Schooling and Health: A Life Course Approach

by

Félice Lê

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2013

Doctoral Committee:

Professor Ana V. Diez Roux, Chair
Research Assistant Professor Yun Li
Professor Hal Morgenstern
Professor Robert Schoeni

Acknowledgments

My sincerest thanks to my advisor, Ana Diez Roux, who knows unfailingly how to help, what to say, and when to push. I couldn't hope for a more inspiring role model of a researcher and mentor, and I am deeply grateful for your unwavering support. Thanks also to the other members of my dissertation committee for your time, expertise, and patience. I thank Hal Morgenstern for his tireless willingness to review and discuss iterations of this work, and Bob Schoeni for repeatedly pulling me out of the trees to find the forest. Brisa Sanchez provided valuable analytical insights on my prospectus. Yun Li generously stepped in to give guidance when I hit analytical stumbling blocks. Joanne Carlisle and Fabian Pfeffer also provided valuable input. This work would not have been possible without the ongoing commitment of thousands of PSID participants, researchers, and staff, including the expert researchers at the Institute for Social Research.

I would like to thank Ana Baylin not only for giving me the opportunity to work with her as a GSI for two semesters but also for her open, frank, and caring conversations about living and working as an academic researcher. Sandro Galea originally brought me to Michigan and has remained an available mentor from afar.

The faculty, staff, and students at the Center for Social Epidemiology and Population Health have been an endless source of academic insight, technical and administrative help, and friendship. Thank you for always being there with a helping hand, no matter what the situation. I am thankful to Gregg Davis, Sarah Reeves, Sarah Lyon-Callo, Mikiko Senga, and Erin Payne for reminding me that we're all in this together. I could not have done this without my family and

friends. Thank you to Peter Zubulake, Liz and Daniel Rodrigues, and the dedicated teachers and dancers at CAS Ballet Theatre School for making Ann Arbor feel like home.

Most of all, I would like to thank my husband Greg Scherban for his endless support. Thank you for not only editing my papers and listening to my presentations, but for shepherding me through this process with such love and patience. This project is yours and Ted's as much as mine.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BMI	Body mass index
CDI	Children's Depression Inventory
CDS	Child Development Supplement
CI	Confidence interval
DAG	Directed acyclic graph
GEE	General estimating equation
MAR	Missing at random
MCAR	Missing completely at random
MNAR	Missing not at random
MSM	Marginal structural model
NHIS	National Health Interview Survey
PR	Prevalence ratio
PSID	Panel Study of Income Dynamics
SRMI	Sequential regression multivariate imputation
TA	Transition into Adulthood Study
WJ-R ACH	Woodcock-Johnson Psycho-Educational Battery-Revised Achievement

Abstract

Health differences between adults with different levels of schooling are among the largest and most wide-ranging of health disparities in the United States. Yet there remains much we do not know about how and when in the course of the lifespan these disparities develop. We used data from the Panel Study of Income Dynamics (PSID) and two supplemental studies of young people, the Child Development Supplement (CDS) and Transition into Adulthood study (TA), to examine three processes with origins early in life that may contribute to adult schooling-related disparities. Our approach emphasized how these processes develop over the life course and used analytical methods to address how bidirectional and time-dependent causation may influence observed associations between schooling and health. First, using CDS and TA data spanning 10 years, we examined how health status throughout childhood and adolescence influences educational progress. Second, again using 10 years of CDS and TA data, we examined how academic achievement was associated with health status, body mass index (BMI), and psychological distress 5 years later. Third, using data from the main PSID, we examined influences on the health of adults by their grandparents' educational attainment. The first study found associations between poorer health status and less completed schooling by the end of follow-up that appeared to accumulate over time and were more evident among children who were older at baseline. The second study found that higher average academic achievement was associated with better health 5 years later among girls but not boys. The third study found that after accounting for parent and participant schooling, there were monotonic inverse associations among Whites between grandparent schooling and poor health status, smoking, and obesity.

Estimates among Blacks were similar to those among Whites for smoking but minimal for health status and obesity. The results all support past evidence that schooling and health are mutually beneficial but highlight the complexity and longitudinal nature of processes relating schooling and health. A better understanding of these processes is imperative to developing interventions both to improve educational outcomes among disadvantaged groups and to reduce educational disparities in adult health.

CHAPTER 1

INTRODUCTION

Background

Health differences between adults with different levels of schooling are among the most wide-ranging and well documented of health disparities.[1-20] Higher educational attainment is related to better health behaviors[10,11,14,18], overall health status[5,6,12], and cognition and memory[8,13]; fewer functional limitations[5,7,17]; lower infection[19] and chronic disease[5] burden; and lower overall and cause-specific mortality.[1,4,9,15,16,20] As of 2006, life expectancy at age 25 in the United States was 9.3 years longer for men with a bachelor's degree or higher than for those without a high school diploma; for women the difference was 8.6 years.[21]

Yet there remains much we do not know about how and when in the course of the lifespan these disparities develop. Although one important way schooling improves future health is by providing access to higher-paying jobs—and the accompanying access to healthier living conditions, nutritious foods, medical care, and time for rest and exercise[10,22,23]—there are likely other processes with origins earlier in life that contribute to adult education-related disparities. The purpose of this dissertation is to examine three of these processes in a contemporary United States sample with an emphasis on how they develop over the life course. The conceptual diagram in Figure 1.1 represents these processes with black arrows; they are shown in the context of interrelated mechanistic pathways that may all contribute to observed associations between adult educational attainment and health. A better understanding of these processes is

imperative to developing interventions both to improve educational outcomes among disadvantaged groups and to reduce educational disparities in adult health.

The first aim (Chapter 2) of this project is to examine how health status throughout childhood and adolescence influences educational progress, as shown by the long-dash line in Figure 1.1. A large number of studies have linked poor childhood and adolescent health with lower educational attainment.[24-28] However, there is disagreement about the magnitude of these effects and how important a role they play in associations between educational attainment and subsequent health.[23,25,29] One reason for these differences in findings may be differences in the ages encompassed by the health measures used: the effects of point-in-time measures health may vary according to the ages to which they pertain and may differ from the effects of global measures describing a predominant health state throughout child and adolescence. Therefore, it may be elucidative to use measures of health that more explicitly portray the health of children and adolescents over time. We use multiple measures of health spanning a 10-year period to more fully characterize health in children and adolescents over time. We hypothesize that poorer health status slows educational progress, and that a pattern of declining health is more deleterious than a pattern of improving health. Besides the inherent public health and educational importance of understanding the degree to which poor health may impede young people's schooling, this aim also helps quantify the potential bias from reverse causation on estimates of educational attainment on adult health.[2]

The second aim (Chapter 3), shown by the short-dash line in Figure 1.1, is to examine longitudinal effects of academic achievement during childhood and adolescence on health several years later. We hypothesize that higher academic achievement leads to better future health. Like the first aim, this aim serves two roles. First, it addresses a question that remains

understudied in extant literature and that may be important for understanding how schooling and health are related in young people. Most previous studies examining links between academic achievement and health in young people have used cross-sectional data or short follow-up times, although a few have examined academic achievement in late adolescence in relation to substance use several years later.[30-37] Second, it may help us better understand mechanisms linking educational attainment and adult health. Specifically, cognitive and psychosocial mechanisms through which schooling may benefit health have been less frequently examined than income-mediated mechanisms. Cognitive mechanisms may include the acquisition of knowledge about how to live healthfully and improvements in abilities to read, understand, or act on health information or better use the resources available to benefit health.[2,38] Psychosocial mechanisms include an increased sense of personal control, the patience and time preference necessary to make healthier choices, access to higher social integration or social support, or exposure to more healthful behavioral norms.[3,10,23,39] Students who more successfully make these cognitive and psychosocial gains—as reflected in their academic achievement—may enjoy better health even before completion of their schooling. This is in contrast to occupation- or income-mediated effects, which would not appear until students enter the workforce. Therefore, evidence that academic achievement is related to later health in young people supports the existence of effects of schooling on health through mechanisms other than adult economic conditions.

The third aim (Chapter 4), represented by the dotted arrows in Figure 1.1, is to examine influences on the health of adults by their grandparents' educational attainment, through pathways other than those mediated by their parents' and own educational attainment. If there is social contact between grandparents and their grandchildren, grandparents may influence their grandchildren's health through many of the same mechanisms exerted by the grandchildren's

parents. Other direct influences—such as paying for schooling or health care, direct monetary gifts, or the facilitation of schooling through legacy admissions policies—do not require direct contact between the grandparent and grandchild, or even that the grandparent be living.[40] We hypothesize that net of their own and their parents' educational attainment, people with higher grandparent educational attainment enjoy better health in adulthood. Persistent social and health inequalities over time in the United States suggest that associations between schooling and adult health may to some degree reflect not only the influence of each person's own schooling but also that of previous generations. This interpretation is supported by existing evidence of intergenerational associations between parents' schooling and their children's schooling and health.[24,41-50] However, only a small handful of studies have explicitly investigated the influence of grandparents' schooling on the health of their grandchildren, and particularly the health of the grandchildren in adulthood. Osler et al. found in a three-generation study that having a higher number of ancestors with some secondary education was related to lower mortality among a cohort of Danish men, and that this relation was robust to adjustment for the occupational class of the ancestors.[51] On the other hand, Ahren-Moonga et al. found that higher education levels of parents and maternal grandmothers were related to a higher risk of hospitalization for an eating disorder among a cohort of Swedish women.[52]

Study Population

We use data from the Panel Study of Income Dynamics (PSID), a longitudinal study of a representative sample of U.S. families conducted by the Survey Research Center at the Institute for Social Research at the University of Michigan.[53] The study was started in 1968 and contains data on about 70,000 individuals. Interviews were conducted annually until 1997, when the study switched to biennial data collection. In most waves, a single adult family member provides

information about him-/herself and all other family members. Since 1973, the large majority of interviews have been conducted over the telephone. The number of families interviewed in each wave has ranged from about 5,000 to over 10,000; per-wave response rates are generally above 95%.[53]

The primary respondent in each family in the PSID is called the Head.[53] The male member of a heterosexual married or unmarried couple is designated as the Head by default, although in some situations—such as if the male partner is incapacitated or the respondents insist—the female partner is designated as Head. The (male) Head's wife or cohabiting female partner is designated as Wife or "Wife" respectively. In households with a single head of household, this person serves as the Head regardless of sex. Information is also collected about other family members, including children. When a family member leaves the household (e.g., after divorce or when a child grows up), his/her new household is added to the PSID sample as a "split-off" family. Because of this, members of many PSID families are related to members of other PSID families. The study currently contains information about up to three generations of any given family.[53]

The analyses for the first two aims use primarily data from the Childhood Development Supplement (CDS) and Transition into Adulthood Study (TA), two supplementary studies to the PSID that focus on children, adolescents, and young adults. The CDS began in 1997; all PSID families with a child aged 0–12 in calendar year 1997 were eligible participate, with up to two children chosen per family.[54] The study consisted of extensive interviews with 3,563 children and their guardians (including absent fathers), teachers, and school administrators. In 2002/2003 a second wave of interviews was carried out with 2,907 children aged 5–18 who had participated in the first wave. In 2007/2008 a third wave of interviews was carried out with children who

were still under the age of 18 in 2007. The primary interviewee in the CDS is the participant's primary caregiver, with whom the child must be living. In the vast majority (over 90%) of cases, this is the child's biological mother. If the biological mother is not living with the child, the primary caregiver is defined using the following order of preference: (1) stepmother, adoptive mother, or foster mother; (2) other female legal guardian at least 18 years old (often the grandmother); (3) biological, step-, adoptive, or foster father; (4) other male legal guardian at least 18 years old; (5) another unpaid adult who lives with the child and takes primary responsibility for his/her care.[55]

When CDS participants turn 18, they become ineligible to continue participation in the CDS. Instead, these teenagers become eligible to participate in the TA. The TA began in 2005 with pilot telephone interviews of 745 young adults aged 18 and over who had participated in the CDS and had finished or left high school.[56] In 2007, a second wave of TA interviews was carried out of the same young adults and additional CDS participants who had turned 18 and finished or left high school since 2005, for a total of 1,118 interviews. When CDS and TA participants move into their own households, their new families are included in the regular PSID biennial interviews.

Analysis Approach

Part of the difficulty in characterizing causal relations between schooling and health is that the complex interplay between schooling, socioeconomic status, and health throughout the life course makes estimates susceptible to biases from confounding and reverse causation.[57] Researchers have employed a range of different methods in attempts to address this problem, including adjusting for confounders in multivariable models[11,20,58], explicitly testing causal and noncausal mechanistic hypotheses[10,11], comparing siblings and twins[12,59,60], using regression

discontinuity to compare health outcomes of cohorts who completed schooling before and after policy reforms that affected the duration of schooling[16], and using instrumental variable (IV) techniques with instruments ranging from IQ, quarter of birth, and parental characteristics to unemployment rates, state-level education expenditures, and school reform laws.[6,7,13,16,61] Because of the challenges of each method, as well as the complexity of the underlying processes being modeled, it can be difficult to determine the extent to which these analytical methods improve on, or even affect, estimates derived from more conventional regression approaches.

In this project, we use methods drawn from the causal inference literature to address potential sources of bias that may not be appropriately handled by conventional regression. In the first aim, we use sibling fixed effects models to help control for observed and unobserved differences that might confound associations between health status and education progress stemming from familial or community factors shared between siblings. In the second aim, we use marginal structural models (MSMs) to account for mutual influence of academic achievement and health over time. In the third aim, we again use an MSM approach, this time to reduce bias in estimating the direct effect of grandparent schooling on adult health independent of effects mediated by parent and own schooling. In each case, we compare estimates from these causal methods with estimates from a conventional regression approach.

Summary

The goal of this dissertation is to examine life course processes that may contribute to schooling-related disparities in adult health. To do this, we use data spanning 41 years from the Panel Study of Income Dynamics, a longitudinal study of American families, to answer three questions.

1. Does poor health status in childhood and adolescence impede educational progress?

2. Does higher academic achievement in childhood and adolescence predict better health several years later?
3. Does grandparent educational attainment influence their grandchildren's adult health, net of the parents' and grandchildren's educational attainment?

References

1. Antonovsky A. Social class, life expectancy, and overall mortality. *The Milbank Memorial Fund Quarterly* 1967;**45**(2):31-73.

2. Cutler DM, Lleras-Muney A. Education and Health: Evaluating Theories and Evidence. *NBER Working Papers*. Cambridge, MA: National Bureau of Economic Research, 2006.

3. Feinstein L, Sabates R, Anderson TM, Sorhaindo A, Hammond C. What are the effects of education on health? In: Desjardins R, Schuller T, eds. *Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium*. Copenhagen: Organisation for Economic Co-operation and Development, 2006;171-353.

4. Olshansky SJ, Antonucci T, Berkman L, Binstock RH, Boersch-Supan A, Cacioppo JT, Carnes BA, Carstensen LL, Fried LP, Goldman DP, Jackson J, Kohli M, Rother J, Zheng Y, Rowe J. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Aff (Millwood)* 2012;**31**(8):1803-13.

5. Adams P, Barnes P, Vickerie J. Summary health statistics for the U.S. population: National Health Interview Survey, 2007. *Vital Health Stat*. Vol. 10 National Center for Health Statistics, 2008.

6. Arendt J. Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review* 2005;**24**:149-160.

7. Berger MC, Leigh JP. Schooling, self-selection, and health. *The Journal of Human Resources* 1989;**24**(3):433-455.

8. Cagney KA, Lauderdale DS. Education, wealth, and cognitive function in later life. *2002* 2002;**57B**(2):P163-P172.

9. Christenson BA, Johnson NE. Educational inequality in adult mortality: an assessment with death certificate data from Michigan. *Demography* 1995;**32**(2):215-229.

10. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. *Journal of Health Economics* 2010;**29**:1-28.

11. Farrell P, Fuchs VR. Schooling and health: the cigarette connection. *Journal of Health Economics* 1982;**1**:217-230.

12. Fujiwara T, Kawachi I. Is education causally related to better health? A twin fixed-effect study in the USA. *Int J Epidemiol* 2009;**38**(5):1310-1322.

13. Glymour MM, Kawachi I, Jencks C, Berkman LF. Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology and Community Health* 2008;**62**:532-537.

14. Harper S, Lynch J. Trends in socioeconomic inequalities in adult health behaviors among U.S. states, 1990-2004. *Public Health Reports* 2007;**122**:177-189.

15. Kitagawa EM, Hauser PM. Education differentials in mortality by cause of death: United States, 1960. *Demography* 1968;**5**(1):318-353.

16. Lleras-Muney A. The relationship between education and adult mortality in the United States. *The Review of Economic Studies* 2005;**72**(1):189-221.

17. Ross CE, Mirowsky J. Refining the assocation between education and health: the effects of quantity, credential, and selectivity. *Demography* 1999;**36**(4):445-460.

18. Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health* 1992;**82**(6):816-20.

19. Zajacova A, Dowd J, Aiello A. Socioeconomic and race/ethnic patterns in persistent infection burden among U.S. adults. *J Gerontol A Biol Sci Med Sci* 2009;**64A**(2):272-279.

20. Zajacova A, Hummer RA. Gender differences in education effects on all-cause mortality for white and black adults in the United States. *Soc Sci Med* 2009;**69**:529-537.

21. National Center for Health Statistics. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. Hyattsville, MD, 2012.

22. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. *Health Affairs* 2002;**21**:60-76.

23. Ross CE, Wu C-L. The links between education and health. *Am Sociol Rev* 1995;**60**(5):719-745.

24. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *J Health Econ* 2005;**24**(2):365-89.

25. Haas SA. Health selection and the process of social stratification: the effect of childhood health on socioeconomic attainment. *Journal of Health and Social Behavior* 2006;**47**(4):339-354.

26. Haas SA. The long-term effects of poor childhood health: an assessment and application of retrospective reports. *Demography* 2007;**44**(1):113-135.

27. Haas SA, Fosse NE. Health and the educational attainment of adolescents: evidence from the NLSY97. *J Health Soc Behav* 2008;**49**(2):178-92.

28. Jackson MI. Understanding links between adolescent health and educational attainment. *Demography* 2009;**46**(4):671-94.

29. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. *Ann N Y Acad Sci* 2006;**896**(1):3-15.

30. Undheim AM, Sund AM. School factors and the emergence of depressive symptoms among young Norwegian adolescents. *Eur Child Adolesc Psychiatry* 2005;**14**(8):446-53.

31. Frojd SA, Nissinen ES, Pelkonen MU, Marttunen MJ, Koivisto AM, Kaltiala-Heino R. Depression and school performance in middle adolescent boys and girls. *J Adolesc* 2008;**31**(4):485-98.

32. Piko BF. Self-perceived health among adolescents: the role of gender and psychosocial factors. *Eur J Pediatr* 2007;**166**(7):701-8.

33. Eide ER, Showalter MH, Goldhaber DD. The relation between children's health and academic achievement. *Children and Youth Services Review* 2010;**32**:231-238.

34. Brolin Laftman S, Modin B. School-performance indicators and subjective health complaints: are there gender differences? *Sociol Health Illn* 2012;**34**(4):608-25.

35. Crosnoe R, Muller C. Body mass index, academic achievement, and school context: examining the educational experiences of adolescents at risk of obesity. *J Health Soc Behav* 2004;**45**(4):393-407.

36. Crosnoe R, Riegle-Crumb C. A life course model of education and alcohol use. *J Health Soc Behav* 2007;**48**(3):267-82.

37. Tucker JS, Martinez JF, Ellickson PL, Edelen MO. Temporal associations of cigarette smoking with social influences, academic performance, and delinquency: a four-wave longitudinal study from ages 13-23. *Psychol Addict Behav* 2008;**22**(1):1-11.

38. Grossman M. On the concept of health capital and the demand for health. *The Journal of Political Economy* 1972;**80**(2):223-255.

39. Fuchs VR. Time Preference and Health: An Exploratory Study. *NBER Working Papers*. Cambridge, MA: National Bureau of Economic Research, 1980.

40. Schiller JS, Lucas JW, Ward BW, Peregoy JA. Summary health statistics for U.S. adults: National Health Interview Survey, 2010. *Vital Health Stat 10* 2012(252):1-207.

41. Harper S, Lynch J, Hsu WL, Everson SA, Hillemeier MM, Raghunathan TE, Salonen JT, Kaplan GA. Life course socioeconomic conditions and adult psychosocial functioning. *Int J Epidemiol* 2002;**31**(2):395-403.

42. Best LE, Hayward MD, Hidajat MM. Life course pathways to adult-onset diabetes. *Soc Biol* 2005;**52**(3-4):94-111.

43. Bowen ME. Childhood socioeconomic status and racial differences in disability: evidence from the Health and Retirement Study (1998-2006). *Soc Sci Med* 2009;**69**(3):433-41.

44. Chichlowska KL, Rose KM, Diez-Roux AV, Golden SH, McNeill AM, Heiss G. Life course socioeconomic conditions and metabolic syndrome in adults: the Atherosclerosis Risk in Communities (ARIC) Study. *Ann Epidemiol* 2009;**19**(12):875-83.

45. Foster HW, Wu L, Bracken MB, Semenya K, Thomas J. Intergenerational effects of high socioeconomic status on low birthweight and preterm birth in African Americans. *J Natl Med Assoc* 2000;**92**(5):213-21.

46. Frank E, Elon L, Hogue C. Transgenerational persistence of education as a health risk: findings from the Women Physicians' Health Study. *J Womens Health (Larchmt)* 2003;**12**(5):505-12.

47. Gilman SE, Abrams DB, Buka SL. Socioeconomic status over the life course and stages of cigarette use: initiation, regular use, and cessation. *J Epidemiol Community Health* 2003;**57**(10):802-8.

48. Guralnik JM, Butterworth S, Wadsworth ME, Kuh D. Childhood socioeconomic status predicts physical functioning a half century later. *J Gerontol A Biol Sci Med Sci* 2006;**61**(7):694-701.

49. Luo Y, Waite LJ. The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. *J Gerontol B Psychol Sci Soc Sci* 2005;**60**(2):S93-S101.

50. Phillips JE, Marsland AL, Flory JD, Muldoon MF, Cohen S, Manuck SB. Parental education is related to C-reactive protein among female middle-aged community volunteers. *Brain Behav Immun* 2009;**23**(5):677-83.

51. Osler M, Andersen AM, Lund R, Holstein B. Effect of grandparent's and parent's socioeconomic position on mortality among Danish men born in 1953. *Eur J Public Health* 2005;**15**(6):647-51.

52. Ahren-Moonga J, Silverwood R, Klinteberg BA, Koupil I. Association of higher parental and grandparental education and higher school grades with risk of hospitalization for eating disorders in females: the Uppsala birth cohort multigenerational study. *Am J Epidemiol* 2009;**170**(5):566-75.

53. Institute for Social Research. The Panel Study of Income Dynamics. http://psidonline.isr.umich.edu/ Accessed April 5, 2010.

54. Institute for Social Research. The PSID-CDS Design At-A-Glance. Ann Arbor, MI: Institute for Social Research, 2006.

55. Institute for Social Research. The Panel Study of Income Dynamics Child Development Supplement: User Guide Supplement for CDS-I. Ann Arbor, MI: The Institute for Social Research, Survey Research Center, University of Michigan, 2010.

56. Institute for Social Research. Transition to Adulthood Study: Overview. Ann Arbor, MI: Institute for Social Research, 2008.

57. Kawachi I, Adler NE, Dow WH. Money, schooling, and health: mechanisms and causal evidence. *Ann N Y Acad Sci* 2010;**1186**:56-68.

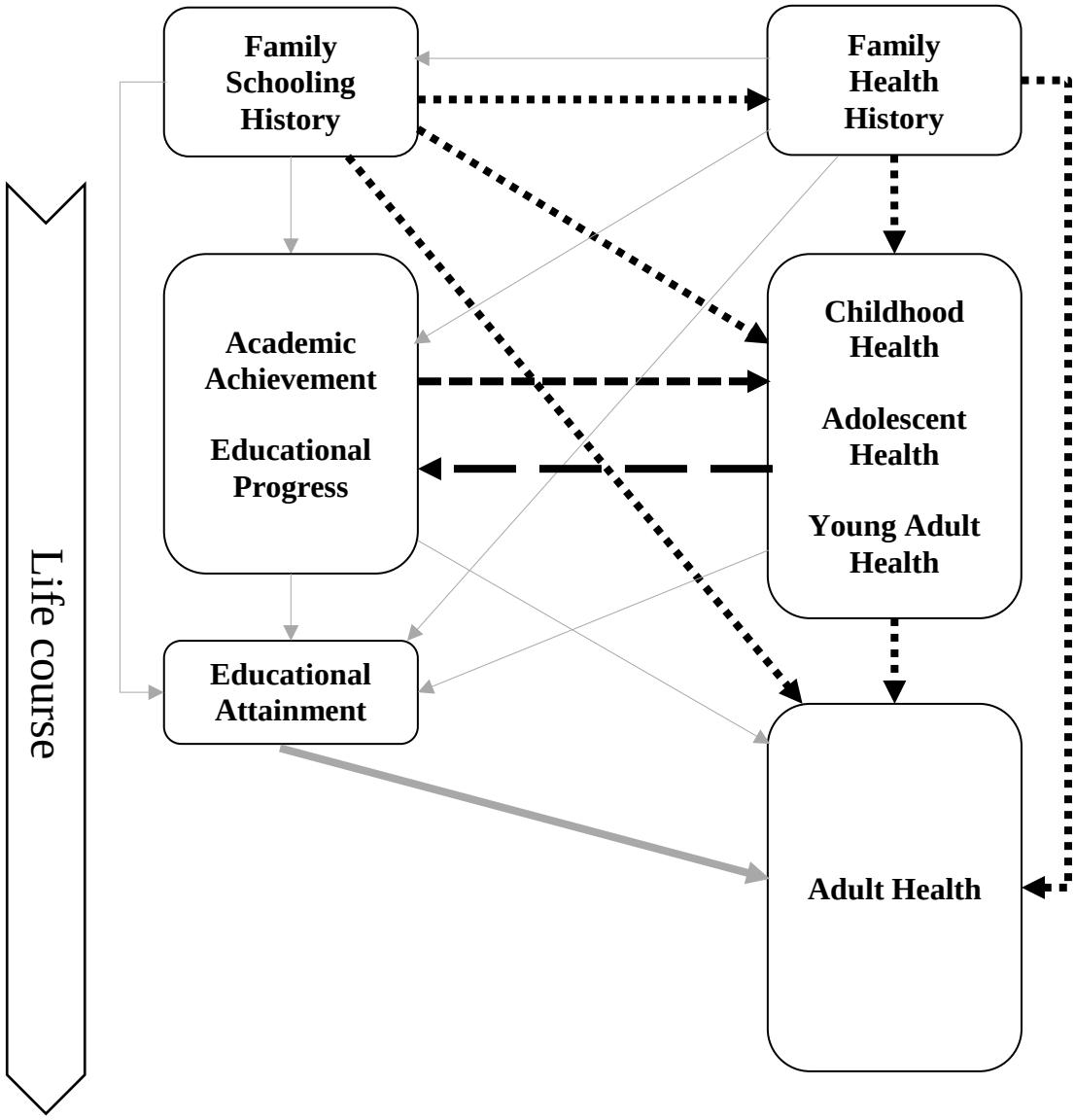
58. Cutler DM, Lange F, Meara E, Richards S, Ruhm CJ. Explaining the Rise in Educational Gradients in Mortality. *NBER Working Paper*. Vol. 15678. Cambridge, MA: National Bureau of Economic Research, 2010.

59. Krieger N, Chen J, Coull B, Selby J. Lifetime socioeconomic position and twins' health: an analysis of 308 pairs of United States women twins. *PLoS Medicine* 2005;**2**(7).

60. Lundborg P. The Health Returns to Education: What Can We Learn from Twins? *IZA Discussion Papers*. Bonn, Germany: Institute for the Study of Labor (IZA), 2008.

61. Arkes J. Does Schooling Improve Adult Health? Santa Monica, CA: RAND, 2003.

Figure 1.1 Conceptual diagram of causal pathways relating schooling and health throughout the life course



The black arrows represent processes addressed in this dissertation. In Chapter 2, we investigate the influence of childhood and adolescent health on educational progress (long-dash arrow). In Chapter 3, we investigate the influence of academic achievement during childhood and adolescence on health 5 years later (short-dash arrow). In Chapter 4, we investigate the influence of grandparent schooling on grandchild adult health through mechanisms other than those mediated by the schooling of the grandchild and his or her parents (dotted arrows).

CHAPTER 2

EFFECTS OF CHILD AND ADOLESCENT HEALTH ON EDUCATIONAL PROGRESS

Introduction

Because health differences between adults with different levels of schooling are among the most wide ranging and well documented of health disparities, factors that influence educational attainment can have a considerable impact on population health and health disparities.[1-7] One such potentially important factor is early-life health (i.e., health during childhood and adolescence), as demonstrated by studies linking poor early-life health with poorer academic performance and lower educational attainment.[8-13] Early-life health may be causally related to both educational attainment and adult health and hence complicate causal inferences regarding the impact of schooling on adult health.[3] A deeper understanding of how early-life health is related to schooling is therefore important not only for understanding the socioeconomic ramifications of poor early-life health but also for understanding the relations between schooling and adult health.

Poor health may impede children and adolescents' progress through the educational system, and their eventual educational attainment, by causing them to miss school days because of illness or medical appointments.[14,15] They may also be less physically or psychologically able to complete assignments and exams or sustain attention during lessons. Students in poor health may also be less able to participate in sports and other school activities, or to interact with their

teachers and peers. These factors may all contribute to poorer academic performance and greater difficulty in developing cognitive, social, and study skills necessary for educational persistence and success. Finally, poor health may affect students' educational attainment by influencing their educational aspirations or expectations, making them less willing or able to invest in long-term educational goals.[12] In these ways, poor health may adversely affect educational progress either by slowing students' advancement (e.g., making them more likely to repeat grades or take time off) or by causing them to start their schooling later or stop their schooling earlier than they would have otherwise.

Studies that span the childhood and adolescent years, when health-related educational differences are first emerging, are valuable tools for investigating how early-life health affects schooling.

