--- 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: March 2011

ABSTRACT

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 based on Vital Statistics data from 1981 to 2004. We generally find 2SLS estimates that are close in magnitude and significance to the OLS estimates. Our preferred 2SLS estimates imply that college completion reduces mortality in our affected cohorts. Effects are largest for deaths from heart disease and substance abuse in the 1980s. However, we find *positive* effects of college completion on mortality from stroke in the 1990s.

(JEL: I12, I23, J24)

Acknowledgments: The authors would like to thank Bill Evans and seminar participants at 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. Moreover, they would imply that policies to increase educational attainment could also 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, 2004, and Cutler and Lleras-Muney, 2006). 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 (2011) 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 represents a significant advance over studies that use the Card and Lemieux measures to identify college going using only year to year variation in induction risk (De Walque 2007, Grimard and Parent 2007, and MacInnis 2006).

Finally, the decomposition of induction risk into its national and state components is a novel use of two instruments to address the endogeneity of both education and military status that is common in studies using changes in military policies as instrumental variables (Angrist and Krueger, 1992; Bound and Turner 2002; Stanley 2003).

We merge our data on national and state-level induction risk with the Vital Statistics Mortality Files from 1981 to 2004 and the U.S. Censuses for 1980, 1990, and 2000 to construct birth state-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 (2011), our instruments strongly predict both veteran status and years of completed education 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 present results using college completion as our measure of educational attainment, but our results are robust to alternative educational definitions.

We first establish that the well-known education-health status gradient is present and statistically significant in our data using OLS specifications. We focus on 10-year mortality rates for the 1980s and 1990s and find that college graduation is associated with lower rates of mortality in the Vital Statistics data. Our instrumental variables estimates, however, present a more complicated picture. After correcting for the endogeneity of veteran status and educational attainment, we find preliminary evidence that higher education has different relationships with mortality depending on the time-horizon of the analysis. Using 10-year mortality rates for the decade of the 1980s, we find 2SLS effects that are generally similar in magnitude to our OLS estimates although somewhat more imprecise (but still significant in many specifications). On the other hand, using 10-year mortality rates over the 1990s we find 2SLS effects that are rarely significant and sometimes with opposite

sign. The causes of death driving the results also differ across decades. Understanding the source of these differences between 1980 and 1990 is an important area for further work on this project.

In ongoing work on this project, we plan to access to restricted data from several years of the National Health Interview Survey, which would allow us to match our induction risk measures with individuals in those surveys. The expanded set of variables available in the NHIS will also allow us to examine the relationship between educational attainment and a wider range of health outcomes than is currently available to us in the Vital Statistics data. Specifically, the NHIS will provide us with detailed information on acute health outcomes, overall health status, health-related behaviors, and insurance utilization. The NHIS will also allow us to observe effects on health outcomes at later ages for our cohorts of interest, which may shed light on why our estimates differ across the earlier decades of the current analysis.

This paper 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.[1] However, whether or not there is a causal impact on health at the margin of dropping out of high school, the causal relationship between education and health may be very 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

[1] 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 tend to be quite imprecise.

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.[2] Our paper extends this identification strategy by incorporating within-cohort variation in induction risk. More importantly, our paper examines a wide range of health outcomes across multiple decades for our cohorts of interest, providing a much broader picture of higher education's potential health impacts.

The findings in this paper are also closely related to research examining the impact of veteran status on health outcomes and behaviors. Angrist, Chen, and Frandson (2009) use variation in veteran status induced by the Vietnam draft lottery to show that military service had no significant effect on work-limiting disability as reported in the 2000 Census, although it slightly increased non-work-limiting disability rates for whites. Dobkin and Shabani (2007) also use the Vietnam draft lottery to examine the effect of veteran status on self-reported health and other health outcomes in the NHIS data, but find 2SLS results that are too imprecise to generate any strong conclusions (although their OLS results indicate that veterans are in systematically worse health than non-veterans).[3] Finally, Bedard and Deschenes (2006) document the impact of military service during WWII and the Korean War on mortality. They show that the large impacts of military service on

[2] 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.

[3] Conley and Heerwig (2009) use Vietnam draft lottery to examine the effect of veteran status on mortality using Vital Statistics Data from 1989-2002 and find no significant effects. This contrasts with early work by Hearst, Newman, and Hulley (1986) showing a short-term increase in mortality associated with draft-lottery induced service in Vietnam.

mortality are due increases in heart disease and lung cancer which are themselves driven by higher rates of smoking among military servicemen.

II. The Vietnam Draft

Approximately 2 million men were drafted through the Selective Service System 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. All men over the age of 18 were required to report to their local draft board for classification. Those who were classified as 'available for service' and passed the pre-induction examinations were liable for induction.[4] Local draft boards were also able to issue deferments for various reasons, such as school attendance or the presence of dependent children. Prior to the introduction of the draft lottery in 1969, the process of determining which men were drafted was set according to the following priorities: highest priority for "delinquents"— those who failed to register or failed to report for the pre-induction physical; second priority for volunteers; and third priority for non-volunteers aged between 19 and 25. Consequently, not everyone who was inducted actually served in Vietnam.

College enrollment was a well known and virtually foolproof way to defer and avoid conscription. Before 1967, deferment for the purposes of four-year college enrollment was essentially guaranteed in practice through the individual decisions of local draft boards.[5] The Military Service Act of 1967 made this official by stating that college students in good standing could defer induction until receipt of an undergraduate degree or age 24, whichever occurred first. Indeed, over

[4] Men who were drafted were generally assigned to serve in the Army for up to 3 years, although they could potentially choose their branch of service by volunteering and thereby also qualify for a shorter period service.
[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. Students at two year programs were only eligible for occupational deferments under the same rules as those already employed. See Rothenberg (1968).

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 ages of 26 and 35 were ever drafted. Moreover, the incentive to enroll in college to avoid the draft during these years was very large. 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.

The introduction of the draft lottery in 1969 led to a substantial change in the induction process. The first lottery, held on December 1, 1969, determined priority for induction in 1970 according to random sequence numbers (RSNs) assigned to the day and month of men born between 1944 and 1950. Additional draft lotteries took place in 1970, 1971 and 1973 which were applicable to men born in 1951, 1952 and 1953 respectively. 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. However, risk of induction during this period was much lower since men were at risk of induction for only a single year and the overall rate of inductions was substantially lower (as documented below).[7]

A. Variation in Inductions over Time

The number of 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. During the period that the draft lottery was in place, the rate of

[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).
[7] Note that the lottery also altered the risk of induction within cohort: men who received high (low) RSNs faced relatively low (high) risk of induction. The net effect of these changes is theoretically ambiguous but likely to be small.

inductions fell rapidly 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.

