An Analysis of Variation in the Health-Related Returns to a College Degree[1]

Shawn Bauldry

University of North Carolina at Chapel Hill

March 2, 2012

[1] This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis. This paper was supported NICHD (R01 HD061622-01, Shanahan PI).

Abstract

Numerous studies have documented a link between educational attainment and health. The health-related benefits of education are known to vary over the life-course, across cohorts, and across racial and ethnic groups. This study contributes to our understanding of variation in the beneficial effects of education on health in two ways. First, the study adopts an innovative approach to detecting meaningful variation in the effects of higher education on health among black and white, female and male subgroups. Second, the study connects the approach to detecting variation in the effects of higher education to competing theories of health inequalities: resource substitution and resource multiplication theory. This work draws on data from the National Longitudinal Study of Adolescent Health to analyze the effects of a college degree on two measures of health behaviors, smoking and BMI, and two measures of health outcomes, self-rated health and systolic blood pressure. The results indicate that there is little evidence of variation in the health-related returns to a college degree among black females and males in a recent cohort of young adults. Among white females and males there is evidence of meaningful variation in the benefits of a college degree that is consistent with the resource multiplication theory.

Introduction

A large body of research documents the many links between education and health (Kitagawa and Hauser 1973; Link and Phelan 1995; Mirowsky and Ross 2003; Preston and Elo 1995). Educational attainment is positively associated, for instance, with physical functioning and subjective health, and negatively associated with morbidity, physical impairment, and mortality (Cutler and Lleras-Muney 2008; Elo 2009; Luo and Waite 2005; Mirowsky and Ross 2003; Smith 2007). These associations arise, in part, from the beneficial effects of education on health behaviors. People with a higher degree are less likely to smoke or to be heavy drinkers, and more likely to have a healthy diet and exercise than people with a high school education (Cutler and Lleras-Muney 2010; Mirowsky and Ross 2003; Pampel, Krueger, and Denney 2010). Better educated people are also more likely to engage in health-related preventative care that can have long-lasting benefits (Cutler and Lleras-Muney 2010).

Recent studies indicate that the benefits of education for health and health behaviors vary systematically across different populations. There has been some debate about how the effects of education on health vary over the life-course and across cohorts (Goesling 2007; House, Lantz, and Herd 2005; Liu and Hummer 2008; Lynch 2003; Lynch 2006; Mirowsky and Ross 2008; Willson, Shuey, and Elder 2007; Zheng and Land 2012). It appears that the effects of education on health increase across cohorts and increase over the life course until a threshold is reached in old age, but more research is needed to make more definitive statements. In addition to variation over the life course and across cohorts, a number of studies find the health-related benefits of education differ across racial and ethnic groups. Research has found that blacks do not gain as much from their education as whites, particularly at higher levels of education (Farmer and Ferraro 2005; Liu and Hummer 2008; Shuey and Willson 2008; Williams and Collins 1995).

This paper makes two contributions to our understanding of the health-related returns to education and variation in the returns. First, the paper adopts an innovative approach to assessing the presence of variation in the benefits of a college degree for health and health behaviors (Morgan and Todd 2008). The analysis centers on attaining a four-year college degree for two reasons. In recent cohorts, a college degree represents a key educational threshold that differentiates the life chances of individuals (Fischer and Hout 2006). The expansion in higher education has led to people attaining degrees who would not have in the past, which suggests that there may be more variation in general in the effects of a college degree among recent cohorts than one would expect among lower levels of education.

The approach to assessing heterogeneity in the effects of a college degree is rooted in the counterfactual framework for analysis (Heckman 2005; Morgan and Winship 2007; Rubin 1974; Splawa-Neyman [1923] 1990) and relies on comparing estimates of the average treatment effect for the treated (ATT) with the average treatment effect for the controls (ATC). In this analysis, the ATT provides an estimate of how much a randomly chosen person with a college degree (the treatment) benefits from that degree with respect to a given health behavior or health outcome. Similarly, the ATC provides an estimate of how much a randomly chosen person without a college degree would benefit from that degree with respect to a given health behavior or health outcome. A difference in the ATT and ATC provides evidence of important dimensions of variation in the health-related returns to education (discussed in detail below).

Second, a comparison of the ATT and ATC also provides a test of the relative significance of two competing theories concerning the role of a college degree in alleviating or exacerbating health inequalities. The *resource substitution* theory holds that a college degree can serve to compensate for other health-related disadvantages (Mirowsky and Ross 2003; Ross and Mirowsky 2006; Ross and Mirowsky 2010). In contrast, the *resource multiplication* theory holds that a college degree simply magnifies other health-related advantages (Mirowsky and Ross 2003; Ross and Mirowsky 2006; Ross and Mirowsky 2010).[2] An estimated ATT that is greater than the ATC implies that people who have the most to gain from a college degree are the ones attaining degrees. This is typically termed “positive selection” and is consistent with the resource multiplication hypothesis. Alternatively, an estimated ATC that is greater than the ATT implies that people who stand to gain the most from a college degree are not the ones who actually attain a degree. This is typically termed “negative selection” and is consistent with the resource substitution hypothesis. As such, this paper provides new evidence for the relative importance of the resource substitution and multiplication hypotheses.

Data for this study come from Waves 1 and 4 of the National Longitudinal Study of Adolescent Health (Add Health), a nationally representative sample of young adults. Wave 4 was collected in 2007 and 2008 when respondents were in their late 20s and early 30s. This represents a recent cohort and an age when most respondents are old enough to have completed their education. The analysis examines two measures of health behaviors, body mass index (BMI)[3] and smoking, and two measures of health outcomes, self-rated health

[2] This is quite similar to the *diminishing returns* hypothesis outlined by Farmer and Ferraro (2005).
[3] BMI is, of course, not a behavior, but rather reflects behaviors related to diet and exercise.

and systolic blood pressure. Excess BMI and smoking are two of the leading behavioral causes of death in the U.S. (Cutler and Lleras-Muney 2010; Mokdad, Marks, Stroup, and Gerberding 2004; Pampel, Krueger, and Denney 2010). Self-rated health has proven to be a reliable leading indicator of morbidity and mortality (Ferraro and Farmer 1999; Idler and Benyamini 1997; Jylhä 2009; Jylhä, Volpato, and Guralnik 2006). Systolic blood pressure provides a more objective measure of health that is associated with cardiovascular disease. Although Add Health respondents are young to experience problems with systolic blood pressure, recent analyses of Add Health respondents have found elevated levels of blood pressure and a higher prevalence of hypertension for their age group than in past cohorts (Nguyen, Tabor, Entzel, Lau, Suchindran, Hussey, Halpern, Harris, and Whitsel 2011). Given the known differences in these health behaviors and outcomes and variation in the effects of education by sex and race/ethnicity, all analyses are conducted separately for black and white, men and women.

Mechanisms Linking Higher Education to Health and Health Behaviors

Health researchers have proposed and tested a number of mechanisms that give rise to the health-related benefits of a higher degree. The mechanisms can be broadly grouped into two categories: (1) economic resources and (2) social psychological and relational resources (Mirowsky and Ross 2003; Ross and Mirowsky 2010). Higher education, particularly in recent cohorts, plays a central role in providing access to high quality jobs with a decent income and a potentially creative and autonomous work environment. The economic resources that accrue from such positions provide a buffer from the stress associated with poverty and economic hardship that can lead to a variety of negative health behaviors and health outcomes (Lantz, House, Mero, and Williams 2005; Mirowsky and Ross 2003; Pampel, Krueger, and Denney 2010). In addition, economic resources may directly lead to better health behaviors and outcomes by facilitating healthy behaviors (e.g., membership to a gym) and allowing greater access to preventative medical care (Cutler and Lleras-Muney 2008; Freese and Lutfey 2011; Mirowsky and Ross 2003; Pampel, Krueger, and Denney 2010). A recent study estimates that economic resources, in particular income, access to health insurance, and resources from one's family background, account for about thirty percent of the relationship between education and health-related behaviors (Cutler and Lleras-Muney 2010).

A college degree may also promote better health behaviors and health outcomes through the development of social psychological and relational resources. Higher education has the potential to instill a sense of mastery and personal control that give people the motivation, knowledge, and ability to address their health-related needs (Cutler and Lleras-Muney 2008; Mirowsky and Ross 1998; Mirowsky and Ross 2003; Pampel, Krueger, and Denney 2010; Ross and Mirowsky 2010). While pursing a college education many people form relationships with other people in college and often maintain relationships with other highly educated people after completing a degree. This leads people with higher degrees to have social networks with other health-oriented people that can support healthy behaviors and develop group-based norms that protect against unhealthy behaviors (Freese and Lutfey 2011; Kawachi, Subramanian, and Kim 2008; Smith and Christakis 2008). For instance, several studies have identified network effects on smoking habits and obesity (Boardman, Onge, Rogers, and Denney 2005; Christakis and Fowler 2007; Christakis and Fowler 2008; Cutler and Glaeser 2007). More generally, social support can help alleviate stress and appears to reduce mortality, though the links between social support and physical health are more contentious (House, Landis, and Umberson 1988; Seeman 1996).