We used three waves of longitudinal data on children, adolescents, and young adults aged 5–24 to investigate how different patterns of health throughout early life are related to educational progress by examining links between the respondents' health over a 10-year period and the amount of schooling they completed during that time. We are unaware of previous studies addressing this question using longitudinal data with as long a follow-up period and young baseline age group as the dataset used in this study. These data allowed us to further understand this important topic in three ways. First, we gained insight about possible mechanistic models that might explain the longitudinal influence of early-life health on schooling by explicitly examining the accumulation risk and trajectory models, drawn from theories on the life-course influences on health.[16] (The data available did not allow us to fully test the critical-period model.) Second, we were able to contrast results between participants at different developmental stages to investigate at which ages longitudinal associations between health and

completed schooling became apparent. Third, we were able to test the robustness of our results to adjustment for a wide variety of sociodemographic, health-related, and behavioral potential confounders. Ethics approval was not required since the study makes a secondary data analysis using a public use data set.

We hypothesized that the effects of child and adolescent health on educational progress can be represented by a combination of the accumulation risk and trajectory models. First, we hypothesized that the effects are cumulative: within a given period, youth reporting consistently poor health complete the least schooling, followed by those with mixed health histories, while those with consistently good health complete the most schooling. Second, in keeping with a trajectory model, we hypothesized that the effects are not only cumulative but also pattern-dependent: children and adolescents reporting declining health over time complete less schooling than do those reporting improving health during the same period. For example, a pattern of improving health may foster higher expectations, greater self-esteem, and accelerated gains in ability that result in quicker educational progress, while a pattern of declining health may put students on a trajectory of lower aspirations, poorer self-esteem, and slowing ability gains that negatively affects educational progress.

Methods

Study Population

The study sample came from the U.S. Child Development Supplement (CDS) and Transition into Adulthood Study (TA), two supplementary cohort studies to the Panel Study of Income Dynamics (PSID) that focus on children, adolescents, and young adults.[17] The PSID, a longitudinal study of a representative sample of U.S. families, was started in 1968 and contains

data on about 70,000 individuals. A primary respondent in each family reports on behalf of the family.

The CDS began in 1997 with a sample of 3,563 children; all PSID families with a child aged 0–12 in calendar year 1997 were eligible to participate, with up to two children chosen per family. Subsequent waves of interviews were carried out in 2002–2003 and 2007–2008, in each case including only children who remained under the age of 18 at the time of the study wave. The primary interviewee in the CDS is the participant's primary caregiver, with whom the child must be living. In the vast majority (over 90%) of cases, this is the child's biological mother. The children themselves are also interviewed, with the content of the interview varying according to the age of the child. The TA, which began in 2005, comprises young adults who participated in the CDS but "aged out" by turning 18 and finishing or leaving high school. The participants themselves answer all questions in the TA.

The sample used in this study combined data from the 1997 CDS, the 2002–2003 CDS (called "2002 CDS" in the remainder of this article), and either the 2007–2008 CDS ("2007 CDS") or the 2007 TA to create a longitudinal sample with three waves of interviews per participant. We excluded children who were younger than 10 years old as of 2002, because these children were not asked to self-report their health status; the measures used in 1997 were reported by the children's primary caregivers. Our sample included 2,368 children aged 5–14 at baseline in 1997. Figure 2.1 shows the sample over the 3 waves of data. At the time of the third and final interview in 2007–2008, the sample ranged in age from 15–24.

Measures

We characterized each participant's global health status over the 10-year follow-up period using responses to two different survey questions: (1) "In general, would you say [child's

name]'s health is excellent, very good, good, fair, or poor?" asked of the participants' primary caregivers in 1997, and (2) "In general, how is your health? Would you say excellent, very good, good, fair or poor?" asked of the participants themselves in 2002 and 2007. We dichotomized the measures into "good" (excellent or very good) or "poor" (good, fair, or poor). Although we named the categories of the dichotomized measures good and poor to simplify presentation of results, they are meaningful only relative to each other and do not correspond to absolute measures of health. Past studies, including studies using PSID data, have demonstrated that a dichotomized version of a categorical health status measure predicts morbidity and mortality and is itself predicted by sociodemographic factors.[18-21] In preliminary analyses of our data, the dichotomized health status measures were negatively associated with low birth weight (< 88 ounces), having spent time in a neonatal intensive care unit, caregiver report at baseline of a physician's diagnosis of one of four serious health conditions (epilepsy, autism, sickle cell anemia, heart condition), number of overnight hospital stays (birth–1997 and 1997–2002), and interviewer-assessed BMI (1997 and 2002).

The outcome of educational progress was a continuous measure of years of completed schooling in 2007 as reported by the participant. The measure included primary, secondary, and postsecondary schooling but did not include kindergarten or preschool, nor did it include vocational or other nonacademic training.[22] We adjusted the measure to reflect academic credentials by assigning 12 years of schooling for a GED or high school degree, 14 years for an associate's degree, and 16 years for a bachelor's degree. The interpretation of our outcome measure differs from that of educational attainment measures used in studies of adults because many of the participants in our study had not completed their schooling by the end of follow-up;

rather than a measure of terminal educational attainment, it measures the amount of schooling completed by the end of the follow-up period.[22]

Models were adjusted for sociodemographic characteristics likely to confound the health–schooling association: baseline age, sex, race/Hispanicity, baseline years completed schooling (reported by the primary caregiver and treated as a continuous measure), kindergarten attendance, and caregiver education (5 categories); and a time-varying measure of family income. Caregiver education and family income information was taken from the main PSID interviews. Caregiver education was measured as the maximum of the respondent's primary and secondary caregivers' years of completed education in 1997 or the most recent information available to that point. Family income for 1996 and 2002 included the sum of taxable, transfer, and social security income for every member of the family. Missing values were imputed prior to data release using a hot-deck imputation procedure.[23] We bottom-coded the income measure to zero, then divided it by the corresponding year-specific U.S. Census Bureau poverty threshold, which takes into account inflation, family size and the ages of family members, to create 1996 and 2002 family income-to-poverty-level ratios.[19,24]

Analysis

All analyses were conducted using the SURVEY procedures in SAS software (version 9.2) to account for clustered sibling pairs.[25] There was substantial unit nonresponse in the data: out of the 2368 children who participated in the 1997 CDS and met our eligibility criteria, 1593 (67%) participated in all three study waves, 541 (23%) participated in two waves, and 234 (10%) participated only in the first wave. There was also a considerable amount of item nonresponse. Multiple imputation with 25 imputations using a sequential regression multivariate imputation (SRMI) approach was used to impute missing information from both unit and item

nonresponse.[26,27] SRMI has performed well in practical applications and allowed us both to specify appropriate distributions for individual variables and to restrict the imputation of variables to relevant observations (an important consideration in this data set because of the variability in questions asked across both study waves and participant age).[28]

We used unadjusted logistic and linear regression to examine bivariable associations between the covariates and health status, and between the covariates and completed schooling, respectively.[29] We then used linear regression to examine crude and confounder-adjusted associations between health status and years of completed schooling in 2007. In order to investigate cumulative independent effects of health status at different waves, the health status variables (caregiver-reported health status in 1997, self-reported health status in 2002, and self-reported health status in 2007) were included as separate terms in the models. To test our second hypothesis, we used interaction (product) terms between the health status variables (three 2-way terms and one 3-way term) to test for departures from additivity that might suggest that the effect of health status at one wave might vary according to health at another wave. This interdependence of health effects at different times would suggest the presence of a trajectory-based process in which specific patterns of health over time influence educational progress differently. As an alternative test of our second hypothesis, we ran models including indicator variables for different patterns of health status over time rather than separate terms for health at each wave. In these models, health was defined as stable good (good in all 3 waves), improving (poor in 1997 or both 1997 and 2002), declining (poor in 2007 or both 2002 and 2007); mixed (poor in 2002 or both 1997 and 2007); or stable poor (poor in all 3 waves).

We used interaction terms between the health status variables and baseline age (3 categories: 5–7, 8–10, and 11–14) to account for differences in the effects of health status on

educational progress among participants at different developmental stages. Continuous and squared terms for baseline age (centered at the mean of 9) were also included in adjusted models to account for residual confounding by age within each age category. The income measures were log transformed both because of evidence that the benefits of higher income accrue from proportional rather than absolute incremental income increases, and to improve model fit.[30,31]

Sensitivity analyses

We conducted a series of sensitivity analyses to assess possible influences of residual confounding, misspecification of the exposure and outcome, and missing data on the results we observed. We examined confounding using two different approaches. First, we tested the robustness of our results to additional adjustment for other baseline potential confounders that were not included in our main models because of sample size considerations. These characteristics were family composition (birth order and number of parent figures in the household), caregiver characteristics (cognitive ability, educational expectations for the child, participation at the child's school, provision of cognitive stimulation and emotional support at home, self-rated health, and smoking status), and the child's health status at birth (low birth weight or having spent time in a neonatal intensive care unit). Second, we conducted a sibling fixed-effects analysis using only the sibling pairs from the original sample (N = 1242, or 621 pairs).[32] The goal of this analysis was to tightly control for measured and unmeasured family- and community-level confounders shared between siblings by comparing siblings to each other.[32] Because this approach controls for any shared family-level confounders by design, we assessed only baseline age, birth order, gender, race, kindergarten attendance, and baseline completed years of schooling as potential covariates in the fixed-effects model.

To evaluate measurement error in the health status measure, we repeated analyses both alternatively dichotomizing it as excellent/very good/good vs. fair/poor and treating it as an interval measure ranging 1–5. We also ran models replacing self-rated health status in 2002 with caregiver-rated health status and, among participants aged 18 and under in 2007 (for whom a 2007 measure of caregiver-rated health status was available), replacing both 2002 and 2007 self-rated health status with caregiver-rated measures. As a check of our schooling measure, we ran a model treating a GED as 11 years of schooling instead of 12. One possible interpretation of apparent cumulative effects of poor health on schooling is that children who have persistently poor health may simply have more severe health problems. To investigate this possibility, we examined models excluding children with one of four serious health conditions at baseline (epilepsy, autism, sickle cell anemia, or heart condition; 2.6% of the sample had at least one of these conditions).

Finally, to examine the role of missing data, we repeated our adjusted analyses (1) including only participants with nonmissing schooling and health status information at every wave and (2) imputing data under the assumption of multivariate normality (SAS PROC MI) rather than using SRMI.

Results

Approximately 54% of participants had complete information; the variables with the most missing information were 2002 self-rated health (28% missing), 2007 self-rated health (27% missing) and 2007 completed years of schooling (24% missing). Variable distributions were very similar in the original and imputed samples (Table 2.1). Most participants' primary caregivers (80%) reported the participant was in good health in 1997. As in previous studies, participants were more likely to self-report poor health than their caregivers reported on their

behalf[33,34]: in 2002, the prevalence of self-reported poor health was 38% while the prevalence of caregiver-reported poor health was 25%. Forty-five percent of participants self-reported poor health in 2007. Thirty-two percent of participants had good reported health in all three waves while only 4.8% had poor health in all three waves.

Table 2.2 shows bivariable associations of health status with covariates. The prevalence of poor health status increased over follow-up. Children with poor health in one time period were also more likely to have poor health in subsequent time periods. Years of completed schooling in 2007 was inversely and monotonically associated with the prevalence of poor health in 1997 and 2007 but not with poor health in 2002.

Table 2.2 also shows the distribution of years of completed schooling in 2007 by category of selected predictors. As expected in a young population such as this one, completed schooling was strongly positively associated with age. There was also more variability in completed schooling in 2007 among participants who were older at baseline: compared to participants aged 5–7 at baseline, the standard deviation of completed years of schooling among participants aged 11–14 at baseline was 51% higher and the interquartile range was 2 years as opposed to 1 year.

Table 2.3 shows associations of health status and covariates with completed schooling in 2007. In Model 1 (unadjusted for covariates), poor caregiver-rated health in 1997 and poor self-rated health in 2007 were both associated with completion of fewer years of schooling by 2007 (Model 1: b = -0.44, 95% confidence interval [CI] [-0.66, -0.23] and b = -0.54, 95% CI [-0.81, -0.27], respectively), while poor self-rated health in 2002 was minimally associated with schooling (b = 0.06, 95% CI [-0.18, 0.31]). The addition of terms for baseline age and health–age interactions (Model 2) revealed a strong positive association between age and completed

schooling. There was also evidence of statistical interaction between health status and age ($p < 0.001$ for joint test of the six interaction terms). The coefficients of all six interaction terms were negative, suggesting a greater effect of poor health on educational progress among participants who were older at baseline. The addition of sociodemographic covariates (Model 3) attenuated the main-effect coefficients for age and 1997 health but did not substantially modify other associations. There was no evidence of interaction between the health status measures at different times (Model 4; $p = 0.83$ for joint test of four interaction terms).

Figure 2.2(a) shows adjusted mean differences in 2007 years of schooling for children in poor health compared to those in good health in each of the 3 waves (adjusted for health in the other waves), as well as the estimated cumulative effect of poor health in all 3 waves compared to good health in all 3 waves. Point estimates were generally negative, suggesting that poor health status reduces educational progress, and were most evident in participants who were older at baseline. Compared to participants with good health in all 3 waves, those with poor health status in all 3 waves were predicted to complete 0.02 (95% CI [-0.35, 0.31]) , -0.50 (95% CI [-0.88, -0.12]), and -1.28 (95% CI [-1.78, -0.78]) years of schooling among those aged 5–7, 8–10, and 11–14 at baseline, respectively.

Figure 2.2(b) shows results from the sibling fixed-effects model. In the interest of parsimony, race/Hispanicity and birth order were not included in the final model; this did not affect estimates. Although standard errors were larger, point estimates were very similar to those obtained from the analogous standard regression model (Table 2.3, Model 3 and Figure 2.2[a]). One exception is that among the oldest age cohort the strongest association was observed for 1997 health status rather than 2002 and 2007 health status.

Figure 2.3 shows results from the models using indicators for different health patterns. In each age group, the estimated differences in schooling were most negative for the group with stable poor health. There was some indication in the full-sample model that a pattern of declining health had a larger effect on schooling than a pattern of improving health (Figure 2.3[a]), but point estimates were similar to each other and were not consistent in the fixed-effects model (Figure 2.3[b]).

Results were not sensitive to the additional adjustments, additional exclusion criteria, alternative specification of completed schooling, use of caregiver-rated health measures, or alternative imputation method implemented in sensitivity analyses. Models using alternate parameterizations of the health status measures also produced results consistent with the ones reported.

Discussion

In a national sample of U.S. children aged 5–14 at baseline, we found that poorer general health status over a 10-year period was associated with fewer years of completed schooling by the end of follow-up. Our results are consistent with an accumulation risk model: we found among older participants that poor health at each wave was associated with less completed schooling independent of health at other waves, and that the health-related difference in educational progress was most pronounced among participants who reported poor health in all 3 waves. A similar pattern was observed for younger participants although associations were weaker. Our results are less consistent with a trajectory model. Contrary to our second hypothesis, we did not find interactions between measures of health status at different waves that would have suggested that the effects of health at one time point varied depending on health at another time. Furthermore, in our models using health pattern indicators we did not find

consistent differences in the estimated effects between patterns representing improving, declining, or mixed health status over time.

Much of the research examining health in early life and schooling has focused on birth weight and other measures of health at birth. These studies, including several studies of siblings and twins[35-37] and at least two studies using PSID data,[13,36] have found consistent associations between poor health at birth and lower educational attainment. The literature examining the influence of health at older ages during childhood on schooling is sparser. Using data from adult PSID participants, Haas found that a retrospective report of poorer health status at age 16 was associated with fewer years of completed schooling and that this association persisted in sibling fixed-effects models.[11] Similarly, several studies have related poor child or adolescent health to lower educational attainment using data from the National Child Development Survey in Britain.[8,38] On the other hand, Warren did not find evidence of a link between three retrospective measures of health before the age of 16 (self-rated health status, a scale of physician-diagnosed conditions, and the presence of activity-limiting health conditions) and years of completed schooling.[39] Two recent studies used data from the U.S. National Longitudinal Survey of Youth 1997 cohort, a group similar in age to our sample; both found that worse self-rated health among adolescents was associated with lower odds of finishing high school before age 20 or enrolling in postsecondary schooling.[12,40] Consistent with our results, Haas and Fosse reported a graded relationship in which schooling outcomes were best among adolescents who reported consistently good health status, followed by those with mixed health histories, and worst among those who reported consistently poor health.[40]

The strong health–age interactions we observed suggest that the impact of health on schooling does not clearly appear until late adolescence or early adulthood. The variability in

schooling was substantially smaller at younger than at older ages. In addition, health-related differences in educational progress may only emerge as youth enter young adulthood and make decisions about college entrance and persistence. Therefore, it is likely that our follow-up period ended before health-related differences in schooling completely manifested themselves. This is particularly true for younger participants but may also apply to older participants, many of whom may go on to complete more schooling. We considered this a worthy tradeoff for the ability to use longitudinal information including very young children, which has not been well studied in this context, but future research will require information both starting at a young age and extending past educational completion to fully characterize the effects of early health on schooling. This research may also be able to disentangle differences in how health affects schooling at different ages from possible period or cohort effects that create differences in health–schooling relations over time. Assuming our estimates represent true effects, we cannot know if these effects will be compounded when the children mature and greater schooling disparities appear or, instead, if children whose educational progress was initially slowed by poor health will eventually catch up. This may also differ by developmental stage and schooling status; for example, the extent to which poor health may discourage a participant who was in high school at the end of follow-up from staying in school likely differs from the extent to which it discourages a participant who had already dropped out to reenter school.

This analysis was subject to several other limitations. Our study relied on self- and proxy-reported data. Despite the robustness of our results to replacing self-reported health status measures with caregiver-rated health, our results may still have been affected by the differing sources of health status information at different waves. In particular, differential measurement error in our health measures from different sources may have hampered our ability to detect

trajectory-dependent differences in the effects of health status on schooling. The design of our study, with a sample of children of various ages at the start of follow-up, also limited our ability to confidently investigate whether health at a specific age was especially predictive of schooling (the critical period model) because of limited sample size at each age and the absence of strictly comparable health status measures across all ages. Future research is needed to investigate longitudinal relations between other health measures and schooling, as well as to more fully characterize the properties of self- and caregiver-rated global health status among children and adolescents.

Despite our efforts to minimize the impact of nonresponse on our results, we cannot rule out that our results may have been affected by missing information or from by the imputation process itself.[27,41] The use of multiple imputation rather than complete-case analysis allowed us to relax the assumption that data were missing completely at random (MCAR) and assume that missingness was random conditional on observed variables (missing at random [MAR]), which in most situations produces less bias, but still required the (untestable) assumption that data missingness was conditionally independent of missing values (i.e., not MNAR).[41] The large number of variables used in the imputation makes this assumption more reasonable. Children who were in very poor health may have been excluded from the initial sample, particularly because of the reliance on self-reported data; our results may not generalize to these children or may be conservative if these children's poorer health is particularly detrimental to their educational progress. Despite the robustness of our results to adjustment by additional potential confounders, they may have still been affected by omitted or mismeasured confounders; in general, one of the persistent challenges of research investigating causal links between schooling and health is the large potential for confounding by common causes of the exposure and

outcome. The similar results from our sibling fixed-effects model were somewhat encouraging in this respect. That said, one potentially important source of confounding we did not address pertains to school-related factors. Finally, the 2007 self-reports of health were collected concurrently with the outcome measure, so we cannot establish their temporal order.

Our results are consistent with prior research linking poor childhood or adolescent health with lower educational attainment and extend it to include younger children, thereby providing information about the ages at which health-related schooling differences begin to emerge. They are also consistent with research demonstrating that poorer health may affect school attendance and academic performance, both of which may slow educational progress through grade repetition or drop-out. Furthermore, the associations we found were evident in a national population of children with a low prevalence of serious health problems, suggesting that variation in health status may have implications for schooling even in the absence of debilitating health conditions. To our knowledge, our analysis is the first to address this topic using longitudinal measures of health status spanning childhood and adolescence; existing literature has focused on perinatal and adolescent health, with few addressing health in the early and middle childhood years. We were therefore able to document the emergence of health-related differences in educational progress at different ages, an important step for identifying mechanisms through which health influences schooling and ultimately for identifying the format and timing of effective interventions to reduce educational disparities. To this end, our use of a contemporary, national sample of children and youth was important in the face of ongoing secular changes in schooling and health. Our estimated effects among the younger cohorts are small. However, our estimated cumulative effect for participants aged 11–14 at baseline with

persistent poor health was larger than a year; prior studies have linked each additional year of schooling to improved health and decreased mortality.[42-44]

Our analysis adds to research suggesting complex, bidirectional relations between health and schooling. Poorer health among those with lower levels of schooling may to some extent reflect a vicious cycle between poor health and limited educational progress throughout the life course. Therefore, a better understanding of how childhood and adolescent health affect educational progress over time will help us better understand the subsequent effects of educational attainment on health. It may also contribute to the development of more effective interventions both to improve educational outcomes among disadvantaged groups and to reduce educational disparities in health. If our results our confirmed, health-related interventions early in life may serve not only to improve children's health but also to reduce the impact of poor early health on educational and health disparities in adulthood.

References

1. Adams P, Barnes P, Vickerie J. Summary health statistics for the U.S. population: National Health Interview Survey, 2007. *Vital Health Stat.* Vol. 10 National Center for Health Statistics, 2008.

2. Antonovsky A. Social class, life expectancy, and overall mortality. *The Milbank Memorial Fund Quarterly* 1967;**45**(2):31-73.

3. Cutler DM, Lleras-Muney A. Education and Health: Evaluating Theories and Evidence. *NBER Working Papers*. Cambridge, MA: National Bureau of Economic Research, 2006.

4. Feinstein L, Sabates R, Anderson TM, Sorhaindo A, Hammond C. What are the effects of education on health? In: Desjardins R, Schuller T, eds. *Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium*. Copenhagen: Organisation for Economic Co-operation and Development, 2006;171-353.

5. Grossman M. Education and nonmarket outcomes. In: Hanushek EA, Welch F, eds. *Handbook of the Economics of Education*. Handbooks in Economics. Vol. 1. Amsterdam, The Netherlands: Elsevier, 2006;577-634.

6. Kitagawa EM, Hauser PM. Education differentials in mortality by cause of death: United States, 1960. *Demography* 1968;**5**(1):318-353.

7. Mirowsky J, Ross CE. *Education, Social Status, and Health*. Hawthorne, NY: Aldine de Gruyter, 2003.

8. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *J Health Econ* 2005;**24**(2):365-89.

9. Ding W, Lehrer SF, Rosenquist JN, Audrain-McGovern J. The impact of poor health on academic performance: New evidence using genetic markers. *J Health Econ* 2009;**28**(3):578-97.

10. Eide ER, Showalter MH, Goldhaber DD. The relation between children's health and academic achievement. *Children and Youth Services Review* 2010;**32**:231-238.

11. Haas SA. Health selection and the process of social stratification: the effect of childhood health on socioeconomic attainment. *Journal of Health and Social Behavior* 2006;**47**(4):339-354.

12. Jackson MI. Understanding links between adolescent health and educational attainment. *Demography* 2009;**46**(4):671-94.

13. Johnson RC, Schoeni RF. The influence of early-life events on human capital, health status, and labor market outcomes over the life course. *The B.E. Journal of Economic Analysis & Policy* 2011;**11**(3 (Advances)):Article 3.

14. Fowler MG, Johnson MP, Atkinson SS. School achievement and absence in children with chronic health conditions. *J Pediatr* 1985;**106**(4):683-7.

15. Taras H, Potts-Datema W. Chronic health conditions and student performance at school. *J Sch Health* 2005;**75**(7):255-66.

16. Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. *J Epidemiol Community Health* 2003;**57**(10):778-83.

17. Institute for Social Research. The Panel Study of Income Dynamics. http://psidonline.isr.umich.edu/ Accessed April 5, 2010.

18. Haas SA. The long-term effects of poor childhood health: an assessment and application of retrospective reports. *Demography* 2007;**44**(1):113-135.

19. Hajat A, Kaufman JS, Rose KM, Siddiqi A, Thomas JC. Long-term effects of wealth on mortality and self-rated health status. *Am J Epidemiol* 2010;**173**(2):192-200.

20. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior* 1997;**38**:21-37.

21. Manor O, Matthews S, Power C. Dichotomous or categorical? Analysing self-rated health and lifetime social class. *Int J Epidemiol* 2000;**29**:149-157.

22. Mincer J. *Schooling, experience, and earnings*. Human behavior and social institutions, 2. New York,: National Bureau of Economic Research; distributed by Columbia University Press, 1974.

23. Hofferth S, Stafford FP, Yeung W-JJ, Duncan GG, Hill MS, Lepkowski J, Morgan JN. A Panel Study of Income Dynamics: Procedures and Codebooks. Guide to the 1993 Interviewing Year. Ann Arbor, MI: Survey Research Center, Institute for Social Research, 1998.

24. Grieger LD, Schoeni RF, Danziger S. Accurately measuring the trend in poverty in the United States using the Panel Study of Income Dynamics. *PSID Technical Series*. Ann Arbor, MI: Institute for Social Research, 2008.

25. SAS Institute Inc. SAS software, version 9.2. Cary, NC: SAS Institute Inc., 2002.

26. Raghunathan TE, Lepkowski J, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 2001;**27**(1):85-95.

27. Spratt M, Carpenter J, Sterne JA, Carlin JB, Heron J, Henderson J, Tilling K. Strategies for multiple imputation in longitudinal studies. *Am J Epidemiol* 2010;**172**(4):478-87.

28. Horton NJ, Kleinman KP. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *Am Stat* 2007;**61**(1):79-90.

29. Berglund P. An introduction to multiple imputation of complex sample data using SAS v9.2. Proceedings of the SAS Global Forum 2010 Conference. Cary, NC: SAS Institute Inc., 2010.

30. Backlund E, Sorlie PD, Johnson NJ. The shape of the relationship between income and mortality in the United States. Evidence from the National Longitudinal Mortality Study. *Ann Epidemiol* 1996;**6**(1):12-20; discussion 21-2.

31. Ecob R, Smith GD. Income and health: what is the nature of the relationship? *Soc Sci Med* 1999;**48**(5):693-705.

32. Allison P. Fixed effects regression methods in SAS. *Proceedings of the Thirty-first Annual SAS Users Group International Conference*. Cary, NC: SAS Institute Inc., 2006.

33. Johnson SB, Wang C. Why do adolescents say they are less healthy than their parents think they are? The importance of mental health varies by social class in a nationally representative sample. *Pediatrics* 2008;**121**(2):e307-e313.

34. Waters E, Stewart-Brown S, Fitzpatrick R. Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. *Child Care Health Dev* 2003;**29**(6):501-9.

35. Behrman JR, Rosenzweig MR. Returns to birthweight. *The Review of Economics and Statistics* 2004;**86**(2):586-601.

36. Conley D, Bennett NG. Is biology destiny? Birth weight and life chances. *Am Sociol Rev* 2000;**65**(3):458-467.

37. Oreopoulos P, Stabile M, Walld R, Roos LL. Short-, medium-, and long-term consequences of poor infant health. *Journal of Human Resources* 2008;**43**(1):88-138.

38. Chandola T, Clarke P, Morris JN, Blane D. Pathways between education and health: a causal modelling approach. *Journal of the Royal Statistical Society* 2006;**169**(2):337-359.

39. Warren JR, Hauser RM. Social stratification across three generations: new evidence from the Wisconsin Longitudinal Study. *American Sociological Review* 1997;**62**(4):561-572.

40. Haas SA, Fosse NE. Health and the educational attainment of adolescents: evidence from the NLSY97. *J Health Soc Behav* 2008;**49**(2):178-92.

41. Rubin DB. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 1996;**91**(434):473-489.

42. Lleras-Muney A. The relationship between education and adult mortality in the United States. *The Review of Economic Studies* 2005;**72**(1):189-221.

43. Ross CE, Mirowsky J. Refining the assocation between education and health: the effects of quantity, credential, and selectivity. *Demography* 1999;**36**(4):445-460.

44. Shuey KM, Willson AE. Cumulative disadvantage and black-white disparities in life-course health trajectories. *Research on Aging* 2008;**30**(2):200-225.

