Center for Human Capital Studies

WORKING PAPER SERIES

The Long Run Health Returns to College Quality

Jason M. Fletcher and David E. Frisvold

CHCS Working Paper 2012-1
July 2012

Abstract: The link between education and health is one of the most robust empirical relationships in the social sciences. However, little research has examined the effects of educational quality on health outcomes. We estimate the long run relationship between health behaviors and graduating from a selective college in the 1960s using the Wisconsin Longitudinal Study, which has tracked siblings for over 50 years. Importantly, we control for measures of health endowments, ability, and time preferences before college enrollment as well as shared family and environmental factors. We find large effects of college selectivity on reducing overweight for individuals in their sixties.

JEL classification: I12, I21, J24

Key words: health production; college quality; obesity

This research was supported by the WLS pilot grant program (Fletcher and Frisvold), the National Institute on Aging (R01 AG027045, Fletcher), the Emory Global Health Institute (Frisvold), and the Robert Wood Johnson Foundation (Fletcher and Frisvold). The authors thank Griffin Edwards and John Zimmerman for excellent research assistance, and participants at the Academy Health, American Society of Health Economists, Annual Meeting on the Economics of Risky Behaviors, APPAM, and Southern Economic Association conferences, Dhaval Dave, Ezra Golberstein, Bo MacInnis, Dave Marcotte, and Jim Marton for helpful comments. This research uses data from the WLS of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. The opinions expressed herein are those of the authors. In particular, the authors thank Carol Roan for assistance with the restricted college data.

Please address questions regarding content to corresponding author Jason M. Fletcher, Yale University, School of Public Health, Division of Health Policy and Administration, 60 College St., New Haven, CT 06510, jason.fletcher@yale.edu, and Robert Wood Johnson Foundation Health & Society Scholar, Columbia University, 420 W 118th St., New York , NY 10027; David E. Frisvold, Emory University, Department of Economics, 1602 Fishburne Dr., Atlanta, GA 30322-2240, 404-727-7833, david.frisvold@emory.edu.

FEDERAL RESERVE BANK *of* ATLANTA

The Long Run Health Returns to College Quality[♦]

Jason M. Fletcher [a]
Yale University
Columbia University

David E. Frisvold [b]
Emory University

Abstract:
The link between education and health is one of the most robust empirical relationships in the social sciences. However, little research has examined the effects of educational quality on health outcomes. We estimate the long run relationship between health behaviors and graduating from a selective college in the 1960s using the Wisconsin Longitudinal Study, which has tracked siblings for over fifty years. Importantly, we control for measures of health endowments, ability, and time preferences before college enrollment as well as shared family and environmental factors. We find large effects of college selectivity on reducing overweight for individuals in their 60s.

JEL Classification: I12, I21, J24

Keywords: health production; college quality; obesity

[♦]This research was supported by the WLS pilot grant program (Fletcher and Frisvold), the National Institute on Aging (R01 AG027045, Fletcher), the Emory Global Health Institute (Frisvold), and the Robert Wood Johnson Foundation (Fletcher and Frisvold). The authors thank Griffin Edwards and John Zimmerman for excellent research assistance and participants at the Academy Health, American Society of Health Economists, Annual Meeting on the Economics of Risky Behaviors, APPAM, and Southern Economic Association conferences, Dhaval Dave, Ezra Golberstein, Bo MacInnis, Dave Marcotte, and Jim Marton for helpful comments.
This research uses data from the WLS of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. The opinions expressed herein are those of the authors. In particular, the authors thank Carol Roan for assistance with the restricted college data.
[a] School of Public Health, Division of Health Policy and Administration, Yale University, 60 College Street, New Haven, CT 06510, USA, email: jason.fletcher@yale.edu.
Robert Wood Johnson Foundation Health & Society Scholar, Columbia University, 420 W 118th Street, New York , NY 10027
[b] Department of Economics, Emory University, 1602 Fishburne Drive, Atlanta, GA 30322-2240, USA; phone: 404-727-7833; email: david.frisvold@emory.edu.

Introduction

The empirical link between education and health is one of the most robust relationships in the social sciences and has been shown across many countries and time periods. Grossman and Kaestner (1997) review the voluminous literature and conclude that the available evidence supports a causal relationship. If the conclusion that education has a causal effect on health is correct, a natural question to ask is what are the mechanisms linking education and later health outcomes. Grossman (1972) suggests that education may improve productive efficiency—individuals with greater education are able to use health inputs more efficiently. Kenkel (1991) focuses on the mechanism of allocative efficiency—individuals with greater education choose different health inputs. Likewise, Cutler and Lleras-Muney (2008) suggest that education increases lead to different thinking and decision making patterns.

One limitation with much current research is that education is almost universally treated as a black box—researchers typically examine the returns to an additional year of schooling or credential but rarely examine *why* the increase in education may lead to better health. One potentially important component of the black box of educational attainment is the quality of schooling. There is almost no current research that examines the effects of quality of education on health (Cutler and Lleras-Muney 2008).

In this paper, we add to the literature by estimating OLS and family fixed effects models of the effects of graduating from a selective college during the 1960s on smoking and obesity as individuals age throughout adulthood. We utilize data from the Wisconsin Longitudinal Study (WLS), which includes measures of these health behaviors as well as IQ and health status measures during adolescence for this sample. We are also able to use the panel nature of the data

to focus on several possible mechanisms, such as occupational characteristics including income, marital histories, and cognition.

Overall, our results suggest that graduating from a selective college reduces the probability of being overweight in later adulthood by approximately 15 percentage points, but we find no consistent evidence for reductions in the probability of smoking. Although we are not able to control for all measures of individual heterogeneity such as non-cognitive traits, the results are shown to be robust to including a rich set of pre-college control variables as well as family fixed effects. We then explore several potential mechanisms that may be responsible for the estimated link between selectivity and health, including occupational characteristics, income/wealth measures, marriage market outcomes, and health at mid-life but find limited evidence that the college quality-health effect operates through these pathways. Thus, we present some of the first evidence linking college quality with later health behaviors and demonstrate that college quality influences these behaviors independent of income, occupation, or marriage.

Background

While there has been very limited research that examines the long run health effects of school quality, a large literature has linked more general measures of schooling quantity with measures of long term health. Grossman and Kaestner (1997) review research linking education and health and suggest that there is considerable evidence of a causal relationship. However, only a few studies use rigorous econometric methods to establish causality. These studies are important because common concerns with attributing causal effects of education on health focus

on issues of reverse causality and omitted variables, including ability (Card 1999), time preferences (Fuchs 1982), or other factors (Arendt 2005, Grossman and Kaestner 1997).

One study that attempts to estimate causal effects of education on health is Lleras-Muney (2005), who uses compulsory schooling law changes to instrument for years of completed schooling in predicting mortality. Overall, she finds that a one year increase in educational attainment increases life expectancy at age 35 by a year or more.[1] Oreopoulos (2007) finds that changes in compulsory schooling throughout the mid-twentieth century in England and Ireland reduce the likelihood of being in poor health or depressed. van Kipperslius, O'Donnell, and van Doorslaer (2011) also find that changes in compulsory schooling laws in the Netherlands in 1928 reduced mortality. On the other hand, Clark and Royer (2010) find that changes in compulsory schooling laws in England in 1947 and 1972 did not influence mortality. Similarly, Albouy and Lequin (2009) find that changes in compulsory schooling laws in France in 1923 and 1953 did not influence mortality.

Arendt (2005) uses Danish school reforms as instruments to examine the effects of schooling quantity on self rated health, smoking behavior, and obesity. He finds large effects but the standard errors do not allow the author to rule out no effect of education on health. On the other hand, using variation induced from English school reforms, Jones, Rice, and Rosa Dias (2011) find that greater educational attainment reduces smoking and improves diet. Additionally, deWalque (2007) and Grimard and Parent (2007) find that changes in education stemming from the Vietnam War influenced smoking. Conti, Heckman, and Urzua (2010) conclude that education decreases smoking, depression, and poor health, but not obesity; although, education does increase the frequency of regular exercise.

[1] Mazumder (2007) provides evidence that Lleras-Muney's estimates may be sensitive to the inclusion of state specific time trends, but also finds evidence consistent with a causal effect of education on mortality.

Fletcher and Frisvold (2009) use an alternative approach to estimate the effects of education quantity on a different set of health outcomes—receipt of preventive health care. The authors use family fixed effects estimators and show that attending college boosts preventive care receipt around age 65 by 5-15%, even controlling for ability, time preferences, and health endowments. Their results provide suggestive evidence that a mechanism linking educational quantity with health was occupational characteristics but not health insurance, income, or wealth. Like Fletcher and Frisvold (2009), Webbink, Martin, and Visscher (2010) use family fixed effects models and extend the set of health outcomes to include obesity. Specifically, the authors compare identical twins in Australia and conclude that an increase in years of schooling reduces the likelihood of being overweight for men and that the size of this relationship increases with age, primarily from the mid 30s to the mid 40s.

While these and other studies suggest a causal link between education quantity and health, there are several reasons to also focus on educational quality and health. First, educational quality has been shown to affect other life outcomes. Behrman et al. (1996) show that several measures of school quality (Ph.D. granting, private universities, faculty salaries) increase adult wages for female twins. Black and Smith (2004) use matching estimators to show large effects of college selectivity on adult wages. Hoekstra (2009) finds, using a regression discontinuity design based on the admission criteria, that attending the flagship state university increases earnings as a young adult. In contrast, Dale and Krueger (2002) compare students with similar patterns of applications and acceptances and show that the wage effects of selective colleges are smaller than previous research. Long (2008) replicates the methods of Black and Smith (2004) and Dale and Krueger (2002) with recent data and concludes that college quality,

using a variety of different measures, has a positive effect on college graduation and household income.

There are several theories suggesting that school quality may directly affect health. Educational quality may provide more information and better skills than lower quality education. Additionally, educational quality, college selectivity in particular, may provide better future employment opportunities through social contacts and better jobs (Ishida et al. 1997, Rosenbaum 1984). Better schools may also have different culture and norms for health behaviors such as smoking, drinking, eating, and health investments (Cockerham, Rutten, and Abel 1997).

On the other hand, rather than a causal relationship, educational quality and health may be linked due to self-selection. Similar to the relationship between educational attainment and health, the correlation between educational quality and health could reflect time preferences (Fuchs, 1982) or non-cognitive abilities (Conti, Heckman, and Urzua, 2010).

There is only a small amount of evidence examining the effects of educational quality on later health. Several recent papers have begun to examine the influence of the quality of primary and secondary schooling on later health outcomes. Frisvold and Golberstein (2011a,b) find that the substantial improvements in the quality of schools attended by blacks in the South in the first half of the twentieth century increased a broad range of health outcomes later in life and reduced health disparities. MacInnis (2009) finds that primary and secondary school quality significantly increases cognitive functioning in old age. Johnson (2010) demonstrates that the increases in school quality following school desegregation improved adult self-reported health status for blacks.

