

--- PRELIMINARY AND INCOMPLETE ---

The Effect of College Education on Health

Kasey Buckles, University of Notre Dame

Ofer Malamud, University of Chicago and NBER

Melinda Morrill, North Carolina State University

Abigail Wozniak, University of Notre Dame, NBER and IZA

This version: May 2012

ABSTRACT

The health consequences of increased college attainment are important for calculating returns to schooling as well as formulating health policy. We examine the causal impact of education on health outcomes using variation in college attainment induced by draft-avoidance behavior during the Vietnam War. We exploit both national and state-level induction risk to identify the effect of educational attainment on cohort-level mortality using decadal U.S. Censuses and Vital Statistics data from 1981 to 2007. Our preferred 2SLS estimates imply that college completion reduces cumulative mortality in our affected cohorts by approximately 100 deaths per 1,000 persons over this 27 year period. This is similar in magnitude and significance to the OLS estimates. The reduction in mortality is larger in the 1990s and 2000s than in the 1980s, and the effects are largest for deaths from cancer and heart disease. We explore some potential mechanisms for these effects using data from the ACS and NHIS.

(JEL: I12, I23, J24)

Acknowledgments: The authors would like to thank Bill Evans, Bob Kaestner and seminar participants at the Population American Association (PAA) Meetings, University of Illinois at Chicago, and the University of Notre Dame for helpful comments. All errors are our own.

I. Introduction

Schooling is highly correlated with subsequent health outcomes. For example, in 2007, the age-adjusted mortality rate of high school graduates aged 25 to 64 was more than twice as large as the mortality rate of those with some college or a collegiate degree (National Vital Statistics Reports, 2010). If these associations between health and education reflect casual effects, they would represent a significant non-pecuniary return to education. They would also imply that policies to increase educational attainment could serve as an important means for improving health. However, there is substantial debate about whether these associations actually represent causal effects (see the reviews by Grossman, 2006, and Cutler and Lleras-Muney, 2008). This paper contributes to the growing literature on the relationship between health and education by estimating the causal impact of education on health behaviors and health outcomes using variation in college attainment induced by draft-avoidance behavior during the Vietnam War.

We use an instrumental variables strategy developed in Malamud and Wozniak (forthcoming) to identify the effect of education on the health status of men who were eligible to be drafted into the Vietnam War. This strategy builds on Card and Lemieux (2000, 2001) who document the excess educational attainment among cohorts induced to enter college in order to defer conscription. While Card and Lemieux focus on differences in induction risk across birth cohorts, we also exploit state level variation in induction risk within cohorts. The existence of state level variation allows us to decompose national induction risk into its constituent parts: induction risk faced by a young man's own state cohort and risk faced by young men of that cohort in the rest of the country. This decomposition yields two instruments, which we use to identify the two endogenous variables—education and veteran status—in our empirical application. This approach is an advance over studies that used the Card and Lemieux measures to identify the effect of college

going on health outcomes using only year to year variation in induction risk (De Walque 2007, Grimard and Parent 2007, and MacInnis 2006).

We merge our data on national and state-level induction risk with the Vital Statistics Mortality Files from 1981 to 2007 and the U.S. Censuses for 1980, 1990, and 2000 to construct birth state by cohort level mortality rates. We also include information about basic demographic characteristics, cohort size, and labor market conditions at the time of entry from a variety of additional sources. As in Malamud and Wozniak (forthcoming), our instruments predict both veteran status and educational attainment for men in the affected cohorts, with the increase in education coming primarily from increased post-secondary schooling attainment. We therefore have a viable instrument for educational attainment at higher levels that can be purged of its correlation with veteran status for the Vietnam cohorts.

We first establish that the well-known gradient between education and health status is present and statistically significant in our Vital Statistics data using OLS specifications. We focus on the cumulative mortality rate between 1981 and 2007 and show that college graduation is associated with 100 fewer deaths per 1,000 persons. Our instrumental variables estimates indicate an effect that ranges from 93 to 172 fewer deaths per 1,000 persons, with our preferred specification yielding a magnitude similar to the OLS estimate. Furthermore, none of these 2SLS estimates are significantly different from the OLS estimates. Using 10-year mortality rates for the 1980s, 1990s and 2000s, we find 2SLS effects that are somewhat larger in magnitude than our OLS estimates, especially in later decades. The largest effects are found for the impact of college completion in lowering death due to cancer and heart disease, which represent the leading causes of mortality among the elderly. We also explore some potential mechanisms for our effects of college education on mortality using the American Community Survey (ACS) and the National Health Interview Survey (NHIS).

The findings in this paper have important implications for both health and education policy. Given that people value health very highly, the health returns to education may represent a substantial fraction of the financial returns. Indeed, Cutler and Lleras-Muney (2010) calculate that their estimates of the health returns to education increase the total returns to education by 15 to 55 percent. If individual investments in college education are suboptimal because of credit constraints, externalities, or lack of information, the presence of additional health returns to college provides an even stronger case for subsidizing education. This is particularly relevant for recent discussions about the rising cost of college and the decline in federal financial aid for college students. On the other hand, a positive causal impact of higher education on health may pose a dilemma for health policy. Health improvements, like smoking cessation, may reduce health care costs in the short run only to increase them in the long run as individuals live longer or as other unforeseen consequences arise.¹ Our analysis will help inform policymakers interested in the connection between education policy and national healthcare spending.

This paper also helps to fill an important gap in the literature on the relationship between education and health. Previous analyses of the causal impacts of education on health outcomes, such as mortality, have relied on variation at the lower part of the schooling distribution. For example, Lleras-Muney (2005) and Clark and Royer (2010) exploit changes in compulsory schooling requirements to examine whether increased schooling improved the health of high school students on the margin of dropping out. Lleras-Muney (2005) finds large and significant effects of increased education on declines in mortality in the United States, whereas Clark and Royer (2010) find no evidence for an impact of education on mortality in England.² However, regardless of the causal impact of schooling on health at the margin of dropping out of high school, the causal relationship

¹ Some of these possibilities are explored in Bearman et al. (2011).

² Arendt (2005) and Albouy and Lequien (2009) also examine the impact of compulsory school reforms on health outcomes in Denmark and France respectively, but their estimated effects have large standard errors.

may be different at the margin between high school and college. Moreover, estimating the effect of education on health at the college margin may be of particular interest given that the largest increase in educational attainment in recent years have occurred among students entering college (Turner, 2004).

The effect of postsecondary education on certain health behaviors, smoking in particular, has been examined in previous literature. De Walque (2007) and Grimard and Parent (2007) exploit year-to-year variation in induction risk faced by cohorts of young men during the Vietnam War to identify the impact of education on smoking. Using different datasets (NHIS and the CPS Tobacco Supplements, respectively) and different specifications, they find that additional education has a negative and significant effect on the likelihood of smoking.³ Our paper extends this identification strategy by incorporating within-cohort variation in induction risk. Moreover, we examine a wide range of health outcomes across multiple decades, providing a much broader picture of higher education's potential health impacts.

Finally, this paper contributes to research examining the impact of military service during the Vietnam War on health outcomes and behaviors (Angrist, Chen, and Frandson, 2009; Dobkin and Shabani, 2007; Conley and Heerwig, 2009; Hearst, Newman, and Hulley, 1986). These studies all exploit variation in veteran status induced by the Vietnam draft lottery, and none reject that the impact of veteran status on health outcomes is zero.⁴ As we explain later, our results might not be directly comparable to these studies because we use a different identification strategy, which likely results in a different local average treatment effect. Nevertheless, our results help to shed light on this important relationship.