Figure 2.1 Study sample from Child Development Supplement (CDS) and Transition into Adulthood (TA) supplemental studies to the Panel Study of Income Dynamics, 1997–2007

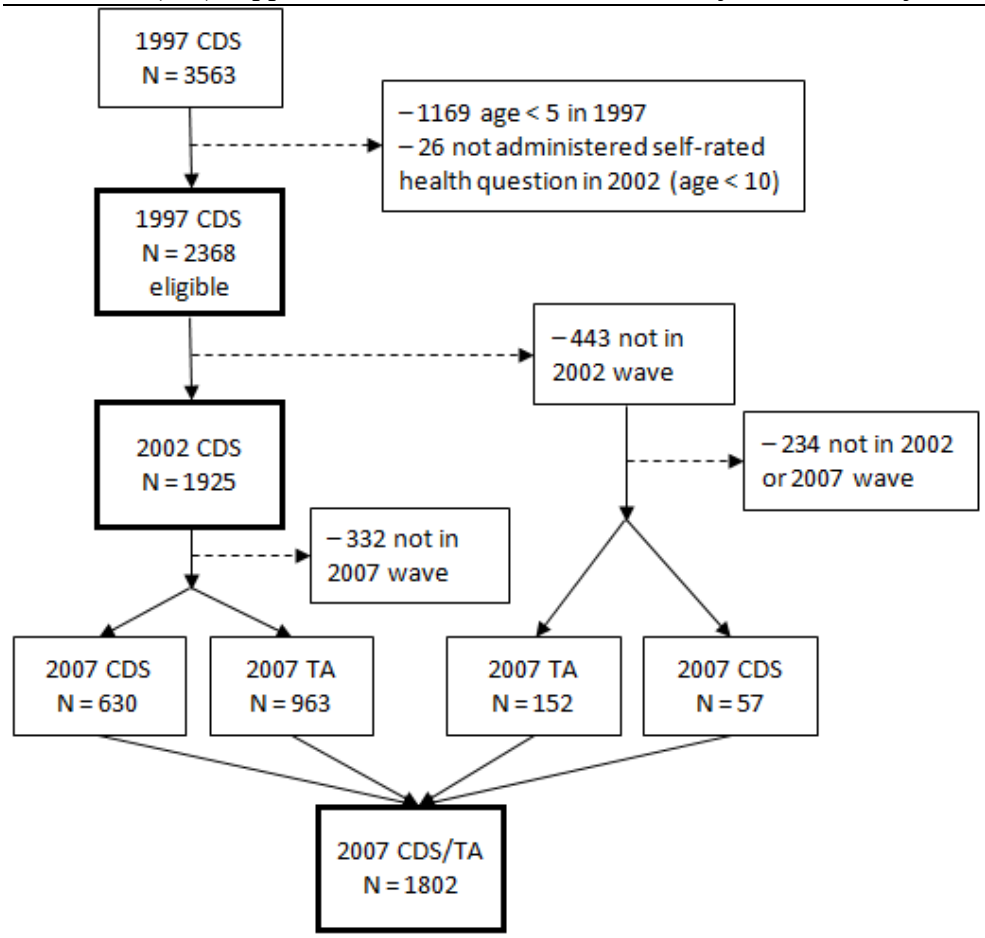
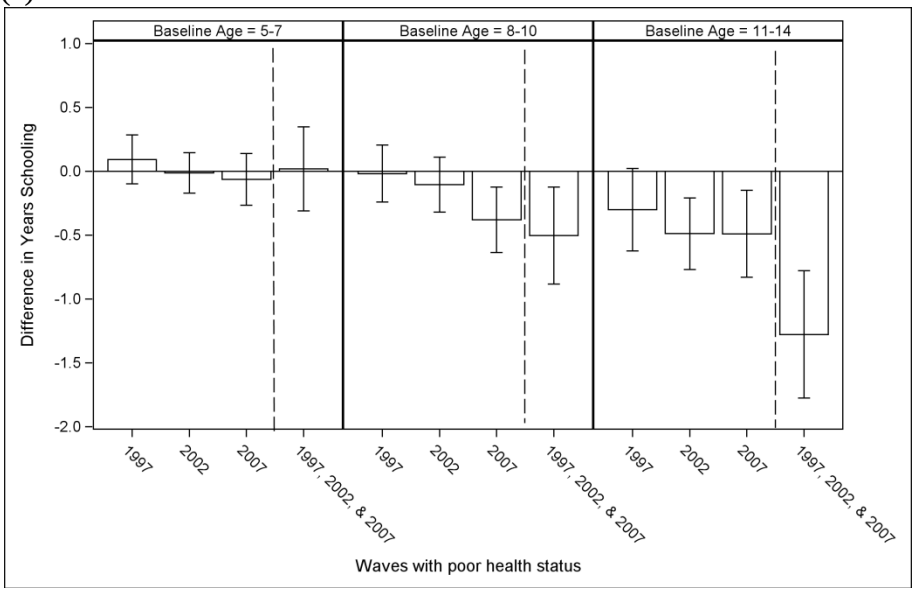
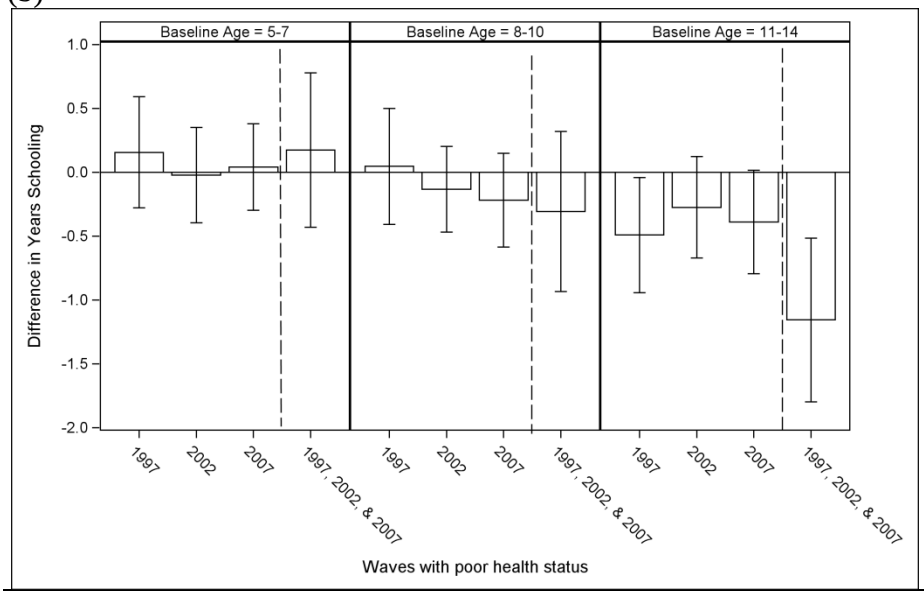


Figure 2.2 Adjusted mean differences in completed years of schooling in 2007, comparing poor health status to good health status separately in each wave, by age in 1997 among (a) entire sample (N = 2368) and (b) siblings only (fixed effects model; N = 1242)

(a)



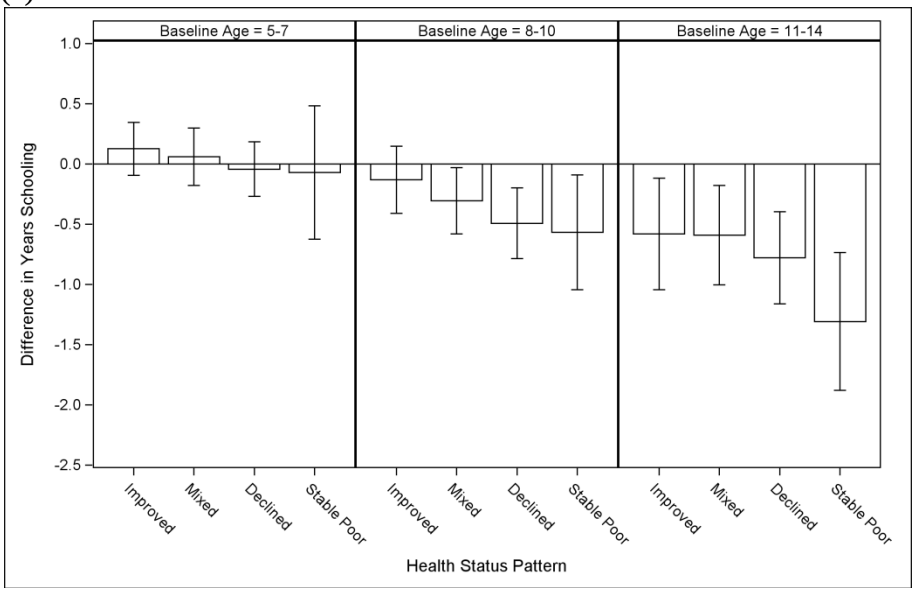
(b)



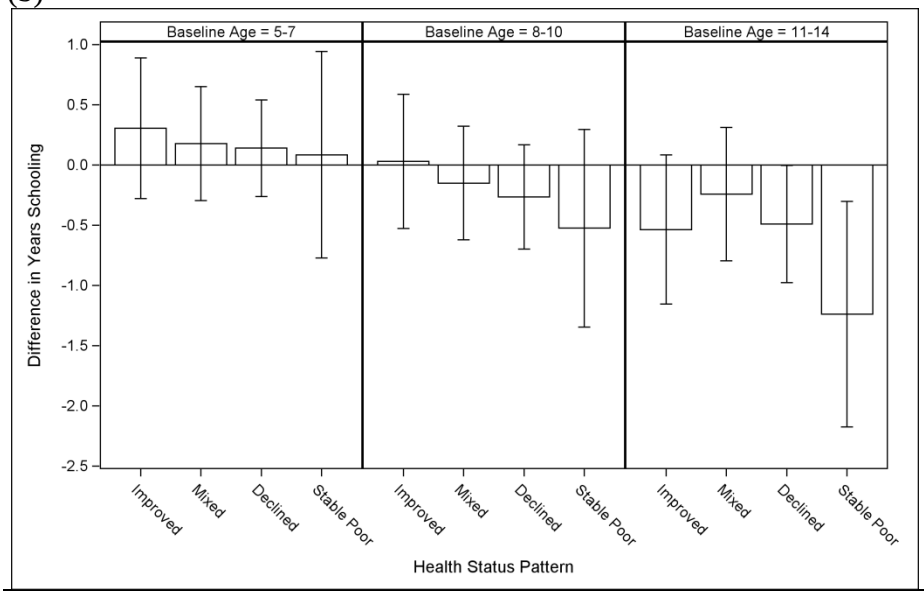
Linear regression models including a separate term for poor health status at each wave. Therefore, single-wave comparisons (poor 1997 health status, poor 2002 health status, and poor 2007 health status) compare participants with poor health status to those with good health status in that wave, adjusting for health status in the other waves. The final comparison (poor 1997, 2002, and 2007 health status) combines coefficients to compare poor health status in all 3 waves to good health status in all 3 waves. (a) Results from Model 3 in Table 3. Model was additionally adjusted for sex, baseline age, race/Hispanicity, kindergarten attendance, baseline years completed schooling, family income in 1996 and 2002, and caregiver education level. (b) Model was additionally adjusted for sex, baseline age, kindergarten attendance, and baseline years completed schooling.

Figure 2.3 Adjusted mean differences in completed years of schooling in 2007, comparing each pattern of health-status change during follow-up to stable good health in all 3 waves, by age in 1997 among (a) entire sample (N = 2368) and (b) siblings only (fixed effects model; N = 1242)

(a)



(b)



Linear regression models with a separate indicator for each health-status pattern, using participants with good health status in all 3 waves as the referent group. (a) Model was additionally adjusted for sex, baseline age, race/Hispanicity, kindergarten attendance, baseline years completed schooling, family income in 1996 and 2002, and caregiver education level. (b) Model was additionally adjusted for sex, baseline age, kindergarten attendance, and baseline years completed schooling.

Table 2.1 Sample characteristics

Characteristic	Original Sample N or range	Original Sample %[a] or mean (SD)	Imputed Sample % or mean (SD)
Total	2368	---	---
Sex			
Male	1195	50	50
Female	1173	50	50
Age in years 1997	4.5–13.9	8.9 (2.6)	8.9 (2.6)
Age in years 2007	14.8–24.0	19.4 (2.5)	19.3 (2.5)
Missing	633	27	---
Race/Hispanicity			
non-Hispanic white	1069	45	45
non-Hispanic black	988	42	42
Hispanic	187	8	8
Other race	121	5	5
Missing	3	< 1	---
Family income/poverty ratio 1996	0.0–37.4	3.0 (2.9)	2.9 (2.9)
Missing	211	9	---
Family income/poverty ratio 2002	0.0–112.5	3.7 (5.0)	3.6 (5.0)
Missing	168	7	---
Maximum of caregivers' completed years schooling 1997			
<9	84	4	4
9–11	232	10	10
12	847	36	36
13–15	631	27	27
≥16	550	23	23
Missing	24	1	---
Caregiver-rated health 1997			
Good	1897	81	80
Poor	456	19	20
Missing	15	< 1	---
Self-rated health 2002			
Good	1157	68	62
Poor	556	32	38
Missing	655	28	---
Self-rated health 2007			
Good	1111	64	55
Poor	627	36	45
Missing	630	27	---
Attended kindergarten 1997			
Yes	2055	89	88
No	246	11	12
Missing	67	3	---
Years completed schooling 1997	0–7	2.3 (2.1)	2.3 (2.1)
Missing	21	1	---
Years completed schooling 2007	5–17	11.5 (2.1)	11.4 (2.0)
Missing	571	24	---

[a] To facilitate comparison with the imputed sample, the denominators for percents for nonmissing categories include only observations with nonmissing information.

Table 2.2 Percent with poor health status and mean years of completed schooling in 2007, by category of selected predictors (N = 2368)[a]

Predictor category	Poor caregiver-rated health status 1997		Poor caregiver-rated health status 2002		Poor self-rated health status 2002		Poor self-rated health status 2007		Years completed schooling 2007	
	%	p[b]	%	p[b]	%	p[b]	%	p[b]	Mean	p[c]
Total	20	--	25	--	38	--	45	--	11.4	
Sex		0.19		0.71		<0.001		0.02		<0.001
Male	21		25		34		42		11.3	
Female	18		25		42		48		11.6	
Age at 1997 interview (years)		0.33		0.07		0.04		0.10		<0.001
5–7	20		23		34		49		9.5	
8–10	21		24		38		42		11.8	
11–14	18		29		42		44		13.1	
Race/Hispanicity		<0.001		<0.001		0.05		0.11		0.007
non-Hispanic white	11		20		35		44		11.6	
non-Hispanic black	27		27		40		44		11.3	
Hispanic	25		39		47		51		11.4	
Other	23		29		41		55		11.1	
Family income/poverty ratio 1996		<0.001		<0.001		<0.001		<0.001		<0.001
0–0.9	32		34		41		50		11.0	
1–1.9	23		29		46		48		11.1	
2–4.9	15		21		37		44		11.7	
≥5	8		17		25		34		12.0	
Family income/poverty ratio 2002		<0.001		<0.001		0.02		0.02		<0.001
0–0.9	34		38		45		51		10.9	
1–1.9	29		29		41		48		11.1	
2–4.9	15		22		38		44		11.5	
≥5	9		17		32		40		12.1	
Maximum of caregivers' completed years education 1997		<0.001		<0.001		<0.001		0.01		<0.001
<9	31		48		51		55		11.3	
9–11	30		38		43		53		10.8	
12	24		27		42		47		11.2	
13–15	15		22		37		42		11.6	
≥16	11		16		30		41		12.0	
Caregiver-rated health 1997		--		<0.001		0.01		0.10		<0.001
Good	--		21		37		44		11.5	
Poor	--		42		45		49		11.1	
Self-rated health 2002		0.01		0.04		--		<0.001		0.67
Good	18		19		--		38		11.5	
Poor	23		34		--		55		11.4	
Self-rated health 2007		0.10		0.03		<0.001		--		<0.001
Good	18		22		31		--		11.7	
Poor	21		28		47		--		11.1	
Attended kindergarten 1997		0.23		0.36		0.04		0.46		<0.001

Yes	19		25		39		45		11.7	
No	22		22		30		47		9.2	
Years completed schooling										
1997		0.24		0.14		0.06		0.14		<0.001
0	20		22		34		48		9.5	
1–4	20		25		40		42		12.0	
5–7	17		29		42		45		13.3	
Completed years schooling										
2007		<0.001		0.39		0.19		<0.001		--
<9	25		29		35		56		--	
9–11	22		25		38		50		--	
12	22		26		42		43		--	
13–15	13		21		36		38		--	
≥16	10		25		27		25		--	

[a] Uses imputed data.

[b] From unadjusted logistic regression with clustering by sibling pair.

[c] From unadjusted linear regression with clustering by sibling pair.

Table 2.3 Mean differences in completed years of schooling 2007 associated with health status at different times and covariates (N = 2368)

	Model 1			Model 2			Model 3			Model 4		
Covariate	Coeff	95% CI	p	Coeff	95% CI	p	Coeff	95% CI	p	Coeff	95% CI	P
Poor caregiver-rated health 1997	-0.44	(-0.66, -0.23)	<0.001	-0.14	(-0.32, 0.05)	0.15	0.09	(-0.10, 0.29)	0.34	0.02	(-0.23, 0.27)	0.89
Poor self-rated health 2002	0.06	(-0.18, 0.31)	0.61	-0.03	(-0.19, 0.12)	0.67	-0.01	(-0.17, 0.15)	0.88	-0.09	(-0.30, 0.13)	0.45
Poor self-rated health 2007	-0.54	(-0.81, -0.27)	<0.001	-0.08	(-0.29, 0.12)	0.42	-0.06	(-0.26, 0.14)	0.54	-0.12	(-0.36, 0.12)	0.34
Age 1997 (per year; centered at 9)				0.58	(0.51, 0.65)	<0.001	0.35	(0.23, 0.48)	<0.001	0.35	(0.28, 0.48)	<0.001
Age 1997 squared term				-0.05	(-0.06, -0.03)	<0.001	-0.06	(-0.08, -0.04)	<0.001	-0.06	(-0.08, -0.04)	<0.001
Age group 1997 (vs. 5–7)												
8–10				0.34	(0.07, 0.60)	0.01	0.19	(-0.08, 0.46)	0.17	0.19	(-0.09, 0.46)	0.18
11–14				0.59	(0.18, 1.01)	0.005	0.41	(0.01, 0.81)	0.04	0.41	(0.01, 0.81)	0.05
(Poor health) × (Age group 1997)												
interactions						<0.001[a]			<0.001[a]			<0.001[a]
1997 × (Age 8–10)				-0.11	(-0.40, 0.18)	0.46	-0.11	(-0.40, 0.18)	0.45	-0.10	(-0.39, 0.19)	0.48
1997 × (Age 11–14)				-0.41	(-0.80, -0.02)	0.04	-0.39	(-0.76, -0.03)	0.03	-0.39	(-0.76, -0.02)	0.04
2002 × (Age 8–10)				-0.13	(-0.39, 0.13)	0.34	-0.09	(-0.34, 0.15)	0.46	-0.08	(-0.33, 0.17)	0.52
2002 × (Age 11–14)				-0.53	(-0.84, -0.22)	0.001	-0.48	(-0.77, -0.18)	0.002	-0.47	(-0.77, -0.19)	0.002
2007 × (Age 8–10)				-0.36	(-0.60, -0.12)	0.003	-0.32	(-0.54, -0.09)	0.006	-0.32	(-0.55, -0.09)	0.006
2007 × (Age 11–14)				-0.45	(-0.80, -0.11)	0.01	-0.43	(-0.75, -0.10)	0.01	-0.43	(-0.75, -0.11)	0.01
Female (vs. male)							0.30	(0.19, 0.41)	<0.001	0.30	(0.19, 0.41)	<0.001
Race/ethnicity (vs. White)									0.79[a]			0.83[a]
Black							-0.02	(-0.16, 0.12)	0.75	-0.02	(-0.16, 0.12)	0.77
Hispanic							0.11	(-0.16, 0.38)	0.44	0.11	(-0.16, 0.38)	0.44
Other race							0.03	(-0.28, 0.34)	0.86	0.03	(-0.28, 0.34)	0.86
Attended kindergarten 1997							0.24	(-0.01, 0.50)	0.06	0.25	(-0.01, 0.50)	0.06
Years completed schooling 1997							0.29	(0.16, 0.43)	<0.001	0.29	(0.16, 0.43)	<0.001
ln(Family income/poverty ratio 1996)							0.04	(-0.11, 0.19)	0.61	0.04	(-0.11, 0.19)	0.59
ln(Family income/poverty ratio 2002)							0.33	(0.19, 0.47)	<0.001	0.33	(0.18, 0.47)	<0.001
Maximum of caregivers' completed years												
schooling 1997 (vs. ≥16)									<0.001[a]			<0.001[a]
13–15							-0.30	(-0.47, -0.13)	<0.001	-0.30	(-0.47, -0.14)	<0.001
12							-0.54	(-0.74, -0.34)	<0.001	-0.54	(-0.74, -0.34)	<0.001
9–11							-0.74	(-1.00, -0.48)	<0.001	-0.74	(-1.01, -0.48)	<0.001
<9							-0.47	(-0.90, -0.03)	0.03	-0.47	(-0.90, -0.03)	0.03
(Poor health) × (Poor health) interactions												0.83[a]
1997 × 2002										0.17	(-0.26, 0.59)	0.44
1997 × 2002										0.17	(-0.20, 0.53)	0.38
2002 × 2007										0.14	(-0.18, 0.46)	0.40
1997 × 2002 × 2007										-0.33	(-0.97, 0.30)	0.30

[a] Joint p-value.

CHAPTER 3

DOES ACADEMIC ACHIEVEMENT DURING CHILDHOOD AND ADOLESCENCE BENEFIT LATER HEALTH?

Introduction

Associations between more schooling and less adult morbidity and mortality are large, persistent, and well documented.[1-3] However, there are still important questions regarding the mechanisms involved. One key mechanism through which schooling may influence health is by providing access to higher-paying jobs, consequent higher income, and the accompanying access to healthier living conditions, nutritious foods, medical care, and time for rest and exercise. [4-6] Numerous studies have documented the many occupational and economic benefits of schooling and many of these factors have in turn been linked to health later in adulthood.[7-9] However, there may also be purely cognitive and psychosocial benefits of learning and academic achievement per se. These mechanisms have been less frequently examined even though the persistence of educational disparities in adult health after adjustment for financial circumstances suggests their existence. Cognitive mechanisms through which schooling may affect health include the acquisition of knowledge about how to live healthfully and improvements in abilities to read, understand, or act on health information or better use the resources available to benefit health.[3,10] Psychosocial mechanisms include an increased sense of personal control; the patience and time preference necessary to make healthier choices; access to higher social integration or social support; or exposure to more healthful behavioral norms.[5,6,11,12]

If these cognitive and psychosocial mechanisms are important, one would expect associations between schooling, or more specifically the academic achievement resulting from the learning process, and health to be apparent early in life, even before the process of schooling is itself completed and socioeconomic sequelae manifest. This is in contrast to occupation- or income-mediated effects, which would not appear until students enter the workforce. Students who more successfully make these cognitive and psychosocial gains—i.e., who experience the benefits of schooling to a greater degree (for example, as reflected in their academic achievement)—may enjoy better health even before completion of their schooling. Yet very few studies have investigated how health in early life relates to academic achievement. Most of these studies have used cross-sectional data or short follow-up times, although a few have examined academic achievement in late adolescence in relation to substance use several years later.[13-20]

An important challenge in studying the impact of academic achievement on health is the possibility of reverse causation: students' health may influence their success in school. [21-25] For example, poor health may affect students' ability to attend school regularly, pay attention in classes, complete assignments, or interact with peers and teachers.[26,27] It may also affect their ability or willingness to invest in long-term educational goals.[24] The causal diagram (DAG) in Figure 3.1 shows this theorized mutual influence of health and academic achievement on each other across time. This mutual influence of achievement and health on each other presents a methodological challenge in longitudinal analyses examining the total effect of achievement on health. Because health status during follow-up (health at Time 2 in Figure 3.1) both mediates the effect of achievement at Time 1 and confounds the effect of achievement at Time 2 (i.e., is a time-dependent confounder), conventional analytical approaches that adjust for early health may underestimate the causal effect of achievement on health (due to overadjustment for a mediator)

while failing to adjust for early health may also lead to incorrect inferences because of confounding bias.[28]

We used a longitudinal data set with three waves of data spanning 10 years and rich demographic, socioeconomic, academic, and health information throughout the life course of a national sample of children, adolescents, and young adults to estimate longer-term effects of academic achievement on health during the schooling process. We addressed potential time-dependent confounding of the academic achievement–health associations by using marginal structural models (MSMs) using inverse-probability weighting instead of adjustment to account for covariate imbalances.[29] We hypothesized that a history of greater academic achievement decreases the risk of poor health among adolescents and young adults. We also hypothesized that conventional models adjusting for health status during follow-up underestimate the effect of academic achievement on health. Finally, we tested differences by sex in associations between academic achievement and future health. Past studies have found differences between boys and girls in how academic achievement is related to health. For example, stronger associations have been observed among girls than among boys between lower academic achievement and poor self-rated health status, depression, and somatic complaints.[13,15,17] These differences may form a basis for differences between the sexes in the effects of schooling on adult health and mortality.[30-32]

Methods

Study Population

The study sample came from the Child Development Supplement (CDS) and Transition into Adulthood Study (TA), two supplementary studies to the Panel Study of Income Dynamics (PSID) that focus on children, adolescents, and young adults.[33] The PSID, a longitudinal study of

a representative sample of U.S. families, was started in 1968 and contains data on about 70,000 individuals. Interviews were conducted annually until 1997, when the study switched to biennial data collection. A primary respondent in each family reports on behalf of the family.

The CDS began in 1997; all PSID families with a child aged 0–12 in calendar year 1997 were eligible to participate, with up to two children chosen per family.[34] Black and low-income families were oversampled. The study consisted of extensive interviews with 3,563 children and their guardians (including absent fathers), teachers, and school administrators. In 2002–2003 a second wave of interviews was carried out with 2,907 children aged 5–18 who had participated in the first wave. In 2007–2008 a third wave of interviews was carried out with children who were still under the age of 18 in 2007. The primary interviewee in the CDS is the participant's primary caregiver, with whom the child must be living. In the vast majority (over 90%) of cases, this is the child's biological mother. If the biological mother is not living with the child, the primary caregiver is defined using the following order of preference: (1) stepmother, adoptive mother, or foster mother; (2) other female legal guardian at least 18 years old (often the grandmother); (3) biological, step-, adoptive, or foster father; (4) other male legal guardian at least 18 years old; (5) another unpaid adult who lives with the child and takes primary responsibility for his/her care.[35]

When CDS participants turn 18, they become ineligible to continue participation in the CDS. Instead, these teenagers become eligible to participate in the TA. The TA began in 2005 with pilot telephone interviews of 745 young adults aged 18 and over who had participated in the baseline 1997 CDS and had finished or left high school.[36] In 2007, a second wave of TA interviews was carried out with the same young adults and additional CDS participants who had

turned 18 and finished or left high school since 2005, for a total of 1,118 interviews. All questions in the TA are answered by the participants themselves.

The sample used in this study combined data from the 1997 CDS, the 2002–2003 CDS (called "2002 CDS" in the remainder of this article), and either the 2007–2008 CDS ("2007 CDS") or the 2007 TA to create a longitudinal sample with three waves of interviews per child spanning 10 years. We included only children who were at least 3 years old (younger children were not administered achievement tests) and enrolled in a childcare center or school as of the 1997 baseline interview. The analysis sample included 2,546 children aged 3–14 years by the time their baseline interviews were completed. At the time of the third and final interview in 2007–2008, the sample ranged in age from 14–24 years (Figure 3.2).

Measures

The primary outcome variable for our analysis was global health status assessed in 2007. Health status in prior waves was considered a time-dependent confounder (see Figure 3.1). Health status at each time-point was measured as a 5-category ordinal variable corresponding to the responses "excellent," "very good," "good," "fair," and "poor." Because the survey questionnaire content varied depending on the age of the participant, and the fact that we combined data from the CDS and TA, the source of the measure varied by wave: it was reported by the participant's primary caregiver in 1997 and 2002 but by the participant himself or herself in 2007.

In secondary analyses we examined two additional health outcomes, body mass index (BMI) and mental health, which might explain an association between academic achievement and global health status. Weight status and mental health have been shown to be predictive of self-rated health status in young people.[37,38] As in the case of global health, the outcome of

interest was the health outcome at the last follow-up (2007), with prior measures of health being treated as time-dependent confounders. BMI was calculated at each wave using height and weight measurements taken by the interviewer in the CDS (with the exception of baseline height, which was reported by the participant's caregiver) and reported by the participant in the TA. As a measure of mental health, we used the K6 non-specific psychological distress scale developed by Kessler et al. for use in the U.S. National Health Interview Survey (NHIS).[39] Analyses using the K6 were limited to participants aged 18 or over in 2007 because the scale was administered in the TA but not the CDS. Because of differences in the measures used in the CDS and TA, for the 2002 study wave the measure of mental health we used was the Children's Depression Inventory (CDI) Short Form, which assesses depressive symptoms over the past two weeks.[40] No scale of mental health was available at baseline.

The primary exposure variable was an averaged measure of academic achievement in 1997–2002. We used academic achievement as a proxy measure of each child's accrual of the cognitive benefits of education. We used an average measure over the two study waves to provide a more stable estimate of academic achievement over 5 years in order to capture the cumulative effects of higher achievement on later health. Similarly, we chose not to conduct a repeated-measures analysis because our goal was to estimate associations between longer-term achievement and later health; in addition, the different sources of the health measures in 2002 and 2007 would have made it inappropriate to combine them into a single outcome measure. We characterized academic achievement by using the Letter-Word Identification and Applied Problems subtests of the Woodcock-Johnson Psycho-Educational Battery-Revised Achievement Tests (WJ-R ACH); these subtests test aspects of the test-taker's stores of acquired knowledge.[40,41] The WJ-R is a battery of 39 tests developed according to the *Gf-Gc* (acronym for

"fluid and crystallized abilities") theory of intellectual processing.[40-42] The Letter-Word Identification test asks participants to identify or correctly pronounce letters or words without context; it is not necessary to know the meanings of the words. The Applied Problems consists of progressively more difficult math questions, ranging from counting a number of objects to word problems. The tests have standardized administration and scoring protocols, were norm-referenced to a nationally standardized sample of 6,359 people aged 2–90 and have been widely used in a variety of populations.[40,41] CDS participants were administered the subtests and assigned age-standardized scores at both the 1997 and 2002 interviews. We averaged each child's Letter-Word Identification and Applied Problems scores to create a measure of overall academic achievement in each wave. We then averaged these measures over the 1997 and 2002 wave to create a measure of average 5-year academic achievement for each child. This measure was normally distributed and ranged 42.0–169.5.

In addition to the time-dependent confounding effects of prior health, we also considered a range of time-invariant and time-varying characteristics that could confound the effects of academic achievement on subsequent health. We capitalized on the richness of the PSID data to incorporate information from nine broad domains that may influence both schooling and future health in young people: demographic characteristics, family composition, family socioeconomic status, neighborhood characteristics, geographical characteristics, perinatal health, health care use, schooling characteristics, and caregiver academic ability. The variables we used to correspond to these domains are listed in Appendix Table 3.A1.

Analysis

To handle missing information (both item and case missingness), multiple imputation with 25 imputations was carried out using the sequential regression method with IVEware

software and simultaneously including information from all three study waves.[43,44] The multiple imputation assumed data were missing at random (MAR) rather than missing completely at random (MCAR) and allowed for standard errors accounting for variability in the imputation process; a recent study found multiple imputation to produce less biased results than other methods of handling missing confounder information in the context of MSMs.[45,46] Using the sequential regression method allowed us to specify an appropriate distribution for each variable, as well as to restrict imputation to relevant observations. The latter is an important consideration in this data set because measures differed by study wave and participant age. In order to improve the imputation, we included variables predictive of the values of analysis variables with large amounts of missing information in the imputation process (Table 3.1).[47]All analyses were then conducted separately in each imputed data set and the estimates were subsequently combined using the SAS MIANALYZE procedure, which averages point estimates across imputations and derives standard errors by combining information about variance within and between imputations.[48]

We estimated three types of models for each outcome. First, we estimated models adjusted for baseline characteristics that might confound the association between average academic achievement and 2007 health. Second, we estimated models that were additionally adjusted for time-varying characteristics that might confound the association between academic achievement at each time-point and subsequent health in order to test our hypothesis that these models would produce biased results. Third, we estimated marginal structural models (MSMs) to address possible time-dependent confounding of our associations by prior health through inverse-probability-of-exposure weighting rather than adjustment.[29] We also tested interactions in all models to investigate differences between males and females in how academic

achievement is related to health. Although we chose these pooled models including interaction terms because of sample size considerations, in sensitivity analyses separate sex-specific models produced very similar findings. All models were estimated with robust standard error estimates and clustered by sibling pair.

