



IZA DP No. 3432

The Impact of College Graduation on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk

Ofer Malamud
Abigail Wozniak

April 2008

The Impact of College Graduation on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk

Ofer Malamud
University of Chicago

Abigail Wozniak
*University of Notre Dame
and IZA*

Discussion Paper No. 3432
April 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Impact of College Graduation on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk^{*}

College-educated workers are twice as likely as high school graduates to make lasting long-distance moves, but little is known about the role of college itself in determining geographic mobility. Unobservable characteristics related to selection into college might also drive the relationship between college education and geographic mobility. We explore this question using a number of methods to analyze both the 1980 Census and longitudinal sources. We conclude that the causal impact of college completion on subsequent mobility is large. We introduce new instrumental variables that allow us to identify educational attainment and veteran status separately in a sample of men whose college decisions were exogenously influenced by their draft risk during the Vietnam War. Our preferred IV estimates imply that graduation increases the probability that a man resides outside his birth state by approximately 35 percentage points, a magnitude nearly twice as large as the OLS migration differential between college and high school graduates. IV estimates of graduation's impact on total distance moved are even larger, with IV estimates that exceed OLS considerably. We provide evidence from the National Longitudinal Survey of Youth (NLSY) 1979 that our large IV estimates are plausible and likely explained by heterogeneous treatment effects. Finally, we provide some suggestive evidence on the mechanisms driving the relationship between college completion and mobility.

JEL Classification: J61, J24, I23

Keywords: geographic mobility, education, college graduates, internal migration, instrumental variables

Corresponding author:

Abigail Wozniak
Department of Economics and Econometrics
University of Notre Dame
441 Flanner Hall
Notre Dame, Indiana 46556
USA
E-mail: a_wozniak@nd.edu

^{*} The authors would like to thank Lawrence Katz, Caroline Hoxby, and Laddie Sula for helpful conversations and comments. Comments from seminar participants at the University of Notre Dame, the Midwest Economics Association Meetings, University of Illinois-Urbana Champaign, Dartmouth College, University of Illinois-Chicago, and Indiana University-Purdue University Indianapolis improved the paper. All errors are our own.

Decades of empirical evidence, as well as simple casual observation, reveal that college-educated workers are more likely than others to make lasting, long-distance moves.¹ Recent papers have confirmed that more educated workers are better able to undertake migration that benefits them economically (Wozniak, 2006; Bound and Holzer, 2000). However in contrast to the burgeoning literature establishing the causal effect of college on wages, little is known about the role of college in determining geographic mobility. Does college itself provide skills which help individuals undertake long-distance moves? Are the skills learned in college more transferable to other locales? Or do individuals who pursue a college education also have a higher propensity to migrate due to unobserved characteristics? In this paper we distinguish these hypotheses—causality versus selection—to better understand the 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, 2001) 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 cohorts to obtain separate instruments for education and veteran status. 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. We provide evidence that state cohort risk was idiosyncratic and unknown *ex ante*, yet strongly predictive of subsequent veteran status. Enrollment in college to avoid the draft was primarily a response to perceived *national* cohort risk, as has been documented by Card and Lemieux.

¹ See Ladinsky (1967), and Greenwood (1975) for early evidence. Wozniak (2006) and Greenwood (1997) provide more recent evidence. The mobility advantage of more educated workers is stable across various types of long-distance migration, i.e. moves between cities or farther.

Given that enrollment responded to national level variation while veteran status was affected by state level variation, we use own state induction risk to instrument for veteran status and average induction risk for young men in all other states to instrument for education. To our knowledge, this is the first use of this variation to solve the problem of endogenous education and veteran status that is common in studies that use changes in military policies as a means of identifying educational attainment.²

Using micro-data from 5% samples of the 1980 U.S. Census of Population, we provide both OLS and 2SLS estimates for the effect of college education on geographic mobility. In cross-sectional OLS regressions, we observe that graduating from college is associated with an increase in an individual's probability of residing outside his birth state at mid-career by fifteen percentage points, or approximately one-half the mean of this variable. Calculating the distance moved based on the distance between state capitals indicates that graduating from college is associated with moving approximately 125 miles further. When we instrument for college with cohort-level induction risk, the causal effect of graduation on geographic mobility is substantially larger at about forty percentage points, while the effect on distance moved is about 700 miles. Our preferred estimation strategy uses national and state-level induction risk to instrument for both graduation and veteran status in 2SLS. This yields slightly smaller coefficient estimates to those in which only education is instrumented, indicating an effect of college on interstate migration of thirty five percentage points.

We then turn to the National Longitudinal Survey of Youth (NLSY) 1979 to assess the plausibility of these estimates. Our finding that 2SLS estimates of the mobility impact of a college education equal or exceed the OLS estimates implies a specific pattern in unobserved characteristics

² 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. In a paper examining the effect of college education on health, morbidities, MacInnis (2006) explores some alternative specifications using the original cohort-level variation induction risk among Vietnam veterans.

and selection, both of which are testable using the more detailed information in the NLSY. First, we examine whether characteristics omitted from our regressions using Census data—which ex ante we might expect to bias the OLS estimates upwards—are in fact unimportant for predicting migration. Next we estimate flexible OLS specifications to identify groups in which the impact of college on migration choices might exceed the average return, which suggests that heterogeneous treatment effects might drive the observed pattern of IV and OLS estimates. We find evidence that heterogeneous treatment effects are a plausible explanation for this pattern.

The notion that education can provide valuable skills which make individuals more productive in the workplace is well-established in the economics literature.³ However, there is less evidence supporting the notion that education may confer an ability to respond to economic disequilibria.⁴ Examining the causal effect of education on migration – the response to economic incentives arising from spatial disequilibria across segmented labor markets – helps fill this gap in the literature on human capital.⁵ This is important because a causal link between education and migration decisions means that geographic mobility constitutes a “return” to higher education—one that can potentially add much to our understanding of what education confers. Our question essentially expands the notion of human capital acquired through schooling to include skills and attributes other than cognitive ability and of returns to human capital beyond the traditional focus on increases in marginal product.

Our research also entails a number of implications for policymakers. First, states’ abilities to benefit from investments in their higher education systems hinge critically on the migration potential of the educated workers these systems produce. If college education increases mobility among those

³ See Card (2001) for causal estimates for the effect of education on wages. Johnson and Neal (1996) and Cascio and Lewis (2006) also provide direct evidence that education can improve skills, as measured by the Armed Forces Qualifying Test (AFQT).

⁴ Nelson and Phelps (1966) postulate a higher return to education in the presence of technological change and Welch (1970) provides some related evidence.

⁵ See Bowles (1970) for this characterization of geographic mobility. Schultz (1975) offers a survey of other studies examining the ability to deal with disequilibria.

who receive it, then the labor force composition effects of more generous state tuition policies (Dynarski, 2005) and the externalities associated with new college construction (Moretti, 2004) may disappoint policymakers who believe their targeted college graduates will remain in state. Second, if geographic adaptability is a product of college education, policymakers may want to reconsider re-training programs for displaced workers that focus on rapid acquisition of specific skills in schools close to where these workers already live. Such programs are unlikely to confer the flexibility and adaptability that are associated, and are perhaps the result of, a typical college education. Finally, the question of what precisely the college experience confers is central to the debate surrounding requirements for colleges that offer distance learning courses. The nature of college may be about to change as a result of online educational opportunities, and it is critical that policymakers know how these changes will affect the next generation of college graduates.

The remainder of the paper proceeds as follows. Section II provides further background on the empirical relationship between education and geographic mobility and the potential mechanisms which might bring about this relationship. Section III explains the empirical strategy used to identify the causal effect of education on mobility and describes the data. Section IV presents our main findings from the empirical analysis. Section V assesses the selection problem using the NLSY79, and Section VI concludes.

II. Background

A. Migration rates by education: evidence from the 1940 to 2000 Censuses

We begin by comparing trends in migration rates and educational attainment over the last 60 years using US Census data. Observing our outcomes of interest over a long time span enables us to assess how the well-documented changes in educational attainment correspond to changes in education group migration rates. One of the two measures of migration we employ in this paper is a

dummy variable indicating that an individual resides outside his birth state at a certain age, typically when he is in his late 20s or early 30s. The main advantage of this measure is its ubiquity: it is observable in the Census, over time, and in several other data sets.⁶ Measuring out of state residence at a point in time captures “lasting” long-distance moves. The migration dummy reflects differences in measured migration rates due to both differential propensities to undertake a move across state lines and to differential probabilities of success in the new location that results in less (unobserved) return migration to the birth state. We introduce a distance-based measure later in the paper.

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.⁷ Thus we interpret migration rate differences observed according to our measure as reflecting migration undertaken by individuals themselves rather than their parents.

Figure 1 shows migration rates for men ages 25 to 45 for four mutually exclusive educational attainment groups during each Census year in which educational attainment is reported. College graduate mobility rates increased in the 1940s and 1950s and then remained stable at or near 50% through 2000. Mobility rates of high school graduates have also been quite stable, remaining near

⁶ Five year migration rates, instead of the “lifetime” rates that we use, are an alternative migration measure one can construct from Census data. 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. Unreported results suggest that the timing (flow) of migration choices varies across education groups, so comparing five-year migration rates across education groups at a particular age could give a misleading picture of lifetime migration differences across the groups.

⁷ Using the NLSY 79, Wozniak (2006) 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.

30% for the entire 60 year period. Men with some (one to three years) of college have migration rates substantially above those of high school graduates but below those of college graduates, and their relative migration rates are roughly stable over the entire 1940 to 2000 period. High school graduates, on the other hand, exhibit migration rates that are indistinguishable from those of high school dropouts for most of the sample period.⁸ How do these patterns compare to trends in educational attainment among men over the same time period? Figure 2 plots the shares of men in this age group at three increasingly restrictive educational attainment levels over the same period. Educational attainment for these men was clearly increasing over the period. Roughly, the patterns in Figures 1 and 2 show that as the share of the population obtaining higher levels of education increased, migration rates at higher levels of attainment flattened or declined.

A model in which educational differences in migration were driven entirely by selection would instead predict declining migration rates at higher levels of educational attainment as individuals with lower propensities to migrate joined the ranks of the more educated. Although the general direction of the trends is consistent with a simple selection story, closer inspection of these trends suggests that this is unlikely to fully explain observed migration differences across education groups. The shortcomings in the selection story are particularly apparent the middle of the data. During the 1960 to 1980 period, the share of men with at least a high school degree rises from (approximately) 50% to 80%, the share of men with at least some college from less than 20% to 40%, and the share of men with a college degree from 13% to 26%. These are dramatic changes in educational attainment, but migration rates for all education groups were flat over this period. Moreover, if migration rates are driven entirely by a fixed personal taste for migration that is heterogeneous across the population, then as education groups become increasingly representative

⁸ Migration rates among African-Americans show some exceptions to these patterns, but these are likely to be largely attributable to the decades-long migration of blacks from the south to the north that concluded in the 1960s. By 1990, migration rates of blacks in both education groups had largely converged to those of whites. All figures are similar for women. We restrict our discussion to men for the sake of brevity.

of the population educational differences in migration rates should disappear. Migration rates of the four education groups in Figure 1 should converge to a single rate equal to the average propensity to migrate. This prediction is also at odds with observed trends.

All in all, the graphical evidence leads to mixed conclusions. The findings do not allow us to unequivocally reject either a selection-based or a causal interpretation of migration rate differences across education groups. Instead, they highlight the need to examine this question in a quasi-experimental setting.

B. Potential Mechanisms: Causality versus Selection

In a typical migration model, an individual who considers moving to another locale compares the costs of undertaking the move to the present value of the expected stream of benefits from moving.⁹ It is certainly possible to write down a structural, market-equilibrium model of an individual's migration decision (See Borjas 2005 for a recent example.) We are interested in determining whether differences in the level of migration investment across education groups are causally driven by educational attainment itself or due to correlation between migration and education choices arising from other unobserved personal characteristics. For this reason we focus on credible identification of college's reduced form effects. Although detailed modeling has the potential to add to our understanding of the results presented here, we leave that for future research.

Education may *directly* confer skills or provide information which enhances the ability to undertake a long distance move. For example, 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

⁹ Schultz (1961) and Becker (1962) were among the first to introduce this approach of analyzing mobility.

itself necessitate a move across state lines. Finally, a college education may 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.

Education may affect the likelihood of undertaking a long distance move in an *indirect* manner as well. Recent research has confirmed that increased schooling is causally related to higher wages (Card, 2001). If moving involves a fixed monetary cost, a higher income may lower the relative cost of moving to another locale. 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.¹⁰ Finally, education may alter other non-economic characteristics such as marriage which can affect the likelihood of geographic mobility (Lefgren and McIntyre, 2006).

