceration and victim costs. Victim costs are taken from Miller, et al. (1996). Incarceration costs per crime equal the incarceration cost per inmate multiplied by the incarceration rate for that crime (approximately $25,000). 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 non-reporting to the police). Incarceration costs per inmate are taken from U.S. Department of Justice (1999). All dollar figures are translated into 2008 dollars using the CPI-U. Source: Lochner and Moretti (2004).

Table 3: Expected Punishment Associated with Incarceration (Uniform Crime Reports)

Crime	Probability Cond. of Arrest	Probability of Conviction Cond. on Arrest	Probability of Incarceration on Conviction	Unconditional Probability of Incarceration	Estimated Months Served if Incarcerated	Expected Days Served per Crime Committed
Violent Crimes	0.25	0.22	0.79	0.043	91	119.4
Murder & Non-Negligent Manslaughter	0.85	0.67	0.95	0.544	248	4,102.4
Forcible Rape	0.15	0.39	0.90	0.051	136	212.2
Robbery	0.15	0.36	0.89	0.047	94	134.8
Aggravated Assault	0.30	0.17	0.71	0.035	59	63.2
Property Crimes	0.06	0.11	0.68	0.004	24	3.2
Burglary	0.07	0.27	0.76	0.015	29	13.2
Larceny-Theft	0.05	0.08	0.61	0.002	20	1.4
Motor Vehicle Theft	0.10	0.08	0.73	0.006	17	3.1

Notes: Probability of arrest computed from crimes and arrests in the U.S. (Uniform Crime Reports, 2000) adjusted for non-reporting to the police (National Criminal Victimization Survey, 2000). It is assumed that all murders are reported to the police. Probability of conviction conditional on arrest divides total arrests in the U.S. by total State and Federal convictions for 2000. Probability of incarceration conditional on conviction is based on reporting of State courts. Estimated months served if incarcerated applies to State prisoners and is estimated by the U.S. Department of Justice based on sentence lengths handed out that year and the average percent of sentences served by prisoners released that year. Unless otherwise noted, all criminal justice figures are for 2000 and are taken from Durose and Langan (2003) "Felony Sentences in State Courts, 2000". Source: Lochner (2004).

Table 4: Regression Discontinuity and Instrumental Variable Studies of the Effects of Years of Schooling on Mortality and Survival

Study	Data	Sample	Instrument	Important Features of Estimation Specification	Mortality or Survival Measure	Estimated Effect	
						No Instruments	RD/IV Methods
Mazumder (2008, 2010)	U.S. Censuses (1960-2000)	Native Whites Born 1901-25	State-Specific Changes in Compulsory Schooling Ages	2SLS with State-Specific Linear Cohort Trends	10-Year Mortality Rate	**-0.036** (0.003)	0.006 (0.031)
Lleras-Muney (2005, 2006)	U.S. Censuses (1960-1980)	Native Whites Born 1901-25	State-Specific Changes in Compulsory Schooling Ages	2SLS with Region-Specific Linear Cohort Trends	10-Year Mortality Rate	**-0.036** (0.004)	**-0.063** (0.024)
Clark & Royer (2010)	Birth & Mortality Data, Office for National Statistics, England and Wales (1970-2007)	Residents of England & Wales in 1970-2004, Born 1925-38	National Compulsory Schooling Age Increase in 1947	RD with Cohorts Defined by Quarter of Birth[a]	Prob. of Dying b/w Ages 45 & 69		Women Men 0.004 0.012 (Reduced Form Not Sig.)
Albouy & Lequien (2009)	French Censuses and Enchantillon Demographique Permanent (Various Years 1968-99), National Registers on Deaths (1968-2005)	French Individuals Born on First 4 Days of October in 1920-25 or 1950-55 and alive in 1968	National Compulsory Schooling Age Increases for Cohorts Born in 1923 & 1953	RD with 1 Year Cohort Window Around Schooling Age Change	Survival from Age 15 to 52 / Survival from Age 45 to 82		0.013 (0.011) / 0.063 (0.620)

Notes: Standard errors in parentheses. Estimates in bold are statistically significant at 0.05 level. [a]Clark and Royer (2010) do not report IV estimates; values reported in table reflect RD estimated effect of 1947 reform on death rates between ages 45 and 69 divided by the effect of the reform on average years of completed schooling.

