

UCSF

UC San Francisco Previously Published Works

Title

Does the Type and Timing of Educational Attainment Influence Physical Health? A Novel Application of Sequence Analysis

Permalink

<https://escholarship.org/uc/item/6391r21x>

Journal

American Journal of Epidemiology, 189(11)

ISSN

0002-9262

Authors

Vable, Anusha M

Duarte, Catherine dP

Cohen, Alison K

et al.

Publication Date

2020-11-02

DOI

10.1093/aje/kwaa150

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

## Original Contribution

# Does the Type and Timing of Educational Attainment Influence Physical Health? A Novel Application of Sequence Analysis

**Anusha M. Vable\***, Catherine dP. Duarte, Alison K. Cohen, M. Maria Glymour, Robert K. Ream, and Irene H. Yen

\* Correspondence to Dr. Anusha M. Vable, Department of Family and Community Medicine, Zuckerberg San Francisco General Hospital, Building 80, Ward 83, 995 Potrero Avenue, San Francisco, CA 94110 (email: anusha.vable@ucsf.edu).

*Initially submitted July 25, 2019; accepted for publication July 13, 2020.*

Nontraditional education trajectories are common, but their influence on physical health is understudied. We constructed year-by-year education trajectories for 7,501 National Longitudinal Survey of Youth 1979 participants aged 14 to 48 years (262,535 person-years of education data from 1979 to 2014). We characterized trajectory similarity using sequence analysis and used hierarchical clustering to group similar educational trajectories. Using linear regression, we predicted physical health summary scores of the participants at age 50 years from the 12-item Short-Form Survey, adjusting for available confounders, and evaluated effect modification by sex, race/ethnicity, and childhood socioeconomic status. We identified 24 unique educational sequence clusters on the basis of highest level of schooling and attendance timing. General education development credentials predicted poorer health than did high school diplomas ( $\beta = -3.07$ , 95% confidence interval:  $-4.07, -2.07$ ), and bachelor's degrees attained at earlier ages predicted better health than the same degree attained at later ages ( $\beta = 1.66$ , 95% confidence interval:  $0.05, 3.28$ ). Structurally marginalized groups benefited more from some educational trajectories than did advantaged groups (e.g., Black vs. White Americans with some college; those of low vs. high childhood socioeconomic status who received an associate's or bachelor's degree). Both type and timing of educational credentials may influence physical health. Literature to date has likely underestimated the impact of educational trajectories on health.

education; educational credential; educational timing; educational trajectories; sequence analysis

Abbreviations: AA, associate's degree; AFQT, Armed Forces Qualification Test; BA, bachelor's degree; cSES, childhood socioeconomic status; GED, general education development; HS, high school.

Educational attainment predicts better health in nearly every domain (1–4). The majority of health research operationalizes education using simple measures of either final credential (e.g., less than high school (HS) vs. HS degree or higher level of education) or total years of schooling. However, given the numerous pathways through which education is likely to influence health—including knowledge, health behaviors, financial and occupational returns, prestige, social networks, and social norms—additional aspects of educational experiences may be relevant (5, 6). Despite extensive research on education as a social determinant of health, little is known about the health effects of factors such as the timing of educational experiences (7). Educational trajectories with periods when schooling was interrupted (e.g.,

started college, then took several years off before completing degree), or delayed (e.g., several-year interlude between HS completion and attending college) are often regarded as nontraditional, but they are increasingly common (8, 9). Differences in education timing may have relevance for financial benefits, social networks and norms, or opportunity costs associated with schooling.

Prior work suggests delayed or interrupted trajectories predict lower educational attainment and degree completion (8, 10, 11); however, little work has evaluated whether these educational trajectories adversely affect health outcomes. Furthermore, women, racially minoritized people, and individuals who are socioeconomically disenfranchised are more likely to follow nontraditional educational trajectories

(10–17); evaluating the effects of educational timing, therefore, is of special relevance to health inequities.

The literature on type of educational credential and health indicate the general education development (GED) credential, although assumed to confer equivalent skills, predicts poorer health than does a HS diploma in several dimensions, including depression, cigarette smoking, cognition, physical illness, and assessments of general health (18–25). Notably, work on GED credentials and health is primarily cross-sectional, resulting in ambiguous temporal ordering, because a GED credential can be attained at later ages. Studies using longitudinal data, in which GED credential attainment precedes assessment of health outcomes, have not examined whether GED credentials attained at later ages confer the same health benefits as those attained at earlier ages.

Women, racially minoritized people, and individuals who are socioeconomically disenfranchised, among others, may benefit more from education than might structurally advantaged groups from mechanisms such as increased access to social, economic, and political agency that higher education affords (26–30). However, differential relationships between timing of educational attainment or type of educational credential by sex, race/ethnicity, or childhood SES (cSES) are understudied.

A challenge to understanding the relationship between educational trajectories and health is characterizing the innumerable possible educational trajectories individuals follow across the life course into a smaller, analytically tractable number of exposure groups while retaining the substantive differences between different trajectories. In this article, we address this problem using sequence analysis and cluster analysis to characterize the type and timing of educational experiences and evaluate the association of educational trajectories with general measures of physical health. As part of our examination of the relationship between educational trajectories and health, we evaluate a broad measure of health, the physical health component summary score from the 12-item Short-Form Survey, which is predictive of death and hospitalizations (31). We also evaluate if the relationships between education and physical health vary by sex, race/ethnicity, and cSES.

## METHODS

### Sample

National Longitudinal Survey of Youth 1979 participants (aged 14–22 years at study initiation) were interviewed annually until 1994 and biennially thereafter. We constructed education trajectories for each participant from age 14 years through 48 years ( $n = 7,912$ ). For individuals who entered the study after age 14 years, we imputed their educational trajectory from age 14 years until study entry, as detailed in the data-cleaning rules explained in Web Table 1 (available at <https://doi.org/10.1093/aje/kwaa150>). After excluding 248 individuals with missing outcome data and 163 with missing data on 1 or more covariates, our final analytic sample consisted of 7,501 participants (94.8%). Web Table 2 shows how included individuals differed from the full cohort.

### Exposure

Because few individuals have identical educational trajectories, it is not meaningful to evaluate distinct associations between each unique educational trajectory and physical health. We sought a principled approach to quantify and collapse similar sequences. We used the following 3 steps to reduce the thousands of possible educational trajectories from ages 14 to 48 years into an analytically tractable number of similar sequences: 1) creation of individual educational trajectories; 2) sequence analysis to quantitatively evaluate how trajectories differed; and 3) cluster analysis to group similar trajectories.

*Creation of individual education trajectories* Each year of respondents' lives from age 14 years to 48 years was classified into 1) enrolled in HS; 2) not completed HS and not enrolled in HS; 3) HS graduate, not enrolled in higher education; 4) completed a GED credential and not enrolled in higher education; 5) completed HS or a GED credential and enrolled in higher education full-time; 6) completed HS or GED credential and enrolled in higher education part-time; 7) completed some college but not a college degree and not enrolled in formal schooling; 8) completed an AA and not enrolled in higher education; 9) completed a bachelor's degree (BA) and not enrolled in higher education; 10) enrolled in or completed graduate school (we collapsed across post-BA categories, because of data sparseness).

*Sequence analysis: quantitative evaluation of how individual trajectories differed* We calculated distances between every pair of education trajectories observed (32) by quantifying the “cost” to transform 1 education trajectory to another via substitutions (i.e., changing 1 state to another state; e.g., enrolled in HS to HS graduate) or insertions or deletions (i.e., inserting or deleting a state to make 2 sequences equivalent). We used the mean probability distance to calculate costs, meaning more common transitions (e.g., enrolled in HS to HS graduate) were assigned lower costs than less frequent transitions (e.g., GED credential to enrolled in college full-time). We used the Halpin optimal matching algorithm (33), which accounts for amount of time in each state (rather than treating each time point as discrete), meaning if a respondent spent longer in a state, transition to a different state had higher costs; this approach differentiated those whose schooling was interrupted for a few versus several years. We calculated distances between every possible pair of trajectories in the data set; the result was a square, symmetric distance matrix for all unique educational trajectories.