Variation in the Health-Related Returns to a College Degree

The dramatic increase in college enrollment over the course of the 20th century has led to changes in who attains college degrees and the nature of colleges and universities granting degrees (Snyder and Dillow 2011). In percentage terms, about a tenth of a percent of the population aged 15 to 24 were enrolled in college in 1900 as compared with roughly 35 percent of the population aged 15 to 24 in 2000. Some people who currently attain degrees would most likely not have in the past. In addition, the recent expansion in higher education has been concentrated among lower tier colleges and universities. Taken together, these changes raise the possibility that there may be more variation in the benefits of a college degree in recent cohorts than in past cohorts.

Changes in the composition of people attaining degrees and in the institutions granting degrees could affect the health-related returns to a college degree through both economic and social psychological and relational resources. For some people, the economic returns to a college degree may not be as high as for others, in which case a college degree may not provide as much protection from economic hardship and resources to afford a healthy lifestyle. In addition, a college degree may not provide the same sense of mastery and personal control, nor does it necessarily provide access to health-promoting social networks among recent cohorts. Thus, an increase in the variation of the benefits of a college degree among recent cohorts has an impact through a variety of mechanisms that link education to health-related behaviors and health outcomes.

Resource Substitution and Multiplication Theories

In addition to assessing the extent of variation in the health-related benefits of a college degree, the analytic approach used in this paper also allows for an assessment of the relative strength of two competing theories of the role of education in alleviating or exacerbating health inequalities: the *resource substitution* and the *resource multiplication* theories (Mirowsky and Ross 2003; Ross and Mirowsky 1989; Ross and Mirowsky 2010). There are potentially multiple resources that people can draw on to support their health. The distinction between resource substitution and resource multiplication theories lies in the relationship between higher education and the other types of resources (e.g., wealth from an advantaged background). If the various resources that can support health serve as substitutes for one another, than having one resource, such as a college degree, can compensate for a lack of other resources, such as wealth obtained from an advantaged background. This is the mechanism that underlies the resource substitution theory. Alternatively, if the various resources that can support health augment each other, then having one resource (e.g., a college degree) can magnify the benefits of other resources (e.g., an advantaged background). This is the mechanism that underlies the resource multiplication theory.

A comparison of the estimates of the health-related returns of a college degree for people who attained a degree (the ATT) with the health-related returns of a college degree for people who did not attain a degree (the ATC) not only provides evidence of variation in the average effect of a college degree but also allows for an assessment of resource substitution and multiplication theories. If the ATC is greater than the ATT, then this indicates that a college degree would provide greater health-related returns to people who did not attain a degree were they to do so than for people who did attain a degree. This pattern is consistent with the resource substitution theory because people who do not attain a college degree have fewer resources on average than people who do attain a college degree. On the other hand, if the ATT is greater than the ATC, then this indicates that a college degree provides greater health-related returns to people who attain a degree than would be realized by people who did not attain a degree. This pattern is consistent with the resource multiplication hypothesis because it suggests that a college degree amplifies pre-existing advantages (other resources).

Methodological Issues in Estimating the Health-Related Returns to a College Degree

The standard approach to analyzing the effect of a college degree on a particular health outcome or behavior is to estimate a regression model of the form

$$y_i = \alpha + \gamma D_i + \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i, \tag{1}$$

where y_i is a health outcome of interest,[4] D_i is an indicator for attaining a college degree, $\mathbf{x}_i$ is a vector of covariates that are associated with both the health outcome and the probability of attaining a college degree, and ε_i is a disturbance term. In this equation, if one assumes that there are no omitted variables (i.e., $\text{Cov}(D_i, \varepsilon_i) = 0$ and $\text{Cov}(\mathbf{x}_i, \varepsilon_i) = 0$), then $\hat{\gamma}$ provides an estimate of the average causal effect of attaining a college degree on a given health outcome.

[4] If the health outcome is not a continuous variable (e.g., an indicator for hypertension) or represents the time to an event (e.g., time to death), then (1) should be understood to refer to an appropriate regression model (e.g., a logistic regression model or a Cox regression model).

The standard approach has a straightforward interpretation in the counterfactual framework (Heckman 2005; Morgan and Winship 2007; Rubin 1974; Splawa-Neyman [1923] 1990). Attaining a college degree can be thought of as a "treatment" – i.e., something that in principle can be manipulated (Holland 1986). The health benefits that arise from receiving the treatment, attaining a college degree, may vary from person to person. It is not, in general, possible or desirable to estimate a separate causal effect from a treatment for all individuals. Instead, analysts are typically interested in an average causal effect defined over a given population or subpopulation (Heckman 2005). This can be reflected in the model in equation (1) by simply adding an index j to γ that references the population or subpopulation of interest.

Identifying Variation in the Health-Related Returns to a College Degree

People who receive the treatment and people who do not receive the treatment are two common subpopulations of interest in the counterfactual framework. In this analysis the treatment is the attainment of a college degree, so the two subpopulations are people who did and did not attain a degree. The average causal effect for people who attained a college degree, the "average treatment effect for the treated" (ATT), is an estimate of how much a randomly selected person who attained a college degree benefited from that degree with respect to their health (Heckman 2005; Morgan and Harding 2006). Similarly, the average causal effect for people who did not attain a college degree, the "average treatment effect for the controls" (ATC), is an estimate how much a randomly selected person who did not attain a college degree would benefit if she were to attain one (Heckman 2005; Morgan and Harding 2006). A comparison of the ATT and ATC provides a means of assessing whether there is consequential variation in the average health-related returns to a college degree.

There are two potential sources of heterogeneity in the effect of a college degree on health that could lead to different estimates for the ATT and ATC (Morgan and Todd 2008; Morgan and Winship 2007). First, a difference in the ATT and ATC could arise if the effect of a college degree depends on one of the other covariates in the model. In other words, there is an interaction effect between the treatment, attaining a college degree, and at least one of the other predictors of the health outcome. Second, a difference in the ATT and ATC could arise if the effect of a college degree depends on an unobserved variable that predicts attaining a college degree. Methodologists divide causal effect heterogeneity due to unobserved variables into two types: (1) heterogeneity due to an omitted variable and (2) heterogeneity due to the possibility that individuals anticipate the potential gains from the treatment (or control) and select into the treatment (or control) to realize these gains (Heckman 2005; Heckman, Urzua, and Vytlacil 2006; Morgan and Todd 2008; Morgan and Winship 2007). A comparison of the ATT and the ATC allows for a more complete assessment of potential variation in the protective effects of college degree and avoids the "curse of dimensionality" that comes with attempting to include all plausible interactions with observed variables.

Data and Methods

Data

This analysis draws on Waves 1 and 4 from the National Longitudinal Study of Adolescent Health (Add Health).[5] The first wave of Add Health was based on a nationally representative sample of youth in grades 7 through 12 in the United States in 1994. The fourth wave of data collection occurred between January 2008 and February 2009, roughly 14 years after the first wave. About 75 percent of the 20,745 youth interviewed at Wave 1 were re-interviewed at Wave 4, resulting in 15,701 respondents with data from both waves.

Past research has found that the effects of education on health behaviors and health outcomes differ by sex and race (Farmer and Ferraro 2005; Ross and Mirowsky 2006; Shuey and Willson 2008; Williams and Collins 1995). Because of this, the analysis sample was restricted to white and black, males and females (N = 11,529) and separate models were estimated for each of the four subgroups. In addition, Add Health provides sample weights to adjust for the differing probabilities of inclusion in the original sample frame and attrition between Waves 1 and 4. Respondents (N = 674) who did not have sample weights were excluded from the analysis sample. Finally, 2 respondents missing information about the highest educational degree completed were excluded from the analysis sample.[6] This results in an analysis sample of N = 10,853.

