

NBER WORKING PAPER SERIES

EDUCATION, COGNITION, HEALTH KNOWLEDGE, AND HEALTH BEHAVIOR

Naci H. Mocan
Duha Tore Altindag

Working Paper 17949
<http://www.nber.org/papers/w17949>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2012

The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Naci H. Mocan and Duha Tore Altindag. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Education, Cognition, Health Knowledge, and Health Behavior
Naci H. Mocan and Duha Tore Altindag
NBER Working Paper No. 17949
March 2012
JEL No. I12,I20

ABSTRACT

Using data from the NLSY97 we analyze the impact of education on health behaviors, measured by smoking and heavy drinking. Controlling for health knowledge does not influence the impact of education on health behaviors, supporting the productive efficiency hypothesis. Although cognition, as measured by test scores, appears to have an effect on the relationship between education and health behaviors, this effect disappears once the models control for family fixed effects. Similarly, the impact of education on smoking and heavy drinking is the same between those with and without a learning disability, suggesting that cognition is not likely to be a significant factor in explaining the impact of education on health behaviors.

Naci H. Mocan
Department of Economics
Louisiana State University
2119 Patrick F. Taylor Hall
Baton Rouge, LA 70803-6306
and NBER
mocan@lsu.edu

Duha Tore Altindag
Auburn University
Department of Economics
0334 Haley Center
Auburn AL, 36849
altindag@auburn.edu

Education, Cognition, Health Knowledge, and Health Behavior

I. Introduction

Schooling impacts health outcomes. More educated people are healthier than the less educated (Grossman 2006, 2008). This positive relationship between education and health is robust whether one analyses aggregates (e.g. mortality or morbidity rates), or micro units (e.g. individuals' self-reported health status, or sick days).

If the effect of education on health is causal, then the impact of education on individual well-being is pronounced. For example, it is well-established that education raises wages (Card, 2000). It is also documented that an improvement in health is associated with increased labor productivity, and that an improvement in health outcomes of a given generation produces an improvement in health of their offspring (see Currie, 2011 and the literature she cites). This means that an increase in education not only has a direct positive impact on the earnings of the individual, but it also has an additional effect on productivity and earnings through an improvement in health. These increases in earnings improve the well-being of the individual in addition to the increase in utility generated by enhanced health. Improved education and health also have an impact on the level of education and health of the individual's children, transmitting the benefit of enhanced education to the second generation (Currie and Moretti, 2003; Sacerdote, 2002).

In standard models of health production, schooling has a causal impact on health because schooling increases the efficiency of health production (Grossman 1972, 1975). An alternative hypothesis, which is also consistent with the observed positive relationship between schooling and health, is that of the allocative efficiency. According to this hypothesis, more educated

individuals choose input allocations that produce more output (better health) than those who have less education (see Rosenzweig and Schultz 1982, and the papers discussed in Grossman 2006). Under allocative efficiency, education expands individuals' knowledge base about health, and an increase in health knowledge alters health behaviors (i.e., consumption of health inputs with both positive and negative marginal products, such as medical care and cigarette smoking), which in turn influence health outcomes.

This paper has two aims. First, it investigates whether education has an impact on input allocation through its impact on health knowledge. Specifically, we employ, for the first time in this literature, a panel data set to analyze the validity of the allocative efficiency hypothesis. Our basic framework is similar to Kenkel (1991), where the impact of schooling on health inputs is estimated. If the influence of schooling on health is working through allocative efficiency (i.e., if schooling improves allocative efficiency by increasing the health knowledge of the individual), schooling should have little or no direct effect on health inputs in a regression that controls for health knowledge. Kenkel (1991) uses cross-sectional data from the 1985 National Health Interview Survey (NHIS) and focuses on health inputs (behavior) such as smoking, drinking and exercise. His data set also contains information about the knowledge of the subjects regarding the health consequences of smoking, drinking and exercise. He finds that inclusion or exclusion of measures of health knowledge does not alter the magnitude of the education coefficients in regressions that explain health behavior, indicating that allocative efficiency is not the main reason schooling is related to health behavior. The same approach was taken by Cutler and Lleras-Muney (2010), who employed cross-sectional data from the NHIS to investigate the impact of knowledge about health risks on the relationship between education and health

behaviors. They too found that health knowledge has only a modest impact on how education impacts health behaviors.