On the other hand, it is possible that no such causal factors drive the observed relationship between education and migration. Instead, individuals with a college education may possess inherent attributes that enable them to consider migrating with greater frequency and success. The education premium in migration rates is said to be driven by selection in this case. For example, individuals with greater ability to accumulate and process information may be more likely to pursue a college education and to consider opportunities away from home, or they may simply be more open to new experiences. Heterogeneity in individual discount rates could also generate a spurious relationship between migration and educational attainment. Standard discounting of future costs and benefits implies that individuals with lower discount rates are more likely to undertake migration since the

¹⁰ See Bound et al. (2004).

costs of moving are borne up front while benefits are paid out gradually over time. Basic human capital theory predicts that these same individuals tend to invest in more education.

III. Assessing Causality using IV: Methods and Data

We are interested in understanding the following relationship:

$$(1) \quad m_{isbt} = g(\text{educ}_i, X_{it}, \gamma_s, \gamma_b, \gamma_t)$$

where m is a measure of individual i 's mobility. Factors that govern mobility include the individual's educational attainment, educ ; other, possibly time-varying, individual characteristics X ; as well as constant state of origin factors indexed by s , cohort-specific factors indexed by b , and time-varying factors indexed by t .¹¹

If we suppose that $g(\bullet)$ is a linear function, then we are interested in the following:

$$(2) \quad m_{isbt} = \beta_0 + \beta_1 \text{educ}_i + X_{it} \Gamma + \theta_s + \theta_b + \theta_t + \varepsilon_{isbt}$$

where the θ 's now represent origin state, birth year cohort, and time fixed effects. Even with this simplification, we are faced with a problem. An individual's educational attainment may be correlated with the unobserved characteristics represented by ε , resulting in biased estimates of β . This is the selection problem discussed in previous sections. We solve this using a novel instrumental variables strategy, which allows us to identify the causal impact of education.

A. Using national and state-level induction risk to identify education and veteran status

We extend an instrumental variables strategy inspired by findings in Card and Lemieux (2000, 2001). They find evidence that Vietnam era induction risk varied significantly across birth year cohorts due to differences in military manpower requirements and cohort size, and show that

¹¹ One might also want to include time-varying state of origin factors to account for relative state economic conditions. We are not able to include flexible state or origin effects in our IV specifications, as will be apparent later. We do include an equivalent of year dummies in a robustness check. These proxy for periods of high and low overall mobility, which will be driven in part by state relative conditions.

young men responded to an increased risk of being drafted by attending college at higher rates. By comparing male college attainment to attainment for females from the same birth year cohort, they find that the risk of induction led to significant deviations in male college attainment gains from the inter-cohort path traced out by females. Had male college attainment followed the same path as that for unaffected women in the same cohorts, Card and Lemieux estimate that male college attendance would have been 4 to 6 percentage points lower and graduation 2 percentage points lower in the affected cohorts.

Like Card and Lemieux, 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. This allows us to estimate the following form of Equation 2 via two stage least squares:

$$(3) \quad I_{isbt} = \beta_0 + \beta_1 \text{educ}_i + X_i \Gamma + \theta_s + \varepsilon_{isbt}$$

Here I is the indicator for whether an individual resides outside his birth state s described in the previous section and educ is the portion of educational attainment predicted by our instrument, which we describe shortly. We define educational attainment as membership in one of four exclusive categories: dropouts, high school graduates, those with some college but no four year degree (hereafter college attenders), and college graduates. Equation 3 reflects additional limitations imposed by the cross-sectional Census data we will use to estimate it. Personal characteristics X are not time-varying and include age. We therefore cannot include birth year dummies, and they are dropped from the specification. Since we use only one cross-section, we also drop the time dummies from earlier specifications.

At this point a second estimation problem arises. An artifact of our identification strategy for education—using draft avoidance behavior—is that we must include veteran status in the set of

personal characteristics X because veteran status varies across cohorts in our sample and is a plausible determinant of mobility.

Hence, we are actually interested in estimating the following specification:

$$(4) \quad I_{isbt} = \beta_0 + \beta_1 educ_i + \beta_2 vet_i + \beta_3 X_i + \theta_s + \varepsilon_{it}$$

where vet is a dummy indicating veteran status. Selection processes may mean that veteran status is correlated with ε in much the same way that education and ε may be correlated. This means our equation of interest contains two endogenous variables. We now require an instrument for veteran status as well, since estimates of β_1 are not guaranteed unbiased if the model contains other endogenous variables.¹²

Equation 4 is the specification we ultimately wish to estimate, but to do so we must obtain instruments for both education and veteran status. As discussed above, work by Card and Lemieux (2000, 2001) suggests a potential instrument for college education. To obtain separate instruments for education and veteran status, we exploit state level variation that exists within the cohort level variation identified by Card and Lemieux. Induction risk varied not only across birth year cohorts but also across states within each cohort. The existence of state cohort level variation allows us to break national induction risk into its constituent parts, some of which we argue exogenously influenced an individual's veteran status while the remainder influenced his likelihood of college attendance and completion. We refer to these components as state cohort risk and national cohort risk, respectively, and explain their construction below. This variation has not, to our knowledge, been employed by economists studying problems in which separate instruments are needed to identify the effects of both education and veteran status.

Our identification strategy has a number of strengths that make it particularly useful in the context of internal migration. First, our instrument increases college attendance in a broader way

¹² Wooldridge (2002), p. 83.

than instruments previously employed in the literature, such as proximity to college and variation in state tuition policy. These instruments have provided useful estimates of the return to college 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. Second, since the exogenous shift in education we consider largely arises at the cohort level, the shifts are unlikely to have affected states’ *relative* supplies of college educated workers. In contrast to variation in state tuition policy, our instrument is not correlated with relative labor supply changes that may have induced outmigration among college educated workers seeking the higher relative wages of states with less generous policies. Finally, our strategy allows us to isolate the effect of going to college from that of military service, a distinction that is not possible with the twentieth century’s other great college influx, the generation of World War II veterans who attended college using GI Bill funding.

B. Sources of variation in national and state-level induction risk

The total number of men to be conscripted from a given birth year cohort was determined at the federal level. Responsibility for devising and meeting this goal rested with the Department of Defense (DoD), which issued monthly “draft calls” to meet manpower needs. Draft calls consisted of assigning quotas to state and local draft boards that did the active work of ordering men for induction. Local boards were required to meet their quotas either with new enlistees (volunteers) or by inducting men from the local pool of draft-eligible men. Adding up all of the local quotas should total the national draft call each month.¹³ Card and Lemieux use the number of men inducted annually through this channel to construct the following measure of induction risk:

$$(i) \quad C \ \& \ L \ induction \ risk = \frac{I_b}{N_b}$$

¹³ Information in this paragraph is based on Shapiro and Striker, *Mastering the Draft*, Chapter 20.

where I is the number of inductees from a birth year cohort, N is the number of men in that cohort, and b indexes birth years:¹⁴

State level variation in induction risk arose within cohorts because of severe communication lags between the federal and state level agencies responsible for inductions. These lags meant that local boards knew the number of registrants available at any given point in time while the DoD generally assigned quotas based on local registrant numbers that were several months old. Thus draft risk for an eligible man at a point in time was not only a function of the number of men in his state currently eligible for the draft but also of the number available several months ago. The current pool could be much larger than the past pool if, for example, a large number of local men graduated high school thus becoming draft eligible, or much smaller if a large number married or aged out of the draft pool in the intervening months.¹⁵

Additionally, local control and the high degree of autonomy enjoyed by the boards generated other sources of variation in draft risk across local boards. Davis and Dolbeare write, “Three kinds of variability have been identified: variation based on differences in socioeconomic characteristics of jurisdictions [leading some jurisdictions to have higher shares of college deferments and presumably also of volunteers]; 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 variation produced by idiosyncratic discretionary decision-making by local boards.”¹⁶

¹⁴ Their measure is actually a 5-year weighted average of induction risk around the year a man turned 19.

¹⁵ In practice, local boards also had an active role in generating local draft risk. Local quotas were determined by the number of registrants the board had in class I-A or I-A-O (eligible for active duty or for active duty in a non-combat role). This number is actually a function of the number of deferments boards elected to grant. The communication lags led to an imperfect adjustment process at the local level, as draft boards attempted to grant just enough deferments to avoid having either an excess or a shortage of eligible men at the time of the next draft call. *Mastering the Draft* Ch. 20 describes the cobweb-type feedback these communication lags generated.

¹⁶ Davis and Dolbeare, *Little Groups of Neighbors*, Page 18.

They later write, “The conclusion seems inescapable: local board autonomy implies both within state and between state variability, even among socioeconomically similar board jurisdictions.”¹⁷

Thus, young men residing in state s faced state cohort risk that is analogous to the national cohort risk measure in equation (i), where s indexes state of residence:

$$(ii) \quad \textit{state cohort risk}_{s,b} = \frac{I_{s,b}}{N_{s,b}}$$

We have found no evidence that young men attempted to exploit local level variation in induction risk by moving between localities.¹⁸ This would be costly as well as highly uncertain, since local risk fluctuated even over the course of a year. The extreme difficulty of judging local risk was likely the primary obstacle to such manipulation. On the other hand, young men so inclined were very likely to attempt to avoid induction by enrolling in college. College enrollment was a well known and virtually foolproof way to defer conscription. Deferment through college enrollment was guaranteed in practice prior to 1967 and by law in 1967.¹⁹ Moreover, the incentive to enroll was great: among “fit” men, risk of induction was [eventually] very high.²⁰

We argue that college-going for draft avoidance purposes was primarily a response to perceived national cohort risk since state cohort risk was unknown *ex ante*. National cohort risk was relatively easy to judge as national draft calls were widely reported. Thus draft age men faced two relevant sources of induction risk. State cohort risk influenced their likelihood of becoming veterans—by being drafted—but since this risk was unknown *ex ante*, it was not itself a determinant

¹⁷ *Ibid.* Page 84.

¹⁸ We have consulted two reference guides on the draft, books written for young men at the time to inform them about their draft duties and obligations: *Selective Service: A Guide to the Draft* by Evers and *Guide to the Draft* by Tatum and Tuchinsky. The two represent opposite ends of the political spectrum, with the first written to encourage young men to meet their obligations to serve and the second written to help them avoid the draft through all possible channels. Both are silent on the possibility of exploiting local variation in draft risk. Only later did advocates of reforming the Selective Service System highlight this variation.

¹⁹ Tatum and Tuchinsky, *Guide to the Draft*, Ch. 3.

²⁰ C&L provide a rough estimate of the risk of being drafted conditional on not enrolling in college.

of college going. National cohort risk was known ex ante and therefore was easy to respond to by enrolling in college.

Provided we can obtain data on inductees and cohort size by state, we can use the measure in Equation (ii) to identify veteran status, reserving national cohort risk for identification of educational attainment. This means our identifying variation arises at the state cohort level, because our first stage equations use the portion of national cohort risk that is “left over” after controlling for an individual’s state cohort risk. Specifically, we use the following measure of national cohort risk:

$$(iii) \quad \textit{national cohort risk}_{s,b} = \frac{I_{-s,b}}{N_{-s,b}}$$

The national cohort risk faced by a man living in state s and born in birth year b is approximated by the number of inductees from all other states $-s$ and birth year b , divided by the total number of such men. This remainder is highly correlated with national cohort risk but does vary across state cohorts.²¹ In our analyses, we use birth state to proxy for residence state at the time of the draft. Like Card and Lemieux, we proxy for total cohort size using data on school enrollments of 17 year olds.

C. 1980 Census sample

We examine the impact of mobility on collegiate attainment among males of “the Vietnam generation” using microdata from the 5% sample of the 1980 Census. (Ruggles et al., 2004). 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. We estimated state level cohort size using state level enrollment numbers spanning 1959 to 1970, the academic years in which our cohorts of interest

²¹ Our results are the same whether we use national risk computed as C&L did or a measure of state-cohort specific national risk that omits own state-cohort induction risk from the computation of “national” risk.

were in 11th grade.²² Thus state-cohort level risk (hereafter state risk) for a young man born in Alabama in 1950 equals the number of inductees from Alabama in 1969 (the year he turned 19) divided by the number of students enrolled in 11th grade in Alabama in 1967. 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 Card and Lemieux, 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.

The main cohorts of interest to Card and Lemieux (2005) were men born between 1935 and 1955. We focus on the same 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 further 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 power by excluding the earlier and later cohorts.

We also examine a second measure of geographic mobility in our regression analysis: total distance moved. We compute this as the distance in hundreds of miles between an individual's birth state and his state of residence at the time of the survey—i.e. the distance between the endpoints that determined our dummy variable measure. Distance is approximated by calculating miles between capitals of the two states. This is the best approximation available to us, since we do not have access to finer data on birth location. The main advantage of the distance measure is that it provides a continuous measure of migration which may yield more precise estimates under our identification strategy than can be obtained by instrumenting for a dummy endogenous variable.

