The Impact of College Education on Geographic Mobility:

Identifying Education Using Multiple Components of Vietnam Draft Risk

Ofer Malamud and Abigail Wozniak

This version: August 2009

ABSTRACT

We examine whether higher education is a causal determinant of geographic mobility using variation in college attainment induced by draft-avoidance behavior during the Vietnam War. We use national and *state-level* induction risk to identify both educational attainment and veteran status among cohorts of affected men observed in the 1980 Census. Our 2SLS estimates imply that the additional years of higher education significantly increased the likelihood that affected men resided outside their birth states later in life. Most estimates suggest a causal impact of higher education on migration that are larger in magnitude but not significantly different from OLS. Our large reduced-form estimates for the effect of induction risk on out-of-state migration also imply that the Vietnam War led to substantial geographic churning in the national labor market. We conclude that the causal impact of college completion on subsequent mobility is large and provide evidence on a range of mechanisms that may be responsible for the relationship between college education and mobility.

(JEL: J61, J24, I23)

Affiliations: Malamud is Assistant Professor at the Harris School of Public Policy at University of Chicago. Wozniak is Assistant Professor of Economics at the University of Notre Dame and a Research Affiliate at IZA.

Acknowledgments: The authors would like to thank Dan Black, Kerwin Charles, Bill Evans, Lawrence Katz, Caroline Hoxby, Thomas Lemieux, Laddie Sula and Lars Lefgren for helpful conversations and comments. Comments from seminar participants at SOLE, the University of Notre Dame, the Midwest Economics Association Meetings, University of Illinois-Urbana Champaign, Dartmouth College, University of Illinois-Chicago, Indiana University-Purdue University Indianapolis, Rutgers University, Princeton University and the University of Maryland improved the paper. All errors are our own.

Data from the 1940-2000 U.S. Censuses show that the long-distance migration rates of college graduates are roughly twice those of high school graduates.[1] This gap is similar in magnitude to the log wage gap between college and high school graduates toward the end of this period. Strikingly, these migration rates by education group have been largely stable over these six decades, despite large increases in post-secondary educational attainment. This is consistent with the possibility that higher education has a causal impact on the geographic mobility of its recipients, but without a credibly identified empirical analysis we cannot rule out other explanations for this pattern, such as falling transportation and communication costs.[2] However, in contrast to the extensive evidence on the causal contribution of education to the wage gap (Card 1999 provides a survey), little is known about the role of education as a causal determinant of geographic mobility. Geographic mobility may afford an individual a larger choice set over jobs and locations and may also reflect an ability to respond appropriately to spatial disequilibria across labor markets. As such, it would constitute an important return to schooling that is distinct from increased wages. In this paper, we use exogenous variation in college going to provide the first credible analysis of the causal relationship between education and geographic mobility.

We use variation in college attainment induced by draft-avoidance behavior among cohorts of men affected by the Vietnam conflict to isolate the causal effect of education on geographic mobility. This approach is inspired by Card and Lemieux (2000, 2001a) who document the excess educational attainment among cohorts induced to enter college as a means of deferring conscription into the Armed Services during the Vietnam conflict. While Card and Lemieux focus on differences in induction risk across birth cohorts, we also exploit *state level* variation in induction risk within

[1] This is true for both migration measures in the Census: out of birth state residence and migration across state/county lines in the past five years. We discuss these measures in more detail later in the paper. For early evidence on educational migration differentials, see Ladinsky (1967), and Greenwood (1975). Wozniak (forthcoming) and Greenwood (1997) provide more recent evidence. Ferrie (2005) provides important evidence on migration rates and the migration-occupation mobility link over the past 150 years, but educational attainment is only available in the Census beginning in 1940.

[2] Migration rates within an education group would fall as educational attainment increased if the propensity to migrate were a normally distributed **unobserved** characteristic correlated with college going, as in a simple selection explanation for the differential in migration rates.

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. The decomposition yields two instruments, which we use to identify the two endogenous variables—education and veteran status—in our empirical application.

This is the first use of a two instrument strategy 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; MacKinnis 2006).[3] The use of two instruments allows us to generate estimates of the causal effect of higher education on geographic mobility that are free of bias arising from the correlation between veteran status and either educational attainment or the individual error term. Of equal importance, our strategy exploits state-year level variation in induction risk to identify the college enrollment decision. This is 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). Our robustness analysis reveals evidence that the assumptions required to identify higher education from the time series variation in induction risk alone are unlikely to be satisfied.

Finally, our identification strategy is particularly appropriate for studying the impact of a college education on internal migration. The increased incentives for college going in our setting leave the relative price of attending any given college unchanged, and thus our instruments encourage college going without affecting which college a student chooses. In other words, our instruments increase college-going in a more general way than other instruments previously employed in the literature, such as proximity to college and variation in state tuition policy (Card 1995; Dynarski 2008). These instruments have provided useful estimates of the return to college

[3] Angrist and Krueger (1992) faced an identical problem in their study of the effects of educational attainment on earnings using draft lottery numbers to identify educational attainment.

education, but because they increase access to college by lowering the price of some colleges relative to others (usually on the basis of geography), they are less appropriate for studying education's effect on mobility because of exclusion restriction concerns. This is an important consideration—particularly when the outcome is mobility—since attending a local college may have different effects on the outcome than attending a distant college. On this basis, our strategy will appeal to those interested in identifying the effect of higher education on migration in the absence of relative price changes. However, we also examine how our results might differ if we used "policy-style" instruments based on college availability and tuition inducements.

Using micro-data from 5% samples of the 1980 U.S. Census of Population, we provide both OLS and 2SLS estimates of the effect of college education on geographic mobility. In cross-sectional OLS regressions, we observe that an additional year of higher education is associated with an increase in an individual's probability of residing outside his birth state at mid-career of three percentage points, or approximately ten percent of the mean of this variable. When we instrument for college with cohort-level induction risk, the causal effect of an additional year of college on geographic mobility is generally larger, but estimates range from 1.9 percentage points to 6.7 percentage points. Moreover, large reduced-form estimates for the effect of induction risk on out-of-state migration suggest that the Vietnam War led to substantial geographic churning in the national labor market.

After presenting our estimates of the causal impact of college going on mobility, we consider the mechanisms through which college may have this effect. We show that our results are driven by increased educational attainment at the post-secondary level alone, not through increases in education generally. This suggests that the particulars of the college experience or skilled labor market are likely drivers of the education-migration relationship, rather than increased cognitive skills. We find evidence of a role for the migration experience of "going away" to college in explaining later adult migration, but other mechanisms, including the geographic diversity of one's peers and the nature of the college labor market, appear to play important roles as well.

The notion that education can provide valuable skills which make individuals more productive in the workplace is well-established in the economics literature. However, the analysis of the relationship between education and labor market outcomes has tended to focus on one of two aspects of this relationship. The first is reduced form, quasi-experimental evidence that education is causally linked to higher earnings (Card 2001). The second is that education improves measured cognitive ability, which is one channel through which education is assumed to affect earnings (Johnson and Neal 1996; Cascio and Lewis 2006). There is less, but growing, evidence that education is causally linked to other important outcomes, such as health (Adams 2002; Currie and Moretti 2003; Lleras-Muney 2005; Jayachandran and Lleras-Muney 2009). Similarly, there is growing interest in and evidence that education confers skills other than cognitive ability that are rewarded in the labor market (Bowles, Gintis, and Osborne 2001; Heckman, Stixrud and Urzua 2006; Borghans et al. 2008; Segal 2008).

Our paper makes contributions to both these growing literatures. First, we test for a reduced-form, causal relationship between education (specifically higher education) and geographic mobility. This has important implications for individual welfare. Geographic mobility, like health, expands the set of choices available to an individual and is therefore an important component of individual utility. Similarly, mobility may have a direct impact on individual earnings, magnifying its importance as a determinant of utility (Wozniak, forthcoming; Bound and Holzer, 2000). In addition to its effect on an individual's choice set of consumption bundles, Shultz (1961) and Bowles (1970) both identified migration as a major form of human capital investment. Our paper can therefore shed light on how one form of human capital investment may encourage another.

This leads to our second contribution to the literature. By focusing on geographic mobility as an outcome, we expand the notion of human capital acquired through schooling to include a broader set of skills and attributes. A causal link between education and migration would imply that geographic mobility is one of the "returns" to higher education. Because migration has long been modeled by economists as the appropriate response to labor market disequilibria, finding a causal link between education and mobility can potentially add much to our understanding of what education confers. There was a period in which leading labor economists considered the question of whether education confers the ability to respond to economic disequilibria, but this line of analysis withered without resolution.[4] Examining the nature of the relationship between education and geographic mobility may provide new insight into this important question.

Finally, concerns about immigration policy's displacement effects on natives (Borjas 2006) and about the need for and effects of geographic reallocation in the recent global downturn also make it a timely outcome for study (Dougherty 2009).

The paper proceeds as follows. Section II explains our identification strategy. Section III describes the data and estimating equations, and IV presents our main results. Section V investigates mechanisms for the causal effect of education on migration, and VI concludes.

II. Identification Strategy

We extend an instrumental variables strategy inspired by Card and Lemieux (2000, 2001); hereafter CL. They show that Vietnam era induction risk varied significantly across birth year cohorts due to differences in military manpower requirements and cohort size. They find that young men responded to an increased risk of induction by attending and completing college at higher rates than pre-draft cohorts.[5] College enrollment was a well known and virtually foolproof way to defer conscription. The Military Service Act of 1967 stated that college students in good

[4] Nelson and Phelps (1966) postulate a higher return to education in the presence of technological change and Welch (1970) provides some related evidence. Schultz (1975) offers a survey of studies examining the ability to deal with disequilibria.

[5] Specifically, increased induction risk led to significant deviations in male college attainment gains from the inter-cohort path traced out by females, whose attainment gains followed a similar trajectory to that of pre-draft males.

standing could defer induction until receipt of undergraduate degree or age 24, whichever occurred first. Prior to 1967, deferment for the purposes of four-year college enrollment was guaranteed in practice as local draft boards issued deferments for a variety of reasons including school attendance.[6] Although new college deferments were eliminated in September of 1971 with the introduction of the draft lottery, those currently enrolled were able to retain their deferments. Thus, the incentive to enroll was great: among "fit" men, risk of induction was [eventually] very high.[7] CL estimate that male college attendance would have been 4 to 6 percentage points lower and graduation 2 percentage points lower had male college attainment in these cohorts followed the female path. Like CL, we use variation in college attendance and completion generated by draft avoidance behavior to isolate the portion of college education that was induced by policy changes on the part of the U.S. government from that which would have been chosen by individuals independent of the policy change.