As a robustness check, we also calculated costs using dynamic hamming, an approach that applies lower substitutions costs when transitions are more frequent (e.g., ages 18–25 years) and does not include insertion or deletion costs. We preferred the Halpin approach because dynamic hamming often failed to capture increases in education that happen later in life (see Web Figure 1 for dynamic hamming modal plot).

*Cluster analysis: grouping similar trajectories* We used cluster analysis to group similar education trajectories together

on the basis of the distance matrix. We used agglomerative clustering (all unique trajectories were initially considered distinct clusters, then similar trajectories—based on the distance matrix—were sequentially clustered together until only 1 cluster remained), with a Wards linkage to calculate distance between clusters. The Wards linkage maximizes similarities within clusters (i.e., minimizes the mean distance within the cluster) and maximizes mean differences between clusters. We used the Duda-Hart cluster-stopping rule, which maximizes distinct cluster structure (34), to help determine the number of clusters. Others have suggested cluster-stopping rules may result in a number of clusters that are either too detailed or not detailed enough for analytic purposes (35). Of particular concern is the potential to miss important differences between educational groups if power were reduced by subdividing a group into multiple smaller subgroups; to address this concern, we retained cluster solutions that produced more clusters, then examined whether physical health associations were similar for the same terminal state using F-tests; application of the F-test is detailed in the Analysis section of Methods.

## Outcome

The outcome for these analyses was the physical health summary score from the 12-item Short-Form Survey assessed at age 50 years. This self-reported measure is used to assess whether physical health limits moderate activities (e.g., moving a table, climbing stairs), pain interfered with work, or activities were limited because of physical health; poorer performance on the physical health summary score is predictive of death and hospitalizations (31). The measure has demonstrated good reliability (2-week test-retest reliability  $\geq 0.89$ ) and validity (36). The National Longitudinal Survey of Youth, 1979 cohort standardized these data to a mean of 50 and a standard deviation of 10, such that a score of 50 corresponds with the US average, and a 1-point difference is a tenth of a standard deviation (37).

## Effect modifiers

We evaluated self-reported sex (male vs. female), race/ethnicity (White, Black, Latino, “other race/ethnicity”), and cSES (high, low) as effect modifiers. Childhood SES was operationalized as mother’s education being less than 12 years or as missing data versus 12 years or more of education in primary analyses, consistent with prior literature (27, 28, 30) and because there were fewer missing data for mother’s education than father’s education (6.4% of mothers vs. 14.9% of fathers). Sensitivity analyses examined cSES operationalized as father’s education being less than 12 years or as missing data versus 12 years or more of education. Because of small numbers and ambiguous interpretation, results for the “other race/ethnicity” category are neither presented nor discussed.

## Confounders

We included the following potential confounders: birth year, race/ethnicity (i.e., White, Black, Latino, Other), birth

in a southern state (38), birth outside of the United States, rural residence at age 14 years, mother’s and father’s education levels (centered at 12 years), indicators for unknown or missing values of mother’s and father’s education ( $n = 436$ , 6.4% of mothers;  $n = 1,013$ , 14.9% of fathers), and the following variables for both parents when the respondent was age 14 years: work for pay (yes/no), missing indicator for parents worked for pay, occupation skilled versus unskilled (dichotomized at 300 using 1970 3-digit census occupation codes), and missing indicator for skilled occupation. We did not include potential confounders that occurred after age 14 years in primary analyses, because they are potential mediators of the education and health relationship; adjusting for mediators may bias estimates (39).

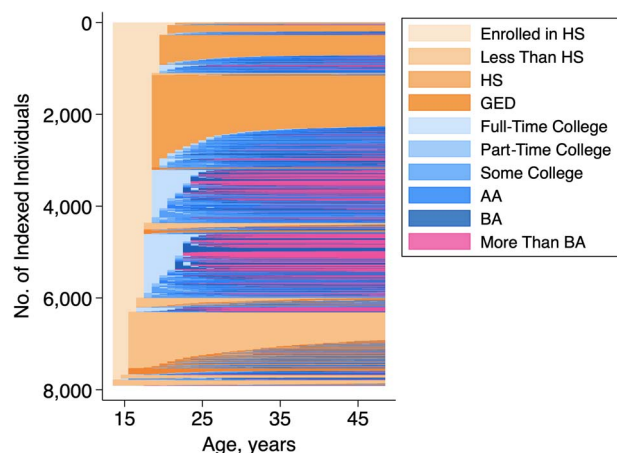
In additional analyses, we adjusted for confounders that were measured after age 14 years, the first age used to construct the educational trajectories. These confounders assessed if health either prevented the respondent from working or limited the type of work the respondent could engage in (as a proxy for childhood health, because direct measures of childhood health were not collected; data were from baseline or the first wave available), poverty status (assessed in 1979, the year of study entry; therefore, poverty was assessed after age 14 years for the respondents older than 14 years at study entry), a missing indicator for poverty status, and the Armed Forces Qualification Test (AFQT) score (administered in 1981), which is purported to reflect aptitude (40), and a missing indicator for the AFQT (critiques of the AFQT argue it actually measures characteristics such as wealth or English language proficiency) (41–43). Because of the ambiguous temporal ordering of these variables compared with the exposure (educational trajectories from 14 to 48 years), results adjusted for the childhood health proxy, poverty, and the AFQT are reported in the Web material.

## Analysis

We used linear regression to determine how different educational sequences predicted physical health at age 50 years, adjusted for previously listed confounders. For any given terminal credential (e.g., HS, AA, BA), we tested if physical health was different based on the timing of degree completion, using the F-test to compare multiple point estimates; if point estimates for the same terminal state were similar ( $P > 0.2$ ), we collapsed across sequences to improve interpretability and our ability to detect meaningful differences (e.g., if all sequences to attain a GED credential were similar based on  $F > 0.2$  in predicting physical health, we collapsed across GED credentials sequences). As a robustness check, we also examined cluster solutions, using the dynamic hamming approach to calculate costs.

Effect modification was assessed using interaction terms to determine if relationships between education trajectories and physical health were meaningfully different. We also conducted analyses stratified by the potential effect modifiers for presentation in the figures; stratification variables were not included as confounders.

Coding of data-cleaning rules, sequence-analysis implementation, and outcome regression were reviewed by a



**Figure 1.** Individual-level educational trajectories represented in the National Longitudinal Survey of Youth 1979 cohort (1979–2016) from ages 14 to 48 years. Each individual is a row on the y-axis; each year of life from ages 14 to 48 years is represented on the x-axis. There were 7,912 individuals followed until age 48 years and included in the sequence analysis. We categorized each year of life from age 14 to 48 years ( $n = 35$  years) into 1 of 10 mutually exclusive states. Of a total possible  $10^{35}$  educational trajectories, there were 3,185 unique educational trajectories. Our challenge was converting these thousands of individual-level trajectories into a meaningful and interpretable number of similar educational sequences. AA, associate's degree; BA, bachelor's degree; GED, general educational development; HS, high school.

second coder. All analyses were weighted to be nationally representative, and all standard errors were adjusted for the clustered sampling design of the National Longitudinal Survey of Youth, 1979 cohort. Data cleaning and analyses were performed using Stata (StataCorp LP, College Station, Texas) (15).