Our study differs from Kenkel (1991) and Cutler and Lleras-Muney (2010) in two important ways. First, we employ a panel data of individuals, rather than a cross-section. Specifically, each person in the National Longitudinal Survey of Youth-97 (NLSY97) was asked questions on health knowledge both in 1997 and 2002. This allows us to investigate the impact of health knowledge on health behavior by netting out time-invariant individual-specific unobservables that may impact both the intensity of the demand for health knowledge and the demand for health behavior. Second, the design of the NLSY97 has generated exogenous increases in the amount of schooling for different individuals in the sample between the two survey years. As explained in detail in the data section, two identical individuals who were both surveyed in 1997 and 2002 could have received significantly differential amounts of schooling (up to 24 months) between these years due to the timing of the 1997 and the 2002 surveys. Thus, individuals are exposed to differential amounts of schooling between the two surveys, which is not related in any way to their personal or family background characteristics.¹ We find that accounting for health knowledge has no impact on the relationship between education and health behaviors.

The second goal of the paper is to investigate the extent to which cognitive ability is responsible for the impact of education on health behavior. Cutler and Lleras-Muney (2010) analyze how the impact of education on health behaviors is influenced by the inclusion of

¹ Also, the NLSY97 allows us to employ the number of months attended to school by the individual as a measure of education. As explained in more detail in the Data section below, the number of months attended to school is measured with a high degree of precision, and it better captures the individual's exposure to schooling. The conventional measure of education (years of completed schooling) contains substantial measurement error, generated by the timing of the survey, in a sample of young adults who are still in school.

various sets of variables to regression models. Using cross-sectional data sets, they find that the impact of education on health behaviors is diminished (but not eliminated) if income, health insurance and family background are controlled for, but that the extent of risk aversion or discounting for the future have no impact on the estimated coefficient of education. They also run regressions of health behavior on education with and without a measure of cognitive ability, and investigate how the estimated coefficient of education is altered. They find evidence suggesting that cognitive ability, measured by the Armed Services Vocational Aptitude Battery (ASVAB) score, accounts for about 20 percent of the impact of education on the demand for health inputs.

The use of the NLSY97 allows us to employ the ASVAB score, as well as another alternative measure of conceptual thinking ability and cognition (PIAT), to investigate the same question. While we obtain similar results as Cutler and Lleras-Muney (2010) using the ASVAB or the PIAT score as a measure of cognitive ability, entertaining the premise that test scores such as ASVAB are impacted by family background (Cunha and Heckman, 2007; Heckman et al., 2006; Hansen et al., 2004) and controlling for family fixed-effects produces the result that cognition, as measured by ASVAB or PIAT, does not influence the impact of education on health behavior.

An arguably exogenous measure of cognitive ability is an indicator of whether the individual suffers from a learning disability such as dyslexia or attention disorder. For any given amount of schooling, individuals with learning disability are expected to learn less in school in comparison to their peers who have no such disability. If learning in school is a determinant of the influence of education on health inputs, then a particular increase in schooling would have a

smaller impact on health behavior for those with learning disability. However, our results show that learning disability does not influence the impact of schooling on health behaviors.

In section II we describe the empirical implementation. Section III presents the data. Section IV includes the discussion of the results, and Section V is the conclusion.

II. Empirical Specification

Consider equation (1) below

$$(1) \quad H_i = \beta_0 + \beta_1 Education_i + X_i \beta_2 + \varepsilon_i$$

where H stands for the demand for various health inputs which are deleterious to health, (such as the demand for cigarettes) for person i . $Education$ represents the level of schooling of the person, X is a vector of control variables, and ε is a standard error term.