³ MacInnis (2006) uses a similar identification strategy to document the effect of education in reducing obesity and its co-morbidities such as hypertension and adult-onset diabetes.

⁴ Angrist, Chen, and Frandson (2009) find no significant effect on work-limiting disability in the 2000 Census, although some increase in non-work-limiting disability rates for whites. Dobkin and Shabani (2007) fail to find any precise 2SLS estimates on self-reported health and other health outcomes using the NHIS data. Conley and Heerwig (2009) find no significant effect of veteran status on mortality using Vital Statistics data from 1989-2000 in contrast to Hearst, Newman, and Hulley (1986) who show a short-term increase in mortality.

II. Background on the Vietnam Draft

Our instrumental variables strategy exploits variation in the risk of induction (“being drafted”) to which young men in the US were exposed during the Vietnam conflict. This section provides a brief overview of the sources of this variation. A more detailed discussion can be found in Malamud and Wozniak (forthcoming); hereafter MW.

Approximately 2 million American men were drafted during the Vietnam War. The Selective Service System, which comprised over 4,000 local draft boards across the nation at that time, was responsible for registering recruits, classifying them for either deferment or selection. Responsibility for devising and meeting the national target number of conscriptions rested with the federal Department of Defense (DoD). To achieve this target, the DoD issued monthly “draft calls” that divided the national number into quotas assigned to state draft boards, which did the active work of ordering men to be inducted.

Faced with an excess supply of eligible draft age men and no way to select among them, draft boards adopted generous deferral policies toward large categories of men. College enrollment was the second most common deferral category, after the exemption for dependents (Semiannual Reports of the Director of the Selective Service System, 1967-1973).⁵ The Military Service Act of 1967 codified the existing de facto arrangement by stating that college students in good standing could defer induction until receipt of an undergraduate degree or age 24, whichever occurred first. Over 1.7 million college deferments were granted in 1967 alone.⁶ Although men who received college deferments were technically eligible for induction until age 35, very few men between the

⁵ Tatum and Tuchinsky, *Guide to the Draft*, Ch. 3. By contrast, enrollment in a two year college was not considered grounds for automatic deferment. See Rothenberg (1968).

⁶ The number of college deferments remained above 1.7 million in 1968 and 1969, and then fell to 1.5 million and 1.3 million in 1970 and 1971 respectively (Semi-annual reports of the director of the Selective Service System, 1967-1973).

ages of 26 and 35 were ever drafted. Card and Lemieux (2000) estimate that, among men born between 1945 and 1947, those with a college degree were only one-third as likely to serve in Vietnam as compared to those without a college degree. Thus, the incentive to enroll in college to avoid the draft during these years was large.

Our identification strategy relies on two sources of variation in induction risk: over time and across states. The existence of intertemporal variation is well-known (Card and Lemieux 2000) and has been used in previous research (De Walque 2007, Grimard and Parent 2007). Inductions varied considerably over the course of the Vietnam War. From 1960 to 1963, inductions were fairly low at approximately 8,000 a month. However, following the Gulf of Tonkin incident in August 2, 1964, Congress authorized an expanded role for the U.S. military in Vietnam. Inductions more than doubled from 1964 to 1965 and again from 1965 to 1966. By the spring of 1968, in the midst of the raging student protests, the rate of inductions reached a peak of almost 42,000 a month.

The introduction of the draft lottery in 1969 led to a substantial change in the induction process. However, college deferments continued to be issued until September 1971, and men who were already enrolled were allowed to retain their deferment until the end of the school year. Risk of induction during this period was also much lower since men were at risk of induction for only a single year and the overall rate of inductions was substantially lower, falling from about 20,000 per month in late 1969 to 2,000 per month in late 1971. In February of 1973, the draft was suspended and no more inductions took place.

The existence of state level variation in induction risk was less well known at the time, and remains so today. This type of variation arose through two channels. The first was uneven application and idiosyncratic application of formal procedures across the thousands of local draft

boards.⁷ In their influential study of the draft, Davis and Dolbeare write, “The conclusion seems inescapable: local board autonomy implies both within state and between state variability, even among socioeconomically similar board jurisdictions.”⁸ Similar idiosyncrasies were reported in the report of the National Advisory Commission on Selective Service, in 1967. A second source of state and year variation in induction risk was communication delay between the federal, state, and local officials. These delays meant that the DoD assigned quotas using registrant numbers that were several months old. Thus, draft risk for an eligible man at a point in time was not only a function of the number of men in his state currently eligible for the draft but also of the number available several months ago. The current pool could be much larger than the past pool if, for example, a large number of local men graduated high school thus becoming draft eligible or much smaller if a large number married or aged out of the draft pool in the intervening months.

III. Empirical Strategy

We use variation in college attainment caused by draft-avoidance behavior during the Vietnam conflict to isolate the causal effect of education on migration. An artifact of this identification strategy is that the likelihood an individual is a veteran also varies systematically across cohorts in our sample. Given that veteran status is a plausible determinant of migration (Pingle 2007), it is important to control for this variable in our estimation strategy. However, selection into military service during the Vietnam War was likely based on characteristics that are unobserved in our data, which would confound our estimates of veteran status and potentially other covariates as

⁷ Davis and Dolbeare, *Little Groups of Neighbors*, Page 18.

⁸Davis and Dolbeare, *Little Groups of Neighbors*, Page 84.

well. To deal with this, we exploit changes in both national and state-level induction risk to generate exogenous variation in both college-going and veteran status.⁹

This strategy is identical to the one described in MW. However, we extend this analysis in two ways. First, we provide a more detailed discussion of the roles of direct and indirect effects in the system that we study. This is particularly important in the context of health, since direct and indirect channels have received much attention from health statisticians – and some attention from economists – trying to understand causal pathways to health. Second, we experiment with a variety of first stage specifications in order to demonstrate that our results are robust to alternative specifications of our estimating equations. This is especially relevant in the context of health where there has been some controversy about alternative specifications in identifying the impact of education on health (Lleras-Muney, 2005 and Mazumder 2008).

A. A Structural Model with Direct and Indirect Effects

We begin by presenting a simple structural model that relates the direct and indirect effects of veteran status and college-going on migration. Ignoring other covariates for conciseness, our model is the following:

$$(1) \text{ health}_{sc} = \alpha_0 + \alpha_1 \text{ vet}_{sc} + \alpha_2 \text{ educ}_{sc} + \varepsilon_{sc}^h$$

$$(2) \text{ vet}_{sc} = \beta_0 + \beta_1 \text{ educ}_{sc} + \varepsilon_{sc}^v$$

$$(3) \text{ educ}_{sc} = \gamma_0 + \gamma_1 \text{ vet}_{sc} + \varepsilon_{sc}^e$$

⁹ A number of related papers face a similar problem. Grimard and Parent (2007) exclude veteran status from their main estimating equation. De Walque (2007) includes veteran status but treats it as an exogenous variable. MacInnis (2006) instruments for veteran status using a quartic polynomial in age (specifically, the difference between 19 and age in 1967).

where s and c refer to birth state and birth cohort respectively. We omit individual subscripts because our source of exogenous variation occurs at the birth state-birth cohort level, and we estimate our specifications using data aggregated accordingly. In this system, α_1 and α_2 are the direct effects of *veteran status* on *health* and *education* on *health*, respectively. However, because of the possible relationships between education and veteran status, a change in either veteran status or schooling may have both a direct effect through the structural parameter in (1) and an indirect effect through its influence on the other variable in either (2) or (3).