An artifact of our identification strategy for education—using draft avoidance behavior—is that the likelihood an individual is a veteran also varies across cohorts in our sample. Veteran status is a plausible determinant of geographic mobility (Pingle 2007), and may be correlated with the individual error term in much the same way that education is. Our equation of interest therefore contains two endogenous variables. To deal with this, we could exclude veteran status from our included controls. This would generate a biased estimate of the causal impact of education on mobility, but the size of the bias might be estimated using formulas for omitted variables bias and assumptions about the magnitude of the causal effect of veteran status on mobility.[8]

Instead, our approach is to deal with the problem of multiple endogenous variables directly. We exploit state level variation within the cohort level variation identified by CL in order to obtain

---

[6] 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).

[7] CL provide a rough estimate of the risk of being drafted conditional on not enrolling in college.

[8] This is the approach taken by Angrist and Kreuger (1992). If the coefficient vector is estimated using instrumental variables, the estimate is not guaranteed unbiased if the model contains more endogenous variables than instruments. Wooldridge (2002), p. 83.

separate instruments that can be used to identify both higher education attainment and veteran status simultaneously. The existence of state-cohort level variation allows us to break national induction risk into its constituent parts, yielding two instruments.

A. Constructing Multiple Instruments: State and National Level Induction Risk

Institutional factors caused induction risk to vary across states within a cohort as well as across birth year cohorts. 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 in turn did the active work of ordering men to be inducted. The monthly state quotas sum to the national draft call each month.[9]

CL use total inductions nationally as the numerator in the following annual measure of induction risk for a birth cohort:

(i) $$CL\ induction\, risk = \frac{I_{bt}}{N_{bt}}$$

where I is the number of inductees from a birth year cohort, N is the number of men in that cohort, b indexes birth years, and t indexes calendar years since a cohort was at risk to be drafted for several years. Similarly, young men residing in state s faced state cohort risk that is analogous to the CL measure, where s indexes state of residence:

(ii) $$state\ cohort\ risk_{sbt} = \frac{I_{sbt}}{N_{sbt}}$$

State cohort risk is our first instrument. (We postpone the details of its construction until Section III.) We then use our state level data on *I* and *N* to construct a second instrument in the following manner:

(iii) $$national\ cohort\ risk_{sbt} = \frac{I_{-sbt}}{N_{-sbt}}$$

---

[9] Information in this paragraph is based on Shapiro and Striker, *Mastering the Draft,* Chapter 20.

This measure defines national cohort risk for a man living in state s and born in birth year b as the number of inductees from all other states –s and birth year b, divided by the total number of such men. This is highly correlated with CL's measure (i) but varies across state-birthyear cohorts. The inclusion of the national and state cohort risk measures is similar to the inclusion of state and national unemployment rates in the right hand side of a regression (as in Ziliak, Wilson and Stone, 1999). Our final measures of state and national cohort risk are actually averages of the measures in (ii) and (iii) over the years a young man was ages 19 through 22. This follows CL, who argue that this average induction risk best reflects the draft-based incentive to stay in college.

State level variation in induction risk—such that the probability of induction was not equal to (i) for all members of cohort *b*—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."[10] They later write, "The conclusion seems inescapable: local board autonomy implies both within state and between state variability, even among socioeconomically similar board jurisdictions."[11]

A second source of state-year variation in induction risk were the severe communication delays between the federal, state, and local officials in charge of the draft. These delays meant that local draft boards knew the number of registrants available in their jurisdiction at a point in time 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

[10] Davis and Dolbeare, *Little Groups of Neighbors*, Page 18.
[11] *Ibid.* Page 84.

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.[12]

We examine our exploited state-year level variation in Figure 2, which graphs the variation in state-level induction risk for a selection of states over our period of interest. Panel A shows raw induction risk (i.e. quantity (ii)) while Panel B shows residual induction risk after controlling for a time trend, state-of-birth fixed effects, and national risk as defined in (iii). It is apparent from Panel A that some states have higher levels of induction risk associated with them on average. However, because we include birth state fixed effects in our main specification, this is not part of our identifying variation. The data in Figure B more closely approximates our identifying variation. (In practice we will add other controls to our 2SLS specifications.)

There are two things to notice in Figure B. First, there is indeed variation across states in the induction risk that is not explained by fixed state characteristics alone. Second, relative induction risk in a state appears to fluctuate smoothly over time. This is consistent with the known sources of cross-state variation in draft risk. While Davis and Dolbeare identify purely idiosyncratic variation in board decision-making (their item [3]), the other sources of state and local variation in risk would tend to generate smooth fluctuations in state relative risk. For example, if the demographics of a local board jurisdiction are changing smoothly such that more men are aging into the draft pool but the local board is slow to adjust the number of deferments it grants because of pressures to achieve standardization, then risk will slowly decline in this jurisdiction over time.

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

These smooth fluctuations do not pose a threat to our identification strategy as long as our instruments meet the IV assumptions. If this is the case, then the problems for statistical inference posed by serial correlation in our instrumental variables (as opposed to purely idiosyncratic fluctuations over time) can be handled through appropriate standard error adjustments along the lines discussed in Bertrand, Duflo, and Mullainathan (2004). (Also see footnote 24 on this.)

B. Do Our Instruments Meet the IV Assumptions?

Our instruments need to meet the fundamental requirements of IV: joint significance in the first stage, the exclusion restriction, and the monotonicity condition for the interpretation of our estimates as local average treatment effects. We present our first stage results in a subsequent section, but as our IV strategy is new to the literature, the other two requirements bear some discussion.

The exclusion restriction requires that the components of draft risk had no influence on migration choices 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 level variation in induction risk by moving between localities. 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.[13]

[13] 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. We find no evidence of this. See Evers (1961), Tatum and Tuchinsky (1969) and Sanders (1966). Only later did advocates of reforming the Selective Service System highlight this variation.

A more likely scenario is that draft risk increased the likelihood that a young man attended college at a significant distance from home since marginal college goers may have had to search more widely to find a college that would accept them. If true, this scenario does not imply a violation of the exclusion restriction. Rather, it implies that local treatment effects we identify are particular to the setting from which we take our exogenous variation (something which is common to all instrumental variables applications). Our effects may be larger than other causal estimates of the same parameter if going away to college is an important mechanism through which college has a causal effect on mobility.

Finally, it is well known that some men enlisted voluntarily in order to secure safer military assignments. (The Texas National Guard is an example.) Again, this possibility does not constitute an exclusion restriction violation. This is simply another channel through which the draft "encouraged" military enlistment. The other channel is the local draft board's threat of incarceration.

Monotonicity requires that there be no "defiers" in the population—i.e. individuals who would have gone to college or become a veteran in the absence of the draft but who did not do these things because of the draft (Angrist and Imbens, 1994; Angrist et al. 1996). It is difficult to test for such a subpopulation directly, and most IV applications assume there are no defiers. We are, however, somewhat concerned that blacks may have been crowded out of college slots by large numbers of whites attempting to avoid conscription, a violation of monotonicity. In light of this, we examine the sensitivity of our results to this assumption by estimating our key models separately for blacks and whites. We find no difference in the key parameters across these subgroups, suggesting that any individual instances that violate the no-defiers assumption do not generalize to the black population as a whole.[14]

[14] For instance as discussed in Kuziemko (2008).

III. Sample Construction and Estimating Equations

Our data on men in the "Vietnam generation" come from the IPUMS microdata 5% sample of the 1980 Census. (Ruggles et al., 2004). The main cohorts of interest to CL were men born between 1935 and 1955. We focus on a similar set of cohorts because these men turned 19, the age of peak draft risk, during the period when induction risk made its dramatic rise and fall with the Vietnam War. However, due to data limitations on state-level induction rates and enrollment, we must restrict our sample to men born between 1942 and 1953. Since these cohorts experienced the largest rise and fall in induction risk, we do not lose much time series variation relative to CL by excluding the earlier and later cohorts.

The measure of migration we employ in this paper is a dummy variable indicating that an individual resides outside his birth state when we observe him in the Census. This is equivalent to an indicator for whether he is residing outside his birth state at a certain age—i.e. his late 20s or early 30s. This measure captures long distance migration that generally involves a change of local labor markets.[15] Because long distance moves involve changing labor markets, we feel they are a more appropriate outcome measure than short distance moves given our interest in the relationship between education and the ability to make changes in the face of disequilibria. We can also show that educational differences in 1-year and 5-year migration rates are similar to those in the out of birth state measure over longer distances.[16] For example, college graduates are roughly twice as likely as high school graduates to have moved across state lines in the last five years. This gap narrows slightly for older individuals and when the one year measure is used, but the educational differential remains substantial. By contrast, both the 1- and 5-year migration rates show that high school graduates are considerably more likely than college graduates to have moved within their county. This is particularly true at younger ages. Based on these comparisons, we conclude that our

[15] See Saks and Wozniak 2009 for a discussion of the measurement error involved in assuming a move across state lines involves a move between local labor markets.

[16] We tabulated 1-year and 5-year migration rates for 20 to 45 year olds in our four education groups from the 1981 Current Population Survey and the 1980 Census, respectively.

measure is a good proxy for the likelihood of a long-distance move over shorter time horizons. However, we prefer the lifetime rates since they more closely approximate a stock of migration choices rather than a flow as do the five-year rates.[17]

The main shortcoming of the dummy variable measure is that it is potentially confounded by parental migration patterns. It is well known that an individual's education level is correlated with that of his parents, and thus some of the observed migration differentials using this measure could arise from more educated parents moving their children (who potentially go on to get more education) across state lines. Fortunately, the migration advantage that more educated parents confer to their children is small compared to migration differentials observed across education groups later in life.[18] Thus we interpret migration rate differences observed according to our measure as reflecting migration undertaken by individuals themselves rather than their parents.

We use birth state to proxy for residence state at the time of the draft. Evidence in Wozniak (forthcoming) 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.

Our identification strategy requires that we choose a measure of education, since we cannot separately identify the effect of one level of higher education from another. To inform our choice of education measure, we explore the impact of our instrument on educational attainment. Figure 3 plots the coefficients and standard errors from separate regressions of educational attainment at each grade level and higher (inclusive). The figure shows that increased induction risk increased male educational attainment at all post-secondary levels. Our main estimates use years of schooling

[17] The one year rates are only available in the CPS.

[18] Using the NLSY 79, Wozniak (forthcoming) shows that the difference in rates of out of birth state residence at age 14 across individuals who go on to become high school graduates and college graduates is small, around four percentage points. Most of the education group differential in migration rates among 30-year-olds in the NLSY using our Census measure represents migration that occurred after age 14. Of course it is possible some migration is driven by the decision about where to attend college. This mechanism could potentially explain education group differences in migration and we consider it later in the paper.

as our education measure but we also present alternative specifications using discrete measures such as graduating or attending college.