Equation (2) is similar to equation (1), but it includes an additional variable, $HealthKnowledge_i$, which measures the extent of the knowledge of person i regarding the health input H . For example, if H stands for consumption of cigarettes, $HealthKnowledge$ measures the extent of the person's knowledge about the health risks associated with smoking.

$$(2) \quad H_i = \delta_0 + \delta_1 Education_i + \delta_2 Health\ Knowledge_i + X_i \delta_3 + \omega_i$$

Kenkel (1991) and Cutler and Lleras-Muney (2010) estimate versions of equations (1) and (2) and investigate the difference between the estimated β_1 and δ_1 ; i.e. the extent to which Health Knowledge alters the impact of education on health inputs. Both papers employ cross-

sectional data sets and they find that health knowledge has a modest (Cutler and Lleras-Muney) or negligible (Kenkel) impact on health behavior; that is, β_1 is not appreciably different from δ_1 .²

In this paper, we employ panel data, which allow us to measure the demand for health inputs, the amount of schooling, and the extent of input-specific health knowledge of individuals in two time periods. Specifically, the respondents of the NLSY97 were asked questions about their health behaviors. Furthermore, information is obtained from survey participants regarding their health knowledge in the 1997 and 2002 waves of the survey along with information on schooling. Time variation in the data allows us to entertain a specification as depicted by equation (3A) where the demand for health inputs for person i at time t depends on the same set of variables as in equation (2), and on an individual-specific, time-invariant heterogeneity component μ_i .

$$(3A) H_{it} = \delta_0 + \delta_1 Education_{it} + \delta_2 Health\ Knowledge_{it} + X_{it} \Psi_1 + \mu_i + \omega_{it}$$

Because the health knowledge questions were administered only in 1997 and 2002, we will employ data from these years. A valuable feature of the data is that among individuals who took the survey in 1997 and again in 2002, there is substantial variation in the distance between the timing of the survey. For example, while some individuals were surveyed as little as 4.5 years apart, the difference between the two surveys was more than 6 years for some others.³ This exogenous variation in the distance between the two interviews translates into variation in schooling received by individuals between the two surveys.

² Kenkel also runs instrumental variables regressions where health knowledge questions are instrumented with whether the individual received advice from a physician on life-style-related topics and for smoking, years of schooling completed after 1964 (the year of surgeon general's report on smoking), as well as indicator variables for occupation and industry and whether the person is employed in a health field. He obtains results similar to OLS (with larger standard errors), and concludes that the OLS results are not biased because of endogeneity.

³ The mean difference between the two surveys is 68 months.

Time-differencing Equation (3A) allows us to eliminate individual-specific unobservables (μ_i) that may be correlated with health behaviors as well as education and health knowledge. In equation (3B) Δ^p stands for p-month difference, where p represents the number of months between the surveys, which is different for different people.

$$(3B) \quad \Delta^p H_{it} = \delta_1 \Delta^p Education_{it} + \delta_2 \Delta^p Health\ Knowledge_{it} + \Delta^p X_{it} \Psi_2 + \Delta^p \omega_{it}$$

The vector X contains time-varying attributes of the individual. Because a higher value of (Δ^p Education) embodies the effect of increased schooling as well as aging, we also control for the difference in age between the two survey years. As mentioned earlier, Cutler and Lleras-Muney (2010) attribute some of the observed relationship between education and health to cognitive ability. They argue that schooling improves cognition and enhanced cognitive skills alter health behaviors and improve health outcomes. Along the same lines, Auld and Sidhu (2005) find that controlling for test scores has an impact on the estimated impact of education on self-reported health.⁴

To test this hypothesis, we estimate regressions very similar to Cutler and Lleras-Muney (2010). Specifically, we run cross-sectional models depicted by Equation 4 below.

$$(4) \quad H_i = \gamma_0 + \gamma_1 Education_i + \gamma_2 Cognitive\ Ability_i + \gamma_3 Health\ Knowledge_i + X_i \Psi_3 + v_i$$

where, following Cutler and Lleras-Muney (2010), *Cognitive Ability* is measured by the ASVAB score. Equation (4) allows us to investigate the sensitivity of the impact of education on health behavior (γ_1) to the inclusion/exclusion of Cognitive Ability. Note that as was the case in Cutler and Lleras-Muney (2010), the ASVAB score of each individual is constant over time.

⁴ Auld and Sidhu (2005) use adjusted-AFQT scores as a measure of ability and find that schooling has an effect on health only for those with low schooling, and in particular with low ability.