## RESULTS

### Educational sequences

Using sequence analysis and cluster analysis, we collapsed the 3,185 unique, individual-level educational trajectories (Figure 1) into 24 clusters (i.e., sequences; Web Figure 1). We refer to these clusters by both the terminal degree and the timing of degree completion, as follows: immediate, indicating continuous enrollment in education until terminal degree attainment; interrupted, indicating the respondent started schooling toward a degree but experienced a period out of school before completing the degree; or delayed, indicating several years with 1 credential before initiating schooling for another credential. After collapsing sequences within the same terminal state that were substantively similar, based on F-tests, we retained 10 sequences: 1) less than a HS education; 2) HS education; 3) GED credential; 4) some college after a GED credential; 5) some college (all other trajectories); 6) AA after less than a HS education; 7) AA (all other trajectories); 8) BA immediate; 9) BA interrupted; and 10) graduate school (Figure 2). Women, White

Americans, those from higher cSES backgrounds, those with higher AFQT results, and those from lower poverty backgrounds were better represented in higher education trajectories (Table 1).

### Overall results

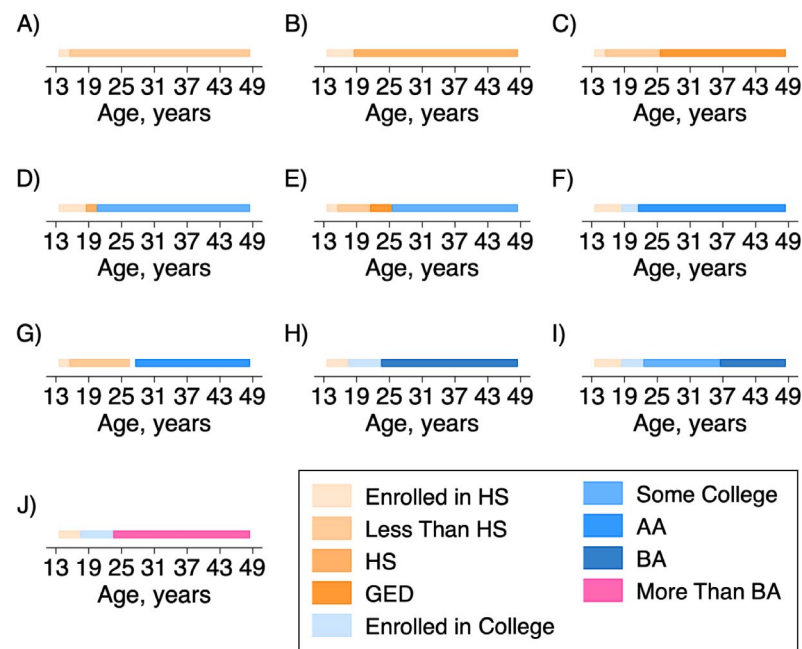
Higher educational attainment predicted better physical health (Figure 3). Compared with those with a HS diploma and no more schooling (the most common sequence), the following groups reported poorer physical health: less than HS ( $\beta = -4.69$ , 95% CI:  $-6.10, -3.28$ ); GED credential ( $\beta = -3.07$ , 95% CI:  $-4.07, -2.07$ ); some college after a GED credential ( $\beta = -3.37$ , 95% CI:  $-5.50, -1.24$ ), and those who completed an AA after less than a HS level of education ( $\beta = -6.94$ , 95% CI:  $-11.56, -2.31$ ). Compared with those with a HS diploma, the following groups reported better physical health: BA immediate ( $\beta = 2.64$ , 95% CI: 1.90, 3.38) and graduate school ( $\beta = 2.68$ , 95% CI: 1.84, 3.51). Those with a GED credential reported slightly better physical health than those with less than a HS education ( $\beta = 1.62$ , 95% CI:  $-0.09, 3.34$ ). For context, a 5-unit difference in the physical health component summary score, approximately the difference between less than a HS education and HS, is equivalent to aging between 10 and 20 years (44, 45).

Physical health at age 50 years differed between people who took different paths to the same final credential. For example, those who completed some college after a GED credential reported worse physical health than all other "some college" sequences ( $\beta = -2.78$ , 95% CI:  $-4.91, -0.66$ ); similarly, those who completed an AA after having less than a HS level of education reported worse physical health than those who completed other AA trajectories ( $\beta = -7.63$ , 95% CI:  $-12.29, -2.97$ ). Respondents who received a BA immediately reported better physical health at age 50 years than those whose BA trajectory was interrupted ( $\beta = 1.66$ , 95% CI: 0.05, 3.28).

### Effect modification

Sex modified the relationship between educational sequences and physical health (Figure 4) such that some college predicted poorer health for women than men ( $n = 715$  women;  $\beta = -1.41$ , 95% CI:  $-2.95, 0.13$ ). Race modified the relationship between educational trajectories and physical health (Figure 5) such that some college predicted better health for Black Americans than for White American ( $n = 473$  Black Americans;  $\beta = 1.90$ , 95% CI: 0.10, 3.71); relationships were in the same direction for Black Americans with less than a HS education ( $n = 252$ ;  $\beta = 2.62$ ; 95% CI:  $-0.16, 5.41$ ), a GED credential ( $n = 330$ ;  $\beta = 1.99$ , 95% CI:  $-0.05, 4.03$ ), and graduate school ( $n = 216$ ;  $\beta = 1.50$ , 95% CI:  $-0.24, 3.24$ ), Latino Americans with some college after a GED credential ( $n = 201$ ;  $\beta = 3.72$ , 95% CI:  $-0.64, 8.13$ ), and AA after less than a HS education ( $n = 27$ ;  $\beta = -7.47$ , 95% CI:  $-0.26, 15.21$ ), although confidence intervals just included the null. Childhood SES modified the relationship between educational trajectories and physical health (Figure 6) such that respondents with low cSES who attained





**Figure 2.** Visual representation of the 10 educational sequences from ages 14 to 48 years used to predict physical health at age 50 years in the National Longitudinal Survey of Youth 1979 cohort (1979–2016). We collapsed the 24 educational sequences from the cluster analysis (Web Figure 1) based on F-test results indicating point estimates were similar for the same terminal educational state (e.g., the 4 GED trajectories were not different, so they were collapsed into 1 GED trajectory). The result was 10 educational sequences: A) less than HS; B) HS; C) GED credential; D) some college; E) some college after a GED credential; F) AA; G) AA after less than a HS education; H) BA immediate; I) BA interrupted; and J) graduate school. AA, associate's degree; BA, bachelor's degree; GED, general educational development; HS, high school.

an AA ( $n = 418$ ;  $\beta = 2.51$ , 95% CI: 0.75, 4.28) or BA immediately ( $n = 169$ ;  $\beta = 1.94$ , 95% CI: 0.26, 3.61) reported better health at age 50 years than respondents who had high cSES.

### Robustness checks

Point estimates were in the same direction and our main findings were unchanged using all 24 educational trajectories recommended by cluster analysis (Web Figure 2), using clusters from the dynamic hamming approach to calculate costs (Web Figure 3), after additional adjustment for the childhood health proxy (Web Table 3), poverty and AFQT (Web Figure 4), with cSES operationalized as father's education (Web Figure 5), and after multiple imputations (Web Table 5).

### DISCUSSION

Higher educational attainment generally predicted better physical health at age 50 years among National Longitudinal Survey of Youth 1979 participants; however, health depended not only on the final credential but on credential type, timing, and the trajectory taken to attain the final credential. Respondents who attained GED credentials averaged poorer physical health at age 50 years than those with a HS diploma. Education completed at younger ages tended to confer greater health benefits than education completed

later in life, and individuals who exited high school without a diploma reported poorer physical health in adulthood than those who completed HS, even if they went on for additional schooling later in life; these relationships were only identifiable because of our use of sequence analysis. Relationships varied by sex, race/ethnicity, and cSES: racially minoritized participants benefitted more from educational attainment than structurally advantaged groups (i.e., White Americans and those from high cSES backgrounds), even if they followed nontraditional trajectories.