Descriptive statistics from the Census samples of the cohorts we consider are given in Tables 1 and 2. Table 1 summarizes the variables used in our analysis using the sample of men born

²² 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.

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. Note that age serves as a linear trend in year of birth when estimating the equations using a single census year. Veteran status is a dummy variable equal to 1 if an individual reported Vietnam era service. Since veteran status is related to induction risk mechanically, but may also influence college-going and interstate migration independently, we take advantage of state-level variation in induction risk to identify the veteran coefficient as explained in the previous section. Furthermore, we allow for state of birth fixed effects because rates of geographic mobility may vary across states.

In addition to personal characteristics, we control for several cohort-level variables that may be related to inter-state mobility levels within a cohort. These include the national unemployment rate at the time a cohort entered the labor force (a moving average around the year a cohort turned 19), log cohort size, as well as the levels of female college graduation. These account for secular trends in college-going that may be correlated with the risk of induction. In particular, female college graduation rates may capture non-linear trends in college-going for men that would have resulted in the absence of any induction risk.²³

Table 2 presents selected statistics separately for each birth year cohort and includes for comparison men born in the earlier and later years of the Card and Lemieux sample. The course of the draft expansion across cohorts is evident in Table 2. The high levels of induction risk and veteran status among men born in the mid and late 1930s are associated with the Korean War, whereas only those cohorts born in the early 1940s and onwards were affected by the Vietnam War.

²³ Our functional form assumptions about the relationship between male and female graduation rates differ somewhat from Card and Lemieux (2000). They examine the ratio of male-female graduation rates, while we include the female graduation rates as controls on the right-hand side. Appendix Figure 1 displays the pattern of male-female graduation trends and show that the time pattern remains remarkably similar when using male-female differences, ratios, or controlling for female trends (and examining the residuals).

The increase in male college going is particularly visible for cohorts born between 1945 and 1951, during the increase in induction risk. This pattern of induction risk across cohorts is quite similar to graphical evidence in Card and Lemieux (2000) and MacInnis (2006). Figure 3 displays these induction rates, as well as both male and female college graduation rates for our sample of cohorts born between 1942 and 1953. Figure 4 shows male and female rates of geographic mobility alongside the induction rates for these same cohorts. The pattern of migration over cohorts is less apparent here, and suggests the use of regression analysis to discern the precise relationship.

Although it is likely that draft avoidance behavior increased male educational attainment at all levels, we focus on instrumental variables estimates of the impact of college graduation on migration since this is the margin of greatest difference in migration rates. Our instrumental variables strategy does not allow us to separately identify the effect of college attendance from that of graduation. Moreover, it is possible that greater measurement error in education at the some college level could bias our IV estimates of the attendance effect upward, along the lines explained by Black et al (2000) for cases of non-classical measurement error.²⁴ We feel that our reduced form estimates provide the best analysis of the overall effect that draft-induced increases in education throughout the distribution had on male migration rates. In appendix tables discussed later, we examine the impact of increased education at the high school level on migration rates using an alternative identification strategy.

IV. Results from the Vietnam Generation

²⁴ Black et al (2000) show that in cases of non-classical measurement error like those found with binary variables, the true parameter estimate lies between the OLS and IV estimates, since the latter is biased upward. Consistent with this, unreported IV estimates of attendance's effect (where those with some college and four year degrees are pooled) are generally much larger than the estimates for graduation. See Black et al (2003) for a discussion of mismeasurement of college completion in the Census.

In this section, we examine the causal effect of education on migration by exploiting draft avoidance behavior during the Vietnam War using data from the 1980 Census. We begin by estimating first stage regressions and show that our measures of induction risk are important determinants of education attainment and veteran status for men but not for women, an important robustness check. We then examine the effect of geographic education on migration in the cross-section and by instrumenting for education with different measures of induction risk. Finally, we discuss the findings and compare our OLS and 2SLS estimates.

A. First stage results for men and women

Table 3 presents results from three first stage specifications. These are estimated using our sample of men in columns (1) through (3). The Panel A specification includes national cohort risk as the only identifying variable in regressions predicting college graduation. This is most similar to other studies in which education is the only instrumented variable. As in Card and Lemieux (2000, 2001), cohorts with higher national risk also tend to have higher rates of college graduation. Our coefficients suggest that a 10 percentage point increase in national cohort risk (roughly the entire range of this variable) raised the probability of completing college by 3.5 to 4 percentage points. The F-statistics suggests that this first stage has substantial power. Although our specification differs from the one employed by Card and Lemieux, the magnitude of our coefficients are relatively similar.²⁵ Moreover, these results confirm that controlling for female graduation more flexibly yields similar results to using male-female relative schooling (see footnote 21).

For the sake of completeness, we repeat the first stage estimates using only state cohort risk as the identifying variable in Panel B of Table 3. When national cohort risk is excluded, state cohort risk positively predicts college graduation rates for men in our sample although the point estimates are smaller, as shown in columns (1) through (3) of this panel. We speculate that the positive

²⁵ Table 1 in Card and Lemieux (2001) shows estimates corresponding to 4.6 percentage point increase for college graduation.

relationship in this specification is largely driven by strong positive collinearity between state and national cohort risk. We discuss the implications of this collinearity for our estimates later in this section.

In Panel C, we fully exploit our two sources of exogenous variation. We estimate two first stage equations, both of which include national *and* state cohort risk as identifying variables. These specifications are used to predict both education and veteran status, consistent with the manner in which 2SLS predicts multiple instrumented variables. Both national and state cohort risk covary positively with veteran status. This is obviously reassuring since higher rates of induction at both the state and national level should lead more young men to go to war. The coefficients on national cohort risk are also positive and significant predictors of education, just as they were when state risk was excluded from the first stage. The estimated effect of national cohort risk on male graduation probabilities is higher in the Panel C specification than it was when state cohort risk was excluded, indicating that a 10 percentage point change in national cohort risk increased male graduation rates by up to 7 percentage points.

State cohort risk is negatively correlated with college graduation and attendance when national induction risk is included. It is difficult to know what we should have expected from this relationship. It seems reasonable to think that men who faced a high state risk of induction are otherwise less likely to go to college—for example, perhaps their local draft boards were less likely to grant college deferments. We might then have expected a negative relationship. On the other hand, we might have expected an insignificant relationship if state variation in induction was completely idiosyncratic. Ultimately, we believe the relationship is partly due to high collinearity between national and state level risk. The coefficient pattern across Panels B and C—in which state level risk positively predicts education when national risk is excluded—is consistent with collinearity

as are the increased point estimates on national cohort risk in Panel C (Verbeek, 2000 and Wooldridge, 2002).

Estimating the same first stage equations on a sample of women sheds light on this issue. Columns 4 and 5 on the right side of Table 3 present these results. Panel A verifies that female college graduation and attendance is not affected by national cohort risk, after controlling for a linear trend in age and additional controls. This is consistent with our contention that induction risk encouraged only men to obtain more higher education. Panel B shows that state cohort risk alone does not affect educational attainment for women either. However including both state and national risk in the first stage equations in Panel C again generates a positive coefficient on national risk and a negative coefficient on state risk. When both variables are included, their coefficients are statistically significant but neither is significant alone in the sample of women.

This pattern suggests strong multicollinearity between national and state cohort risk, especially since the coefficients from the women's sample are almost equal and opposite in sign. 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.²⁶ The collinearity does not introduce systematic bias since the predicted values of the endogenous variables in the first stage are correct in expectation. We suspect that it does, however, contribute to imprecision in our second stage estimates.

B. Estimates of the effect of college graduation on migration

Now that we understand the first stage specifications, we turn to our estimates of interest. We drop specifications using only state cohort risk to identify education in our subsequent analysis. Table 4 presents reduced form estimates of the impact of induction risk on our two measures of geographic mobility. These results are striking. The estimates show that increases in national cohort

²⁶ Verbeek (2000).

risk (our instrument for college graduation) are associated with strong positive increases in both migration measures. Once national cohort risk is included, state cohort risk (our instrument for veteran status) has generally weak negative impacts on migration. This pattern implies that cohorts of men in states who face higher levels of external national cohort risk were more likely to move out of their birth states and on average lived farther from their birth states than men in states facing lower levels of national risk. The magnitudes are large—a 10 percentage point change in national cohort risk increases the probability that a man resides outside his birth state by 16 percentage points. The latter is equal to the OLS coefficient on college graduation in Table 5; recall that 10 percentage points is roughly the range of the national cohort risk variable in our data. The reduced form impact of national cohort risk on this migration measure is even larger once we control for state cohort risk, again possibly as a result of collinearity. The same pattern appears using distance moved as our migration measure. In this case, the reduced form impact exceeds the OLS coefficient on college graduation (in Table 6). The remainder of the paper is devoted to demonstrating that higher levels of college graduation are the mechanism driving these reduced form estimates.

The next two tables present 2SLS estimates. Both tables contain two sets of 2SLS specifications. The Panel A specification uses only national cohort risk as the exogenous variable and instruments only for college graduation. The Panel B specification uses our preferred identification strategy to instrument for both graduation and veteran status. Table 5 presents estimates for the effect of college graduation on geographic mobility measured using our dummy variable for residence outside one's birthstate. We report OLS estimates in columns (1), (2) and (3) since they provide a useful means of assessing the magnitude of our IV estimates and serve as a valuable comparison to OLS estimates from other data sources (such as the NLSY79 that we use later in the paper). Column (1) shows that male college graduates were about 15.2 percentage points more likely to reside outside their birth state than their counterparts with less education. Not merely

significant, the magnitude of this effect is approximately 50% of the average probability of residing outside one's state of birth. Adding a trend in female graduation in column (2), or additional state-specific trends in column (3) does not affect the estimate. In columns (4), (5) and (6) of Panel A, we instrument for college graduation with national cohort risk. These point estimates are considerably larger but significant at conventional levels although they have large standard errors. There is a trend toward larger coefficient estimates across the columns as controls are added. The magnitude of these results thus range from roughly equal to the probability of out of birth state residence for the average man (column 4) to the probability for the average college graduate (column 6).

In Panel B of Table 5, we add veteran status as a control in the OLS specifications and as an endogenous variable in the 2SLS specifications. Adding the veteran control leaves the OLS coefficients on education in columns (1), (2), and (3) essentially unchanged. OLS coefficients on veteran status suggest that veterans were about 5 percentage points more likely to reside outside their birth state than their counterparts who did not serve in Vietnam. This is consistent with concerns about the endogeneity of veteran status. It is reasonable to think that men who are willing to serve in the military are also willing to move geographically. If so, this relationship could bias our estimated coefficient vector and it is the reason we need to instrument for veteran status as well as education.

In Columns (4), (5), and (6), we instrument for both college graduation and veteran status. The magnitudes of the estimated college graduation effect are actually quite similar to those from the IV with only national cohort risk. Consistent with the added imprecision of using two instrumental variables instead of one, the Panel B estimates have larger standard errors than the analogous Panel A estimates. After instrumenting, veteran status appears to have no significant impact on subsequent mobility, consistent with our concerns about its endogeneity.

Table 6 shows results using approximate distance moved as our measure of geographic mobility. OLS results in the left three columns show that college graduate men live an average of 140 miles from their birth state, regardless of whether veteran status is an included control. As in the previous tables, the OLS results are robust to a variety of additional trend controls. Instrumenting for the endogenous variable produces 2SLS estimates (on the right side of the table, Panels A and B) that are considerably larger than the OLS estimates. These results suggest a significant causal impact of college graduation on distance moved that are 4-5 times as large as the OLS estimate.

C. Discussion and Robustness Checks

We consistently find estimated treatment effects that equal or exceed OLS estimates in applying our instrumental variables strategy for college graduation to Vietnam cohorts in the 1980 Census data. Although this pattern is not uncommon in the literature on instrumented returns to education, the magnitudes we find are a little surprising.²⁷ What should we make of the large causal impact of college graduation on mobility estimated in Tables 5 and 6? For one thing, it is important to remember the relatively lumpy nature of our distance variable in comparing results across Tables 5 and 6. Although distance moved has the appearance of a continuous variable—and it is “more continuous” than our dummy variable measure—moving from one state over from the birth state to two states over produces a discontinuous jump in the measure. This potentially accounts for the fact that 2SLS estimates in Table 6 are 4-5 times as large as OLS while those in Table 5 are only twice as large.

²⁷ Card (1995), Card (2001) and Ashenfelter and Krueger (1992) consider the differences between OLS and IV estimates of returns to schooling.

Still, we are left with 2SLS estimates that are roughly twice as large as OLS. The possibility of heterogeneous treatment effects immediately comes to mind.²⁸ We examine evidence for this in the next section using an alternative data source covering different cohorts. It is also worth considering the possibility that the historical setting of our instrument leads to larger treatment effects in the cohorts we examine in Tables 5 and 6. 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. We simply observe that the impacts estimated using our IV strategy are higher than what one might expect. While we allow that the historical setting we use may contribute to this, we do not believe that the causal impact of college on mobility is potentially zero. The array of evidence in previous and subsequent sections, combined with our 2SLS estimates, strongly suggests a large causal effect of college graduation on mobility.