Thus, although each individual contributes two observations (one from 1997, the other from 2002), equation (4) is a pooled cross-section. In addition to ASVAB, we also employ the Peabody Individual Achievement Test (PIAT) as an alternative correlate of cognition.

To test this conjecture in a different framework, we hypothesize that if cognition matters, the impact of education on health behaviors should be different between those who have a learning disability and those who do not. That is, if education improves cognition which in turn impacts health behavior, an additional amount of education should have a *smaller* impact on health behavior among those who have a *learning disability*. More specifically, the coefficient γ_2 should be negative in Equation (5) below, mitigating the impact of education on health behavior.

$$(5) \quad \Delta^p H_{it} = \gamma_1 \Delta^p Education_{it} + \gamma_2 (\Delta^p Education_{it} \times Learning\ Disability_{it}) \\ + \gamma_3 \Delta^p Health\ Knowledge_{it} + \Delta^p X_{it} \Psi + v_{it}$$

We have information, obtained from parents, on whether the individuals in the sample have a learning disability such as dyslexia or attention deficit disorder. This is arguably an exogenous indicator of the extent of cognitive difficulty of the individual. Note that the main effect of learning disability on health behavior cannot be identified in this specification because the indicator of learning disability is time-invariant.

III. Data and Measurement of Variables

The data are obtained from the NLSY97, which contains a nationally representative sample of 8,984 youths who were aged 12–16 as of December 31st 1996. The respondents have been followed annually since the survey was initiated. The cohort born in 1983 was asked health

knowledge questions in the 1997 and 2002 waves of the survey. Therefore, the bulk of our analysis uses data from these two waves.

The 1997 wave of the NLSY97 was administered between January 1997 and May 1998, and the 2002 survey was administered between November 2002 and July 2003. This means, for example, that a 9th grader in the 1997 wave could have been interviewed 54 months later in the 2002 wave of the survey, while another 9th grader could have been interviewed 78 months after the first survey. As the timing of the surveys is random and, therefore, not correlated with student or parent attributes, this design implies that the second student could have been exposed to 20 additional months of schooling in comparison to the first student (Altindag, Cannonier and Mocan, 2011).⁵

Table 1 provides the summary statistics of the variables employed in the analysis. Health behavior variables are cigarette smoking and alcohol consumption. *Cigarettes per Day* stands for the average number of cigarettes smoked by the individual during the last 30 days. *Smoker* is an indicator for smoking participation (smoked at least one cigarette per day). *Cigarettes per Day among Smokers* gives the number of cigarettes smoked among smokers. *One Pack per Day* is a measure of heavy smoking. It takes the value of one if an individual has smoked at least twenty cigarettes per day in the thirty days prior to the interview, and zero otherwise. *Heavy Drinker* is an indicator that takes the value of one if an individual has consumed more than sixty alcoholic drinks in the last thirty days⁶. Summary statistics in Table 1 suggest an increase between the two waves in smoking participation, number of cigarettes smoked per day and heavy alcohol

⁵ The difference in exposure to schooling in this example is 20 months rather than 24 because there is no schooling in summer months.

⁶ This cut-off of sixty drinks is not arbitrary. According to Dawson, Grant and Chou (1995), individuals who consume more than 2 drinks every day are considered heavy drinkers.

consumption for the individuals in our sample. Note also that average age has increased from 13 to 19 between the two survey waves.

We measure schooling by the number of *Months Attended*, which is the cumulative number of months the individual has attended any type of school (kindergarten to college) since the first interview in the 1997 wave.⁷ This variable is created using monthly schooling status information available in the schooling event history of each wave of the NSLY97 between 1997 and 2002 waves. Another schooling measure we employ is *Highest Grade Completed*. We prefer *Months Attended* to *Highest Grade Completed* because the latter does not measure schooling with precision. For example, consider the case where some respondents are interviewed right after the end of the school year and others are interviewed right before the end of the school year. Those who are interviewed when the school was in session (but close to the end) will report a value for the number of years of completed schooling which is one year fewer in comparison to those who are interviewed right after the end of the school year. However, the actual difference in terms of schooling is much smaller than one full year of schooling. Similarly, years of completed schooling will not reflect the true difference in schooling for two students who are interviewed in different months of the same academic year. Such measurement error would generate attenuation bias in the estimated coefficient of years of completed schooling. Nevertheless, we use both measures because it is not possible to employ *Months Attended* in cross sectional regressions as explained below.