In the context of our Vietnam-era setting, we assume that β_1 is equal to zero. In other words, we assume that education only has a direct effect on migration. Although college students could delay and eventually avoid conscription by staying in school (indeed, this is a key component of our identification strategy), the overall induction rates for a particular birth state-birth cohort were unlikely to be affected by draft-avoidance behavior. This is because local draft boards needed to fulfill specific manpower requirements set by the Department of Defense. So if a certain individual avoided the draft, someone else from his state-cohort would need to be drafted instead. Thus, to a first approximation, veteran status at the state-year cohort level was unaffected by college-going decisions.

On the other hand, γ_1 is not equal to zero. Recent work by Angrist and Chen (2011) indicates that veterans were more likely to attend college than non-veterans, primarily as a result of the educational benefits conferred to them by the GI Bill. If we substitute (3) into (1), we can derive an expression for the total effect of veteran status on migration:

$$(4) \text{ health}_{sc} = (\alpha_0 + \alpha_2\gamma_0) + (\alpha_1 + \alpha_2\gamma_1)\text{vet}_{sc} + \alpha_2\varepsilon_{sc}^e + \varepsilon_{sc}^h$$

Thus, the total effect of veteran status on migration is the sum of the direct effect, α_1 , and the indirect effect, $\alpha_2\gamma_1$. Exogenous variation in veteran status, such as the Vietnam draft lottery used in

Angrist and Chen (2011), would yield an unbiased estimate of the total effect, $(\alpha_1 + \alpha_2\gamma_1)$, of veteran status on migration in (4). However, without additional assumptions, estimates of the direct effects of veteran status and college in (1) may be biased.¹⁰

To summarize, our system contains one causal factor (education) that has a *direct effect* on health and another causal factor (veteran status) that has both direct and indirect effects. The labels “direct” and “indirect” refer to pathways within the specified system. It is therefore straightforward to use random assignment to identify the direct causal effect of education on health (Pearl 2005; Mealli and Rubin 2003). For veteran status, there are multiple causal pathways in the model, and identification is more complicated. While the total effect is identified, the direct and indirect effects are not separately identified without further assumptions, even in the ideal instance of random assignment of all causal factors in the model, (Robins and Greenland 1992). Note, it is possible that education impacts health works through other mediating variables and thereby has both direct and indirect effects as well. In that case, α_2 would represent the total effect of both direct and any other unspecified indirect causal pathways.

We conclude that exogenous variation in both education and veteran status allows us to identify the causal impact of education on health but not necessarily the direct causal impact of veteran status – at least not without further functional form assumptions. Our estimates on veteran status should therefore be viewed conservatively as reflecting both direct and indirect effects of veteran status, with the latter operating through increased educational attainment made possible by the GI Bill. We will therefore refrain from interpreting our estimates of the impact of veteran status as direct causal impacts of veteran status. This is important for two reasons. First, in our setting it is not clear that any returns to veteran status reflected in the 2SLS coefficients can be attributed solely

¹⁰ This is analogous to the potential bias due associated with controlling for a post-treatment variable (Rosenbaum, 1984). A derivation of this bias in the context of an earnings equation where the measure of ability is affected by schooling can be found in Appendix 1 of Angrist and Krueger (1999).

to veteran status. It is therefore inappropriate to use our estimates on veteran status to inform policy toward veterans. Second, different instruments operate on different segments of the population, and the possibility of heterogeneous treatment effects makes it difficult to directly compare IV estimates across studies that use different instruments. This is true of any two studies, but the comparison of estimates that involve both direct and indirect effects (and heterogeneous treatment effects within both) is likely to be even more problematic.

B. Identification

To provide identification for estimating the direct causal effect of college on health, we employ the same strategy as MW, who in turn extend an instrumental variables strategy inspired by Card and Lemieux (2000, 2001); henceforth CL. Like CL, we assume that draft avoidance was proportional to the risk of induction. To account for the mechanical relationship between inductions and veteran status, we exploit state level variation within the cohort level variation identified by CL. The existence of state-cohort level variation allows us to break national induction risk into its constituent parts, and obtain two separate instruments that can be used to identify both college attainment and veteran status. Thus, young men faced state cohort risk that is analogous to the CL measure, where s indexes state of residence:

$$(5) \quad \textit{state cohort risk}_{sc} = \frac{\sum^t I_{sct}}{N_{sc}}$$

This measure of state cohort risk is our first instrument. We then use our state level data on I and N to construct a second instrument in the following manner:

$$(6) \quad \textit{national cohort risk}_{sc} = \frac{\sum^t I_{-sct}}{N_{sc}}$$

This measure defines national cohort risk for a man living in state s and born in the (annual) birth cohort c as the number of inductees from all other states, $-s$, and six month birth cohort c , divided by the total number of such men.¹¹

To construct the measures in equations (5) and (6), we obtained data on the number of inductees from each state for each six-month period spanning 1961 to 1972 from reports of the Selective Service. We estimate state cohort size using enrollment numbers spanning 1959 to 1970, the academic years in which our cohorts of interest were in 11th grade. Thus state-cohort level risk (henceforth state risk) for a young man born in Alabama in the first half of 1950 equals the number of inductees from Alabama in the first half of 1969 (the year he turned 19) divided by half the number of students enrolled in 11th grade in Alabama in 1967. National level risk for the same young man roughly equals the number of men inducted nationally in the first half of 1969 divided by the size of his six month birth cohort; more precisely, we subtract own state inductions from the numerator and own state cohort size from the denominator. Like CL, we construct an average draft risk for the years a man was 19 to 22 since draft risk was non-trivial for men ages 20 to 22.

Figure 1 graphs the variation in state-level induction risk for a selection of states over our period of interest. Panel A shows raw induction risk as defined in (5) while Panel B shows residual induction risk after controlling for a cohort trend, state-of-birth fixed effects, and national risk. Given that we include these controls in our main empirical specifications, the patterns in Panel B more closely approximate our identifying variation. For a more detailed discussion, see MW.

C. Validity of Our Instruments

We present our first stage estimates in the results section, but given the novelty of our identification strategy, some of the IV assumptions bear more discussion. Our instruments could fail

¹¹ MW were able to use variation at the six month birth cohort level.

if young men attempted to exploit local variation in induction risk by moving between localities. In this case, risk would not be truly randomly assigned. Our risk measures would only bind for men who were unwilling or unable to move to low risk jurisdictions, which might in turn be correlated with other unobservable characteristics related to health. MW document that this this type of “local board shopping” was prohibited by draft board regulations.

Identification further requires the assumption that induction risk only affected health through either education or veteran status. There are two ways in which this might fail. First, the health screenings required to determine draft eligibility might uncover an important health condition earlier than it might otherwise have been detected, thereby encouraging individuals to treat the condition and improving future health. Our reading of the historical literature suggests that this was unlikely.¹² The required exams were very standard and did not involve any invasive testing for health conditions not easily observed in a doctor’s office. It is therefore unlikely these exams provided young men with information they did not already have. Second, in some instances, marriage was considered a way to reduce one’s chances of being drafted, and marriage may confer health benefits. However, there is little evidence that marriage or fatherhood rates increased with our measures of induction risk. We conclude that the exclusion restriction is valid in this context. For a more detailed discussion of how other IV assumptions are met under our identification strategy, we again refer to MW.