It is possible that measurement error in education (in any of the measures) could bias our IV estimates, along the lines explained by Black et al. (2000) and Kane, Rouse, and Staiger (1999) for cases of non-classical measurement error in an endogenous variable.[19] Unfortunately we are not able to implement the bounding procedures developed by these authors, since we do not have the second independent observation on an individual's educational attainment these corrections require. We simply note this is a concern in most IV estimates of the returns to schooling. However, we will show that our conclusions are robust to alternative measures of educational attainment. This is reassuring since these measures suffer from potential bias to different degrees. 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). Years of schooling, while suffering from non-classical measurement error to some degree, is unlikely to be plagued by error that is purely negatively correlated with the recorded value, as is the case with the dummy variable schooling measures. Measurement error is likely of greatest concern in IV results using the college attendance dummy to measure education.

To construct the measures in Equations (ii) and (iii), we obtained data on the number of inductees from each state for each six-month period spanning 1952 to 1972 from reports of the Selective Service and converted these into electronic format. Like CL, we measure cohort size (at both the national and state levels) using data on school enrollments of 17 year olds. We estimated state cohort size using enrollment numbers spanning 1959 to 1970, the academic years in which our cohorts of interest were in 11th grade.[20] Thus state-cohort level risk (hereafter state risk) for a young

[19] Kane et al. (1999) note that the direction of the bias is not clear in the general case of non-classical measurement error in the endogenous variable, although Black et al. show that in some 2SLS specifications the IV estimates constitute an upper bound on the true parameter.

[20] We also check our estimates using enrollment in 10th grade. For 1959 and 1960, we only have information on enrollment in all high school grades so we divide this figure by 4.

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. National level risk for the same young man roughly equals the number 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. Like CL, we construct an average draft risk for the years a man was 19-22 since draft risk was non-trivial for men ages 20 to 22.

Our basic second stage estimating equation is the following:

(1) $move_{isb} = \beta' X_{isb} + \delta yrs_{isb} + \lambda vet_{isb} + \tau trend_{isb} + \Theta_i + \varepsilon_{isb}$

move is our indicator for out of birth state residence, *yrs* is years of schooling completed, *vet* is a dummy variable for veteran status, *trend* is a linear time trend entered as *i*'s birth-year cohort, X is our set of demographic controls (which includes indicator variables for black and Hispanic), and Θ is a complete set of state of birth dummies. As before, s indexes state of birth and b indexes birth year. Since our sample is drawn from only a single cross-section of Census data, we omit time subscripts. *yrs* and *vet* are predicted from first stage equations that includes the remaining right hand side covariates in Equation (1) plus $staterisk_{isb}$ and $nationalrisk_{isb}$ as defined in (ii) and (iii), respectively. All estimation is done via standard linear 2SLS.[21]

All covariates in (1) are included in all subsequent specifications; in our tables, we refer to these as the "demographic controls". The inclusion of age is equivalent to a year of birth trend in our cross-sectional data. Age also accounts for national (linear) trends in male college completion. We allow for state of birth fixed effects to account for different rates of geographic mobility across states. This also removes variation arising from states that have persistently higher or lower than average induction rates, which may be correlated with other state characteristics which are in turn correlated with migration rates. Industrial composition, which may affect both migration choices

[21] 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.

and induction rates if deferments are routinely granted for certain types of work, is one example of a potential confounding factor that state dummies remove.

We also show results for two additional specifications that include an expanded set of controls. The first adds two variables to capture labor market conditions facing a cohort at the time of the college enrollment decision: a measure of employment conditions and the log of cohort size. Employment conditions are measured as the employment to population (epop) ratio in the individual's state of birth the year his cohort turned 19. Cohort size is measured as the log of the number of respondents from a birth year cohort in the 1960 Census. The epop measure approximates state labor demand conditions facing a birth year cohort at the time they decided to enter college. Log cohort size approximates the labor supply available to meet those conditions. Together, we believe these measures capture labor market conditions that may have occurred alongside changes in state-level induction risk and that are thought to influence college enrollment decisions.[22] Both factors may also be independently related to inter-state mobility levels within a cohort. Saks and Wozniak (2009) show that internal migration rates, including inter-state migration rates, are procyclical. The literature on cohort size suggests that large cohorts have relatively poorer educational and behavioral outcomes than large cohorts. Although we know of no work that links cohort size to mobility, we include it since cohort differences in other outcomes may be related to mobility potential. The second additional specification adds nine linear birth region trends to this expanded set of controls. These trends further account for changes in college-going or migration that may have been correlated with changes in state level induction risk.

We restrict our analysis to men. As women were not subject to induction at the time of the Vietnam War, we do not expect to find large treatment effects of male induction risk on female college going. This implies that women might make a useful comparison group, and this is how CL treat women in their analysis. They assume that the trend in female college attainment is the

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

appropriate counterfactual trend for male college attainment in the absence of the Vietnam War. This allows them to calculate excess college going by men. However, there are reasons to believe that female college attendance may have been affected by male college going during the Vietnam years, with large inflows of men into college either crowding out women who would otherwise have attended or encouraging more women to attend to take advantage of marriage market prospects. Moreover, while women may or may not make an appropriate counterfactual for men in the context of college going, marriage markets that link men and women in their geographic location decisions mean that they certainly are not a useful counterfactual in the context of migration. For these reasons, we prefer to control directly for counterfactual male trends in college going and its co-occurring determinants—as outlined above—rather than indirectly as is suggested by the CL approach. Finally, our use of state-year variation means identification is generated in part by comparing men from different birth states to their peers in a given birth year cohort. This within-cohort identifying variation reduces the need to find comparison groups outside our men in cohorts of interest. We also estimate our main specifications on two placebo groups of men to show that our results are unlikely to be explained by changes across birth states and over time in college going other than those captured by our instrumental variables.

Descriptive statistics from the Census samples of the cohorts we consider are given in Table 1 which summarizes the variables used in our analysis using the sample of men born between 1942 and 1953. Migration is a common outcome in our sample, with roughly one-third of men living outside their birth states and the average man living over 300 miles from his place of birth. Our main demographic controls include age, veteran status, and state of birth. Veteran status is a dummy variable equal to 1 if an individual reported Vietnam era service.

Figure 1 plots the means of selected variables separately for each birth year cohort. The course of the draft expansion across cohorts is evident in Panel A, which also shows the high correlation between our measure of national risk in (iii) and CL's measure in (i). The increase in male college going is also visible in Panel A, particularly for cohorts born between 1944 and 1950, during the main rise and fall in induction risk. Panel B of Figure 1 plots our measures of education and migration (adjusted for a basic set of controls) alongside our national risk measure. This is essentially a graph of our reduced form results. It is clear that migration and educational attainment increase along with national induction risk across cohorts.

The inclusion of controls in our 2SLS regressions removes a large amount of variation from our instruments. To assess the remaining variation, we regressed both risk measures, (ii) and (iii), separately on the sets of controls just described. The R-squared from all six versions of this exercise—two outcome variables times three sets of controls, used in Tables 3 through 6 below—ranged between 0.25 and 0.42. This means that we are using approximately two thirds of the variation in the risk measures to identify our endogenous variables.[23]

IV. Results from the Vietnam Generation

A. First stage and reduced form results

Table 2 presents first stage results using the sample of men described in Table 1 based on our basic Equation 1 specification as well as the two augmented specifications. We maintain the convention of showing results for all three specifications for our main results. We also maintain the convention of clustering our standard errors at the birth year level in specifications using only a single instrument or national aggregates for cohort size and the employment rate. In specifications using both instruments or state level measures of cohort size and employment rates, we cluster the standard errors at the birth state-birth year level.[24] In Panel A, national and state cohort risk are entered separately as the only variable identifying years of schooling. This is similar to the approach

[23] Most of the variation is explained by the inclusion of the linear trend in age and another non-linear variable. In our exercise, this was the national unemployment rate, but we did not perform an exhaustive set of regressions that examined all possible subsets of our control variables.
[24] We have also explored some specifications using multi-way clustering on both birth year and birth state, following procedures suggested by Cameron, Gelbach, and Miller (2006). The standard errors on the coefficients increase somewhat, but the qualitative results remain similar.

in other studies in which education is assumed to be the only endogenous variable. This is not our preferred approach to identification, but we present the results for the purpose of comparison.

The results in columns (1) through (3) of Panel A show that a 10 percentage point increase in national cohort risk (roughly the entire range of this variable) increased the average years of completed schooling in our sample by 0.35 to 0.45 years with the impact declining somewhat as we add additional controls to the basic specification in column (1). These estimates show that cohorts with higher national risk also attain more years of schooling. This is consistent with Figure 1 and the results in CL (2000, 2001). Although our specification differs from that of CL, the magnitude of our coefficients is similar and the F-statistics suggest that this first stage has substantial power.[25]

We repeat the first stage estimates using only state cohort risk as the identifying variable in columns (4) through (6) of Panel A. When national cohort risk is excluded, state cohort risk positively predicts college graduation rates for men in our sample although the point estimates are slightly smaller than those generated by national risk alone. We speculate that the similarity in the estimates of the effects of the two risk measures is driven by the strong correlation (of 0.93) between our state and national cohort risk measures.

Panel B shows first stage equations from our preferred approach of identifying both years of schooling and veteran status in Equation 1. For transparency, we estimate two separate first stage equations—predicting college graduation and veteran status—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.