Table 5: Recent U.S.-Based Instrumental Variable Studies of the Effects of Schooling on Health

Study	Data	Sample	Instrument	Schooling Measure	Health Outcome	Estimated Effect	
						No Instruments	**IV Methods**
						OLS/Probit (Marg. Eff.)	IV/2SCML (Marg. Eff.)
Mazumder (2008)	Survey of Income and Program Participation (1984, 1986-88, 1990-93, 1996)	Native Whites Born 1901-25	State-Level Compulsory Schooling Laws	Years of Education	Self-Reported Heath Fair or Poor	**-0.036** (0.001)	**-0.082** (0.034)
					Number of Nights in Hospital Last Year	**-0.073** (0.0186)	-1.083 (0.767)
					Trouble with Stairs	**-0.025** (0.001)	-0.007 (0.032)
					Health Limitation	**-0.025** (0.0013)	**-0.074** (0.035)
Oreopoulos (2006)	U.S. Censuses (1950-2000)	Natives Born 1901-56, Ages 25-84 at Time of Survey	State-Level Compulsory Schooling Laws	Average Years of Education		OLS	2SLS
					Disability that Limits Personal Care	**-0.014** (0.0003)	**-0.025** (0.006)
					Disability that Limits Mobility	**-0.020** (0.0004)	**-0.043** (0.007)

Study	Data	Sample	Instrument	Schooling Measure	Health Outcome	No Instruments		IV Methods	
Kenkel, Lillard & Mathios (2006)	National Longitudinal Survey of Youth, 1979 Cohort	Born 1957-64 Observed in 1998 (Oversamples Minorities & Poor Whites)	High School & GED Requirements & Policies	High School Graduate Indicator (GED results not reported here)		OLS		2SLS	
						Women	Men	Women	Men
					Currently Smoke	**-0.194** (0.030)	**-0.226** (0.029)	-0.102 (0.124)	**-0.229** (0.088)
					Obese	0.005 (0.029)	0.013 (0.026)	-0.021 (0.139)	-0.008 (0.082)

Study	Data	Sample	Instrument	Schooling Measure	Health Outcome	OLS	Risk of Induction	Risk of Induction x Risk of Being Killed
de Walque (2007)	National Health Interview Survey - Smoking Supplements (Various years 1983-1995)	Natives born 1937-56, aged 25+ at the time of the survey	Risk of Induction in Vietnam War	Number of Years of Education Above High School	Currently Smoke	**-0.040** (0.004)	-0.038 (0.020)	**-0.040** (0.019)
					Stopped Smoking	**0.041** (0.002)	0.051 (0.031)	**0.066** (0.029)
				Indicator for College Graduate or More	Currently Smoke	**-0.173** (0.015)	-1.169 (0.642)	**-0.181** (0.087)
					Stopped Smoking	**0.178** (0.010)	2.190 (1.395)	**0.297** (0.133)

Study	Data	Sample	Instrument	Schooling Measure	Health Outcome	OLS	2SLS
Grimard & Parent (2007)	Current Population Survey Tobacco Supplements (1995, 1996, 1998, 1999)	White native-born citizens at least age 25, born 1935-74	Vietnam Draft Cohort Indicator (Born 1945-50)	Years of Education	Currently Smoke Every Day	**-0.055** (0.002)	**-0.136** (0.024)
					Ever smoked	**-0.043** (0.002)	**-0.111** (0.029)

Notes: Standard errors in parentheses. de Walque (2007) provides t-statistics rather than standard errors; standard errors are calculated as the reported estimates divided by reported t-statistics. 2SCML reflects the two-stage conditional maximum likelihood estimator. Estimates in bold are statistically significant at 0.05 level.

Table 6: Recent European Instrumental Variable Studies of the Effects of Years of Schooling on Health