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. The monthly state quotas sum to the national draft call each month.[8] Moreover, the sum of monthly state and national draft quotas over consecutive six month periods were reported in the Semi-Annual Reports of the Director of the Selective Service System to Congress. We use this information to construct a measure of induction risk based on annual induction rates in our empirical analysis.

B. Variation in Risk of Induction across States

While the national level rise and fall in the risk of induction over the course of the Vietnam War is well known in the U.S, institutional factors led to additional variation in induction risk at the state level. This within cohort, state level variation in induction risk arose for several reasons. First, the high degree of autonomy enjoyed by local draft boards generated variation in draft risk across local boards. Davis and Dolbeare identify three sources of variation in local board risk, "…[1] variation based on differences in socioeconomic characteristics of jurisdictions; [2] variation among states based on differences in policy interpretations provided to local boards and exaggerated by success in achieving standardization around such particular practices; and [3] variation produced by idiosyncratic discretionary decision-making by local boards."[9] They later write, "The conclusion seems inescapable: local board autonomy implies both within state and between state variability, even among socioeconomically similar board jurisdictions."[10]

[8] Information in this paragraph is based on Shapiro and Striker, *Mastering the Draft,* Chapter 20.
[9] Davis and Dolbeare, *Little Groups of Neighbors*, Page 18.
[10] *Ibid.* Page 84.

A second source of state-year variation in induction risk was severe communication delays between the federal, state, and local officials in charge of the draft. These delays meant that local draft boards knew the current number of registrants available in their jurisdiction while 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. In practice, boards were encouraged to have just enough eligible registrants to match the number of inductees they would be asked to produce in response to the next call. Under this pressure, the communication lags led to a cobweb-type feedback loop as boards struggled to achieve the successful "standardization" identified in item [2] of the Davis and Dolbeare list.[11]

III. Empirical Strategy

To isolate the causal effect of education on health, we use variation in college attainment caused by draft-avoidance behavior during the Vietnam conflict. An artifact of our identification strategy is that the likelihood an individual is a veteran also varies across cohorts in our sample. Insofar as veteran status is a plausible determinant of health, it is important to control for this variable in our estimation strategy. However, selection into veteran status may also lead to bias since the coefficient vector is not guaranteed to be unbiased unless all covariates are either identified or exogenous. Therefore, our main equation of interest contains both college-going and veteran status

[11] See Shapiro and Striker, *Mastering the Draft* Ch. 20.

as endogenous variables, and we exploit both national and state-level induction risk to generate exogenous variation in these variables.[12]

A. A Structural Model with Direct and Indirect Effects

We begin by presenting a rather general structural model that relates the direct and indirect effects of veteran status and college-going on health (ignoring other covariates for conciseness):

(1) $health_{sc} = \alpha_0 + \alpha_1 vet_{sc} + \alpha_2 educ_{sc} + \varepsilon_{sc}^{h}$

(2) $vet_{sc} = \beta_0 + \beta_1 educ_{sc} + \varepsilon_{sc}^{v}$

(3) $educ_{sc} = \gamma_0 + \gamma_1 vet_{sc} + \varepsilon_{sc}^{e}$

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, as in Malamud and Wozniak (2011; hereafter MW). In other words, we assume that education only has a direct effect on heath. 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 likely to be unaffected by draft-avoidance behavior. This is because local draft boards needed to fulfill specific manpower requirements. So if a certain individual avoided the draft, it was likely that someone else from his state-cohort would be drafted

[12] 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).

instead. Thus, to a first approximation, veteran status at the state-year cohort level was unaffected by college-going decisions.

On the other hand, $\gamma_1$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 health:

(4) $$health_{sc} = (\alpha_0 + \alpha_2\gamma_0) + (\alpha_1 + \alpha_2\gamma_1)vet_{sc} + \alpha_2\varepsilon^e_{sc} + \varepsilon^h_{sc}$$

Exogenous variation in veteran status, such as the Vietnam draft lottery used in Angrist and Chen (2010), would yield an unbiased estimate of the total effect, $(\alpha_1 + \alpha_2\gamma_1)$, of veteran status on health in (4). However, any correlation between ε^e_{sc} and ε^h_{sc} would potentially lead to bias in the estimates of $\alpha_1$and α_2. We explain how our IV strategy avoids this potential bias in the following two sections.

B. Identification

We use the same two-instrument identification strategy as MW, which was inspired by Card and Lemieux (2000, 2001). Like CL, we assume that draft avoidance was proportional to the risk of induction. Figure 1 plots the pattern of induction risk across cohorts, along with means of selected education variables separately for each birth cohort. The increase in college going is visible in the figure, particularly for cohorts born between 1944 and 1950, during the main rise and fall in induction risk.

However, in order 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 the following induction risk at the level of their state and cohort:

$$\text{(i)} \qquad state\ \ cohort\ \ risk_{sc} = \frac{\sum^{t} I_{sct}}{N_{sc}}$$

where I is the number of inductees from a birth year cohort, N is the number of men in that cohort, s indexes state of residence at 17 (which we approximate with state of birth for now), c indexes birth cohorts, and t indexes calendar years since a cohort was at risk to be drafted for several years between the ages of 19 and 22. 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:

$$\text{(ii)} \qquad national\ \ cohort\ \ risk_{sc} = \frac{\sum^{t} I_{-sct}}{N_{sc}}$$

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

To construct the measures in equations (i) and (ii), we obtained data on the number of inductees from each state for each year between 1961 and 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 1950 equals the number of inductees from Alabama in 1969 (the year he turned 19) divided by the number of students enrolled in 11th grade in Alabama in 1967.[13] National level risk for the same young man roughly equals the number

[13] We use birth state to proxy for residence state at the time of the draft. Evidence in Wozniak (2010) shows that although this assumption generates measurement error by misclassifying individuals into incorrect states of residence at draft age, the misclassification error is not substantially different across education groups.

of men inducted nationally in 1969 divided by the size of his birth cohort. More precisely, we subtract own state inductions from the numerator and own state cohort size from the denominator. 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 (Card and Lemieux 2000, 2001). MW provide a thorough description of the state and national level identifying variation.