The variables *Smoking Knowledge* and *Drinking Knowledge* indicate the proportion of correctly answered questions about health risks of smoking and drinking, respectively. For *Smoking Knowledge*, the questions gauge whether the individual has correct information about

⁷ However, none of the individuals in our sample is in kindergarten or in primary school at the time of the first interview.

the connection between smoking and heart disease, and smoking and AIDS. For *Drinking Knowledge*, the questions are based on the connection between drinking and liver disease, heart disease, arthritis, addiction to alcohol, and harm on unborn child. The list of the questions and the correct answers are listed in the Appendix. Summary statistics in Table 1 indicate that most of the individuals have high levels of health knowledge about both smoking and drinking. The proportion of correct answers has increased between 1997 and 2002 in case of smoking.

We use the ASVAB score, as is the case in Cutler and Lleras-Muney (2010). About 80 percent of the respondents in the NLSY97 sample took the computer-adaptive form of the Armed Services Vocational Aptitude Battery (ASVAB) test. The ASVAB test consists of 12 subtests that measure vocational aptitude in areas such as arithmetic reasoning, assembling objects, auto information and so on. The variable used in our analysis is constructed based on age adjusted test scores of individuals in four sub tests: mathematical knowledge, arithmetic reasoning, word knowledge, and paragraph comprehension as obtained from the NLSY97 data set.⁸ The final variable is the percentile in which the individual's test scores fall in comparison to other ASVAB takers.

As an alternative test score, we utilize individual's PIAT (Peabody Individual Achievement Test) math assessment scores. Specifically, we use the individual's percentile score for the PIAT. The version of PIAT administered for the NLSY97 respondents involved answering several mathematics questions. The difficulty of the questions is age-adjusted. Ninety-four percent of the individuals in our sample took the PIAT test during the 1997 wave.

Learning Disability is an indicator for whether the individual has a learning disability. This variable is constructed based on parents' report, who were asked the following question.

⁸ These four subtests are used by the Department of Defense to calculate AFQT scores (Armed Forces Qualification Test scores).

“Does your child now have or has [he/she] ever had a learning or emotional problem that limits or has limited the kind of schoolwork or other daily activities [he/she] can perform, the amount of time [he/she] can spend on these activities or [his/her] performance in these activities?” If the parent answered in the affirmative, a second question was asked as follows. “What (is/are) the condition(s)? (Select all that apply.) Learning disability (i.e., dyslexia) or attention disorder; Emotional/mental problem or behavior problem; Eating disorder like anorexia or bulimia; Mental retardation; Other (Specify).” We coded our *Learning Disability* variable to take the value of one if the parent declared the existence of learning disability (i.e., dyslexia) or attention disorder. In our sample about nine percent of the individuals have learning disability. This is consistent with the findings of a CDC report by Pastor and Reuben (2008) who find that about eight to nine percent of all children aged between six and eleven have learning disorders.

Time-dependent variables shown in Table 1 are included as control variables in the empirical analyses. All individuals in the sample are born in 1983.⁹ However, due to the differences in the interview date, there is variation in *Age*. On average, respondents age by about 6 years between the two survey waves. *Household Income* is deflated by 1,000. Unsurprisingly, none of the individuals in the 1997 wave were married, and very few were married as of the 2002 wave. *Cumulative Hours Worked* measures the total number of hours an individual has worked in the labor market. *Household Size* gives the number of individuals in the respondent’s household.

The remaining variables in Table 1 are time-invariant individual characteristics. They are included in cross sectional sections as control variables. We present summary statistics of these variables to provide information about the sample. For example, about half of the sample consist

⁹ This is because of the design of the survey. Only individuals in the cohort born in 1983 are asked health knowledge questions. These individuals make up our estimation sample.

of males. Individuals who identify themselves as *Hispanic* and non-Hispanic *Black* make up 20% and 26% of the whole sample, respectively.