An alternative possibility is that there is some unknown mechanical relationship that leads to estimates of overly large causal impacts when we instrument for education in a migration equation. We explore this issue using an alternative identification strategy. So far, we have focused on the variation in induction rates during the Vietnam War to identify the effect of college completion on geographic mobility. The main motivation for using this source of variation has been twofold: (i) other sources of variation in college-going such as proximity to college and variation in state tuition

²⁸ As is also the case in the returns to schooling context, measurement error might bias OLS estimates downwards, particularly in the case of older survey respondents, who may be more likely to misreport educational attainment. Eliminating bias due to measurement error can increase IV estimates above OLS estimates without an underlying assumption of heterogeneous treatment effects.

policy also alter the relative price of moving out of state, and therefore affect migration; (ii) most of the cross-sectional difference in migration rates occurred between college and high school graduates rather than between high school graduates and high school drop-outs (as shown in Figure 1).

It is a valuable specification check to estimate the causal effect of changes in schooling on migration for the lower segment of the education distribution, even if this difference seems less economically significant in the cross-section. In an appendix to the paper, we examine the variation in schooling generated by quarter of birth which should affect education on the margin of dropping out of high school. We show that instrumenting for education using quarter of birth reveals no evidence for a positive effect of education on geographic mobility. Given that we see very little difference in migration rates between high school graduates and high school dropouts, it is not surprising that using an instrument which affects individuals on this margin shows no corresponding migration effect.

We conducted a number of additional specification checks that bear mentioning. Our results are largely robust to estimating the 2SLS equations on data collapsed to birthstate-birthyear cells. This is level of our variation, so it is natural to consider obtaining estimates using collapsed data. However, the resulting sample size when using cell-level data makes it difficult to support a flexible set of covariates in 2SLS estimates (especially state-level trends). For that reason we prefer estimates obtained from the individual level data. Adding birth year dummies to the Table 5 and 6 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. The statistical significance of the main results is also largely retained when clustering at the broader birth year level, rather than the birthstate-birthyear level as reported. All results are available upon request. Finally, we can compare the estimates from our OLS and 2SLS estimates more formally using a Durbin-Wu-

Hausman test for endogeneity. We reject the null hypothesis that OLS and IV are the same for estimates in Panel B of Table 5. In the other three cases OLS and IV estimates are not statistically distinguishable.²⁹

V. Assessing the Magnitude of Causal Impacts Using the NLSY79

We are able to provide evidence on the plausibility of large causal impacts of college graduation using detailed longitudinal data from the National Longitudinal Survey of Youth 1979. The NLSY is a widely-used resource for social scientists studying the transition from school to the labor force. Although it does not cover the Vietnam cohorts that were the focus of our IV estimates, the NLSY contains information on respondents and their parents that enables us to examine the role of omitted variables in driving the geographic mobility differences we observe between education groups in the Census data. Specifically we can control for ability, parental migration histories, and other personal characteristics that are unobserved in the Census data but may be important determinants of both migration and college decisions.

The NLSY follows a representative sample of approximately 12,000 individuals who were aged 14 to 21 at the outset of the survey in 1979. Statistics summarizing our NLSY sample and included covariates are given in Table 7. In order to match the Census sample, we restrict the NLSY sample to 30-year old men.³⁰ All respondents are U.S. natives, and education is defined as the highest level achieved in the life of the survey. One of its strengths is that the NLSY79 administered the Armed Forces Qualifying Test to its respondents and made the scores on all components of the test available as respondent characteristics. Like other researchers (Johnson and Neal 1996), we use

²⁹ The F-statistic for this test across the different specifications in Table 5 range from 4.11 to 5.51 (p-values 0.0675 to 0.0387) when instrumenting for national induction risk only and from 0.06 to 0.25 (p-values 0.8117 to 0.6256) when instrumenting for both national and state-level risk. For the specifications in Table 6, the F-statistics range from 47.56 to 45.18 (p-values of less than 0.0001) when instrumenting for national induction risk only and from 12.15 to 9.87 (p-values 0.0051 to 0.0094) when instrumenting for both national and state-level risk.

³⁰ All results are robust to using 28- or 35-year olds. 25-year olds are young enough that college graduates have not reached the stable mid-career point in their migration patterns.

these scores as a measure of ability. We normalize the scores to have mean zero and standard deviation one within the age groups at which the test was taken. In our analysis we will also control for mother's education level and for having a parent who was not born in the respondent's birth state. The remaining variables have definitions and means that are similar to those used in the Census sample.³¹ An exception is the veteran status variable as the NLSY cohorts had much lower rates of military participation than those examined in the Census data.

We begin by using NLSY79 data to estimate the same OLS specification used with the Census data. We regress a dummy variable indicating whether a respondent resides outside his birth state at the time of the survey on controls for race, ethnicity, and educational attainment.³² Estimates from this regression are shown in column [1] of the top panel in Table 8. The mobility advantage of college graduates over high school graduates is nearly 17 percentage points, an estimate very close to that obtained in the Census. Estimates of the advantage for those with some college, on the other hand, are somewhat smaller than those obtained in the Census.

In the remaining columns, we add controls that are not available in the Census but are plausible determinants of both the decision to attend college and the decision to move out of state. We progressively add controls both for ability using the AFQT measure (column [2]) and for parental education and migration histories (column [3]).³³ College graduation is still associated with a significantly higher likelihood of out of state residence even after these controls are included. The premium is reduced by about one third, from 17 percentage points to 12 percentage points, but still reflects a sizable difference. Measured ability and parental characteristics are both important for this decline; about half the reduction is due to inclusion of the ability measure and half to the parental

³¹ High shares of Black and Hispanic respondents are due to the NLSY's oversample of disadvantaged youth.

³² Other cohort level controls used with Census data vary so little over the NLSY cohorts that they are dropped due to collinearity.

³³ Maternal education was entered as dummy variables for the same four education groups into which respondents were classified. Parental migration history controls included dummy variables for whether a parent was born outside the respondent's birth state and for whether the respondent had moved out of his birth state by age 14.

controls. The small mobility premium for college attenders in this specification is fully explained by these additional controls; in fact, it becomes insignificant after the addition of ability alone. Finally, we add an expanded set of individual controls (column [4]) that reflect personal traits and life experiences that may ultimately relate to geographic mobility. These include a dummy for residing with both parents at age 14, number of siblings, ASVAB coding speed test score, and measures of self-esteem, sense of control over one's circumstances, traditional family attitudes, and educational aspirations.³⁴ Of these, only the psychometric measures of control and self-esteem were significant. The college graduate coefficient remains large and significant at 0.10.

The bottom panel of Table 8 shows that results are qualitatively unchanged when migration is measured as distance from the birth state. The unexplained portion of the college graduation premium is also robust to using information in the NLSY to construct a dependent variable more closely resembling independent moves on the part of respondents, rather than their parents. When the dependent variable is an indicator for residing outside the state the respondent lived in at age 14, the coefficient on college graduation is somewhat larger in all specifications but the difference is not statistically or economically significant.

What do these results tell us about the instrumented estimates of college's mobility impact obtained from the Census data? First, "unobservables" are working the way we expect for the average person in the sense that some of the college premium in OLS estimates from the Census is due to correlation of college graduation with other characteristics that positively influence mobility. The evidence for this is that adding controls for characteristics observed in the NLSY but unobserved in the Census reduces the coefficients on college graduation in Table 8.

³⁴ Work by Segal (2006) has shown that the coding speed score is as important as the AFQT score in wage equations. Regarding the last four measures, the NLSY asked a set of psychometric questions in 1979 to assess self-esteem and to measure "locus of control," i.e. a measure of how much control respondents feel they have over their lives. It also asked a series of questions about women's roles (used here to measure traditional family attitudes) and about the highest grade a respondent expected to complete. We use the responses to these questions from the 1979 wave.

Second, the fact that two-thirds of the college migration premium in the NLSY remains unexplained by the very broad array of additional controls means that positive omitted variables bias is unlikely to be the sole cause of the large OLS estimates of this premium. A large positive causal impact of college on migration, like that estimated via our IV strategy in the Census, is plausible in light of these NLSY results.

We can use the NLSY to evaluate the plausibility of heterogeneous treatment effects in addition to the plausibility of large average causal impacts. The fact that the IV estimates obtained from Census data exceed the OLS estimates suggests that the marginal man induced to graduate from college experienced a boost to his migration propensity in excess of the average impact. We estimate the specification in column [3] of Table 8, our preferred specification, separately for three sets of groups representing individuals who we hypothesize may be more and less likely to be marginal college graduates in a situation with altered incentives, like those faced by young men during the Vietnam draft.

Table 9 shows the results of this exercise. The cells in Panel A report coefficients on the college graduate dummy, i.e. the college graduation premium, from eight group-specific regressions. The dependent variable in this case is a dummy for out of birth state residence, but as before the results are unchanged if we use the migration measure based on age 14 state of residence. Panel B shows estimates using the distance-based migration measure. It is important to note that the coefficients represent the OLS impact of college graduation on migration for the average graduate in each group. A change in the incentives to college going would change both the composition within the groups shown in the table as well as changing the shares of college graduates from each group. In interpreting the evidence as suggestive of heterogeneous treatment effects, we assume that adding additional college graduates within each group would not change the within-group OLS impact substantially. We assume instead that larger returns for the marginal Vietnam era graduate were

driven by compositional changes in the pool of graduates due to increased shares of graduates with high mobility premia to college.

The results in Table 9 shed light on the type of compositional changes that would be required to bring about IV estimates of the impact of college graduation that equal or exceed the OLS estimates. The evidence is suggestive of heterogeneous treatment effects if one believes the marginal college graduate in the Vietnam era was either more likely to come from the top of the ability distribution than the average graduate or more likely to have a mother with only some college education as opposed to a high school or four year college degree. Are such compositional shifts plausible? We cannot provide direct evidence on this, but unweighted means from our NLSY sample show that less than 60% of those who score in the highest AFQT quartile obtain a college degree while less than half of those with mothers in the some college group obtain one. It appears then that on average there are certainly large numbers of comparatively qualified men who forgo a college degree. If similar men both opted to and found it easiest to obtain a college deferment during the Vietnam era, it is likely that the marginal college graduate experienced higher migration returns to his degree than the average graduate.

VI. Mechanisms behind the Causal Relationship and Concluding Remarks

In this paper, we examined the causal role of college graduation in determining geographic mobility, in particular, the likelihood of lasting, long-distance moves. Using two major data sets, 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 instrumental variables estimates of the 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 to identify the causal effect of a college degree on subsequent mobility. We find that college completion 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 instrument both for college completion and veteran status among our sample of Vietnam era young men.

We also use rich data from the NLSY 1979 to show that roughly two-thirds of the mobility advantage of college graduates is robust to controlling for ability and family characteristics that are typically unobservable in large data sets. We also find that the mobility premium to college completion in OLS regressions is higher among some groups of young men in the NLSY than others. We argue that these groups are likely to be disproportionately represented among marginal college graduates affected by our instrument. This composition effect would generate IV estimates of the effect of college graduation on mobility that are higher than OLS estimates, which is consistent with the evidence we find in the Census.

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 datasets moves out of his birthstate with probability 0.33, and simple OLS estimates put the college graduation premium on out-of-birthstate migration at roughly 0.15. Once we correct for the potential correlation of college completion with characteristics unobserved in the simple regression, our lowest estimate suggests that the college graduation

mobility premium using this migration measure is 0.11 (Table 8). Our highest estimates from our preferred 2SLS specification put this premium at 0.37 (Table 5).

Future work on this topic should delve more deeply into the potential mechanisms behind the effect of college completion on geographic mobility. We provide preliminary evidence on several of these in Tables 10 and 11. We consider two broad classes of causal mechanisms: direct and indirect. Direct effects of college completion are mechanisms through which the college experience contributes to subsequent mobility. These include general cognitive skills that assist in gathering and processing information as well as specific skills in migration that emerge because attending college may itself necessitate a move across state lines. In principle policy makers might expose young workers to any of these direct mechanisms outside the college setting and thereby increase migration rates. Indirect effects are channels through which college completion increases mobility, but not as a result of any component of college experience. These include other benefits to college such as increased income and occupational sorting or upgrading which may contribute to higher migration rates. Indirect mechanisms cannot be separated from college completion itself.

Table 10 presents evidence on two indirect mechanisms. We add controls for income and two-digit occupation to our simple linear probability migration model and estimate this expanded equation using our Census sample. Adding income controls has no impact on the estimated college premium. This suggests that increased income is not the channel through which college affects migration. While migration may be a normal good, the large differences in “migration consumption” across education groups is not likely due to income. On the other hand, about a third of the college mobility advantage is explained by including controls for detailed occupational categories, suggesting that indeed college graduates may be more mobile because the careers they choose require it.