Figure 2 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 (i) while Panel B shows residual state level 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.

C. Estimating Equations

Based on the identification strategy described in the previous section, our estimating equations can be expressed as follows (again ignoring additional covariates for conciseness):

(1a) $health_{sc} = a_0 + a_1 vet_{sc} + a_2 educ_{sc} + \mu^h_{sc}$

with the following first stage equations, estimated jointly:

(2a) $vet_{sc} = b_0 + b_1 natrisk_{sc} + b_2 staterisk_{sc} + \mu^v_{sc}$

(3a) $educ_{sc} = k_0 + k_1 natrisk_{sc} + k_2 staterisk_{sc} + \mu^e_{sc}$

National and state-level induction risk yield exogenous variation in both veteran status and education. Our exogenous variation identifies the parameters in Equations (1) and (3) by eliminating endogeneity between these variables and the unobserved error components. Essentially, our IV strategy replaces both *veteran* and *education* with their predicted values in the structural equations. Since the predicted values are driven entirely by exogenous variation in the risk measures, they are uncorrelated with the unobserved error terms in both the main *health* equation (1) and the feedback

equations (2) and (3). Our estimates of the parameters in (1a) are therefore free of confounding effects of correlated errors across the structural equations.

Having exogenous variation in both veteran status and education is the central assumption necessary to identify the direct and indirect effects on health. However, it is not a sufficient condition. Robins and Greenland (1992) explain that "randomization of such interventions [of the intermediate cofactor, or mediator]…can allow separation of direct and indirect exposure effects if only [certain] types [of effects]…are present." Our interpretation is that the assumptions regarding the types of effects required for identification in the Robins and Greenland context are standard in 2SLS and that therefore their assumptions hold in our context.

Another challenge to interpreting our results arises because our instruments affect both veteran status and education. Thus, some marginal college students will have been "sent directly" to college because of the desire to avoid induction risk, and others will have been "sent indirectly" to college because the instruments affected their veteran status which gave them access to the GI Bill and, in turn, college. The estimates in (1a) are therefore free from omitted variables bias, but our estimate of a_2 will reflect the average treatment effect of college going for both the direct and indirect college students. To the extent that the treatment effect varies across these different types of exogenous variation, our estimates reflect the average of two different local average treatment (LATE) effects.

D. Validity of Our Instruments

We present our first stage estimates in the results section, but here we present evidence that our instruments satisfy the exclusion restriction. The exclusion restriction requires that draft risk had no influence on health outcomes except through the channels of college attainment and veteran status. The fact that our identification works through variation in induction risk at the state-cohort level is particularly important for meeting this assumption. In this context, the exclusion restriction

means that the health of *cohorts* cannot have been affected by induction risk, except through cohorts' exposure to different levels of educational attainment or veteran status.

There are two main ways this assumption might fail. First, induction risk may have directly affected health or some other outcome that influences health (other than education and veteran status). Because fitness for duty was a major element in selecting men for duty from among those eligible to serve, it is theoretically possible that some men attempted to avoid induction by intentionally reducing their own health. In surveying the literature both of the time and subsequent studies, we find no evidence that this was a common occurrence. Moreover, the penalties for malingering, if caught, were severe. The offender would be classified as "delinquent," the category with the highest probability of service in the most dangerous service branch.

Alternatively, the draft may have led men at risk to take other actions that in turn influenced their health. MW rule out the possibility that men moved to lower risk states to reduce induction risk. To the extent that migration in young adulthood influences health, it is unlikely that such migration across cohorts differed with their induction risk. MacKinnis (2006) documents that the number of men who emigrated to Canada to avoid conscription was very small compared to deferments via other channels. While dependency deferments, such as those for married men and those with children, were quite common, the available evidence suggests that small numbers of men were induced to marry earlier in response to certain deferment policies and that outcomes for such induced early marriages were similar to those for regular marriages (Bailey, 2010). In sum, enrolling in a four year institution was by far the most certain and most common way for at risk men to avoid conscription. We conclude that the risk of substantive exclusion restriction violations arising from effects of the draft on factors other than schooling and veteran status is small.

A second possibility is that induction risk might be a function of the pre-existing health status of a cohort. We can test for this possibility by looking for a relationship between pre-war

cohort health and subsequent induction risk. We have examined this by regressing infant mortality on our measures of induction risk, and we find no evidence that this early health measure is related to subsequent induction risk at the birth state-birth cohort level (available upon request).

The exclusion restriction requires that the components of draft risk have no influence on health except through the channels of college attainment and veteran status. Our instruments could fail to meet this restriction if young men attempted to exploit local variation in induction risk by moving between localities. But draft board policies prohibited this type of "local board shopping." Rothenberg (1968) notes that it is a myth that "you [can] change draft boards the way you change your patronage of a supermarket or a bank." A young man was required to register with his local board at age 18, and this remained his local board for his entire period of draft eligibility. In the event that he did move away, his original local board always maintained final decision making authority over his draft eligibility although some particulars of draft processing may have been handled by a board closer to his new residence.[14] We find no evidence that young men responded to higher state-level draft risk by moving away. Moreover, our reading of the literature both of the period and subsequent studies suggests that Americans only became aware of local disparities in induction risk after the war concluded.[15] This means that, in addition to their mathematical differences, state and national risk operated differently on the choices of young men in the period because national risk was known with some precision while state risk was largely unknown.

IV. Data and Health Outcome Measures

[14] Rothenberg (1968) p. 48. We have also consulted several reference guides on the draft—books written for young men at the time to inform them about their draft duties and obligations—to determine whether the practice of board shopping may have been common even though it was prohibited by board policy.
[15] See Evers (1961), Tatum and Tuchinsky (1969) and Sanders (1966). Only later did advocates of reforming the Selective Service System highlight this variation.