In the only other papers focused on post-secondary school quality and health in the literature, Ross and Mirowsky (1999) examine educational quality by estimating the effects of

attending a selective college on health outcomes, including physical functioning and self rated health. They find suggestive correlation evidence that attending a more selective college is associated with better health. They also present evidence that educational quality affects health through health behaviors rather than income, employment and other factors. However, one limitation of this paper is that the authors do not allow for the possibility that college selectivity is correlated with the determinants of health outcomes. Fletcher and Frisvold (2012) examine the impact of attending a selective college on young adult health behaviors during and shortly after college and find that selective college attendance decreases tobacco and marijuana use, but increases binge drinking. Our paper contributes to the literatures on the returns to education, the impact of school quality, and the determinants of health by using a variety of empirical strategies and a rich dataset to examine the impact of graduating from a selective college on health behaviors throughout adulthood.

Data

In this study, we use data from the Wisconsin Longitudinal Study (WLS) to examine the impact of college quality on later health outcomes. The WLS is a longitudinal study of a one-third random sample of the graduating high school class of 1957 in Wisconsin. Survey information from 10,317 of the graduates was collected in 1957, 1964, 1975, 1992-1993, and 2003-2004. Important for our research design, the WLS sample also includes information from a randomly selected sibling that was collected in 1977, 1993-1994, and 2005-2007. The sample is broadly representative of white, non-Hispanic American men and women who have completed at

least a high school education.[2] The WLS includes extensive information about schooling, social background, labor market experiences, and health behaviors.[3]

Important features of the WLS for this study are the longitudinal nature of the data set that combines information about the health behaviors of individuals aged 63 to 67 years old with information about the family economic characteristics prior to college attendance. In addition to the rich sociodemographic and health information contained in the survey, using the name of the Bachelor's degree-granting college, we merge information on college selectivity taken from Barron's Profiles of American Colleges (Fine, 1969).[4] This information ranks schools based on the median SAT scores, high school rank, and high school grade point average of the freshman class and produces categories of college selectivity.[5] Our analysis compares individuals who graduated from colleges that were at least very competitive (including colleges classified as very competitive, highly competitive, and most competitive) to individuals who attended competitive, less competitive, and non-competitive colleges.[6] [7] Thus, our college selectivity variables

[2] As noted in the WLS Handbook (2007, p.14), "among Americans aged 50 to 54 in 1990 and 1991, approximately 66 percent were non-Hispanic white persons who completed at least 12 years of schooling. Some strata of American society are not well represented. The WLS sample is mainly of German, English, Irish, Scandinavian, Polish, or Czech ancestry. It is estimated that about 75 percent of Wisconsin youth graduated from high school in the late 1950s – everyone in the primary WLS sample graduated from high school; about seven percent of their siblings did not graduate from high school. Minorities are not well-represented: there are only a handful of African American, Hispanic, or Asian persons in the sample. … About 19 percent of the WLS sample is of farm origin, and that is consistent with national estimates of persons of farm origin in cohorts born in the late 1930s. … In 1964, 1975, and again in 1992, about two-thirds of the sample lived in Wisconsin, and about one-third lived elsewhere in the U.S. or abroad."

[3] Full information can be found online: http://www.ssc.wisc.edu/wlsresearch/

[4] Barron's Profile of American Colleges is a reference for students considering college attendance, their parents, and guidance counselors, which provides an overview of 4-year colleges in the United States that includes their admissions standards, expenses, and programs of study. We use the 1969 edition of this resource, which provides the earliest measures of college quality, to our knowledge. Because most survey respondents attended college between 1958 and 1963, there is likely to be measurement error in the college quality variables. However, due to the stability of college rankings, the measurement error is likely minimal.

[5] The information used to construct these selectivity categories is not available for all colleges; thus, we focus our analysis on the selectivity categories: (1) most competitive, (2) highly competitive, (3) very competitive, (4) competitive, (5) less competitive, and (6) non-competitive.

[6] Examples of colleges in the very competitive category include University of Wisconsin, University of Michigan, and University of Minnesota.

includes colleges with a median SAT score above 1100 that accept students who are in the top half of their high school graduating class and have a minimum of a B- to B grade point average. Over 30% of the graduates from a 4-year college in the WLS graduated from a very competitive college and nearly 3% graduated from a highly competitive college. We use this information as our key explanatory variable in our analysis below to examine whether college quality affects health behaviors, controlling for important pre-college information as well as family environments though the use of fixed effects.

Of the 8,493 original respondents and their 5,365 siblings who were surveyed between 1992 and 1994, approximately 27 percent graduated from a 4-year college. We restrict the sample to the 2,356 original respondents and 1,362 siblings who graduated from a 4-year college. The selectivity of the college is based on the name of the college that the individual received a Bachelor's degree, measured in the 1992-1993 and 1993-1994 survey years if available, or the 1975-1977 survey if the later college code is missing. Excluding the six individuals who graduated from a college outside of the United States, we are able to determine the selectivity of the college for 3,708 individuals. We further exclude from the sample individuals with missing values for BMI and current smoking status in 1993 or 2004, which

[7] As shown in Table 1, for the sample of all college graduates (n=3210), only 3.6% graduated from a highly competitive college. As shown in Table 2, among the sibling sample (n=530), only 3.2% graduated from a highly competitive college. Thus, the number of individuals graduating from a highly competitive college is very small, and, for this reason, we do not focus our analysis on individuals graduating from a highly competitive college. As shown in Table 1, for the sample of all college graduates, 64.2% graduated from a competitive college. As shown in Table 2, among the sibling sample, 64.9% graduated from a competitive college. The selection criteria for schools to be classified as competitive is a median SAT score of 900-1100 and students should have an average GPA of B- or better and be in the top half of their high school class. The comparison group for students graduating from at least a competitive college would be students graduating from a less competitive college or a non-competitive college. As would be expected because the cutoff for a competitive college is not very selective, the magnitudes of the estimates are smaller and most are not statistically significant. Thus, because the highly competitive colleges are so selective as to only include a few people in the sample and the competitive colleges are not very selective and include over half the sample, we focus our analysis on graduating from a very selective college.

reduces the sample to 3,210 individuals.[8, 9] In order to maximize our sample size, we impute missing values of the control variables and include missing-value indicator variables in our specifications.[10]

These reductions in the sample based on non-response and mortality could pose sample selection issues for our main results, particularly if college quality is related to appearing in the sample. We show in Appendix Table 2 that our results are robust to using inverse probability weights to adjust for attrition bias.[11] We also have investigated whether college quality influences mortality in our sample and found no evidence.[12]

Table 1 presents the means and standard deviations of the individual and family background characteristics of the sample of college graduates with non-missing college selectivity and at least one non-missing health behavior. The summary statistics show that 26 percent of respondents who graduated from college are obese in the 2004 wave of the data, 69

[8] In 2004, 7,732 original respondents completed the survey for a response rate of 75 percent; however, only 6,845 respondents or 66 percent of the original respondents completed the mail questionnaire that included the height, weight, and smoking questions. Among the 3,472 respondents who did not complete the mail questionnaire in 2004, 1,288 were deceased, 785 were not able to be contacted, and 1,399 refused to complete the questionnaire (WLS 2009). Of the 7,928 siblings of the original respondents, 4,004 or 51% completed the mail questionnaire in 2005. Among the 3,924 respondents who did not complete the mail questionnaire in 2004, 1,226 were deceased, 1,365 were not able to be contacted, and 1,333 refused to complete the questionnaire (WLS 2009). For comparison, the response rate of survey respondents in the 1988 wave of the Panel Study of Income Dynamics who lived in the original 1968 households is 56.1 percent (PSID User Guide, available at http://psidonline.isr.umich.edu/Guide/ug/chap5.html).

[9] Appendix Table 1 displays the characteristics of WLS respondents who were not included in this analysis. As expected, the individuals who graduated from college during this time period are relatively advantaged. Compared to the full sample of respondents, college graduates have mothers and fathers with more years of schooling completed, were raised in smaller families with higher incomes, and have higher IQ scores. College graduates with missing health data have similar pre-college characteristics to the college graduates included in Table 1.

[10] We also assign siblings with missing values the value of the non-missing sibling for living with both parents and assign sibling pairs with missing values the imputed value of the graduate for mother's education, father's education, and family income; these missing values will difference out in the fixed effects model.

[11] The weights equal the inverse of the predicted probability of an individual being in the sample. Additionally, graduating from a very competitive college is not correlated with whether the respondent is included in this analysis sample. In a regression of whether the respondent is in the sample on college quality and individual and family characteristics, the coefficient for graduating from a very competitive college is 0.008 with a standard error of 0.012.

[12] Of the 3,708 individuals with non-missing college quality values, who must have been alive through 1975 for us to be able to assign a college quality value, 376 died by 2004. One hundred twenty seven, or one-third, graduated from a selective college. The correlation between graduating from a very competitive college and mortality by 2004 is -0.01 and is not statistically different from zero. Similarly, graduating from a very competitive college is not predictive of early mortality in a regression.

percent are overweight, and 8 percent currently smoke cigarettes. In 1993, 19 percent of the college graduates were obese, 61 percent were overweight, and 11 percent smoked cigarettes. The sample is 64 years old on average in 2004 and 43 percent are female. As discussed above, the data also include information preceding college attendance, including IQ, childhood health status, and family background information. For example, over 3% of the sample reported poor or fair health as a child.

Table 1 also displays the characteristics of college graduates who attended a selective college and college graduates who did not. Individuals who graduated from a selective college have lower rates of obesity, overweight, and smoking in 1993 and 2004 and have higher total household income and assets in 2004. Graduates from a selective college also have a more advantaged family background; mothers have one additional year of schooling and fathers have an additional year and a half of schooling, on average. Family incomes during high school are over $10,000 (in 2004 dollars) greater and the number of siblings is lower for graduates from a selective college. The substantial difference in family background suggests that comparisons between individuals with different college choices that fail to adequately control for family backgrounds may lead to biased estimates of the impact of college quality.

To take advantage of the use of family fixed effects, we must constrain our analysis sample to sibling pairs in which both siblings graduated from college.[13] Additionally, to examine the influence of college quality as individuals age on a consistent sample, we restrict the

[13] Only the name of the college that the individual graduated from is available for both the original WLS respondents and their siblings. Thus, we focus on the college selectivity of the college that individuals graduated from, as opposed to the college that individuals initially attended. In Appendix Table 3, we show the results from OLS and linear probability models of the relationship between attending a very competitive college and health behaviors for the sample of original WLS respondents. These results show that there is not a statistically significant relationship between attending a selective college and later health behaviors; thus, based on our results below, college graduation is the more relevant margin for influencing health behaviors. We also show that our results are very similar if we compare the WLS respondents with a sibling in the sample with WLS respondents without a sibling in the sample. We will return to these results below.

sample to sibling pairs with non-missing health measures in 1993 and 2004. Our analysis sample consists of 265 original respondents and their 265 siblings for a total of 530 observations.

We present means and standard deviations of our analysis sample of 530 individuals in Table 2. Among this sample of sibling pairs, 23 percent are obese in the 2004 wave of the data, 67 percent are overweight, and 7 percent smoke cigarettes. The sample is 64 years old on average and 44 percent are female. The table shows that our analysis sample is similar, though slightly more advantaged, than the full sample of college graduates in this data set with slightly higher IQ scores, higher family income during childhood and higher parental education levels.