All of the remaining missing data were addressed using multiple imputation (Little and Rubin 2002). Ten complete data sets were constructed using the chained equation approach implemented in Stata 12 (StataCorp 2011). Most of the measures used in the

[5] See Harris, Halpern, Whitsel, Hussey, Tabor, Entzel, and Udry (2009) for a detailed description of the construction of the Add Health sample.
[6] This missing data could not be handled with multiple imputation because the estimation of the propensity score weights and subsequent models does not allow for potential differences in educational level across for the same case across imputed data sets.

analysis were missing for less than 4 percent of the cases. The exceptions were some of the measures related to parental SES and the respondent's birth weight. These variables were missing for 20 to 25 percent of the cases.[7] Multiple imputation appeared to perform reasonably well for these variables. With the exception of a small number of outliers (i.e., less than 10), the range of the imputed values was quite similar to the range of the non-imputed values for all of these variables. In addition, the mean and variance of the measures for the non-imputed cases were quite similar (i.e., within a couple hundredths of a point) to the mean and variance of the measures with the imputed cases included.

Variables

Health Measures. This analysis focuses on two measures of health and two measures of health behaviors. The first measure, self-rated health, is a general measure of health based on responses to the question "In general, how is your health?" that include "excellent" to "poor." Self-rated health has proven to be a reliable leading indicator of morbidity and mortality (Ferraro and Farmer 1999; Idler and Benyamini 1997; Jylhä 2009; Jylhä, Volpato, and Guralnik 2006). Self-rated health is treated as a dichotomous variable ("excellent" or "very good" versus "good," "fair," or "poor"), which is consistent with the majority of past research using this item.[8] The second measure is systolic blood pressure

---

[7] Auxiliary variables that could improve the handling of missing data were considered, but none met the required associations outlined in chapter 4 to have an impact on reducing bias.

[8] Supplemental analyses were also run treating self-rated health as a continuous measure and as an ordinal measure that maintained all of the categories and the results were essentially the same (available on request).

(SBP),[9] a more objective measure of health associated with cardiovascular disease. Recent analyses of Add Health respondents have found elevated levels of blood pressure and a higher prevalence of hypertension for their age group than in past cohorts (Nguyen et al. 2011).

The analysis focuses on two measures of health behaviors that are among the leading behavioral causes of death in the U.S. (Cutler and Lleras-Muney 2010; Mokdad, Marks, Stroup, and Gerberding 2004; Pampel, Krueger, and Denney 2010). The first, body mass index (BMI), is a commonly used measure that reflects behaviors related to diet and exercise. The second measure is smoking. Smoking is coded as an indicator that takes a value of 1 for respondents who reported smoking at least once a day over the past month.

College Education. In Wave 4 respondents were asked the highest educational level of education they had achieved to date. This information was used to construct an indicator for respondents who had completed a four-year degree. Of the 10,853 respondents in the analysis sample, 33 percent had attained a four-year degree.[10] Given the age range of respondents at Wave 4 (24 to 34 years old), this is likely a slight underestimate of the percentage who will ultimately attain a college degree.

---

[9] Certified field interviewers measured respondents' resting, seated systolic and diastolic blood pressures (mmHg) and pulse rate (beats/minute). Following a five-minute seated rest, three serial measurements were performed at 30-second intervals. SBP reflects the average of measures 2 and 3. When either the second or third measure was missing, the other single measure was used. In cases where both measures 2 and 3 were missing, the first measure was used.

[10] This is the unweighted percentage. The weighted percentage is 31 percent.

Table 1: Descriptive Statistics for Health Behaviors and Health Measures; N = 10,853.

	Black Female N = 1,681		Black Male N = 1,295		White Female N = 4,135		White Male N = 3,742	
	No Deg Mean	Degree Mean	No Deg Mean	Degree Mean	No Deg Mean	Degree Mean	No Deg Mean	Degree Mean
Self-rated health	0.42	0.62	0.49	0.64	0.51	0.78	0.54	0.76
Sys. blood pres.	124.07	120.33	130.01	130.14	120.21	118.78	130.19	129.19
BMI	33.06	31.00	28.94	29.85	29.66	26.41	29.29	27.65
Daily smoker	0.14	0.04	0.26	0.06	0.37	0.10	0.38	0.09

Notes: Based on 10 complete data sets. Weighted data.

Table 1 provides weighted descriptive statistics for each of the measures of health and health behaviors for respondents with and without college degrees across the subgroups defined by sex and race/ethnicity. There are substantial differences in the proportions of people reporting very good or excellent self-rated health and smoking daily between those with college degrees and those without across the groups. Among black and white, females and males, respondents with degrees were more likely to report being in very good or excellent health than respondents without degrees. The proportion of daily smokers is higher among respondents without a college degree across all of the groups. The differences in BMI and systolic blood pressure between people with and without a college degree are more muted, but still present for black females and white males and females. Among black males, those with a college degree have slightly higher systolic blood pressure and BMI than those without a college degree, though the difference is minimal.

Additional Covariates. It is important to adjust for a number of covariates when estimating the effect of education on health and health behaviors.[11] The covariates can be

[11] Many of these covariates are likely to be subject to measurement error. The primary goal in using the covariates is to construct weights based on propensity scores of attaining a

generally divided into three groups: (1) measures of family socioeconomic status, (2) measures of ability, and (3) measures of childhood and adolescent health and health behaviors. All three groups of covariates have well-established associations with educational attainment and with health and health behaviors. In addition, all of the models include region (West, Midwest, South, Northeast) and the age of the respondents at Wave 4. Descriptive statistics for all of the covariates are available in Appendix 1.

One of the benefits of using Add Health data is that it includes a number of measures of family socioeconomic status gathered at Wave 1 when respondents were in high school. This analysis draws on mother's and father's education (ten category measures ranging from "8th grade or less" to "professional training beyond a 4-year degree"), logged family income, family structure, and an interviewer assessment of respondent's living environment. Family structure is a five category variable that differentiates respondents living with two biological parents, two parents with at least one non-biological, a single mother, a single father, or some other arrangement (Harris 1999). The measure of the respondent's living environment is constructed as the average of the responses to the following two questions completed by Add Health interviewers: (1) "How well kept is the building in which the respondent lives?" and (2) "How well kept are most of the buildings on the street?" Responses to both questions ranged from "very poorly kept (needs major repairs)" (1) to "very well kept" (4).

---

college. As such, interest does not center on interpreting the structural effects of the covariates on either attaining a college degree or the given health outcomes, which would be affected by measurement error. Measurement error in the covariates may still have an effect in the models for the health outcomes, though the effect is not likely to be large given that the weights by design render the covariates largely insignificant predictors of the outcomes.

Cognitive ability is often proposed as a source of spuriousness in analyses of the effects of education on health (Gottfredson 2004; Hirschi and Gottfredson 1994). Recent work suggests that it is unlikely that cognitive ability unrelated to education can account for the beneficial effects of education on health (Link, Phelan, Miech, and Westin 2008). Nonetheless, this analysis draws on two measures of ability to address this possibility. At Wave 1 Add Health administered an abbreviated version of the Peabody Picture Vocabulary Test (PPVT) that serves as a measure of verbal ability or scholastic aptitude (Dunn and Dunn 1981). In addition, at Wave 1 respondents reported their grades in four subjects from the most recent semester (English, Math, History or Social Science, and Science). Self-reported GPA was constructed as the average of the grades from the available subjects.

The third group of covariates includes measures of health and health behaviors from childhood and adolescence. At Wave 1 parents reported the birth weight of respondents. Birth weight has been associated with educational attainment and a range of adult health outcomes (Almond and Currie 2011; Dahly, Adair, and Bollen 2009). Add Health did not include a clinical measure of systolic blood pressure at Wave 1, but it did include self-rated health, self-reported measures of height and weight that can be used to construct BMI, and the number of cigarettes smoked in the last month. In addition, the analysis adjusts for a measure of adolescent physical activity and measures of adolescent drinking. The measure of physical activity is taken as the maximum number of days per week respondents report engaging in any of a series of activities.[12] The measure of drinking

[12] Wave 1 activities included: (a) roller-blading, roller-skating, skate-boarding, bicycling, (b) active sports such as baseball, softball, basketball, soccer, swimming, or football, (c) exercise such as jogging, walking, karate, jumping rope, gymnastics, or dancing.

stems from the questions “During the past 12 months, on how many days did you drink alcohol?” with responses ranging from “never” (1) to “every day or almost every day” (6).

Analysis Strategy

Following the approach outlined by Morgan and Todd (2008), the analysis to detect variation in the effects of education on health behaviors and outcomes proceeds in four steps for black and white, males and females. The first step in the analysis is to estimate a propensity score model. The second step is to use the propensity score models to construct weights that allow for estimates of different treatment effects. The third step is to estimate the average treatment effect (ATE), the average treatment effect of the treated (ATT), and the average treatment effect for the controls (ATC). The final step is to compare the estimates of the ATE, ATT, and ATC and make a determination whether they are all roughly the same or whether the ATT differs from the ATC. This section provides a brief description of each step.