Columns (1) through (3) in Panel B show that, as in Panel A, increased national induction risk is strongly associated with increases in years of schooling. The point estimates here are somewhat larger than when national risk is included alone, indicating increases of 0.5 to 0.6 years of

[25] Table 1 in Card and Lemieux (2001a) shows estimates corresponding to 4.6 percentage point increase for college graduation. Using our specification with a dummy variable for college graduation as left hand side variable suggests an increase of 3.5 to 4 percentage points in college completion for the same change in induction risk.

schooling over the range of in induction risk in our sample. On the other hand, years of schooling are insignificantly or negatively related to state cohort risk when national induction risk is included. This is in sharp contrast to the relationship between years of schooling and state risk in Panel A, which suggests that increases in schooling attainment are largely driven by variation in national risk.[26]

It is not clear a priori what relationship we should expect between state risk and years of schooling in the Panel B specifications. If young men were unaware of their state level risk, we would expect no relationship, as we see in column (1). On the other hand, if young men were aware of their state risk, we might expect a positive relationship between state risk and years of schooling as men "dodged in to college" more in riskier state-years. We find no evidence of this. Instead, in the augmented specifications we find a negative relationship between state risk and years of schooling. We can think of at least two reasons for this. First, it may be the case that state risk increased a young man's likelihood of becoming a veteran and that being a veteran in turn decreased his likelihood of attending college. This is not unlikely, given that the well known hostility toward the war that prevailed on many college campuses may have deterred some veterans from attending. Alternatively, the negative relationship in columns (2) and (3) may be an artifact of the high correlation (collinearity) between national and state level risk, which may be exacerbated by including the non-linear controls for cohort size and employment rate. The coefficient pattern across Panels A and B—in which state level risk positively predicts education when national risk is excluded and the point estimates on national risk increase when state risk is included—is consistent with collinearity (Verbeek 2000 and Wooldridge 2002).[27]

Neither interpretation of the results in columns (2) and (3) of Panel B invalidates our IV strategy. Ex ante, increased state and national risk both increase the likelihood that a man becomes

[26] National cohort risk was relatively easy to judge as national draft calls were widely reported while deviations of state cohort risk from national were unknown ex ante. These facts suggest to us that state cohort risk had relatively more influence on an individual's veteran status while national cohort risk influenced his likelihood of college attainment. This independence in the roles of the two instruments is not necessary for identification. All that is required in the 2SLS method we apply is that the number of instruments equal or exceed the number of endogenous regressors.

[27] Authors' simulations show that the probability of collinearity alone generating coefficient estimates of opposing signs on two variables that are significant at the 5% level is about 1 in 20.

a veteran and that he attends college; this is the required monotonicity condition. Conditional on becoming either a veteran or a college attender, a young man may then be less likely to also become the other. Indeed, this is why the draft drove young men into college. This simply highlights the desirability of including controls for education and veteran status simultaneously in our setting. Multicollinearity does not affect the predicted value of the left hand side variable, although it has obvious effects on the coefficients estimated for the collinear variables.[28] Thus collinearity does not introduce systematic bias since the predicted values of the endogenous variables in the first stage are correct in expectation, although it may contribute to imprecision in our second stage estimates.

Columns (4) through (6) address the second endogenous variable in Equation 1. These show that national and state cohort risk both covary positively with veteran status. This is reassuring since higher rates of induction 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.

Table 3 presents reduced form estimates of the impact of induction risk on our measure of geographic mobility.[29] These results are striking. The estimates show that increases in national cohort risk are associated with a large positive increase the likelihood than an individual is residing outside his birth state. Men from the highest risk cohort are 8 to 13 percentage points more likely to reside outside their states of birth than men in the lowest risk cohort. This difference is the equivalent of roughly one-fourth to one-third of the mean of this variable in Table 1. Once state cohort risk is included, the impacts of national risk on mobility are even larger and state cohort risk has generally weak negative impacts on migration. The remainder of the paper is devoted to

[28] Verbeek (2000).

[29] Note that we drop specifications using only state cohort risk to identify education in our subsequent analysis. This was only presented as an alternative first stage in Table 2.

demonstrating that higher levels of college graduation are the mechanism driving these reduced form estimates.

B. 2SLS estimates of the effect of higher education on migration

Table 4 presents our main estimates of interest. We show results from the estimation of Equation 1 via both OLS and 2SLS and with and without a control for veteran status. When we follow the single instrument strategy (for comparison), as we do in Panel A, we exclude the veteran status control since it is believed to be endogenous. We use our national risk measure as the single instrument for years of schooling in this case. Our preferred estimates are in Panel B, where we instrument for both years of schooling and veteran status using our two risk measures.

In columns (1) through (3) of both panels, we report our OLS estimates. An additional year of schooling is associated with a 3 percentage point increase in the likelihood that a man resides outside of his birth state. This estimate is robust to the inclusion of additional controls across specification columns and to the addition of the veteran status control in Panel B. Consistent with our concern that veterans may be more geographically mobile than non-veterans, we find that veteran status is associated with a 4 percentage point increase in our migration measure. If men who are willing to serve in the military are also willing to move geographically, this relationship could bias our estimated coefficient vector and it is the reason we need to instrument for veteran status as well as education.

Columns (4) through (6) of Panel A show results from 2SLS estimation of Equation 1 and its augmented versions using national risk to instrument for years of schooling and omitting veteran status from the equation. Years of schooling have an insignificant relationship to migration in column (4), but the augmented specifications in (5) and (6) show a positive and significant – and precisely estimated – relationship. As is often the case in the literature on instrumented returns to schooling, the IV estimates exceed the OLS estimates, in our case by a little over 25%.

Columns (4) through (6) of Panel B show our preferred estimates from 2SLS estimation of Equation 1 using our two induction risk instruments to correct for endogeneity in both years of schooling a veteran status. The estimates of the causal impact of years of schooling on migration are larger than in the single IV estimates of Panel A, although they also cover a fairly large range. The smallest estimate implies that an additional year of schooling raises the probability of an out of state residence by 4.4 percentage points while the largest implies an increase of 6.7 percentage points. The 2SLS estimates in Panel B have large standard errors, but this is not surprising given the added imprecision our two instrument strategy entails and all are significant at conventional levels.[30] The estimates across all three 2SLS specifications in Panel B consistently imply a large causal role for years of higher education schooling in determining geographic mobility. Interestingly, after instrumenting veteran status has no significant impact on subsequent mobility. This indicates that the significant relationship between veteran status and mobility in the OLS equations is in fact driven by selection of more geographically mobile men into the military.

The OLS coefficient on college graduation typically lies within the 95% confidence interval around our 2SLS estimates because the latter are estimated with large standard errors. We can compare the estimates from our OLS and 2SLS estimates more formally using a Durbin-Wu-Hausman test for endogeneity. We can reject the null hypothesis that OLS and IV are the same at the 5 percent level for estimates in column 4 of Panel B of Table 5. In the other five cases, OLS and IV estimates are not statistically distinguishable.[31]

C. Discussion of 2SLS Magnitudes and Specification Checks

We consistently find estimated treatment effects that equal or exceed OLS estimates. On the one hand, this pattern is not uncommon in the literature on instrumented returns to education, and

[30] Additionally, there is some cause for concern about the consistency of the standard errors in the single-instrument estimates, since the number of clusters in that case is small. Our preferred estimates alleviate this concern since the number of clusters is sufficiently large. See discussions in Bertrand et al. (2004) and Wooldridge (2003).

[31] The F-statistic of this test for the 2SLS specifications in Panel A of Table 5 are 0.78, 0.99, and 1.00 respectively (p-values of 0.3960, 0.3407, and 0.3395); the F-statistic of this test for the 2SLS specifications in Panel B of Table 5 are 6.11, 1.82, and 1.90 respectively (p-values 0.0310, 0.2048, 0.1958).

we generally cannot reject equality of the OLS and 2SLS impacts.[32] Two standard explanations immediately come to mind. First, as is also the case in the returns to schooling context, measurement error might bias OLS estimates downwards. This bias might be more likely in our sample if older respondents are more likely to misreport educational attainment in retrospective surveys. Eliminating bias due to measurement error can increase IV estimates above OLS estimates. A second possibility is that heterogeneity in the response to treatment leads to local average treatment effects on the marginal college graduate (identified in our 2SLS estimates) that differ from the average effect (identified in OLS). We examined evidence for this in the National Longitudinal Survey of Youth by dividing an NLSY sample of men into groups whom we considered a priori to have different propensities to complete college. We found that the OLS returns to college completion were higher among groups who were more likely to be marginal college attenders in the Vietnam era—specifically, men in the upper quartile of the AFQT distribution and with mothers who had some college education. We interpret this evidence as suggestive of the possibility of heterogeneous treatment effects.[33]

It is also worth considering the possibility that the historical setting of our instrument leads to larger treatment effects in the cohorts we examine than would be found in a more culturally neutral experiment. The atmosphere surrounding the decision to avoid Vietnam by going to college may have served to heighten an existing difference between marginal and inframarginal college graduates in a manner that increases the estimated causal impact of college on mobility. For example, if marginal college graduates are less similar to their family and high school friends than inframarginal college graduates, they may find it easier to relocate over long distances leading to heterogeneous treatment effects and the resulting 2SLS > OLS coefficient relationship. Unfortunately we have no way of testing this claim. While we allow that the historical setting we use may contribute to our large estimates, we do not believe that the causal impact of college on

[32] Card (1999) addresses the differences between OLS and IV estimates of returns to schooling.
[33] Results and analysis available from the authors upon request.

mobility is potentially zero. The array of evidence strongly suggests a large causal effect of college graduation on mobility.

We conducted a number of tests to examine the robustness of the results in Table 4. First, we repeated the Table 4 analysis using alternative educational attainment measures: dummy variables for college attendance (measured as 13 or more years of schooling) and for college graduation (16 or more years of schooling). The results are reported in Appendix Table 1.[34] College graduation is associated with a higher probability of out of state residence among men in our sample than is college attendance. Consistent with this, our IV estimates of the causal impact of college completion are somewhat larger than our estimates of the causal impact of college attendance (although these are not statistically distinguishable due to the large standard errors). In general, our estimates tell the same story regardless of the measure of education we use.

We also estimated the 2SLS equations on data collapsed to birthstate-birthyear cells. This is the level of our variation, so it is natural to consider obtaining estimates using collapsed data. We present the results of this exercise in Appendix Table 2. The point estimates are nearly identical to those in Table 4 – particularly in Panel A – but the increased standard errors associated with a smaller sample size mean that only the estimates in columns (5) and (6) of Panel A retain their significance. We interpret the results in Appendix Table 2 to mean that our Table 4 point estimates are valid but that we simply lack the sample size to generate precise estimates with our two-instrument strategy in the collapsed data. Additionally, dropping national induction risk and replacing the trend with birth year dummies in the Table 4 specifications—which forces our identifying variation to come solely from cross-state differences in induction rates—does not qualitatively alter our conclusions, nor does using overall national cohort risk as defined in (i) instead of (iii) as our education instrument.

[34] Since the inclusion of veteran status does not affect the OLS coefficient on education in these specifications either, we only report one set of OLS results. 2SLS estimates with national risk only exclude veteran status from the equations. 2SLS estimates using both instruments include veteran status and treat it as endogenous.

Finally, in Table 5 we report the results of two placebo style analyses. These are intended to verify that the causal relationship we identify is in fact driven by the response of young men in our affected cohorts to the induction risk they faced. First, we re-estimate our reduced form specifications for men in our same sample of cohorts using data from the 1960 Census. At this point in their lives, these men had not yet been affected by the incentives for higher education associated with the Vietnam draft. If the cohorts who would eventually face the greatest draft risk were somehow also more likely to have migrated by our measure *before* the draft, this might mean our causal estimates in Table 4 are spurious. It would suggest high-risk cohorts were somehow different from low-risk cohorts in a way that might lead them to have both higher educational attainment and migration rates.