To estimate the weighted models (MSMs), a stabilized treatment (i.e., exposure) weight sw_i was constructed for each participant in a manner analogous to the one used by Cerdà et al.[49] Details about the estimation of the weights and the variables used are located in the Appendix. The weights calculation produced a small number of extreme values. Out of a total of 63,650 weights (2546 participants with 25 imputed observations each), there were seven weights larger than 100, with the largest being 5.6×10^{7}, and four weights smaller than 0.01, with the smallest being 1.2×10^{-8}. In order to produce a reasonable distribution of weights for the MSM models of global health status and BMI, (i.e., mean approximately 1 and reasonable range), we trimmed the weights separately for each imputation at their 3rd-highest and 3rd-lowest values. In other words, within each imputation the observations with the highest and second-highest weights were reassigned the third-highest weight while the observations with the lowest and second-lowest weights were reassigned the third-lowest weight. This translated into changing the value of four weights per imputation, or 0.16% of the sample.[50] The imputation-specific means of the trimmed weights ranged 1.00–1.05; the range of weights over all imputations was 0.07–32.97. We had to trim the weights for the psychological distress models at the 1st and 99th percentiles because the distribution of weights was more unstable, resulting in imputation-specific means ranging 0.99–1.01 and a total range 0.25–2.73. Further trimming the weights produced nearly identical results in our analyses. Final marginal structural outcome models of 2007 health with average academic

achievement as the exposure were estimated by incorporating the exposure weights.[49,51] The weighted (MSM) models were adjusted for baseline characteristics.[52]

We dichotomized health status into "very good" (excellent or very good) and "poorer" (good, fair, or poor) and used the modified Poisson regression method developed by Zou and Donner[53,54] to calculate prevalence ratios for poor health status. We dichotomized the health status measures from 1997 and 2002 in the same way. We chose this dichotomization instead of excellent/very good/good vs. fair/poor because in this young, relatively healthy population, the number of participants with fair or poor health was very small, particularly in the earlier study waves. Ordinal logistic regression models provided consistent results but violated the proportional odds assumption.

We log-transformed the BMI measure because of its skewed distribution and then modeled it using linear regression. It is likely that academic achievement relates differently to BMI among the underweight. This interpretation is supported by the inverted J-shaped bivariable relations between BMI status and academic achievement we observed in our sample (Table 3.2). Therefore, we repeated analyses of the BMI outcome excluding observations underweight in 1997 ($< 5^{th}$ percentile according to the Centers for Disease Control & Prevention 2000 growth; table 2).

We used modified Poisson regression[53,54] to calculate prevalence ratios for being in the highest quartile of the K6 scale ("serious psychological distress") in 2007. We were not able to use the cutoff of 13 often used to identify severe psychological distress in the published literature because of the limited sample size and low prevalence in our sample of participants meeting this threshold (about 4%).[55] However, the sensitivity of the scale using the cutoff of 13 was quite low in validation studies and there is evidence that a lower cutoff identifies clinically relevant

distress.[55,56] It was not possible to estimate risk ratios with these data because the differences in health measures across study waves precluded measurement of incidence of the health outcomes.

Results

There was substantial and nonmotonic case missingness in the data: 67% of participants completed all three study waves, 23% missed either wave 2 or wave 3, and 10% missed both wave 2 and wave 3 (Figure 3.2). Twenty-nine and 27% of participants were missing academic achievement scores in 1997 and 2002, respectively (Table 3.2). In 2007, 26% of participants were missing self-rated health information, 28% were missing BMI information, and 21% of participants aged 18 years or over were missing psychological distress information.

Variable distributions were very similar in the original and imputed data sets (Appendix Table 3.A2). About 70% of participants lived with two parents at baseline and two-thirds had a sibling in the sample. About 46% were non-Hispanic White, 42% were non-Hispanic Black, and 8% were Hispanic. At baseline, 21% lived below the Federal poverty line and 18% had no caregiver with a high school degree. Most participants had good overall health status, were of normal weight, and reported few mental distress symptoms. Table 3.1 shows associations of academic achievement and 2007 health status with other variables. Higher academic achievement was associated with higher family income and caregiver education, a higher caregiver achievement score, a more supportive home environment, a better-quality neighborhood, residence in the Northeast, private school attendance, better health status, and normal BMI status. Lower academic achievement was associated with living with fewer than two parents, low birth weight, a greater number of hospital stays, more behavior problems, changing schools, having repeated a grade, and more mental distress symptoms. Poorer self-rated health status was associated with female sex, lower family income and caregiver schooling, poorer

previous caregiver-rated health status, high BMI at all time-points and underweight BMI status in 2007, higher levels of 2002 depression symptoms and 2007 psychological distress symptoms, and more problem behaviors in 1997 and 2002. Self-rated health status also varied in the expected directions with the measures of perinatal health and overnight hospital stays.

In regression models, lower average academic achievement 1997–2002 was associated with poorer health status in 2007 (Table 3.2). There was evidence of effect measure modification by sex, such that the association was stronger among girls than among boys (p-interaction = 0.12 in the weighted model). In the weighted model, the prevalence ratio (PR) for one standard deviation higher academic achievement was 0.87 (95% confidence interval: 0.78-0.97) in girls and 0.96 (0.86-10.8) in boys. The standard deviation of average academic achievement was calculated separately in each imputed data set; it ranged 14.7–15.1, with a mean of 14.9. Compared to the unadjusted model, the adjusted models and MSMs produced wider confidence intervals but only minimally different point estimates.

Models of the other health outcomes produced results generally consistent with the ones of health status. Models excluding observations underweight in 1997 produced nearly identical estimates. In the weighted model, each standard deviation higher academic achievement was weakly associated with a 1.4% lower BMI (Table 3.3; % difference = -1.36 [-3.12, 0.44]) among girls but a negligible difference among boys (% difference = -0.14 [-1.60, 1.35]). In the combined models and among girls, estimates from the different models of BMI were indicative of the pattern we would expect in the presence of time-dependent confounding: the estimated differences produced by the weighted model were smaller in magnitude than those produced by the unweighted baseline-adjusted model but larger than those produced by the unweighted model additionally adjusted for time-varying covariates. Among boys, the point estimate from the MSM

was slightly smaller than from the unweighted model adjusting for time-varying covariates, but the estimates were very similar to each other and also very small. Among participants aged 18 and over in 2007, in the weighted model, each standard deviation higher academic achievement was associated with a lower prevalence of serious psychological distress among girls (Table 3.4; PR = 0.82 [0.64–1.05]), but not among boys (PR = 0.98 [0.79–1.22]). Similar to the models of health status, point estimates were very similar between the different adjusted models but less precise.

In sensitivity analyses, additionally incorporating baseline measures of caregiver-reported diagnosis by a medical professional of a chronic physical condition (asthma, epilepsy, heart condition, diabetes, or sickle cell anemia), sensory or movement impairment (speech impairment, orthopedic impairment, difficulty seeing, or difficulty hearing), or emotional or developmental condition (developmental delay, emotional disturbance, autism, mental retardation, or hyperactivity) produced nearly identical results. Including only observations with nonmissing exposure and outcome information and incorporating censoring weights into the marginal structural models did somewhat alter point estimates but produced qualitatively similar results. Associations among girls were larger in magnitude for health status and BMI than in the models using fully imputed data. For poor health status, PR = 0.77 (0.65–0.90) among girls and PR = 0.98 (0.83–1.16) among boys for 1-standard-deviation higher average academic achievement. The percent difference in BMI associated with 1-standard-deviation higher achievement was -3.47 (-5.49– -1.41) among girls and 0.37 (-1.60–2.37) among boys. For serious psychological distress, PR = 0.81 (0.55–1.19) among girls and PR = 0.84 (0.59–1.19) among boys.

Discussion

In a national sample of U.S. youth aged 3–14 at baseline, higher average academic achievement 1997–2002 was associated with better health in 2007(better global health status, lower BMI, less serious psychological distress) among girls but consistently less so for boys. The gender pattern we observed mirrors results from studies of similar outcomes in adults. Ross et al.[32] and Liu and Hummer[30] both found using National Health Interview Survey (NHIS) data that associations between more schooling and better self-rated health status were stronger among women than among men. Similarly, there is evidence that lower schooling levels are more strongly related to depression[3,57] and obesity[3,58,59] in women than men, although this finding has not been universal for obesity.[60] Studies of gender differences in associations between academic achievement and health among youth are rarer and their evidence is mixed. Brolin Laftman and Modin found that school performance was slightly more strongly associated with subjective health complaints among girls than boys in a sample of Hungarian 9th-graders.[17] Among adolescents in Norway, Undheim and Sund found that school grades were more predictive of depressive symptoms a year later among girls than boys.[13] On the other hand, gender differences were not reported in relations between lower school performance and worse self-rated health among students in Hungary[15] or overweight among American adolescents.[61] Our results suggest the need for future research examining gender differences in how schooling relates to health. These differences may reflect objective mechanistic differences in how schooling influences health, as well as differences in how men and women incorporate information when reporting their health.[62,63] They also likely vary by age, health outcome, and time.[30,32,58]

Contrary to our hypothesis, accounting for mutual influence of academic achievement and health over time did not produce meaningfully different results: differences in estimates

between the unweighted and weighted models were minimal, particularly for the health status and psychological distress outcomes. One explanation is that there may simply not have been substantial time-dependent confounding in the associations between academic achievement and 2007 health. There are four preconditions that can be used to determine the potential usefulness of MSMs for a given research question: (1) time-varying covariates of interest predict the exposure, (2) the exposure predicts time-varying covariates of interest, (3) time-varying covariates of interest predict the outcome independent of exposure, and (4) exposure and covariates of interest vary over time.[49]

Appendix Table 3.A3 and Figure 3.A1 demonstrate how these preconditions apply to our question for the health status outcome. Although our measures met the preconditions when tested in unadjusted models, after adjustment for relevant covariates, many of the associations were weak, although in expected directions. With respect to precondition (3), after adjustment the outcome of 2007 poorer self-rated health status was only weakly predicted by the previous health measures. It is possible that in the relative absence of serious chronic illnesses, reported health status measures in this young, healthy population reflect primarily short-term states rather than the influence of chronic health conditions as defined here. Another consideration is that 2007 health status was reported by the participant himself or herself while the earlier measures were reported by the participant's caregiver; the two sources may have based their assessments of the participant's health on different facets of health. That said, in preliminary analyses of our data, the health-status measures reported by the care-giver and participant were highly correlated and shared many predictors in common. Furthermore, in a sensitivity analysis we restricted the sample to participants aged 5 years or older in 1997; these participants were old enough to self-

report their own health status in 2002. Replacing the 2002 caregiver-reported measure with the self-reported measure in the marginal structural model did not affect our results.

The reliance on self- and proxy-reported information in general was a limitation of this analysis, as it may have results in large and/or correlated errors in measurement of different variables. However, exposure information came from validated and interviewer-administered academic achievement tests. Despite the breadth of information on confounders available, we also cannot rule out the possibility that our results are biased because of unmeasured confounders. Our results may also have been affected by missing data and attrition if these characteristics were related to unobserved variables.

The associations we observed—although small in magnitude and weak for some outcomes among girls, and minimal for all outcomes among boys—are consistent with the hypothesis that higher academic achievement provides future health benefits. More generally, schooling may benefit health through cognitive and psychosocial mechanisms related to the learning process itself that are distinct from income-mediated mechanisms associated with educational attainment and that accrue during childhood and adolescence. Academic achievement was also predictive of completed schooling in this sample: after adjustment for confounders, 1-standard-deviation higher average academic achievement was associated with completion of 0.25 (0.17–0.32) more years of schooling in 2007. This is consistent with the interpretation that relations between completed schooling and health may to some extent reflect earlier achievement gains.

This analysis was, to our knowledge, the first to address the question of whether academic achievement has longer-term effects on adolescent and young adult health after accounting for the possible time-dependent confounding effects of prior health. We found that

greater academic achievement is associated with better health, especially in girls. However, a number of questions remain unanswered, including the reasons for the gender differences, the specific mechanisms for these effects and the health outcomes for which they are most important, and the ways in which these academic-achievement-related health disparities may progress over the life course and interact with other social determinants of health.

References

1. Schiller JS, Lucas JW, Ward BW, Peregoy JA. Summary health statistics for U.S. adults: National Health Interview Survey, 2010. *Vital Health Stat 10* 2012(252):1-207.

2. Kitagawa EM, Hauser PM. Education differentials in mortality by cause of death: United States, 1960. *Demography* 1968;**5**(1):318-353.

3. Cutler DM, Lleras-Muney A. Education and Health: Evaluating Theories and Evidence. *NBER Working Papers*. Cambridge, MA: National Bureau of Economic Research, 2006.

4. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. *Health Affairs* 2002;**21**:60-76.

5. Ross CE, Wu C-L. The links between education and health. *Am Sociol Rev* 1995;**60**(5):719-745.

6. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. *Journal of Health Economics* 2010;**29**:1-28.

7. Julian T, Kominski R, Educational Resources Information Center (U.S.), US Census Bureau. Education and Synthetic Work-Life Earnings Estimates. American Community Survey Reports. ACS-14. [S.l.]: Distributed by ERIC Clearinghouse, 2011;16 p.

8. Lynch J, Kaplan GA. Socioeconomic position. In: Berkman LF, Kawachi I, eds. *Social Epidemiology*. New York, NY: Oxford University Press, 2000;13-25.

9. Kawachi I, Adler NE, Dow WH. Money, schooling, and health: mechanisms and causal evidence. *Ann N Y Acad Sci* 2010;**1186**:56-68.

10. Grossman M. On the concept of health capital and the demand for health. *The Journal of Political Economy* 1972;**80**(2):223-255.

11. Fuchs VR. Time Preference and Health: An Exploratory Study. *NBER Working Papers*. Cambridge, MA: National Bureau of Economic Research, 1980.

12. Feinstein L, Sabates R, Anderson TM, Sorhaindo A, Hammond C. What are the effects of education on health? In: Desjardins R, Schuller T, eds. *Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium*. Copenhagen: Organisation for Economic Co-operation and Development, 2006;171-353.

13. Undheim AM, Sund AM. School factors and the emergence of depressive symptoms among young Norwegian adolescents. *Eur Child Adolesc Psychiatry* 2005;**14**(8):446-53.

14. Frojd SA, Nissinen ES, Pelkonen MU, Marttunen MJ, Koivisto AM, Kaltiala-Heino R. Depression and school performance in middle adolescent boys and girls. *J Adolesc* 2008;**31**(4):485-98.

15. Piko BF. Self-perceived health among adolescents: the role of gender and psychosocial factors. *Eur J Pediatr* 2007;**166**(7):701-8.

16. Eide ER, Showalter MH, Goldhaber DD. The relation between children's health and academic achievement. *Children and Youth Services Review* 2010;**32**:231-238.

17. Brolin Laftman S, Modin B. School-performance indicators and subjective health complaints: are there gender differences? *Sociol Health Illn* 2012;**34**(4):608-25.

18. Crosnoe R, Muller C. Body mass index, academic achievement, and school context: examining the educational experiences of adolescents at risk of obesity. *J Health Soc Behav* 2004;**45**(4):393-407.

19. Crosnoe R, Riegle-Crumb C. A life course model of education and alcohol use. *J Health Soc Behav* 2007;**48**(3):267-82.

20. Tucker JS, Martinez JF, Ellickson PL, Edelen MO. Temporal associations of cigarette smoking with social influences, academic performance, and delinquency: a four-wave longitudinal study from ages 13-23. *Psychol Addict Behav* 2008;**22**(1):1-11.

21. Needham BL, Crosnoe R, Muller C. Academic failure in secondary school: the inter-related role of health problems and educational context. *Soc Probl* 2004;**51**(4):569-586.

22. McLeod JD, Uemura R, Rohrman S. Adolescent mental health, behavior problems, and academic achievement. *J Health Soc Behav* 2012;**53**(4):482-97.

23. Haas SA, Fosse NE. Health and the educational attainment of adolescents: evidence from the NLSY97. *J Health Soc Behav* 2008;**49**(2):178-92.

24. Jackson MI. Understanding links between adolescent health and educational attainment. *Demography* 2009;**46**(4):671-94.

25. Sabia JJ, Rees DI. The effect of sexual abstinence on females' educational attainment. *Demography* 2009;**46**(4):695-715.

26. Fowler MG, Johnson MP, Atkinson SS. School achievement and absence in children with chronic health conditions. *J Pediatr* 1985;**106**(4):683-7.

27. Taras H, Potts-Datema W. Chronic health conditions and student performance at school. *J Sch Health* 2005;**75**(7):255-66.

28. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;**40 Suppl 2**:139S-161S.

29. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;**11**(5):550-60.

30. Liu H, Hummer RA. Are educational differences in U.S. self-rated health increasing?: an examination by gender and race. *Soc Sci Med* 2008;**67**(11):1898-906.

31. Zajacova A, Hummer RA. Gender differences in education effects on all-cause mortality for white and black adults in the United States. *Soc Sci Med* 2009;**69**:529-537.

32. Ross CE, Masters RK, Hummer RA. Education and the gender gaps in health and mortality. *Demography* 2012;**49**(4):1157-83.

33. Institute for Social Research. The Panel Study of Income Dynamics. Ann Arbor, MI: University of Michigan.

34. Institute for Social Research. The PSID-CDS Design At-A-Glance. Ann Arbor, MI: Institute for Social Research, 2006.

35. Institute for Social Research. The Panel Study of Income Dynamics Child Development Supplement: User Guide Supplement for CDS-I. Ann Arbor, MI: The Institute for Social Research, Survey Research Center, University of Michigan, 2010.

36. Institute for Social Research. Transition to Adulthood Study: Overview. Ann Arbor, MI: Institute for Social Research, 2008.

37. Fosse NE, Haas SA. Validity and stability of self-reported health among adolescents in a longitudinal, nationally representative survey. *Pediatrics* 2009;**123**(3):e496-501.

38. Vingilis ER, Wade TJ, Seeley JS. Predictors of adolescent self-rated health. Analysis of the National Population Health Survey. *Can J Public Health* 2002;**93**(3):193-7.

39. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, Walters EE, Zaslavsky AM. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002;**32**(6):959-76.

40. Institute for Social Research. The Panel Study of Income Dynamics Child Development Supplement: User Guide for CDS-II. Ann Arbor, MI: Institute for Social Research, Survey Research Center, University of Michigan, 2010.

41. Dean RS, Woodcock RW. The WJ-R and Bateria-R in Neuropsychological Assessment: Report Number 3. Rolling Meadows, IL: Riverside Publishing, 1999.

42. Woodcock RW. Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment* 1990;**8**:231-258.

43. Raghunathan TE, Lepkowski J, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 2001;**27**(1):85-95.

44. Institute for Social Research. IVEware: Imputation and Variance Estimation Software. http://www.isr.umich.edu/src/smp/ive/ Accessed December 8, 2011.

45. Raghunathan TE. What do we do with missing data? Some options for analysis of incomplete data. *Annu Rev Public Health* 2004;**25**:99-117.

46. Moodie EEM, Delaney JAC, Lefebvre G, Platt RW. Missing confounding data in marginal structural models: a comparison of inverse probability weighting and multiple imputation. *Int J Biostat* 2008;**4**(1):Article 13.

47. Spratt M, Carpenter J, Sterne JA, Carlin JB, Heron J, Henderson J, Tilling K. Strategies for multiple imputation in longitudinal studies. *Am J Epidemiol* 2010;**172**(4):478-87.

48. Sas Institute Inc. *SAS/STAT 9.2 User's Guide*. Vol. SAS Institute Inc. Cary, NC, 2008.

49. Cerda M, Diez-Roux AV, Tchetgen ET, Gordon-Larsen P, Kiefe C. The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology*;**21**(4):482-9.

50. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;**168**(6):656-64.

51. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;**11**(5):561-70.

52. Yang W, Joffe MM. Subtle issues in model specification and estimation of marginal structural models. *Pharmacoepidemiol Drug Saf* 2012;**21**(3):241-5.

53. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**(7):702-6.

54. Zou G, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res* 2011;**0**(0).

55. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand SL, Manderscheid RW, Walters EE, Zaslavsky AM. Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 2003;**60**(2):184-9.

56. Prochaska JJ, Sung HY, Max W, Shi Y, Ong M. Validity study of the K6 scale as a measure of moderate mental distress based on mental health treatment need and utilization. *Int J Methods Psychiatr Res* 2012;**21**(2):88-97.

57. Ross CE, Mirowsky J. Sex differences in the effect of education on depression: resource multiplication or resource substitution? *Soc Sci Med* 2006;**63**(5):1400-13.

58. Yu Y. Educational differences in obesity in the United States: a closer look at the trends. *Obesity (Silver Spring)* 2012;**20**(4):904-8.

59. Molarius A, Seidell JC, Sans S, Tuomilehto J, Kuulasmaa K. Educational level, relative body weight, and changes in their association over 10 years: an international perspective from the WHO MONICA Project. *Am J Public Health* 2000;**90**(8):1260-8.

60. Cohen AK, Rehkopf DH, Deardorff J, Abrams B. Education and obesity at age 40 among American adults. *Soc Sci Med* 2012.

61. Florin TA, Shults J, Stettler N. Perception of overweight is associated with poor academic performance in US adolescents. *J Sch Health* 2011;**81**(11):663-70.

62. Ross CE, Mirowsky J. Gender and the health benefits of education. *The Sociological Quarterly* 2010;**51**:1-19.

63. Benyamini Y, Leventhal EA, Leventhal H. Gender differences in processing information for making self-assessments of health. *Psychosom Med* 2000;**62**(3):354-64.

64. Hofferth S, Davis-Kean PE, Davis J, Finkelstein J. The Child Development Supplement to the Panel Study of Income Dynamics: 1997 User Guide. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, 1997.

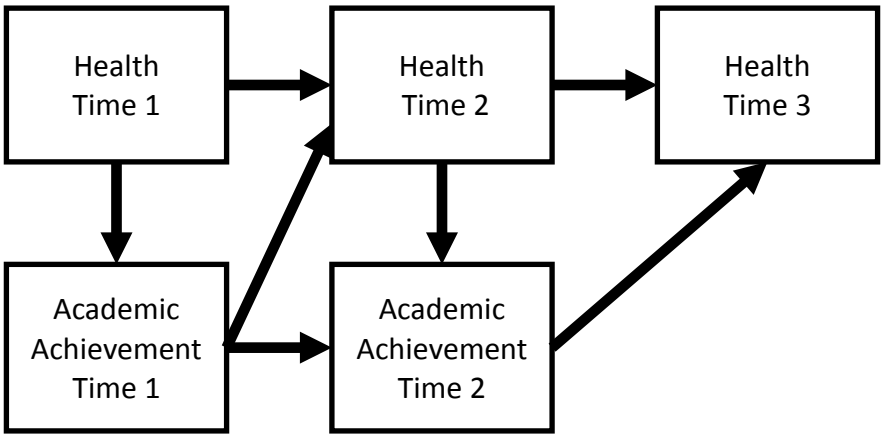
65. Baker PC, Keck CK, Mott FL, Quinlan SV. NLSY Child Handbook: A Guide to the 1986-1990 National Longitudinal Survey of Youth Child Data. Revised Edition. Columbus, OH: Center for Human Resource Research, The Ohio State University, 1993.

66. Grieger LD, Schoeni RF, Danziger S. Accurately measuring the trend in poverty in the United States using the Panel Study of Income Dynamics. *PSID Technical Series*. Ann Arbor, MI: Institute for Social Research, 2008.

67. Hajat A, Kaufman JS, Rose KM, Siddiqi A, Thomas JC. Long-term effects of wealth on mortality and self-rated health status. *Am J Epidemiol* 2010;**173**(2):192-200.

68. Hofferth S, Stafford FP, Yeung W-JJ, Duncan GG, Hill MS, Lepkowski J, Morgan JN. A Panel Study of Income Dynamics: Procedures and Codebooks. Guide to the 1993 Interviewing Year. Ann Arbor, MI: Survey Research Center, Institute for Social Research, 1998.

Figure 3.1 Assumed causal structure of health and academic achievement



Academic achievement is assumed to be influenced by health at each time-point, and in turn to influence health at future time-points.

Figure 3.2 Study sample from Child Development Supplement (CDS) and Transition into Adulthood (TA) supplemental studies to the Panel Study of Income Dynamics, 1997–2007

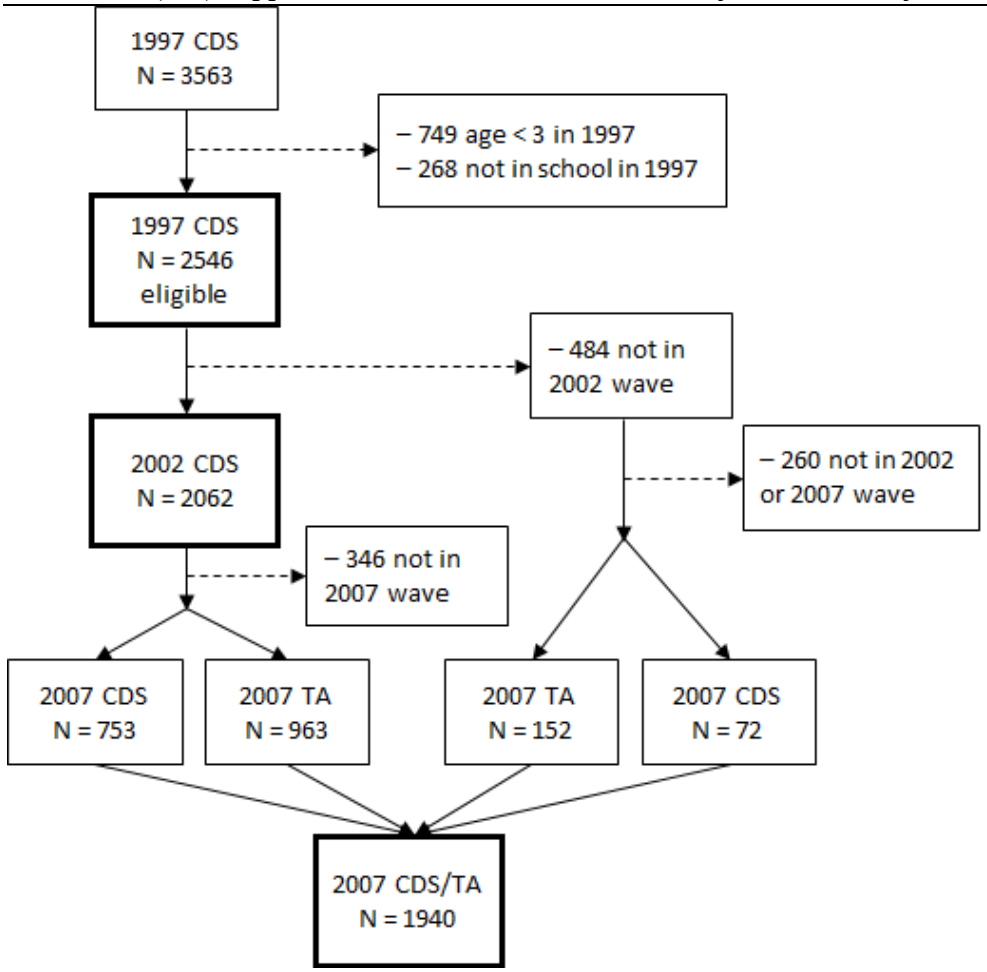


Table 3.1 Mean academic achievement score and 2007 health status by category of selected predictors (N = 2546)[a]

	Sample percent	Mean academic achievement 1997–2002		Poorer self-rated health status 2007[b]	
		Mean	p[c]	Percent	p[d]
Total	--	101.9	--	43	--
Sex			0.28		0.05
Male	51	102.8		40	
Female	49	103.5		46	
Age at 1997 interview (years)			0.20		0.21
3–5	18	103.1		47	
6–8	31	103.0		43	
9–11	31	103.9		39	
12–14	20	102.0		44	
Race/Hispanicity			<0.001		0.33
non-Hispanic white	46	108.4		42	
non-Hispanic black	42	97.7		42	
Hispanic	8	100.5		47	
Other	5	104.4		50	
Marital status of head of household 1997			<0.001		0.44
Married	63	105.6		42	
Never married	16	96.6		44	
Divorced/Separated/Widowed	21	100.7		45	
Number of parent figures live with 1997			<0.001		0.40
0	1	96.2		50	
1	30	98.7		45	
2	69	105.1		42	
Number of parent figures live with 2002			<0.001		0.27
0	7	103.3		49	
1	31	99.5		45	
2	62	104.9		41	
HOME scale quartile 1997			<0.001		0.11
1 (8.8–17.29)	25	97.2		46	
2 (17.3–18.99)	27	102.4		43	
3 (19.0–20.89)	22	102.8		44	
4 (20.9–24.0)	26	109.9		39	
Primary caregiver achievement score quartile 1997			<0.001		0.59
1 (4.0–26.99)	23	96.3		43	
2 (27.0–30.99)	25	100.7		44	
3 (31.0–34.27)	27	105.2		43	
4 (34.28–43.0)	25	109.6		40	
Family income/poverty ratio 1996			<0.001		0.003
0–0.9	21	96.8		48	
1–1.9	22	99.9		46	
2–4.9	44	104.9		42	
≥5	13	112.2		33	

Table 3.1 Mean academic achievement score and 2007 health status by category of selected predictors (N = 2546)[a]