Propensity score models are used to estimate the conditional probability of a “treatment” (Guo and Fraser 2010; Rosenbaum and Rubin 1983). For this analysis the “treatment” is attaining a college degree. The probability of attaining a college degree conditional on the covariates is estimated using a logistic regression model specified as

$$p_i = \Pr[D_i = 1|\mathbf{x}_i] = \frac{\exp(\mathbf{x}_i\phi)}{1+\exp(\mathbf{x}_i\phi)}, \quad (2)$$

where p_i is the propensity score for person i, D_i in an indicator for attaining a college degree, and $\mathbf{x}_i$ is a vector of covariates with coefficients $\boldsymbol{\phi}$. The covariates in the propensity score model should include all of the relevant predictors of the measures of health behaviors and outcomes (Brookhart, Schneeweiss, Rothman, Glynn, Avorn, and Stürmer

2006; Guo and Fraser 2010; Rubin 1997). Square terms for mother's education, father's education, and logged family income were included in the propensity score model to capture their potential non-linear relationships with attaining a college degree and with health and health behaviors. The propensity score models were weighted by the sample weights provided by Add Health and estimated separately on each of the 10 imputed data sets.

The three groups of covariates, family SES, ability, and childhood and adolescent health and health behaviors, account for the main predictors of adult health that could be confounded with educational attainment. It is, however, likely that at least some relevant predictors have been omitted from the model. A sensitivity analysis is included to assess this concern.

The second step in the analysis involves calculating appropriate weights to obtain the various treatment effect estimators. The weight for the average treatment effect (ATE) is calculated as

$$w_{i,ATE} = \begin{cases} \dfrac{1}{\hat{p}_i} & \text{if } D_i = 1 \\ \dfrac{1}{1-\hat{p}_i} & \text{if } D_i = 0 \end{cases} . \tag{3}$$

Weighting by the inverse of the propensity score for the people who attained college degree and the inverse of one minus the propensity score for people who did not attain a college degree attempts to balance the sample with respect to the distribution of covariates between people who attained a college degree and people who did not. In other words, after weighting the data by the propensity scores, the distributions of the covariates among the people who attained a college degree should be similar to the distribution of the covariates among people who did not attain a college degree. In theory, this should hold for all moments of the distributions of the covariates, but in practice just the mean and the standard deviation are typically assessed.

One common metric to assess the balance of the means is the average of the standardized mean differences across the treatment and control groups (Morgan and Todd 2008; Rubin 1973). The standardized mean difference (StdMD) for a given variable is calculated as

$$\text{StdMD} = \frac{|\,\bar{x}_{i,D_i=1} - \bar{x}_{i,D_i=0}\,|}{\sqrt{\frac{1}{2}Var(x_{i,D_i=1}) + \frac{1}{2}Var(x_{i,D_i=0})}}\,. \tag{4}$$

Similarly, a common metric to assess the balance of the standard deviations is the average of the standardized standard deviation differences across the treatment and control groups. The standardized standard deviation difference (StdSD) for a given variable is calculated as

$$\text{StdSD} = \frac{|\,sd(x_{D_i=1}) - sd(x_{D_i=0})\,|}{\sqrt{\frac{1}{2}Var(x_{i,D_i=1}) + \frac{1}{2}Var(x_{i,D_i=0})}}\,. \tag{5}$$

The StdSD is typically only used to assess balance among continuous covariates (Morgan and Todd 2008). In practice, perfect balance is never achieved and instead researchers assess how close the StdMD and StdSD are to 0 and how much balance is improved by the propensity score weight (Rubin 2006).

The weight for the average treatment effect for the treated (ATT) is calculated as

$$w_{i,ATT} = \begin{cases} 1 & \text{if } D_i = 1 \\ \dfrac{\hat{p}_i}{1-\hat{p}_i} & \text{if } D_i = 0 \end{cases}\,. \tag{6}$$

This weight treats the population-level group of people who attained a college degree as the target population by leaving the members of the sample who attained a college degree unweighted and weighting the members of the sample who did not attain a college degree according to their odds of having done so. This is an attempt to transform the sample of people who did not attain a degree into a representative sample of the population-level group of people who did attain a college degree (Morgan and Todd 2008). As with the ATE, the extent to which the weight balances the distribution of the covariates can be assessed.

Finally, the weight for the average treatment effect for the controls (ATC) is calculated as

$$w_{i,ATC} = \begin{cases} \dfrac{1-\hat{p}_i}{\hat{p}_i} & \text{if } D_i = 1 \\ 1 & \text{if } D_i = 0 \end{cases} . \tag{7}$$

Similar to the ATT, the purpose of this weight is to transform the sample into a representative sample of people who did not attain a college degree.

The propensity score weights have the same properties as other types of sample weights and can thus be combined in the standard way that different sample weights are combined (Kalton and Flores-Cervantes 2003). All of the weights outlined above were multiplied by the Add Health sample weights to incorporate the adjustments for unequal probability of sample selection and attrition into the estimates of the various treatment effects.

The third step is to estimate models for each of the measures of health and health behaviors to obtain the ATE, ATT, and ATC for attaining a college a degree for each of the subgroups defined by sex and race. Depending on the outcome, the models are either weighted linear regression models or weighted logistic regression models. The models take the general form

$$y_i = \alpha + \gamma D_i + \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i, \tag{8}$$

where y_i is either a health measure or the logit of a health measure, D_i is an indicator for attaining a four-year college degree, γ gives the effect of a college degree with the precise interpretation depending on the weights, and $\mathbf{x}_i$ is a vector of covariates (the same covariates used in the propensity score model), and ε_i is a disturbance term. In theory the propensity score weights should balance the data such that it is unnecessary to adjust the covariates that went into the creation of the propensity scores when estimating the treatment effects. If, however, the propensity score model is misspecified, including the covariates in the regression models can help mitigate the effects of this misspecification (Bang and Robins 2005; Morgan and Winship 2007; Robins and Rotnitzky 1997). The estimated standard errors are robust standard errors that adjust for heteroskedasticity and the clustering of respondents in schools in Wave 1. The standard errors do not specifically adjust for the additional uncertainty due to the estimation of the propensity scores; however, the robust standard errors should at least adjust for some of this uncertainty. Finally, the models are estimated on each of the imputed data sets and the results are combined according to Rubin's formulas using the multiple imputation suite of commands in Stata 12 (Little and Rubin 2002; StataCorp 2011).

The final step in the analysis is to assess whether the estimates of the ATT and ATC are different. Unfortunately, there is not a simple statistical test for the difference. It is, of course, possible to calculate the difference between the estimates, but it is not clear how to calculate an accurate standard error for the difference (Morgan and Todd 2008). Instead, this analysis relies on a substantive assessment of the magnitude of the differences and the consistency of the results across the different measures among black and white, females and males.

Results

Propensity Scores

The first step in the analysis is to estimate propensity scores for attaining a college degree for black and white, males and females. The parameter estimates for the propensity score models are provided in Appendix 2. Figure 1 illustrates the distribution of propensity scores for people who attained and did not attain a college degree for each subgroup based on the first complete data set.[13] Among black females and males, the propensity for attaining a college degree is relatively flat for those who did attain a degree and shows a sharp drop around 0.1 for those who did not attain a degree. White females and males who did not attain a degree have a similar propensity that peaks around 0.1 and then sharply drops afterward. In contrast, white females who did attain a degree have a steadily increasing propensity until around 0.9. There is a similar pattern for males, but the increase is not as pronounced.

Although the distributions of propensity scores are quite different for people who did and did not attain a degree, the distributions largely span the same range of propensities (see Table 2). The only particularly sparse region is among the highest decile of propensity scores. There are only a few people who did not attain a college degree,

---

[13] The distributions were examined in all 10 of the complete data sets and were found to be quite similar.

particularly among black males, with estimated propensity scores greater than or equal 0.9.[14]

Figure 1: Distribution of Estimated Propensity Scores.