We focus on mortality as our measure of health outcomes. As outlined in the previous 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 measure of mortality at the birth state-cohort level is the 10-year mortality rate, constructed as follows::

(iii) $$MR_{sc}^{T} = \frac{deaths_{sc}\ between\ years\ T{+}1\ and\ T{+}10}{cohort\ size_{sc}^{T}}$$

where T is the Census year (1980 or 1990), s is state of birth, and c is year of birth. Thus the 10-year mortality rate gives the probability of dying in the ten years following a Census, conditional on being alive at the Census time. One can think of other health and even mortality outcomes that could be the outcome of interest in our analysis. We use the ten-year mortality rate for several practical reasons. First, we only observe population size of our cohorts every ten years in the Census data.[16] It is therefore natural to have deaths from our cohorts over a ten year period as a numerator while using the population size at the beginning of the decade as the denominator. Second, our cohorts are still relatively young in these decades and deaths are rare. Moreover, any effect of education is likely to be gradual and lasting. It therefore seems reasonable to aggregate deaths over a period, rather than looking for different year-to-year effects of education on mortality. Finally, while annual level mortality rates seem to fine, a twenty-plus year mortality rate seems too coarse. Leading causes of death change as our cohorts move from the 1980s into the 1990s. We therefore prefer to examine cause-specific mortality at the ten-year level to determine whether the role of education changes as likely reasons for mortality evolve.[17]

[16] We plan to experiment with using annual population estimates from the National Cancer Institutes-SEER program in future work.

[17] We have produced estimates using twenty-four (1980-2004) and fourteen-year (1990-2004) mortality rates. Results are available upon request.

The data for our analysis come from two sources. First, we use data from the IPUMS microdata 5% samples of the 1980 and 1990 Censuses (Ruggles et al., 2004) to construct denominators in the mortality rate measures as well as the 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.[18]

Our second source is the Vital Statistics data for the period 1981 to 2004. 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 have information on the cause of death, and we use this to construct cause-specific mortality rates.

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.[19]

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

[19] 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).

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.[20] We exclude anyone from our sample who is on active duty in the military, and we define Reservists or National Guard members as non-veterans in our 1990 data, to match the veteran definition in the 1980 Census. We also exclude observations with imputed values for a number of key variables.

Building on the second stage estimating equation described in the previous section, we include a number of additional covariates to estimate our main econometric model as follows:

$$(1c)\ MR_{sc}^{T} = \lambda' X_{sb} + a_1 vet_{sb} + a_2 educ_{sb} + \tau trend_{sb} + \delta_s + \epsilon_{sc}$$

where *s* indexes state of birth and *c* indexes birth year, MR_{sc}^{T} is the 10-year mortality rate as defined in (iii), *vet* is the fraction of veterans, and *educ* is the fraction with a college degree. In addition, *trend* is a linear trend in birth cohort, *X* is a set of state-cohort level controls, and δ represents a full 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.[21]

[20] 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.

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

Our endogenous variables, *educ* and *vet*, are predicted from first stage equations that include the remaining right hand side covariates in Equation (1c) plus $staterisk_{sc}$ and $nationalrisk_{sc}$ as defined in (i) and (ii), respectively. Estimation is done via standard linear 2SLS, weighted by the number of observations in each state-cohort cell.[22] Standard errors are clustered at the birth state-cohort cell level. For the purposes of comparison with the previous literature, we also present results from specifications in which we predict education using only national induction risk, $nationalrisk_{sc}$ (and where we cluster at the birth cohort level).

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 10-year mortality rate as defined in (iii) is 0.0255 for the 1980s and 0.0307 for the 1990s. Table 1 also gives mortality rates by cause-of-death for ten important causes of death for these cohorts, in order of prevalence in the 1980s. For both decades, the three leading causes of death are external causes (including accidents, murders, and suicides), cancers, and heart disease. In the 1980s, external causes accounted for over a third of deaths among these cohorts, while in the 1990s cancer and heart disease were more common causes (CDC 2005).

V. Results

A. First stage results

Table 2 presents the first stage results obtained by estimating versions of Equations (2a) and (3a). Results for the 1980s are in Panel A. The first two columns show estimates of the effect of national induction-risk alone on the likelihood of college graduation; this is similar to the approach taken in previous studies in which education is assumed to be the only endogenous variable. In the

[22] 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.

specifications with and without birth region trends, we see that a 10 percentage point increase in national cohort risk (roughly the entire range of this variable) increased the percent with a college degree by 6.7 percentage points. In other words, cohorts with higher national risk were more likely to graduate college, consistent with Figure 1.[23] The F-statistics suggest that this first stage has substantial power.

The remaining columns show estimates from our preferred approach of identifying both college graduation and veteran status in Equation (1c). For transparency, we estimate two separate first stage equations—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. Again, increased national induction risk is strongly associated with increases in years of college. The F-tests reported for these columns are calculated according to the (corrected) Angrist and Pischke (2009) procedure and indicate that our instruments have sufficient power. The point estimates are somewhat larger than when using national risk alone, indicating increases in the probability of college graduation of 9.3 to 11.2 percentage points over the range of induction risk in our sample. 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.

Columns (4) through (6) address the second endogenous variable in Equation (1c). These show that veteran status positively varies with both national and state cohort risk. 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

[23] 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.

fluctuation in military manpower demands is responsible for more of the variation in veteran status than are the differences in induction risk across states.

Results for the 1990s (Panel B) are qualitatively similar to those for the 1980s. When only national induction risk is included, the estimated effect of draft risk on college graduation is smaller, but in all cases the F-statistics indicate that draft risk was a strong predictor of college graduation and of veteran status through the 1990s.

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

Table 3 presents our main results based on both OLS and 2SLS estimates of Equation (1c). Panel A reports results for mortality over the 1980s. For purposes of comparison, the first four columns show results using only national risk as a single instrument for years of college (without controls for veteran status since it is likely to be endogenous). Our preferred estimates are in columns (5)-(8), where we include both college graduation and veteran status, using our two risk measures as instruments for the IV results. The OLS estimate in column (6), which includes birth region trends, indicates that college graduation is associated with a 0.025 decrease in the mortality rate. This is approximately 40 percent smaller than the effect estimated when veteran status is omitted in column (2). Veteran status is also negatively associated with mortality risk in the OLS specifications.

Columns (7) and (8) show our preferred estimates from 2SLS estimation of Equation (1c), which use both national and state-level induction risk as instruments for college graduation and veteran status. These point estimates of the causal effect of college graduation on mortality are remarkably similar to the OLS estimates, though they are less precise; the same is true for estimates of the effect of veteran status. Thus the 2SLS results imply a large causal role for both college graduation and veteran status in reducing mortality over the 1980s.