In Table 11 we present evidence on two direct mechanisms. In Panel A, we estimate our linear probability model separately by quartile of “out-of-state college going.” We used data on public and private four-year college enrollment from the 1980 HEGIS survey (conducted by the National Center for Education Statistics) to rank states according to the shares of their college-attending natives getting their degrees out of state. Quartile 1 has the lowest shares of its native attending college out of state, 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. We find weak evidence of this, suggesting in fact a penalty for graduates from the lowest quartile states. In Panel B we again divide states into quartiles, this time according to the geographic diversity of colleges in their birth states measured using a Herfindahl index. 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. Here we find no clear pattern, since the college premium varies across quartiles of state college diversity in different directions using our two migration measures.

This evidence suggests that occupation choice and the experience of having moved to go to college both contribute to the large causal impact of college completion on migration we estimate here. We leave more detailed explorations of these mechanisms for future research.

References

- 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).
- Ashenfelter, Orley and Krueger, Alan. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *The American Economic Review*. 84(1994): 1157-1173.
- Becker G.S. (1962) "Investment in Human Capital: A Theoretical Analysis" in *Journal of Political Economy*. 70(5), pp. 9-49.
- 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.
- Borjas, George. "Native Internal Migration and the Labor Market Impact of Immigration." *NBER Working Paper #11610* (Sep. 2005).
- 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, David A. Jaeger, and Regina M. Baker.. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90(1995): 430:443-50.
- 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.
- Card, David. "Earnings, Schooling, and Ability Revisited." *Research in Labor Economics*. 14(1995): 23-48.
- 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." *American Economic Review*. 91(2001): 97-102.

- 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).
- Cascio, Elizabeth and Lewis, Ethan. "Schooling and the Armed Forces Qualifying Test." *Journal of Human Resources* 41(2006): 294-318.
- Davis, Jr., James W., and Dolbeare, Kenneth M. *Little Groups of Neighbors: The Selective Service System*. Chicago: Markham Publishing Company (1968).
- Dynarski, Susan. "State Tuition Policy and Degree Completion." Manuscript. Massachusetts: Kennedy School of Government, Harvard University (2005).
- Evers, Alf. *Selective Service: A Guide to the Draft*. New York: Lippincott Company (1961).
- 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
- 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.
- 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.
- Nelson, Richard R. and Phelps, Edmund S. "Investment in Humans, Technological Diffusion and Economic Growth." *American Economic Review* 56(1966): 69-75.
- 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.
- 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." *Mimeo*. Universitat Pompeu Fabra (2006).

Shapiro, Andrew O. and Striker, John M. *Mastering the Draft: A Comprehensive Guide to Solving Draft Problems*. Boston: Little, Brown and Company (1970).

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 Waggoner. "Making Their Own Luck: Educational Differences in the Migration Responses in Young Workers to Local Labor Market Conditions." Manuscript. Notre Dame, Indiana: University of Notre Dame (2006).

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

Appendix A: Estimates Using Quarter of Birth Instruments

This paper has focused on the variation in induction rates during the Vietnam War to identify the effect of college on geographic mobility. This source of variation is particularly valuable because the relative price of moving out of state remains unchanged (in contrast to other sources of variation in college-going such as proximity to college and variation in state tuition policy) and because the first-stage effects are on college going which represent most of the cross-sectional difference in migration rates. Nevertheless, as mentioned earlier, it may still be instructive to estimate the causal effect of changes in schooling on migration for the lower segment of the education distribution, even if this difference seems less significant in the cross-section. Consequently, this appendix considers similar estimates using the variation in schooling generated by quarter of birth which should affect education on the margin of dropping out of high school.

Angrist and Krueger (1991) use the interaction of birth quarter with compulsory schooling laws to identify the effect of schooling on wages. They show that children born in the fourth quarter during 1960 started school about 0.4 years later than students born in the first quarter. Furthermore, they use variation in compulsory schooling laws across states to show that these laws influence the relationship between quarter of birth and school enrollment. Using data from the 1980 Census, Angrist and Krueger find that the return to schooling when instrumenting for education with quarter of birth is similar those from OLS. More recently, Lefgren and McIntyre (2006) have used variation in quarter of birth to effect of education on women's marital outcomes. Nevertheless, there are criticisms concerning this identification strategy. Most notable among them is the problem of weak instruments which arises because Angrist and Krueger use quarter of birth interacted with state fixed effects. (Bound et. al., 1995) Similar to Lefgren and McIntyre, we attempt to address this criticism by using a more parsimonious specification.

We implement the quarter of birth identification strategy using the same 1980 Census samples used by Angrist and Krueger: men in cohorts born from 1930 to 1939 and men in cohorts born from 1940 to 1949. Summary statistics of mobility rates, years of education, and age are shown in Panels B and C of Table 1. The second stage regression equation is similar in structure to equation (4) in the main text of the paper:

$$(5) \quad I_i = \beta_0 + \beta_1 \mathit{educ}_i + X_i \Gamma + \varepsilon_i$$

where I is the indicator for whether an individual resides outside his birth state and educ is years of schooling (as opposed to an indicator for college going). The first stage regression equation instruments for years of schooling using indicator variables for quarter of birth.

$$(6) \quad \mathit{educ}_i = \alpha_0 Q2_i + \alpha_1 Q3_i + \alpha_2 Q4_i + X_i \Gamma + v_i$$

All specifications include controls for age and age squared. Our preferred specifications do not interact quarter of birth with state of birth because of concerns with weak instruments. In some specifications we also include state of birth fixed effects, as in our main regressions for the Vietnam generation. However, we do include one specification which interacts quarter of birth with state of birth and includes age fixed effects to compare with Angrist and Krueger's main specification.

Appendix Table 1 displays the results of using quarter of birth to identify the effect of schooling on migration. The OLS coefficients in columns (1), (2), and (3) show the effect of schooling for the entire range of the education distribution. They indicate that an additional year of schooling is associated with an approximately 2 percentage point increase in the probability of residing outside one's birth state for men 40-49 years of age. The corresponding effect for men aged 30-39 years is slightly higher. In contrast, the 2SLS estimates shown in columns (4), (5), and (6) are all over the map. For the majority of the specifications, the effect of schooling on migration appears to be insignificant. However, the effect also comes in negative and positively significant in certain specifications. We therefore conclude that there is no robust effect of education on

migration at the margin where compulsory schooling laws operate. This is not very surprising given that the effect of education on geographic mobility is not substantially different between high school graduates and high school drop-outs. But it does confirm that the causal effect of schooling on migration is mainly a phenomenon that results from higher education.

Table 1: Summary Statistics

	Mean	Stand. deviation	Min	Max	Observations
Living outside of birth state	0.361	0.480	0	1	916,297
Distance moved ('000 miles)	0.307	0.601	0	4.933	916,297
College graduation	0.274	0.446	0	1	916,297
Veteran status	0.332	0.471	0	1	916,297
National induction risk	0.059	0.038	0.002	0.153	916,297
State-level risk	0.059	0.036	0.004	0.112	916,297
Age	31.379	3.430	26	38	916,297
Black	0.099	0.299	0	1	916,297
Other nonwhite	0.048	0.213	0	1	916,297
Unemployment rate	4.800	0.678	3.802	5.992	916,297
Log cohort size	12.056	0.114	11.869	12.171	916,297
Female graduation	0.198	0.018	0.162	0.218	916,297

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS, and restricted to men born between 1942 and 1953 in the US. Cohort size estimated from 1980 5% sample. Mobility is equal to 1 if respondent resides outside his state of birth at the time of the Census, and 0 otherwise. Distance moved is calculated from the capital of respondent's birth state to capital of the respondent's state of residence at the time of the Census. Unemployment is the average national unemployment rate over the years a respondent's birth year cohort was ages 19-22.

Table 2: Mean of Selected Variables by Birth Year

Year of birth	Out of birth-state	Distance moved	College Graduation	Induction risk	Veteran Status
1935	0.399	0.322	0.214	0.082	0.504
1936	0.399	0.327	0.214	0.068	0.482
1937	0.394	0.320	0.216	0.060	0.456
1938	0.402	0.325	0.223	0.053	0.448
1939	0.393	0.320	0.226	0.049	0.416
1940	0.401	0.329	0.235	0.042	0.382
1941	0.393	0.316	0.245	0.044	0.375
1942	0.396	0.328	0.261	0.038	0.377
1943	0.401	0.331	0.266	0.039	0.382
1944	0.397	0.336	0.274	0.082	0.397
1945	0.394	0.340	0.288	0.096	0.428
1946	0.378	0.323	0.303	0.097	0.464
1947	0.370	0.311	0.304	0.112	0.455
1948	0.361	0.306	0.296	0.094	0.413
1949	0.353	0.295	0.287	0.075	0.351
1950	0.344	0.291	0.272	0.051	0.272
1951	0.337	0.287	0.266	0.025	0.201
1952	0.331	0.284	0.247	0.012	0.185
1953	0.326	0.280	0.227	0.004	0.144
1954	0.320	0.274	0.210	0	0.130
1955	0.306	0.262	0.192	0	0.122

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS, and restricted to men born between 1942 and 1953 in the US. Mobility is equal to 1 if respondent resides outside his state of birth at the time of the Census, and 0 otherwise. Distance moved is calculated from the capital of respondent's birth state to capital of the respondent's state of residence at the time of the Census.

Table 3: 1st-Stage Estimates of Induction Risk on Male and Female College Graduation

Panel A: National-level induction risk					
	(1)	(2)	(3)	(4)	(5)
<i>endogenous variable:</i>	<i>Male Graduation</i>			<i>Female Graduation</i>	
National induction risk	0.410** [0.035]	0.365** [0.036]	0.359** [0.025]	0.03 [0.029]	0.025 [0.021]
F-Test	137.96	104.42	199.52	1.04	1.41
Main controls	X	X	X	X	X
Female graduation trend		X	X		
State-level trends			X		X
Panel B: State-level induction risk					
	(1)	(2)	(3)	(4)	(5)
<i>endogenous variable</i>	<i>Male Graduation</i>			<i>Female Graduation</i>	
National induction risk	0.289** [0.036]	0.245** [0.036]	0.292** [0.025]	-0.029 [0.029]	0.002 [0.022]
Joint F-test	63.41	47.71	131.44	1.04	0.01
Main controls	X	X	X	X	X
Female graduation trend		X	X		
State-level trends			X		X
Panel C: National and state-level induction risk					
	(1)	(2)	(3)	(4)	(5)
<i>endogenous variable 1:</i>	<i>Male Graduation</i>			<i>Female Graduation</i>	
National induction risk	0.773** [0.081]	0.727** [0.078]	0.543** [0.072]	0.370** [0.068]	0.188** [0.057]
State-level induction risk	-0.365** [0.138]	-0.364** [0.139]	-0.185** [0.126]	-0.288** [0.065]	-0.120* [0.059]
Joint F-test	95.11	76.46	110.77	15.20	5.72
<i>endogenous variable 2:</i>	<i>Veteran Status</i>			<i>Veteran Status</i>	
National induction risk	1.049** [0.082]	1.072** [0.077]	1.287** [0.072]	-0.021 [0.012]	-0.018 [0.012]
State-level induction risk	0.835** [0.137]	0.834** [0.137]	0.618** [0.129]	0.013 [0.011]	0.01 [0.012]
Joint F-test	436.16	426.70	813.91	2.26	2.21
Main controls	X	X	X	X	X
Female graduation trend		X	X		
State-level trends			X		X

Notes: Robust standard errors clustered by birth year-state in brackets. ** and * indicate statistical significance at the 1 and 5 percent level, respectively. Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. Specifications in columns (1), (2) and (3) are estimated on a sample restricted to men; in columns (4) and (5) on a sample restricted to women. All samples are restricted to individuals born between 1942 and 1953 in the US. Number of observations is roughly 900,000 in each regression. Main controls include race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, and a linear trend.

Table 4: Reduced-form Estimates for the Impact of Induction Risk on Geographic Mobility

Panel A: Dependent variable is living outside state of birth						
	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	0.160** [0.037]	0.161** [0.039]	0.165** [0.029]	0.282** [0.084]	0.283** [0.084]	0.209** [0.078]
State-level induction risk				-0.124 [0.078]	-0.124 [0.078]	-0.045 [0.072]
Main controls	X	X	X	X	X	X
Female graduation trend		X	X		X	X
State-level trends			X			X
R ²	0.03	0.03	0.03	0.03	0.03	0.03
Mean of dep. Variable	0.361	0.361	0.361	0.361	0.361	0.361
Panel B: Dependent variable is distance moved ('000 miles)						
	(1)	(2)	(3)	(4)	(5)	(6)
National induction risk	0.265** [0.047]	0.260** [0.048]	0.261** [0.040]	0.499** [0.097]	0.494** [0.098]	0.332** [0.091]
State-level induction risk				-0.238** [0.092]	-0.237** [0.091]	-0.072 [0.085]
Main controls	X	X	X	X	X	X
Female graduation trend		X	X		X	X
State-level trends			X			X
R ²	0.02	0.02	0.02	0.02	0.02	0.02
Mean of dep. variable	0.307	0.307	0.307	0.307	0.307	0.307

Notes: Robust standard errors clustered by birthyear-state in brackets. ** and * indicate statistical significance at the 1 and 5 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 in the US. Number of observations is 916,297 in each regression. Main controls include race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, and a linear trend.