	Sample percent	Mean academic achievement 1997–2002		Poorer self-rated health status 2007[b]	
		Mean	p[c]	Percent	p[d]
Family income/poverty ratio 2002			<0.001		0.007
0–0.9	20	98.9		51	
1–1.9	19	98.5		45	
2–4.9	41	103.7		41	
≥5	19	111.0		37	
Maximum of caregivers' completed years schooling 1997			<0.001		0.03
<9	3	97.1		51	
9–11	9	96.0		51	
12	35	99.7		44	
13–15	28	103.1		40	
≥16	24	111.7		39	
Neighborhood dangerous to walk in alone after dark 1997			<0.001		0.81
Yes	15	97.5		42	
No	85	104.1		43	
Neighborhood dangerous to walk in alone after dark 2002			<0.001		0.84
Yes	15	98.1		43	
No	85	104.0		43	
Neighborhood as place to raise children 1997			<0.001		0.73
Excellent	27	105.2		40	
Very good	30	102.3		43	
Good	27	97.7		44	
Fair/Poor	17	97.3		44	
Neighborhood as place to raise children 2002			<0.001		0.19
Excellent	31	106.2		40	
Very good	29	104.8		41	
Good	22	100.2		47	
Fair/Poor	17	98.7		45	
Urbanicity 1997			0.16		0.76
Central county, metropolitan area ≥1M	30	102.6		44	
County, metropolitan area ≥250K	41	103.9		43	
Area <250K	29	102.5		42	
Region 1997			<0.001		0.45
Northeast	14	107.8		40	
North Central	25	103.3		45	
South	45	101.4		42	
West	16	103.5		45	
Primary caregiver-rated health compared to other babies at birth			<0.001		0.34
Better	26	105.2		41	
Same	65	102.8		43	
Worse	9	99.6		48	

Table 3.1 Mean academic achievement score and 2007 health status by category of selected predictors (N = 2546)[a]

	Sample percent	Mean academic achievement 1997–2002		Poorer self-rated health status 2007[b]	
		Mean	p[c]	Percent	p[d]
Low birth weight (<2500 g)			<0.001		0.68
Yes	9	98.6		41	
No	91	103.6		43	
Spent time in NICU after birth			<0.001		0.08
Yes	12	99.5		48	
No	88	103.6		42	
Number overnight hospital stays birth–1997			<0.001		0.66
0	75	103.9		42	
1–2	20	101.2		44	
≥3	4	99.4		46	
Number overnight hospital stays 1997–2002			0.001		0.14
0	89	103.4		42	
1–2	9	101.0		48	
≥3	2	99.5		53	
Routine physician check-up 1996–1997			0.001		0.10
Yes	78	102.5		44	
No	22	105.2		39	
Routine physician check-up in 2001–2002			0.30		0.54
Yes	83	103.0		43	
No	17	104.0		41	
Primary caregiver-rated health 1997			<0.001		0.03
Very good	81	104.1		42	
Poorer	19	99.1		48	
Primary caregiver-rated health 2002			<0.001		0.03
Very good	78	104.1		41	
Poorer	22	99.7		49	
Self-rated health 2002			0.00		<0.001
Very good	62	104.2		36	
Poorer	38	101.3		54	
Self-rated health 2007			0.001		--
Very good	57	104.3		--	
Poorer	43	101.6		--	
BMI percentile 1997			<0.001		0.008
< 5 (underweight)	13	99.8		41	
5–84.9 (normal)	52	104.6		40	
85–94.5 (overweight)	14	103.3		41	
≥ 95 (obese)	21	101.2		50	
BMI percentile 2002			0.001		<0.001
< 5 (underweight)	5	101.7		38	
5–84.9 (normal)	55	104.1		37	
85–94.5 (overweight)	17	103.6		46	
≥ 95 (obese)	23	100.7		56	

Table 3.1 Mean academic achievement score and 2007 health status by category of selected predictors (N = 2546)[a]

	Sample percent	Mean academic achievement 1997–2002		Poorer self-rated health status 2007[b]	
		Mean	p[c]	Percent	p[d]
BMI status 2007 (age ≤ 21/21+)			<0.001		<0.001
< 5th percentile/<18.5 (underweight)	5	102.4		45	
5–84.9th percentile/18.5–24.9 (normal)	53	104.6		36	
85–94.5/25–29.9 (overweight)	22	101.8		46	
≥ 95th percentile/≥ 30 (/obese)	20	100.9		57	
Children's Depression Inventory quartile 2002 (age 12+)			0.03		0.04
1	19	105.5		36	
2	31	103.7		39	
3	24	103.0		44	
4	25	101.8		47	
K-6 Non-Specific Psychological Distress Scale quartile 2007 (age 18+)			0.03		<0.001
1	25	103.0		34	
2	27	104.9		40	
3	23	102.6		45	
4	25	101.7		52	
Behavior Problems Index quartile 1997			<0.001		0.17
1	25	105.2		40	
2	27	104.2		42	
3	22	103.2		43	
4	26	99.8		46	
Behavior Problems Index quartile 2002			<0.001		0.02
1	25	105.9		39	
2	26	104.8		39	
3	23	102.7		45	
4	25	99.1		48	
Ever repeated a grade 1997			<0.001		0.43
Yes	7	91.0		46	
No	93	104.0		43	
Changed school during current school year 1997			0.002		0.92
Yes	5	98.8		43	
No	95	103.4		43	
Changed school during current school year 2002 (among those in school)			<0.001		0.74
Yes	6	97.0		41	
No	94	103.8		43	
Type of school 1997			0.002		0.68
Public	81	102.5		43	
Private	8	110.6		38	
Home	11	102.7		47	

Table 3.1 Mean academic achievement score and 2007 health status by category of selected predictors (N = 2546)[a]

	Sample percent	Mean academic achievement 1997–2002		Poorer self-rated health status 2007[b]	
		Mean	p[c]	Percent	p[d]
Type of school 2002			<0.001		0.13
Public	78	102.9		42	
Private	5	110.5		35	
Home	9	103.4		45	
Not in school	8	100.1		54	

[a] Uses imputed data.
[b] Good/fair/poor (vs. excellent/very good).
[c] From unadjusted linear regression with clustering by sibling pair.
[d] From unadjusted logistic regression with clustering by sibling pair.

Table 3.2 Prevalence ratios of poorer self-rated health 2007 for 1-standard-deviation[a] higher average academic achievement 1997–2002, by sex

		Combined			Boys			Girls			
Model Type	Covariates	PR	95% CI	p	PR	95% CI	p	PR	95% CI	p	Interaction p
Unweighted	None	0.90	(0.84, 0.97)	0.004	0.94	(0.85, 1.03)	0.20	0.86	(0.78, 0.95)	0.002	0.17
Unweighted	Baseline[b]	0.91	(0.84, 0.99)	0.03	0.95	(0.85, 1.06)	0.35	0.87	(0.79, 0.97)	0.01	0.20
Unweighted	Baseline and time-varying[c]	0.93	(0.85, 1.01)	0.08	0.96	(0.86, 1.07)	0.47	0.89	(0.81, 0.99)	0.04	0.27
MSM	Baseline[b]	0.91	(0.83, 1.00)	0.05	0.96	(0.86, 1.08)	0.49	0.87	(0.78, 0.97)	0.01	0.12

[a] Standard deviation was calculated separately for each imputation. Values ranged 14.7–15.1, with a mean of 14.9.

[b] Model includes baseline age, sex, race, perinatal health, caregiver education, number of parents in household, HOME scale, household income, region, urbanicity, whether the child had repeated a grade, and the primary caregiver's achievement score as covariates.

[c] Model includes baseline adjustment variables + time-varying measures of family income, neighborhood rating, health status, overnight hospital stays, BMI percentile, problem behaviors, school type, and whether the child switched schools in the current year.

Table 3.3 Percent difference in 2007 BMI associated with 1-standard-deviation higher average academic achievement 1997–2002[a], by sex

(a)

		Combined			Boys			Girls			
Model Type	Covariates	Diff	95% CI	p	Diff	95% CI	p	Diff	95% CI	p	Interaction p
Unweighted	None	-1.83	(-2.89, -0.77)	<0.001	-1.08	(-2.51, 0.37)	0.14	-2.70	(-4.13-1.24)	<0.001	0.11
Unweighted	Baseline[b]	-0.88	(-2.07, 0.31)	0.15	-0.26	(-1.76, 1.25)	0.73	-1.58	(-3.07, -0.07)	0.04	0.16
Unweighted	Baseline and time-varying[c]	-0.52	(-1.51, 0.50)	0.30	-0.19	(-1.35, 0.97)	0.74	-0.90	(-2.13, 0.35)	0.16	0.31
MSM	Baseline[b]	-0.71	(-1.98, 0.58)	0.28	-0.14	(-1.60, 1.35)	0.86	-1.36	(-3.12, 0.44)	0.14	0.24

[a] Calculated from coefficients and 95% confidence intervals from linear regression of the natural log of BMI using the formula: % difference = 100*(exp(b) - 1)). Standard deviation of average academic achievement was calculated separately for each imputation. Values ranged 14.7–15.1, with a mean of 14.9.

[b] Model includes baseline age, sex, race, perinatal health, caregiver education, number of parents in household, HOME scale, household income, region, urbanicity, whether the child had repeated a grade, and the primary caregiver's achievement score as covariates.

[c] Model includes baseline adjustment variables + time-varying measures of family income, neighborhood rating, health status, overnight hospital stays, BMI percentile, problem behaviors, school type, and whether the child switched schools in the current year.

Table 3.4 Prevalence ratios of serious psychological distress for 1-standard-deviation higher average academic achievement 1997–2002[a], by sex

		Combined			Male			Female			
Model Type	Covariates	PR	95% CI	p	PR	95% CI	p	PR	95% CI	p	Interaction p
Unweighted	None	0.84	(0.73, 0.97)	0.02	0.92	(0.76, 1.12)	0.39	0.77	(0.61, 0.95)	0.02	0.24
Unweighted	Baseline[b]	0.89	(0.76, 1.05)	0.16	0.95	(0.77, 1.17)	0.65	0.82	(0.65, 1.04)	0.10	0.34
Unweighted	Baseline and time-varying[c]	0.91	(0.77, 1.07)	0.26	0.97	(0.78, 1.20)	0.76	0.85	(0.67, 1.07)	0.17	0.40
MSM	Baseline[b]	0.91	(0.76, 1.07)	0.25	0.98	(0.79, 1.22)	0.85	0.82	(0.64, 1.05)	0.12	0.28

[a] Includes only participants aged at least 18 years in 2007: N = 36,293 in imputed sample, including information for 1475 participants. Standard deviation was calculated separately for each imputation based on the entire sample. Values ranged 14.7–15.1, with a mean of 14.9.
[b] Model includes baseline age, sex, race, perinatal health, caregiver education, number of parents in household, HOME scale, household income, region, urbanicity, whether the child had repeated a grade, and the primary caregiver's achievement score as covariates.
[c] Model includes baseline adjustment variables + time-varying measures of family income, neighborhood rating, health status, overnight hospital stays, BMI percentile, problem behaviors, childhood depression scale (2002), school type, and whether the child switched schools in the current year.

Appendix

Multiple imputation

Table 3.A1 Variables used in multiple imputation and estimation of treatment weights

Variables	Multiple imputation Years used	Treatment weights Years used
Demographic		
Sex	1997	1997
Age	1997, 2002, 2007	1997
Race	1997	1997
Family composition		
Family size	1997, 2002	
Head of household's marital status	1997	
Number of parent figures/guardians lives with	1997, 2002, 2007	1997, 2002
Lives with a partner (age ≥ 18)	2007	
Birth order to mother	1997	
Sibling in 1997 CDS	1997	
Family socioeconomic status & environment		
Maximum of parents' years schooling at birth	1997	
Maximum of caregivers' years schooling	1997, 2002	1997
Family poverty threshold	1996, 2002, 2006	
ln(Family income-to-poverty ratio)	1996, 2002, 2006	1996, 2002
HOME scale	1997	1997
Neighborhood characteristics		
Years lived in current neighborhood	1997, 2002	
Rating of neighborhood as a place to raise children	1997, 2002	1997, 2002
Neighborhood dangerous to walk around alone after dark	1997, 2002	
Geographical characteristics		
Urbanicity	1997, 2001, 2003, 2007	1997
Region	1997, 2001, 2003, 2007	1997
Perinatal health		
Birth weight	1997	
Low birth weight		1997
Spent time in neonatal intensive care unit	1997	
Primary caregiver's rating of health at birth	1997	1997
Current physical health and health care use		
Primary caregiver's rating of health status	1997, 2002	1997, 2002
Self-rated health status	2002, 2007	
Body mass index	1997, 2002, 2005 (TA), 2007	1997, 2002
Number of overnight hospital stays	1997, 2002, 2007	1997, 2002
Routine physician check-up in past 2 years	1997, 2002	1997, 2002
Primary caregiver's self-rated health status	1997, 2002	
Mental health and well-being		
Children's Depression Inventory (age 12+)	2002, 2007	2002
K-6 Nonspecific Psychological Distress Scale (TA only; age 18+)	2005, 2007	
Physician diagnosis of serious emotional disturbance	1997, 2002	
Physician diagnosis of depression (TA only; age 18+)	2007	
Global self-concept (CDS version; age 8+ in 1997,	1997, 2002, 2007	

Table 3.A1 Variables used in multiple imputation and estimation of treatment weights

Variables	Multiple imputation Years used	Treatment weights Years used
age 10+ in 2002, 2007)		
Global self-concept compared to others (TA version; age 18+)	2005, 2007	
Emotional well-being subscale (age 12+)	2002, 2007	
Psychological well-being subscale (age 12+)	2002, 2007	
Social well-being subscale (age 12+)	2002, 2007	
Behaviors		
Behavior Problems Index (age ≤ 19)	1997, 2002, 2007	1997, 2002
Internalizing Behaviors subscale	1997, 2002, 2007	
Externalizing Behaviors subscale	1997, 2002, 2007	
Categories of # arrests (age 12+)	2002, 2005 (TA), 2007	
Drunk driving past 6 mos (age 12+)	2007	
Rode w/drunk driver past 6 mos (age 12+)	2007	
How often wear seatbelt (age 10+)	2002	
Did something dangerous past 6 mos just for thrill (TA only)	2005, 2007	
Regular smoker (age 11+)	2002, 2007	
Friend influences (CDS version; 9.5 ≤ age ≤ 19)	2002, 2007	
Friend influences (TA version; age 18+)	2005, 2007	
Friend influences (CDS & TA combined)	2007	
Schooling		
Years completed schooling	1997, 2002, 2007	
Degrees completed	2007	
Ever repeated a grade	1997	1997
Whether in school	2002, 2007	2002
Type of school	1997, 2002	1997, 2002
Changed school during current school year	1997, 2002	1997, 2002
Academic achievement		
Academic achievement score	1997, 2002	1997, 2002
Primary caregiver's achievement score	1997	1997

Table 3.A2 Distributions of characteristics in original and imputed samples

Characteristic	Original sample		Imputed sample
	N or range	%[a] or mean (SE[b])	% or mean (SE[b])
Total	2546	---	63,650
Female	1248	51	51
Age in years 1997	3.0–13.9	8.5 (0.059)	8.5 (0.059)
Missing	0	0	---
Age in years 2002	7.8–19.3	14.0 (0.068)	14.1 (0.060)
Missing	679	679	---
Age in years 2007	12.8–24.0	19.0 (0.068)	18.9 (0.059)
Missing	671	671	---
Race/Hispanicity			
non-Hispanic white	1160	46	46
non-Hispanic black	1063	42	42
Hispanic	193	8	8
Other race	127	5	5
Missing	3	< 1	---
Number of parent figures live with 1997			
0	24	1	1
1	756	30	30
2	1762	69	69
Missing	4	< 1	---
Number of parent figures live with 2002			
0	3	< 1	7
1	621	30	31
2	1438	70	62
Missing	484	19	---
Marital status of head of household 1997			
Married	1606	63	63
Never married	406	16	16
Divorced/Separated/Widowed	533	21	21
Missing	1	< 1	---
HOME scale 1997	8.8–24.0	18.9 (0.058)	18.7 (0.058)
Family income/poverty ratio 1996	0–36.9	2.9 (0.068)	2.9 (0.068)
Missing	0	0	---
Family income/poverty ratio 2002	0–111.6	3.7 (0.13)	3.4 (0.13)
Missing	186	7	---
Maximum of caregivers' completed years education 1997	3–17	13.2 (0.062)	13.2 (0.061)
<9	84	3	3
9–11	234	9	9
12	897	36	35
13–15	694	28	28
≥16	613	24	24
Missing	24	1	---
Caregiver achievement score	4–43	30.6 (0.16)	30.4 (0.14)
Missing	631	25	---
Neighborhood as place to raise children 1997			
Excellent	427	27	27
Very good	481	30	30
Good	408	26	27
Fair/Poor	263	17	17

Table 3.A2 Distributions of characteristics in original and imputed samples

Characteristic	Original sample		Imputed sample
	N or range	%[a] or mean (SE[b])	% or mean (SE[b])
Missing	967	38	---
Neighborhood as place to raise children 2002			
Excellent	621	30	31
Very good	633	31	29
Good	450	22	22
Fair/Poor	345	17	17
Missing	497	20	---
Urbanicity 1997			
Central county, metropolitan area ≥1M	754	30	30
County, metropolitan area ≥250K	1044	41	41
Area <250K	748	29	29
Region 1997			
Northeast	362	14	14
North Central	624	25	25
South	1152	45	45
West	407	16	16
Missing	1	< 1	---
Primary caregiver-rated health compared to other babies at birth			
Same	1633	65	65
Better	662	26	26
Worse	225	9	9
Missing	26	1	---
Low birth weight (<2500 g)	219	9	9
Missing	62	2	---
Spent time in NICU after birth	313	12	12
Missing	38	1	---
Number overnight hospital stays birth–1997			
0	1912	75	75
1–2	519	20	20
≥3	107	4	4
Missing	8	< 1	---
Number overnight hospital stays 1997–2002			
0	1859	90	89
1–2	171	8	9
≥3	31	2	2
Missing	485	19	---
Routine physician check-up 1996–1997	1952	78	78
Missing	40	2	---
Routine physician check-up in 2001–2002	1713	84	83
Missing	507	20	---
Primary caregiver-rated health 1997			
Excellent	1237	49	49
Very good	814	32	32
Good	410	16	16
Fair	65	3	3
Poor	5	< 1	<1
Missing	15	< 1	---
Primary caregiver-rated health 2002			

Table 3.A2 Distributions of characteristics in original and imputed samples

Characteristic	Original sample		Imputed sample
	N or range	%[a] or mean (SE[b])	% or mean (SE[b])
Excellent	1090	53	47
Very good	653	32	32
Good	257	12	13
Fair	56	3	3
Poor	6	< 1	5
Missing	484	19	---
Self-rated health 2007			
Excellent	464	25	25
Very good	725	39	32
Good	518	28	28
Fair	152	8	10
Poor	14	1	5
Missing	673	26	---
BMI percentile 1997	0–100	59 (0.78)	59 (0.78)
< 5 (underweight)	271	12	13
5–84.9 (normal)	1238	54	52
85–94.5 (overweight)	325	14	14
≥ 95 (obese)	459	20	21
Missing	253	10	---
BMI percentile 2002	0–100	66 (0.75)	66 (0.79)
< 5 (underweight)	49	3	5
5–84.9 (normal)	1102	60	55
85–94.5 (overweight)	288	16	17
≥ 95 (obese)	399	22	23
Missing	708	28	---
BMI status 2007 (age < 21/age 21+)			
< 5th percentile/<18.5 (underweight)	60	3	5
5–84.9th percentile/18.5–24.9 (normal)	1035	56	53
85–94.5th percentile/25–29.9 (overweight)	394	21	22
≥ 95th percentile/≥ 30 (obese)	347	19	20
Missing	710	28	---
Children's Depression Inventory 2002 (age 12+)	0–18	3.0 (0.092)	3.5 (1.27)
Missing	566	22	---
Too young to be asked question[c]	725	28	21
K-6 Non-Specific Psychological Distress Scale 2007 (age 18+)	0–23	5.2 (0.11)	5.2 (0.17)
Missing	522	21	---
Too young to be asked question[c]	911	36	35
Behavior Problems Index 1997	0–27	8.0 (0.13)	8.0 (0.13)
Missing	102	4	---
Behavior Problems Index 2002	0–30	8.7 (0.16)	8.7 (0.15)
Missing	509	20	---
Ever repeated a grade 1997	159	6	7
Missing	67	3	---
Changed school during current school year 1997	129	5	5
Missing	49	2	---
Changed school during current school year	122	6	6

Table 3.A2 Distributions of characteristics in original and imputed samples

Characteristic	Original sample		Imputed sample
	N or range	%[a] or mean (SE[b])	% or mean (SE[b])
2002			
Not in school	29	1	8
Missing	513	20	---
Type of school 1997			
Public	1867	90	81
Private	188	9	8
Home	17	1	11
Missing	474	19	---
Type of school 2002			
Public	1784	91	78
Private	126	6	5
Home	28	1	9
Not in school	29	1	8
Missing	579	23	---
Academic achievement 1997	40.5–163.0	104.3 (0.42)	104.4 (0.40)
Missing	745	29	---
Academic achievement 2002	42.0–169.5	102.2 (0.42)	101.9 (0.39)
Missing	691	27	---
Mean academic achievement 1997–2002	42.0–169.5	103.1 (0.37)	103.1 (0.36)
Missing	316	12	---
Change in academic achievement 1997–2002	-48.0–72.5	-2.6 (0.35)	-2.5 (0.32)
Missing	1120	44	---

[a] To facilitate comparison with the imputed sample, the denominators for percents for nonmissing categories include only observations with nonmissing information. To facilitate interpretation, denominators also exclude participants ineligible to be asked the question because of age.

[b] Standard error of the mean.

[c] Estimated from age in 1997 for participants missing from this wave.

Construction of treatment weights

To estimate the marginal structural models, each participant was assigned a treatment weight sw_i that was constructed using the equation:

$$sw_i = \prod_{t=1}^{2} \frac{f(A(t) = a_i(t) | \boldsymbol{A}(t-1) = a_i(t-1),\ \boldsymbol{V} = \boldsymbol{v}_i)}{f(A(t) = a_i(t) | \boldsymbol{A}(t-1) = a_i(t-1),\ \boldsymbol{L}(t) = \boldsymbol{l}_i(t))}$$

where each term's numerator $f(A(t) = a_i(t) | \boldsymbol{A}(t-1) = a_i(t-1),\ \boldsymbol{V} = \boldsymbol{v}_i)$ is the probability density of participant i's academic achievement score at time-point t, conditional on his or her history of academic achievement through time-point $t-1$ (so $\boldsymbol{A}(1)$ was 1997 academic achievement and $\boldsymbol{A}(0)$ was set to zero) and values of time-invariant baseline covariates $\boldsymbol{V}$. In the denominator, $\boldsymbol{V}$ is replaced by $\boldsymbol{L}(t)$, the history of both baseline and time-varying covariates through time t ($\boldsymbol{V}$ is included in $\boldsymbol{L}(1)$).[29,51]

Included in $\boldsymbol{V}$ were sex, race, baseline age, perinatal health (whether the child was low birth weight [< 88 ounces] and the primary caregiver's assessment of the child's health at birth), an interviewer-assessed scale of cognitive stimulation and emotional support provided in the participant's home at baseline (using selected items from Caldwell and Bradley's Home Observation for the Measurement of The Environment [HOME] scale),[64,65] whether the participant had ever repeated a grade in school, and the primary caregiver's achievement score. In addition, the following variables with time-varying information but very high correlation or agreement between 1997 and 2002 were treated as time invariant and therefore also included in $\boldsymbol{V}$: the maximum of the child's caregivers' years of schooling, geographical region, and urbanicity. In addition to the variables included in $\boldsymbol{V}$, $\boldsymbol{L}(t)$ also included time-varying information on family income (measured as a family income-to-poverty-level ratio in which the annual household income was divided by the corresponding annual U.S. Census Bureau poverty

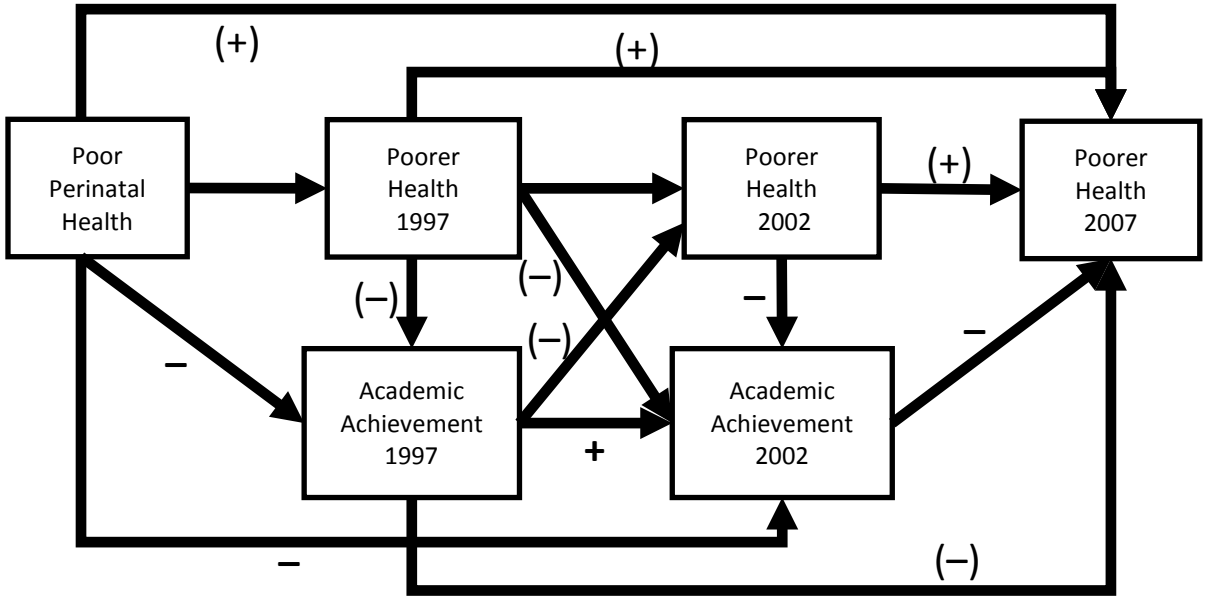
threshold, which takes into account family size and the ages of family members[66,67]), whether the child lived with two parent figures, a caregiver-rated measure of how good a place the neighborhood where the family resided was to raise children, health (global health status, body mass index [BMI], number of overnight hospital stays), whether the child had a routine physician check-up in the previous two years, a scale of the caregiver's report of problem behaviors (including both aggressive and withdrawn or sad behaviors)[68], school type (public, private, home, not in school), and whether the child changed schools during the current school year. In models of psychological distress, the log of the 2002 CDI score was additionally included as a time-dependent confounder. Since academic achievement was a normally distributed continuous measure, the conditional numerator and denominator probability densities were estimated using linear regression and the normal probability density function.[29,49]

Preconditions for usefulness of marginal structural models

Table 3.A3 Tests of preconditions for the outcome of 2007 poorer self-rated health status

Precondition	In This Case	Tests	Conclusions
Key time-varying covariates predict exposure	Does health predict academic achievement?	Adjusted linear regression of 1997 academic achievement on perinatal health (low birth weight, spent time in NICU, caregiver-rated health compared to other babies) & 1997 health (caregiver-rated health status, nights spent in hospital birth-1997).	In general, poorer health predicts lower academic achievement. For 1997 achievement, point estimates are driven by NICU and are stronger for perinatal health than 1997 health. For 2002 achievement, perinatal health is stronger predictor than 1997 or 2002 health.
		Adjusted linear regression of 2002 academic achievement on perinatal health, 1997 health, and 2002 health (caregiver-rated health status, nights spent in hospital 1997-2002).	
Exposure predicts key time-varying covariates	Does academic achievement predict future health?	Adjusted logistic regression of 2002 poorer caregiver-rated health status on 1997 academic achievement.	In general, higher academic achievement predicts lower odds of poorer health. Stronger evidence for 2007 self-rated health status than 2002 caregiver-rated health status.
		Adjusted logistic regression of 2007 poorer self-rated health status on 1997 and 2002 academic achievement.	
Key time-varying covariates predict outcome independent of exposure	Does past health predict 2007 health independent of academic achievement?	Adjusted logistic regression of 2007 poorer self-rated health status on perinatal, 1997, and 2002 health measures, with average 1997-2002 academic achievement included as a covariate.	Poorer previous health weakly predicts 2007 poorer self-rated health independent of academic achievement.
Exposure and key time-varying covariates vary over time	Do academic achievement and health vary over time?	Correlation between 1997 and 2002 academic achievement.	Academic achievement and health status do vary over time.
		Agreement between 1997, 2002, and 2007 poorer health status measures.	

Figure 3.A1 Empirical support for hypothesized causal paths linking academic achievement and health status



Results are from tests described in Table 3.A3 and describe the direction and strength of associations. Parentheses indicate weak associations.

CHAPTER 4

INFLUENCE OF GRANDPARENT SCHOOLING ON ADULT HEALTH STATUS, SMOKING, AND OBESITY

Introduction

Persistent social and health inequalities over time in the United States suggest that a person's adult health may be shaped not only by his or her own characteristics and experiences or even those of his or her parents, but also by those of previous generations. Schooling may be one such characteristic with multigenerational effects on health, given its strong associations with health within each generation and plausible mechanistic pathways linking schooling of prior generations to health of later generations.[1-12] Multigenerational effects of schooling on health may operate through mechanisms mediated by the intervening generation(s). For example, grandparents with more schooling may be able to facilitate higher levels of schooling in their children, which in turn will benefit the health of the children's children (i.e., grandchildren). Grandparents may also accrue income and wealth as a result of their own schooling that they pass on to their children, who in turn pass these health-benefiting economic advantages on to the next generation. The intergenerational effects of grandparents' schooling on their children's health may also in turn affect these children's ability to parent effectively. Similarly, grandparents with more schooling may model better health and parenting behaviors that their children emulate, to the benefit of the grandchildren.