The results for mortality over the 1990s are shown in Panel B. Again, we focus on our preferred specifications that include both college graduation and veteran status. The OLS results again indicate a negative relationship between college graduation and mortality, though the magnitude is smaller and less precisely estimated in the latter decade. Veteran status is also negatively associated with mortality. However, the 2SLS results in columns (7) and (8) are different from what we find for the 1980s. For the 1990s, it appears that college graduation has a positive effect on mortality, though the standard errors are large and in the case of column (8) the 95% confidence interval includes the point estimate from the 1980s. Veteran status continues to have a negative effect on mortality for these cohorts.

The imprecision of the 1990 estimates may be driven by the weaker first stage for college graduation; the F-statistics in Table 2 indicate that induction risk is a better predictor of college graduation in the 1980s than in the 1990s. However, we also consider whether the effect of college completion on mortality might change as the cohort ages. For both decades, the mortality rate is low, but different causes (e.g. cancer, heart disease, and stroke) become more prevalent in the 1990s. If the effect of college education is not consistent across causes of deaths, we would expect the effect of college on total mortality to change as certain causes become more important. 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 4 shows the effect of college graduation on the three leading causes of death for men in these cohorts in the 1980s and 1990s. All specifications include birth region trends and a control for veteran status. The OLS results in Panel A show that college graduation is negatively associated with the risk of dying from external causes, cancers, or heart disease in the 1980s. The 2SLS estimates of the effect of college on death due to external causes or cancer are also negative though

not statistically significant. The effect for heart disease, however, is statistically significant and indicates a decrease in mortality due to this cause of 0.014.

For the 1990s, on the other hand, the OLS results show no statistically significant relationship between college graduation and death due to external causes, cancer, or heart disease. And while 2SLS results continue to be imprecise, estimates of the effect of college graduation on deaths due to cancer and heart disease are in fact negative for the 1990s.

In Table 5, we show results for two other interesting causes of death—substance abuse and stroke. For both OLS and IV specifications, estimates for substance abuse are negative and statistically significant for the 1980s. Interestingly, this is the only cause for which we show a positive effect of veteran status on mortality. Neither education nor veteran status is significantly linked to death by stroke. For the 1990s, however, education has no causal effect on death by substance abuse, but does appear to increase death by stroke.

The results in Tables 3-5 indicate that for the 1980s, college education had a negative effect on total mortality, and in particular on mortality due to heart disease. For the 1990s, the picture is more complex—we estimate no statistically significant effect of college education on mortality, and find a positive effect for stroke and external causes but negative effects for cancers and heart disease (though results are generally imprecise). In future work, we plan to explore how college education might influence risk factors and behaviors associated with specific causes of death.

VI. Next Steps: Individual-level Data from the National Health Interview Survey

The National Health Interview Survey (NHIS) is a nationally representative data set collected annually by the Center for Disease Control and Prevention's National Center for Health Statistics since 1957. There are roughly 60,000 households included each year, with detailed data included for

one randomly selected adult within the household in survey years 1997 and onwards. In addition to a wide range of self-reported medical data, the survey includes detailed demographics and socioeconomic information. For this portion of our analysis we will pool data from the 1997 through 2008 NHIS, focusing on the medical information available in the adult health supplement. We are currently applying for access to the restricted version of the NHIS, which will allow us to identify each individual's state of birth. The analysis will be conducted at the individual level, although the two induction risk instruments will be defined at the state by cohort level.

The detailed health measures available in the NHIS will complement the mortality results discussed in Section V. First, we will consider the prevalence of health outcomes that are associated with heart disease and stroke, such as hypertension. We will also model how diabetes, asthma, and lung cancer prevalence is related to educational attainment and veteran status. Similarly, we will use the height and weight information in the NHIS to consider how BMI is affected.

In addition to these measures of morbidity, we will consider mental health, health behaviors, and health care usage as outcomes. The mental health outcomes we will consider include probes for depression and a five-point scale of mental health. The health behaviors available in the data include smoking, exercise and vigorous activity, and alcohol consumption. Finally, there is extensive information in the NHIS on usual places of care and preventative care usage.

Using the more acute measures available in the NHIS data will allow us to detect any small differences in morbidity that might not be reflected in the mortality results. In addition, we can explore the mechanisms through which educational attainment and veteran status affect health by considering health behaviors.

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.

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 Econoimcs*. 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

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

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*, revise and resubmit.

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.