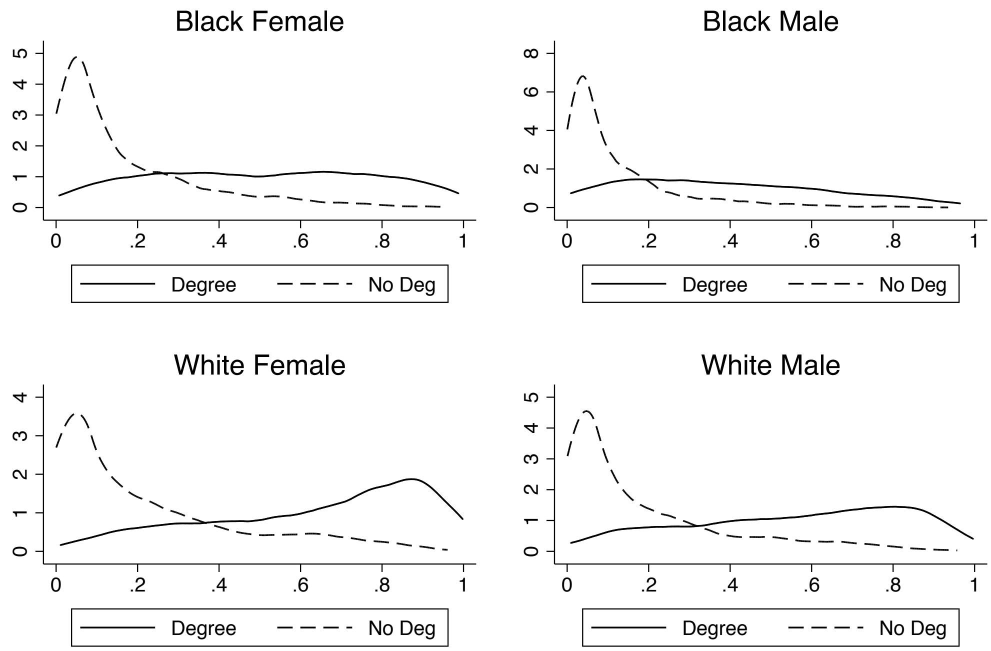

Table 2: Number of Cases in Deciles of Propensity Scores.

	Black Female		Black Male		White Female		White Male	
Propensity	No Deg	Degree	No Deg	Degree	No Deg	Degree	No Deg	Degree
[0.0 - 0.1)	500.4	32.2	582.7	40.8	1083.1	46.3	1237.9	59.5
[0.1 - 0.2)	225.8	42.8	189.8	35.2	467.8	77.2	471.6	84.5
[0.2 - 0.3)	134.0	54.5	94.0	36.7	300.4	114.2	295.8	98.8
[0.3 - 0.4)	86.4	50.3	56.0	27.2	200.7	122.8	184.5	110.9
[0.4 - 0.5)	74.4	56.9	35.1	28.1	142.7	135.2	124.8	114.8
[0.5 - 0.6)	52.7	56.1	23.5	26.3	114.3	144.4	107.5	126.9
[0.6 - 0.7)	32.4	66.8	14.4	22.5	104.6	175.8	81.4	146.5
[0.7 - 0.8)	23.6	55.5	10.1	25.7	70.5	217.5	52.6	155.8
[0.8 - 0.9)	10.0	66.2	8.7	21.5	47.7	289.0	27.2	154.5
[0.9 - 1.0]	2.3	57.7	1.7	15.0	10.2	270.6	7.7	98.8

Notes: Unweighted average counts across 10 complete data sets.

[14] Even in the sparse regions, none of the complete data sets had 0 people in any cell.

Given that all deciles have at least one case and most have many more cases, the main analysis proceeds using all of the cases representing the entire distribution of propensity scores. A sensitivity analysis in the following section assesses the robustness of the results with the sample limited to cases with greater overlap in the distribution of propensity scores.

The second step in the analysis is to construct the propensity score weights and assess the degree to which the weights balance the covariates. Table 3 provides estimates for the average standardized mean difference and average standardized standard deviation difference for the covariates using only the sample weights and the covariates using each of the propensity score weights. The distributions of the covariates differ substantially across people who did and did not attain college degrees for all of the groups as evidenced by the average difference in the standardized means. For instance, among black females, the average difference in the standardized means across all of the covariates is 0.356. Each of the propensity score weights substantially improves the balance in the covariates across the treatment and control groups with respect to the mean in all of the groups. For black females, the average difference in the standardized means across all of the covariates drops from 0.356 to 0.063 when applying the ATE weight, 0.039 when applying the ATT weight, and 0.091 when applying the ATC weight. The propensity score weights do less to improve balance with respect to the standard deviations of the covariates, particularly among black males where the ATE and ATC weights actually increase the average difference in the standardized standard deviation from 0.070 to 0.086. This suggests that the ATE and ATC weights may not provide consistent estimates of the ATE and ATC for black males and caution is warranted in interpreting these results.

Table 3: Assessment of Mean and Standard Deviation Balance.

	Black Female		Black Male	
	Avg, StdMD	Avg, StdSD	Avg, StdMD	Avg, StdSD
Sample weight	0.356	0.084	0.337	0.070
ATE	0.063	0.060	0.084	0.086
ATT	0.039	0.038	0.037	0.035
ATC	0.091	0.068	0.101	0.103
	White Female		White Male	
	StdMD	StdSD	StdMD	StdSD
Sample weight	0.423	0.116	0.404	0.106
ATE	0.055	0.051	0.059	0.053
ATT	0.041	0.040	0.015	0.022
ATC	0.071	0.080	0.083	0.061

Notes: Balance metrics averaged over 10 complete data sets. ATE, ATT, and ATC incorporate sample weights.

Table 4 provides the estimates for the ATE, ATT and ATC along with a standard regression estimate (SRE) as a point of comparison for the two health outcomes and two measures of health behaviors.

Black Females. Attaining a college degree is significantly associated with an increased likelihood of reporting very good or excellent self-rated health, a substantial reduction in systolic blood pressure, and a decreased likelihood of being a daily smoker. The estimates from the standard regression models are quite similar to the estimates of the ATE using the propensity score weights. The largest difference between the standard regression estimate and the ATE is with effect of a college degree on smoking (SRE = -1.339, ATE = -1.875). The ATE is larger than the standard regression estimate, but it is still within two standard errors.

Among black females the ATC is greater (in absolute value) than the ATT for both self-rated health (ATT = 0.557, ATC = 0.751) and for smoking (ATT = -1.179, ATC = -2.223), while the ATT is greater than the ATC for BMI (ATT = -1.414, ATC = -0.813) and systolic blood pressure (ATT = -3.424, ATC = -3.114). The differences between the ATT and ATC for BMI and systolic blood pressure are not substantively large and the confidence intervals around the estimates allow for a high degree of overlap. The differences between the ATT and ATC for self-rated health and smoking are substantively large. In terms of the odds ratios (OR), for self-rated health the ATT OR is 1.75 compared to the ATC OR of 2.12, while for smoking the ATT OR is 0.31 compared to the ATC OR of 0.11. In addition, there is much less overlap in the confidence intervals of the estimates, particularly for smoking, than with systolic blood pressure. These results suggest that among black females there is evidence of variation in the health-related returns to a college degree and that returns would be greater for people who do not attain a degree were they to attain one than for those who attain a degree.

Black Males. Attaining a college degree is significantly associated with an increased likelihood of reporting very good or excellent health (SRE = 0.689) and a decreased likelihood of being a daily smoker (SRE = -1.510) based on the SRE. The ATE for smoking (ATE = -1.436) is similar to the SRE, but the ATE for self-rated health is not statistically significant. In addition, the ATT is greater (in absolute value) than the ATC for self-rated health and smoking. Given the issues with covariate balance among black males for the ATE and ATC as well as the large degree of overlap in the confidence intervals of the ATE and ATC for smoking, there is insufficient evidence to conclude that the health-related returns to a college degree differ for black males who do and do not attain a college degree.

Table 4: Estimates of Effects of a College Degree on Health Behaviors and Health Measures.