However, there may also be processes through which grandparents' schooling influences their grandchildren's health directly (i.e., through mechanisms not mediated by the

grandchildren's parents). This is most readily evident when there is direct social contact between the grandparent and grandchild. Decreasing mortality rates make it increasingly common for grandparents and grandchildren to share lifespans. In 1996, Uhlenberg estimated that in 2000, 76% of 30-year-olds in the U.S. would have at least one living grandparent, compared to just 51% in 1960.[13] Furthermore, as of 2005, 23% of children under 5 years of age in the U.S. were regularly cared for by a grandparent [14] and in 2010, 5.4 million children under age 18 in the U.S. lived in the same household as a grandparent.[15] These grandparents may influence their grandchildren's health through many of the same mechanisms exerted by the grandchildren's parents. Other direct influences—such as paying for schooling or health care, direct monetary gifts, or the facilitation of schooling through legacy admissions policies—do not require direct contact between the grandparent and grandchild, or even that the grandparent be living.[16] Grandparent influences on health may be beneficial—for example, facilitating access to health care—or harmful—for example, smoking in the presence of the grandchild.

Despite substantial literatures addressing the intergenerational transmission of education and health separately, as well as studies examining the effects of parental schooling on their children's health in childhood, there is relatively little research addressing the intergenerational effects of schooling on adult health. Most existing studies examine the associations of parental schooling with child health, and use parental schooling as a proxy for childhood socioeconomic status, often in combination with other measures such as family income, parents' occupation, poverty level, or housing quality. Results of these studies have linked higher parental educational attainment with better self-rated health status, less disability and better physical function, better cognitive and psychosocial status, better birth outcomes, better health behaviors, fewer chronic conditions, less inflammation, and lower mortality.[1-8,10-12]

Even fewer studies have examined the impact of grandparents' schooling on the health of their grandchildren. Osler et al. found in a three-generation study that having more ancestors with some secondary education was related to lower mortality among a cohort of Danish men, and that this relation was robust to adjustment for the occupational class of the ancestors.[17] Krzyzanowka found that among university students in Poland, those with low grandparent and parent schooling were the shortest while those with high grandparent and parent schooling were the tallest.[18] Ahren-Moonga et al. found that higher schooling levels of parents and maternal grandmothers were related to a higher risk of hospitalization for an eating disorder among a cohort of Swedish women.[19] On the other hand, in a U.S. study Foster et al. found that the education levels of women in two cohorts did not predict the preterm birth or low birth weight of their grandchildren.[6]

The purpose of this analysis was to investigate associations of the educational attainment of grandparents with the health status, smoking status, and body mass index (BMI) of their grandchildren in adulthood. We estimated associations representing both total effects of grandparent schooling and direct effects through pathways unmediated by parent and grandchild schooling.[20] We also examined whether the effects of grandparent schooling differed by the geographical proximity of the grandchild to living grandparents. We would expect the amount of direct social interaction between grandparents and grandchildren to be higher if they live closer to each other. Therefore, stronger associations between grandparent schooling and grandchild health among grandchildren whose grandparents live geographically closer would support the existence of direct effects of grandparent schooling on grandchild adult health through pathways involving social interaction.

Racial Differences in the Effects of Schooling on Health

Two competing hypotheses have been postulated regarding racial differences in intragenerational schooling–health associations in the U.S.[21] The first, the *minority poverty hypothesis*, proposes a synergistic harmful effect between low socioeconomic status and exposure to racism and discrimination. Applied to schooling, this hypothesis predicts that the deleterious effects of relatively little schooling would be greater in minority populations than in non-Hispanic white populations. This would result in a steeper schooling-related health gradient in minority populations, and racial health disparities would be greatest at low schooling levels. The second hypothesis, the *diminishing returns hypothesis*, focuses on the smaller income and occupational returns to education historically experienced among minority populations in the U.S.[22,23] According to this hypothesis, the disparities in material gains from schooling dampen the beneficial effects of schooling on health in minority populations. Therefore, the diminishing returns hypothesis predicts that the beneficial effects of more schooling would be greater in minority populations than in non-Hispanic white populations. This would result in a weaker schooling-related health gradient in minority populations, and racial health disparities would be greatest at high schooling levels.

The minority poverty and diminishing returns hypotheses can be readily extended to multigenerational effects. The steeper schooling–health gradient implied by the minority poverty hypotheses could be translated into greater health returns among blacks than whites from higher educational attainment of previous generations. Similarly, the shallower gradient implied by the diminishing returns hypothesis could also be applied to the effects of multigenerational schooling effects, and in fact could be compounded if the smaller income returns to one generation result in less accumulated wealth being transferred to subsequent generations.

Existing research addressing racial differences in multigenerational schooling effects on health is limited, particularly for adult health outcomes. The body of research addressing racial/ethnic differences in intragenerational schooling–health does not give a clear answer about the nature of racial differences in the associations. Some studies have found a steeper educational gradient in health among minority populations, supporting the minority poverty hypothesis,[24,25] while others have found a shallower gradient, supporting the diminishing returns hypothesis.[26,27] Other studies have found no racial/ethnic differences, or have found differences only among some groups or for some outcomes.[21,28,29] It is likely that differences between races in the effects of schooling on health differ according to the specific sociocultural context and health outcome.

For the reasons outlined above, in addition to examining the overall associations of grandparent schooling with health, we examined whether these associations differed in non-Hispanic black and non-Hispanic white participants.

Methods

Study Population

Data came from the Panel Study of Income Dynamics (PSID), a longitudinal study started in 1968 of a representative sample of U.S. families conducted by the Survey Research Center at the Institute for Social Research at the University of Michigan.[30] Interviews were conducted annually until 1997, when the study switched to biennial data collection. In most waves, a single adult family member provides information about him-/herself and all other family members. Since 1973, the large majority of interviews have been conducted over the telephone. The number of families interviewed in each wave has ranged from about 5,000 to over 10,000; per-wave response rates are generally above 95%.[30]

In each PSID study wave, one adult member of each family serves as the primary respondent; in most cases this is the head of household, called the "Head."[30] The male member of a heterosexual married or unmarried couple is designated as the Head by default, although in some situations—such as if the male partner is incapacitated or the respondents insist—the female partner is designated as Head. The Head's wife or cohabiting female partner is designated as Wife or "Wife" respectively. In households with a single head of household, this person serves as the Head regardless of sex. When a family member leaves the household (e.g., after divorce or when a child grows up), his/her new household is added to the PSID sample as a "split-off" family. Because of this, members of many PSID families are related to members of other PSID families. The study currently contains information about up to three generations of any given family.[30]

The study sample for this analysis comprised PSID Heads and Wives (including cohabiting female partners) from the 2009 study wave who had at least one parent who was also a PSID Head or Wife. In keeping with previous research, the sample was restricted to participants aged 25 and older; additionally excluding participants who were in school in 2009 (7% of the sample) did not affect results.[21,25,31-33] Participants aged over 55 years in 2009 were also excluded to limit age variation in the sample and permit meaningful estimation of childhood socioeconomic measures starting in 1968, the first year of the PSID. Finally, the sample was limited to participants who reported being of non-Hispanic white or non-Hispanic black race. This is because the very small number of participants with multigenerational information who reported being of other races or of Hispanic ethnicity precluded deriving meaningful estimates for these groups. The study sample included 4,648 participants (Figure 4.1).

Measures

The measure of schooling for the study sample (third generation, G3) was drawn from the 2009 study wave, in which education information or all Heads and Wives was updated. We treated the measure as an ordinal variable referring to the number of years of completed schooling and ranging 1–17; alternative specifications, including those allowing for nonlinear associations with health, did not affect the results.

Categorical schooling measures for the parents of participants (second generation, G2) were retrospectively reported by the family primary respondent (usually the Head) on behalf of both the Head and, if applicable, the Wife. Rather than years of schooling, these measures described categories of schooling ranging from no education/could not read or write to completion of graduate work or a professional degree. This information was collected when a (G3) participant first became a Head or Wife. We combined information on the participants' mothers and fathers to create a single measure of the highest schooling category ever reported for either parent, and treated the measure as an ordinal variable ranging 0–8. As with the measure of G3 schooling, alternate specifications of this measure of G2 schooling, including allowing for nonlinear associations with G3 health, did not affect our results. The same retrospective measures were used for grandparent (first generation, G1) schooling, as reported by G2 Heads. As with the measure of G2 schooling, we combined information on all four grandparents to create a single measure of the highest schooling category ever reported for any grandparent. Because of our a priori interest in potential nonlinearity in the effects of grandparent schooling, our exposure of interest, on grandchild health, we treated this measure of G1 schooling as a set of indicators categorizing the maximum schooling achieved by any grandparent as less than a high school degree, a high school degree, some college education but no college degree, or a

college degree. For example, a value of “high school” can be interpreted as having at least one grandparent who completed high school but no grandparent who completed schooling beyond a high school degree.

We used measures of G2 and G1 schooling reported retrospectively by G3 and G2, respectively, instead of schooling information collected directly from G2 and G1 PSID sample members because (1) the retrospective measures contain more complete information, including information about both parents irrespective of whether they were PSID sample members and (2) the prospective measures were more likely to be outdated because after initial collections this information was only updated during several specific PSID study waves. The retrospective information was collected later, when the child of the PSID sample member first became a Head or Wife. In addition, using the retrospective measure of grandparent schooling reported by the parent more than doubled the size of the sample because it did not require that any grandparent be a PSID sample member. Finally, selecting G3 participants based on the availability of prospective grandparent information might also create a sample very distinct from the general population because of the multigenerational survival and fertility patterns required to make three adult generations available to participate in the PSID during its 41 years of follow-up. In preliminary exploratory analyses of G3 participants with both prospective and retrospective parent and grandparent schooling information, agreement between prospective and retrospective measures of parent high school and college degree attainment were 87% and 90%, respectively. Agreement between prospective and retrospective measures of grandparent high school and college degree attainment were 73% and 86%, respectively. Additional comparisons of the prospective and retrospective schooling measures can be found in the appendix.

The outcome variables were global health status, current smoking, and obesity. The health status question was asked of each Head on behalf of both him- or herself and, if applicable, his Wife. It was, "Would you say your/her health in general is excellent, very good, good, fair, or poor?" Health status was dichotomized into excellent/very good/good ("good") vs. fair/poor ("poor").[34-36] The PSID smoking variables pertain specifically to cigarettes. Participants were categorized as current smokers, with former and never smokers combined into a single referent group, as reported by the Head. We calculated each participant's BMI, based on height and weight reported by the Head, using the formula $BMI = \frac{weight\ in\ pounds}{(height\ in\ inches)^2} \times 703$. We then categorized participants with BMI $\geq$ 30 as obese.[37]

Covariates included sex, age, Head vs. Wife status, and whether the participant ever lived in the same state as any grandparent. Participants were categorized as ever living in the same state as a grandparent if (1) both the participant and a grandparent were present in the PSID and living in the same state during any study wave or (2) a parent of the participant reported having a parent living in the same state in a special module of the 1988 study wave. We were not able to include the years of birth of the grandparents because this information was missing for the large majority of observations.

In order to estimate the direct effect of grandparent schooling, we additionally added parent and participant schooling, as well as variables which might confound associations between the mediators (parent and participant schooling) and participant health outcomes. Potential confounder variables of the association between parent schooling and participant health included parental reports of poverty in childhood (as an indicator variable) and the years of birth of the parents. Potential confounder variables of the association between participant schooling and health included an indicator for poverty in childhood reported retrospectively, a prospective

measure of average family income when the participant was aged less than 18 years (measured as an income-to-poverty-level ratio in which the annual household income was divided by the corresponding annual U.S. Census Bureau poverty threshold, which takes into account family size and the ages of family members[38,39]), having a mother who was unmarried when the participant was born, having a mother aged less than 20 years when the participant was born, living with both natural parents most of the time while growing up, or participant fair or poor health status while growing up. We also included relevant measures of parent health: either parent ever reporting fair or poor health, a parent smoking while the participant was a child, and either parent ever being obese. These measures of parent health were collected prospectively from the parents and therefore only reflect PSID study waves in which this information was collected: 1984–2009 for health status, and 1986 and 1999–2009 for obesity and smoking. We supplemented the prospective parent smoking information with information from the question, "Did your/her parents smoke during your childhood?" asked of Heads on behalf of Heads and Wives in 2007 and 2009. Finally, we examined the robustness of our models to the addition of measures of participant adult income-to-poverty-level ratio and marital status.

Analysis

To handle missing information (both item and case missingness), multiple imputation with 25 imputations was carried out using the sequential regression method with IVEware software and simultaneously including all variables.[40,41] The multiple imputation assumed data were missing at random (MAR) rather than missing completely at random (MCAR) and allowed for standard errors accounting for variability in the imputation process.[42,43] Using the sequential regression method allowed us to specify an appropriate distribution for each variable, as well as to restrict imputation to relevant observations. All analyses were then conducted separately in

each imputed data set and the estimates were subsequently combined using the SAS MIANALYZE procedure, which averages point estimates across imputations and derives standard errors by combining information about variance within and between imputations.[44]

We assessed bivariate associations of covariates with grandparent schooling and the health outcomes using unadjusted logistic regression. We then estimated a series of models using the modified Poisson regression method developed by Zou and Donner (equivalent in this case to general estimating equation [GEE] Poisson regression with robust standard error estimation) to calculate prevalence ratios for poor health status, current smoking, and obesity using the household identifier from the 1968 baseline PSID study wave to account for correlated observations between family members.[45,46]

Figure 4.2 shows the assumed causal structure we used to guide the selection of variables to include in the models. To estimate the total effect of grandparent schooling, we estimated models adjusted for the variables in C1. To estimate the direct effect, including only the pathway not including parent and participant schooling, we first used the conventional regression method of adjusting for parent and participant schooling and the variables in C1, C2, and C3.[20,47,48] However, this method can produce biased estimates of the direct effect if there are consequences of exposure that confound the mediator–outcome associations (shown by the dashed lines in figure 2).[20,47] This situation is likely in our context. For example, grandparents' economic circumstances are a consequence of grandparent schooling that may confound associations between parent schooling and grandchild adult health. Therefore, we also estimated the direct effects of grandparent schooling using an alternative approach in which marginal structural models (MSMs) were employed to account for confounders using inverse probability weighting rather than adjustment.[20,49]

To estimate the direct effect using an MSM, we first estimated a stabilized weight sw_i for each observation i of the form

$$sw_i = \prod_{g=1}^{3} \frac{\Pr[S_g = s_{gi} | \boldsymbol{S_{g-1}} = \boldsymbol{s_{(g-1)i}}, race]}{\Pr[S_g = s_{gi} | \boldsymbol{S_{g-1}} = \boldsymbol{s_{(g-1)i}}, \boldsymbol{C_g} = \boldsymbol{c_{gi}}, race]}$$

where g is the generation (G1, G2, or G3), S_g is the schooling of that generation, $\boldsymbol{S_{g-1}}$ is the schooling of the previous generation(s) ($\boldsymbol{S_0}$was set to zero), and $\boldsymbol{C_g}$ is the appropriate set of confounders (C1, C2, or C3 from Figure 4.2). Race, defined as the race of the participant (G3), was included in the numerator of each term to allow us to test our hypothesis of differential effects of grandparent schooling by race. We estimated the numerator and denominator probabilities by dividing each generation's schooling into four categories (less than high school, high school, some college, college degree) and using multinomial logistic regression to estimate the conditional probability of the observed category. We then incorporated the weights into a GEE Poisson regression model of the form

$$\begin{aligned} \ln[E(Y_i)|S_1 = s_{1i}, S_2 = s_{2i}, S_3 = s_{3i}, race)] \\ = \beta_0 + \boldsymbol{\beta_1 S_1} + \beta_2 S_2 + \beta_3 S_3 + \beta_4 race + \boldsymbol{\beta_5 S_1} \times \boldsymbol{race} \\ + \beta_6 S_2 \times race + \beta_7 S_3 \times race \end{aligned}$$

where Y_i is the outcome for person i, $\boldsymbol{S_1}$ is a set of indicators for G1 schooling (high school, some college, and college degree; less than high school is the reference group) with a corresponding vector of coefficients $\boldsymbol{\beta_1}$, S_2 is the ordinal variable representing categories of G2 schooling, S_3 is the ordinal variable representing years of completed schooling by G3, and $race$ is an indicator for black race (vs. white). The vector of coefficients $\boldsymbol{\beta_5}$ represents differences in the coefficients of G1 schooling for blacks compared to whites. In a sensitivity analysis, separate race-specific models produced nearly identical but somewhat less precise estimates.

To assess whether the strength of the direct effect of grandparent schooling differed by geographical proximity, we estimated MSMs subcategorizing the indicators of grandparent schooling by whether the most-educated grandparent lived in the same state. For example, we created separate indicators for having a grandparent with a college degree who lived in the same state and having a grandparent with a college degree but who did not live in the same state. Because the additional parameters created by subcategorizing grandparent schooling made the models less stable, we combined the high school and some-college categories of grandparent schooling for this secondary analysis.

Results

Variable distributions in the original and imputed samples were very similar (Table 4.1). Forty-four percent of participants had at least one grandparent in the PSID sample. Black participants were more likely to have only one parent in the sample and to be unmarried or low income. Schooling levels were higher in each successive generation and were lower among blacks than whites in each generation. Black participants were also more likely to have poor health status or be obese, and less likely to be current smokers.

In bivariable comparisons, younger participants were more likely to have more highly educated grandparents (Table 4.2); this is not surprising given secular increases in schooling in the United States throughout the 20th century. Higher grandparent schooling was also associated with higher G2 and G3 schooling, better financial conditions across generations, and better health outcomes, but these associations were generally stronger among Whites than among Blacks. Younger, married, and more highly educated participants were less likely to report poor health status, as were participants with a grandparent who lived in the same state (Table 4.3). Obesity was more prevalent among females and was inversely related to G2 and G3 schooling

among Whites. Smoking was more prevalent among males and among participants with low schooling and income. Poor health status was positively associated with both obesity and smoking while obesity and smoking were inversely associated.

There was a graded estimated total effect (i.e., through pathways both mediated and unmediated by parent and participant schooling) of higher grandparent schooling on health status among Whites: prevalence ratios (PR) and 95% confidence intervals of poor health status comparing each schooling category to less than a high school degree were: PR = 0.82 (0.58–1.15) for high school degree, PR = 0.75 (0.51–1.11) for some college, and PR = 0.54 (0.36–0.82) for college degree (Table 4.4[a]). This association was not evident among Blacks (PR = 1.11 [0.84–1.46]; PR = 1.01[0.69–1.49]; and PR = 0.97 [0.60–1.56] for high school, some college, and a college degree, respectively). The direct effects (i.e., only through pathways not mediated by parent and participant schooling) of grandparent schooling among Whites were smaller than the estimated total effects. The direct effects estimated using marginal structural models were larger than those estimated using the conventional adjustment approach suggesting that the conventional approach may have biased the estimates of the direct effects toward the null. The direct effects estimated by the MSM were PR = 0.84 (0.57–1.22) for high school, PR = 0.81 (0.52–1.27) for some college, and PR = 0.71 (0.44–1.15) for a college degree.

Results for obesity followed the same pattern as those for poor health status (Table 4.4[b]). Among Whites, the total effect estimates showed a graded association between higher grandparent schooling and lower probability of obesity. Estimated direct effects were slightly smaller, and were more attenuated using conventional regression adjustment than using MSM. The direct effects among Whites estimated by the MSM were, compared to less than high school, PR = 0.85 (0.68–1.05) for high school, PR = 0.81 (0.64–1.03) for some college, and PR = 0.73

(0.56–0.94) for a college degree. As with health status, estimated effects of grandparent schooling on obesity among Blacks were minimal.

Higher grandparent schooling was associated with lower probabilities of current smoking but, unlike the other outcomes, the associations were stronger among Blacks than among Whites (Table 4.4[c]). Among Whites, only having a grandparent with a college degree was associated with a lower probability of current smoking compared to having only grandparents with less than high school degrees (estimated total effect PR = 0.71 [0.55–0.92]). Estimates of direct effects for the college-degree category were PR = 0.88 (0.66–1.17) for conventional adjustment and PR = 0.78 (0.58–1.04) for the MSM. In Blacks there was a suggestion of a slightly graded relation between higher grandparent schooling less smoking; estimated total effects were PR = 0.86 (0.69–1.08) for high school, PR = 0.84 (0.63, 1.12) for some college, and PR = 0.81 (0.56–1.17) for a college degree. Direct effects estimated with MSM were PR = 0.97 (0.75–1.25) for high school, PR = 0.89 (0.62–1.27) for some college and PR = 0.71 (0.45–1.13) for a college degree.

To facilitate comparison of results between Blacks and Whites, Figure 4.3 shows predicted probabilities for each of the three outcomes by grandparent schooling categories. The probabilities were adjusted for parent and participant schooling using MSMs (with parent and participant schooling set to a high school degree) and allow for modification of the effects of schooling of all three generations by race. Probabilities of poor health status and obesity were higher among Blacks, and the grandparent-schooling gradient evident among Whites was absent among Blacks. In contrast, predicted probabilities of current smoking were similar between the race groups and in both groups the probability of smoking was lower among participants who had a grandparent with a college degree.

Figure 4.4 shows prevalence ratios for MSM estimates of direct effects of each category of grandparent schooling (compared to less than high school) separated by whether the most-educated grandparent ever lived in the same state as the participant. Among both Blacks and Whites, the inverse association between grandparent schooling and poor health status was more pronounced when the most-educated grandparent lived in the same state (Figure 4.4[a]). In fact, point estimates among Blacks were in the unexpected direction when the most-educated grandparent lived in another state, although the imprecision of the estimates makes it difficult to determine how meaningful this finding is. The pattern of results was similar for obesity but the differences in estimated direct effects between participants whose most-educated grandparent lived in the same state and in another state were smaller (Figure 4.4[b]). For smoking, estimates for having a grandparent with a college degree were larger in magnitude when the grandparent lived in a different state (Figure 4.4[c]).

In sensitivity analyses, models incorporating only observed (i.e., not imputed) measures of grandparent schooling produced very similar estimates, as did models incorporating PSID sample weights.

Discussion

In a national sample of adults in the United States aged 25–55 in 2009, higher grandparent schooling was associated with better health status, less obesity, and less current smoking. Results were consistent with the possibility of a "direct effect" of grandparent schooling on health that is not entirely mediated by parent and participant schooling, although the difference in magnitude between estimated total and direct effects differed by outcome and level of grandparent schooling.

Our results are consistent with the limited past research linking higher grandparent schooling to better grandchild health and may also have implications for research documenting multigenerational family histories of socioeconomically patterned diseases.[17,18] For example, grandparent obesity and asthma have been associated with grandchild overweight and asthma, respectively, in the PSID.[50,51] In general, this analysis points to the need for more research investigating multigenerational influences of schooling on health, including influences unmediated by the schooling of the younger generations. It also suggests the need for more research about the ways grandparents influence grandchild health even when the grandparent and grandchild do not live in the same household. Much past research has focused on co-residing grandparents and grandparents who serve as primary caretakers of their grandchildren.[52] These topics are important but attention should also be given to the case of grandparents who do not reside with their grandchildren, which represents the majority of grandparents in the U.S.

We found that for health status and obesity, estimates were larger when the most-educated grandparent lived in the same state as the participant, while the opposite was true for smoking. The differences we observed provide preliminary evidence that social contact with grandparents—facilitated by geographical proximity—may play a role in how grandparents directly influence their grandchildren's health. However, the role of social contact may differ across outcomes and schooling levels. For example, it is possible that spending time with grandparents is beneficial for health status and, compared to grandparents with little schooling, those who had more schooling but lived in another state spent relatively little time with their grandchildren. On the other hand, college-educated grandparents living in the same state may have been more likely to smoke in the presence of their grandchildren, leading to a dampening of the beneficial influence of their schooling on grandchildren's smoking. Future research is needed

to identify mechanisms specific to different outcomes through which grandparent schooling may influence health. For example, our results are consistent with the interpretation that social contact with more highly educated grandparents may benefit health chiefly through psychosocial mechanisms, as reflected in better perceived health status, as opposed to through modeling of healthy behaviors.

The race differences we observed for health status and obesity support the diminishing returns hypothesis rather than the minority poverty hypothesis: graded associations between grandparent schooling and health were stronger among Whites than among Blacks and predicted probabilities of poor health outcomes differed more between the race groups at high grandparent schooling levels than low grandparent schooling levels. This pattern mirrors findings in other national samples for within-generation associations between schooling and these outcomes.[21,27,53] Higher education has also been more predictive of mortality and life expectancy in Whites than Blacks in national data.[28,54] In contrast, our estimates for smoking were similar among Black and White participants. This may reflect differences over time in how race modifies schooling–health relations depending on the specific health outcome. Currently in the U.S., poor health status and overweight are more prevalent among Blacks than Whites while smoking prevalence is approximately equal between Blacks and Whites.[16] Historically, Whites (in particular, White men) were more likely to smoke than Blacks early in the 20th century but this pattern reversed mid-century; Whites have also historically been more likely to quit smoking than Blacks.[55] The results here may reflect differences both in smoking initiation and cessation.

There was some evidence that conventional regression adjustment may underestimate direct effects in this context, although the differences were not large. MSMs have not been widely used to address questions in life course epidemiology despite the fact that the problem

they address—how to account for variables that are both effects of the exposure and confounders of mediator–outcome associations—is common in this field. It may be fruitful to consider MSMs in future research addressing life-course socioeconomic influences.

The reliance on self- and proxy-reported data in this analysis was a limitation. That said, distributions of health measures in the PSID align reasonably well with national estimates.[56,57] Another concern is that despite high wave-to-wave response rates in the PSID, there has been substantial attrition since the start of the study.[58] However, using data on PSID participants aged 0–16 in 1968, Fitzgerald did not find strong evidence of attrition bias in models of adult socioeconomic and health outcomes.[58] Attrition is higher among participants with lower education, lower income, and worse health.[58] It is therefore possible that the estimates here may be conservative if low education and poor health were strongly correlated in individuals and families who dropped out of the sample. Another consideration is that our analysis sample represents not only families that have persisted in participating in the PSID but also families that have successfully produced living adult members over three generations—this likely varies with schooling levels. For example, the infant mortality rate in the United States among mothers with less than a high school education is twice that among mothers with a college degree.[59]

We were also limited in our ability address cohort differences, which may be important for this analysis because of strong secular trends in schooling, obesity, and smoking throughout the 20th century. We tried to limit the influence of cohort effects by restricting the age range of our sample; we also accounted not only for participants' ages but also for the years of birth of both parents. Including grandparents' years of birth in our models did not change results but these values were imputed (i.e., missing in the original sample) for the large majority of observations (see Table 4.1). In a sensitivity analysis, we estimated separate models for

participants aged 25–34, 35–44, and 45–55. Although there was some evidence of variation in estimates between age groups for some outcomes, overall associations between grandparent schooling and the health outcomes persisted. Furthermore, because the outcomes were measured at only a single point in time, it is not possible to know whether any variation stemmed from age or cohort differences. Finally, we were not able to account for the health or health behaviors of the grandparents.

We made several simplifying assumptions that enabled us to address our question with these data, but that warrant further investigation. First, we used a single measure of grandparent schooling. It is possible that grandparent schooling influences grandchild health differently depending on lineage. For example, evolutionary biology literature suggests that maternal grandparents invest more strongly in their grandchildren than paternal grandparents.[60] Second, we did not address possible differences depending on the genders of the grandparents, parents, and grandchildren. Third, we used the maximum of grandparent schooling; other aspects of combined grandparent schooling—such as the minimum of grandparent schooling or the amount of variation between grandparents in their schooling—may be meaningful for grandchild health. Fourth, we did not examine possible interactions between grandparent schooling and the schooling of the parents and grandchildren. This is a methodological challenge because the choice of marginal structural models, based on our assumed causal structure, did produce somewhat different point estimates from conventional regression but precluded examination of interactions between the schooling of the different generations.[20] It is also a data challenge: in preliminary analyses, some combinations of schooling across three generations produced very small cell sizes. The rarest combinations were those in which there were large intergenerational differences in schooling; these may in fact be of particular importance for health.[61]

The identification of multigenerational effects of schooling on health may have important intervention and policy implications. The United States holds equal access to education as a fundamental societal tenet: the first statement of the U.S. Department of Education's mission is to "strengthen the Federal commitment to assuring access to equal educational opportunity for every individual."[62] Implicit in this statement is that through equal access to education, individuals should also have equal access to the benefits of education, including benefits to their health. If our health is affected not only by our own schooling but also by that of our parents and grandparents, the benefits we gain depend not only on our own educational opportunities but on those conferred to previous generations. Therefore, policies to reduce inequities in education may also serve not only to reduce inequities in health of the current generation but also to reduce health inequities in future generations. By the same token, failure to reduce current inequities in educational opportunities may contribute to health inequities in future generations.

References

1. Harper S, Lynch J, Hsu WL, Everson SA, Hillemeier MM, Raghunathan TE, Salonen JT, Kaplan GA. Life course socioeconomic conditions and adult psychosocial functioning. *Int J Epidemiol* 2002;**31**(2):395-403.

2. Best LE, Hayward MD, Hidajat MM. Life course pathways to adult-onset diabetes. *Soc Biol* 2005;**52**(3-4):94-111.

3. Bowen ME. Childhood socioeconomic status and racial differences in disability: evidence from the Health and Retirement Study (1998-2006). *Soc Sci Med* 2009;**69**(3):433-41.

4. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *J Health Econ* 2005;**24**(2):365-89.

5. Chichlowska KL, Rose KM, Diez-Roux AV, Golden SH, McNeill AM, Heiss G. Life course socioeconomic conditions and metabolic syndrome in adults: the Atherosclerosis Risk in Communities (ARIC) Study. *Ann Epidemiol* 2009;**19**(12):875-83.

6. Foster HW, Wu L, Bracken MB, Semenya K, Thomas J. Intergenerational effects of high socioeconomic status on low birthweight and preterm birth in African Americans. *J Natl Med Assoc* 2000;**92**(5):213-21.

7. Frank E, Elon L, Hogue C. Transgenerational persistence of education as a health risk: findings from the Women Physicians' Health Study. *J Womens Health (Larchmt)* 2003;**12**(5):505-12.

8. Gilman SE, Abrams DB, Buka SL. Socioeconomic status over the life course and stages of cigarette use: initiation, regular use, and cessation. *J Epidemiol Community Health* 2003;**57**(10):802-8.