The results from this exercise are in Panel A of Table 5. In four of the six specifications we find no relationship between the draft risk a man will face in the future and his out of state residence in 1960. In two of the specifications, however, men who will face high national risk are also more likely to have moved out of their birth states prior to 1960. We believe this suggests estimates using national risk alone to identify years of schooling in our augmented specifications should be interpreted with caution. However, our preferred specifications and the basic specification using the single IV do not suffer from this concern.

In Panel B, we estimate our 2SLS models on a set of placebo cohorts. Unfortunately, there is no ideal placebo group on which to run such a test. We use men born between 1930 and 1941—the twelve cohorts preceding those in our sample—who were too old to be affected by induction orders during the Vietnam conflict although by virtue of their age these men were not subject to the same non-draft related cohort factors experienced by the men in our main sample. Nevertheless, the estimates in Panel B show no significant relationship between education and migration when we use our Vietnam draft risk measures to identify years of schooling. This is consistent with our claim that our instruments affect educational attainment only for our main sample of cohorts and that it is the higher education of these cohorts that leads to increased migration.

V. Mechanisms behind the Causal Relationship

Our results suggest a large causal role for higher education in geographic mobility. This causal effect may operate through either direct or indirect channels. On the one hand, higher education may *directly* confer skills or provide information which enhances the ability to undertake a long distance move. Individuals may develop general cognitive skills that facilitate the accumulation of information about alternative employment possibilities in other places. Moreover, individuals may actually gain information about new opportunities for migration either through attending college away from home or by interacting with peers who come from other places. In addition, individuals may gain some moving-specific skills because attending college may itself necessitate a move across state lines. A college education may also lower the psychic costs associated with migration by fostering openness to new experiences and awareness of national or global issues may lessen the difficulties of adjustment in a new place.

On the other hand, higher education may *indirectly* affect the likelihood of undertaking a long distance move. Research shows that increased schooling is causally related to higher wages (Card, 2001). If moving involves a fixed monetary cost, higher incomes may increase migration. Obtaining a college degree also alters the set of possible occupations that are available for recent graduates. Across the United States, the market for college graduates is often considered more geographically integrated than the market for lower-educated workers. If the skills associated with a college education are marketable in many regions, this may make it easier to consider moving out of state.[35] Finally, education may alter other non-economic characteristics such as marriage which can affect the likelihood of geographic mobility (Lefgren and McIntyre, 2006).

[35] See Bound et al. (2004).

We are able to provide credibly identified evidence on two of the mechanisms mentioned above: exposure to a geographically diverse set of peers and a move out of state to attend college. We obtain exogenous variation on both dimensions using data on public and private four-year college enrollment from the 1980 HEGIS survey (conducted by the National Center for Education Statistics). We created two separate rankings of states: the "out of state" ranking and the "diversity" ranking. The first ranks states according to the shares of their college-attending natives getting their degrees out of state, with quartile 1 having the lowest shares of its natives attending college out of state and quartile 4 the highest. If the experience of moving to go to college is an important channel behind the causal impacts we estimate, we might expect the college migration premium to be larger for natives of the highest quartile states. The "diversity" ranking divide states into quartiles according to the geographic diversity of their college populations measured using a Herfindahl index.[36] The idea in this case is that exposure to a geographically diverse set of fellow students might increase one's awareness of other labor markets or even improve one's network in distant locations, thereby increasing the migration premium to higher education in these states.

We use our sample of Vietnam era men from the 1980 Census to estimate versions of our main 2SLS specification in which years of schooling has been interacted with quartile indicators for each of the two state rankings. Since this specification now contains five endogenous variables—the four interactions with schooling plus veteran status—we instrument using interactions of the four quartiles with national risk plus (non-interacted) state risk.[37] The results, including OLS estimates, are shown in Table 6.

We find that for natives of states in the bottom quartile of out-of-state college going, the OLS migration premium to an extra year of schooling in our sample, of 0.019, is significantly lower

[36] The states in each set of quartiles are listed in Table 6. While there is some overlap between a state's quartile ranking on one dimension and its ranking on the other, there seem to be enough differences for the measures to capture different dimensions of the college experience.
[37] We find qualitatively similar results when estimate these effects using interactions of the four quartiles with national risk as well as interactions of the four quartiles with state-;level risk (an overidentified specification).

than that for natives in the higher quartiles who get returns of above 0.03. As was the case with the 2SLS estimates of the migration premium in earlier tables, the estimates in Table 6 have large standard errors. These prevent us from saying definitively whether the causal effect of higher education on migration is larger for natives of one quartile than another. The point estimates suggest, however, that the premium may be higher for natives of states in the middle of the out-of-state college going distribution. In the case of the geographic diversity of one's college peers, the OLS estimates show no consistent difference in the schooling migration premium across natives from states in the four quartiles of this distribution. Again, the 2SLS estimates are noisy, but they suggest that the causal impact of higher education on subsequent migration may be higher for natives of states in which the college system enrolls a more geographically diverse set of students.

From the Table 6 results we conclude that while there is some evidence that going to college out of state increased a man's likelihood of residing out of state later, this is only part of the story. The relationship between out-of-state college going propensity across states and subsequent migration of that state's college educated natives is far from linear. Moreover, our estimates using variation in the geographic diversity of a state's college population suggest that exposure to peers from other places may also play a role in subsequent mobility. While these conclusions are limited, we nevertheless believe they are a helpful step in understanding the nature of the relationship between higher education and migration.[38]

Lastly, we turn to alternative instrumental variables techniques to shed further light on the mechanisms behind the causal relationship we identify. As explained in the introduction, we feel our strategy is the most appropriate of several alternatives for identifying a "pure" effect of [higher] education on geographic mobility. However, because alternative identification strategies use policies that affected educational attainment in particular ways, we can compare the results from our main

[38] These results are consistent with what is known about the causal effect of attending college out of state (Groen 2004) or studying abroad (Parey and Waldinger 2009) on subsequent labor mobility choices. In our analysis, however, we attempt to distinguish between the two distinct but correlated channels of "going away" for higher education and exposure to a more geographically diverse set of peers.

specifications to those using alternative IV strategies in order to assess how different policies affect the migration-education relationship. If the manner in which educational attainment was increased matters for our results, this would suggest that the particular policy may have also affected one of the channels through which education has its effect on migration.

In Table 7 we present results from our main specification (or its nearest approximation, given data constraints) estimated using three leading alternative instruments for educational attainment: quarter of birth, new college openings, and state merit aid policies.[39] Details of the estimation are given in the table notes. Panel A shows the results of estimating our three main 2SLS specification using quarter of birth to instrument for years of schooling. We find no evidence of a causal impact of years of schooling on migration using this identification strategy. This has two important implications for our main results. First, since quarter of birth increases educational attainment at the high school margin and not the college margin, these results imply that it is years of college only that causally impact adult migration, not simply education at any level. Second, since education alone is not enough to increase mobility, these results downplay the importance of cognitive ability as a mechanism in the education-migration relationship.[40]

Panel B shows the 2SLS results for our main equations using the Currie and Moretti (2003) college openings instrument to identify years of schooling. In this case, the instrument is number of colleges in a man's state of birth in the year he turned 17, divided by his state cohort size at age 17. Higher educational attainment is identified using variation in the per capita number of colleges available to 17 year olds across states and over time. The results show a large, negative impact of an additional year of schooling on migration. This is in striking contrast to our main results and to the OLS estimate. We believe the difference is due to the way in which the instrument affects

[39] These were originally exploited for identification purposes in Angrist and Krueger (1991); Currie and Moretti (2003); and Dynarski (2008).

[40] The estimates in Panel A also serve as a valuable specification check for those concerned that some unknown mechanical relationship leads to estimates of overly large causal impacts when we instrument for education in a migration equation. On the other hand, there are important concerns about the consistency of 2SLS estimates obtained using the quarter of birth instrument, particularly in a setting such as ours with very limited controls for family background (Buckles and Hungerman, 2008).

educational attainment. The college openings instrument operates by encouraging marginal college attenders to go to a college nearer their home and specifically in their state of birth. If going away to college is an important channel through which higher education increases mobility (as the results in Table 6 suggest), then an instrument that effectively shuts down that channel should be expected to produce different estimates of the education premium on migration. Moreover, if college graduates require labor markets with more high skilled jobs, and the presence of colleges creates such jobs, then additional in-state colleges may further decrease a man's propensity to leave his birth state after completing a college education. (See Moretti 2004a and 2004b for related discussions.)

Finally, Panel C shows results from estimating our main specifications using two state merit aid policies to identify higher education in the 2000 Census. Following Dynarski (2008), we use a dummy variable for college completion as our education measure and estimate the equations for a younger sample of cohorts who were eligible to receive assistance through these policies. These state policies increase college completion by lowering college costs through tuition scholarships to any student who maintains a threshold grade point average in both high school and college. Large percentages of high school seniors in both states with these programs qualify for this assistance. As was the case with our main estimates, we find large, positive OLS and 2SLS premiums to college completion on subsequent migration.

The implications of these results are less straightforward than with the instruments in Panels A and B because of the fact that the merit aid instrument operates on two margins. First, by lowering the relative cost of in-state colleges, it likely increased in-state college attendance by students who would otherwise have gone out of state to college. This would tend to reduce the migration premium estimated via 2SLS in a manner similar to that of the college openings instrument. On the other hand, lowering the relative costs of college (without increasing the per capita number of colleges) encourages college attendance overall among a state's natives although the biggest effect is probably on in-state college attendance. We conclude that in the case of the merit aid instrument, the positive effect of increased college attainment swamps the negative effect of in-state college attendance on subsequent migration. This is consistent with the results of Table 6, which show that while out of state college going may be an important channel through which higher education increases mobility, it is unlikely to be the only channel. This is also consistent with the possibility that college openings in a state discourage later outmigration of college educated natives by providing attractive local labor markets for high skill workers.

VI. Conclusion

In this paper, we examined the causal role of college graduation in determining geographic mobility, in particular, the likelihood of lasting, long-distance moves indicated by a man's residence outside his birthstate in early- to mid-adulthood. We show that selection is unlikely to explain the large increases in migration rates associated with college completion.