Our results, consistent with prior evidence (18–20, 25), suggest health researchers should not conflate GED credentials with HS diplomas. Recipients of a GED credential often have similar health as those who did not complete HS; potential explanations suggest the GED credential may confer less social capital (46) or may be differentially valued in the labor market (47). Major public health surveillance studies, such as the National Health and Nutrition Examination Survey and the Behavioral Risk Factor Surveillance Surveys, either have historically (the former) or continue (the latter) to categorize HS diplomas and GED credentials together (18). Aggregating the GED and high school diploma credentials under the assumption that they confer equivalent benefits masks important heterogeneities.

Our work also demonstrates the relevance of timing: educational credentials attained at later ages appear to offer fewer health benefits than credentials attained earlier. In predicting physical health, those whose educational trajectory to a BA was interrupted had poorer physical health than

**Table 1.** Distribution of Covariates Arranged by Educational Trajectories in the National Longitudinal Survey of Youth 1979 Cohort (n = 7,912), 1979–2016

Characteristic	<HS (n = 823)	HS (n = 1,745)	GED (n = 898)	Some College After GED (n = 209)	Some College (n = 1,370)	AA After < HS (n = 95)	AA (n = 1,054)	BA Imme- diate (n = 810)	BA Inter- rupted (n = 168)	Graduate School (n = 740)
Birth year (linearized SE)	1960; 35 (0.10)	1960; 28 (0.07)	1960; 81 (0.10)	1960; 35 (0.18)	1960; 41 (0.08)	1960; 57 (0.31)	1960; 36 (0.09)	1960; 53 (0.10)	1960; 68 (0.19)	1960; 59 (0.10)
Female sex <sup>a</sup>	0.44	0.49	0.47	0.57	0.52	0.62	0.59	0.5	0.68	0.53
White (non-Hispanic) <sup>a</sup>	0.59	0.77	0.61	0.73	0.7	0.55	0.72	0.79	0.72	0.84
Black (non-Hispanic)	0.21	0.14	0.24	0.15	0.18	0.21	0.16	0.09	0.14	0.07
Latino	0.13	0.04	0.09	0.08	0.05	0.12	0.06	0.02	0.05	0.02
Other race	0.06	0.05	0.06	0.03	0.04	0.06	0.05	0.08	0.07	0.06
Mother's education (linearized SE)	8.18 (0.23)	10.29 (0.11)	9.24 (0.18)	9.71 (0.38)	11.13 (0.13)	9.09 (0.78)	11.45 (0.13)	12.70 (0.13)	11.71 (0.21)	13.27 (0.15)
Mother's education data missing	0.12	0.06	0.09	0.09	0.05	0.17	0.04	0.02	0.02	0.03
Father's education (linearized SE)	6.94 (0.29)	9.72 (0.17)	7.93 (0.22)	9.28 (0.42)	10.85 (0.18)	8.14 (0.82)	11.03 (0.20)	13.42 (0.20)	11.28 (0.39)	13.66 (0.25)
Father's education data missing	0.23	0.09	0.19	0.12	0.08	0.26	0.09	0.04	0.06	0.05
Southern-state birth	0.45	0.32	0.4	0.36	0.31	0.36	0.28	0.29	0.27	0.27
Birth outside the United States	0.07	0.02	0.07	0.07	0.05	0.11	0.04	0.03	0.03	0.04
Rural residence at age 14	0.24	0.28	0.21	0.17	0.21	0.17	0.24	0.19	0.2	0.16
Mother worked for pay at age 14 years	0.47	0.5	0.51	0.53	0.52	0.56	0.54	0.52	0.62	0.57
Mother's work-for-pay status at age 14 years missing	0.05	0.02	0.05	0.03	0.01	0.04	0.02	0.01	0	0.02
Mother's job skilled	0.57	0.57	0.55	0.57	0.58	0.53	0.58	0.67	0.55	0.68
Mother's job unskilled	0.39	0.4	0.41	0.41	0.39	0.45	0.4	0.31	0.43	0.3
Mother's job-skill information missing	0.03	0.03	0.04	0.03	0.03	0.02	0.02	0.02	0.01	0.02
Father worked for pay at age 14 years	0.67	0.85	0.73	0.74	0.83	0.69	0.83	0.89	0.78	0.87
Father's work-for-pay status at age 14 years missing	0.22	0.1	0.21	0.19	0.13	0.19	0.12	0.09	0.17	0.1

Table continues

Table 1. Continued

Characteristic	< HS (n = 823)	HS (n = 1,745)	GED (n = 898)	Some College After GED (n = 209)	Some College (n = 1,370)	AA After < HS (n = 95)	AA (n = 1,054)	BA Imme- diate (n = 810)	BA Inter- rupted (n = 168)	Graduate School (n = 740)
Father's job unskilled	0.55	0.64	0.58	0.61	0.54	0.41	0.47	0.32	0.44	0.28
AFQT score (linearized SE)	16.69 (0.90)	36.54 (0.97)	25.22 (1.01)	35.64 (1.63)	47.57 (1.19)	41.63 (3.32)	52.33 (1.16)	71.18 (0.96)	57.12 (2.21)	76.23 (1.01)
AFQT score missing	0.06	0.05	0.07	0.1	0.04	0.21	0.03	0.03	0.03	0.02
Poverty <sup>a</sup>	0.22	0.09	0.18	0.16	0.1	0.17	0.08	0.06	0.09	0.04
Poverty information missing	0.22	0.09	0.18	0.16	0.1	0.17	0.08	0.06	0.09	0.04
Mother's education low	0.72	0.45	0.6	0.61	0.31	0.54	0.29	0.14	0.23	0.11
Father's education low	0.78	0.49	0.66	0.53	0.35	0.49	0.35	0.16	0.34	0.16
Physical health at age 50 years (linearized SE)	44.50 (0.68)	49.66 (0.28)	46.35 (0.50)	46.06 (1.01)	49.29 (0.42)	42.38 (2.29)	50.52 (0.40)	53.02 (0.29)	50.77 (0.85)	53.10 (0.31)

Abbreviations: < HS, less than a high school education; AA, associate's degree; AFQT, Armed Forces Qualification Test; BA, bachelor's degree; GED, general education development; HS, high school; SE, standard error.

<sup>a</sup> Women, White Americans, and those from higher childhood socioeconomic status backgrounds were better represented in the higher educational attainment trajectories.



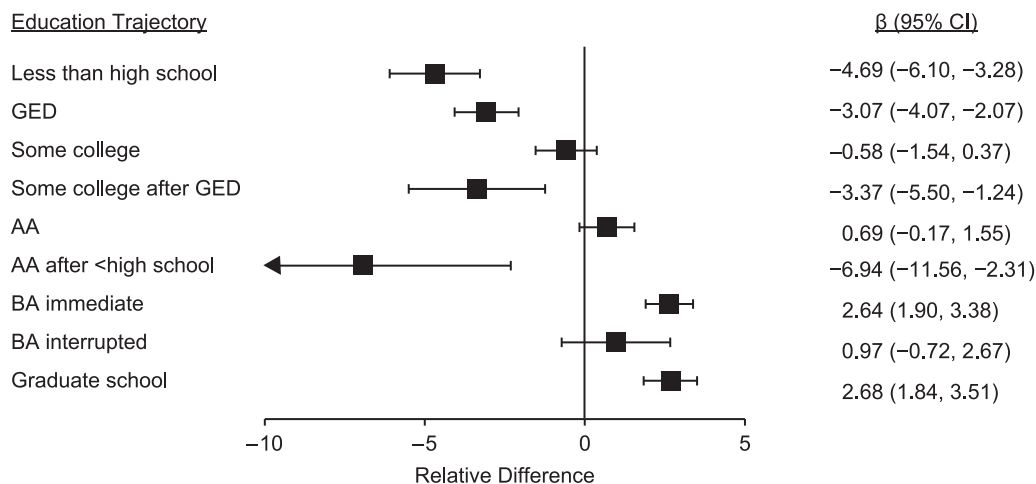
those who received a BA immediately. This is consistent with the very limited prior work in this area (48). We also found that the educational trajectory to the final credential matters, such that those who drop out of high school do not attain health parity with those who have HS diplomas, even if they attain post-HS education. These results suggest the conventional approach of operationalizing education on the basis of final credential obscures meaningful heterogeneities in educational trajectories.