Study	Data	Sample	Instrument	Health Outcome	Estimated Effect	
					No Instruments	IV
Clark & Royer (2010)	Health Survey of England (1991-2004)	Residents of England in 1991-2004, Born 1918-72	National Compulsory Schooling Age Increases in 1947 & 1973		OLS	2SLS
				Self-Reported Health Fair or Bad	**-.081** (0.003)	-0.005 (0.014)
				Long Illness	**-0.033** (0.003)	0.023 (0.016)
				Reduced Activity	**-0.021** (0.002)	0.014 (0.012)
				Obese	**-0.024** (0.003)	0.028 (0.016)
				Hypertension	**-0.015** (0.003)	-0.025 (0.018)
				Currently Smoke	**-0.072** (0.003)	**-0.035** (0.014)
				Physically Active	**0.055** (0.004)	-0.001 (0.019)
Powdthavee (2010)	Health Survey of England (1991-2007)	Residents of England in 1991-2007, Born 1929-39 or 1952-65	National Compulsory Schooling Age Increases in 1947 & 1973		Probit (Marg. Eff.) Women / Men	IV Probit (Marg. Eff.) Women / Men
				Hypertension (1947 Law Change)	**-0.017** / **-0.015**	**-0.084** / **-0.070**
				Hypertension (1973 Law Change)	**-0.009** / **-0.008**	0.014 / 0.075
Oreopoulos (2006, 2008)	Annual U.K. General Household Surveys (1983-98) & Northern Ireland Continuous Household Surveys (1985-98)	Native-Born British & N. Ireland Residents Born 1921-51, Ages 25-84 at Time of Survey	Increased Compulsory Schooling Ages in G. Britain (1947) & N. Ireland (1957)		OLS	2SLS
				Self-Reported Poor Heath	**-0.037** (0.002)	0.007 (0.008)
				Self-Reported Good Heath	**0.065** (0.002)	-0.010 (0.011)
Silles (2009)	General Household Survey for England, Scotland and Wales (1980-2004)	U.K. Individuals Ages 25-60	National Compulsory Schooling Age Increases in 1947 & 1973		OLS / 2SLS / RD	
				Self-Reported Good Heath	**0.026** (0.001) / **0.045** (0.009)	**0.064** (0.003)
				No Long-Term Illness	**0.008** (0.001) / **0.055** (0.009)	**0.075** (0.003)
				No Activity-Limiting Illness	**0.008** (0.001) / **0.046** (0.008)	**0.051** (0.002)
Kemptner, Jurges & Reinhold (2010)	German Microcensus (1989, 1995, 1999, 2002, 2003)	Germans Born 1930-60, Living and Educated in Former West German States	Increased State-Specific Compulsory Schooling from 8 to 9 Years Between 1949 and 1969		OLS Women / Men	2SLS Women / Men
				Long-Term Illness	**-0.012** (0.002) / **-0.029** (0.001)	0.013 (0.024) / **-0.039** (0.019)
				Work Limiting Illness	**-0.013** (0.001) / **-0.029** (0.001)	0.036 (0.022) / **-0.037** (0.017)
				Overweight	**-0.047** (0.002) / **-0.031** (0.001)	-0.016 (0.016) / **-0.034** (0.015)
				Obese	**-0.020** (0.001) / **-0.018** (0.001)	0.007 (0.012) / -0.028 (0.015)
				Ever Smoked	**-0.005** (0.002) / **-0.026** (0.001)	0.014 (0.011) / -0.012 (0.013)
				Currently Smoke	**-0.021** (0.002) / **-0.032** (0.001)	0.001 (0.01) / 0.005 (0.012)
Brunello, Fabbri & Fort (2009)	ECHP (1998), SHARE (2004), GSOEP (2002), BHPS (2003), Enquete sur la Sante (2003)	European Women (from 10 Countries) Ages 25-65 Born Within 7 Years of Cohort First Affected by School Law Changes	Increased Compulsory Schooling Ages by Country		Probit (Marg. Eff.)	IV Probit (Marg. Eff.)
				Overweight	**-0.022** (0.001)	**-0.044** (0.021)
				Obesity	**-0.012** (0.001)	-0.012 (0.013)

Notes: Standard errors in parentheses. Estimates in bold are statistically significant at 0.05 level.

Table 7: Effects of Years of Education on Health and Mortality by Outcome (Best IV/RD Estimates from Selected Studies)