IV. Data Sources and Empirical Specification

We focus on mortality as our measure of health outcomes. As outlined in the previous

¹² In fact, Tatum and Tuchinsky describe the exams this way, “Since the [Army] examines large numbers of men each day, and since the doctors orderlies, and clerks assigned there must process them in assembly-line fashion as quickly as possible, the examinations are often careless. If you have a medical or other condition which should disqualify you, bring letters and other evidence from your own doctors...” (Ch. 6). A 1964 report to the president noted that “The current published medical standards are roughly the same as those which were in effect at the close of World War II.” (U.S. Department of Health, Education, and Welfare, 1964, Ch. 3)

section, we perform all analysis on data aggregated to the birth state-birth year cohort level because this is the level of our exogenous variation. We will refer to these cells as birth state-cohorts. Our main measure of mortality at the birth state-cohort level is the cumulative mortality rate per 1000 persons between 1980 and 2007, constructed as follows:

$$(7) \quad MR_{sc} = \frac{\text{deaths}_{sc} \text{ between 1981 and 2007}}{\text{cohort size}_{sc}^{1980}} * 1,000$$

where s is state of birth and c is year of birth. Thus the mortality rate gives the probability of dying by 2007, conditional on being alive in 1980. We also construct cumulative mortality rates by cause. These are defined as above, but MR_{sc}^j is mortality due to cause j , and the numerator is all deaths due to cause j over the period.

The data for our analysis come from two sources. First, we use data from the IPUMS microdata 5% samples of the 1980 Census (Ruggles et al., 2004) to construct denominators in the mortality rate measures as well as the birth state-cohort rates of college completion and veteran status. We restrict our sample to men born between 1942 and 1953, following MW. These are the years for which both inductions and enrollments are available at the state level, which are the two components of our induction risk measures. We further restrict the sample we use in this paper to white men only. This is to alleviate concerns about missing African-American men due to disproportionate increases in incarceration rates for this group over our period of study.¹³

Our second source is the Vital Statistics mortality data for the period 1981 to 2007, which contains observations on all deaths in the United States, at the annual level. We use this data to construct the numerators in our mortality rate measures, after applying the same sample restrictions applied to the Census data. We match numbers of deaths to the appropriate state-cohort information using year and state of birth. We also have information on the primary (or underlying) cause of death, and we use this to construct cause-specific mortality rates.

¹³ This is relevant for survey data sets, in which incarcerated black men are not observed.

Our primary measure of educational attainment is a dummy variable for college completion. To inform this choice, we explored the impact of our instruments on educational attainment. Figure 3 plots the coefficients and standard errors from separate regressions of educational attainment at each grade level and higher (inclusive), and confirms that higher national and state-level induction risk increased male educational attainment at all post-secondary levels, including completion. Nevertheless, we have experimented with alternative specifications using years of schooling and years of college as our measure of educational attainment. The substance of our conclusions appears robust to the choice of education measure.¹⁴

Our measure of veteran status is based on veteran information in the Census. Specifically, we define a veteran in our cohorts as someone who answered affirmatively that he was a Vietnam veteran.¹⁵ We exclude anyone from our sample that is on active duty in the military. We also exclude observations with imputed values for a number of key variables.¹⁶

We estimate all of our regressions at the birth state by birth cohort level because our source of exogenous variation occurs at this level. Our main econometric model as follows:

$$(8) MR_{sc} = \lambda'X_{sc} + a_1vet_{sc} + a_2educ_{sc} + \tau trend_{sc} + \delta_s + \epsilon_{sc}$$

where s indexes state of birth and c indexes birth year, and the dependent variable MR_{sc} is the mortality rate per 1,000 persons as defined in (7)—either in total or by cause. The variable vet is the fraction of veterans, and $educ$ is the fraction with a college degree. In addition, $trend$ is a region-specific linear trend in birth cohort, X is a set of state-cohort level controls, and δ represents a full

¹⁴ College completion has been shown to be the most accurately measured higher education outcome and to contain little measurement error overall (Black et al. 2003). On the other hand, years of post-secondary schooling is unlikely to be plagued by error that is negatively correlated with the recorded value, as is the case with the dummy variable schooling measures (Kane, Rouse, and Staiger, 1999 and Black et al., 2000).

¹⁵ Census respondents are asked whether they are veterans, and if they answer yes, they are then asked to identify a specific period of conflict. A small number of men in our cohorts report that they are veterans but not Vietnam veterans.

¹⁶ Specifically, we drop observations with imputed values for age, education, birth place, and veteran status.

set of state of birth dummies. Allowing for state of birth fixed effects removes variation arising from states with persistently higher or lower than average induction rates, which may be associated with other state characteristics (e.g. industrial composition) that are correlated with mortality rates. X includes the shares of blacks and other non-whites in a state-cohort. X also includes two variables to capture labor market conditions facing a cohort at the time of the college enrollment decision: (a) the employment to population (epop) ratio in the individual's state of birth the year his cohort turned 19, and (b) the log of the number of respondents from a birth state and year cohort in the 1960 Census. Together, these approximate the changes in labor demand and labor supply which may have occurred alongside changes in state-level induction risk.¹⁷

Our endogenous variables, *educ* and *vet*, are predicted from first stage equations that include the remaining right hand side covariates in Equation (8) plus functions of *staterisk_{sc}* and *nationalrisk_s*, as defined in (5) and (6), respectively. Our main results show estimates from three different first stage specifications. In the first, *staterisk* and *nationalrisk* are entered linearly. This is the same specification used in MW. The second and third specifications include quadratic and cubic functions of both risk measures, respectively. Our inspection of the relationship between our identifying state risk variation and our covariates of interest (veteran status and college completion) suggests that the relationship may be non-linear, in which case including higher order risk terms is appropriate. This non-linearity can be seen in Figure 2. We plot regression adjusted veteran and college graduate against regression adjusted state risk measures; the regression adjustment helps to isolate the identifying variation in state risk by removing variation driven by included control variables. Figure 2 shows that the relationship between veteran status and state risk, in particular, is non-linear. Estimation is done via standard linear 2SLS, weighted by the number of observations in each state-

¹⁷ The literature tends to find no consistent, significant relationship between local labor market conditions and college attendance (Wozniak, 2010; Card and Lemieux 2001b).

cohort cell.¹⁸ Standard errors are clustered at the birth state-cohort cell level.¹⁹

Descriptive statistics are shown in Table 1, which summarizes the variables used in our analysis for the sample of white men born between 1942 and 1953. The average cumulative mortality rate per 1,000 persons as defined in (7) is 138.58. Table 1 also gives mortality rates by cause-of-death for ten important causes for these cohorts, in order of prevalence. The most common causes of death were cancers (36.5 per 1,000 white men) and heart disease (32.3/1,000). But as Figure 1 also shows, the causes of death change dramatically over the period. In the 1980s, external causes (accidental injury, suicide, and homicide) accounted for over a third of deaths among these cohorts, while in the later periods cancer and heart disease were more common causes (CDC 2005).

V. Results

A. First stage results

Table 2 presents F-statistics from estimation of the three first stage specifications: linear, quadratic and cubic functions of the risk measures. For transparency, we estimate two separate first stage equations for each specification—predicting college graduation and veteran status separately—although 2SLS estimates these equations jointly. Consistent with the manner in which 2SLS identifies endogenous variables, both equations include national and state cohort risk as identifying variables. The F-statistics are from tests of the null hypothesis that the induction risk variables are jointly insignificant, and are calculated according to the (corrected) Angrist and Pischke (2009) procedure.

¹⁸ See Wooldridge (2002) pp. 622-624 concerning 2SLS versus an approach with a probit first stage when the endogenous variable is a dummy variable. In some cases, the latter is more efficient but may tend to produce larger point estimates. Given our concerns about possible upward bias, we implement 2SLS estimation.