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

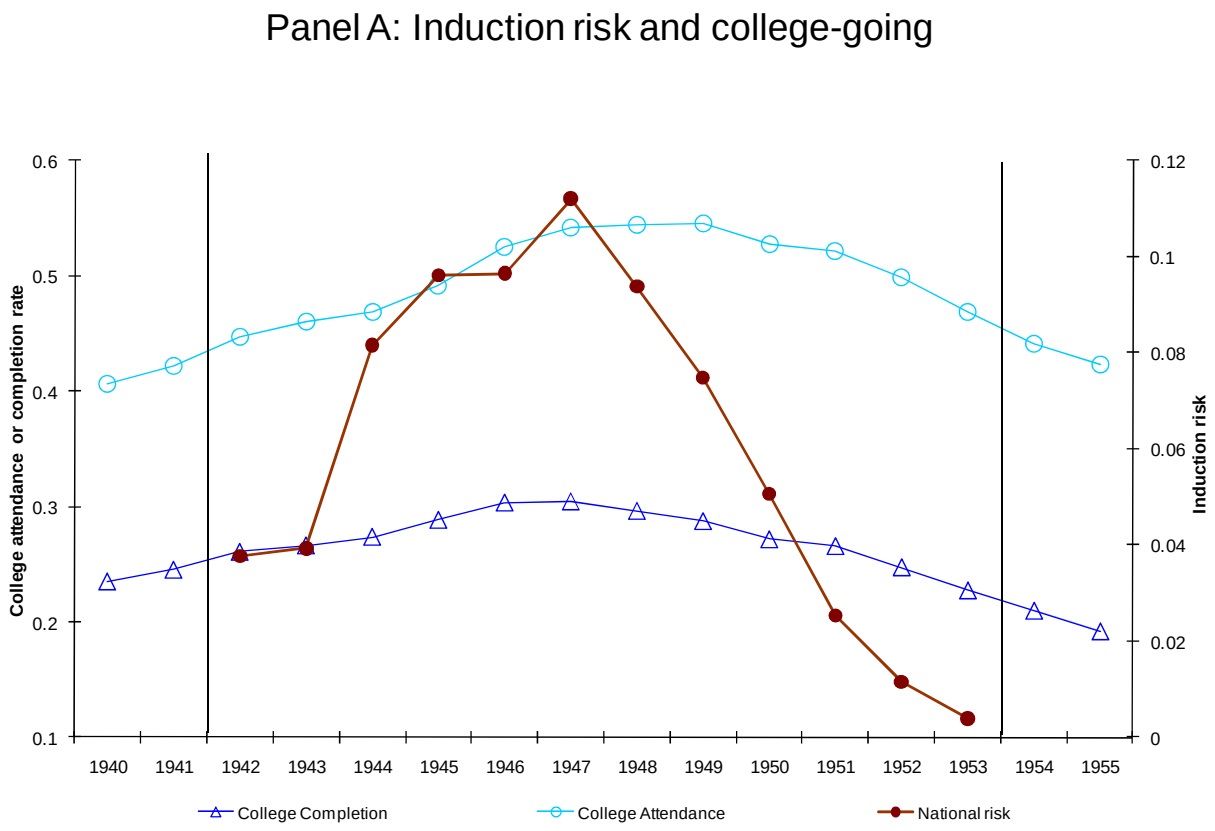

Figure 1. Induction risk and college-going. The figure plots shares of each birth year cohort with 1+ and 4+ years of post-secondary schooling (left axis) and national induction risk as defined in Equation (ii) (right axis).

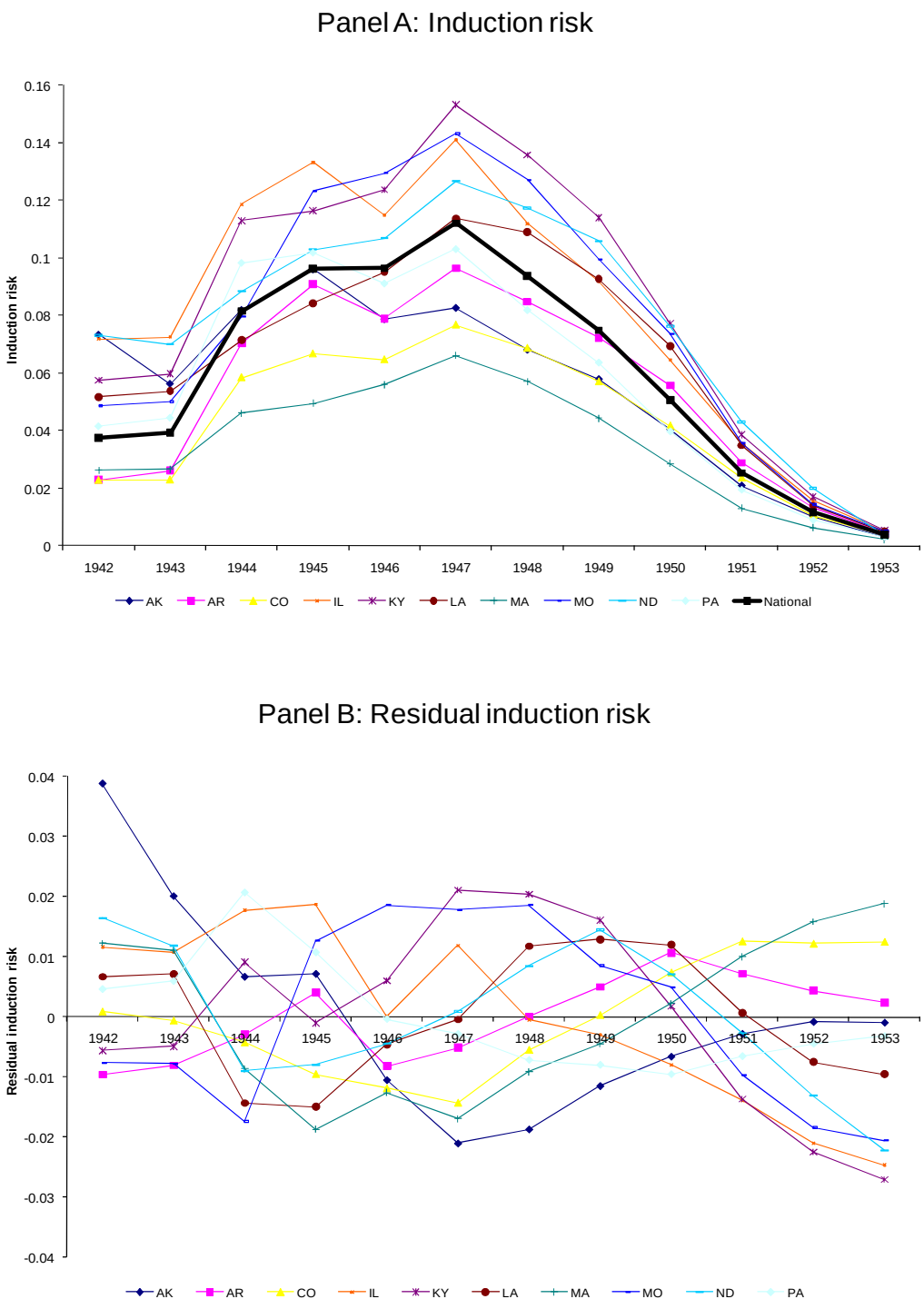

Figure 2. Birth state – birth year variation in induction risk. Panel A plots state risk as defined in (i). Panel B plots (i) adjusted for birth state fixed effects, birth year trend, and national risk as defined in (ii).