Study	Country	Outcome Details	Sample Average		Estimated Effect	Effect as % of Average
A. Mortality						
Mazumder (2008)	U.S.	10-Year Mortality Rates	0.213		0.006	2.8%
Lleras-Muney (2005, 2006)	U.S.	10-Year Mortality Rates	0.106		-0.063*	-59.4%
Clark & Royer (2010)[a]	U.K.	Prob. of Dying b/w Ages 45 & 69	Women: 0.146		0.004	2.7%
			Men: 0.221		0.012	5.4%
Albouy & Lequien (2009) [b]	France	37-Year Mortality Rate	Ages 15-42:	0.07	-0.013	-18.6%
			Ages 45-82:	0.40	-0.063	-15.8%
B. Self-Reported Health						
Mazumder (2008)	U.S.	Self-Report Fair or Poor	0.357		-0.082*	-23.0%
Clark & Royer (2010)	U.K.	Self-Report Fair or Bad	0.25		-0.005	-2.0%
		Long Illness	0.46		0.023	5.0%
Oreopoulos (2006, 2008)	U.K.	Self-Report Poor Health	0.150		0.007	4.7%
		Self-Report Good Health	0.564		-0.01	-1.8%
Silles (2009)	U.K.	Self-Report Good Health	0.661		0.064*	9.7%
		No Long-Term Illness	0.685		0.075*	10.9%
Kemptner, Jurges & Reinhold (2010)	Germany	Long-Term Illness	Women: 0.15		0.013	8.7%
			Men: 0.20		-0.039*	-19.5%
C. Disability, Limited Mobility, Activity						
Mazumder (2008)	U.S.	Health Limitation	0.423		-0.074*	-17.5%
Oreopoulos (2006)	U.S.	Disability Limiting Personal Care	0.092		-0.025*	-27.2%
		Disability Limiting Mobility	0.128		-0.043*	-33.6%
Clark & Royer (2010)	U.K.	Reduced Activity	0.16		0.014	8.8%
		Physically Active	0.44		-0.001	-0.2%
D. Smoking						
Clark & Royer (2010)	U.K.	Currently Smoke	0.27		-0.035*	-13.0%
Kenkel, et al. (2006)[c]	U.S.	Currently Smoke	Women: 0.26		-0.229*	-88.1%
			Men: 0.27		-0.102	-37.8%
de Walque (2007)[d]	U.S.	Currently Smoke	0.40		-0.040*	-10.0%
Grimard & Parent (2007)	U.S.	Currently Smoke Every Day	0.31		-0.136*	-43.9%
		Ever Smoked	0.58		-0.111*	-19.1%
Kemptner, Jurges & Reinhold (2010)	Germany	Currently Smoke	Women: 0.24		0.001	0.4%
			Men: 0.36		0.005	1.4%
		Ever Smoked	Women: 0.39		0.014	3.6%
			Men: 0.63		-0.012	-1.9%
E. Obesity						
Clark & Royer (2010)	U.K.	Obese	0.21		0.028	13.3%
Kemptner, Jurges & Reinhold (2010)	Germany	Overweight	Women: 0.44		-0.016	-3.6%
			Men: 0.66		-0.034*	-5.2%
		Obese	Women: 0.13		0.007	5.4%
			Men: 0.16		-0.028	-17.5%
Brunello, Fabbri & Fort (2009)	10 European Countries	Overweight (Women)	0.387		-0.044*	-11.4%
		Obese (Women)	0.114		-0.012	-10.5%
Kenkel, et al. (2006)[c]	U.S.	Obese	Women: 0.27		-0.021	-7.8%
			Men: 0.25		-0.008	-3.2%

Notes: [a]Clark and Royer (2010) do not report IV estimates for the effect of education on mortality; values reported in table reflect RD estimated effect of 1947 reform on death rates between ages 45 and 69 divided by the effect of the reform on average years of completed schooling. [b]Mortality estimates for Albouy & Lequien (2009) are calculated from survival rates; mortality rates inferred from from Figures 1 and 2. [c]Kenkel, et al. (2006) estimates are for high school graduation rather than years of schooling. [d]Average smoking rates for de Walque (2007) inferred from Figure 1 for those with a high school degree. * denotes statistical significance at 0.05 level.

Table 8: Channels through Which Education May Improve Health

Channel	Private Costs	Benefits Private or Public?
1. Reduce Stress	None	Private
2. Better Decisionmaking Ability/Use of Inputs	None	Private
3. Better at Gathering/Interpreting Information	None	Private
4. Health Insurance	Financial	Private
5. Healthier lifestyle		
a. Safety Precautions (e.g. seatbelts, smoke alarms)	Utility, Financial	Private
b. Diet	Utility, Financial	Private
c. Exercise	Utility, Financial	Private
d. Non-Smoking, Alcohol Moderation, Avoiding Drugs	Utility, Saves Money	Private
6. Healthier/Safer Employment	Lower Wages	Private
7. Healthier Neighborhoods	Housing Prices	Private
8. Healthier Peers and Friends	None	Public