Identification in family fixed effects specifications comes from discordant sibling pairs, where one sibling graduated from a selective college and the other sibling graduated from a less selective college. Table 2 includes the descriptive statistics for the 91 discordant sibling pairs.[14] Comparing the individual and family background characteristics of the graduates who attended a selective college to the graduates who did not for the discordant sibling pairs and the full sample of college graduates suggests that restricting the sample to sibling pairs improves the comparability of the groups. For example, the difference in future expectations about college attendance is reduced. By definition, parental education, family income, and other family background characteristics are the same. In a set of family fixed effects regressions of each characteristic separately on graduating from a selective college, childhood health, plans to attend college, and birth order are not related to college selectivity. On the other hand, males, older siblings, and siblings with a higher IQ were more likely to graduate from a selective college; however, IQ is not predictive of these health measures and the interaction terms between college quality and these characteristics are never statistically significant in family fixed effects

[14] Although this effective sample size is less than ideal, we note that it is comparable to many other studies using sibling comparisons and the sample size should influence the precision but not the magnitude of the estimates (Garces, Thomas, and Currie, 2002).

regressions. Overall, the differences in pre-college characteristics of selective college graduates and less selective college graduates are reduced by focusing on siblings, which provides suggestive evidence that estimates based on sibling comparisons are less likely to be biased due to selection on observed and unobserved characteristics.

Methods

In this paper, we follow Fletcher and Frisvold (2009) and use a variety of empirical strategies to examine the links between college quality and later health behaviors. First, we examine the determinants of obesity and smoking using regression analysis. We estimate the likelihood of each health behavior, B_i, as a function of an individual's college selectivity, C_i, individual and family characteristics, X_i, and an idiosyncratic shock, ε_i:

$$B_i = \alpha_0 + \alpha_1 C_i + \alpha_2 X_i + \varepsilon_i . \quad (1)$$

Individual and family characteristics included in X are mother's education, father's education, family income during high school, number of siblings, sex, age, birth order, whether the individual lived with both parents during high school, and a dummy variable indicating whether the individual is in the graduate sample (vs. the sibling sample). Including a wide array of family background characteristics measured prior to college is important due to the influence of family background on education and health outcomes (Case et al., 2005; Wolfe and Behrman, 1987).

Next, to further explore whether the observed correlation that exists between college quality and the health behaviors is the result of reverse causality from health to college quality, we include measures of childhood health (whether at least one month of school was missed due

to illness and an indicator variable for poor or fair self-reported health in childhood).[15] We thus augment the above equation to include past health, $H_{i,t-1}$:

$$B_{i,t} = \beta_0 + \beta_1 H_{i,t-1} + \beta_2 C_i + \beta_3 X_i + \varepsilon_i . \quad (2)$$

Additionally, to examine the possibility that third factors, which are commonly unobserved, lead to a spurious correlation between college and health behaviors, we augment the previous demand equation with additional characteristics that may jointly influence both college and behaviors. We include the Henmon-Nelson measure of IQ as a proxy for innate ability.[16] We also include whether the individual planned to attend college when they were 16 years old as a proxy for time preferences and future expectations prior to college attendance.[17]

In our preferred specifications, we further saturate the empirical model to control for unobserved characteristics that are common to siblings by using a family fixed effects estimator that compares the health behaviors of siblings with different college selectivity. To implement this strategy, we compare the obesity and smoking status of WLS graduates to their siblings' behaviors using the 1993 or 2004 data by adding a family fixed effect, μ_f, to the previous equation:

[15] These measures of childhood health are derived from questions asked of respondents in the latest survey wave, which introduces the possibility of substantial recall error. On the other hand, significant events, such as illness for at least one month, are less likely to be subject to recall error (Garces, Thomas, and Currie, 2002). The most commonly reported illness is infectious and parasitic diseases. In the WLS, childhood measures of health behaviors are unavailable. It could be possible that individuals are overweight in childhood but do not report poor or fair childhood health or a prolonged childhood illness and that childhood overweight is related to both adult overweight and college selectivity. However, Kaestner and Grossman (2009) and Fletcher and Lehrer (2009) suggest that there is no relationship between childhood overweight and educational achievement.

[16] This variable is based on tests of mental ability conducted for all high school students in the state by the Wisconsin State Testing Service in 9th and 11th grade (Hauser, 2005). We use the recommended measure constructed by the WLS that consists of the 11th grade score and a transformation of the 9th grade score if the 11th grade score is missing.

[17] Further information about high school experiences is available for the original WLS respondents, but not the siblings. We examine the influence of the extent to which the respondent discussed their future (post-high school) plans with teachers, counselors, and parents, as an alternative proxy for future expectations, and whether the respondent participated in high school sports on the estimated impact of college selectivity using the sample of original WLS respondents only. The inclusion of these additional variables does not influence the results.

$$B_{i,f,t} = \delta_0 + \delta_1 H_{i,f,t-1} + \delta_2 C_{i,f} + \delta_3 X_{i,f} + \mu_f + \varepsilon_{i,f} . \quad (3)$$

By controlling for a wide array of exogenous individual characteristics that influence college selectivity and might differ between siblings, in addition to any fixed family characteristics that determine selectivity, we hope to identify the nature of the relationship between college quality and health behaviors.[18]

This estimate is derived by comparing sibling pairs that consist of one sibling who graduated from a selective college and one sibling who graduated from a less selective college. The identifying assumption is that the underlying reasons why one sibling attended the selective college and the other did not are conditionally uncorrelated with later health behaviors. Thus, consistent with the majority of the literature examining the impact of college quality on labor market outcomes, our identification strategy relies on the conditional independence of college quality. A concern with this identifying assumption is the possibility of unobserved heterogeneity that is correlated with college quality; in other words, what is essential for our identification strategy is that we capture important differences between siblings in our measurable characteristics. Our results below suggest that, in order to confound our results, there would need to be a critical measure of individual heterogeneity between siblings that is both correlated with college selectivity and later health that explains more of the variance in health behaviors than our measured characteristics, such as IQ, time preferences, childhood health, to eliminate the estimated effect of college quality on obesity.[19] One possibility for such

[18] As noted in Fletcher and Frisvold (2009), a limitation of the fixed effects strategy is the exaggeration of the influence of measurement error, which could bias the estimate of δ_2 towards zero. Additionally, while the fixed effects strategy is implemented to reduce endogenous variation in college attendance, exogenous variation may also be reduced (Bound and Solon, 1999). Further, sibling spillovers, where the behavior of one sibling is influenced by the other, would bias the estimates towards zero.

[19] To further examine this possibility, we include personality characteristics (extraversion, agreeableness, conscientiousness, neuroticism, and openness) as additional covariates and find that our results are robust to including these characteristics. A caveat to the importance of this robustness check is that the personality

a critical measure of individual heterogeneity would be non-cognitive traits. Conti, Heckman, and Urzua (2010) suggest that half of the association between the quantity of schooling and health reflects selection into schooling and that non-cognitive skills are more important determinants of healthy behaviors than cognitive skills. Although there are not direct measures of non-cognitive skills prior to college in the WLS, we examine this possibility by controlling for the respondents expected job status as an adult when the respondent was in high school, based on the Duncan SEI measure, and note that the results reported below are robust to including this individual-specific characteristic.

In addition to unmeasured between-sibling heterogeneity, another concern about our strategy is that unobserved parental investments that differ among siblings influence both college attendance and later health behaviors. To examine this possibility, we estimate additional specifications that include the interaction of birth order, as a proxy variable for parental investments, and college quality. Birth order is related to the amount of time that parents spend with their children, which is an important measure of parental investment (Price 2008). All results are robust to including this interaction term, which suggests that unobserved parental investments are not influencing the estimates of the impact of college quality.

Finally, to examine the robustness of the results from the above methods, we use matching estimation methods to estimate the impact of graduating from a selective college on health behaviors. Black and Smith (2004), in their study of the impact of college quality on earnings, use matching methods to relax the parametric assumption of linearity embedded in the

---

characteristics are measured during adulthood in 1993, not prior to college graduation. In their meta-analysis of the psychology literature, Roberts, Walton, and Viechtbauer (2006) conclude that these personality traits are not fixed throughout adulthood (see also Almlund et al., 2011). We note, however, that college selectivity is statistically significantly related only to conscientiousness and openness at the 10 percent significance level in family fixed effects models. These results also suggest that changes in personality are not a mechanism through which graduating from a very selective college influences health behaviors.

OLS specifications that are commonly estimated in the college quality literature and to explicitly focus on the overlap in the distributions of observable characteristics between the samples of students who attended selective and less selective colleges. These authors find that matching estimates are similar to OLS estimates for men, but are smaller for women. Similar to Black and Smith (2004), we examine the robustness of our results to the parametric assumption of linearity that is imposed in the previous specifications. Specifically, we use the bias-corrected nearest neighbor matching estimator described in Abadie and Imbens (2002) as well as several alternative propensity-score matching estimators, such as stratification and kernel matching. We estimate these matching estimators for the full sample of observations, including individuals without a sibling in the sample, in the common support and the trimmed sample of observations with propensity scores in the range of [0.1, 0.9], which is the optimal subsample for estimating the average treatment effect on the treated under a wide range of distributions (Crump et al. 2006). We also use matching methods for the "thick support" sample, which includes observations with propensity scores in the range of [0.33, 0.67], similar to Black and Smith (2004). These techniques compare the health behaviors of individuals with similar observable characteristics, but whose college attendance and graduation choices differed. If individuals choose college selectivity based on the extensive list of observable characteristics that we can match on in the WLS, then matching estimates the causal effect of college selectivity on obesity, overweight, BMI, and smoking.

Results

Table 3 presents our main set of findings for log adult body mass index in 1993 (approximately aged 54) and 2004 (approximately aged 65).[20] The baseline results in columns 1 and 4 show that graduating from a very competitive college in the 1960s reduces BMI by 2.3 and 4.7 percent, respectively, which suggests that the effects of college quality are important in magnitude and grow over time. In columns 2 and 5, we further control for pre-college measures of health status, IQ, and a proxy measure of time preference, which does not influence the link between college selectivity and BMI. In columns 3 and 6, we control for all shared factors of the two siblings, such as family background and 50% of their genetic endowments, and find that the sibling who graduates from a very competitive college has a 3.6 percent lower BMI in 1993 and a 6.1 percent lower BMI in 2004. While the sibling fixed effects specifications are our preferred results, we also present robustness checks using a variety of matching estimators in Appendix Tables 4 and 5; these results are smaller than the fixed effects results and suggest a reduction in BMI at age 54 of approximately 2 percent and a reduction in BMI at age 65 of approximately 3 percent.[21] As a result of the improvement of the comparability of the observed characteristics between selective college graduates and other college graduates in the family fixed effects specification and the robustness of the results to controlling for expected job status as an adult when in high school, the family fixed effects specifications are our preferred models. The increase in the magnitude of the college quality estimates from the OLS specification to the family fixed effects specification could reflect differential parental investments post college

[20] The results are qualitatively similar if BMI is measured in levels.