¹⁹ Because our mortality measures combine Census data with data on the universe of mortality outcomes, we do not further correct our standard errors to account for multiple data sources, as for example, in TSIV.

The first two columns show results from the linear specification. A 10 percentage point increase in national cohort risk (roughly the entire range of this variable) increased the percent with a college degree by 9.4 percentage points. In other words, cohorts with higher national risk were more likely to graduate college, consistent with Figure 1.²⁰ Coefficients from the quadratic and cubic specifications also indicate an overall positive relationship between higher national risk and college graduation. On the other hand, college graduation is negatively related to state cohort risk when national induction risk is included. As discussed in MW, this may be an artifact of the high correlation (collinearity) between national and state level risk. Across specifications, the F-statistics suggest that this first stage has substantial power.

Columns (4) through (6) address the second endogenous variable in Equation (8). These show that veteran status positively varies with both national and state cohort risk, and the F-statistics are large. This is reassuring since higher rates of induction risk at both the state and national level should lead more young men to go to war. We interpret the fact that the coefficient on national risk exceeds that of state risk in the veteran equation to mean that the time series variation in draft risk generated by the massive fluctuation in military manpower demands is responsible for more of the variation in veteran status than are the differences in induction risk across states.

B. OLS and 2SLS effects of college education on total mortality

Table 3 presents our OLS and 2SLS estimates of the effect of college graduation on cumulative mortality (Equation 8). For the IV results, we show specifications in which the induction risk variables are linear, quadratic, and cubic. For the quadratic and cubic specifications the model is over-identified, so we can estimate the model using both 2SLS and limited-information maximum likelihood (LIML). We do this to further investigate the validity of our instruments—LIML is less

²⁰ The comparable estimate from Table 1 in Card and Lemieux (2001a) shows a 4.6 percentage point increase. Our estimate is slightly larger because we restrict the sample to whites.

precise than 2SLS but is also less biased in the presence of confounding variables. If the coefficients from the two approaches are similar, this is evidence that 2SLS is unbiased (Angrist and Pischke 2009).

The OLS coefficient for college graduation is -101.97. As the average cumulative mortality rate is 138.58, this is a large decrease in mortality relative to the mean. Veteran status is also negatively associated with mortality for these cohorts. Column [2] shows the point estimates of the causal effect of college graduation and veteran status on mortality using a linear specification for the instruments, and the results are remarkably similar to the OLS estimates, though they are less precise. With the quadratic specification, the estimated effect of college graduation is substantially larger and is statistically significant; with the cubic specification the point estimates are again very close to OLS and are statistically significant. In columns [3] through [6], the 2SLS and LIML results are nearly identical, further supporting our identification strategy. Taken together, the 2SLS results imply a large causal role for both college graduation and veteran status in reducing cumulative mortality for these cohorts. For the linear and cubic specifications, IV results are very similar to OLS.

One might wonder whether the effect of education on mortality changes over time, since the causes of death vary considerably over the period (as Figure 1 shows). In Table 4, we show OLS and IV estimates on mortality by decade. In these and all remaining results, we show results for the cubic 2SLS specification only.²¹ For each decade, this measure is the total number of deaths over the decade (1981-1990, 1991-2000, 2001-2007) divided by the cohort size in the beginning of the decade (taken from the 1980, 1990, and 2000 Censuses, respectively). These shorter-term mortality

²¹ We choose to show the cubic specification because the estimates are generally more precise. Also, DeWalque (2007) uses a non-linear specification of induction risk to instrument for both education and veteran status. The 2SLS and LIML results for decadal mortality are similar using the cubic functional form (though more so for 1980 and 2000 than for 1990), which again suggests that 2SLS using this specification is unbiased. Results are qualitatively similar across specifications.

rates will allow us to determine whether the role of education changes as likely reasons for mortality evolve.

For each decade, college graduation and veteran status are negatively associated with mortality.²² For the 1980s (where the first stage is strongest), IV results are again very similar to OLS. The estimated effect of a college degree on mortality during the 1980s is -24.38, or a 96% decrease relative to the mean. For the 1990s and 2000s, the IV estimates are larger than the OLS estimates, and represent a 108% and 72% decrease in mortality, respectively. Results for these decades are also less precise, however, which may be driven by the weaker first stage for college graduation (first-stage F-statistics are reported in the table). While the results are imprecise, there is little evidence that the impact of college education on total mortality (as a percentage of the mortality rate) varied widely across decades. Veteran status has a negative and statistically significant negative effect on mortality in all periods. In the next section, we consider the effect of college graduation on specific causes of death.

C. OLS and 2SLS effects of college education on mortality by cause of death

Table 5 shows the effect of college graduation on leading causes of death for men in these cohorts. Causes are shown in descending order of prevalence from 1980-2007. OLS results are in Panel A and 2SLS results are in Panel B. Again, all specifications include birth region trends and a control for veteran status. For cancer and heart disease—the two leading causes of death—the point estimates on college graduate are negative and statistically significant for both OLS and IV, and are larger for IV. A college degree is estimated to reduce mortality due to cancer by -61.90 and to heart disease by -39.88. The effect of veteran status on mortality due to these causes is also negative, though not statistically significant.

The next two causes are external: accidental injury and suicide. The OLS results show a

²² For 1990, we define Reservists or National Guard members as non-veterans to match the veteran definition in the 1980 and 2000 Censuses.

negative association between a college degree and death by accidental injury, but the IV coefficient is smaller and statistically insignificant. There is also no evidence that college decreases deaths by suicide. Veteran status, however, does decrease deaths due to these two external causes.

Infectious and parasitic diseases and liver diseases are next, and for both causes the point estimate for college graduate is negative in the OLS specification. But in the 2SLS specifications, the effect is positive and statistically significant. We think there are at least two possible explanations for finding a positive effect of college education on these two causes. First, since the first two columns in the table indicate that college prevented some men from dying of cancer and heart disease, these men may now be more susceptible to death by other causes due to competing risks (Honore and Lleras-Muney, 2006). Second, the causes of these deaths specifically may be linked to college education. Many fatal infections are contracted in hospitals; if more educated men are more likely to be treated in hospitals, they may be more at risk for infection.²³ Liver diseases, on the other hand, include cirrhosis which can be caused by excessive alcohol consumption—48% of deaths to cirrhosis in 2007 were alcohol-related (Yoon and Yi 2010). If a college degree increases this behavior, graduates may be more at risk for liver diseases. The effect of veteran status on deaths by infectious diseases and liver disease is negative.

For the remaining causes of death—diabetes, stroke, chronic low respiratory conditions, and homicide—the point estimates are negative in both the OLS and IV specifications. A college degree decreases deaths per 1,000 white men from diabetes and stroke by 7 and from chronic low respiratory conditions by 6. There is no effect of a college education on deaths by homicide. Veteran status also decreases deaths due to chronic low respiratory conditions; the effect on the diabetes, stroke, and homicide is small and statistically insignificant.

The results in Tables 3, 4, and 5 indicate that college education had a negative effect on total

²³ Infections and parasitic diseases account for about 5% of the deaths we observe. Klevens et al. (2007) estimate that in 2002, 98,987 deaths were caused by or associated with infections contracted in hospitals.

mortality, that the effect was not confined to any particular decade, and that deaths due to heart disease and cancer are particularly affected. In the next section, we explore mechanisms.