Using a 5% sample from the 1980 U.S. Census, we provide estimates of the causal impact of college education on the probability of a long-distance move. We use state-cohort level variation in college completion arising from draft avoidance behavior among men at risk for conscription into the Armed Forces during the Vietnam conflict as a source of exogenous variation in the probability that a man completed college. We show that this variation increased migration rates substantially among affected cohorts. It also significantly increased the probability that a young man completed some post-secondary education. We then use induction risk to identify the causal effect of higher education on subsequent mobility. We find that college education increases the probability of a long-distance move for the marginal college graduate significantly. Moreover, our instrumental variables strategy enables us to purge our estimates of potential bias due to correlation of unobservables with a relevant control variable other than education: veteran status. Our preferred estimates use two instruments to account for the potential endogeneity of both college completion and veteran status among our sample of Vietnam era young men.

The use of two instruments to identify separate variables in our equations of interest results in large standard errors. Our 2SLS estimates of the causal impact of college are significant at conventional levels, but we cannot say with certainty that the causal effect is larger or smaller than the OLS estimate. However, our estimates strongly suggest that the causal impact is economically significant. The average man in our dataset moves out of his birthstate with probability 0.33, and simple OLS estimates put the premium to an additional year of higher education on out-of-birthstate migration at 0.3 while the highest estimate from our preferred 2SLS specification puts this premium at 0.7.

References

Adams, Scott. 2002. "Educational Attainment and Health: Evidence from a Sample of Older Adults." *Education Economics*. 10(1): 97-109.

Angrist, Joshua and Imbens, Guido. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*. 62(1994): 467-475.

Angrist, Joshua; Guido Imbens; and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association: Applications*.

Angrist, Joshua and Kreuger, Alan. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106(1991): 979-1014.

Angrist, Joshua and Kreuger, Alan. "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." *NBER Working Paper #4067* (1992).

Bertrand, Marianne; Esther Duflo; and Sendhil Mullainathan. 2004. "How Much Should We Trust Difference-in-Differences Estimates?" *Quarterly Journal of Economics*. 119(1): 249-275.

Black, Dan A.; Mark C. Berger; and Frank A Scott. "Bounding parameter estimates with nonclassical measurement error." *Journal of the American Statistical Association*. Sep 2000. Vol. 95, Iss. 451; p. 739-738.

Black, Dan; Seth Sanders; and Lowell Taylor. "Measurement of higher education in the census and current population survey." *Journal of the American Statistical Association*. Sep 2003. Vol. 98, Iss. 463; p. 545-554.

Borghans, Lex; Angela L. Duckworth; James J. Heckman; and Bas Ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources*. 43(4): 972-1059.

Borjas, George. 2006. "Native Internal Migration and the Labor Market Impact of Immigration." *Journal of Human Resources*. 41(2): 221-258.

Bound, John; Groen, Jeffrey; Kezdi, Gabor; and Turner, Sarah. "Trade in University Training: Cross-state Variation in the Production and Use of College-Educated Labor." *Journal of Econometrics*. 121(2004).

Bound, John and Holzer, Harry. "Demand Shifts, Population Adjustments and Labor Market Outcomes during the 1980s." *Industrial and Labor Relations Review*. 18(2000): 20-54.

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.

Bowles, Samuel. "Migration as Investment: Empirical Tests of the Human Investment Approach to Geographic Mobility." *The Review of Economics and Statistics* 52(1970): 356-362.

Bowles, Samuel; Herbert Gintis; and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*. 39(4): 1137-1176.

Buckles, Kasey and Daniel Hungerman. 2008. “Season of Birth and Later Outcomes: Old Questions, New Answers.” *NBER Working Paper #14573*.

Cameron, C., Gelbach, J., and D. Miller (2006) “Robust inference with Multi-Way Clustering” *NBER Technical Working Paper* #T0327.

Card, David. “Earnings, Schooling, and Ability Revisited.” *Research in Labor Economics*. 14(1995): 23-48.

Card, David. 1999. “The Causal Effect of Education on Earnings.” *Handbook of Labor Economics*. Vol 3. Ed. Orley Ashenfelter and David Card. Elsevier Science B.V.

Card, David. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica*. 69(2001): 1127-1160.

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.

Card, David and Thomas Lemieux. 2001. “Dropout and Enrollment Trends in the Post-war Period: What Went Wrong in the 1970s?” ed. John Gruber. *Risky Behavior Among Youth*. Cambridge, MA.

Cascio, Elizabeth and Lewis, Ethan. “Schooling and the Armed Forces Qualifying Test.” *Journal of Human Resources* 41(2006): 294-318.

Currie, Janet and Enrico Moretti. 2003. “Mother’s Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings.” 118(4): 1495-1532.

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.

Dougherty, Conor. “U.S. Migration Falls Sharply.” *The Wall Street Journal*. Page A3. March 19, 2009.

Dynarski, Susan. 2008. “Building the Stock of College Educated Labor.” *Journal of Human Resources*. 43(3): 576-610.

Evers, Alf. *Selective Service: A Guide to the Draft*. New York: Lippincott Company (1961).

Ferrie, Joseph. 2005. “The End of American Exceptionalism? Mobility in the U.S. Since 1850.” *Journal of Economic Perspectives*. 19(Summer): 199-215.

Greenwood, M.J. (1975) "Research on Internal Migration in the United States: A Survey." *Journal of Economic Literature*. 13: 397-433.

Greenwood, M.J. (1997) "Internal Migration in Developed Countries." *Handbook of Population and Family Economics*. Mark R. Rosenzweig and Oded Stark, eds. New York: Elsevier Science

Grimard, F. and D. Parent (2007) "Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking?" in *Journal of Health Economics*, Vol. 26. No. 5: pp. 896-926.

Groen, J. 2004. "The Effect of College Location on Migration of College-educated Labor." *Journal of Econometrics*. 121(1-2): 125-142.

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

Kane, Thomas; Cecilia Rouse and Douglas Staiger. 1999. "Estimating Returns to Schooling When Schooling is Misreported." *Industrial Relations Section Working Paper #419*. Princeton University.

Kuziemko, Ilyana. "Dodging Up to College or Dodging Down to Jail: Behavioral Responses to the Vietnam Draft by Race and Class." *Manuscript*, Princeton University (2008).

Ladinsky, J. (1967) "The Geographic Mobility of Professional and Technical Manpower." in *Journal of Human Resources*: Vol. 2, No. 4, pp. 475-494

Lefgren, Lars and McIntyre, Frank. "Examining the Relationship between Women's Education and Marriage Outcomes." *Journal of Labor Economics* 24(2006): 787-730.

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the U.S." 72(1): xx-xx.

Jayachandran, Seema and Adriana Lleras-Muney. 2009. "Longevity and Human Capital Investments: Evidence from Maternal Mortality Declines in Sri Lanka." *Quarterly Journal of Economics*. 124(1): 349-397.

Johnson, William R. and Neal, Derek. "The Role of Pre-Market Factors in Black-White Wage Differences." *Journal of Political Economy*. 104(1996):869-895.

MacInnis, Bo (2006) "The Long-Term Effects of College Education on Morbidities: New Evidence From the Pre-Lottery Vietnam Draft," *mimeo*, University of Michigan.

Moretti, Enrico. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-sectional Data." *Journal of Econometrics*. 121(2004): 175-212.

Moretti, Enrico. 2004. "Human Capital Externalities in Cities." *Handbook of Urban and Regional Economics*. North Holland-Elsevier.

Nelson, Richard R. and Phelps, Edmund S. "Investment in Humans, Technological Diffusion and Economic Growth." *American Economic Review* 56(1966): 69-75.

Parey, Matthias and Fabian Waldinger. 2009. "Studying Abroad and the Effect on International Labor Market Mobility: Evidence from the Introduction of Erasmus." *Manuscript*. London School of Economics.

Pingle, Johnathan. 2007. "A note on measuring internal migration in the United States." 94(1): 38-42.

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.

Saks, Raven and Wozniak, Abigail. 2009. "Labor Reallocation over the Business Cycle: New Evidence from Internal Migration." *Manuscript*, University of Notre Dame.

Sanders, Jacquin. *The Draft and the Vietnam War*. New York: Walker and Company (1966).

Schultz, T.W. "Investment in Human Capital" in *American Economics Review* 51(1961), pp. 1-17.

Schultz, T.W. "The Value of the Ability to Deal with Disequilibria." *Journal of Economic Literature* 13(1975): 827-846.

Segal, Carmit. "Motivation, Test Scores and Economic Success." Universitat Pompeu Fabra Working Paper #1124 (2008).

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).

VerBeek, Marno. *A Guide to Modern Econometrics*. Chichester, UK: John Wiley and Sons Ltd. (2000).

Welch, Finis. "Education in Production," *Journal of Political Economy* 78(1970), 35-59

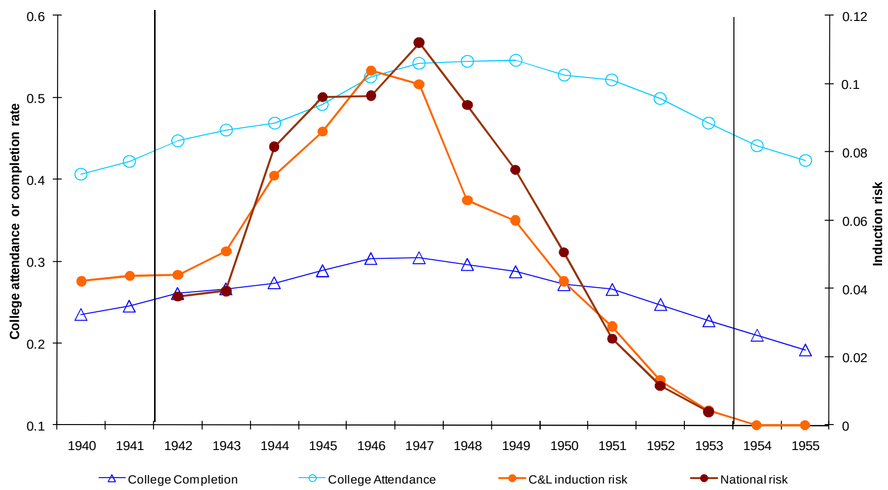
Wozniak, Abigail. Forthcoming. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*.

Wooldridge, Jeffrey. *Econometric Analysis of Cross-section and Panel Data*. MIT Press: Cambridge, Massachusetts (2002).

________. 2003. "Cluster-Sample Methods in Applied Econometrics." *American Economic Review*. 93(2):122-138.

Ziliak, James; Beth Wilson; and Joseph Stone. "Spatial Dynamics and Heterogeneity in the Cyclicality of Real Wages." *Review of Economics and Statistics* , May 1999, 81(2): 227-236.