Future research should explore why timing matters. Delayed receipt of a credential may be a marker of marginalization: individuals who complete a BA at later ages may come from more structurally marginalized backgrounds compared with those who complete a BA directly after HS, creating confounding. The credential itself may be valued less in the labor market when received later, or there may be high opportunity costs to pursuing additional schooling later in adulthood. Individuals who return to school after an interruption may juggle additional demands, such as child care or paid work, making the process of schooling itself particularly aversive (49). The opportunity cost of degree completion later in adulthood may offset any physical health advantage conferred by the degree. Alternatively, respondents who had interrupted schooling before attaining a BA may not have been followed long enough to present health benefits but may manifest advantages later in life.

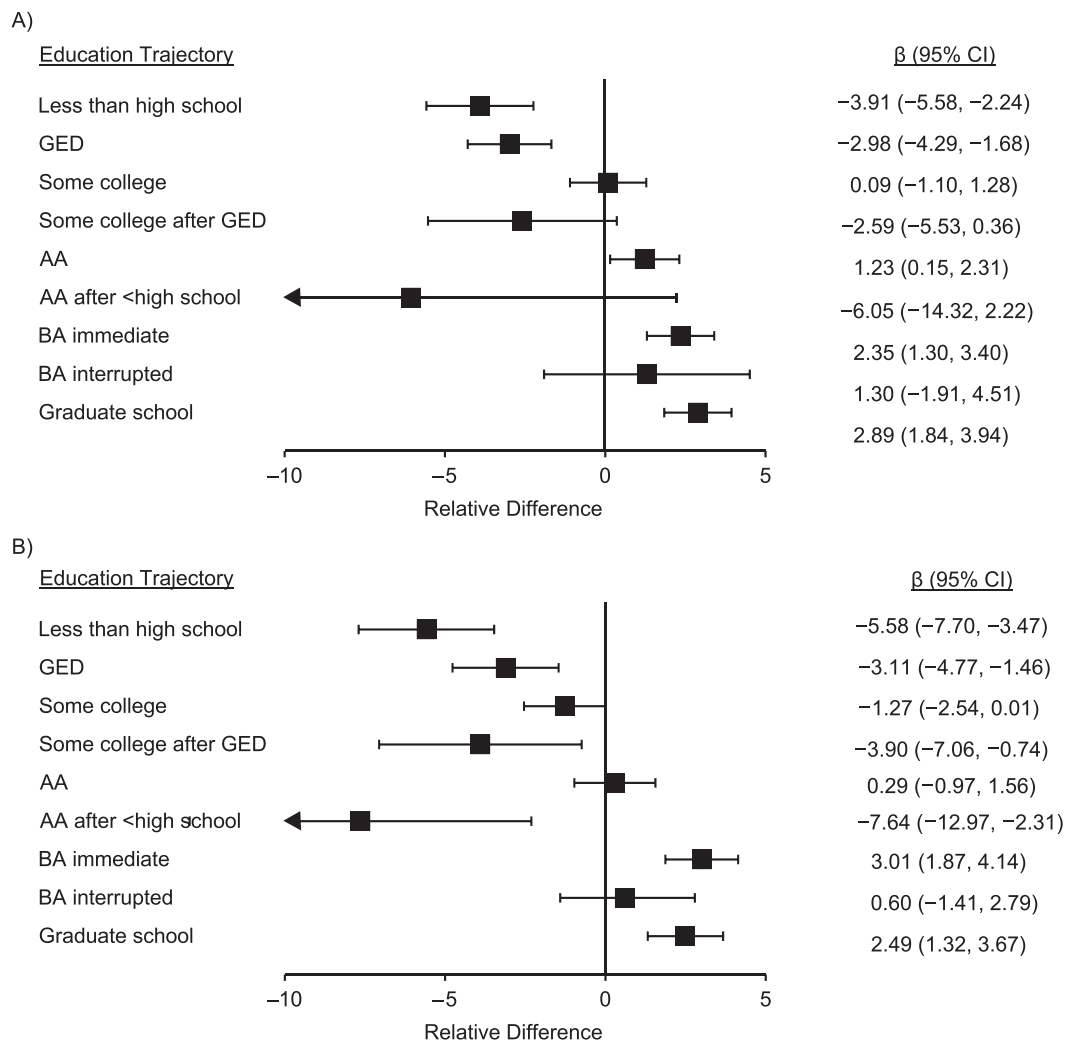
With the exception of Black Americans, no groups benefited from completing some college if it did not culminate in a degree, suggesting that educational initiatives that recruit individuals into college without providing support for degree completion are unlikely to benefit health. Future

work should evaluate if these findings are causal (e.g., completing some college results in the accrual of debt without an increase in salary, harming health) or are due to confounding (e.g., structurally marginalized groups or those with lower skills begin but do not complete college). The finding that Black Americans benefitted from some college compared with getting a HS diploma, whereas all other groups either did not benefit or were harmed, may speak to the uniquely minoritized experience of Black Americans in American society. That is, because Black Americans are uniquely minoritized (e.g., targeted by anti-Black racism), the agency and access conferred by “some college” trajectories may provide consequential benefits, though still insufficient to attain the health returns of degree completion (27, 50). Chance is an alternative explanation for these findings.

Across the range of educational trajectories, the health benefits of education are especially pronounced among structurally marginalized groups, such as racially minoritized participants and those from low cSES backgrounds compared with socially advantaged participants. This was true even when participants followed nontraditional trajectories. Our results add to a growing body of literature suggesting structurally marginalized groups benefit more from education than do socially advantaged groups and suggest education may be a means to increase access to the social, economic, and political agency from which structurally marginalized populations have been systematically excluded (26–30). Thus, policies and programs to increase educational attainment among these groups in particular, such as the Pell Grant program (51) or affirmative action



**Figure 3.** Coefficient plots for the association between educational trajectories from ages 14 to 48 years and physical health at age 50 years compared with respondents who completed high school only in the National Longitudinal Survey of Youth 1979 cohort (1979–2016). Overall, physical health was patterned such that higher educational attainment predicted better health. Physical health varied by type of educational credential such that GED credentials did not confer the same health benefits as a high school diploma. Physical health varied by trajectory to final credential such that those who followed most some college trajectories had better physical health than those who received some college after a GED credential; similarly, those who followed most AA trajectories had better physical health than those who received an AA after less than a high school education. Physical health also varied by timing of educational attainment such that those who received a BA immediately had better physical health at age 50 years than those whose BA trajectory was interrupted. Results are adjusted for birth cohort, race, birth in a southern state, birth outside the United States, rural residence at age 14 years, parental education, if parents worked for pay, and parental occupation. AA, associate’s degree; BA, bachelor’s degree; CI, confidence interval; GED, general educational development.