[21] We also examined the robustness of these results to selection on unobservables using the method developed by Altonji et al. (2005). Their method involves estimating a bivariate probit model for the binary outcomes with specific values for the correlation parameter. By allowing the degree of selection on observables to equal the degree of selection on unobservables and setting the correlation parameter equal to the amount of selection on observables, it is possible to examine the influence of selection on unobservables on the results. This method suggests that the potential bias due to selection on unobservables is negligible. However, the method is only slightly informative due to the low predictive power of the pre-college explanatory variables of the smoking and weight outcomes at old age.

between siblings that reinforce each sibling's college selectivity decision.[22] On the other hand, the increase, as opposed to a decrease, in the magnitude of the estimates could suggest that selection on unobservables is greater among sibling pairs that share the same environment and parents than among non-siblings with similar observable characteristics. Overall, though, this set of results suggests that graduating from a very competitive college is associated with at least a 2 percent reduction in BMI that increases as individuals age.

Tables 4 and 5 extend our analysis of later life weight outcomes to examine overweight and obesity status. In Table 4, the baseline (columns 1 and 4) results suggest a nearly 8 and 12 percentage point reduction in the probability of being overweight at ages 54 and 65, respectively. Controlling for childhood health, IQ, and measures of time preference in columns 2 and 5 does not influence the result at age 54 and suggests a decrease of 13 percentage points at age 65. In columns 3 and 6 we estimate sibling fixed effects models, which suggest a nearly 15 and 18 percentage point reduction in overweight status for the sibling who graduated from a more selective college. We also present robustness checks using a variety of matching estimators in Appendix Tables 4 and 5; these results suggest between a 8 and 10 percentage point reduction in overweight at age 54 and a 7 to 9 percentage point reduction in overweight at age 65.

In Table 5, the results for obesity are lower in magnitude and are not always statistically significant. However, the results suggest a 4 and 11 percentage point reduction in obesity in the baseline results in columns 1 and 4, which are virtually identical when additional controls are used in columns 2 and 5. In columns 3 and 6, our sibling fixed effects estimates suggest a 3 and 9 percentage point reduction in obesity for the sibling who graduated from a more selective college, however neither result is statistically significant. These finding are suggestive that

[22] As a way to provide suggestive evidence about this potential mechanism, we estimated models that adjusted for individual inheritance. Our results are robust to this specification.

college quality has an impact throughout the right-hand side of the body mass index distribution. We also present robustness checks using a variety of matching estimators in Appendix Tables 4 and 5; these results again suggest between a 3 and 4 percentage point reduction in obesity at age 54 and a 5 to 7 percentage point reduction in obesity at age 65. These findings for overweight and obese suggest that the impact of college quality increases with age, which is similar to the results of Webbink, Martin, and Visscher (2010) for years of schooling.

In Table 6, we examine tobacco use around age 54 and 65. Here we find that individuals who graduate from a very competitive college are between 4 (baseline) and 2 (fixed effects) percentage points less likely to smoke near age 54 than individuals who graduate from a less selective college, although the result is not statistically significant. For individuals near retirement age in columns 4-6, the estimate is near zero and not statistically significant, although the point estimate is a large fraction of the baseline rate of approximately 7%. We also present robustness checks using a variety of matching estimators in Appendix Tables 4 and 5; these results suggest that there is not a relationship between graduation from a selective college and smoking at age 54 and suggest a 2 to 4 percentage point reduction in obesity at age 65.

Mechanisms

The existing literature on the impact of college quality almost exclusively focuses on the impact on labor market outcomes (e.g., Long 2008). Thus, to begin our examination of the potential mechanisms that link graduating from a selective college and health behaviors later in life, we examine whether the additional income from college quality can explain the estimated impacts shown in Tables 3 through 6. In Table 7, we examine the impact of graduating from a selective college on total household income with a sibling fixed effects specification. Consistent

with the existing literature, our results show that college quality increases total household income, but only in 2004. The estimated increase in total household income is 36 percent at age 65 in 2004.

In Table 8, we examine whether the impact of college quality on health behaviors is due to the increase in income. To do this, we include log total household income in the preferred fixed effects specifications and examine the change in the estimate of the impact of college quality once this variable is included. The results suggest that total household income explains very little of the estimated impact of college quality on health.[23] [24]

In Tables 7 and 8, we also examine whether graduating from a selective college increases graduate school attendance and the years of schooling completed after college and whether these changes explain part of the relationship between graduating from a selective college and health behaviors. In Table 7, our results show that graduating from a selective college increases the probability of graduate school attendance by 16 percentage points and increase the years of schooling completed by 0.36 years. As shown in Table 8, controlling for graduate school attendance explains a small portion of the estimated impact of college quality. For example, the coefficient for graduating from a selective college changes from -0.061 to -0.055 for log BMI and from -0.176 to -0.162 for overweight in 2004.[25]

In Table 9, we seek to examine some additional potential mechanisms that may link graduating from a selective college and later life health. As graduating from a selective college could confer advantages in occupation, wealth, the marriage market, future health status, and

[23] Bratti and Miranda (in press) also find that income mediates a minor part of the relationship between education quantity and adult smoking.
[24] Since the average age is 64 years in 2004, a potential implication from Table 7 is that graduating from a very selective college induces individuals to work longer. Similar to total household income, we find that controlling for employment status has little impact on the results.
[25] Similar changes occur when controlling for years of schooling completed instead of a dummy variable for graduate school attendance or when controlling for years of schooling completed and graduate school attendance.

future cognition, we examine some of these potential mechanisms by controlling for characteristics and choices of the individuals that occurred between the time the individuals graduated from college and the time we measure health behaviors around age 65. We also further examine occupational characteristics and advantages, such as health insurance and access to care. Overall, the results in Table 9 suggest that the mechanisms we are able to examine do not explain the link between college selectivity and later health behaviors.[26] In unreported results, we also include measures of occupation-level smoking rates and obesity rates to examine whether broad "peer effects" may be an occupational mechanism but find no evidence of this effect.[27]

Although our results are quite robust in our inability to account for the effect of college quality on health in old age, one potential issue with sequentially adding covariates in an attempt to "explain" the college quality effects has recently been discussed by Gelbach (2009).[28] In unreported results, we find that in our case, the main conclusion that the mechanisms we examine are not able to explain the effects of college quality are invariant to using the approach in this paper in comparison to approach outlined in Gelbach (2009).[29]

Thus, while we provide evidence that largely rules out many of the potential mechanisms linking college quality and health behaviors in old age, several alternative mechanisms that we

[26] In contrast, Cutler and Lleras-Muney (2010) present suggestive evidence that they can "explain" up to 30% of the educational *quantity* gradient and health behaviors (smoking, heavy drinking, and obesity) with measures of resources (e.g. income) and up to 30% with measures of cognitive ability. However, they are unable to examine the relationship between educational *quality* and health behaviors.

[27] We calculate occupation-specific obesity rates within-sample so that the measure represents the proportion of WLS respondents (excluding the focal individual) who are obese in each occupation. We also use respondent's reports of whether they work with smokers at their current job (or last job for those who are retired).

[28] An additional potential issue is that the inclusion of possibly endogenous variables could bias the college quality estimate; however, based on the results shown in Table 9, this possibility does not seem likely in this context.

[29] Gelbach points out that sequential addition can lead to biased estimates of the impact when examining the robustness of the impact to the inclusion of additional covariates. He develops an alternative approach that provides consistent estimates of the influence of additional variables on the relationship of interest, but notes that this method may not be appropriate when the effect of the primary variable of interest operates through the additional variables. Gelbach also has created Stata code to implement his estimator "b1x2.ado", which we use: http://gelbach.eller.arizona.edu/papers/b1x2/

are unable to examine include: increases in health knowledge as a result of college quality, increases in productive efficiency, the impact of college peers on health behaviors, and other characteristics of occupation such as workplace policies.[30] We leave investigation of these potential channels to future work with alternative data.

Discussion and Conclusion

In this paper, we present some of the first evidence on the long term association between educational quality and health behaviors. We utilize a rich dataset that has followed respondents and their siblings for over fifty years, which allows us to control for an important set of pre-college factors as well as any experience or characteristic shared by siblings (50% of genetic endowments, family background, components of neighborhood and school factors experienced as a child).[31] Although we are not able to control for all sources of individual heterogeneity such as non-cognitive traits, this rich set of controls does not affect our principle findings—college selectivity is associated with a reduction in several measure of weight during old age. Specifically, we find a reduction of approximately 4 to 6 percent of BMI and a 15 to 18 percentage point reduction in overweight status. In contrast, graduation from a selective college is not associated with smoking. However, this finding is consistent with evidence that

[30] As we discussed above, when we examined the effects of *attending* a selective college rather than focusing on *graduating* from a selection college on health behaviors for the WLS respondents, we found smaller and insignificant effects. Although these results were not able to use sibling fixed effects because the survey does not contain attendance information of the siblings of the WLS respondents, we posit that the differential results for attendance versus graduation could suggest a limited role for some classes of mechanisms, such as exposure to college peers and changes in preferences/knowledge. That is, a discrete effect of college that appears at college graduation but not before may be inconsistent with the (presumably continuous) effects of college peers, preference formation, and knowledge acquisition during the college years. However, we note that this argument is necessarily speculative.

[31] Although a potential limitation is the external validity of this sample, which is drawn from a single state in 1957, as described above, this sample is broadly representative of white, non-Hispanic adults with at least a high school degree who were born during this period. Additionally, if the influences of college quality do not vary substantially across states, then results from this sample would be similar to results from a nationally representative sample (Jencks, Crouse, and Mueser, 1983). Further, we view the wealth of information available in the WLS that is not available in nationally representative data sets as a benefit that exceeds the costs of any reduced external validity.

individuals with greater educational attainment are more likely to quit smoking at earlier ages and in response to the 1964 Surgeon General Report on Smoking and Health (Sander, 1995; Aizer and Stroud, 2010). We then attempted to find potential mechanisms that may link the large effects of college quality with long term health behaviors, such as occupational characteristics (access to care, insurance coverage, prestige), income/wealth measures, marriage market outcomes, as well as later cognition and self-rated health status. Interestingly, we find no evidence that these typically hypothesized pathways explain any of the effects of college selectivity.

Although our focus is the long-run relationship between college quality and health behaviors, it is possible that the reduction in overweight during the 50s and 60s will translate to improvements in a variety of health outcomes. Obesity is associated with a higher number of chronic conditions and greater risk of mortality than normal weight; however, overweight is only associated with a greater number of chronic conditions for women and is not related to mortality risk (Sturm and Wells, 2001; Flegal et al., 2005). In unreported results, we examined the relationship between graduating from a selective college and self-reported health status. The results show that there is an improvement in self-reported health status and an increase in the likelihood of reporting excellent health that is statistically significant in regressions without family fixed effects, but that is not statistically significant after including fixed effects.

While our results focus on college quality from the 1960s, it is interesting to speculate on how these results might differ for more recent cohorts. For this question, we turn to recent research by Hoxby (2009), who demonstrates that although the selectivity of the top 10 percent of colleges has increased since the 1960s, the selectivity of all other colleges has remained the same or decreased during this time. This finding suggests that our results could be a lower

bound of the impact on long-run health behaviors of graduating from a selective college today. Finally, while the specific mechanisms linking college quality and health remain blurry, our results are likely the best evidence to date on the presence of substantial health returns to college quality.

References

Abadie, Alberto and Guido Imbens. 2002. "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects." NBER Technical Working Paper 283.

Aizer, Anna and Laura Stroud (2010) "Education, Knowledge and the Evolution of Disparities in Health," National Bureau of Economic Research Working Paper 15840.