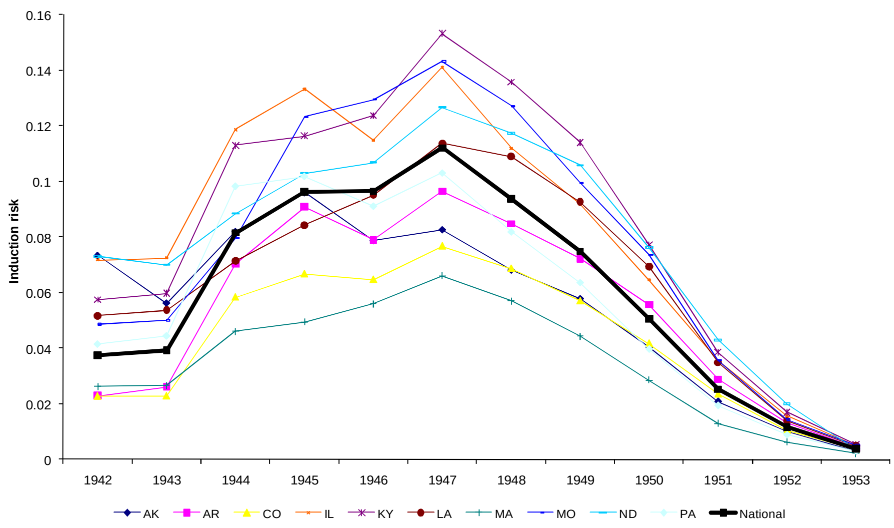
Panel A: Induction risk and college-going



Panel B: Induction risk and residual schooling and mobility

Figure 1. Panel A plots shares of each birth year cohort with 1+ and 4+ years of post-secondary schooling (left axis) and induction risk defined in Equations (i) and (iii) (right axis). Panel B plots birth year cohort's average years of schooling and migration rate adjusted for state of birth fixed effects, birth year trend and national risk as defined in (iii) (left axis) and national risk (right axis).

Panel A: Induction risk



Panel B: Residual induction risk

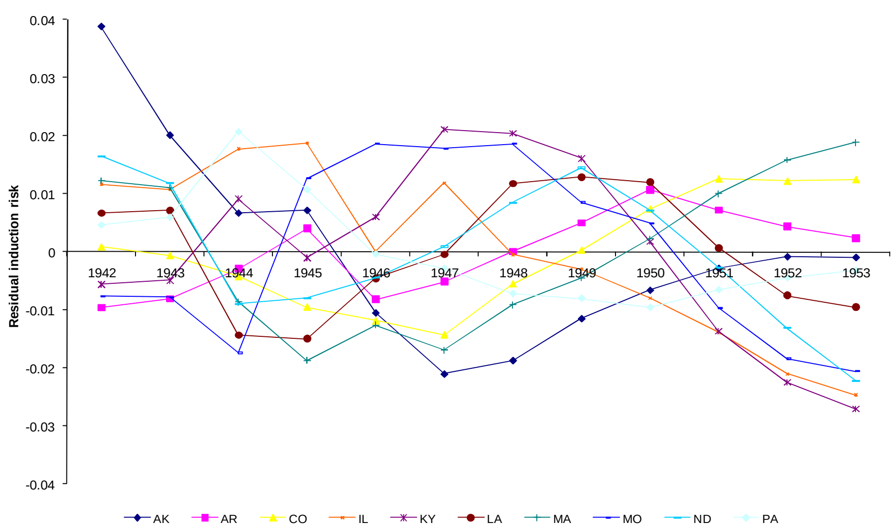


Figure 2. Birth state – birth year variation in induction risk. Panel A plots state risk as defined in (ii). Panel B plots (ii) adjusted for birth state fixed effects, birth year trend, and national risk as defined in (ii).

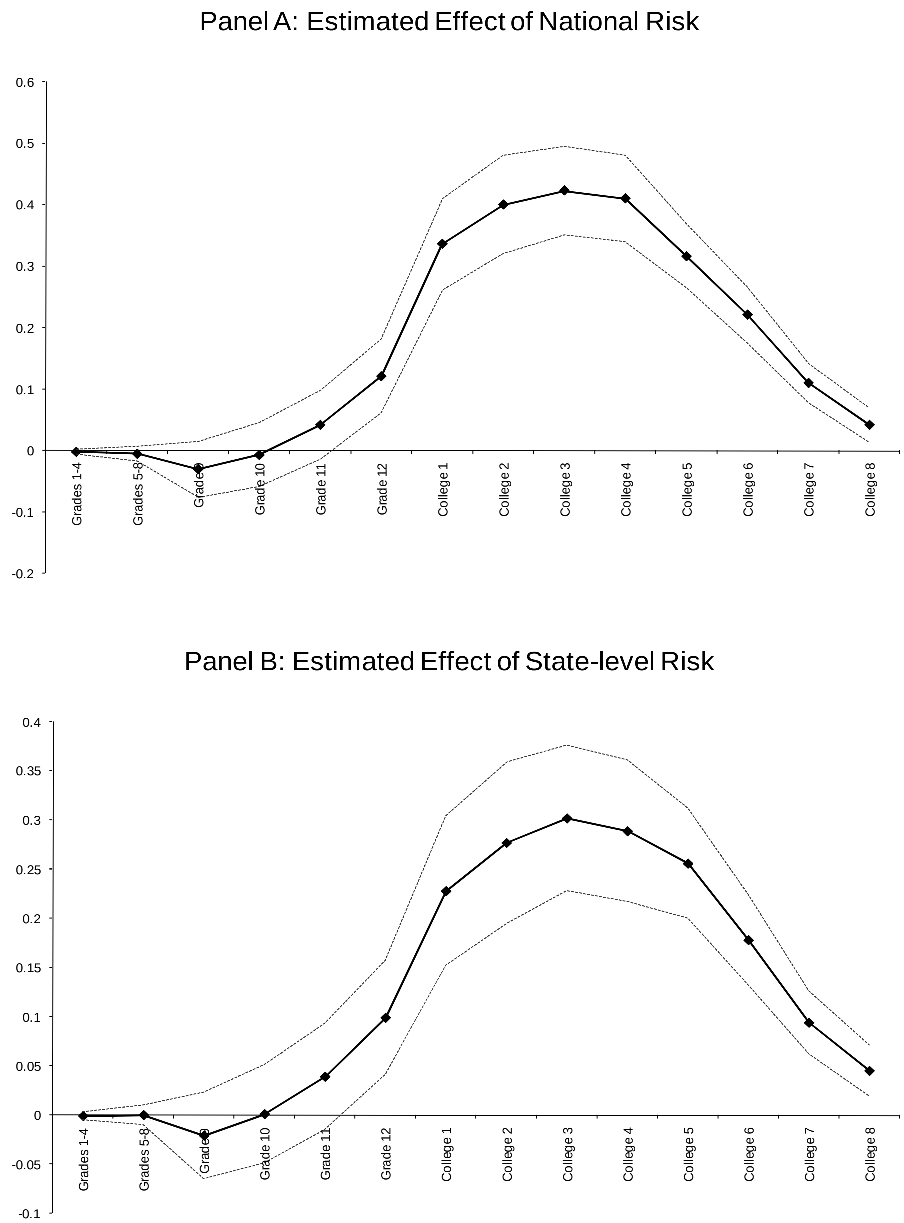


Figure 3. Coefficients and standard errors from OLS regressions of dummy variables for completed education of x-axis grade and higher, inclusive, on national risk and state risk as defined in Equations (iii) and (ii), respectively.

Table 1: Summary Statistics for 1980 Census

	Mean	Standard deviation	Min	Max	Observations
Living outside of birth state	0.358	0.479	0	1	794,331
Distance moved ('000 miles)	0.300	0.593	0	4.933	794,331
Years of schooling	13.425	2.952	0	20	794,331
College graduation	0.281	0.450	0	1	794,331
College attendance	0.428	0.495	0	1	794,331
Veteran status	0.329	0.470	0	1	794,331
National induction risk	0.059	0.036	0.004	0.112	794,331
State-level induction risk	0.060	0.039	0.002	0.153	794,331
Age	31.412	3.430	26	38	794,331
Black	0.089	0.285	0	1	794,331
Other nonwhite	0.045	0.208	0	1	794,331
Employment/population (birth-year level)	327.489	20.462	294.579	352.073	794,331
Employment/population (birth-year-state)	323.285	43.010	185.279	415.559	794,331
Log cohort size (birth-year level)	10.298	0.237	9.835	10.518	794,331
Log cohort size (birth-year-state level)	6.760	0.821	2.398	8.039	794,331

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Cohort size is derived from the 1960 Census. Mobility is equal to 1 if respondent resides outside his state of birth at the time of the Census, and 0 otherwise. The employment-population ratio is defined for the year in which individuals are 19 years of age. Log cohort size is defined as the logarithm of the number of respondents to the 1960 Census in the corresponding year of birth.

Table 2: 1st-Stage Estimates of Induction Risk

Panel A: National and state-level induction risk, entered *separately*

dependent variable:	*years of schooling*			*years of schooling*		
	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	4.441*** [0.512]	3.748*** [0.394]	3.707*** [0.392]			
State-level induction risk				4.031*** [0.208]	3.396*** [0.208]	3.375*** [0.167]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
F-test on IV	75.27	90.32	89.51	377.26	267.30	408.30

Panel B: National and state-level induction risk, entered *together*

dependent variable:	*years of schooling*			*veteran status*		
	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	5.420*** [0.785]	6.049*** [0.651]	5.237*** [0.615]	1.234*** [0.151]	1.256*** [0.136]	1.332*** [0.116]
State-level induction risk	-0.978 [0.784]	-2.198*** [0.649]	-1.459** [0.593]	0.890*** [0.148]	0.733*** [0.139]	0.689*** [0.117]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
Joint F-test on IVs	233.00	217.82	260.30	1251.49	1229.23	1651.41

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of observations is 794,331 in each regression. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6.

Table 3: Reduced-Form Estimates of Induction Risk on Migration

dependent variable: living outside state of birth

	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	0.084*	0.139***	0.138***	0.238***	0.226***	0.230***
	[0.043]	[0.031]	[0.031]	[0.088]	[0.086]	[0.088]
State-level induction risk				-0.155*	-0.118	-0.125
				[0.083]	[0.083]	[0.086]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.02	0.02	0.02	0.02	0.02	0.02
Mean of dep. variable	0.358	0.358	0.358	0.358	0.358	0.358

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of observations is 794,331 in each regression. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6.