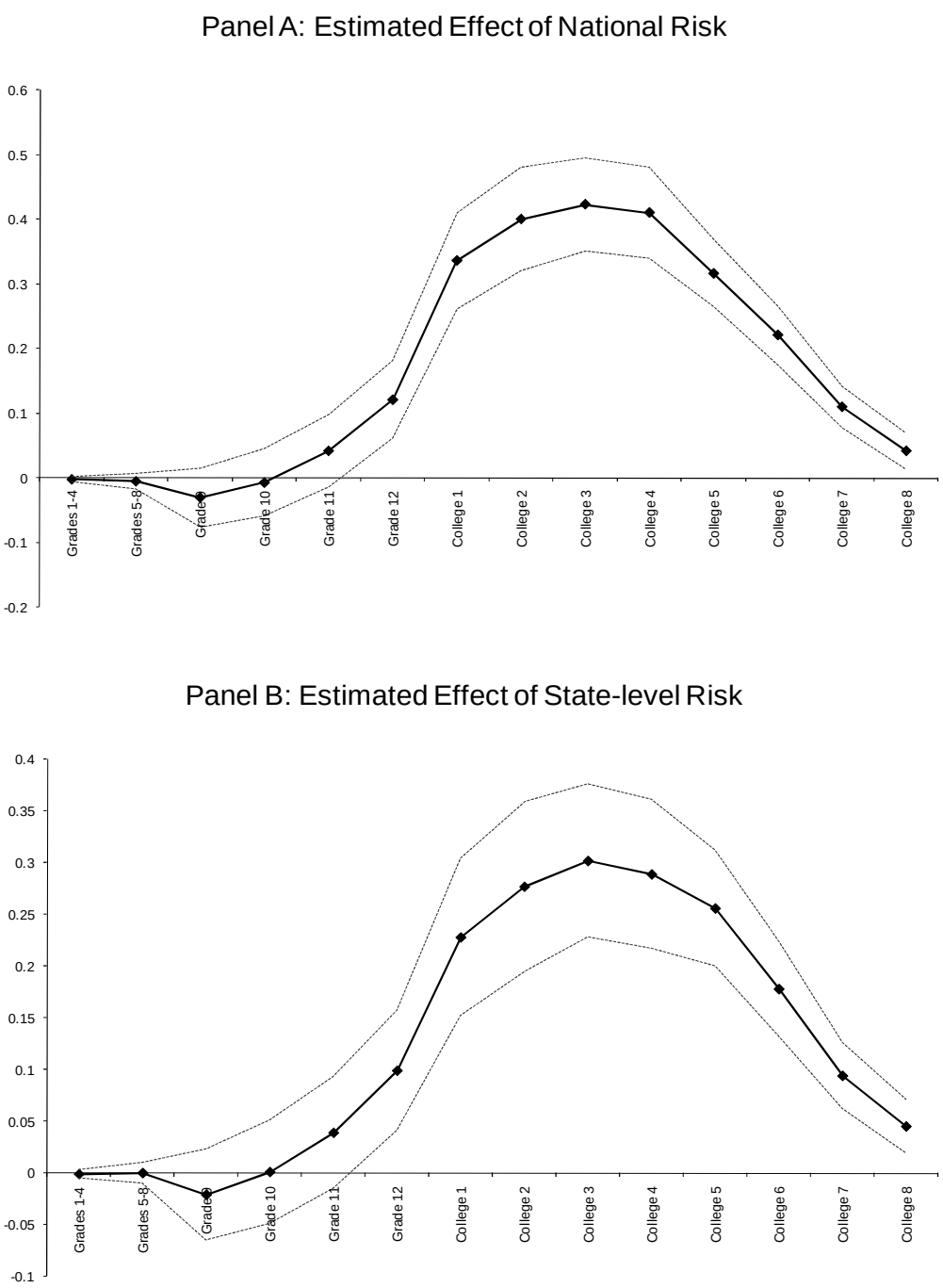

Figure 3. Coefficients and standard errors from OLS regressions of dummy variables for completed education of x-axis grade and higher, inclusive, on national risk and state risk as defined in Equations (ii) and (i), respectively.

Table 1: Summary Statistics

	1980		1990	
VARIABLES	Mean	SD	Mean	SD
Census Data				
Years of Schooling	13.61	0.57	13.60	0.46
College Graduate	0.3508	0.0666	0.3298	0.0657
Veteran	0.3092	0.1231	0.3283	0.1142
Individual Observations	14,392,122		13,247,934	
Mortality Data				
10-Year Mortality Rates by Cause of Death:				
Total	0.0255	0.0062	0.0307	0.0138
External Causes	0.0092	0.0024	0.0054	0.0023
Cancers	0.0039	0.0019	0.0062	0.0039
Heart Disease	0.0038	0.0025	0.0057	0.0035
Liver Disease	0.0011	0.0006	0.0014	0.0007
Cerebrovasc. Disease (Stroke)	0.0005	0.0003	0.0009	0.0007
Diabetes	0.0004	0.0002	0.0007	0.0004
Nephritis (Kidney Disease)	0.0001	0.0001	0.0002	0.0001
Substance Abuse	0.0001	0.0001	0.0001	0.0001
Alzheimer's	0.0001	0.0001	0.0001	0.0001
Hypertension	0.0000	0.0000	0.0001	0.0001
Infectious & Parasitic Diseases	0.0000	0.0000	0.0000	0.0000
Total Deaths	366,365		407,193	
State/Birth Year Cells	600		600	

Notes: Census data are from the 5% sample of the 1980 and 1990 U. S. Census, available from IPUMS. Mortality data are from the Vital Statistics Multiple Cause of Death files from 1980 to 1999. 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. 10-year mortality rates are constructed by dividing the total number of deaths in the decade by the cohort size in the first year of the decade.

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

Panel A: 1980s

	Dependent Variable:					
	College Graduate		College Graduate		Veteran	
National Induction Risk	0.6845***	0.6776***	1.1215***	0.9345***	2.2447***	2.3491***
	(0.0624)	(0.0622)	(0.1039)	(0.0914)	(0.1634)	(0.1579)
State-level Induction Risk			- 0.4071***	-0.2399***	0.8481***	0.7684***
			(0.1023)	(0.0894)	(0.1611)	(0.1560)
Birth Region Trends		X		X		X
F-Stat for Instruments	120.5	118.7	35.1	21.3	238.8	133.1

Panel B: 1990s

	Dependent Variable:					
	College Graduate		College Graduate		Veteran	
National Induction Risk	0.4227***	0.4162***	0.7643***	0.6243***	2.3827***	2.3715***
	(0.0702)	(0.0700)	(0.0932)	(0.0859)	(0.1294)	(0.1360)
State-level Induction Risk			-0.3117***	-0.1900**	0.6259***	0.6426***
			(0.0877)	(0.0800)	(0.1251)	(0.1327)
Birth Region Trends		X		X		X
F-Stat for Instruments	36.2	35.3	21.3	12.3	408.3	205.3

Notes: Standard errors are clustered by birth-year and are in parenthesis. ***, **, 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. Number of observations is 600 in each regression and regressions are weighted by cell size. Trend is a linear trend in age. Additional controls include birth state fixed effects, the employment to population ratio, and the cohort size (derived from the 1960 Census and defined at the birth-year level). F-statistics are calculated following the Angrist-Pischke (2009) procedure for the case of multiple endogenous variables.