Table 5: OLS and IV Estimates for the Impact of College Graduation on Likelihood of Living Outside of One's Birth State

Dependent variable: living outside state of birth

Panel A: IV with national-level induction risk

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Graduation	0.161** [0.004]	0.161** [0.004]	0.161** [0.004]	0.387** [0.092]	0.443** [0.110]	0.462** [0.077]
Main controls	X	X	X	X	X	X
Female graduation trend		X	X		X	X
State-level trends			X			X
R ²	0.05	0.05	0.05	0.05	0.05	0.05
Mean of dep. variable	0.361	0.361	0.361	0.361	0.361	0.361

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

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Graduation	0.168** [0.004]	0.168** [0.004]	0.168** [0.004]	0.355* [0.145]	0.371* [0.152]	0.327 [0.221]
Veteran Status	0.051** [0.002]	0.051** [0.002]	0.051** [0.002]	0.007 [0.035]	0.014 [0.033]	0.026 [0.041]
Main controls	X	X	X	X	X	X
Female graduation trend		X	X		X	X
State-level trends			X			X
R ²	0.05	0.05	0.05	0.02	0.01	0.03
Mean of dep. variable	0.361	0.361	0.361	0.361	0.361	0.361

Notes: Robust standard errors clustered by birth year-state in brackets. ** and * indicate statistical significance at the 1 and 5 percent level, respectively. Data are from 5% samples of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953 in the US. Number of observations is 916,297 in each regression. Main controls include race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, and a linear trend in age. Columns 4, 5, and 6 of Panel A use national-level induction risk to instrument for graduation from college. Columns 4, 5, and 6 of Panel B use national and state-level induction risk to instrument for graduation and veteran status.

Table 6: OLS and IV Estimates for the Impact of College Graduation on Distance Moved
(between Birth State and State of Residence)

dependent variable: distance of move ('000 miles)

Panel A: IV with national-level induction risk

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Graduate	0.137** [0.003]	0.137** [0.003]	0.137** [0.003]	0.642** [0.109]	0.711** [0.126]	0.729** [0.109]
Main controls	X	X	X	X	X	X
Female attendance trend		X	X		X	X
State-level trends			X			X
R ²	0.03	0.03	0.03	0.02	0.02	0.02
Mean of dep. variable	0.307	0.307	0.307	0.307	0.307	0.307

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

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Graduate	0.144** [0.003]	0.144** [0.003]	0.144** [0.003]	0.648** [0.161]	0.670** [0.168]	0.523 [0.269]
Veteran Status	0.056** [0.002]	0.056** [0.002]	0.056** [0.002]	-0.001 [0.040]	0.008 [0.038]	0.039 [0.052]
Main controls	X	X	X	X	X	X
Female attendance trend		X	X		X	X
State-level trends			X			X
R ²	0.03	0.03	0.03	0.02	0.02	0.02
Mean of dep. variable	0.307	0.307	0.307	0.307	0.307	0.307

Notes: Robust standard errors clustered by birth year-state in brackets. ** and * indicate statistical significance at the 1 and 5 percent level, respectively. Data are from 5% samples of the 1980 U.S. Census, available from IPUMS. All specifications are restricted to men born between 1942 and 1953 in the US. Number of observations is 916,297 in each regression. Main controls include race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, and a linear trend in age. Columns 4, 5, and 6 of Panel A use national-level induction risk to instrument for graduation from college. Columns 4, 5, and 6 of Panel B use national and state-level induction risk to instrument for graduation and veteran status.

Table 7: Variable Means in the NLSY79 Sample

Means of Key Regressors		Means of Migration Measures by Educational Attainment	
Black	0.28	Moved out of Birth State	
		<i>Dropouts</i>	0.30
Hispanic	0.11	<i>High School Grad</i>	0.29
		<i>Some College</i>	0.33
Veteran/AF Experience	0.05	<i>College Graduate</i>	0.48
Dropout	0.23	<i>Sample</i>	0.34
High School Graduate	0.35	Moved out of Age 14 State	
		<i>Dropouts</i>	0.20
Some College	0.20	<i>High School Grad</i>	0.18
		<i>Some College</i>	0.23
College Graduate	0.22	<i>College Graduate</i>	0.38
Normalized AFQT	0.01	<i>Sample</i>	0.24
Mother Dropout	0.39	Moved out of Birth State by Age 14	
		<i>Dropouts</i>	0.21
Mother HS Graduate	0.43	<i>High School Grad</i>	0.18
		<i>Some College</i>	0.19
Mother Some College	0.10	<i>College Graduate</i>	0.23
Mother College Graduate	0.08	<i>Sample</i>	0.20
Parent Born Out of State	0.45	Miles from Birth State (1000's)	
		<i>Dropouts</i>	0.20
		<i>High School Grad</i>	0.21
		<i>Some College</i>	0.30
		<i>College Graduate</i>	0.41
		<i>Sample</i>	0.27

Notes: Sample is 30-year old men in the NLSY79.

Table 8: OLS Estimates of Migration Equations using the NLSY 1979

Panel A Dependent Variable: Out of Birth State Residence				
	[1]	[2]	[3]	[4]
Black	-0.04 [0.02]**	-0.02 [0.02]	0.01 [0.02]	0.02 [0.02]
Hispanic	-0.05 [0.03]*	-0.03 [0.03]	0.04 [0.03]	0.05 [0.03]*
Veteran/AF Experience	0.10 [0.03]***	0.10 [0.04]***	0.07 [0.03]**	0.07 [0.04]*
Dropout	0.00 [0.02]	0.03 [0.02]	0.00 [0.02]	0.00 [0.02]
Some College	0.04 [0.02]**	0.03 [0.02]	0.01 [0.02]	0.01 [0.02]
College Graduate	0.17 [0.02]***	0.14 [0.02]***	0.12 [0.02]***	0.11 [0.02]***
Normalized AFQT		0.03 [0.01]***	0.01 [0.01]	0.01 [0.01]
Maternal Ed. & Family Migration Controls	No	No	Yes	Yes
Expanded Individual Controls	No	No	No	Yes
Birth State & Cohort FEs	Yes	Yes	Yes	Yes
Observations	4234	4038	3119	3008
R-squared	0.09	0.10	0.38	0.39

(Table 8 continued on next page)

Table 8, continued

Panel B Dependent Variable: Distance from Birth State (1000's of miles)				
	[1]	[2]	[3]	[4]
Black	-0.04 [0.02]**	-0.02 [0.02]	0.01 [0.02]	0.02 [0.03]
Hispanic	-0.05 [0.03]*	-0.03 [0.03]	0.04 [0.03]	0.09 [0.05]*
Veteran/AF Experience	0.1 [0.03]***	0.1 [0.04]***	0.07 [0.03]**	0.09 [0.05]
Dropout	0 [0.02]	0.03 [0.02]	0 [0.02]	-0.01 [0.02]
Some College	0.04 [0.02]**	0.03 [0.02]	0.01 [0.02]	0.03 [0.03]
College Graduate	0.17 [0.02]***	0.14 [0.02]***	0.12 [0.02]***	0.10 [0.03]***
Normalized AFQT		0.03 [0.01]***	0.01 [0.01]	0.02 [0.01]
	[1]	[2]	[3]	[4]
Maternal Ed. & Family Migration Controls	No	No	Yes	Yes
Expanded Individual Controls	No	No	No	Yes
Birth State & Cohort FEs	Yes	Yes	Yes	Yes
Observations	4234	4038	3119	3008
R-squared	0.09	0.10	0.38	0.27

Notes: Dependent variable is one if respondent resides out of birth state at the time of the survey, zero otherwise. Data are from the National Longitudinal Survey of Youth, 1979. Sample is all 30-year old U.S. native men. Respondent's education variables indicate highest level of schooling achieved in the life of the survey, not current education. Robust standard errors are in brackets. *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level.

Table 9: Coefficients on College Graduate in Separate Regressions by Group

Panel A Dependent Variable: Out of Birth State Residence				
<u>Groups</u>				
AFQT Quartile	<i>Bottom</i>	<i>Second</i>	<i>Third</i>	<i>Top</i>
	0.05	0.13	0.10	0.17
	[0.12]	[0.06]**	[0.04]**	[0.04]***
Mother's Education	<i>Dropout</i>	<i>HS Grad</i>	<i>Some College</i>	<i>College Grad</i>
	0.10	0.12	0.20	0.12
	[0.05]*	[0.03]***	[0.07]***	[0.10]
Race	<i>Black</i>	<i>White</i>		
	0.13	0.12		
	[0.05]***	[0.03]***		
Panel B Dependent Variable: Distance from Birth State (1000's of miles)				
<u>Groups</u>				
AFQT Quartile	<i>Bottom</i>	<i>Second</i>	<i>Third</i>	<i>Top</i>
	0.02	0.02	0.15	0.14
	[0.092]	[0.055]	[0.061]**	[0.046]***
Mother's Education	<i>Dropout</i>	<i>HS Grad</i>	<i>Some College</i>	<i>College Grad</i>
	0.07	0.07	0.37	0.13
	[0.055]	[0.040]	[0.107]***	[0.154]
Race	<i>Black</i>	<i>White</i>		
	0.07	0.12		
	[0.049]	[0.036]***		

Notes: Cells show coefficient and standard error on college graduate dummy variable in separate regressions of specification [3] from Table 8 by group. Dependent variable is one if respondent resides out of birth state at the time of the survey, zero otherwise. Data are from the National Longitudinal Survey of Youth, 1979. Sample is all 30-year old U.S. native men. Robust standard errors are in brackets. *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level.

Table 10: OLS Estimates for Indirect Effects of College Graduation on Geographic Mobility

<i>dependent variable</i>	<i>living outside state of birth</i>			<i>distance moved ('000 miles)</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
College Graduation	0.168** [0.004]	0.160** [0.004]	0.111** [0.003]	0.144** [0.003]	0.141** [0.003]	0.095** [0.003]
Main controls	X	X	X	X	X	X
Wage income controls		X	X		X	X
2-digit occupation controls			X			X
R ²	0.05	0.05	0.07	0.03	0.03	0.05
Mean of dep. Variable	0.361	0.361	0.361	0.307	0.307	0.307

Notes: Robust standard errors clustered by birth year-state in brackets. ** and * indicate statistical significance at the 1 and 5 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 in the US. Number of observations is 916,297 in each regression. Main controls include veteran status, race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, female graduation rate, and a linear trend in age.

Table 11: OLS Estimates of the Impact of College Graduation on Mobility by Quartiles of Out-of-state College Going and Geographic Diversity of the State College System

Panel A: Out-of-state College Going (by quartile)

<i>dependent variable</i>	<i>living outside state of birth</i>				<i>distance moved ('000 miles)</i>			
	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)	Q1 (5)	Q2 (6)	Q3 (7)	Q4 (8)
College Graduation	0.124** [0.008]	0.184** [0.005]	0.184** [0.005]	0.201** [0.005]	0.117** [0.005]	0.151** [0.005]	0.156** [0.005]	0.164** [0.007]
Observations	296,551	272,271	251,471	96,004	296,551	272,271	251,471	96,004
R ²	0.05	0.04	0.04	0.06	0.02	0.03	0.04	0.04
Mean of dependent variable	0.312	0.391	0.372	0.396	0.274	0.292	0.352	0.329

Panel B: Geographic Diversity of the State College System (by quartile of the Herfindahl Index)

<i>dependent variable</i>	<i>living outside state of birth</i>				<i>distance moved ('000 miles)</i>			
	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)	Q1 (5)	Q2 (6)	Q3 (7)	Q4 (8)
College Graduation	0.167** [0.007]	0.184** [0.004]	0.155** [0.006]	0.136** [0.007]	0.145** [0.004]	0.145** [0.004]	0.132** [0.006]	0.160** [0.011]
Observations	504,424	216,249	146,741	48,883	504,424	216,249	146,741	48,883
R ²	0.04	0.04	0.05	0.06	0.02	0.03	0.04	0.1
Mean of dependent variable	0.329	0.367	0.427	0.471	0.299	0.300	0.294	0.448

Notes: Robust standard errors clustered by birth year-state in brackets. ** and * indicate statistical significance at the 1 and 5 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 in the US. Number of observations is 916,297 in each regression. Main controls include veteran status, race (white, black, other), national cohort size, national-level unemployment rate, state-level fixed effects, female graduation rate, and a linear trend in age.