Table 9: Effects of Education on Voter Registration and Voting

Study	Data	Education Measure	Instrumental Variable	Voting/Registration Measure	Sample Mean	IV Estimate	% Effect
A. United States							
Dee (2004)	HS&B	college enroll.	local 2-yr colleges	Registered to vote	0.67	0.218*	32.5
				Voted in last 12 months	0.36	0.176*	48.9
				Voted in 1988 presidential election	0.55	0.178	32.4
Dee (2004)	GSS	years of school	state school/labor laws	Voted in last presidential election	0.73	0.068*	9.3
Milligan et al. (2004)	NES	high school grad.	state school/labor laws	Registered to vote	0.82	0.093	11.3
				Voted in last election	0.68	0.288*	42.3
Milligan et al. (2004)	CPS	high school grad.	state school/labor laws	Voted in last election	0.62	0.435*	70.4
B. United Kingdom							
Milligan et al. (2004)	BES	years of school	national min. school age	Voted in last election	0.85	-0.008	-0.9
Milligan et al. (2004)	EB	years of school	national min. school age	Name on electoral list	0.92	-0.014	-1.5
C. Germany							
Siedler (2007, 2010)	ALLBUS	years of school	national min. school age	Voted in last election	0.93[a]	0.005	0.5
Siedler (2007, 2010)	ForsaBus	years of school	national min. school age	Voted in last election	0.93[a]	0.000	0.0

Notes: HS&B = High School and Beyond; GSS = General Social Surveys NES = National Election Survey; CPS = Current Population Survey November Voting Supplements; BES = British Election Studies; EB = Eurobarometer surveys for Britain and Northern Ireland. [a]Sample means taken from ALLBUS data in Table 3 of Siedler (2007) for those 'not affected by the reform'. * reflects statistically significant at 0.05 level.

Table 10: Effects of Education on Political Interest, Information, Views, and Participation

Study	Country	Data	Education Measure	Instrumental Variable	Outcome	Sample Mean	IV Estimate	% Effect
A. Political Interest								
Milligan et al. (2004)	U.S.	NES	high school grad.	state school/labor laws	Interested in election	0.30	0.270*	90.0
Siedler (2007, 2010)	Germany	ALLBUS	years of school	national min. school age	Interested in politics	0.30[a]	-0.096*	-32.0
Siedler (2007, 2010)	Germany	ForsaBus	years of school	national min. school age	Interested in politics	0.30[a]	-0.008	-2.7
B. Political Information								
Dee (2004)	U.S.	GSS	years of school	state school/labor laws	Newspaper readership index (0-4)	3.2	0.113*	3.5
Milligan et al. (2004)	U.S.	NES	high school completion	state school/labor laws	Follow campaign on TV	0.79	0.392*	49.6
					Follow campaign in newspapers	0.66	0.852*	129.1
					Follow public affairs	0.66	0.544*	82.4
Milligan et al. (2004)	U.S.	EB	years of school	national min. school age	Follow news every day	0.9	-0.007	-0.8
C. Political Views/Philosophy								
Dee (2004)	U.S.	GSS	years of school	state school/labor laws	Allow homosexual to speak	0.73	0.123*	16.8
					Allow militarist to speak	0.59	0.036*	6.1
Siedler (2007, 2010)	Germany	ALLBUS	years of school	national min. school age	Demonstrating is civil right	0.67[a]	-0.001	-0.1
D. Political/Social Participation								
Dee (2004)	U.S.	HS&B	college enroll.	local 2-yr colleges	Volunteered in last 12 months	0.37	-0.047	-12.7
Dee (2004)	U.S.	GSS	years of school	state school/labor laws	Group memberships	1.8	0.164	9.1
Milligan et al. (2004)	U.S.	NES	high school grad.	state school/labor laws	Attend political meetings	0.07	0.132	188.6
Milligan et al. (2004)	U.K.	EB	years of school	national min. school age	Often or occasionally try to persuade others to share views	0.45	0.066*	14.7
					Discuss political matters with friends at least occasionally	0.67	0.095*	14.2
Siedler (2007, 2010)	Germany	ALLBUS	years of school	national min. school age	Active in citizen group	0.17[a]	-0.050	-29.4
					Participated in demonstration	0.09[a]	-0.016	-17.8

Notes: HS&B = High School and Beyond; GSS = General Social Surveys NES = National Election Survey; EB = Eurobarometer surveys for Britain and Northern Ireland.

[a]Sample means taken from ALLBUS data in Table 3 of Siedler (2007) for those 'not affected by the reform'. * reflects statistically significant at 0.05 level.