Table 4: OLS and IV Estimates for the Impact of Years of Schooling on Migration

dependent variable: living outside state of birth

Panel A: IV with national-level induction risk

	OLS			*2SLS*		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.029***	0.029***	0.029***	0.019	0.037***	0.037***
	[0.001]	[0.001]	[0.001]	[0.011]	[0.008]	[0.009]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.05	0.05	0.05			
Mean of dep. variable	0.358	0.358	0.358	0.358	0.358	0.358

Panel B: IV with national and state-level induction risk

	OLS			*2SLS*		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.029***	0.029***	0.029***	0.067**	0.044*	0.059*
	[0.001]	[0.001]	[0.001]	[0.033]	[0.024]	[0.032]
Veteran Status	0.040***	0.041***	0.041***	-0.101	-0.03	-0.057
	[0.002]	[0.002]	[0.002]	[0.069]	[0.048]	[0.062]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.06	0.06	0.06			
Mean of dep. variable	0.358	0.358	0.358	0.358	0.358	0.358

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of observations is 794,331 in each regression. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Columns 4, 5, and 6 of Panel A use national-level induction risk to instrument for years of schooling. Columns 4, 5, and 6 of Panel B use national and state-level induction risk to instrument for schooling and veteran status.

Table 5: Placebo Regressions

dependent variable: living outside state of birth

Panel A: Reduced-Form Estimates of Induction Risk on Childhood Migration (1960 Census)

	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	0.024	0.078***	0.076***	0.027	0.053	-0.003
	[0.026]	[0.010]	[0.009]	[0.109]	[0.111]	[0.114]
State-level induction risk				-0.003	-0.003	0.049
				[0.105]	[0.107]	[0.108]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.02	0.02	0.02	0.02	0.02	0.02
Mean of dep. variable	0.163	0.163	0.163	0.163	0.163	0.163

Panel B: IV Estimates for Schooling on Migration for older cohorts born 1930-1941(1980 Census)

	2SLS with national risk			2SLS with national & state-level risk		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.121	-0.008	-0.026	0.016	-0.871	-0.069
	[0.947]	[0.183]	[0.148]	[0.132]	[7.094]	[0.255]
Veteran Status				-0.02	-0.662	0.023
				[0.260]	[5.646]	[0.215]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
Mean of dep. variable	0.396	0.396	0.396	0.396	0.396	0.396

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data in Panel A is from the 1% sample of the 1960 U.S Census; data in Panel B is from the 5% sample of the 1980 U.S. Census. Specifications in Panel A are restricted to men born between 1942 and 1953; specifications in Panel B are restricted to men born between 1930 and 1941. Number of observations is 167,475 in each regression of Panel A; 531,138 in each regression of Panel B. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census for both panels and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Columns 1, 2, and 3 of Panel B use national-level induction risk to instrument for years of schooling while columns 4, 5, and 6 use national and state-level induction risk to instrument for schooling and veteran status.

Table 6: Potential Mechanisms

dependent variable: living outside state of birth	Out-of-State College Going (by quartile)		Geographic Diversity of State Colleges (by quartile of Herfindahl Index)	
	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)
Years of schooling*Quartile 1	0.019***	0.067	0.029***	0.05
	[0.001]	[0.057]	[0.001]	[0.041]
Years of schooling*Quartile 2	0.032***	0.131**	0.033***	0.047
	[0.001]	[0.053]	[0.001]	[0.036]
Years of schooling*Quartile 3	0.034***	0.084**	0.025***	0.145***
	[0.001]	[0.042]	[0.001]	[0.049]
Years of schooling*Quartile 4	0.037***	0.015	0.027***	0.060*
	[0.001]	[0.044]	[0.001]	[0.033]
Veteran Status	0.043***	-0.108	0.041***	-0.069
	[0.002]	[0.089]	[0.002]	[0.071]
Trend	X	X	X	X
Demographic controls	X	X	X	X
Cohort size	X	X	X	X
Employment rate	X	X	X	X
Birth region trends	X	X	X	X
R^2	0.06		0.06	
Mean of dep. variable	0.358	0.358	0.358	0.358

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of observations is 794,331 in each regression. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Column 2 uses national-level induction risk interacted with quartile and state-level induction risk to instrument for years of schooling interacted with quartiles and for veteran status.

Out-of-state college going quartiles: Quartile 1 (AL, AZ, CA, LA, MI, MS, NC, OK, SC, TX, UT, WI); Quartile 2 (AR, FL, IL, IN, IA, KS, KY, MO, NE, OH, OR, TN, WV); Quartile 3 (CO, GA, HI, ID, MA, MN, MT, NY, ND, PA, RI, WA, WY); Quartile 4 (AK, CT, DE, DC, ME, MD, NV, NH, NJ, NM, SD, VT, VA). Herfindahl index quartiles: Quartile 1 (AL, CA, IL, MA, MI, NY, NC, OH, PA, TN, TX, VA); Quartile 2 (CO, CT, FL, GA, IN, KS, MN, MO, NH, NJ, SC, WA, WI); Quartile 3 (AZ, AR, IA, KY, LA, ME, MD, MS, OK, SD, UT, VT, WV); Quartile 4 (AK, DE, DC, HI, ID, MT, NE, NV, NM, ND, OR, RI, WY)

Table 7: Alternative IV Strategies

dependent variable: living outside state of birth				
Panel A: Quarter of Birth (cohorts born 1942-1953)				
		2SLS		*OLS*
	(1)	(2)	(3)	(4)
Years of schooling	-0.052	-0.051	-0.051	0.029***
	[0.054]	[0.051]	[0.052]	[0.001]
Trend	X	X	X	X
Demographic controls	X	X	X	X
Cohort size		X	X	X
Employment rate		X	X	X
Birth region trends			X	X
Mean of dep. variable	0.358	0.358	0.358	0.358
Panel B: College Openings (cohorts born 1942-1953)				
		2SLS		*OLS*
	(1)	(2)	(3)	(4)
Years of schooling	-0.045**	-0.045**	-0.062**	0.029***
	[0.018]	[0.018]	[0.027]	[0.001]
Birth-year dummies	X	X	X	X
Demographic controls	X	X	X	X
Cohort size		X	X	X
Employment rate		X	X	X
Birth region trends			X	X
Mean of dep. variable	0.358	0.358	0.358	0.358
Panel C: State Merit Aid Policies (cohorts born 1966-1978)				
		2SLS		*OLS*
	(1)	(2)	(3)	(4)
College degree attainment	0.800**	0.629**	0.523*	0.117***
	[0.332]	[0.274]	[0.356]	[0.004]
Birth-year dummies	X	X	X	X
Demographic controls	X	X	X	X
Cohort size		X	X	X
Employment rate		X	X	X
Birth region trends			X	X
Mean of dep. variable	0.333	0.333	0.333	0.333

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data in Panels A and B are from the 5% sample of the 1980 Census; data in Panel C is from the 1% sample of 2000 Census. Number of observations in Panels A, B, and C are 798,508, 792,169 and 349,578 respectively. Columns 1, 2, and 3 of Panel A use indicators for quarter of birth to instrument for years of schooling; columns 1, 2, and 3 of Panel B use number of 2 and 4 year colleges per capita to instrument for years of schooling; columns 1, 2, and 3 of Panel C use an indicator for Arkansas after 1991 and Georgia after 1993 to instrument for years of schooling. The standard set of controls is defined as in previous tables.

Appendix Table 1: OLS and 2SLS Estimates for the Impact of College Attendance and Graduation on Migration

dependent variable: living outside state of birth

Panel A: College Attendance

	OLS			2SLS with national risk			2SLS with national and state-level risk		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Attendance	0.152***	0.153***	0.153***	0.096	0.190***	0.191***	0.268**	0.203*	0.297*
	[0.008]	[0.008]	[0.008]	[0.055]	[0.044]	[0.045]	[0.126]	[0.109]	[0.162]
Trend	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X
Cohort size		X	X		X	X		X	X
Employment rate		X	X		X	X		X	X
Birth region trends			X			X			X
R^2	0.03	0.05	0.05						
Mean of dep. var.	0.428	0.428	0.428	0.428	0.428	0.428	0.428	0.428	0.428

Panel B: College Graduation

	OLS			2SLS with national risk			2SLS with national and state-level risk		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Graduation	0.166***	0.166***	0.166***	0.132*	0.242***	0.244***	0.304**	0.237*	0.319*
	[0.007]	[0.006]	[0.006]	[0.072]	[0.054]	[0.055]	[0.135]	[0.124]	[0.166]
Trend	X	X	X	X	X	X	X	X	X
Race & Birth state	X	X	X	X	X	X	X	X	X
Cohort size		X	X		X	X		X	X
Employment rate		X	X		X	X		X	X
Birth region trends			X			X			X
R^2	0.03	0.05	0.05						
Mean of dep. var.	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of observations is 794,331 in each regression. Trend is a linear trend in age. Demographic controls include race (white, black, other) and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1-6 and birth year-state in columns 7-9. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1-6 and birth year-state in columns 7-9.

Appendix Table 2: OLS and IV Estimates for the Impact of Years of Schooling on Migration (Collapsed to the birth-state level)

dependent variable: living outside state of birth

Panel A: IV with national-level induction risk

	OLS			*2SLS*		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.027*** [0.006]	0.039*** [0.005]	0.045*** [0.006]	0.021 [0.017]	0.037** [0.012]	0.037** [0.012]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.92	0.92	0.93			
Mean of dep. variable	0.408	0.408	0.408	0.408	0.408	0.408

Panel B: IV with national and state-level induction risk

	OLS			*2SLS*		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.035*** [0.010]	0.040*** [0.011]	0.042*** [0.012]	0.073 [0.051]	0.053 [0.042]	0.077 [0.056]
Veteran Status	-0.037 [0.030]	-0.041 [0.029]	-0.039 [0.031]	-0.111 [0.105]	-0.063 [0.087]	-0.111 [0.113]
Trend	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
Cohort size		X	X		X	X
Employment rate		X	X		X	X
Birth region trends			X			X
R^2	0.92	0.92	0.92			
Mean of dep. variable	0.408	0.408	0.408	0.408	0.408	0.408

Notes: Robust standard errors in brackets, clustered by birth-year in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953. Number of state-year observations is 600 in each regression. Trend is a linear trend in age. Demographic controls include the percentage of whites, blacks, and others in each state-year and birth state fixed effects. Cohort size is derived from the 1960 Census and defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Employment rate is the employment to population ratio, also defined at the birth-year level in columns 1, 2, and 3 and birth year-state in columns 4, 5, and 6. Columns 4, 5, and 6 of Panel A use national-level induction risk to instrument for years of schooling. Columns 4, 5, and 6 of Panel B use national and state-level induction risk to instrument for schooling and veteran status.