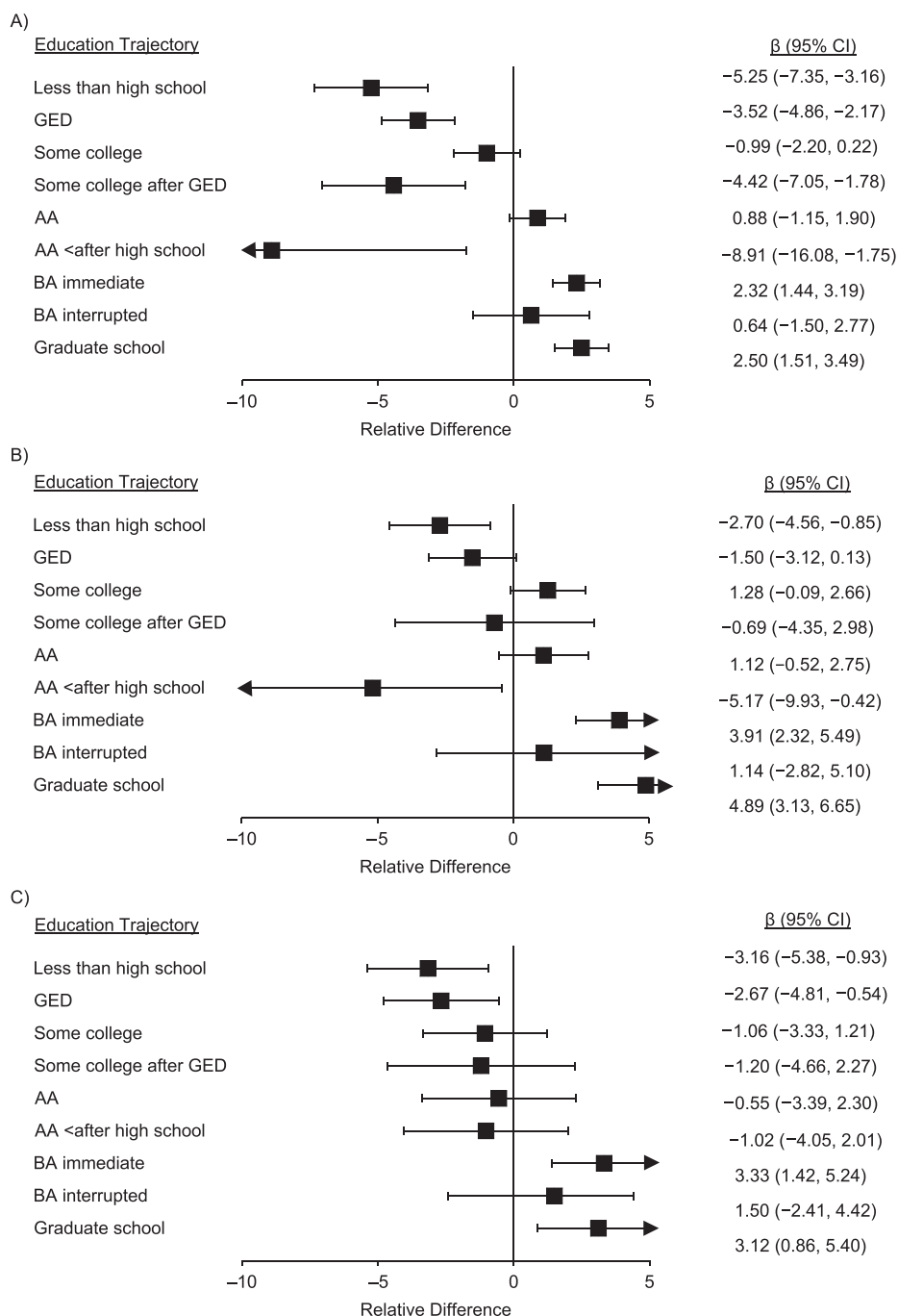
**Figure 4.** Coefficient plots for the association between educational trajectories from ages 14 to 48 years and physical health at age 50 years compared with respondents who completed high school only among A) men and B) women in the National Longitudinal Survey of Youth 1979 cohort (1979–2016). Stratified results are presented. Having some college predicted poorer physical health at age 50 years for women than men; all other relationships between educational trajectories and physical health did not meaningfully differ by sex. Results are adjusted for birth cohort, race, birth in a southern state, birth outside the United States, rural residence at age 14 years, parental education, if parents worked for pay, and parental occupation. AA, associate's degree; BA, bachelor's degree; CI, confidence interval; GED, general educational development.

programs (52, 53), could be powerful mechanisms to reduce racial and socioeconomic health inequities.

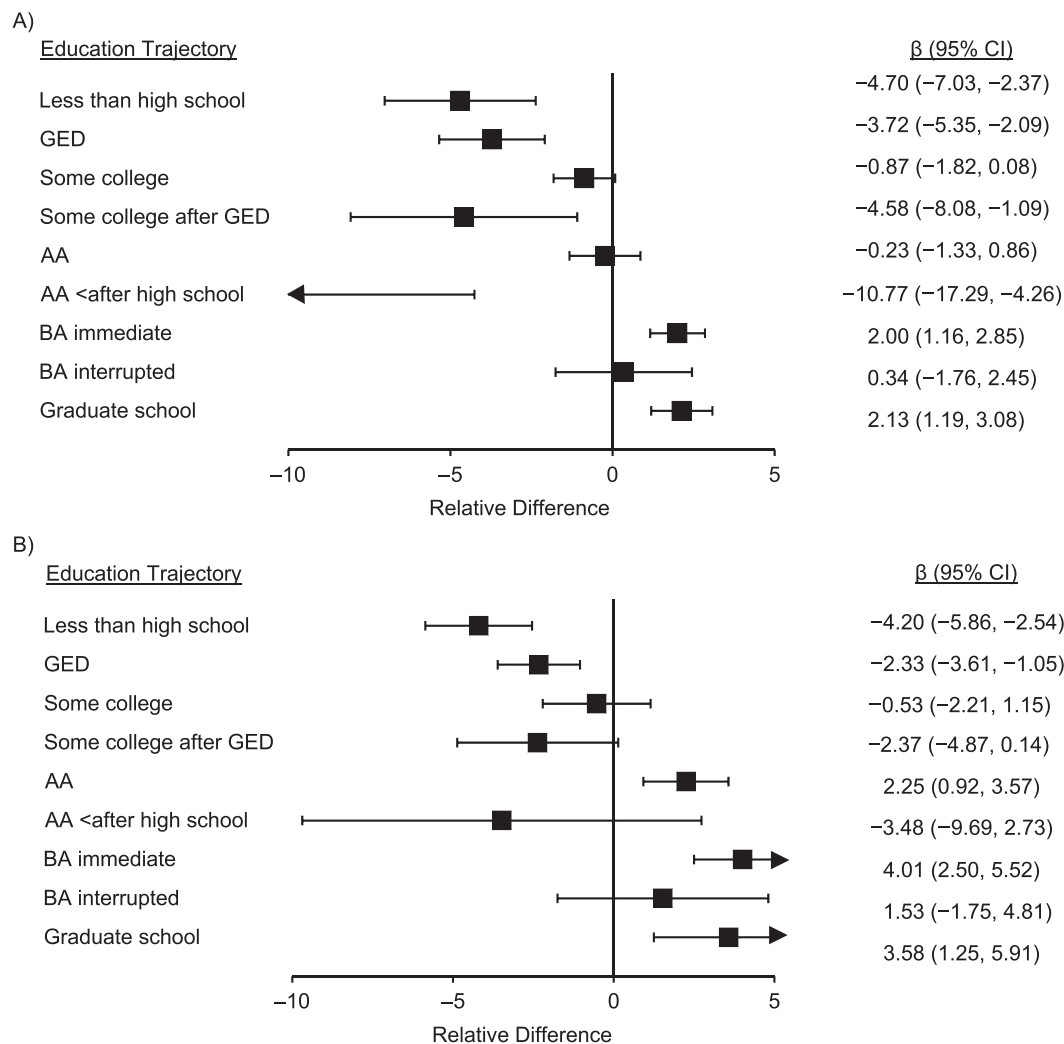
Nearly all health outcomes are strongly patterned by educational attainment, a relationship that has remained robust across time, disease conditions, and populations (54). Huge investments in increasing educational attainment have been made in the United States and globally over the past century. These investments may reduce inequities, but our findings suggest that to redress inequity, we need a better understanding of what types of educational experiences improve health. This is especially important for developing structural interventions to improve the health of women, racially minoritized groups, or first-generation college students who are more likely to experience nontraditional educational trajectories (12–16).