Outline of final sections

VI. Potential Mechanisms

- Income/hourly wage – Census (Table 6)
- Access to insurance – ACS and NHIS (Table 7)
- NHIS data on health behaviors/diagnoses: Hospital stay, surgery – minor or major, alcohol consumption, diagnosed liver disease, chronic low respiratory, diagnosed diabetes, stroke risk factors, heart disease risk factors, cancer risk factors. These will answer the question of whether there is a college vs non-college differential in the direction of the mortality results while men in our cohorts are still alive.

VII. Conclusions

- Comparison of our results on education's impact on health with the literature
- Conclusion of large, positive causal impact of college on health
- Conclusions on mechanisms
- Comparison of effects of veteran status on health with those in the literature
- Outstanding questions

References

- Albouy, Valerie, & Lequien, Laurent. 2009. Does Compulsory Education Lower Mortality? *Journal of Health Economics*, 28(1), 155-168.
- Angrist, Joshua, Chen, Stacy, and Brigham Frandsen. 2010. "Did Vietnam Veterans Get Sicker in the 1990s? The Complicated Effects of Military Service on Self-Reported Health," *Journal of Public Economics* 94, 11-12, 824-837
- Angrist, Joshua and Kreuger, Alan. "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." *NBER Working Paper #4067* (1992).
- Angrist, Joshua and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.
- Arendt, Jacob Nielsen. 2005. Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification. *Economics of Education Review*, 24(2), 149-160.
- Bearman, Peter; Kathryn M. Neckerman; and Leslie Wright. 2011. *After Tobacco: What Would Happen if Americans Stopped Smoking?* New York: Columbia University Press.
- Bedard, Kelly; Deschênes, Olivier. 2006. "The Long-Term Impact of Military Service on Health: Evidence from World War II and Korean War Veterans". *American Economic Review*, Volume 96, Number 1, pp. 176-194(19)
- Bound, John and Sarah Turner. 2002. "Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?" *Journal of Labor Economics*. 20(4): 784-815.
- Cameron, C., Gelbach, J., and D. Miller (2006) "Robust inference with Multi-Way Clustering" *NBER Technical Working Paper #T0327*.
- Card, David, and Lemieux, Thomas. "Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War." *Manuscript*. Vancouver, Canada: University of British Columbia (2000).
- Card, David, and Lemieux, Thomas. "Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War." *American Economic Review*. 91(2001): 97-102.
- Clark, Damon and Heather Royer. 2010. "The Effect of Education on Adult Health and Mortality: Evidence from Britain" NBER Working Paper No. 16013
- Centers for Disease Control and Prevention. Office of Statistics and Programming, National Center for Injury Prevention and Control. 2005. "10 Leading Causes of Death by Age Group." http://www.cdc.gov/injury/images/LC-Charts/10LC_overall_2005b-a.pdf
- Cutler, David M. and Adriana Lleras-Muney. 2010. "Understanding Differences in Health Behaviors by Education." *Journal of Health Economics*, 29(1): 1-28.
- Davis, Jr., James W., and Dolbeare, Kenneth M. *Little Groups of Neighbors: The Selective Service System*. Chicago: Markham Publishing Company (1968).

- De Walque, D. 2007. "Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education," *Journal of Health Economics* Vol. 26(5), pages 877-895.
- Dobkin, Carlos and Shabani, Reza. 2007. "The Health Effects of Military Service: Evidence from the Vietnam Draft" *Economic Inquiry*, Volume 47, Issue 1, pages 69–80
- Evers, Alf. *Selective Service: A Guide to the Draft*. New York: Lippincott Company (1961).
- Grimard, F. and D. Parent. 2007. "Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking?" *Journal of Health Economics*, Vol. 26. No. 5: pp. 896-926
- Grossman, Michael, 2006. "Education and Nonmarket Outcomes," *Handbook of the Economics of Education*, Elsevier.
- Honoré, Bo and Adriana Lleras-Muney. 2006. "Bounds in Competing Risks Models and the War on Cancer." *Econometrica*, 74(6): 1675-1698.
- Klevens, R. Monina, et al. 2007. "Estimating Health Care-Associated Infections and Deaths in U.S. Hospitals, 2002." *Public Health Reports*, 122(2): 160-166.
- Lleras-Muney, Adriana. 2005. The Relationship between Education and Adult Mortality in the United States. *Review of Economic Studies*, 72, 189-221.
- MacInnis, Bo. 2006. "The Long-Term Effects of College Education on Morbidities: New Evidence from the Pre-Lottery Vietnam Draft," *mimeo*, University of Michigan.
- Malamud, Ofer and Wozniak, Abigail. 2011. "The Impact of College on Migration: Evidence from the Vietnam Generation." *Manuscript*.
- Robins, James and Sander Greenland. 1992. "Identifiability and Exchangeability for Direct and Indirect Effects." *Epidemiology*. 3(2): 143-155.
- Rothenberg, Leslie S. *The Draft and You: A Handbook on Selective Service*. New York: Doubleday & Company, Inc. Anchor Books (1968).
- Ruggles, Steven; and Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004. URL: www.ipums.org.
- Sanders, Jacquin. *The Draft and the Vietnam War*. New York: Walker and Company (1966).
- Shapiro, Andrew O. and Striker, John M. *Mastering the Draft: A Comprehensive Guide to Solving Draft Problems*. Boston: Little, Brown and Company (1970).
- Stanley, Marcus. 2003. "College Education and the Midcentury GI Bills." *Quarterly Journal of Economics*. 118(2): 671-708.
- Tatum, Arlo and Tuchinsky, Joseph S. *Guide to the Draft*. 2nd Edition. Boston: Beacon Press (1969).

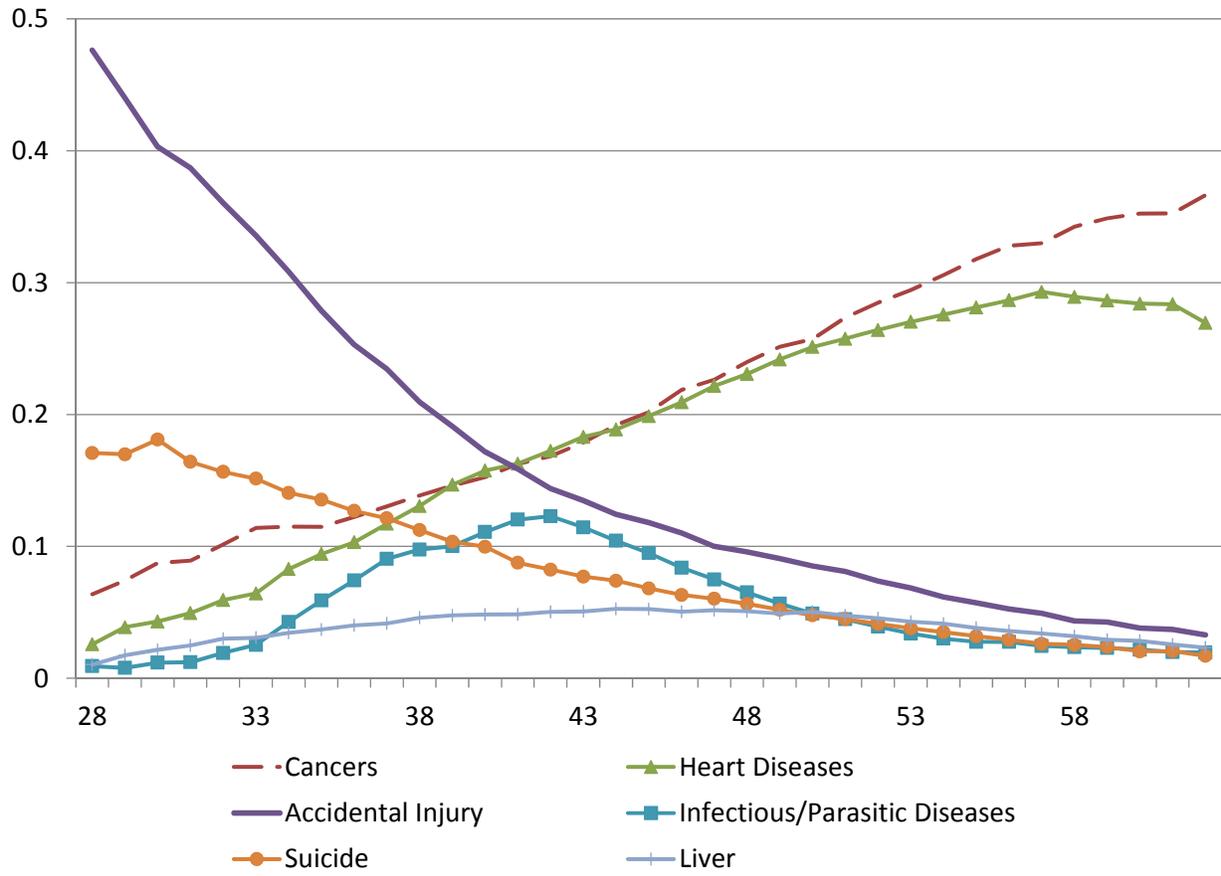
Turner, Sarah. 2004. "Going to College and Finishing College: Explaining Different Educational Outcomes." In *College Decisions: How Students Actually Make Them and How They Could*, ed. Caroline Hoxby. University of Chicago Press for NBER.