	SRE			ATE			ATT			ATC		
	Est	SE		Est	SE		Est	SE		Est	SE	
	Black Female (N = 1,681)											
Self-rated health	0.687	(0.187)	***	0.683	(0.222)	**	0.557	(0.205)	**	0.751	(0.257)	**
Systolic blood pressure	-3.291	(1.192)	**	-3.191	(1.549)	*	-3.424	(1.274)	**	-3.114	(1.796)	
BMI	-0.835	(0.603)		-0.995	(0.724)		-1.414	(0.644)	*	-0.813	(0.820)	
Daily smoker	-1.339	(0.332)	***	-1.875	(0.377)	***	-1.179	(0.343)	***	-2.223	(0.494)	***
	Black Male (N = 1,295)											
Self-rated health	0.689	(0.251)	**	0.445	(0.302)		0.938	(0.259)	***	0.315	(0.336)	
Systolic blood pressure	0.242	(1.502)		0.492	(1.925)		0.420	(1.364)		0.404	(2.119)	
BMI	-0.101	(0.614)		-0.163	(0.623)		-0.491	(0.660)		0.007	(0.666)	
Daily smoker	-1.510	(0.375)	***	-1.436	(0.374)	***	-1.503	(0.401)	***	-1.484	(0.404)	***
	White Female (N = 4,135)											
Self-rated health	0.664	(0.118)	***	0.600	(0.161)	***	0.711	(0.149)	***	0.547	(0.198)	**
Systolic blood pressure	-0.139	(0.637)		-0.123	(0.758)		-0.004	(0.863)		-0.126	(0.902)	
BMI	-1.283	(0.303)	***	-1.114	(0.333)	***	-1.084	(0.365)	**	-1.092	(0.383)	**
Daily smoker	-1.084	(0.365)	**	-1.092	(0.383)	**	-1.162	(0.151)	***	-1.075	(0.199)	***
	White Male (N = 3,742)											
Self-rated health	0.487	(0.122)	***	0.381	(0.162)	*	0.493	(0.150)	***	0.325	(0.192)	
Systolic blood pressure	-0.704	(0.634)		-0.256	(0.790)		-0.025	(0.755)		-0.357	(0.989)	
BMI	-0.896	(0.240)	***	-0.654	(0.272)	*	-0.883	(0.293)	**	-0.548	(0.313)	
Daily smoker	-1.446	(0.166)	***	-1.252	(0.192)	***	-1.557	(0.180)	***	-1.153	(0.223)	***

Notes: SRE refers to the "standard regression estimate" based on models using just the sample weights. Logit models used for self-rated health and daily smoker (reported estimates are log odds). Linear regression models used for systolic blood pressure and BMI. All results based on 10 complete data sets. Heteroskedasticity and cluster robust standard errors are

White Females. Attaining a college degree is significantly associated with an increased likelihood of very good or excellent self-rated health (SRE = 0.664, ATE = 0.600), lower BMI (SRE = -1.283, ATE = -1.114), and a decreased likelihood of smoking daily (SRE = -1.084, ATE = -1.092). For both BMI and the likelihood of being a daily smoker, the ATT and ATC estimates are roughly the same. For self-rated health, however, the ATT is greater than the ATC (ATT = 0.711, ATC = 0.547). The magnitude of the difference is substantively significant (ATT OR = 2.03 compared with ATC OR = 1.73), which suggests that for self-rated health, white females who attain a college degree benefit more from than degree than one would expect from white female who did not attain a degree were they to do so. This is consistent with the resource multiplication hypothesis.

White Males. There is a similar pattern of significant results among white males as was observed among white females. Attaining a college degree has a significant association with an increased likelihood of reporting very good or excellent self-rated health (SRE = 0.487, ATE = 0.381), reduced BMI (SRE = -0.896, ATE = -0.654), and a reduced likelihood of being a daily smoker (SRE = -1.446, ATE = -1.252). The estimates for the ATEs are generally smaller than SREs, but all of the estimates are within a standard error of each other. Among white males, the ATT estimates are consistently greater (in absolute value) than the ATC estimates (self-rated health: ATT = 0.493, ATC = 0.325; BMI: ATT = -0.883, ATC = -0.548; smoke: ATT = -1.557, ATC = -1.153). With the exception of smoking (ATT OR = 0.21, ATC OR = 0.32), the differences between the ATT and ATC are not particularly large; however, the pattern is consistent and the confidence intervals have relatively little overlap. This suggests that across several health outcomes, white males who attain a college degree accrue more health-related benefits than would white males who do not attain a college degree were they to do so.

Overall. Taken as a whole, the results suggest that among whites there is evidence that variation in the health-related returns to a college degree is greater for people who attain a college degree than for people who do not attain a college degree were they to do so. This pattern of results is consistent with the resource multiplication hypothesis that whites are able to magnify other advantages related to health by attaining a college degree.

For blacks there is generally less evidence of variation in the health-related returns to a college degree. The one case where there appears to be some variation is in the effect of a college degree on smoking among black women where the ATC is greater than the ATT. This is consistent with resource substitution hypothesis, which suggests that attaining a college degree could be an important resource among black women to compensate for other health-related disadvantages. It is important to recall, however, that, particularly for blacks, the estimated propensity scores resulted in some sparse regions (i.e., there were few cases of people who did not attain a college degree but had a high propensity to do so) and the metrics assessing whether balance in the covariates was achieved were not ideal. The next section reports a series of sensitivity analyses to help determine the robustness of these results.

Sensitivity Analysis

The first sensitivity analysis involves only retaining people who have estimated propensity scores of attaining a college degree less than 0.8 for the analysis. Limiting the analysis sample to these people addresses the small number of cases with high estimated propensity scores who did not attain a college degree (particularly among blacks). Although this limits the generalizability of the results for people with high propensities for attaining a college degree, it provides a more robust basis for the estimated treatment effects for the remaining cases (roughly 90 percent of the analysis sample). This analysis revealed the same pattern of results for black females, black males, and white males (results available upon request). For white females, the pattern of results is the same except for BMI. The ATT for BMI is estimated to be greater than the ATC (ATT = -1.177, ATC = -1.041), whereas in the main analysis the opposite pattern is observed. In both cases, however, the ATT and ATC are not substantively different and the same conclusion as in the main analysis is warranted. Limiting the sample to cases with propensity scores less than 0.8 does not alter the conclusions and suggests that the results do not hinge on the small number of cases with high propensity scores who did not attain a college degree.

As indicated in Table 3, there is room for improvement in the extent to which the estimated propensity score weights balance the data. Add Health contains a number of additional potential predictors of attaining a college degree that could be used in the propensity score models. Although in theory one should use the same predictors in the propensity score model as in the outcome model (Brookhart et al. 2006; Rubin 1997), it is possible that additional covariates if even weakly associated with the health outcomes could help reduce bias. The second sensitivity analysis reestimates the propensity scores using additional covariates.

The additional covariates used in the reestimation of the propensity scores included a number of measures of school behaviors from Wave 1 – number of times skipped school, number of grades repeated, ever received a suspension, and a scale consisting four measures related to creating problems in the classroom. The propensity score models also included college expectations, college aspirations, and attachment to school at Wave 1. These variables are all known to predict educational attainment, though they are likely to have a weak association, at best, with health outcomes (especially conditional on the covariates already in the models).

Rerunning all of the analyses using the weights constructed from the updated propensity scores resulted in modest improvements in the balance metrics and the same substantive pattern of results for black females, white females, and white males (results available upon request). For black males, the balance metrics improved slightly, but in contrast to the main analysis, the ATC for daily smoking is estimated to be greater than the ATT for daily smoking in absolute value (ATT = -1.480, ATC = -2.224). The confidence intervals for the two estimates still show a significant degree of overlap and the ATC weights still do a relatively poor job of balancing the covariates. Therefore, the conclusion from the main analysis of no evidence of causal effect heterogeneity remains.

Overall, the two sensitivity analyses, restricting the sample to people with propensity scores less than 0.8 and adding additional covariates to the propensity score model, confirmed the results from the main analysis. It is, of course, still possible that the propensity score models suffer from misspecification or omitted variables, but the results of the sensitivity analyses suggest that these potential problems are not likely to be severe.

Discussion

This paper reports the results of an innovative approach to assessing variation in the health-related returns to a college degree. Past studies have found that the benefits of a college degree for health are unevenly distributed across racial and ethnic groups and sexes, over the life-course, and across cohorts. This analysis contributes to our understanding of variation in the benefits of a college degree through an assessment of whether there is additional heterogeneity in the effects of higher education within a recent cohort of young adults among black and white, males and females. In addition to identifying variation in the health-related benefits of a college degree, the analytic strategy also allowed for an assessment of two competing theories of whether education alleviates or exacerbates health inequalities.

The results suggest that there is variation in the health-related returns to a college degree among white females and males. In particular, the benefits of higher education for health are greater for people who attained a degree than for people who did not attain a degree were they to do so. This is consistent with the resource multiplication theory and suggests that a college education contributes to widening health inequalities among advantaged and disadvantaged whites. Past research examining health inequalities has focused on disparities across rather than within racial and ethnic groups. These results indicate that it is also important to consider within group inequality.

The results suggest that there is also variation in the health-related returns to a college degree for black females. Higher education is associated with better self-rated health, lower systolic blood pressure, and a reduced likelihood of smoking among black women. Furthermore, there is evidence of variation in the returns to a college degree for self-rated health and smoking that is consistent with resource substitution theory. Among black women, the estimated health-related benefits of higher education are greater for people who did not attain a degree were they to do so than those who did attain a degree.