Albouy, Valerie, and Laurent Lequien (2009). "Does Compulsory Education Lower Mortality?" *Journal of Health Economics* 28(1):155–68.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz (2011) "Personality Psychology and Economics," IZA Discussion Paper 5500.

Altonji, Joseph, Todd Elder, and Christopher Taber (2005) "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113 (1), 151-184.

Arendt, Jacob. (2005). "Does Education Cause Better Health? A Panel Data Analysis using School Reforms for Identification." Economics of Education Review, 24: 149-160

Behrman, Jere, Mark Rosenzweig, and Paul Taubman. (1996). "College Choice and Wages: Estimates Using Data on Female Twins." Review of Economics and Statistics, 78(4): 672-685

Berger, M. C., & Leigh, J. P. (1989). Education, self-selection and health. Journal of Human Resources, 24(3), 433–455.

Black, Dan and Jeffrey Smith. (2004). "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." Journal of Econometrics 121: 99-124

Bratti, Massimiliano and Alfonso Miranda. (in press). "Non-pecuniary Returns to Higher Education: The Effect of Smoking Intensity in the UK." *Health Economics*

Bound, J. & Solon, G. 1999. "Double trouble: on the value of twins-based estimation of the return to schooling." Economics of Education Review, 18: 169-182.

Case, Anne, Angela Fertig, and Christina Paxson (2005). "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24, 365-389.

Clark, Damon and Heather Royer (2010). "The Effect of Education on Adult Health and Mortality: Evidence from Britain." NBER Working Paper 16013.

Cockerman, W.C., A. Rutten, and T. Abel. (1997). "Conceptualizing Contemporary Health Lifestyles: Moving Beyond Weber." Sociological Quarterly, 38: 601-22

Conti, Gabriella, James Heckman, and Sergio Urzua (2010), "The education-health gradient," *American Economic Review, papers and proceedings*, 100:2, 234–8.

Cutler, David and Adriana Lleras-Muney (2008). "Education and Health; Evaluating Theories and Evidence." In Robert F. Schoeni, James S. House, George Kaplan and Harold Pollack (Eds.). Making Americans Healthier: Social and Economics Policy as Health Policy. New York: Russell Sage Foundation.

Cutler, David and Adriana Lleras-Muney (2010). "Understanding Differences in Health Behaviors by Education." Journal of Health Economics 29(1): 1-28.

Dale, Stacey and Alan Krueger. (2002). "Estimating the Payoff to Attending A More Selective College: An Application of Selection on Observables and Unobservables." Quarterly Journal of Economics

Fine, Benjamin (1969) Barron's Profiles of American Colleges, Woodbury, New York: Barron's Educational Series, Inc.

Flegal, Katherine M., Barry I. Graubard, David F. Williamson, and Mitchell H. Gail. (2005). "Excess Deaths Associated with Underweight, Overweight, and Obesity." Journal of the American Medical Association. 293: 1861-1867.

Fletcher, JM and DE Frisvold. (2009). "Higher Education and Health Investments: Does More Schooling Affect Preventive Care Use?" Journal of Human Capital. 3(2): 144-176.

Fletcher, JM and DE Frisvold. (2009). "College Selectivity and Young Adult Health Behaviors." Economics of Education Review, 30, 826-837.

Fletcher, JM and SF Lehrer (2009). "The Effects of Adolescent Health on Educational Outcomes: Causal Evidence using 'Genetic Lotteries' between Siblings." *Forum for Health Economics & Policy* 12 (2), Article 8

Frisvold, David and Ezra Golberstein (2011a). "School Quality and the Education-Health Relationship: Evidence from Blacks in Segregated Schools," Journal of Health Economics, 30, 1232-1245.

Frisvold, David and Ezra Golberstein (2011b). "The Effect of School Quality on Black-White Health Differences: Evidence from Southern Segregated Schools," working paper.

Garces, E., Thomas, D., Currie, J. (2002). Longer-term effects of Head Start. American Economic Review 9 (4), 999-1012.

Gelbach, J. (2009). When Do Covariates Matter? And Which Ones, and How Much? UA Department of Economics Working Paper #09-07.
http://gelbach.eller.arizona.edu/~gelbach/papers/

Grimard, Franque and Daniel Parent (2007). "Education and Smoking: Were Vietnam War Draft Avoiders also more Likely to Avoid Smoking?" *Journal of Health Economics* 26(5):896-926.

Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political Economy, 80(2), 223–255.

Grossman, M., & Kaestner, R. (1997). Effects of education on health. In J. Behrman, & N. Stancey (Eds.). The social benefits of education. Ann Arbor: The University of Michigan Press.

Hoekstra, Mark (2009). "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." Review of Economics and Statistics, 91(4): 717-724.

Hoxby, Caroline M. (2009). "The Changing Selectivity of American Colleges." National Bureau of Economic Research Working Paper 15446.

Ishida, H, S. Spilerman, and K.H. Su (1997). "Educational Credentials and Promotions Chances in Japanese and American Organizations." American Sociological Review, 62: 866-82

Jencks, C, J Crouse, and P Mueser (1983). "The Wisconsin Model of Status Attainment: A National Replication with Improved Measures of Ability and Aspiration." Sociology of Education, 56: 3-19.

Johnson, Rucker. (2010). "Long-Run Impacts of School Desegregation and School Quality on Adult Attainments." University of California, Berkeley Working Paper.

Jones, Andrew M., Nigel Rice and Pedro Rosa Dias (2011), "Long-Term Effects of School Quality on Health and Lifestyle: Evidence from Comprehensive Schooling Reforms in England", *Journal of Human Capital* Vol. 5, No. 3, pp. 342-376.

Kaestner, Robert and Michael Grossman. (2009). "Effects of Weight on Children's Educational Achievement." Economics of Education Review, 28(6), 651-661.

Kenkel, D. (1991). Health behavior, health knowledge, and education. Journal of Political Economy, 99(2), 287–305.

Kenkel, Donald. (2007). "The Evolution of the Schooling-Smoking Gradient." Cornell University Working Paper

Lleras-Muney, A. (2004). The relationship between education and adult mortality in the US. The Review of Economic Studies

Long, Mark C. (2008). "College Quality and Early Adult Outcomes." *Economics of Education Review*, 27(5): 588-602.

Mazumder, Bhashkar. (2007). "How Did Schooling Laws Improve Long-Term Health and Lower Mortality?" Federal Reserve Bank of Chicago Working Paper 2006-23

MacInnis, Bo. (2009) "Returns to School Quality on Elderly Cognition and Cognitive Aging" working paper.

Oreopoulos, Philip. (2007). "Do Dropouts Drop Out Too Soon? Wealth, Health, and Happiness from Compulsory Schooling." Journal of Public Economics, 91(11-12), 2213-2229.

Price, Joseph. 2008. "Parent-Child Quality Time: Does Birth Order Matter?" Journal of Human Resources, 43(1): 240-265.

Roberts, Brent W., Kate E. Walton, and Wolfgang Viechtbauer (2006) "Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies," Psychological Bulletin, 132(1), 1-25.

Rosenbaum, J.E. (1984). Career Mobility in a Corporate Hierarchy. New York: Academic Press

Ross, Catherine and John Mirowsky. (1999). "Refining the Association between Education and Health: The Effects of Quantity, Credential, and Selectivity." Demography, 36(4): 445-460

Sander, William (1995) "Schooling and Quitting Smoking," *Review of Economics and Statistics*, 77(1), 191-199.

Sturn, R. and KB Wells. (2001). "Does Obesity Contribute as Much to Morbidity as Poverty or Smoking?" Public Health 115: 229-235.

Van Kippersluis, Hans, Owen O'Donnell, and Eddy van Doorslaer (2011), "Long Run Returns to Education: Does Schooling lead to an Extended Old Age?", *Journal of Human Resources* 46(4): 695-721.

Webbink, Dinand, Nicholas G. Martin, Peter M. Visscher (2010). "Does Education Reduce the Probability of Being Overweight?" Journal of Health Economics.

Wisconsin Longitudinal Study. 2009. "Retention and Response Rates of Wisconsin Longitudinal Study 1957-2005." Available at http://www.ssc.wisc.edu/wlsresearch/documentation/retention/cor1004_retention.pdf.

Wolfe, Barbara and Jere Berhman. (1987), "Women's Schooling and Children's Health: Are the Effects Robust with Adult Sibling Control for the Women's Childhood Background?" *Journal of Health Economics*, 6.

Table 1: Descriptive Statistics of College Graduates

	College Graduates		College Graduates of a Very Competitive College		College Graduates of Less Selective College	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
BMI in 2004	27.159	(4.641)	26.688	(4.297)	27.407	(4.795)
Obese in 2004	0.259	(0.438)	0.220	(0.415)	0.279	(0.449)
Overweight in 2004	0.693	(0.461)	0.658	(0.475)	0.711	(0.453)
Smoke in 2004	0.077	(0.267)	0.069	(0.254)	0.082	(0.274)
BMI in 1993	26.168	(4.295)	25.846	(4.091)	26.344	(4.393)
Obese in 1993	0.190	(0.393)	0.163	(0.369)	0.206	(0.404)
Overweight in 1993	0.610	(0.488)	0.573	(0.495)	0.630	(0.483)
Smoke in 1993	0.107	(0.310)	0.099	(0.298)	0.112	(0.316)
Graduated from a Highly Competitive College	0.036	(0.187)	0.103	(0.304)	0.000	--
Graduated from a Very Competitive College	0.352	(0.478)	1.000	--	0.000	--
Graduated from a Competitive College	0.642	(0.479)	1.000	--	0.448	(0.497)
Mother's years of schooling in 1957	11.664	(2.869)	12.350	(2.858)	11.291	(2.806)
Father's years of schooling in 1957	11.420	(3.849)	12.439	(4.032)	10.867	(3.628)
Family income during high school ($10,000s)	4.905	(2.530)	5.715	(2.693)	4.466	(2.321)
Number of siblings	2.652	(2.030)	2.329	(1.733)	2.827	(2.155)
Female	0.430	(0.495)	0.359	(0.480)	0.468	(0.499)
Age in 2004	63.943	(3.870)	64.146	(3.770)	63.833	(3.920)
Birth order	2.198	(1.490)	1.984	(1.260)	2.314	(1.589)
Lived with both parents in high school	0.928	(0.258)	0.931	(0.254)	0.926	(0.261)
Graduate from Class of 1957	0.639	(0.480)	0.647	(0.478)	0.635	(0.482)
Self-reported childhood health status is poor or fair	0.035	(0.170)	0.039	(0.180)	0.032	(0.164)
Missed school for >= 1 month because of health	0.083	(0.254)	0.087	(0.258)	0.080	(0.252)
IQ score during high school	112.116	(12.933)	116.790	(12.222)	109.580	(12.600)
Planned to attend college at age 16	0.892	(0.310)	0.948	(0.223)	0.862	(0.345)
Total household income in 2004 ($10,000s)	9.715	(10.208)	11.371	(11.723)	8.816	(9.163)
Total household assets in 2004 ($100,000s)	9.549	(14.533)	12.511	(17.401)	7.942	(12.421)
Observations	3210		1129		2081	

Notes: The college graduates sample is the sample of individuals who graduated from college with non-missing values for college selectivity and either obesity or smoking in 2004 or 1993. All dollar values are converted to 2004 dollars using the Consumer Price Index for All Urban Consumers.