Appendix Table 1: Summary Statistics by State of Birth

State	Out of birth state	Distance moved	College Graduation	Induction risk	Veteran
Alabama	0.41	0.259	0.18	0.06	0.30
Alaska	0.52	1.040	0.19	0.04	0.31
Arizona	0.45	0.383	0.22	0.05	0.32
Arkansas	0.51	0.361	0.19	0.06	0.32
California	0.25	0.315	0.30	0.05	0.33
Colorado	0.44	0.377	0.30	0.05	0.35
Connecticut	0.36	0.337	0.36	0.05	0.33
Delaware	0.38	0.251	0.29	0.06	0.36
D.C.	1.00	0.514	0.40	0.06	0.33
Florida	0.38	0.336	0.24	0.05	0.34
Georgia	0.32	0.189	0.19	0.06	0.31
Hawaii	0.34	0.972	0.31	0.04	0.34
Idaho	0.56	0.413	0.29	0.04	0.34
Illinois	0.36	0.311	0.32	0.07	0.34
Indiana	0.35	0.268	0.25	0.06	0.35
Iowa	0.44	0.325	0.29	0.06	0.35
Kansas	0.50	0.361	0.31	0.05	0.35
Kentucky	0.40	0.188	0.17	0.08	0.32
Louisiana	0.31	0.255	0.22	0.06	0.30
Maine	0.39	0.365	0.21	0.05	0.38
Maryland	0.34	0.262	0.26	0.05	0.33
Massachusetts	0.35	0.364	0.34	0.04	0.33
Michigan	0.28	0.284	0.26	0.07	0.35
Minnesota	0.32	0.308	0.29	0.06	0.36
Mississippi	0.51	0.328	0.18	0.06	0.27
Missouri	0.41	0.294	0.27	0.07	0.37
Montana	0.49	0.423	0.31	0.05	0.36
Nebraska	0.47	0.368	0.30	0.06	0.35
Nevada	0.68	0.707	0.29	0.05	0.36
New Hampshire	0.41	0.352	0.26	0.04	0.37
New Jersey	0.38	0.337	0.37	0.05	0.31
New Mexico	0.52	0.397	0.22	0.05	0.32
New York	0.38	0.402	0.38	0.06	0.31
North Carolina	0.31	0.187	0.19	0.07	0.31
North Dakota	0.55	0.474	0.30	0.06	0.38
Ohio	0.33	0.297	0.27	0.06	0.35
Oklahoma	0.45	0.341	0.28	0.05	0.36
Oregon	0.39	0.308	0.29	0.05	0.38
Pennsylvania	0.35	0.270	0.29	0.06	0.35
Rhode Island	0.43	0.412	0.33	0.05	0.35
South Carolina	0.36	0.213	0.18	0.06	0.31
South Dakota	0.57	0.447	0.31	0.05	0.34
Tennessee	0.36	0.210	0.20	0.07	0.32
Texas	0.25	0.255	0.24	0.06	0.31
Utah	0.36	0.267	0.33	0.05	0.30
Vermont	0.43	0.342	0.21	0.04	0.36
Virginia	0.37	0.260	0.22	0.07	0.32
Washington	0.37	0.424	0.30	0.05	0.37
West Virginia	0.50	0.246	0.17	0.07	0.36
Wisconsin	0.30	0.257	0.28	0.06	0.34
Wyoming	0.58	0.428	0.30	0.05	0.37

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS, and restricted to men born between 1942 and 1953 in the US.

Appendix Table 2: Summary Statistics by Education Group

Education level	Out of birth-state	Distance moved	Induction risk	Veteran Status
HS Drop-out	0.281	0.193	0.059	0.224
HS Graduate	0.298	0.242	0.057	0.395
Some college	0.377	0.349	0.059	0.407
College grad	0.476	0.414	0.062	0.246

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS, and restricted to men born between 1942 and 1953 in the US. Cohort size estimated from 1980 5% sample. Mobility is equal to 1 if respondent resides outside his state of birth at the time of the Census, and 0 otherwise. Distance moved is calculated from the capital of respondent's birth state to capital of the respondent's state of residence at the time of the Census. Unemployment is the average national unemployment rate over the years a respondent's birth year cohort was ages 19-22.

Appendix Table 3: Quarter of Birth Estimates of Schooling on Geographic Migration

Panel A: living outside state of birth

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.024** [0.000]	0.027** [0.000]	0.027** [0.000]	0.003 [0.024]	0.004 [0.021]	0.008 [0.018]
Age and age squared	X	X	X	X	X	X
State-fixed effects		X	X		X	X
Age indicator controls			X			X
R ²	0.03	0.07	0.07	0.01	0.04	0.04
Mean of dep. variable	0.400	0.400	0.400	0.400	0.400	0.400

Panel B: distance moved ('000 miles)

	<i>OLS</i>			<i>2SLS</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	0.028** [0.000]	0.026** [0.000]	0.026** [0.000]	-0.006 [0.029]	-0.004 [0.026]	0.001 [0.022]
Age and age squared	X	X	X	X	X	X
State-fixed effects		X	X		X	X
Age indicator controls			X			X
R ²	0.02	0.04	0.04		0.01	0.02
Mean of dep. variable	0.383	0.383	0.383	0.383	0.383	0.383

Notes: Data are from the 5% sample of the 1980 U.S. Census, available from IPUMS. Number of observation is 252,899. Sample includes men born between 1930 and 1939. Age dummy controls represent indicator variables for each age year. Columns 4 and 5 of Panels A and B use quarter-of-birth to instrument for years of schooling. Column 6 of Panels A and B use the full set of quarter-of-birth times year of birth interactions as instruments for years of schooling.

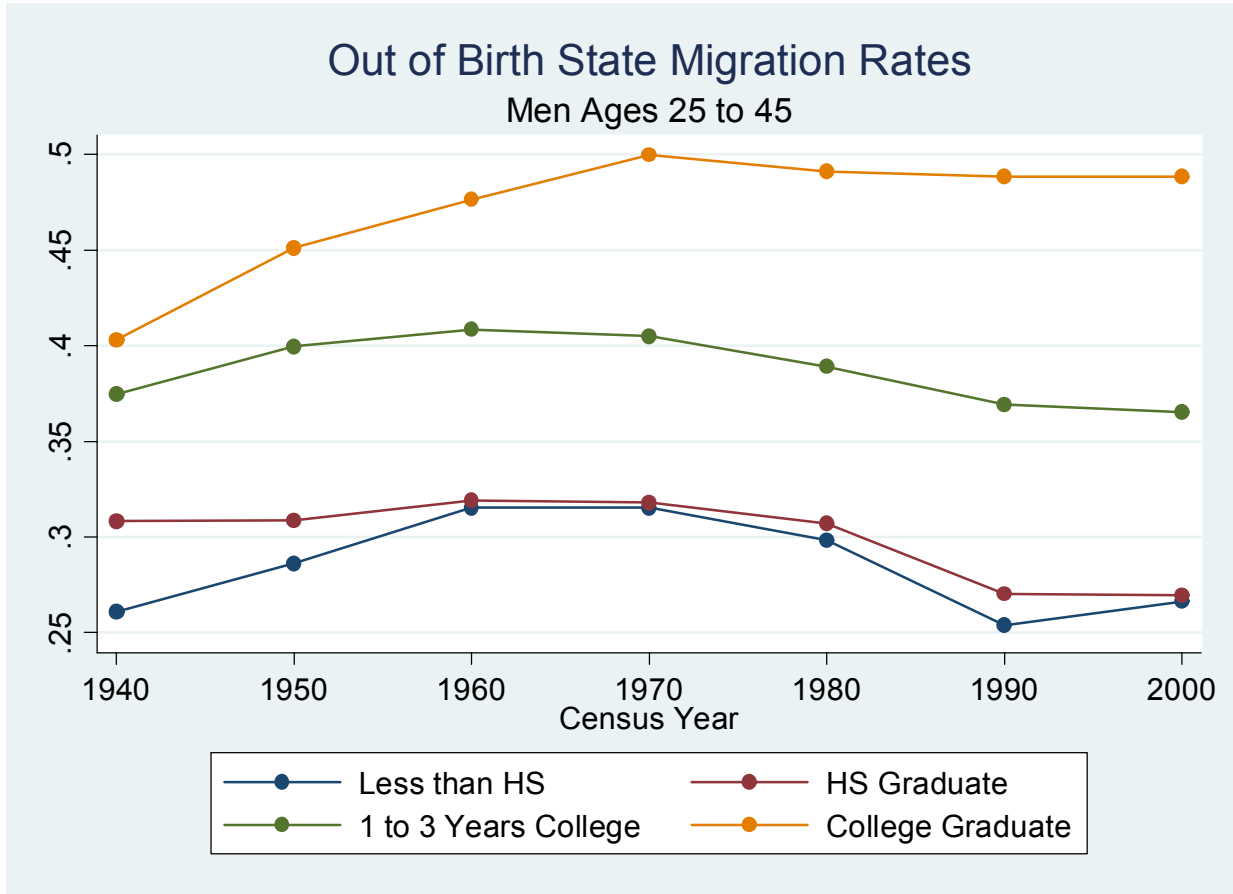


Figure 1: Data from IPUMS U.S. Census extracts, Census years 1940 to 2000. Sample in each year is U.S. native men, ages 25 to 45 not living in group quarters or abroad. Migration rates indicate the share of men in each Census year living outside their states of birth at the time of the Census.

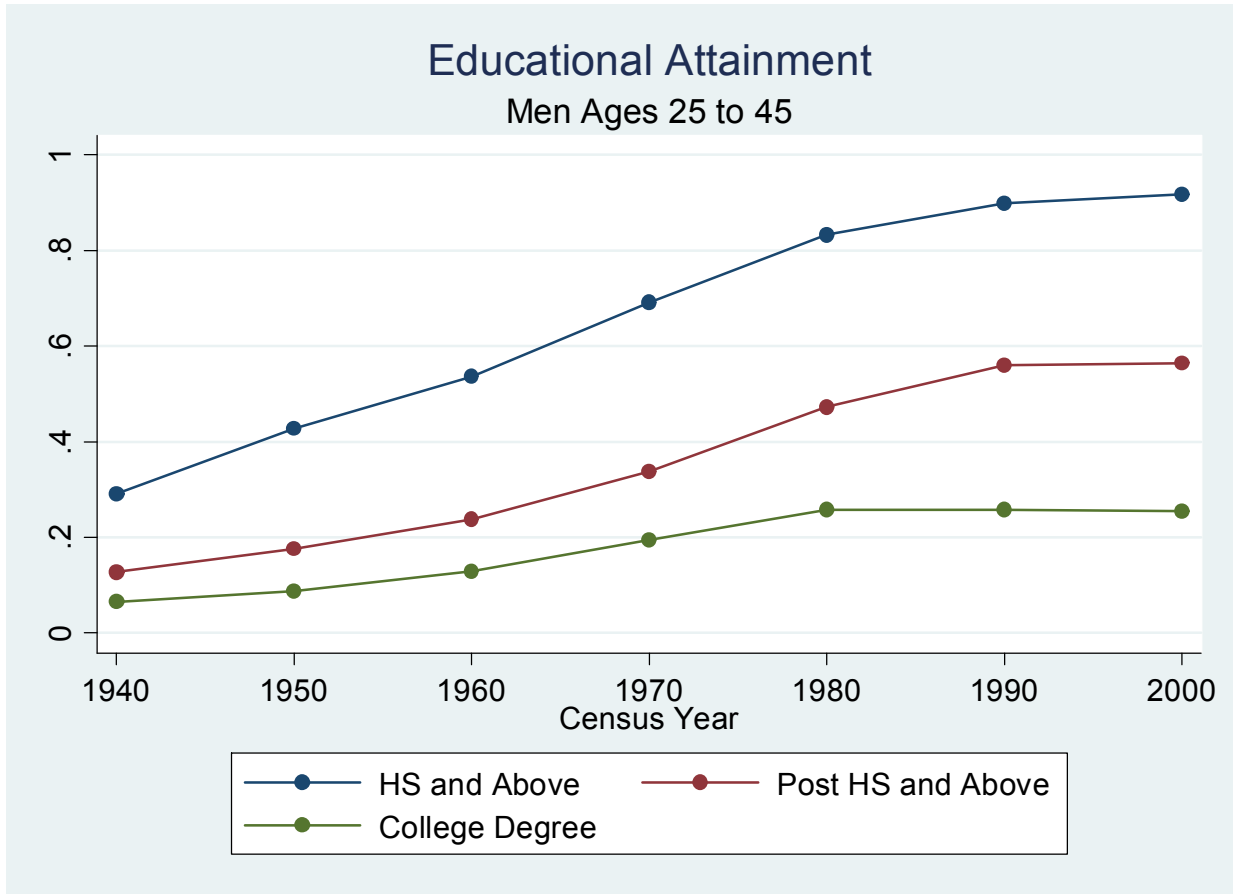
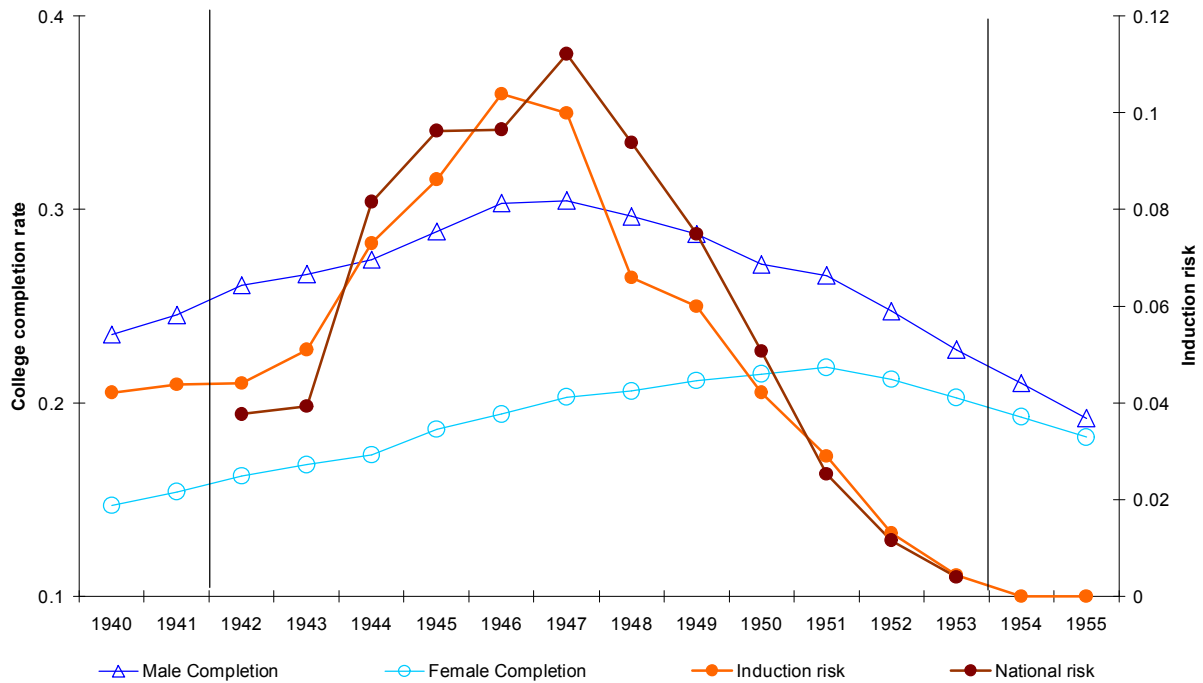