Finally, the results for black males are inconclusive. The diagnostics from the construction of the propensity score weights indicated that it is difficult to construct weights that balance the covariates for those who did and did not attain a degree. In particular, it is hard to identify black males who did not attain a college degree but had a high propensity to do so. This is most likely an indication that there are unobserved factors that differentiate black males who do and not attain college degrees. These results indicate not only that it is difficult to identify variation in the health-relate returns to higher education among black males, but that even the average estimate of the benefits of higher education is likely to be biased.

There are some limitations to this study worth keeping in mind that could be addressed in future research. First, the assessment of differences in the ATT and ATC and consequent determination of meaningful variation in the effects of a college degree relies on a substantive rather than statistical determination. Some of the differences in the ATT and ATC were deemed to be substantively meaningful, but it is unlikely they would be statistically significant were an appropriate test available. This suggests that more work is needed to see if these patterns replicate in other data sets and across other measures of health behaviors and outcomes.

Second, although the Add Health respondents show signs of future health problems, particularly related to obesity and hypertension, they are relatively young to experience significant health declines. Past research suggests that the health-related returns to education increase over the life course until a threshold is reached in old age, so it is possible that as the beneficial effects of education strengthen variation in the effects will also increase. If so, it will be easier to detect evidence of variation in the health-related returns to education and, in particular, higher education. Future work should look to examine this possibility among older age groups.

Third, this study focused on a small number of health outcomes and health-related behaviors. The measures of health-related behaviors, BMI and smoking, are known to be among the leading behaviors causes of death in the US. One of the health outcomes, self-rated health, is known to be a reliable leading indicator of morbidity and mortality, while the other health outcome, systolic blood pressure, is closely related to cardiovascular disease. It would be useful, however, to extend the analysis to other health-related behaviors, particularly measures of diet and exercise, as well as other health outcomes, such as allostatic load. A consideration of these measures could help establish the robustness of the patterns observed in this analysis, particularly if assessed among older adults.

Appendix 1

Appendix Table 1: Descriptive Statistics for Predictors of Education and Health Outcomes; N = 10,853.

	Mean
w4age	28.50
west	0.14
midwest	0.31
south	0.42
northeast	0.13
2 bio parents	0.51
2 parents	0.19
single mother	0.22
single father	0.03
other	0.05
Mother's edu.	5.68
Father's edu	5.58
Log income	3.57
Lived envir.	3.37
PVT score	102.18
GPA	2.80
birth weight	117.13
W1 SRH	3.90
W1 BMI	22.57
W1 daily smoke	0.01
W1 physical act.	2.08
W1 drinking	2.09

Note: Based on 10 complete data sets. Weighted data.

Appendix 2

Appendix Table 2: Propensity Score Models.

	Black Female			Black Male			White Female			White Males		
	Est	SE		Est	SE		Est	SE		Est	SE	
w4age	-0.074	(0.054)		-0.037	(0.068)		0.086	(0.032)	**	0.109	(0.035)	**
midwest	-0.364	(0.361)		0.383	(0.455)		0.256	(0.174)		0.598	(0.174)	***
south	0.055	(0.304)		0.107	(0.396)		0.075	(0.172)		0.290	(0.175)	
northeast	0.069	(0.436)		0.425	(0.618)		1.053	(0.197)	***	0.791	(0.196)	***
2 parents	-0.699	(0.278)	*	-1.001	(0.372)	**	-0.649	(0.142)	***	-0.416	(0.152)	**
single mother	-0.386	(0.241)		-0.521	(0.310)		-0.043	(0.187)		0.035	(0.188)	
single father	-0.564	(0.524)		0.704	(0.832)		0.335	(0.388)		-0.432	(0.323)	
other	-1.093	(0.380)	**	0.129	(0.427)		-1.100	(0.330)	***	-0.370	(0.425)	
Mother's edu.	0.147	(0.240)		-0.651	(0.330)	*	0.058	(0.163)		0.018	(0.203)	
Mother's edu2	0.004	(0.022)		0.075	(0.029)	**	0.011	(0.014)		0.011	(0.017)	
Father's edu	-0.180	(0.270)		0.089	(0.324)		-0.076	(0.150)		-0.167	(0.161)	
Father's edu2	0.028	(0.025)		0.008	(0.030)		0.020	(0.013)		0.029	(0.014)	*
Log income	0.528	(0.687)		-0.199	(0.713)		-0.161	(0.446)		-0.636	(0.307)	*
Log income2	-0.031	(0.108)		0.051	(0.123)		0.092	(0.066)		0.173	(0.048)	***
Lived envir.	0.091	(0.152)		0.118	(0.230)		0.452	(0.132)	**	0.378	(0.142)	**
PVT score	0.048	(0.009)	***	0.029	(0.010)	**	0.027	(0.005)	***	0.031	(0.005)	***
GPA	1.107	(0.161)	***	1.202	(0.188)	***	1.424	(0.092)	***	1.213	(0.092)	***
birth weight	0.006	(0.005)		0.001	(0.007)		0.005	(0.003)		0.004	(0.003)	
W1 SRH	0.249	(0.101)	**	0.426	(0.144)	**	0.240	(0.077)	**	0.163	(0.075)	*
W1 BMI	0.007	(0.019)		0.032	(0.027)		-0.024	(0.016)		-0.008	(0.015)	
W1 daily smoke	0.622	(0.697)		-3.103	(1.133)	**	-0.196	(0.386)		-0.623	(0.482)	
W1 physical act.	-0.207	(0.096)	*	-0.052	(0.148)		0.044	(0.061)		0.126	(0.066)	
W1 drinking	0.054	(0.068)		-0.124	(0.083)		0.024	(0.042)		-0.041	(0.039)	
_cons	-10.229	(2.326)	***	-8.879	(2.879)	**	-14.930	(1.400)	***	-14.630	(1.542)	***

References

Almond, Douglas and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypoth
Journal of Economic Perspectives 25:153-172.
Bang, Heejung and James M. Robins. 2005. "Doubly Robust Estimation in Missing Dat
Causal Inference Models." *Biometrics* 61:962-972.
Boardman, Jason D., Jarron M. Saint Onge, Richard G. Rogers, and Justin T. Denney. 20
"Race Differentials in Obesity: The Impact of Place." *Journal of Health and Soci
Behavior* 46:229-243.
Brookhart, M. Alan, Sebastian Schneeweiss, Kenneth J. Rothman, Robert J. Glynn, Jerr
Avorn, and Til Stürmer. 2006. "Variable Selection for Propensity Score Models
American Journal of Epidemiology 163:1149-1156.
Christakis, Nicholas A. and James H. Fowler. 2007. "The Spread of Obesity in a Large S
Network Over 32 Years." *New England Journal of Medicine* 357:370-379.
—. 2008. "The Collective Dynamics of Smoking in a Large Social Network." *New Engla
Journal of Medicine* 358:2249-2258.
Cutler, David M. and Edward L. Glaeser. 2007. "Social Interactions and Smoking." in
Harvard Institute of Economic Research Discussion Paper Number 2153.
Cutler, David M. and Adriana Lleras-Muney. 2008. "Education and Health: Evaluating
Theories and Evidence." Pp. 29-60 in *Making Americans Healthier: Social and
Economic Policy as Healthy Policy*, edited by R. F. Schoeni, J. S. House, G. A. Kapl
and H. Pollack. New York: Russell Sage Foundation.
—. 2010. "Understanding Differences in Health Behaviors by Education." *Journal of H
Economics* 29:1-28.
Dahly, Darren L., Linda S. Adair, and Kenneth A. Bollen. 2009. "A Structural Equation
of the Developmental Origins of Blood Pressure." *International Journal of
Epidemiology* 38:538-548.
Dunn, L. and L. Dunn. 1981. *PPVT-R Manual*. Circle Pines, MN: American Guidance Ser
Elo, Irma T. 2009. "Social Class Differentials in Health and Mortality: Patterns and
Explanations in Comparative Perspective." *Annual Review of Sociology* 35:553-
Farmer, Melissa M. and Kenneth F. Ferraro. 2005. "Are Racial Disparities in Health
Conditional on Socioeconomic Status?" *Social Science & Medicine* 60:191-204.
Ferraro, Kenneth F. and Melissa M. Farmer. 1999. "Utility of Health Data from Social
Surveys: Is there a Gold Standard for Measuring Morbidity?" *American Sociolo
Review* 64:303-315.
Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Chang
the Last One Hundred Years*. New York: Russell Sage Foundation.
Freese, Jeremy and Karen Lutfey. 2011. "Fundamental Causality: Challenges of an
Animating Concept for Medical Sociology." Pp. 67-81 in *Handbook of the Sociol
health, Illness, and Healing: A Blueprint for the 21st Century*, edited by B. A.
Pescosolido, J. K. Martin, J. D. McLeod, and A. Rogers. New York: Springer.
Goesling, Brian. 2007. "The Rising Significance of Education for Health?" *Social Force
85:1621-1643.
Gottfredson, Linda S. 2004. "Intelligence: Is It the Epidemiologists' Elusive "Fundame
Cause" of Social Class Inequalities in Health?" *Journal of Personality and Social
Psychology* 86:174-199.