IV. Results

The Influence of Health Knowledge

The results obtained from estimating equation (3B) are presented in Table 2. In this specification, all variables are in first-differences. For example, *Months Attended* stands for the change in the number of months the individual attended school between the two survey years. For each health behavior, two columns of results are presented. The odd-numbered (even-numbered) columns exclude (include) individual's health knowledge about the health behavior. For example, columns (1) and (2) report the regression results where the dependent variable is whether the person is a smoker. Both columns are based on the same specification except that column (2) controls for smoking knowledge, and column (1) omits it.

Education has negative impact on smoking, both at the extensive and intensive margins, and on the propensity to drink heavily. Note again, that we analyze the propensity for heavy drinking because questions on drinking knowledge are based on heavy drinking (see the Appendix). An increase in *Months Attended* by one school year (9 months) decreases the propensity to smoke by 2.7 percentage points (0.3×9), which translates into a 12% decline. A one-year increase in schooling (9 months) reduces the daily number of cigarettes smoked by one cigarette for everyone. The same increase in schooling reduces the daily number of cigarettes smoked among smokers by 0.8 cigarettes (by about 11%), and it reduces the propensity for being a heavy drinker by about one percentage point (about 20%).

Marriage has a negative impact on the number of cigarettes smoked as well as on the propensity to drink heavily. The number of hours worked in the labor market is positively associated with smoking and also with heavy drinking.

An increase in knowledge about smoking has a negative impact on cigarettes smoked per day, and an increase in drinking knowledge has a negative impact on heavy drinking. However, inclusion of the knowledge variables does not change the estimated coefficients of education.

The last row of Table 2 presents the percent change in the estimated coefficient of education due to the inclusion of health knowledge. Inclusion of smoking knowledge reduces the size of the education coefficient trivially (by about one percent) in two of the five models. The change is virtually zero for smoking participation, heavy smoking and heavy alcohol consumption.

Table 3 presents the same specifications as in Table 2 with one difference: here the change in education is measured by the change in the highest grade completed . As discussed earlier, this is not as accurate a measure of acquired schooling as *Months Attended* which is used in Table 2. However, we present the results using highest grade completed as well, because later in the paper cross-sectional regressions can only be run using highest grade completed as the measure of schooling. As can be seen, the results are consistent between tables 2 and 3.

These results show that accounting for health knowledge does not eliminate or reduce the impact of education on health behavior. Thus, they indicate that allocative efficiency is not likely a primary mechanism through which education impacts health inputs.

Cognition

In this section, we present the results of the analyses that investigate whether variations in cognitive ability is the reason behind the impact of education on health behaviors. Table 4A presents the results obtained from estimating versions of Equation (4). ASVAB is a measure of cognitive ability; it stands for the percentile ranking of the individual's ASVAB score, ranging from 0 to 100 where higher scores represent higher ability (Cutler and Lleras-Muney, 2010). Education is measured by the highest grade completed. We cannot use *Months Attended* as the measure of education in these regressions. This is because, to replicate Cutler and Lleras-Muney specifications we run cross-sectional regressions in this specification using data from 1997 (the first wave) as well as from 2002. These are the two years in which health knowledge questions were administered. The *Months Attended* variable measures the number of months the individual has attended any school *since* the first interview; and the number of months of school attendance prior to the first wave of interviews in the NSLY97 has not been recorded. Thus, we employ *Highest Grade Completed* as the measure of schooling, which is reported both in 1997 and in 2002.

As Table 4A demonstrates, education has a negative impact on unhealthy behavior, but controlling for the ASVAB score reduces the magnitude of the coefficient of education. For example, an additional year of schooling reduces the propensity to smoke by 4.4 percentage points in column (1) when the model does not include ASVAB, but the marginal effect of schooling is -0.037 when the model contains ASVAB (column 2). The same is true for cigarettes smoked per day and cigarettes smoked per day among smokers. These results are consistent with those reported by Cutler and Lleras-Muney (2010).