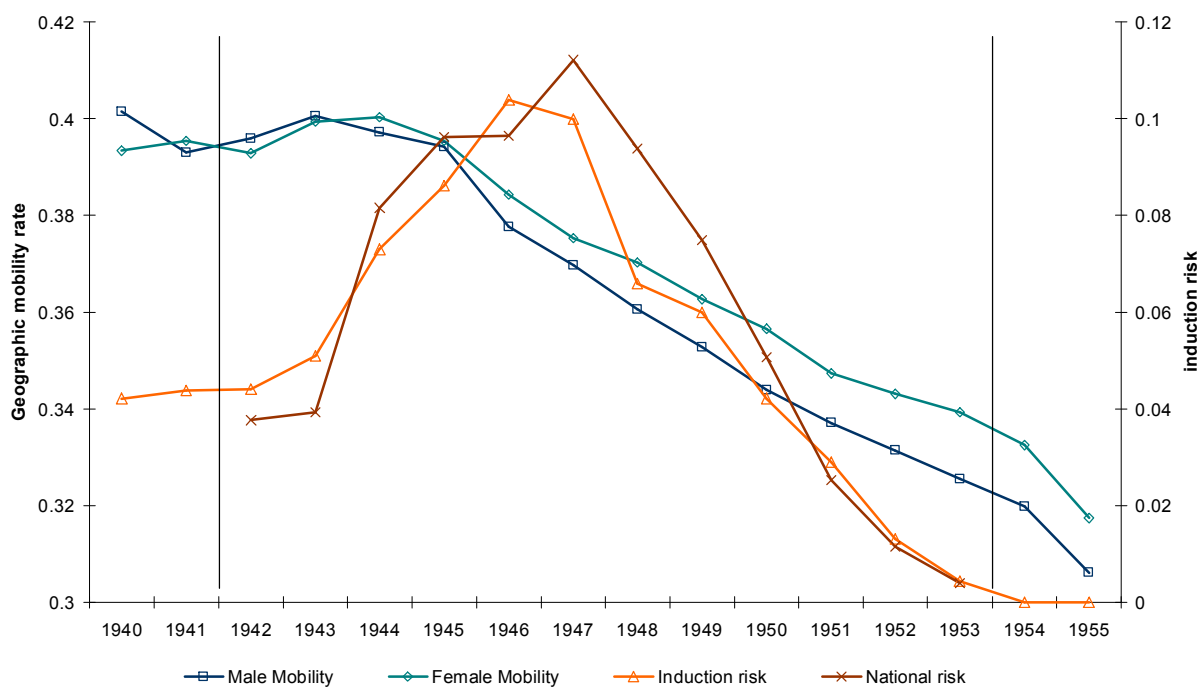
Figure 2: Data from IPUMS U.S. Census extracts, Census years 1940 to 2000. Sample in each year is U.S. native men, ages 25 to 45 not living in group quarters or abroad. Educational attainment rates indicate the share of men in each Census year reporting educational achievement at or above the indicated level. Categories are not mutually exclusive.

Figure 3: Induction risk and college-going for men and women



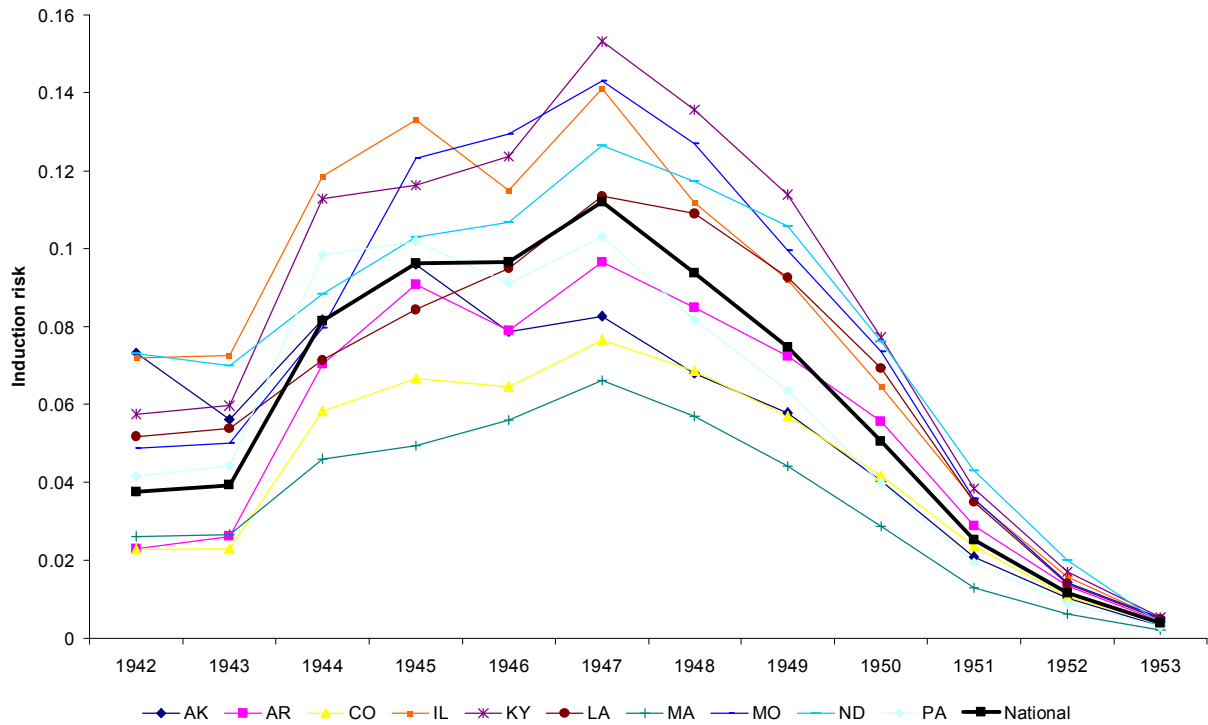
Notes: Each point represents an unadjusted mean from the 1980 Census among U.S. born birth cohorts 1940–1955. Male and female completion indicates the fraction with college graduates among males and females respectively. Induction risk represents the Card and Lemieux measure of induction risk, while national induction risk portion of national cohort risk that is “left over” after controlling for an individual’s state cohort risk. See the text for more details.

Figure 4: Induction risk and mobility for men and women



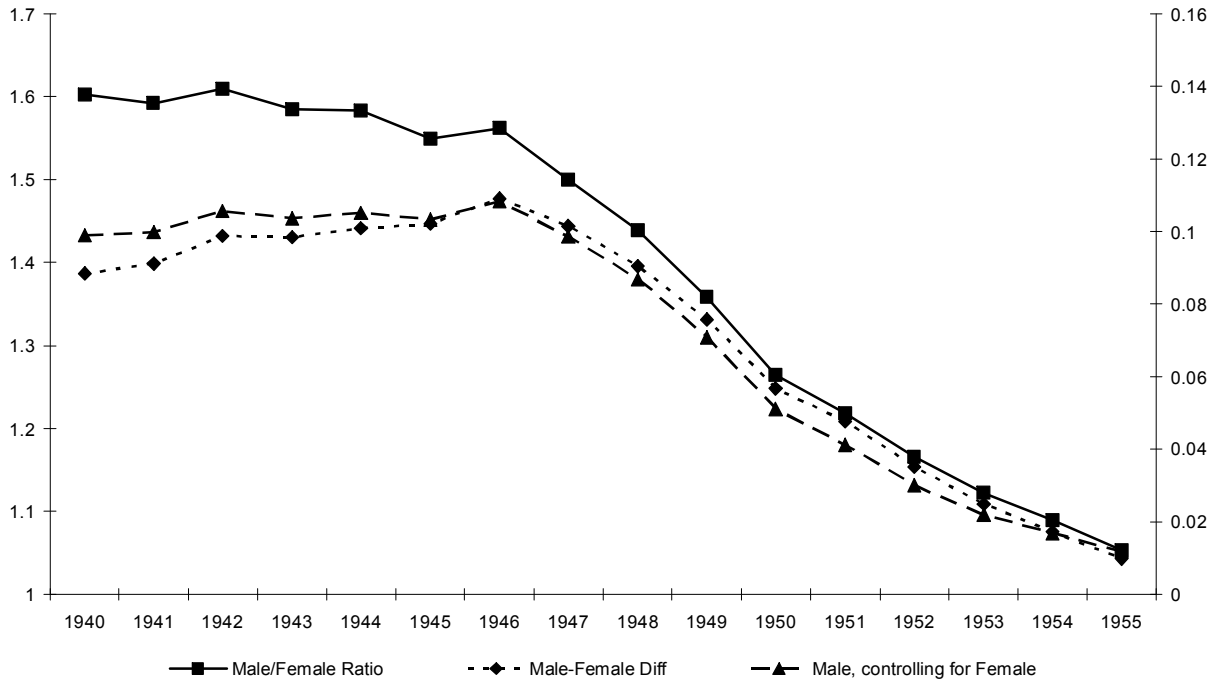
Notes: Each point represents an unadjusted mean from the 1980 Census among U.S. born birth cohorts 1940–1955. Male and female mobility indicates the fraction residing outside their birth state among males and females respectively. Induction risk represents the Card and Lemieux measure of induction risk, while national induction risk is the portion of national cohort risk that is “left over” after controlling for an individual’s state cohort risk. See the text for more details.

Figure 5: Vietnam War induction risk for selected states



Notes: Each line indicates induction risk over time for one of the following states: AK – Alaska, AR – Arkansas, CO – Colorado, IL – Illinois, KY – Kentucky, LA – Louisiana, MA – Massachusetts, ND – North Dakota, PA – Pennsylvania, as well as the national induction risk across all states.

Appendix Figure 1: Alternative Specifications for the Male-Female College Graduation Rate



Notes: The pattern of male-female graduation trends is shown using male-female ratios, differences, and residuals after controlling for a female trend. The scale on the left of the y-axis corresponds to the ratio, while the scale on the right of the y-axis corresponds to the differences and residuals.