Guo, Shenyang and Mark W. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: Sage.
Harris, Kathleen Mullan. 1999. "The Health Status and Risk Behavior of Adolescents in Immigrant Families." Pp. 286-347 in *Children of Immigrants: Health, Adjustment, and Public Assistance*, edited by D. J. Hernandez. Washington, D.C.: National Academy Press.
Harris, Kathleen Mullan, Carolyn T. Halpern, Eric A. Whitsel, Jon M. Hussey, Joyce W. Tabor, Pamela P. Entzel, and J. Richard Udry. 2009. "The National Longitudinal Study of Adolescent Health: Research Design." URL: http://www.cpc.unc.edu/projects/addhealth/design.
Heckman, James J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35:1-97.
Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88:389-432.
Hirschi, Travis and Michael R. Gottfredson. 1994. "The Generality of Deviance." New Brunswick, NJ: Transaction Publishers.
Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945-970.
House, James S., Karl R. Landis, and Debra Umberson. 1988. "Social Relationships and Health." *Science* 241:540-545.
House, James S., Paula M. Lantz, and Pamela Herd. 2005. "Continuity and Change in the Social Stratification of Aging and Health Over the Life Course: Evidence From a Nationally Representative Longitudinal Study From 1986 to 2001/2002 (Americans' Changing Lives Study)." *Journals of Gerontology (Social Sciences)* 60B:15-26.
Idler, Ellen L. and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38:21-37.
Jylhä, Marja. 2009. "What Is Self-Rated Health and Why Does It Predict Mortality? Towards a Unified Conceptual Model." *Social Science & Medicine* 69:307-316.
Jylhä, Marja, Stefano Volpato, and Jack M. Guralnik. 2006. "Self-Rated Health Showed a Graded Association with Frequently Used Biomarkers in a Large Population Sample." *Journal of Clinical Epidemiology* 59:465-471.
Kalton, Graham and Ismael Flores-Cervantes. 2003. "Weighting Methods." *Journal of Official Statistics* 19:81-97.
Kawachi, Ichiro, S. V. Subramanian, and Daniel Kim. 2008. "Social Capital and Health." New York: Springer.
Kitagawa, Evelyn M. and Philip M. Hauser. 1973. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge: Harvard University Press.
Lantz, Paula M., James S. House, Richard P. Mero, and David R. Williams. 2005. "Stress, Life Events, and Socioeconomic Disparities in Health: Results from the Americans' Changing Lives Study." *Journal of Health and Social Behavior* 46:274-288.
Link, Bruce G. and Jo Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* 35:80-94.

Link, Bruce G., Jo C. Phelan, Richard Miech, and Emily Leckman Westin. 2008. "The Resources That Matter: Fundamental Social Causes of Health Disparities and the Challenge of Intelligence." *Journal of Health and Social Behavior* 49:72-91.
Little, Roderick J. A. and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. Hoboken, NJ: Wiley.
Liu, Hui and Robert A. Hummer. 2008. "Are Educational Differences in U.S. Self-Rated Health Increasing?: An Examination by Gender and Race." *Social Science & Medicine* 67:1898-1906.
Luo, Ye and Linda J. Waite. 2005. "The Impact of Childhood and Adult SES on Physical, Mental, and Cognitive Well-Being in Later Life." *Journals of Gerontology (Social Sciences)* 60B:S93-S101.
Lynch, Scott M. 2003. "Cohort and Life-Course Patterns in the Relationship Between Education and Health: A Hierarchical Approach." *Demography* 40:309-331.
—. 2006. "Explaining Life Course and Cohort Variation in the Relationship between Education and Health: The Role of Income." *Journal of Health and Social Behavior* 47:324-338.
Mirowsky, John and Catherine E. Ross. 1998. "Education, Personal Control, Lifestyle and Health: A Human Capital Hypothesis." *Research on Aging* 20:415-449.
—. 2003. *Education, Social Status, and Health*. New York: Aldine De Gruyter.
—. 2008. "Education and Self-Rated Health: Cumulative Advantage and Its Rising Importance." *Research on Aging* 30:93-122.
Mokdad, Ali H., James S. Marks, Donna F. Stroup, and Julie L. Gerberding. 2004. "Actual Causes of Death in the United States, 2000." *Journal of the American Medical Association* 291:1238-1245.
Morgan, Stephen L. and David J. Harding. 2006. "Matching Estimators of Causal Effects: Prospects and Pitfalls of Theory and Practice." *Sociological Methods & Research* 35:3-60.
Morgan, Stephen L. and Jennifer J. Todd. 2008. "A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects." *Sociological Methodology* 38:231-281.
Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.
Nguyen, Quynh C., Joyce W. Tabor, Pamela P. Entzel, Yan Lau, Chirayath M. Suchindran, Jon M. Hussey, Carolyn T. Halpern, Kathleen Mullan Harris, and Eric A. Whitsel. 2011. "Discordance in National Estimates of Hypertension Among Yound Adults." *Epidemiology* 22:532-541.
Pampel, Fred C., Patrick M. Krueger, and Justin T. Denney. 2010. "Socioeconomic Disparities in Health Behaviors." *Annual Review of Sociology* 36:349-370.
Preston, Samuel H. and Irma T. Elo. 1995. "Are Educational Differentials in Adult Mortality Increasing in the United States?" *Journal of Aging and Health* 7:476-496.
Robins, James M. and Andrea Rotnitzky. 1997. "Comment on 'Inference for Semiparametric Models: Some Questions and an Answer' by Bickel and Kwon." *Statistica Sinica* 11:920-936.
Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41-55.

Ross, Catherine E. and John Mirowsky. 1989. "Explaining the Social Patterns of Depression: Control and Problem Solving - Or Support and Talking." *Journal of Health and Social Behavior* 30:206-219.
—. 2006. "Sex Differences in the Effect of Education on Depression: Resource Multiplication or Resource Substitution?" *Social Science & Medicine* 63:1400-1413.
—. 2010. "Why Education Is the Key to Socioeconomic Differentials in Health." Pp. 33-51 in *Handbook of Medical Sociology (6th ed)*, edited by C. E. Bird, P. Conrad, A. M. Fremont, and S. Timmermans. Nashville, TN: Vanderbilt University Press.
Rubin, Donald B. 1973. "Matching to Remove Bias in Observational Studies." *Biometrics* 29:159-183.
—. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688-701.
—. 1997. "Estimating Causal Effects from Large Data Sets Using Propensity Scores." *Annals of Internal Medicine* 127:757-763.
—. 2006. *Matched Sampling for Causal Effects*. New York: Cambridge University Press.
Seeman, Teresa E. 1996. "Social Ties and Health: The Benefits of Social Integration." *Annals of Epidemiology* 6:442-451.
Shuey, Kim M. and Andrea E. Willson. 2008. "Cumulative Disadvantage and Black-White Disparities in Life-Course Health Trajectories." *Research on Aging* 30:200-225.
Smith, James P. 2007. "The Impact of Socioeconomic Status on Health over the Life-Course." *Journal of Human Resources* 42:739-764.
Smith, Kristen P. and Nicholas A. Christakis. 2008. "Social Networks and Health." *Annual Review of Sociology* 34:405-429.
Snyder, Thomas D. and Sally A. Dillow. 2011. "Digest of Education Statistics 2010." National Center for Education Statistics, Washington, D.C.
Splawa-Neyman, Jerzy. [1923] 1990. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. (translated, with discussion)." *Statistical Science* 5:465-480.
StataCorp. 2011. *Stata Statistical Software: Release 12*. College Station, TX: StataCorp, LP.
Williams, David R. and Chiquita Collins. 1995. "US Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology* 21:349-386.
Willson, Andrea E., Kim M. Shuey, and Glen H. Elder. 2007. "Cumulative Advantage Processes as Mechanisms of Inequality in Life Course Health." *American Journal of Sociology* 112:1886-1924.
Zheng, Hui and Kenneth C. Land. 2012. "Composition and Decomposite in US Gender-Specific Self-Reported Health Disparities, 1984-2007." *Social Science Research* 41:477-488.