Table 3: OLS and IV Estimates for the Impact of Education on 10-Year Mortality

Panel A: 1980s

	No Control for Veteran Status, IV with National-Level Induction Risk				Control and Instrument for Veteran Status, IV with National- and State-Level Induction Risk			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	OLS	OLS	IV	IV	OLS	OLS	IV	IV
College Graduate	-0.0535***	-0.0518***	-0.0678***	-0.0679***	-0.0317***	-0.0245***	-0.0297*	-0.0249
	(0.0097)	(0.0108)	(0.0148)	(0.0149)	(0.0056)	(0.0047)	(0.0166)	(0.0242)
Veteran Status					-0.0088**	-0.0099**	-0.0080*	-0.0091*
					(0.0036)	(0.0035)	(0.0041)	(0.0055)
Birth Region Trends		X		X		X		X
R-squared	0.9052	0.9117	0.9012	0.9072	0.9137	0.9209	0.9116	0.9191

Panel B: 1990s

	No Control for Veteran Status, IV with National-Level Induction Risk				Control and Instrument for Veteran Status, IV with National- and State-Level Induction Risk			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	OLS	OLS	IV	IV	OLS	OLS	IV	IV
College Graduate	-0.0673***	-0.0615**	-0.1588***	-0.1599***	-0.0188*	-0.0099	0.0600	0.0428
	(0.0185)	(0.0199)	(0.0259)	(0.0259)	(0.0103)	(0.0071)	(0.0381)	(0.0464)
Veteran Status					-0.0209***	-0.0202***	-0.0299***	-0.0280***
					(0.0041)	(0.0045)	(0.0063)	(0.0070)
Birth Region Trends		X		X		X		X
R-squared	0.9209	0.9415	0.9002	0.9202	0.9329	0.9521	0.9327	0.9485

Notes: Standard errors are in parenthesis and are clustered by birth year for specifications 1-4 and by birth year-state for specifications 5-8. ***, **, 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. Specifications 3 and 4 use the national induction risk as an instrument for college graduate; specifications 7 and 8 use national and state-level induction risk to instrument for college graduate and veteran status. Number of observations is 600 in each regression and regressions are weighted by cell size. Trend is a linear trend in age. Additional controls include birth state fixed effects, the employment to population ratio, and the cohort size (derived from the 1960 Census).

Table 4: OLS and IV Estimates for the Impact of Education on 10-Year Mortality, by Leading Causes of Death

Panel A: 1980s

	External Causes		Cancers		Heart Disease	
VARIABLES	OLS	IV	OLS	IV	OLS	IV
College Graduate	-0.0048*	-0.0026	-0.0043***	-0.0066	-0.0069***	-0.0139**
	(0.0022)	(0.0105)	(0.0013)	(0.0059)	(0.0019)	(0.0068)
Veteran Status	-0.0014	-0.0015	-0.0036***	-0.0030**	-0.0032***	-0.0017
	(0.0010)	(0.0024)	(0.0007)	(0.0014)	(0.0009)	(0.0015)
Birth Region Trends	X	X	X	X	X	X
R-squared	0.8831	0.8810	0.9312	0.9309	0.9507	0.9484

Panel B: 1990s

	External Causes		Cancers		Heart Disease	
VARIABLES	OLS	IV	OLS	IV	OLS	IV
College Graduate	0.0022	0.0055	-0.0058**	-0.0050	-0.0032	-0.0020
	(0.0014)	(0.0108)	(0.0026)	(0.0150)	(0.0027)	(0.0128)
Veteran Status	-0.0022**	-0.0028*	-0.0069***	-0.0071***	-0.0041***	-0.0045**
	(0.0007)	(0.0016)	(0.0011)	(0.0022)	(0.0010)	(0.0020)
Birth Region Trends	X	X	X	X	X	X
R-squared	0.9232	0.9228	0.9346	0.9334	0.9411	0.9407

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. IV results use national and state-level induction risk to instrument for college graduate and veteran status. Number of observations is 600 in each regression and regressions are weighted by cell size. Trend is a linear trend in age. Additional controls include birth state fixed effects, the employment to population ratio, and the cohort size (derived from the 1960 Census). The dependent variable is the number of cause-specific deaths in the decade divided by the cohort size in the initial year of the decade.

Table 5: OLS and IV Estimates for the Impact of Education on 10-Year Mortality, Additional Causes of Death

Panel A: 1980s

	Substance Abuse		Cerebrovascular Diseases (Stroke)	
VARIABLES	OLS	IV	OLS	IV
College Graduate	-0.0005**	-0.0017**	-0.0006	0.0007
	(0.0002)	(0.0008)	(0.0005)	(0.0016)
Veteran Status	0.0001**	0.0004**	-0.0003	-0.0006
	(0.0001)	(0.0002)	(0.0002)	(0.0004)
Birth Region Trends	X	X	X	X
R-squared	0.4824	0.4378	0.6795	0.6733

Panel B: 1990s

	Substance Abuse		Cerebrovascular Diseases (Stroke)	
VARIABLES	OLS	IV	OLS	IV
College Graduate	0.0002	0.0002	0.0003	0.0079*
	(0.0002)	(0.0010)	(0.0008)	(0.0042)
Veteran Status	-0.0001	-0.0001	-0.0006***	-0.0018***
	(0.0000)	(0.0001)	(0.0002)	(0.0006)
Birth Region Trends	X	X	X	X
R-squared	0.6629	0.6632	0.8695	0.8395

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. IV results use national and state-level induction risk to instrument for college graduate and veteran status. Number of observations is 600 in each regression and regressions are weighted by cell size. Trend is a linear trend in age. Additional controls include birth state fixed effects, the employment to population ratio, and the cohort size (derived from the 1960 Census). The dependent variable is the number of cause-specific deaths in the decade divided by the cohort size in the initial year of the decade.