Sources: Wisconsin Longitudinal Study, Fine (1969)

Table 2: Descriptive Statistics of the Sibling Sample

	Sibling Sample		Discordant Sibling Pairs					
			Both Siblings		Very Competitive College Graduate		Less Selective College Graduate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
BMI in 2004	26.781	(4.430)	26.692	(4.584)	26.077	(4.132)	27.308	(4.941)
Obese in 2004	0.230	(0.421)	0.209	(0.408)	0.176	(0.383)	0.242	(0.431)
Overweight in 2004	0.674	(0.469)	0.659	(0.475)	0.604	(0.492)	0.714	(0.454)
Smoke in 2004	0.062	(0.242)	0.066	(0.249)	0.066	(0.250)	0.066	(0.250)
BMI in 1993	25.830	(4.022)	25.736	(4.102)	25.538	(4.064)	25.934	(4.152)
Obese in 1993	0.160	(0.367)	0.148	(0.356)	0.143	(0.352)	0.154	(0.363)
Overweight in 1993	0.585	(0.493)	0.593	(0.493)	0.560	(0.499)	0.626	(0.486)
Smoke in 1993	0.079	(0.270)	0.082	(0.276)	0.066	(0.250)	0.099	(0.300)
Graduated from a Highly Competitive College	0.032	(0.176)	0.033	(0.179)	0.066	(0.249)	0.000	--
Graduated from a Very Competitive College	0.391	(0.488)	0.500	(0.501)	1.000	--	0.000	--
Graduated from a Competitive College	0.649	(0.478)	0.692	(0.463)	1.000	--	0.385	(0.489)
Mother's years of schooling in 1957	12.042	(2.938)	12.209	(3.259)	12.209	(3.268)	12.209	(3.268)
Father's years of schooling in 1957	12.196	(4.008)	12.418	(3.885)	12.418	(3.896)	12.418	(3.896)
Family income during high school ($10,000s)	5.290	(2.735)	5.599	(2.723)	5.599	(2.731)	5.599	(2.731)
Number of siblings	2.581	(1.912)	2.319	(1.628)	2.319	(1.632)	2.319	(1.632)
Female	0.436	(0.496)	0.412	(0.494)	0.319	(0.469)	0.505	(0.503)
Age in 2004	64.068	(4.753)	64.308	(5.299)	65.099	(5.993)	63.516	(4.393)
Birth order	2.198	(1.451)	2.148	(1.474)	2.088	(1.575)	2.209	(1.370)
Lived with both parents in high school	0.945	(0.228)	0.967	(0.179)	0.967	(0.180)	0.967	(0.180)
Graduate from Class of 1957	0.500	(0.500)	0.500	(0.501)	0.451	(0.500)	0.549	(0.500)
Self-reported childhood health status is poor or fair	0.032	(0.176)	0.044	(0.206)	0.055	(0.229)	0.033	(0.180)
Missed school for >= 1 month because of health	0.107	(0.308)	0.099	(0.299)	0.121	(0.328)	0.078	(0.268)
IQ score during high school	113.494	(12.847)	113.800	(12.878)	117.396	(12.116)	110.205	(12.673)
Planned to attend college at age 16	0.923	(0.267)	0.929	(0.258)	0.945	(0.229)	0.912	(0.285)
Observations	530		182		91		91	

Notes: The sibling sample is the sample of individuals included in the college graduate sample from Table 1 with at least one sibling in the sample. The sample of discordant sibling pairs includes sibling pairs from the sibling sample in which one sibling graduated from a very competitive college and the other sibling graduated from a college that was not very competitive. Columns 3 and 4 show the means and standard deviations for both siblings in the sample of discordant sibling pairs. Columns 5 and 6 show the means and standard deviations for the sibling in the discordant pair who graduated from a very competitive college. Columns 7 and 8 show the means and standard deviations for the sibling in the discordant pair who graduated from a college that was not very competitive.
Sources: See Table 1.

Table 3: Estimates of the Impact of Graduating from a Selective College on Log BMI

	BMI in 1993			BMI in 2004		
Very Selective College Graduate	-0.023	-0.022	-0.036*	-0.047***	-0.047***	-0.061***
	(0.014)	(0.015)	(0.020)	(0.015)	(0.016)	(0.020)
Mother's Education	-0.004	-0.004		-0.001	-0.001	
	(0.003)	(0.003)		(0.003)	(0.003)	
Father's Education	-0.002	-0.001		-0.001	-0.001	
	(0.002)	(0.002)		(0.002)	(0.002)	
Family Income	0.001	0.001		-0.002	-0.002	
	(0.003)	(0.003)		(0.003)	(0.003)	
Number of Siblings	0.006	0.007		0.004	0.004	
	(0.005)	(0.005)		(0.005)	(0.005)	
Female	-0.079***	-0.077***	-0.074***	-0.065***	-0.064***	-0.065***
	(0.013)	(0.014)	(0.017)	(0.014)	(0.015)	(0.018)
Age	0.000	-0.000	0.006**	-0.002	-0.002	0.004
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Birth Order	-0.007	-0.007	0.016	-0.005	-0.005	0.015
	(0.006)	(0.006)	(0.010)	(0.007)	(0.007)	(0.010)
Live with Both Parents	-0.066**	-0.065**	-0.072	-0.022	-0.021	-0.018
	(0.032)	(0.032)	(0.066)	(0.031)	(0.030)	(0.056)
WLS graduate	-0.018	-0.013	-0.004	-0.015	-0.014	0.013
	(0.012)	(0.018)	(0.021)	(0.012)	(0.018)	(0.021)
Poor Childhood Health		-0.061*	-0.069		-0.025	-0.026
		(0.037)	(0.043)		(0.038)	(0.035)
Missed School as Child		0.004	-0.009		0.001	-0.010
		(0.020)	(0.026)		(0.022)	(0.028)
IQ		-0.000	-0.000		0.000	0.000
		(0.001)	(0.001)		(0.001)	(0.001)
Plan to Attend College		-0.014	-0.003		-0.028	-0.002
		(0.023)	(0.033)		(0.026)	(0.039)
Constant	3.412***	3.458***	2.980***	3.499***	3.517***	3.035***
	(0.120)	(0.146)	(0.224)	(0.123)	(0.150)	(0.217)
Observations	530	530	530	530	530	530
R-squared	0.099	0.111	0.111	0.066	0.071	0.093
Fixed Effects?	No	No	Yes	No	No	Yes

Notes: Heteroskedasticity-robust standard errors that allow for clustering within families in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

Sources: See Table 1.

Table 4: Estimates of the Impact of Graduating from a Selective College on Overweight

	Overweight in 1993			Overweight in 2004		
Very Selective College Graduate	-0.077*	-0.078*	-0.145**	-0.120***	-0.130***	-0.176**
	(0.044)	(0.045)	(0.063)	(0.044)	(0.046)	(0.068)
Mother's Education	-0.009	-0.009		-0.005	-0.004	
	(0.009)	(0.009)		(0.008)	(0.008)	
Father's Education	-0.007	-0.006		0.006	0.006	
	(0.007)	(0.007)		(0.006)	(0.006)	
Family Income	-0.002	-0.003		-0.005	-0.006	
	(0.010)	(0.010)		(0.009)	(0.009)	
Number of Siblings	0.002	0.004		0.027*	0.027*	
	(0.015)	(0.015)		(0.015)	(0.015)	
Female	-0.253***	-0.249***	-0.246***	-0.238***	-0.240***	-0.240***
	(0.041)	(0.042)	(0.055)	(0.041)	(0.041)	(0.058)
Age	0.002	0.001	0.020**	0.002	0.002	0.019*
	(0.005)	(0.005)	(0.009)	(0.005)	(0.005)	(0.010)
Birth Order	-0.007	-0.007	0.058*	-0.007	-0.006	0.037
	(0.021)	(0.021)	(0.034)	(0.020)	(0.020)	(0.036)
Live with Both Parents	-0.068	-0.066	0.066	-0.044	-0.041	-0.217
	(0.080)	(0.077)	(0.217)	(0.076)	(0.075)	(0.168)
WLS graduate	-0.043	0.003	0.049	0.005	0.004	0.067
	(0.040)	(0.061)	(0.073)	(0.038)	(0.058)	(0.069)
Poor Childhood Health		-0.259**	-0.386**		0.033	0.047
		(0.124)	(0.151)		(0.119)	(0.169)
Missed School as Child		-0.014	0.006		-0.034	-0.040
		(0.072)	(0.098)		(0.069)	(0.103)
IQ		0.000	0.002		0.002	0.000
		(0.002)	(0.003)		(0.002)	(0.002)
Plan to Attend College		-0.002	-0.039		-0.002	0.010
		(0.078)	(0.103)		(0.075)	(0.095)
Constant	0.917**	0.889**	-0.932	0.708**	0.489	-0.331
	(0.381)	(0.450)	(0.742)	(0.358)	(0.414)	(0.735)
Observations	530	530	530	530	530	530
R-squared	0.083	0.099	0.127	0.089	0.093	0.104
Fixed Effects?	No	No	Yes	No	No	Yes

Notes: See Table 3.
*** p<0.01, ** p<0.05, * p<0.1.
Sources: See Table 1.

Table 5: Estimates of the Impact of Graduating from a Selective College on Obesity

	Obese in 1993			Obese in 2004		
Very Selective College Graduate	-0.039	-0.041	-0.029	-0.107***	-0.104**	-0.087
	(0.036)	(0.038)	(0.048)	(0.038)	(0.041)	(0.058)
Mother's Education	-0.005	-0.005		0.004	0.003	
	(0.008)	(0.008)		(0.007)	(0.008)	
Father's Education	-0.005	-0.005		-0.004	-0.003	
	(0.006)	(0.006)		(0.006)	(0.007)	
Family Income	-0.000	-0.001		-0.011	-0.011	
	(0.007)	(0.007)		(0.008)	(0.008)	
Number of Siblings	0.020	0.019		0.011	0.011	
	(0.012)	(0.012)		(0.013)	(0.013)	
Female	-0.016	-0.016	0.030	-0.052	-0.050	-0.074
	(0.034)	(0.034)	(0.043)	(0.039)	(0.039)	(0.051)
Age	0.001	0.001	0.013**	-0.010**	-0.011**	-0.001
	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.008)
Birth Order	-0.013	-0.011	0.033	-0.015	-0.014	0.019
	(0.019)	(0.018)	(0.021)	(0.019)	(0.020)	(0.027)
Live with Both Parents	-0.147	-0.153	-0.229*	0.016	0.016	0.105
	(0.102)	(0.101)	(0.120)	(0.084)	(0.085)	(0.159)
WLS graduate	-0.050*	-0.010	0.003	-0.036	-0.030	0.005
	(0.028)	(0.046)	(0.057)	(0.033)	(0.052)	(0.063)
Poor Childhood Health		-0.075	-0.104		-0.063	-0.065
		(0.089)	(0.071)		(0.099)	(0.076)
Missed School as Child		0.064	0.089		0.052	0.023
		(0.055)	(0.072)		(0.060)	(0.075)
IQ		0.000	0.001		-0.001	0.000
		(0.001)	(0.002)		(0.002)	(0.002)
Plan to Attend College		-0.053	-0.097		-0.067	-0.020
		(0.072)	(0.088)		(0.083)	(0.106)
Constant	0.408	0.362	-0.549	1.032***	1.201***	0.148
	(0.316)	(0.401)	(0.484)	(0.352)	(0.453)	(0.626)
Observations	530	530	530	530	530	530
R-squared	0.031	0.041	0.048	0.044	0.054	0.031
Fixed Effects?	No	No	Yes	No	No	Yes

Notes: See Table 3.
*** p<0.01, ** p<0.05, * p<0.1.
Sources: See Table 1.