9. Grossman M, Kaestner R. Effects of education on health. In: Behrman JR, Stacey N, eds. *The Social Benefits of Education*. Ann Arbor: University of Michigan Press, 1997;69-124.

10. Guralnik JM, Butterworth S, Wadsworth ME, Kuh D. Childhood socioeconomic status predicts physical functioning a half century later. *J Gerontol A Biol Sci Med Sci* 2006;**61**(7):694-701.

11. Luo Y, Waite LJ. The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. *J Gerontol B Psychol Sci Soc Sci* 2005;**60**(2):S93-S101.

12. Phillips JE, Marsland AL, Flory JD, Muldoon MF, Cohen S, Manuck SB. Parental education is related to C-reactive protein among female middle-aged community volunteers. *Brain Behav Immun* 2009;**23**(5):677-83.

13. Piko BF. Self-perceived health among adolescents: the role of gender and psychosocial factors. *Eur J Pediatr* 2007;**166**(7):701-8.

14. Tucker JS, Martinez JF, Ellickson PL, Edelen MO. Temporal associations of cigarette smoking with social influences, academic performance, and delinquency: a four-wave longitudinal study from ages 13-23. *Psychol Addict Behav* 2008;**22**(1):1-11.

15. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chronic Dis* 1987;**40 Suppl 2**:139S-161S.

16. Schiller JS, Lucas JW, Ward BW, Peregoy JA. Summary health statistics for U.S. adults: National Health Interview Survey, 2010. *Vital Health Stat 10* 2012(252):1-207.

17. Osler M, Andersen AM, Lund R, Holstein B. Effect of grandparent's and parent's socioeconomic position on mortality among Danish men born in 1953. *Eur J Public Health* 2005;**15**(6):647-51.

18. Krzyzanowska M. Inter-generational educational advancement and body height. *J Biosoc Sci* 2007;**39**(3):321-39.

19. Ahren-Moonga J, Silverwood R, Klinteberg BA, Koupil I. Association of higher parental and grandparental education and higher school grades with risk of hospitalization for eating disorders in females: the Uppsala birth cohort multigenerational study. *Am J Epidemiol* 2009;**170**(5):566-75.

20. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* 2009;**20**(1):18-26.

21. Farmer MM, Ferraro KF. Are racial disparities in health conditional on socioeconomic status? *Soc Sci Med* 2005;**60**:191-204.

22. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S. Socioeconomic status in health research: one size does not fit all. *JAMA* 2005;**294**(22):2879-88.

23. Crissey SR. Educational Attainment in the United States: 2007. *Population Characteristics*. Washington, DC: U.S. Census Bureau, 2009.

24. Goldman N, Kimbro RT, Turra CM, Pebley AR. Socioeconomic gradients in health for white and Mexican-origin populations. *Am J Public Health* 2006;**96**(12):2186-93.

25. Pappas G, Queen S, Hadden W, Fisher G. The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. *N Engl J Med* 1993;**329**(2):103-9.

26. Crimmins EM, Mayward MD, Seeman TE. Race/ethnicity, socioeconomic status, and health. In: Anderson NB, Bulatao RA, Cohen B, eds. *Critical Perspectives on Racial and*

Ethnic Differences in Health in Late Life. Washington, DC, USA: National Academies Press, 2004;310-352.

27. Shuey KM, Willson AE. Cumulative disadvantage and black-white disparities in life-course health trajectories. *Research on Aging* 2008;**30**(2):200-225.

28. Zajacova A, Hummer RA. Gender differences in education effects on all-cause mortality for white and black adults in the United States. *Soc Sci Med* 2009;**69**:529-537.

29. Cagney KA, Lauderdale DS. Education, wealth, and cognitive function in later life. *2002* 2002;**57B**(2):P163-P172.

30. Institute for Social Research. The Panel Study of Income Dynamics. http://psidonline.isr.umich.edu/ Accessed April 5, 2010.

31. Christenson BA, Johnson NE. Educational inequality in adult mortality: an assessment with death certificate data from Michigan. *Demography* 1995;**32**(2):215-229.

32. Kim J. Intercohort trends in the relationship between education and health: examining physical impairment and depressive symptomatology. *J Aging Health* 2008;**20**(6):671-93.

33. Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health* 1992;**82**(6):816-20.

34. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior* 1997;**38**:21-37.

35. Dowd J, Zajacova A. Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *Int J Epidemiol* 2007;**36**:1214-1221.

36. Manor O, Matthews S, Power C. Dichotomous or categorical? Analysing self-rated health and lifetime social class. *Int J Epidemiol* 2000;**29**:149-157.

37. Centers for Disease Control and Prevention. Defining overweight and obesity. http://www.cdc.gov/obesity/adult/defining.html Accessed January 30, 2013.

38. Grieger LD, Schoeni RF, Danziger S. Accurately measuring the trend in poverty in the United States using the Panel Study of Income Dynamics. *PSID Technical Series*. Ann Arbor, MI: Institute for Social Research, 2008.

39. Hajat A, Kaufman JS, Rose KM, Siddiqi A, Thomas JC. Long-term effects of wealth on mortality and self-rated health status. *Am J Epidemiol* 2010;**173**(2):192-200.

40. Raghunathan TE, Lepkowski J, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 2001;**27**(1):85-95.

41. Institute for Social Research. IVEware: Imputation and Variance Estimation Software. http://www.isr.umich.edu/src/smp/ive/ Accessed December 8, 2011.

42. Raghunathan TE. What do we do with missing data? Some options for analysis of incomplete data. *Annu Rev Public Health* 2004;**25**:99-117.

43. Moodie EEM, Delaney JAC, Lefebvre G, Platt RW. Missing confounding data in marginal structural models: a comparison of inverse probability weighting and multiple imputation. *Int J Biostat* 2008;**4**(1):Article 13.

44. Sas Institute Inc. *SAS/STAT 9.2 User's Guide*. Vol. SAS Institute Inc. Cary, NC, 2008.

45. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**(7):702-6.

46. Zou G, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res* 2011;**0**(0).

47. Petersen ML, Sinisi SE, van der Laan MJ. Estimation of direct causal effects. *Epidemiology* 2006;**17**(3):276-284.

48. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**(6):1173-82.

49. Nandi A, Glymour MM, Kawachi I, VanderWeele TJ. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology* 2012;**23**(2):223-32.

50. Davis MM, McGonagle K, Schoeni RF, Stafford F. Grandparental and parental obesity influences on childhood overweight: implications for primary care practice. *J Am Board Fam Med* 2008;**21**(6):549-54.

51. Valerio MA, Andreski PM, Schoeni RF, McGonagle KA. Examining the association between childhood asthma and parent and grandparent asthma status: implications for practice. *Clin Pediatr (Phila)* 2010;**49**(6):535-41.

52. Thomas JL, Sperry L, Yarbrough MS. Grandparents as parents: research findings and policy recommendations. *Child Psychiatry Hum Dev* 2000;**31**(1):3-22.

53. Cohen AK, Rehkopf DH, Deardorff J, Abrams B. Education and obesity at age 40 among American adults. *Soc Sci Med* 2012.

54. Olshansky SJ, Antonucci T, Berkman L, Binstock RH, Boersch-Supan A, Cacioppo JT, Carnes BA, Carstensen LL, Fried LP, Goldman DP, Jackson J, Kohli M, Rother J, Zheng Y, Rowe J. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Aff (Millwood)* 2012;**31**(8):1803-13.

55. Escobedo LG, Peddicord JP. Smoking prevalence in US birth cohorts: the influence of gender and education. *Am J Public Health* 1996;**86**(2):231-6.

56. Andreski PM, McGonagle KA, Schoeni RF. An analysis of the quality of the health data in the Panel Study of Income Dynamics. *Technical Paper Series*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, 2009.

57. National Center for Health Statistics. Health, United States, 2009: With Special Feature on Medical Technology. Hyattsville, MD, 2010.

58. Fitzgerald JM. Attrition in Models of Intergenerational Links Using the PSID with Extensions to Health and to Sibling Models. *B E J Econom Anal Policy* 2011;**11**(3).

59. Robert Wood Johnson Foundation Commission to Build a Healthier America. America's Health Starts With Healthy Children: How do States Compare? Washington, DC, 2008.

60. Coall DA, Hertwig R. Grandparental investment: past, present, and future. *Behav Brain Sci* 2010;**33**(1):1-19; discussion 19-40.

61. Morgenstern H. Socioeconomic factors: concepts, measurement, and health effects. In: Ostfeld AM, Eaker ED, eds. *Measuring Psychosocial Variables in Epidemiologic Studies of Cardiovascular Disease* National Institutes of Health, 1985.

62. U.S. Department of Education. About ED. Overview: Mission. http://www2.ed.gov/about/overview/mission/mission.html?src=rt Accessed October 26, 2010.

Figure 4.1 Study sample

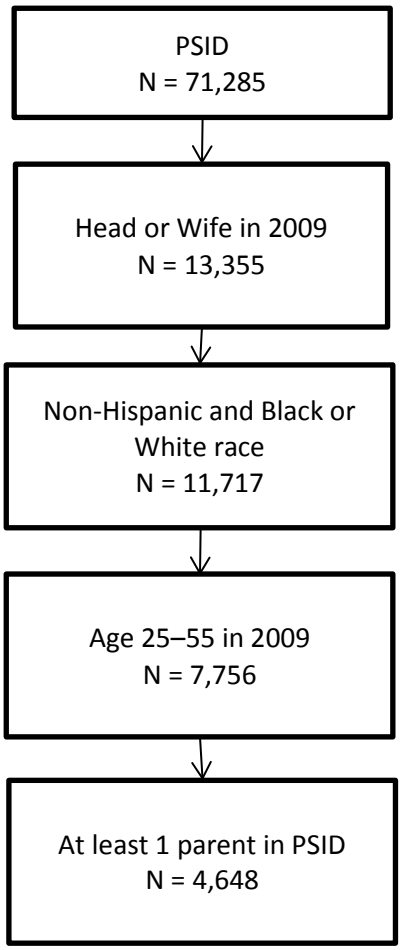
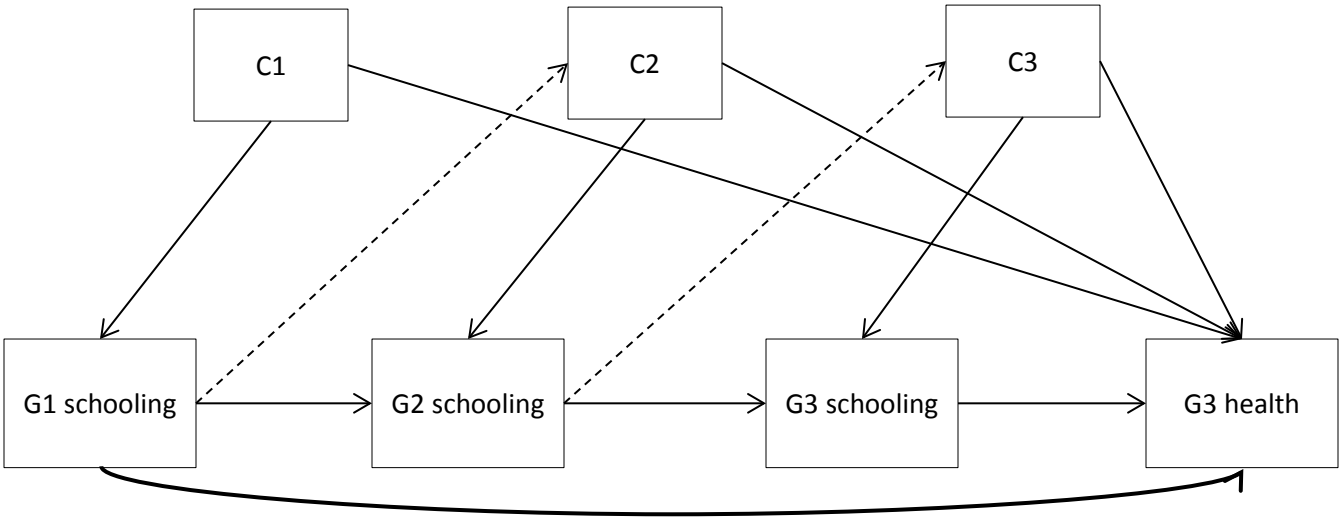


Figure 4.2 Causal diagram of relationships between grandparent (G1), parent (G2), and grandchild (G3) schooling and grandchild adult health



The thick arrow represents the direct effect grandparent schooling on grandchild health. The dotted arrows represent consequences of grandparent schooling that confound associations between parent and grandchild schooling and grandchild health. Although they are shown separately to facilitate interpretation of the diagram, C1, C2, and C3 are not mutually exclusive. In our models, C1 includes participant sex, race, age, Head vs. Wife status, index exam wave, and presence of a grandparent in the same state. C2 includes C1 as well as parent years of birth and an indicator for at least one parent being poor while growing up. C3 includes C2 as well as mother's age and marital status at the time of the participant's birth, whether the participant lived with both parents growing up, whether the participant was poor while growing up, average family income-to-poverty ratio when the participant was less than 18 years old, parent health status, parent obesity, whether a parent smoked when the participant was a child, and participant health status during childhood.

Table 4.1 Characteristics of original and imputed sample (G3), by race

	Total			Black			White		
	Original sample		Imputed sample	Original sample		Imputed sample	Original sample		Imputed sample
	N or range	% or median	% or median	N or range	% or median	% or median	N or range	% or median	% or median
Total	4648		16,200	1858		46,450	2790		69,750
Head/Wife & sex status									
Male Head	2018	43%	43%	683	37%	37%	1335	48%	48%
Female Head	1239	27%	27%	804	43%	43%	435	16%	16%
Female Wife	1391	30%	30%	371	20%	20%	1020	37%	37%
Age	25–55	38	38	25-55	39	39	25-55	37	37
Race									
non-Hispanic Black	1858	40%	40%	--	--	--	--	--	--
non-Hispanic White	2790	60%	60%	--	--	--	--	--	--
Marital status				--	--	--	--	--	--
Married	2362	51%	51%	575	31%	31%	1787	64%	64%
Never married	1395	30%	30%	823	44%	44%	572	21%	21%
Divorced/separated/widowed	891	19%	19%	460	25%	25%	431	15%	15%
Income-to-poverty ratio	0–303	3.4	3.4	0–29	2.3	2.3	0–303	4.3	4.3
< 1	568	12%	12%	403	22%	22%	165	6%	6%
1–1.9	707	15%	15%	424	23%	23%	283	10%	10%
2–4.9	1992	43%	43%	775	42%	42%	1217	44%	44%
5+	1381	30%	30%	256	14%	14%	1125	40%	40%
Years completed schooling									
< 12	399	9%	8%	252	14%	13%	147	5%	5%
12	1685	38%	36%	790	43%	42%	895	33%	32%
13–15	1326	30%	29%	554	30%	30%	772	29%	28%
16+	1129	25%	27%	239	13%	15%	890	33%	35%
Missing	109	2%	---	23	1%	---	86	3%	---
Maximum of parents' schooling									
< 12 grades	554	12%	12%	369	20%	19%	185	7%	7%
12 grades; high school	1519	33%	33%	745	41%	40%	774	28%	28%
> 12 grades, no BA	1179	26%	26%	448	25%	26%	731	26%	26%
BA+	1356	29%	29%	265	15%	15%	1091	39%	39%
Missing	40	1%	---	31	2%	---	9	0.3%	---
Maximum grandparents' schooling									
< 12 grades	1386	30%	25%	905	50%	40%	481	17%	15%

Table 4.1 Characteristics of original and imputed sample (G3), by race

	Total			Black			White		
	Original sample		Imputed sample	Original sample		Imputed sample	Original sample		Imputed sample
	N or range	% or median	% or median	N or range	% or median	% or median	N or range	% or median	% or median
12 grades; high school	1492	33%	34%	575	32%	35%	917	33%	33%
> 12 grades, no BA	857	19%	22%	205	11%	17%	652	24%	26%
BA+	843	18%	18%	138	8%	8%	705	26%	26%
Missing	70	2%	---	35	2%	---	35	1%	---
Grandparent in same state	3163	74%	74%	1241	74%	72%	1922	74%	76%
Missing	361	8%	---	170	10%	---	191	7%	---
Average income-to-poverty ratio when aged < 18	0.1–29	2.4	2.4	0.1–7	1.3	1.3	0.4–29	3.2	3.1
Missing	867	19%	---	288	16%	---	579	21%	---
Poor while growing up	1260	27%	27%	648	35%	35%	612	22%	22%
Missing	9	0.2%	---	3	0.2%	---	6	0.2%	---
At least one parent poor while growing up	2620	57%	57%	1268	70%	70%	1352	49%	49%
Missing	84	2%	---	43	2%	---	41	1%	---
Mother's year of birth	1908–1969	1946	1946	1911–1969	1947	1947	1908-1967	1946	1946
Missing	75	2%	---	38	2%	---	37	1%	---
Father's year of birth	1893–1974	1943	1941	1893-1974	1940	1938	1901-1965	1943	1943
Missing	985	21%	---	725	39%	---	260	9%	---
Mother's marital status when born									
Married	3378	81%	80%	945	58%	59%	2433	95%	94%
Never married	627	15%	15%	559	35%	33%	68	3%	3%
Other	168	4%	5%	112	7%	8%	56	2%	3%
Missing	475	10%	---	242	13%	---	233	8%	---
Lived w/both natural parents most of time until age 16	2665	64%	64%	904	52%	52%	1761	73%	73%
Missing	489	11%	---	109	6%	---	380	14%	---
Mother's age when born	10–45	26	25	11-45	24	24	10-45	26	26
Missing	137	3%	---	46	2%	---	91	3%	---
Mother age < 20 when born	643	14%	14%	397	22%	22%	246	9%	9%
Missing	137	3%	---	46	2%	---	91	3%	---
Maternal grandmother year of birth	1886–1968	1927	1925	1889–1954	1928	1926	1886–1968	1926	1925

Table 4.1 Characteristics of original and imputed sample (G3), by race

	Total			Black			White		
	Original sample		Imputed sample	Original sample		Imputed sample	Original sample		Imputed sample
	N or range	% or median	% or median	N or range	% or median	% or median	N or range	% or median	% or median
Missing	3148	68%	---	1095	59%	---	2053	74%	---
Maternal grandfather year of birth	1884–1951	1924	1923	1893–1951	1924	1923	1884–1949	1924	1923
Missing	3643	78%	---	1414	76%	---	2229	80%	---
Paternal grandmother year of birth	1895–1954	1924	1923	1895–1954	1926	1924	1898–1945	1923	1923
Missing	3709	80%	---	1585	85%	---	2124	76%	---
Paternal grandfather year of birth	1887–1946	1921	1921	1887–1946	1921	1921	1896–1943	1922	1921
Missing	3956	85%	---	1708	92%	---	2248	81%	---
Parent smoked during childhood	2784	60%	60%	1079	58%	58%	1705	61%	61%
Missing	16	0.3%	---	11	1%	---	5	0.2%	---
Parent ever reported fair/poor health status 1984–2009	2690	61%	78%	1339	77%	85%	1351	50%	72%
Missing	206	4%	---	109	6%	---	97	3%	---
Parent ever obese 1986, 1999–2009	1221	39%	63%	621	45%	66%	600	34%	62%
Missing	1493	32%	---	482	26%	---	1011	36%	---
Poor health age 0–16	180	5%	5%	333	23%	6%	93	4%	4%
Missing	864	19%	---	409	22%	---	455	16%	---
Health status									
Excellent	895	19%	19%	282	15%	15%	613	22%	22%
Very good	1726	37%	37%	589	32%	32%	1137	41%	41%
Good	1429	31%	31%	654	35%	35%	775	28%	28%
Fair	496	11%	11%	286	15%	15%	210	8%	8%
Poor	101	2%	2%	47	3%	3%	54	2%	2%
Missing	1	0.02%	---	0	0%	---	1	0.04%	---
Current smoker	1176	25%	25%	502	27%	27%	674	24%	24%
Missing	3	0.1%	---	1	0.1%	---	9	0.3%	---
Obese	1500	33%	33%	761	42%	42%	739	27%	27%
Missing	90	2%	---	36	2%	---	54	2%	---

Table 4.2 Percents or median values of selected characteristics among participants (G3) with each level of grandparent schooling, by race[a]

	Black					White				
	< HS	HS	Some college	College degree		< HS	HS	Some college	College degree	
	% or med	% or med	% or med	% or med	p[b]	% or med	% or med	% or med	% or med	p[b]
Total	40%	35%	17%	8%		15%	33%	26%	26%	
Wife (vs. Head)	21%	20%	18%	22%	0.81	31%	40%	38%	34%	0.01
Age	44	36	34	33	<0.001	45	40	35	33	<0.001
Female	64%	62%	62%	67%	0.73	49%	54%	54%	50%	0.20
Marital status					0.008					<0.001
Married	34%	30%	27%	28%		63%	69%	62%	61%	
Never married	38%	46%	54%	48%		13%	15%	24%	30%	
Divorced/separated/widowed	28%	24%	20%	24%		24%	17%	15%	9%	
Income-to-poverty ratio	2.3	2.3	2.2	2.3	0.77	3.7	4.3	4.3	4.6	0.03
Years completed schooling	12	12	12	13	0.36	12	13	14	15	<0.001
Maximum of parents' completed years schooling					<0.001					<0.001
< 12 grades	30%	14%	10%	10%		22%	6%	4%	1%	
12 grades; high school	44%	42%	33%	26%		45%	35%	21%	15%	
> 12 grades, no BA	17%	30%	36%	31%		18%	30%	32%	21%	
BA+	9%	14%	20%	34%		14%	30%	42%	63%	
Grandparent ever lived in same state	64%	75%	79%	83%	<0.001	63%	77%	78%	79%	<0.001
Grandparent with maximum schooling ever lived in same state	64%	52%	43%	50%	0.001	63%	61%	54%	56%	0.08
Average income-to-poverty ratio when aged < 18	1.2	1.4	1.5	1.5	0.003	2.4	3.0	3.3	3.8	<0.001
Poor while growing up	39%	32%	33%	29%	0.05	27%	19%	24%	22%	0.03
At least one parent poor while growing up	81%	64%	58%	63%	<0.001	74%	55%	45%	31%	<0.001
Mother's year of birth	1939	1951	1951	1951	<0.001	1937	1944	1948	1948	<0.001
Father's year of birth	1934	1940	1943	1945	<0.001	1934	1941	1945	1947	<0.001
Mother unmarried when born	37%	40%	52%	45%	0.02	6%	5%	8%	3%	0.04
Lived w/both natural parents most of time until age 16	59%	51%	40%	45%	<0.001	76%	76%	67%	72%	0.003
Mother age < 20 when born	16%	24%	28%	29%	0.001	13%	9%	10%	6%	0.003
Parent smoked during childhood	54%	63%	59%	62%	0.10	69%	64%	64%	51%	<0.001

Table 4.2 Percents or median values of selected characteristics among participants (G3) with each level of grandparent schooling, by race[a]

	Black					White				
	< HS	HS	Some college	College degree		< HS	HS	Some college	College degree	
	% or med	% or med	% or med	% or med	p[b]	% or med	% or med	% or med	% or med	p[b]
Parent ever reported fair/poor health status 1984–2009	90%	83%	81%	79%	0.06	84%	77%	70%	62%	<0.001
Parent ever obese 1986, 1999–2009	67%	65%	66%	63%	0.96	57%	64%	62%	61%	0.87
Ever reported poor health age 0–16	6%	6%	5%	4%	0.84	5%	4%	5%	4%	0.71
Poor health status	19%	19%	16%	15%	0.55	14%	10%	9%	6%	<0.001
Obese	45%	41%	39%	38%	0.29	36%	29%	26%	21%	<0.001
Current smoker	28%	26%	26%	25%	0.91	27%	25%	27%	18%	0.004

[a] Uses imputed data.

[b] From unadjusted multinomial logistic regression with clustering by 1968 PSID family.

Table 4.3 Percents of poor health outcomes among participants (G3) by category of selected predictors, by race[a]

	Black						White					
	Poor health status		Obese		Current smoker		Poor health status		Obese		Current smoker	
	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]
Total	18		42		27		9		27		24	
PSID sample member		0.81		0.02		<0.001		0.35		<0.001		<0.001
Head	18		41		30		10		30		28	
Wife	18		47		16		9		23		17	
Age		<0.001		0.08		0.03		<0.001		0.003		0.008
25–34	15		39		28		7		24		28	
35–44	14		42		21		10		30		21	
45–55	25		45		31		13		29		21	
Sex		0.24		<0.001		<0.001		0.21		0.04		<0.001
Female	19		46		24		10		29		21	
Male	17		36		33		9		25		27	
Marital status		0.02		0.10		<0.001		<0.001		0.27		<0.001
Married	14		46		18		7		27		17	
Never married	17		40		32		11		25		34	
Divorced/separated/widowed	23		42		30		16		30		38	
Income-to-poverty ratio		<0.001		0.09		<0.001		<0.001		0.02		<0.001
< 1	29		40		41		28		29		52	
1–1.9	23		45		34		22		34		42	
2–4.9	13		44		20		9		28		25	
≥ 5	8		35		13		4		25		15	
Years completed schooling		<0.001		0.19		<0.001		<0.001		<0.001		<0.001
< 12 grades	26		38		53		26		35		50	
12 grades; high school	19		43		29		13		33		36	
> 12 grades, no BA	17		45		21		9		30		25	
BA+	10		38		12		4		19		9	
Maximum of parents' completed years schooling		<0.001		0.19		0.04		<0.001		<0.001		<0.001
< 12 grades	27		45		30		24		39		38	
12 grades; high school	17		44		28		9		31		28	
> 12 grades, no BA	13		38		27		10		30		23	
BA+	16		42		20		7		20		20	

Table 4.3 Percents of poor health outcomes among participants (G3) by category of selected predictors, by race[a]

	Black						White					
	Poor health status		Obese		Current smoker		Poor health status		Obese		Current smoker	
	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]
Grandparent ever lived in same state		<0.001		0.41		0.19		0.01		0.92		0.82
No	23		40		29		12		27		24	
Yes	16		43		26		9		27		24	
Grandparent with maximum schooling ever lived in same state		0.003		0.47		0.40		0.002		0.66		0.47
No	21		41		28		12		28		23	
Yes	15		43		26		8		27		25	
Average income-to-poverty ratio when aged < 18		0.26		0.18		0.33		<0.001		<0.001		<0.001
< 1	21		40		27		24		36		49	
1–1.9	16		45		28		14		36		34	
2–4.9	18		40		25		9		27		22	
≥ 5	16		40		14		4		17		18	
Poor while growing up		0.04		0.56		0.29		0.04		0.03		0.09
No	17		43		26		9		26		23	
Yes	21		41		29		12		31		27	
At least one parent poor while growing up		0.06		0.82		0.20		<0.001		0.18		0.08
No	15		42		29		7		26		23	
Yes	19		42		26		12		29		26	
Mother unmarried when born		0.34		0.18		0.14		0.004		0.001		0.001
No	19		43		26		9		26		23	
Yes	17		40		29		17		41		36	
Lived w/both natural parents most of time until age 16		0.51		0.44		0.19		0.03		0.53		<0.001
No	19		41		29		12		28		31	
Yes	17		43		26		9		27		21	
Mother age < 20 when born		0.54		0.58		0.85		0.31		0.001		0.03
No	18		42		27		9		26		24	
Yes	17		43		27		11		37		30	

Table 4.3 Percents of poor health outcomes among participants (G3) by category of selected predictors, by race[a]

	Black						White					
	Poor health status		Obese		Current smoker		Poor health status		Obese		Current smoker	
	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]	%	p[b]
Parent smoked during childhood		0.33		0.99		0.08		0.002		0.002		<0.001
No	17		42		25		7		24		18	
Yes	19		42		29		11		30		28	
Parent ever reported fair/poor health status 1984–2009		0.16		0.10		0.77		0.44		0.45		0.94
No	12		35		28		7		24		23	
Yes	19		43		27		11		28		25	
Parent ever obese 1986, 1999–2009		0.68		<0.001		0.50		0.45		0.13		0.87
No	17		35		26		8		23		24	
Yes	18		46		28		10		30		24	
Ever reported poor health age 0–16		0.002		0.15		0.47		<0.001		0.12		0.006
No	17		42		27		9		30		24	
Yes	29		50		30		24		34		36	
Poor health status		--		<0.001		<0.001		--		<0.001		<0.001
No	--		39		25		--		25		23	
Yes	--		54		36		--		46		37	
Obese		<0.001		--		<0.001		<0.001		--		0.07
No	14		--		31		7		--		25	
Yes	23		--		22		16		--		21	
Current smoker		<0.001		<0.001		--		<0.001		0.07		--
No	16		45		--		8		28		--	
Yes	24		34		--		15		24		--	

[a] Uses imputed data.

[b] From unadjusted logistic regression with clustering by 1968 PSID family.