Semiannual Reports of the Director of the Selective Service System, 1967-1973.

Wozniak, Abigail. (2010) "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*.

Yoon, Young-Hee, and Hsiao-ye Yi. 2010. "Surveillance Report #88: Liver Cirrhosis Mortality in the United States, 1970–2007." Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research.

Figure 1: Fraction of Deaths Attributable to Common Causes, by Age



Source: Mortality Detail Files, 1981-2007.

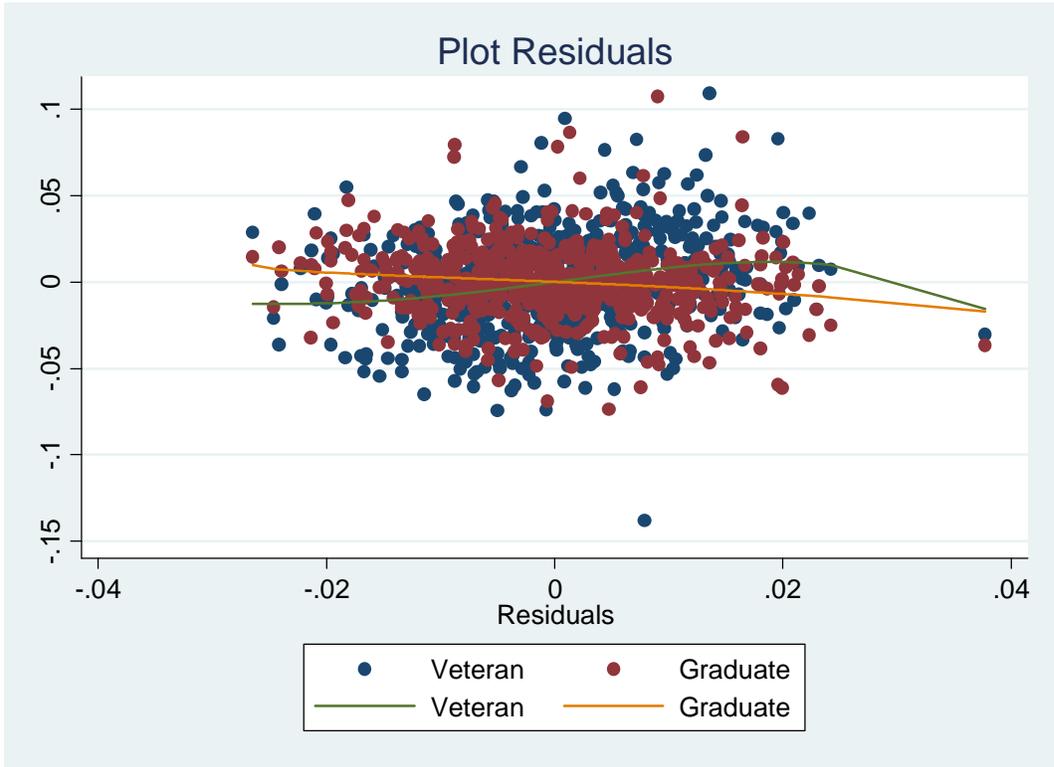


Figure 2. Plot of regression adjusted veteran status, graduate status, and state risk (i.e. residuals). State risk residuals plotted on Y-axis, veteran and graduate status residuals plotted on X-axis. Regression adjusts for linear trend in cohort birthyear, birth state fixed effects, and birth state-cohort epop ratio at labor market entry.

Table 1: Summary Statistics

VARIABLES	Mean	SD
Census Data		
Years of Higher Education	1.99	0.38
College Graduate	0.3508	0.0666
Veteran	0.3092	0.1231
Individual Observations	14,392,122	
Mortality Data		
<u>1981-2007 Mortality Rates by Cause of Death:</u>		
Total	138.58	38.03
Cancers	36.52	15.18
Heart Disease	32.33	12.57
Accidental Injury	13.67	3.49
Suicide	7.31	1.37
Infectious and Parasitic Diseases	7.20	2.41
Liver Disease	5.75	1.93
Diabetes	3.82	1.44
Cerebrovascular Disease (Stroke)	3.61	1.47
Chronic Low Respiratory Disease	2.83	1.83
Homicide	1.98	0.97
Total Deaths	1,994,481	
State/Birth Year Cells	600	

Notes: Census data are from the 5% sample of the 1980, 1990, and 2000 U. S. Census, available from IPUMS. Mortality data are from the Vital Statistics Multiple Cause of Death files from 1980 to 2007. The sample is restricted to white men born between 1942 and 1953. Veterans include any respondent that served in active duty in the Vietnam War. Respondents currently in active duty are excluded. Means are weighted by cell size. Mortality rates are deaths over the period per 1,000 population, where population is the cohort size in the first year of the period.

Table 2: First Stage Estimates of Effect of Induction Risk on Education

	Dependent Variable:					
	College Graduate	Veteran	College Graduate	Veteran	College Graduate	Veteran
National Induction Risk	0.935** (0.091)	2.349** (0.158)	1.899** (0.255)	2.865** (0.462)	2.368** (0.558)	1.237 (0.992)
National Risk ^2			-7.484** (1.529)	-3.897 (2.673)	-22.748** (7.795)	25.194* (13.281)
National Risk ^3					101.276** (37.535)	-154.080** (63.983)
State Induction Risk	-0.240** (0.089)	0.768** (0.156)	-0.380* (0.229)	0.653 (0.419)	0.588 (0.490)	1.171 (0.971)
State Risk ^ 2			0.770 (1.146)	0.627 (2.069)	-12.254 (6.095)	-5.907 (11.608)
State Risk ^ 3					54.507** (25.124)	25.502 (46.891)
F-Stat	21.30	133.10	18.63	120.60	18.94	137.60

Notes: F-statistic is for the test of the null hypothesis that all national and state induction risk variables are jointly insignificant. Specification refers to the functional form of the induction risk variables in the first stage regression. F-statistics are calculated using the Angrist-Pischke method for the case with multiple endogenous variables. See Table 1 notes for data sources, sample restrictions and variable definitions. Number of observations is 600 in each regression and regressions are weighted by cell size. Additional controls include birth state fixed effects, the employment to population ratio, the cohort size (derived from the 1960 Census and defined at the birth-year level), and birth-region trends.