Table 6: Estimates of the Impact of Graduating from a Selective College on Smoking

	Smoking in 1993			Smoking in 2004		
Very Selective College Graduate	-0.039	-0.044	-0.021	-0.012	-0.018	0.008
	(0.029)	(0.030)	(0.043)	(0.023)	(0.025)	(0.036)
Mother's Education	-0.003	-0.003		-0.004	-0.003	
	(0.005)	(0.005)		(0.005)	(0.005)	
Father's Education	-0.003	-0.004		-0.001	-0.001	
	(0.003)	(0.004)		(0.003)	(0.003)	
Family Income	0.012**	0.012**		0.012**	0.011**	
	(0.006)	(0.006)		(0.005)	(0.005)	
Number of Siblings	0.002	0.002		0.005	0.005	
	(0.007)	(0.007)		(0.007)	(0.007)	
Female	0.021	0.019	0.047	0.001	0.000	0.023
	(0.024)	(0.025)	(0.040)	(0.022)	(0.023)	(0.040)
Age	-0.002	-0.001	0.002	-0.001	-0.000	0.001
	(0.003)	(0.003)	(0.007)	(0.002)	(0.002)	(0.007)
Birth Order	-0.006	-0.006	0.000	0.002	0.004	0.009
	(0.012)	(0.012)	(0.028)	(0.010)	(0.011)	(0.028)
Live with Both Parents	-0.060	-0.061	-0.096	0.000	0.006	-0.093
	(0.060)	(0.060)	(0.159)	(0.037)	(0.038)	(0.086)
WLS graduate	0.005	0.009	-0.000	-0.003	0.027	0.006
	(0.023)	(0.032)	(0.042)	(0.021)	(0.028)	(0.031)
Poor Childhood Health		-0.027	0.015		0.004	0.008
		(0.065)	(0.092)		(0.061)	(0.089)
Missed School as Child		-0.017	-0.033		-0.044*	-0.021
		(0.037)	(0.058)		(0.024)	(0.033)
IQ		0.000	-0.000		0.000	-0.000
		(0.001)	(0.001)		(0.001)	(0.002)
Plan to Attend College		0.060*	0.041		0.057***	0.020
		(0.032)	(0.055)		(0.017)	(0.035)
Constant	0.269	0.140	0.029	0.100	-0.049	0.016
	(0.200)	(0.264)	(0.519)	(0.166)	(0.207)	(0.512)
Observations	530	530	530	530	530	530
R-squared	0.020	0.027	0.024	0.018	0.032	0.015
Fixed Effects?	No	No	Yes	No	No	Yes

Notes: See Table 3.
*** p<0.01, ** p<0.05, * p<0.1.
Sources: See Table 1.

Table 7: Family Fixed Effects Estimates of the Impact of Graduating from a Selective College on Total Household Income and Graduate School Attendance

	Log Total Household Income in 1993	Log Total Household Income in 2004	Graduate School Attendance	Years of Schooling Completed
Very Selective College	-0.030	0.363**	0.160**	0.362*
	(0.192)	(0.165)	(0.075)	(0.194)
Observations	530	530	530	530

Notes: Heteroskedasticity-robust standard errors that allow for clustering within families in parentheses. Additional explanatory variables that are not shown include sex, age, birth order, whether the individual lived with both parents during high school, a dummy variable indicating whether the individual is an original WLS graduate (vs. a sibling), an indicator variable for poor childhood self-reported health status, an indicator variable for missing at least one month of school as a child due to health problems, IQ, whether the individual planned to attend college at age 16, and family fixed effects.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Sources: See Table 1.

Table 8: Family Fixed Effects Estimates of the Impact of Graduating from a Selective College on Health Behaviors Controlling for Income and Graduate School Attendance

	Log BMI			Obese			Overweight			Smoking		
Panel A: 1993 Outcomes												
Very Selective College	-0.036*	-0.036*	-0.031	-0.029	-0.028	-0.022	-0.145**	-0.144**	-0.124**	-0.021	-0.021	-0.014
	(0.020)	(0.020)	(0.020)	(0.048)	(0.048)	(0.048)	(0.063)	(0.063)	(0.062)	(0.043)	(0.043)	(0.041)
Log Household Income		0.003			0.016			0.026			-0.006	
		(0.007)			(0.019)			(0.022)			(0.014)	
Graduate School Attendance			-0.027			-0.047			-0.129**			-0.044
			(0.017)			(0.041)			(0.055)			(0.033)
Observations	530	530	530	530	530	530	530	530	530	530	530	530
R^2	0.111	0.112	0.120	0.048	0.051	0.053	0.127	0.131	0.145	0.024	0.025	0.031
Panel B: 2004 Outcomes												
Very Selective College	-0.061***	-0.064***	-0.055***	-0.087	-0.086	-0.067	-0.176**	-0.190***	-0.162**	0.008	0.009	0.007
	(0.020)	(0.021)	(0.021)	(0.058)	(0.058)	(0.059)	(0.068)	(0.068)	(0.067)	(0.036)	(0.036)	(0.035)
Log Household Income		0.008			-0.001			0.038*			-0.002	
		(0.006)			(0.019)			(0.020)			(0.009)	
Graduate School Attendance			-0.038**			-0.123**			-0.089			0.010
			(0.019)			(0.049)			(0.054)			(0.028)
Observations	530	530	530	530	530	530	530	530	530	530	530	530
R^2	0.093	0.097	0.110	0.031	0.031	0.055	0.104	0.114	0.113	0.015	0.015	0.015

Notes: See Table 7.

*** p<0.01, ** p<0.05, * p<0.1.

Sources: See Table 1.

Table 9: Examining the Mechanisms of the Impact of Graduating from a Selective College on Health Behaviors in 2004

Outcome:	Obese	Overweight	Log BMI	Smoke
Baseline	-0.087	-0.176**	-0.061***	0.008
	(0.058)	(0.068)	(0.020)	(0.036)
Baseline with prior Health Insurance types	-0.083	-0.183***	-0.060***	0.003
	(0.057)	(0.069)	(0.020)	(0.036)
Baseline with Access to Care	-0.091	-0.173**	-0.063***	0.006
	(0.059)	(0.068)	(0.020)	(0.036)
Baseline with Marital Status	-0.092	-0.175**	-0.061***	0.007
	(0.058)	(0.068)	(0.020)	(0.038)
Baseline with Cognition	-0.087	-0.171**	-0.061***	0.009
	(0.058)	(0.069)	(0.020)	(0.035)
Baseline with prior Self Reported Health	-0.090	-0.184***	-0.064***	0.004
	(0.059)	(0.068)	(0.021)	(0.035)
Baseline with Occupational Prestige	-0.087	-0.174**	-0.061***	0.011
	(0.058)	(0.069)	(0.021)	(0.036)
plus prior Health Insurance types	-0.082	-0.180**	-0.058***	0.006
	(0.057)	(0.070)	(0.020)	(0.036)
plus Access to Care	-0.087	-0.176**	-0.060***	0.003
	(0.058)	(0.069)	(0.020)	(0.036)
plus Marital Status	-0.089	-0.174**	-0.060***	0.002
	(0.059)	(0.069)	(0.020)	(0.038)
plus Cognition	-0.090	-0.170**	-0.060***	0.003
	(0.059)	(0.071)	(0.020)	(0.038)
plus Total Household Income	-0.090	-0.186***	-0.063***	0.001
	(0.059)	(0.070)	(0.020)	(0.038)
plus Assets	-0.093	-0.179**	-0.061***	-0.003
	(0.059)	(0.071)	(0.020)	(0.039)
Plus Graduate School Attendance	-0.077	-0.163**	-0.055***	-0.005
	(0.060)	(0.070)	(0.020)	(0.039)
plus prior Self Reported Health	-0.081	-0.180***	-0.063***	-0.016
	(0.061)	(0.069)	(0.020)	(0.039)
Observations	530	530	530	530

Notes: Heteroskedasticity-robust standard errors that allow for clustering within families in parentheses. Each cell in this table is the estimated coefficient for graduating from a very competitive college from separate regressions that also control for sex, age, birth order, whether the individual lived with both parents during high school, a dummy variable indicating whether the individual is in the graduate sample (vs. the sibling sample), an indicator variable for poor childhood self-reported health status, an indicator variable for missing at least one month of school as a child due to health problems, IQ, whether the individual planned to attend college at age 16, and family fixed effects. The baseline results are the fixed effects results shown in Tables 3 through 6. Each row that begins with the title "plus" report estimates that add the denoted variable to the specification estimated in the row directly above. The categories of health insurance type are employer-provided, privately-purchased, other insurance, and no insurance. The access to care variables are the three measures: difficulty obtaining health care, access to health care satisfaction, and usual source of care. The variables denoted as "prior" are measured in the prior survey wave in 1993.

*** p<0.01, ** p<0.05, * p<0.1

Sources: See Table 1.

Appendix Table 1: Summary Statistics of Respondents Excluded from the Analysis Sample

	Full WLS Sample			Respondents who Did Not Graduate from College			Respondents with Missing Health Outcomes			Respondents without Siblings		
Variable	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.
BMI in 1993	26.768	(4.589)	10652	26.987	(4.673)	7796	0.000	--		26.213	(4.280)	1858
Obese in 1993	0.234	(0.423)	10652	0.249	(0.433)	7796	0.000	--		0.192	(0.394)	1858
Overweight in 1993	0.658	(0.474)	10652	0.676	(0.468)	7796	0.000	--		0.615	(0.487)	1858
Smoke in 1993	0.175	(0.380)	10759	0.200	(0.400)	7882	0.000	--		0.105	(0.307)	1861
BMI in 2004	27.875	(4.987)	9412	28.145	(5.085)	6833	0.000	--		27.237	(4.716)	1794
Obese in 2004	0.312	(0.464)	9412	0.333	(0.471)	6833	0.000	--		0.265	(0.442)	1794
Overweight in 2004	0.736	(0.441)	9412	0.752	(0.432)	6833	0.000	--		0.698	(0.459)	1794
Smoke in 2004	0.128	(0.335)	10618	0.146	(0.353)	7901	0.000	--		0.084	(0.278)	1830
Graduated from a Very Competitive College	0.354	(0.478)	3708	0.000	--		0.371	(0.484)	498	0.338	(0.473)	1926
Mother's years of schooling in 1957	10.470	(2.812)	20296	10.200	(2.725)	16596	11.775	(2.905)	498	11.581	(2.806)	1921
Father's years of schooling in 1957	9.791	(3.396)	19938	9.409	(3.155)	16271	11.906	(3.952)	490	11.346	(3.809)	1902
Family income during high school ($10,000s)	3.948	(2.187)	19310	3.727	(2.028)	15737	4.969	(2.566)	478	4.791	(2.468)	1856
Number of siblings	3.247	(2.572)	19126	3.395	(2.664)	15423	2.499	(2.047)	495	2.646	(2.033)	1924
Female	0.506	(0.500)	19095	0.527	(0.499)	15387	0.355	(0.479)	498	0.447	(0.497)	1926
Age in 2004	64.202	(4.365)	11536	64.317	(4.448)	8503	62.885	(4.789)	174	63.964	(3.711)	1852
Birth order	2.580	(1.843)	18334	2.673	(1.907)	14631	2.315	(1.635)	495	2.177	(1.484)	1924
Lived with both parents in high school	0.909	(0.288)	14402	0.903	(0.295)	10762	0.904	(0.295)	446	0.919	(0.272)	1922
Self-reported childhood health status is poor or fair	0.037	(0.189)	10773	0.037	(0.190)	8027	0.250	(0.463)	8	0.035	(0.184)	1829
Missed school for >= 1 month because of health	0.081	(0.273)	10661	0.081	(0.272)	7940	0.000	(0.000)	7	0.083	(0.275)	1814
IQ score during high school	101.001	(15.317)	16936	98.204	(14.530)	13434	109.902	(14.166)	468	112.057	(12.994)	1842
Planned to attend college at age 16	0.423	(0.494)	11219	0.284	(0.451)	8473	0.861	(0.346)	382	0.855	(0.352)	1465

Notes: This table the characteristics of the WLS respondents who were excluded from the analysis sample according to the reason that the sample was restricted.
Sources: See Table 1.