Table 4.4 Prevalence ratios of (a) poor health status, (b) obesity, and (c) current smoking by grandparent schooling (vs. less than high school) and race

	(a) Poor Health Status						(b) Obesity					
	High school		Some college		College degree		High school		Some college		College degree	
	PR	95% CI	PR	95% CI	PR	95% CI	PR	95% CI	PR	95% CI	PR	95% CI
Black												
Total effect[a]	1.11	(0.84, 1.46)	1.01	(0.69, 1.49)	0.97	(0.60, 1.56)	0.93	(0.77, 1.12)	0.94	(0.73, 1.21)	0.92	(0.68, 1.23)
Adjusted direct effect[b]	1.16	(0.87, 1.54)	1.06	(0.71, 1.57)	1.03	(0.63, 1.69)	0.92	(0.76, 1.11)	0.92	(0.71, 1.20)	0.91	(0.67, 1.24)
MSM direct effect[c]	0.98	(0.73, 1.32)	1.15	(0.78, 1.68)	1.09	(0.68, 1.74)	0.91	(0.74, 1.11)	0.92	(0.68, 1.23)	1.02	(0.74, 1.39)
White												
Total effect[a]	0.82	(0.58, 1.15)	0.75	(0.51, 1.11)	0.54	(0.36, 0.82)	0.86	(0.70, 1.06)	0.80	(0.64, 1.00)	0.65	(0.51, 0.82)
Adjusted direct effect[b]	0.91	(0.64, 1.29)	0.90	(0.59, 1.37)	0.76	(0.49, 1.20)	0.92	(0.74, 1.15)	0.90	(0.71, 1.15)	0.79	(0.61, 1.02)
MSM direct effect[c]	0.84	(0.57, 1.22)	0.81	(0.52, 1.27)	0.71	(0.44, 1.15)	0.85	(0.68, 1.05)	0.81	(0.64, 1.03)	0.73	(0.56, 0.94)

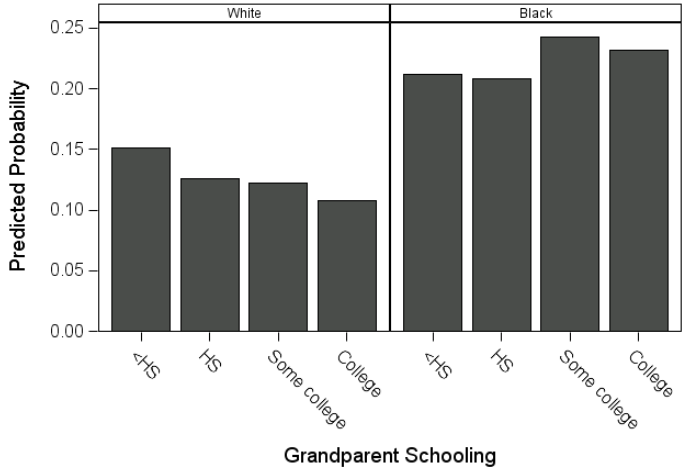
	(c) Current Smoking					
	High school		Some college		College degree	
	PR	95% CI	PR	95% CI	PR	95% CI
Black						
Total effect[a]	0.86	(0.69, 1.08)	0.84	(0.63, 1.12)	0.81	(0.56, 1.17)
Adjusted direct effect[b]	0.87	(0.69, 1.09)	0.85	(0.63, 1.15)	0.84	(0.57, 1.22)
MSM direct effect[c]	0.97	(0.75, 1.25)	0.89	(0.62, 1.27)	0.71	(0.45, 1.13)
White						
Total effect[a]	0.98	(0.78, 1.24)	1.00	(0.79, 1.28)	0.71	(0.55, 0.92)
Adjusted direct effect[b]	1.07	(0.84, 1.36)	1.13	(0.87, 1.45)	0.88	(0.66, 1.17)
MSM direct effect[c]	0.95	(0.75, 1.19)	0.99	(0.77, 1.28)	0.78	(0.58, 1.04)

[a] Adjusted for sex, race, age, Head vs. Wife status, and ever having a grandparent living in the same state.

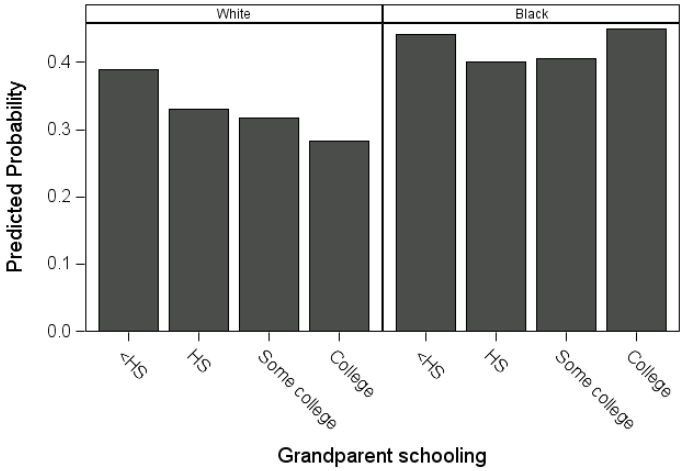
[b] Adjusted for sex, race, age, Head vs. Wife status, ever having a grandparent living in the same state, parent schooling, participant schooling, having a parent who grew up poor, whether reported growing up poor, average family income-to-poverty ratio before age 18, unmarried mother at birth, teen mother, whether lived with both parents growing up, parent health status, parent obesity, parent smoking during participant's childhood, and childhood health status.

[c] Marginal structural model (MSM) adjusted for race, parent schooling, and participant schooling.

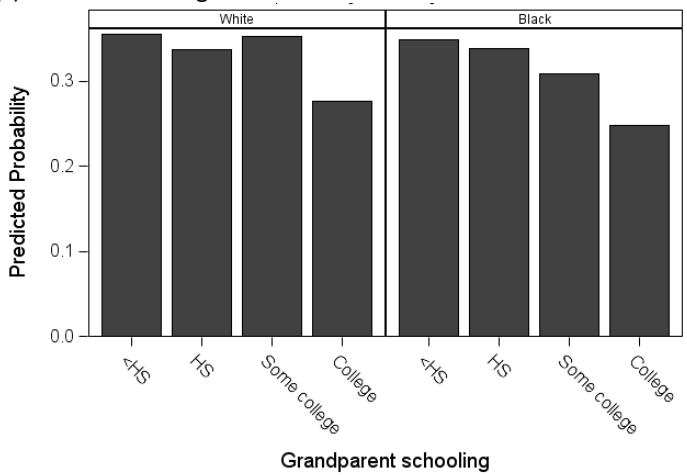
Figure 4.3 Predicted probabilities of (a) poor health status, (b) obesity, and (c) current smoking, by grandparent schooling and race



(b) Obesity



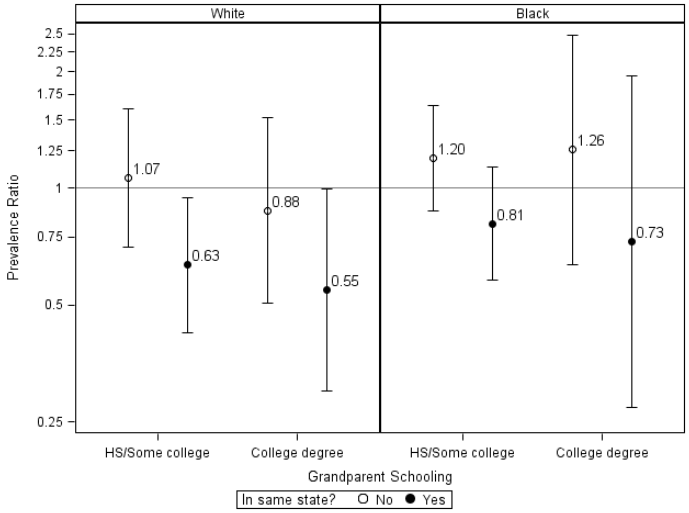
(c) Current Smoking



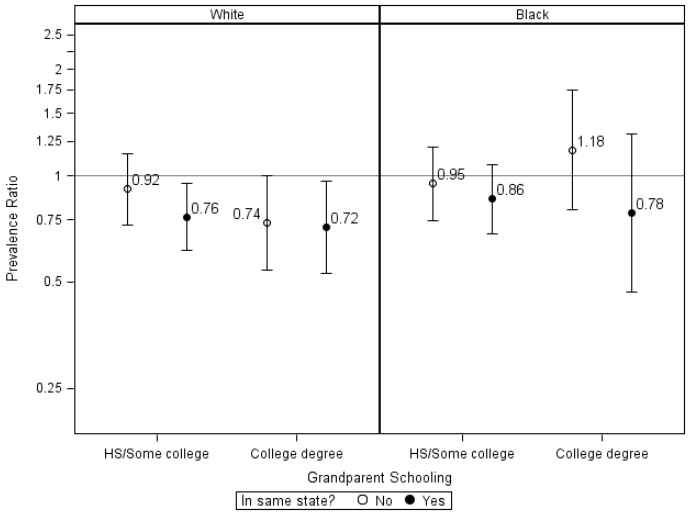
Probabilities estimated using marginal structural models adjusted for race, parent schooling, and participant schooling. Parent and participant schooling were both set to a high school degree and allowed to vary by race.

Figure 4.4 Prevalence ratios of (a) poor health status, (b) obesity, and (c) current smoking for grandparent schooling (vs. less than high school), by whether the highest-educated grandparent ever lived in the same state as the participant and race

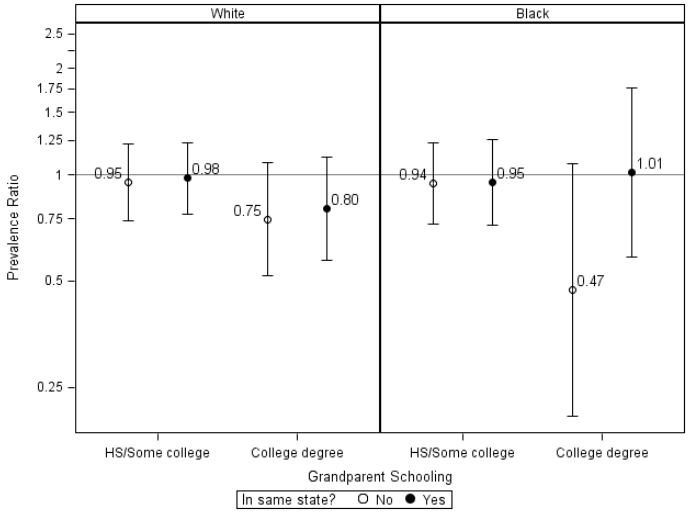
(a) Poor Health Status



(b) Obesity



(c) Current Smoking



From marginal structural models adjusted for race, parent schooling, and participant schooling.

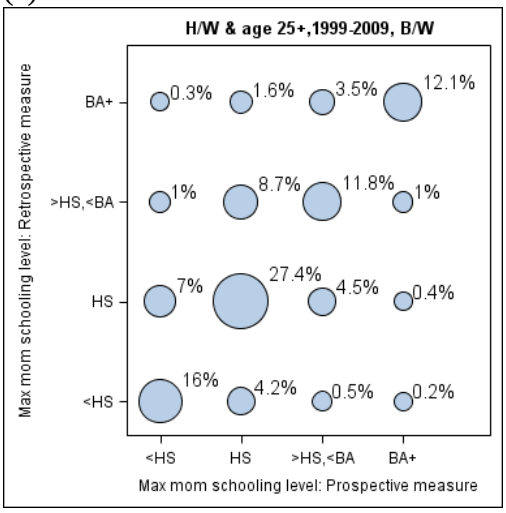
Appendix

Comparisons of prospective and retrospective measures of G1 and G2 schooling

In most cases, information for the prospective measure was collected from the Head on behalf of the Head and Wife in the first study wave in which the person was a Head or Wife, and was updated for all PSID Heads and Wives in several update study waves. Information for the retrospective measure was collected from the Head on behalf of the parents of the Head and Wife in the first study wave in which the person was a Head or Wife. The prospective and retrospective measures agree fairly well (Figure 4.A1). There is a tendency for the retrospective measures to be higher than the prospective measures. This is consistent with the fact that the information for the retrospective measures was collected later in time (Figure 4.A2), suggesting that the prospective measures had not been updated. However, it is also possible that the discrepancies result from a tendency by the G2 and G3 generation to overreport the schooling of the previous generation.

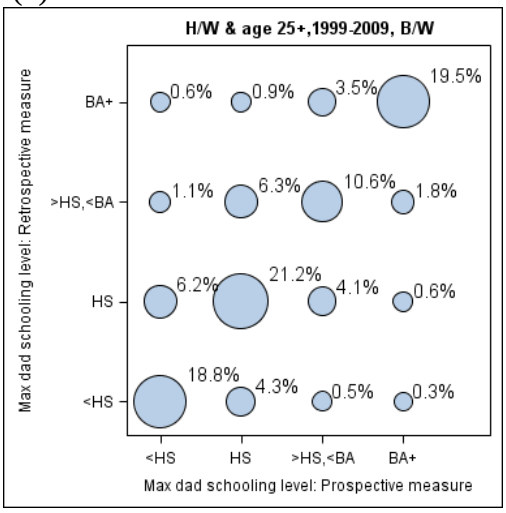
Figure 4.A1 Comparison of retrospective and prospective measures of (a) mother's schooling, (b) father's schooling, (c) maximum of parents' schooling, and (d) maximum of grandparents' schooling

(a)



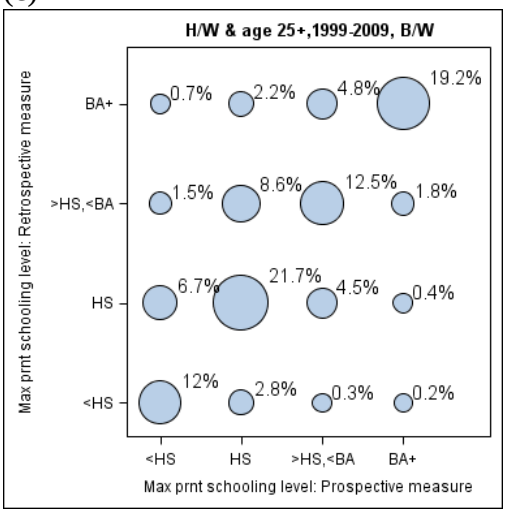
Weighted kappa = 0.54
Effective sample size = 6113

(b)



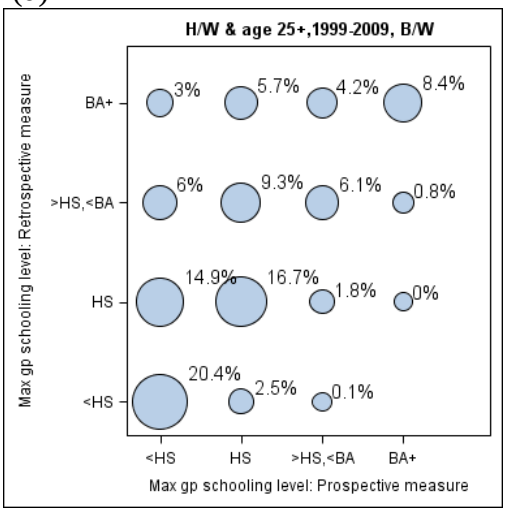
Weighted kappa = 0.72
Effective sample size = 4666

(c)



Weighted kappa = 0.65
Effective sample size = 6287

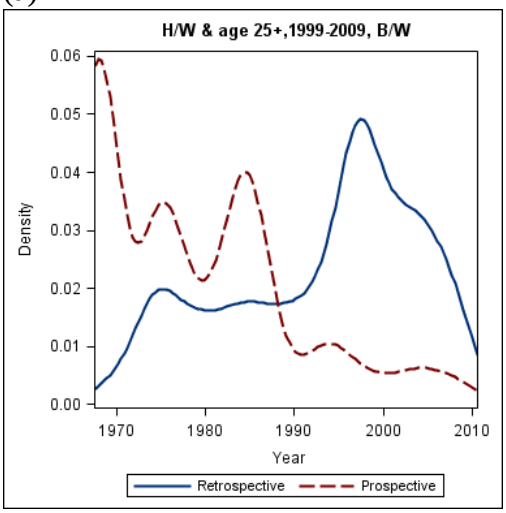
(d)



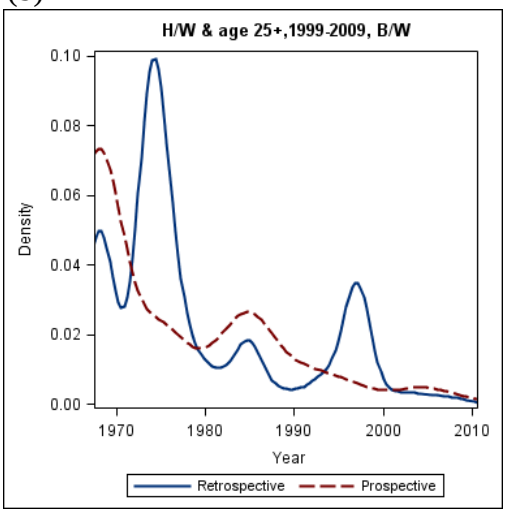
Weighted kappa = 0.45
Effective sample size = 2668

Figure 4.A2 Year in which retrospective and prospective measure of (a) G2 and (b) G1 schooling was reported

(a)



(b)



The years in which information was updated are clearly visible as peaks: 1976, 1985, and 1997.

CHAPTER 5

DISCUSSION

Introduction

This project used a life course approach to examine longitudinal processes that contribute to the strong associations between higher educational attainment and better health in the United States using data from the Panel Study of Income Dynamics (PSID). We specifically addressed how bidirectional causation and time-dependent causation may influence observed associations between schooling and health. In Chapter 2, poor health over a 10-year period among young people aged 5–14 at baseline was associated with less completed schooling at the end of follow-up, and the health-related difference in educational progress was most pronounced among participants who reported poor health in all three waves. In Chapter 3, higher average academic achievement 1997–2002 in young people aged 3–14 at baseline was associated with better health in 2007(better global health status, lower BMI, less serious psychological distress), although the associations were small in magnitude and more consistent among girls than among boys. In Chapter 4, there were graded inverse associations among non-Hispanic White adults aged 25–55 in 2009 between higher grandparent schooling and poor global health status, current and ever smoking, and obesity. For all the outcomes except ever smoking, graded associations persisted after accounting for parent and participant schooling. Associations among Blacks were similar to those among Whites for smoking but minimal for health status and obesity.

Links to Associations between Educational Attainment and Adult Health

These results reflect the complexity and lifelong nature of processes underlying associations between educational attainment and health in adults. The results in Chapter 2, which document the emergence in late adolescence of health-related disparities in completed schooling, demonstrate the possible contribution of reverse causation to estimates of effects of educational attainment on health. The magnitude of the associations suggest that poorer overall health status may have a meaningful impact on education progress: among participants aged 11–14 at baseline, those with poorer health status over all study waves completed on average 1.28 fewer years of schooling over the course of a decade. However, only 5% of participants had poor health status in all three waves. In general, the associations were not large enough to conclude that reverse causation may fully explain, or even be the major driver of, adult education attainment–health associations in the general population. Furthermore, it remains an open question whether the differences observed here will be compounded when the cohort matures further and greater schooling disparities appear or, instead, if children whose educational progress was initially slowed by poor health will eventually catch up.

The Chapter 3 results are broadly consistent with the hypothesis that higher academic achievement provides future health benefits and, more generally, that schooling benefits health through cognitive or psychosocial mechanisms that are distinct from the benefits of higher adult income and may accrue during childhood and adolescence. By using a relatively long follow-up time and a study sample that spanned early childhood through early adulthood, our goal was to examine the emergence of schooling-related health disparities that may persist into adulthood. In this way, our results help link existing studies relating academic achievement to health in young people in the short term to studies relating schooling to longer-term health in adults. However, as

is the case in the Chapter 2 analysis, we cannot know if the associations observed in Chapter 3 will in fact persist as this young cohort continues to mature. In addition, the associations were small in magnitude among girls and minimal among boys.

The effects of grandparent schooling on adult health estimated in Chapter 4 suggest multigenerational transmission of health benefits to schooling. Intragenerational schooling–health associations, therefore, may to some extent be amplified by intergenerational effects. By the same token, apparent blunting of the beneficial effects of schooling for health in some populations may to some extent reflect multigenerational schooling inequities. However, the results here also suggest that the degree of intergenerational influence, as well as the extent to which influence is exerted through pathways mediated by the schooling of later generations, varies depending on the health outcome. At least for some outcomes, the degree to which grandparent schooling contributes to intragenerational schooling–health associations may also depend on the amount of contact between the grandchild and grandparent. Finally, there may be synergistic effects of high educational attainment across multiple generations; I was not able to fully address this issue with these data and analysis methods.

Demographic Variation in Schooling–Health Relations

The associations observed in this project varied considerably between demographic groups. In Chapter 2, health-related differences in completed schooling were largest in participants who were older at baseline. In Chapter 3, academic achievement predicted future health status, body mass index, and psychological distress in girls but not boys. In Chapter 4, grandparent schooling was more strongly associated with health status and obesity in White participants than Black participants, while having a grandparent with a college degree was more strongly associated with smoking in Blacks than Whites.

Given the complex and multifaceted nature of relations between schooling and health, these differences are not surprising. As demonstrated in the analyses here, demographic variation may also differ depending on the specific health and schooling measures, as well as by time period and geographical context. Not only will accounting for these differences allow us to describe demographic and schooling-related health disparities more accurately, but it may also provide insights into underlying mechanisms. For example, examining gender differences in the effects of schooling on self-rated health status may shed light on specific cognitive, psychosocial, or economic mechanisms that differ between men and women and therefore explain the difference in the effects of schooling.

An important implication of demographic differences in how schooling and health are related is that health inequities may be compounded in groups that are both socially and educationally disadvantaged. One hope is that addressing educational disadvantage may help reduce other sociodemographic health disparities. According to the resource substitution theory, schooling may be particularly important for disadvantaged groups who have few other resources available to them for fostering better health.[1] However, evidence supporting this theory is mixed, as demonstrated by the results for obesity and health status in Chapter 4 here. In situations in which disadvantaged demographic groups gain less benefit from schooling than others, blanket efforts to improve schooling may in fact compound health disparities between groups. Joint demographic and educational health disparities may therefore require solutions that explicitly address both and account for demographic variation in schooling–health relations.

Application of Alternative Analytical Methods

Comparing results from these analyses estimated with conventional regression methods and those estimated from alternative methods with the aim of better addressing the complex

nature of relations between schooling and health highlights both the potential utility and the potential complications associated with employing alternative methods. In Chapter 2, estimates using sibling fixed effects models were consistent with those from conventional regression. The similarity of the estimates serves to some extent as reassurance that the results from conventional regression may not have been driven by unmeasured confounding, a substantial concern for this topic. At the same time, the imprecision of the estimates from the fixed effects models reflect the tradeoff between bias and precision inherent to choosing between the two approaches.[2] It is also necessary to remember that, at least in this analysis, the estimates from the fixed effects models may have still been subject to confounding by unmeasured factors not necessarily shared between siblings, such as school characteristics. The fixed effect analysis, like the conventional regression analysis, may also have been affected by measurement error. Nonetheless, the results demonstrated the utility of fixed effects models as a tool for addressing unmeasured confounding, particularly in a context such as this one where unmeasured confounding is a substantial concern and the inclusion of sibling pairs in the data set permitted the estimation of these models.

Unlike fixed effects models, marginal structural models (MSMs) do not address unmeasured confounding. They may also, like conventional regression and fixed effects models, be affected by measurement error and model misspecification. In fact, model misspecification is of particular concern when using MSMs because the method requires specification of both the exposure model (used to estimate inverse probability weights) and the structural model (incorporating the weights and used to estimate the exposure–outcome association).[3] Rather, the usefulness of MSMs lies in their ability to account for time-dependent confounding. In Chapter 3, MSMs were used to address confounding of estimated effects of academic achievement on

health at the end of the follow-up period by interim health. In Chapter 4, MSMs were used to address confounding of estimated direct effects of grandparent schooling on grandchild adult health by factors such as grandparent income that are effects of grandparent schooling but confounders of associations between parent and grandchild schooling and grandchild health.

The estimates produced by MSMs in Chapters 3 and 4—particularly Chapter 3—were similar to those produced using conventional adjusted regression. This is not unusual in empirical analyses using MSMs. In a 2011 review, Suarez et al. found that out of 164 exposure–outcome associations (from 65 papers) in which both MSM and conventional regression estimates were reported, the MSM produced an estimate that was within 20% of the conventional regression estimate in 88 (54%).[4] In another example of the tradeoff between bias and precision, MSM estimates also tend to have larger standard errors, as was the case here. It is difficult to determine whether the similarity in estimates in Chapter 3 reflects a lack of time-dependent confounding, in which case the fact that two different estimation methods produced similar results may be reassuring, or if it reflects biases shared between the two types of models. For example, the different sources of the global health status measures at different study waves may have reduced the ability to detect the influence of interim health status on estimates in both types of models. In Chapter 4, MSM estimates of direct effects in Whites were systematically larger in magnitude than conventional regression estimates but the differences were not large. One possibility is that the intermediate variables we included, such as childhood socioeconomic status, function primarily as confounders of mediator–outcome associations and only secondarily as mediators of exposure–outcome associations. Therefore conventional regression estimates adjusting for these variables may not be severely biased. This interpretation is plausible if direct effects of grandparent schooling are primarily mediated by other mechanisms such as social

contact. A final consideration is that the degree of similarity between MSM and conventional regression models differed by outcome in both Chapter 3 and Chapter 4.

Novel analytical approaches can be powerful tools for addressing problems with conventional approaches for addressing specific questions. We demonstrated here their potential utility for addressing some of the difficulties in characterizing the causal nature of schooling–health associations. They may also be useful for examining other research questions in social epidemiology that present similar challenges. However, it is important to bear in mind not only the challenges addressed by alternative methods but also those not addressed, as well as new challenges that they may present.

Strengths and Limitations

The use of a contemporary, national data set with longitudinal information across the lifespan and over multiple generations makes the estimates here relevant for understanding these processes as they apply to the United States in current times. The diversity of socioeconomic and health measures in the PSID across time made it possible to account for many potential confounders in the associations estimated and will facilitate future examination of the mechanisms driving them. At the same time, in interpreting the findings here we should remain cognizant of their reliance on self- and proxy-reported information, the presence for some variables of measurement inconsistencies over time, and the possible role of attrition and other missing data.

The use of alternative analytic approaches to address specific challenges to estimating causal effects was also a strength of this project, as well as the direct comparison of results estimated using these approaches and using conventional regression methods. However, as stated above, these methods are imperfect and present their own challenges. Ultimately, this project

remains a study of observational data and the causal nature of the associations observed here should not be overstated. Rather, these findings should be interpreted as complementary to, and in the context of, the existing body of research addressing relations between schooling and health.

Future Directions

The analyses here highlight the complexity and longitudinal nature of processes relating schooling and health. They all support the interpretation that schooling and health are mutually beneficial. Chapter 2 suggests that poor health in childhood and adolescence is detrimental to educational progress, while Chapter 3 suggests that higher academic achievement is beneficial for future health. Chapter 4 suggests that the health benefits of schooling may cross generations.

The ultimate goal for this research is not only to further understanding of how schooling and health are related but also to help inform the development of policies and interventions that will improve population health and reduce health disparities. This will require a deeper understanding of the mechanisms driving the associations observed here. For example, while it is important to know that poor childhood health interferes with educational progress, effective intervention requires additional knowledge about which aspects of health are most important for educational progress and how they impede schooling. In particular, additional research is needed in young people, for whom educational interventions that will benefit health—as well as health interventions that will benefit education—are both most feasible and potentially most effective in reducing future health disparities.

A challenge to this research is that a comprehensive mechanistic understanding may require information at many levels ranging from the macrosocial to the physiological. For this reason, combining information of different types is imperative. For example, the PSID contains

information on a variety of potential mediators of effects of academic achievement on health that may lend themselves to intervention, including peer influences, psychological well-being, educational expectations and aspirations, substance use, and risk-taking behavior. Future research examining these mechanisms can provide valuable insights into how achievement relates to health. However, additional research is also needed on physiological mechanisms through which these factors "get under the skin" to affect health.[5] Focusing this research on young people who have not yet reached their terminal education levels or developed serious health problems is important for developing interventions but, as demonstrated here, poses its own set of challenges. Because terminal educational attainment is not a meaningful metric in this population, alternative measures of schooling are required. Similarly, age-appropriate measures of health and health risk are required. Because of the relative lack of clinical morbidity, research examining how physiological health relates to schooling in young people may require measures of subclinical risk factors that are less easily acquired.

Additional research is also needed on the role of educational policy in population health. In the face of ongoing policy debate, examining the health effects of emerging education reforms is necessary to fully understand their implications. For example, high school exit exams, which are being mandated in an increasing number of states, may increase the credentialing effects of a high school degree on health. This may intensify the role of educational differences in health disparities, especially if exam implementation affects a disproportionate number of economically disadvantaged students or those belonging to racial or ethnic minority groups because of geographical population distributions.

One current barrier to understanding how educational policy and schooling interventions may influence health is simply that evaluation efforts often do not include collection of health

data. Incorporating even simple health or health behavior data into policy and program evaluation efforts could provide valuable information about potential effects of these efforts beyond the academic achievement gains generally assessed. This information would permit more comprehensive evaluation of the benefits and cost-effectiveness of the programs. Beyond that, it might also help us identify which aspects of schooling are most valuable for promoting health. This is important because while reducing overall disparities in educational attainment is itself a worthy goal, and would likely lead to reductions in health disparities, it is also worth considering more targeted interventions to address those aspects of schooling that are most important for students' current and future health. Furthermore, these interventions need not necessarily be limited to young people or currently enrolled students.

Besides incorporation of health data into the evaluation of new educational policies and programs, more systematic and routine collection of health information in schools about students could be an important step in preventing early-life health disparities from resulting in later schooling and adult health disparities. Identification of at-risk students and coordination of health services when necessary may have large effects in reducing schooling disparities on a population level if, as suggested by the results in Chapter 2, poorer health can impede educational progress even in the absence of serious medical illness.[6]

In the book *Making Americans Healthier: Social and Economic Policy as Health Policy*, Schoeni et al. describe the "three Cs," three challenges to increased research and practice on the health effects of social policy: causality, determining the causal nature of associations between social factors and health; cost-effectiveness, determining whether a particular social policy is a cost-effective means of improving population health; and "Can we do it?", our ability to overcome the political, ethical, or other barriers to enacting policies.[7 p.11-13] Despite these

challenges, our hope is that the educational system may serve as a means of reducing health disparities and improving population health on a broad scale, and our understanding is that this effort must be specific and interdisciplinary, and will exist at the vanguard of research and analytical methodologies.

References

1. Ross CE, Mirowsky J. Gender and the health benefits of education. *The Sociological Quarterly* 2010;**51**:1-19.

2. Allison P. Fixed effects regression methods in SAS. *Proceedings of the Thirty-first Annual SAS Users Group International Conference*. Cary, NC: SAS Institute Inc., 2006.

3. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;**168**(6):656-64.

4. Suarez D, Borras R, Basagana X. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. *Epidemiology*;**22**(4):586-8.

5. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. *Ann N Y Acad Sci* 2006;**896**(1):3-15.

6. Basch CE. Healthier students are better learners: high-quality, strategically planned, and effectively coordinated school health programs must be a fundamental mission of schools to help close the achievement gap. *J Sch Health* 2011;**81**(10):650-62.

7. Schoeni RF, House JS, Kaplan GA, Pollack H. *Making Americans Healthier: Social and Economic Policy as Health Policy*. New York: Russell Sage Foundation, 2008.