It is plausible that the ASVAB score is not a reliable indicator of cognitive ability. For example, Cunha and Heckman (2007), Heckman et al. (2006) and Hansen et al. (2004) stress that a person's schooling and family background at the time tests are taken affect test scores. Although we control for some family background characteristics in the regressions reported in table 4A, it is likely that important family attributes are omitted. Thus, we re-estimate Equation (4) by controlling for family fixed-effects. Identification is obtained off of variation between siblings who have different cognitive ability as measured by ASVAB scores. The results are displayed in Table 4B. The impact of education on being a smoker is substantially reduced in this specification; and the same is true to a lesser extent about the number of cigarettes consumed. However, in this specification the coefficient of education is insensitive to the inclusion of the ASVAB score. Put differently, after controlling for family fixed effects, variations in cognitive ability does not influence the impact of education on health behavior.

We repeat the same exercise using the PIAT (the percentile score of the individual's Peabody Individual Achievement Test) score, instead of ASVAB.¹⁰ The results are displayed in Table 5A. Inclusion of the PIAT score reduces the magnitude of the estimated education coefficient slightly. For example, in column 1 of Table 5A we observe that an additional year of education reduces the propensity to smoke by 3.9 percentage points. The regression result, reported in column 2 controls for the PIAT score; and in this specification the impact of an additional year of education is to reduce the smoking propensity by 3.5 percentage points. Similar results are obtained for most outcomes reported in Table 5A. However, when we add family fixed-effects to these specifications, we observe that inclusion or exclusion of the PIAT

¹⁰ PIAT and ASVAB are highly correlated with a correlation coefficient of 0.72.

score has no impact on the estimated education coefficient (Table 5B). In other words, controlling for cognition, as measured by the PIAT score, does not alter the relationship between education and health behaviors in models that control for unobserved family attributes.

To analyze whether the results are altered if we employ a different, arguably more exogenous measure of cognition, we estimate models where an indicator for learning disability of the individual is employed. To make the results comparable to those obtained from the regressions with ASVAB and PIAT, we first use cross-sectional data from 1997 and 2002 and employ *Highest Grade Completed* as the measure of schooling. The results, which are presented in Table 6 show that controlling for learning disability does not alter the estimated coefficient of education.

Table 7A presents the models that perform the same analysis, but here the panel nature of the data is exploited. These models are based on equation (5). We use the *Months Attended* in these regressions. Once again, exposure to additional months of schooling between the two survey years reduces the propensity to smoke, the number of cigarettes smoked and the propensity to be a heavy drinker. However, the impact of schooling is not different between students with and without learning disability. Table 7B estimates the same specification as in Table 7A, but uses *Highest Grade Completed* as the measure of schooling to investigate if the results are sensitive to the measure of schooling. The inference does not change. Controlling for learning disability does not alter the relationship between education and health behavior. Furthermore, the marginal effect of education on health behaviors is the same between those who have a learning disability and those who don't. These results suggest that cognition does not impact the education gradient in health behaviors. More specifically, the results don't lend

support to the hypothesis that education increases cognition, and enhanced cognition and intelligence enable people to make better health decisions.

V. Summary and Conclusion

Using a panel data set of young individuals from the NLSY97, we pose two questions. The first question is whether schooling increases the efficiency of health production. Productive efficiency hypothesis suggests that education has a direct impact on health, much like the impact of technology on production. More educated people are more efficient producers of health, perhaps because the marginal products of health inputs differ by education. An alternative hypothesis is that of allocative efficiency, where more educated people make different choices about health inputs; i.e., they allocate inputs differently which in turn produce more health output. Under allocative efficiency, education has no direct influence on health as the impact of education is only working through the pathway of health inputs. For example, education provides knowledge about the benefits or harmful effects of health inputs (such as nutrition or smoking) and this knowledge alters health behavior and health outcomes.

To investigate the relative validity of these hypotheses, we estimate models of health behavior, where the change in various measures of smoking and heavy drinking between the two survey years are regressed on arguably exogenous increases in educational attainment between the same years and on the change in the relevant health knowledge. We find that health knowledge has an impact on health behavior, but that accounting for health knowledge does not eliminate or reduce the impact of education on health behavior. This finding supports the productive efficiency hypothesis.