Appendix Table 2: Family Fixed Effects Estimates of the Impact of Graduating from a Selective College Using Inverse Probability Weights to Account for Attrition

	Log BMI		Obese		Overweight		Smoking	
	Without Weights	With Weights	Without Weights	With Weights	Without Weights	With Weights	Without Weights	With Weights
Panel A: 1993 Outcomes								
Very Selective College	-0.036*	-0.036*	-0.029	-0.030	-0.147**	-0.146**	-0.021	-0.021
	(0.020)	(0.020)	(0.048)	(0.047)	(0.063)	(0.063)	(0.043)	(0.042)
Observations	530	530	530	530	530	530	530	530
R^2	0.111	0.111	0.048	0.048	0.127	0.127	0.025	0.025
Panel B: 2004 Outcomes								
Very Selective College	-0.061***	-0.062***	-0.087	-0.087	-0.176**	-0.178***	0.008	0.008
	(0.020)	(0.020)	(0.058)	(0.058)	(0.068)	(0.068)	(0.036)	(0.035)
Observations	530	530	530	530	530	530	530	530
R^2	0.093	0.093	0.031	0.031	0.104	0.104	0.015	0.015

Notes: Inverse probability weights are calculated as the inverse of the predicted probability that the respondent is included in the sample described in Table 1. The predicted probabilities are calculated using a probit regression that controls for all individual and family characteristics displayed in columns 2 and 5 of Tables 3, 4, 5, and 6.

*** p<0.01, ** p<0.05, * p<0.1.

Sources: See Table 1.

Appendix Table 3: Estimates of the Impact of Attending a Selective College without Family Fixed Effects for the Sample of Original WLS Respondents

	Log BMI		Obese		Overweight		Smoking	
	1993	2004	1993	2004	1993	2004	1993	2004
Panel A: Original WLS Respondents in the College Graduates Sample								
Attending a Very Competitive College	-0.007	-0.007	-0.010	-0.021	-0.041	-0.031	-0.011	-0.025*
	(0.008)	(0.009)	(0.021)	(0.025)	(0.025)	(0.026)	(0.017)	(0.014)
Observations	1837	1690	1837	1690	1837	1690	1851	1761
R2	0.076	0.040	0.026	0.018	0.097	0.052	0.024	0.014
Panel B: Original WLS Respondents in the Sibling Sample								
Attending a Very Competitive College	-0.017	-0.027	-0.074	-0.075	-0.005	-0.026	-0.053	-0.027
	(0.018)	(0.021)	(0.047)	(0.058)	(0.065)	(0.064)	(0.037)	(0.034)
Observations	265	265	265	265	265	265	265	265
R2	0.200	0.098	0.071	0.048	0.150	0.110	0.046	0.090
Panel C: Original WLS Respondents in the College Graduates Sample without a Sibling								
Attending a Very Competitive College	-0.007	-0.004	-0.003	-0.015	-0.049*	-0.033	-0.002	-0.023
	(0.009)	(0.010)	(0.023)	(0.027)	(0.027)	(0.028)	(0.019)	(0.015)
Observations	1572	1425	1572	1425	1572	1425	1586	1496
R2	0.067	0.035	0.025	0.020	0.095	0.047	0.028	0.016
Panel D: Original WLS Respondents in the College Graduates Sample								
Graduating from a Very Competitive College	-0.020**	-0.021**	-0.049**	-0.043*	-0.066***	-0.054**	-0.019	-0.013
	(0.008)	(0.009)	(0.020)	(0.024)	(0.025)	(0.025)	(0.016)	(0.014)
Observations	1837	1690	1837	1690	1837	1690	1851	1761
R2	0.078	0.043	0.028	0.019	0.099	0.054	0.024	0.012
Panel E: Original WLS Respondents in the Sibling Sample								
Graduating from a Very Competitive College	-0.019	-0.032	-0.049	-0.098*	-0.023	-0.061	-0.053	-0.020
	(0.019)	(0.021)	(0.050)	(0.058)	(0.066)	(0.063)	(0.038)	(0.034)
Observations	265	265	265	265	265	265	265	265
R2	0.201	0.100	0.066	0.053	0.151	0.113	0.046	0.089
Panel F: Original WLS Respondents in the College Graduates Sample without a Sibling								
Graduating from a Very Competitive College	-0.020**	-0.020**	-0.051**	-0.034	-0.075***	-0.052*	-0.012	-0.010
	(0.009)	(0.010)	(0.022)	(0.027)	(0.027)	(0.028)	(0.018)	(0.015)
Observations	1572	1425	1572	1425	1572	1425	1586	1496
R2	0.069	0.037	0.028	0.021	0.097	0.048	0.028	0.015

Notes: Heteroskedasticity-robust standard errors that allow for clustering within families in parentheses for those with siblings. Additional explanatory variables that are not shown include sex, age, birth order, whether the individual lived with both parents during high school, a dummy variable indicating whether the individual is in the graduate sample (vs. the sibling sample) if appropriate, an indicator variable for poor childhood self-reported health status, an indicator variable for missing at least one month of school as a child due to health problems, IQ, and whether the individual planned to attend college at age 16.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Sources: See Table 1.

Appendix Table 4: Matching Estimates of the Impact of Graduating from a Selective College on Health Behaviors in 1994

	Log BMI	Overweight	Obese	Smoking
Sample	Common Support	Common Support	Common Support	Common Support
Nearest Neighbor Matching (bias-corrected)	-0.023***	-0.090***	-0.043**	-0.033**
	(0.007)	(0.023)	(0.020)	(0.017)
Stratification Matching	-0.023***	-0.080***	-0.040**	-0.016
	(0.007)	(0.021)	(0.019)	(0.014)
Kernel-Based Matching	-0.023***	-0.079***	-0.045**	-0.019
	(0.007)	(0.024)	(0.018)	(0.014)
Sample Size	2848	2848	2848	2873
Sample	Trimmed	Trimmed	Trimmed	Trimmed
Nearest Neighbor Matching (bias-corrected)	-0.021***	-0.087***	-0.035*	-0.034**
	(0.008)	(0.024)	(0.020)	(0.017)
Stratification Matching	-0.023***	-0.083***	-0.040**	-0.017
	(0.007)	(0.024)	(0.017)	(0.014)
Kernel-Based Matching	-0.023***	-0.082***	-0.044**	-0.020
	(0.007)	(0.021)	(0.019)	(0.013)
Sample Size	2629	2629	2629	2651
Sample	Thick Support	Thick Support	Thick Support	Thick Support
Nearest Neighbor Matching (bias-corrected)	-0.020**	-0.100***	0.019	-0.014
	(0.010)	(0.032)	(0.028)	(0.022)
Stratification Matching	-0.021**	-0.081***	-0.028	-0.012
	(0.009)	(0.029)	(0.022)	(0.020)
Kernel-Based Matching	-0.023**	-0.081***	-0.031	-0.012
	(0.010)	(0.028)	(0.021)	(0.019)
Sample Size	1185	1185	1185	1193

Notes: Heteroskedasticity-robust standard errors in parentheses. Standard errors are bootstrapped for stratification and kernel-based matching. The sample in the top panel includes individuals from the full sample in the common support. The sample in the middle panel is the subset of individuals from the full sample in the common support with propensity scores in the range [0.1, 0.9]. The sample in the bottom panel is the subset of individuals from the full sample in the common support with propensity scores in the range [0.33, 0.67]. The first row of estimates is based on the bias-adjusted nearest neighbor matching with replacement estimator developed by Abadie and Imbens (2002). These estimates are based on a minimum of three matches per observation; similar results are obtained using a minimum of two and four matches per observation.

*** p<0.01, ** p<0.05, * p<0.1

Sources: See Table 1.

Appendix Table 5: Matching Estimates of the Impact of Graduating from a Selective College on Health Behaviors in 2004

	Log BMI	Overweight	Obese	Smoking
Sample	Common Support	Common Support	Common Support	Common Support
Nearest Neighbor Matching (bias-corrected)	-0.027***	-0.081***	-0.055**	-0.039***
	(0.008)	(0.024)	(0.023)	(0.015)
Stratification Matching	-0.029***	-0.075***	-0.052***	-0.024**
	(0.007)	(0.025)	(0.019)	(0.012)
Kernel-Based Matching	-0.031***	-0.077***	-0.059***	-0.025*
	(0.008)	(0.020)	(0.021)	(0.014)
Sample Size	2576	2576	2576	2713
Sample	Trimmed	Trimmed	Trimmed	Trimmed
Nearest Neighbor Matching (bias-corrected)	-0.027***	-0.081***	-0.055**	-0.037**
	(0.009)	(0.024)	(0.023)	(0.015)
Stratification Matching	-0.030***	-0.078***	-0.054***	-0.023*
	(0.008)	(0.019)	(0.021)	(0.014)
Kernel-Based Matching	-0.031***	-0.078***	-0.061***	-0.024*
	(0.008)	(0.024)	(0.020)	(0.013)
Sample Size	2346	2346	2346	2476
Sample	Thick Support	Thick Support	Thick Support	Thick Support
Nearest Neighbor Matching (bias-corrected)	-0.017	-0.075**	-0.052	-0.033
	(0.012)	(0.033)	(0.032)	(0.021)
Stratification Matching	-0.028***	-0.088***	-0.066**	-0.024
	(0.011)	(0.031)	(0.027)	(0.020)
Kernel-Based Matching	-0.029***	-0.092***	-0.068**	-0.022
	(0.009)	(0.028)	(0.026)	(0.018)
Sample Size	1048	1048	1048	1111

Notes: See Appendix Table 4.
*** p<0.01, ** p<0.05, * p<0.1
Sources: See Table 1.