There are limitations to these analyses. All data are self-reported. The available health measures may not be sensitive to health conditions that have not yet developed, remain asymptomatic at age 50 years, or are undiagnosed; however, finding meaningful differences in this relatively young and healthy cohort suggests health differentials could become even wider with age. Residual confounding is a concern in this observational analysis; although our substantive research questions are clearly causal, strong, untestable assumptions are needed for causal inference. These results may not generalize to more recent cohorts, because of secular changes in education over time. Our results should be confirmed in other data sources.

To our knowledge, this is the first analysis to rigorously examine the combined consequences of different types



**Figure 5.** Coefficient plots for the association between educational trajectories from ages 14 to 48 years and physical health at age 50 years compared with respondents who completed high school only among A) White Americans, B) Black Americans, and C) Latino Americans in the National Longitudinal Survey of Youth 1979 cohort (1979–2016). Stratified results are presented. Some college predicted better physical health at age 50 years for Black Americans than for White Americans. Findings were similar, although interaction models included the null, for Black Americans in the following educational trajectories: less than high school education, GED credential, graduate school; and for Latino Americans in the following educational trajectories: some college after GED credential, and AA after less than high school education. Results are adjusted for birth cohort, race, birth in a southern state, birth outside the United States, rural residence at age 14 years, parental education, if parents worked for pay, and parental occupation. AA, associate's degree; BA, bachelor's degree; CI, confidence interval; GED, general educational development.



**Figure 6.** Coefficient plots for the association between educational trajectories from ages 14 to 48 years and physical health at age 50 years compared with respondents who completed high school only among those from A) high childhood socioeconomic (cSES) backgrounds and B) low childhood socioeconomic backgrounds in the National Longitudinal Survey of Youth 1979 cohort (1979–2016). Stratified are results presented. Low cSES respondents who attained AAs or BAs immediately had better physical health at age 50 years than high cSES respondents. Results are adjusted for birth cohort, race, birth in a southern state, birth outside the United States, rural residence at age 14 years, parental education, if parents worked for pay, and parental occupation. AA, associate's degree; BA, bachelor's degree; CI, confidence interval; GED, general educational development.

and timing of educational credentials for health. Despite strong theoretical reasons to question whether alternative or delayed credentials would have the same benefit as traditional timing to credential, there has been little empirical evaluation of these questions. Nearly all educational programs focus on promoting school attendance and credential attainment, without attention to the timing of the credential. Many programs, such as for-profit colleges, promote college enrollment but confer limited labor market mobility and saddle students with debt (55). Evaluating whether health benefits depend on educational timing has substantial policy relevance. A challenge has been characterizing thousands of possible educational trajectories across the life course into a smaller, analytically tractable number of exposure groups.

We addressed this problem using sequence analysis, which led to unique insights such as situations in which the timing of schooling differentially predicts health for individuals who receive the same final credential; conventional analyses of terminal state would miss the importance of life-course educational trajectories.

As delayed or interrupted educational trajectories become more common in the United States, particularly among structurally marginalized groups, examining the associations between such educational trajectories and health is a critical research domain. Without evidence on the health effects of educational experiences as lived by contemporary cohorts, we cannot provide relevant policy guidance or anticipate health inequities. Our findings suggest that the

literature to date has likely underestimated the impact of heterogeneous educational trajectories on population health and health inequities.

## ACKNOWLEDGMENTS

Author affiliations: Department of Family and Community Medicine, University of California, San Francisco, San Francisco, California (Anusha M. Vable); Department of Epidemiology, University of California, Berkeley, Berkeley, California (Catherine d. P. Duarte); School of Management, University of San Francisco, San Francisco, California (Alison K. Cohen); Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California (M. Maria Glymour); Graduate School of Education, University of California, Riverside, Riverside, California (Robert K. Ream); and Department of Public Health, School of Social Sciences, Humanities and Arts, University of California, Merced, Merced, California (Irene H. Yen).

This research was supported by the National Institute on Aging (award R01AG056360 to IHY).

We thank Rae Wannier for help with revision analyses and the creation of figures.

Portions of the work were presented at the 52nd Annual Meeting of the Society for Epidemiologic Research in Minneapolis, Minnesota, June 18–21, 2019.

Conflict of interest: none declared.

## REFERENCES

- Cohen AK, Rai M, Rehkopf DH, et al. Educational attainment and obesity: a systematic review. *Obes Rev*. 2013; 14(12):989–1005.
- Gilman SE. Review: there is marked socioeconomic inequality in persistent depression. *Evid Based Ment Heal*. 2003;6(3):75.
- Sharp ES, Gatz M. Relationship between education and dementia: an updated systemic review. *Alzheimer Dis Assoc Disord*. 2011;25(4):289–304.
- Galea S, Tracy M, Hoggatt KJ, et al. Estimated deaths attributable to social factors in the United States. *Am J Public Health*. 2011;101(8):1456–1465.
- Cohen AK, Syme SL. Education: a missed opportunity for public health intervention. *Am J Public Health*. 2013;103(6): 997–1001.
- Rehkopf DH, Glymour MM, Osypuk TL. The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. *Curr Epidemiol Rep*. 2016;3(1):63–71.
- Walsemann KM, Bell BA, Hummer RA. Effects of timing and level of degree attained on depressive symptoms and self-rated health at midlife. *Am J Public Health*. 2012;102(3): 557–563.
- Milesi C. Do all roads lead to Rome? Effect of educational trajectories on educational transitions. *Res Soc Stratif Mobil*. 2010;28(1):23–44.
- Horn LJ, Carroll CD. *Nontraditional Undergraduates: Trends in Enrollment from 1986 to 1992 and Persistence and Attainment Among 1989 Beginning Postsecondary Students*. National Center for Education Statistics. 1996. <https://nces.ed.gov/pubs/web/97578.asp>. Accessed April 17, 2019.
- Bozick R, DeLuca S. Better late than never? Delayed enrollment in the high school to college transition. *Soc Forces*. 2005;84(1):531–554.
- Taniguchi H, Kaufman G. Degree completion among nontraditional college students. *Soc Sci Q*. 2005;86(4): 912–927.
- National Center for Education Statistics. *Associate's degrees conferred by postsecondary institutions, by sex of student and discipline division: 2005 - 06 through 2015-16*. 2018. [https://nces.ed.gov/programs/digest/d17/tables/dt17\\_321.10.asp](https://nces.ed.gov/programs/digest/d17/tables/dt17_321.10.asp). Accessed February 23, 2019.
- National Center for Education Statistics. *Fast Facts: Degrees conferred by race and sex in 2000-01 and 2015-16*. 2018. <https://nces.ed.gov/fastfacts/display.asp?id=72>. Accessed February 23, 2019.
- National Center for Education Statistics. *Associate's degrees conferred by postsecondary institutions, by race/ethnicity and sex of student: Selected years, 1976-77 through 2015-16*. 2018. [https://nces.ed.gov/programs/digest/d17/tables/dt17\\_321.20.asp](https://nces.ed.gov/programs/digest/d17/tables/dt17_321.20.asp). Accessed February 24, 2019.
- Malkus N, Sen A. *Characteristics of GED Recipients in High School: 2002–06*. Vol. 2002. Washington, DC: National Center for Education Statistics; 2011:1, 3. <https://nces.ed.gov/pubs2012/2012025.pdf>. Accessed February 26, 2019.
- Carnevale AP, Strohl J. *How Increasing College Access Is Increasing Inequality, and What to Do About It*. 2010. <https://www.luminafoundation.org/files/resources/increasing-inequality.pdf>. Accessed August 21, 2020.
- Goldrick-rab S. Following their every move: differences in college pathways. *Sociol Educ*. 2006;79:61–79.
- Liu SY, Chavan NR, Glymour MM. Type of high-school credentials and older age ADL and IADL limitations: is the GED credential equivalent to a diploma? *Gerontologist*. 2013;53(2):326–333.
- Zajacova A, Montez JK. The health penalty of the GED: testing the role of noncognitive skills, health behaviors, and economic factors. *Soc Sci Q*. 2017;98(1):1–15.
- Zajacova A, Everett BG. The nonequivalent health of high school equivalents. *Soc Sci Q*. 2014;95(1):221–238.
- Caputo RK, Race S, Causes P, et al. The GED as a signifier of later life health & economic well-being. *Race, Gender, & Class*. 2005;12(2):81–103.
- Caputo R. The GED as a predictor of mid-life health and economic well-being. *J Poverty*. 2005;9(4):73–97.
- Kenkel D, Lillard D, Mathios A, et al. The roles of high school completion and GED receipt in smoking and obesity. *J Labor Econ*. 2006;24(3):635–660.
- Ou S, Do GED. Recipients differ from graduates and school dropouts? Findings from an inner-city cohort. *Urban Educ*. 2008;43(1):83–117.
- Zajacova A. Health in working-aged Americans: adults with high school equivalency diploma are similar to dropouts, not high school graduates. *Am J Public Health*. 2012; 102(Suppl 2):284–290.
- Liu SY, Manly JJ, Capistrant BD, et al. Historical differences in school term length and measured blood pressure: contributions to persistent racial disparities among US-born adults. *PLoS One*. 2015;10(6):1–13.
- Vable AM, Cohen AK, Leonard SA, et al. Do the health benefits of education vary by sociodemographic subgroup? Differential returns to education and implications for health inequities. *Ann Epidemiol*. 2019;28(11):759–766.e5.
- Vable AM, Kiang MV, Basu S, et al. Military service, childhood socio-economic status, and late-life lung function:



- Korean war era military service associated with smaller disparities. *Mil Med.* 2018;183(9–10):e576–e582.
29. Vable AM, Canning D, Glymour MM, et al. Can social policy influence socioeconomic disparities? Korean War GI bill eligibility and markers of depression. *Ann Epidemiol.* 2016; 26(2):129–135.e3.
  30. Vable AM, Eng CW, Mayeda ER, et al. Mother's education and late-life disparities in memory and dementia risk among US military veterans and non-veterans. *J Epidemiol Community Health.* 2018;72(12): 1162–1167.
  31. Laeson E, Xu J, Lin S, et al. A comparison of SF-36 and SF-12 composite scores and subsequent hospitalization and mortality risks in long-term dialysis patients. *Clin J Am Soc Nephrol.* 2010;5(2):252–260.
  32. Brzinsky-fay C, Brzinsky-fay C, Kohler U. Sequence analysis with Stata sequence analysis with Stata. *Stata J.* 2006;6(4): 435–460.
  33. Halpin B. Optimal matching analysis and life-course data: the importance of duration. *Sociol Methods Res.* 2009;38(3): 365–388.
  34. Stata.com. Cluster-analysis stopping rules. In: *Stata Manuals.* <https://www.stata.com/manuals/mvclusterstop.pdf>. Accessed June 16, 2019.
  35. Halpin B. *Cluster Analysis Stopping Rules in Stata, Working Paper WP2016-01.* Limerick, Ireland: University of Limerick. 2016. <http://ulsites.ul.ie/sociology/sites/default/files/wp2016-01.pdf>.
  36. Ware J, Kosinski M, Keller SD. A 12-item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34(3): 220–233.
  37. National Longitudinal Survey of Youth 1979. *Appendix 19: SF-12 Health Scale Scoring.* 1978. <https://www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/codebook-supplement/nlsy79-appendix-19-sf-12-health-scale>. Accessed March 9, 2018.
  38. Glymour MM, Avendaño M, Berkman LF. Is the “Stroke Belt” worn from childhood? Risk of first stroke and state of residence in childhood and adulthood. *Stroke.* 2007;38(9): 2415–2421.
  39. Victora C, Huttly S, Fuchs S, et al. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *Int J Epidemiol.* 1997;26(1):224–227.
  40. Bureau of Labor Statistics. Aptitude, achievement, and intelligence scores. In: *National Longitudinal Surveys.* 2019. <https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/education/aptitude-achievement-intelligence-scores>. Accessed June 16, 2019.
  41. Kanarek JL. Examining the AFQT as a proxy for human capital. *Undergrad Econ Rev.* 2015;11(1).
  42. Cascio EU, Lewis EG. Schooling and the AFQT: evidence from school entry laws. In: *NBER Working Paper 11113.* 2005. <https://www.nber.org/papers/w11113.pdf>. Accessed November 26, 2019.
  43. Dean RA, Hoffman LM, Hoffman DM, et al. *Evaluation of English Language Needs, Preparation, and Screening Among Non-Native English Speaking Officers and Soldiers.* 1988. <https://apps.dtic.mil/dtic/tr/fulltext/u2/a203611.pdf>. Accessed February 13, 2020.
  44. Utah Health Department. *Interpreting the SF-12.* 2001. [http://health.utah.gov/opha/publications/2001hss/sf12/SF12\\_Interpreting.pdf](http://health.utah.gov/opha/publications/2001hss/sf12/SF12_Interpreting.pdf). Accessed February 13, 2020.
  45. Hopman WM, Harrison MB, Coe H, et al. Associations between chronic disease, age and physical and mental health status. *Chronic Dis Can.* 2009;29(2):108–117.
  46. Tyler JH, Lofstrom M. Is the GED an effective route to postsecondary education for school dropouts? *Econ Educ Rev.* 2010;29(5):813–825.
  47. Heckman JJ, Humphries JE, Mader NS. The GED. *NBER Working Paper No. 16064.* 2010. <https://www.nber.org/papers/w16064>. Accessed November 12, 2018.
  48. Edwards RD. Health, SES, and the timing of education among military retirees. *Educ Econ.* 2016;24(4):393–410.
  49. Pallas AM. Schooling in the course of human lives: the social context of education and the transition to adulthood in industrial society. *J Indexing Metrics.* 1993;63(4):409–447.
  50. Ross CE, Mirowsky J. Sex differences in the effect of education on depression: resource multiplication or resource substitution? *Soc Sci Med.* 2006;63(5):1400–1413.
  51. Denning JT, Marx BM, Turner LJ. *ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare.* In: *Upjohn Institute Working Paper 17–280.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. 2017. [https://research.upjohn.org/up\\_workingpapers/280/](https://research.upjohn.org/up_workingpapers/280/). Accessed June 24, 2019.
  52. Leonard JS. The impact of affirmative action regulation and equal employment law on black employment. *J Econ Perspect.* 1990;4(4):47–63.
  53. Venkataramani AS, Cook E, O'Brien IRL, et al. College affirmative action bans and smoking and alcohol use among underrepresented minority adolescents in the United States: a difference-in-differences study. *PLoS Med.* 2019; 16(6):e1002821.
  54. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav.* 2010; 51(1\_suppl):S28–S40.
  55. The New York Times Editorial Board. Exploiting veterans for profit. In: *New York Times.* 2017. [https://www.nytimes.com/2017/11/24/opinion/exploiting-veterans-profit.html?\\_r=1](https://www.nytimes.com/2017/11/24/opinion/exploiting-veterans-profit.html?_r=1). Accessed March 3, 2018.