Table 3: OLS and IV Estimates for the Impact of Education on Cumulative Mortality, 1980-2007

	OLS [1]	IV Specification:				
		Linear-2SLS [2]	Quadratic-2SLS [3]	Quadratic-LIML [4]	Cubic-2SLS [5]	Cubic-LIML [6]
College Graduate	-101.97*** (21.16)	-99.26 (120.51)	-170.76*** (65.76)	-172.46*** (63.89)	-93.93* (50.76)	-93.53* (50.37)
Veteran Status	-46.36*** (7.53)	-42.93 (27.09)	-27.21* (15.12)	-26.82* (14.66)	-44.45*** (13.04)	-44.50*** (12.83)
Observations	600	600	600	600	600	600
R-squared	0.9516	0.9515	0.9505	0.9504	0.9515	0.9515

Notes: Standard errors are in parenthesis and are clustered by birth year-state. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively. See Table 1 notes for data sources, sample restrictions and variable definitions. The IV specifications use the national and state-level induction risk to instrument for college graduate and veteran status. The column headings indicate the functional form of the instruments in the first stage equation (linear, quadratic, or cubic) and the choice of model (two-staged least squares or limited-information maximum likelihood). Number of observations is 600 in each regression and regressions are weighted by cell size. Additional controls include birth state fixed effects, the employment to population ratio, the cohort size (derived from the 1960 Census), and birth region trends.

Table 4: OLS and IV Estimates for the Impact of Education on Mortality by Decade

VARIABLES	1981-1990		1991-2000		2001-2007	
	OLS	IV	OLS	IV	OLS	IV
College Graduate	-24.51*** (4.68)	-24.38** (10.64)	-20.00** (8.25)	-79.46** (38.76)	-38.58*** (10.88)	-100.37*** (36.11)
Veteran Status	-9.93** (3.53)	-9.33*** (2.69)	-35.54*** (6.84)	-27.32*** (5.87)	-36.99*** (4.87)	-28.80*** (5.59)
First Stage F-Stat for:						
College Grad	18.94		6.04		5.93	
Veteran	137.6		109.51		133.04	
Mean Mortality Rate	25.46		54.88		73.33	
Total Deaths	366,365		727,108		901,008	

Notes: Standard errors are in parenthesis and are clustered by birth year-state. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively. See Table 1 notes for data sources, sample restrictions and variable definitions. The IV specifications use a third-order polynomial in the national and state-level induction risks to instrument for college graduate and veteran status using a 2SLS model. Number of observations is 600 in each regression and regressions are weighted by cell size. Additional controls include birth state fixed effects, the employment to population ratio, the cohort size (derived from the 1960 Census), and birth region trends.

**Table 5: OLS and IV Estimates for the Impact of Education on Mortality,
by Leading Causes of Death**

	Cancers	Heart Disease	Accidental Injury	Suicide	Infectious/ Parasitic	Liver Disease	Diabetes	Stroke	Chronic Low Resp.	Homicide
Panel A: OLS										
College Graduate	-38.34*** (11.20)	-29.51*** (7.37)	-5.21* (2.54)	-0.68 (1.67)	-2.62* (1.25)	-1.19 (2.66)	-2.41 (1.36)	-2.98** (1.10)	-4.04*** (0.99)	-1.62* (0.90)
Veteran Status	-15.88* (7.93)	-8.82 (6.96)	-3.20 (1.91)	-1.71 (1.12)	-0.98 (0.96)	-1.25 (1.26)	-0.73 (0.83)	-1.79** (0.78)	-3.32*** (0.61)	0.04 (0.36)
Panel B: IV										
College Graduate	-61.90*** (15.27)	-39.88*** (14.93)	-0.71 (6.50)	4.24 (3.43)	14.75** (6.64)	12.20*** (3.96)	-7.04** (2.89)	-6.61*** (2.28)	-6.08** (2.41)	-0.10 (1.64)
Veteran Status	-9.72** (3.83)	-5.83 (3.79)	-3.81** (1.59)	-2.74*** (0.88)	-4.56*** (1.56)	-4.04*** (0.96)	0.35 (0.69)	-0.88 (0.58)	-2.78*** (0.55)	-0.22 (0.40)

Notes: Standard errors are in parenthesis and are clustered by birth year-state. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively. See Table 1 notes for data sources, sample restrictions and variable definitions. The IV specifications use a third-order polynomial in the national and state-level induction risks to instrument for college graduate and veteran status using a 2SLS model. Number of observations is 600 in each regression and regressions are weighted by cell size. Additional controls include birth state fixed effects, the employment to population ratio, the cohort size (derived from the 1960 Census), and birth region trends.

Table 6: The Causal Impact of College on Earnings and Wages

dependent variable: log earnings/wages

	Log earnings		Log wages	
	<i>OLS</i>	<i>2SLS</i>	<i>OLS</i>	<i>2SLS</i>
College Graduate	0.427*** [0.086]	0.780*** [0.140]	0.152*** [0.054]	0.393*** [0.085]
Veteran Status	0.170*** [0.035]	0.07 [0.044]	0.140*** [0.031]	0.075** [0.034]

Notes: Data are birth state-cohort cell averages from white men in the 1980 Census IPUMS extracts, born between 1942 and 1953. First stage specifications use cubic functions of induction risk and are identical to those in the last two columns of Table 2. Dependent variable is log real earnings or log real wages. Underlying microdata sample is therefore restricted to those with valid earnings in 1980. We truncate the top and bottom 1 percent of the hourly wage distribution to deal with measurement error in hours and weeks worked. Specifications are otherwise identical to those in Table 3. Number of observations is 600 in each regression and regressions are weighted by cell size. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.

Table 7: College Graduation and Health Insurance Access

	OLS		IV	
	College Graduate	Veteran	College Graduate	Veteran
<i>ACS Sample</i>				
Any Insurance	0.061*** (0.001)	0.048*** (0.001)	0.156* (0.087)	0.079*** (0.020)
Private Coverage	0.149*** (0.002)	0.012*** (0.002)	0.187 (0.135)	0.000 (0.028)
Public Coverage	-0.129*** (0.002)	0.136*** (0.002)	0.005 (0.129)	0.272*** (0.032)
<i>NHIS Sample</i>				
Any Insurance	0.091*** (0.007)	0.061*** (0.008)	-	-
Private Coverage	0.196*** (0.010)	-0.042*** (0.011)	-	-
Public Coverage	-0.105*** (0.008)	0.103*** (0.010)	-	-

Notes: ACS and NHIS data are from survey years 2008-2010. Both samples are restricted to white, male respondents born between 1942 and 1953, and aged 64 or younger at the time of the survey. Each “College graduate” and “Veteran” pair are coefficients from a single regression. ACS specifications are estimated using microdata but are otherwise identical to those in Table 4, with the addition of survey year controls. NHIS specifications include survey year controls and age controls only. First stage F statistics are 6.3 and 88.4 for college graduate and veteran status, respectively, in the IV specifications using ACS data. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level respectively.