We also investigate whether cognitive ability is responsible for the impact of education on health behavior. Using the ASVAB and PIAT scores as alternative measures of cognitive ability we find that accounting for ability reduces the impact of education on health behavior, suggesting that cognition partly explains the relationship between education and health behaviors. However, test scores such as ASVAB and PIAT are known to be noisy measures of cognition that may be impacted by family attributes (Cunha and Heckman, 2007; Heckman 2006; Hansen, Heckman and Mullen 2004). In fact, regressions that control for family fixed-effects show that inclusion or exclusion of ASVAB or PIAT scores have no impact on the estimated impact of education on health behavior.

In addition, we perform another test to investigate how cognitive ability impacts the relationship between education and health behaviors. The test involves a comparison of health input demands of two individuals who are observationally identical except for one dimension: one of them has a learning disability such as dyslexia or attention disorder. The individual with the learning disability is expected to learn less in school compared to the individual without the disorder for a given level of schooling. If what is learned in school is a determinant of the influence of education on health inputs, then a particular increase in schooling would have a smaller impact on health behavior for those with learning disability. However, our results show that learning disability does not influence the impact of schooling on health behaviors. An increase in schooling has the same impact on health behaviors for those who have a learning disability as for those who don't have a learning disability. These findings suggest that cognition is unlikely to be a primary factor in explaining the relationship between education and the demand for health inputs.

Table 1: Summary Statistics

Variable	1997 Wave			2002 Wave			Whole Sample		
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Cigarettes per day	1,790	0.399	1.895	1,629	3.050	6.618	3,419	1.662	4.949
Smoker 0/1	1,790	0.112	0.315	1,629	0.341	0.474	3,419	0.221	0.415
Cigarettes per day among Smokers	200	3.570	4.572	556	8.935	8.706	756	7.516	8.175
One pack per day 0/1	1,790	0.003	0.058	1,629	0.053	0.225	3,419	0.027	0.163
Heavy drinker 0/1	1,789	0.006	0.078	1,627	0.090	0.286	3,416	0.046	0.209
Months Attended	1,791	0.958	0.200	1,631	50.536	11.719	3,422	24.588	26.053
Highest Gr. Completed	1,791	6.788	0.757	1,631	11.514	1.971	3,422	9.041	2.779
Drinking Knowledge	1,791	0.829	0.170	1,631	0.817	0.187	3,422	0.823	0.178
Smoking Knowledge	1,791	0.899	0.209	1,631	0.930	0.176	3,422	0.914	0.195
Learning Disability	1,587	0.089	0.285	1,457	0.089	0.285	3,044	0.089	0.285
ASVAB	1,421	45.378	29.213						
PIAT	1,689	49.233	34.457						
Age	1,791	13.372	0.507	1,631	19.028	0.315	3,422	16.068	2.857
Household Income	1,353	46.356	40.613	1,311	56.143	56.327	2,664	51.172	49.215
Married	1,791	0.000		1,631	0.039	0.194	3,422	0.019	0.106
Cumulative Hours Worked (1,000s)	1,785	0.003	0.049	1,524	2.519	1.949	3,309	1.162	1.823
Household Size	1,791	4.584	1.521	1,631	4.000	1.775	3,422	4.306	1.672
Mother High School Graduate	1,791	0.733	0.442	1,631	0.735	0.442	3,422	0.734	0.442
Male	1,791	0.516	0.500	1,631	0.514	0.500	3,422	0.515	0.500
Black	1,791	0.256	0.437	1,631	0.261	0.439	3,422	0.258	0.438
Hispanic	1,791	0.204	0.403	1,631	0.203	0.402	3,422	0.203	0.403

Table 2
Health Knowledge, School Attendance and Health Behaviors – First Differences

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Months Attended	-0.003*** (0.001)	-0.003*** (0.001)	-0.117*** (0.018)	-0.116*** (0.018)	-0.092*** (0.034)	-0.091*** (0.034)	-0.003*** (0.001)	-0.003*** (0.001)	-0.001* (0.001)	-0.001* (0.001)
Health Knowledge		-0.022 (0.049)		-1.282* (0.672)		-1.889 (1.644)		-0.031 (0.021)		-0.065** (0.031)
Age	0.028 (0.025)	0.029 (0.025)	0.369 (0.363)	0.393 (0.366)	0.524 (0.849)	0.522 (0.844)	0.001 (0.011)	0.002 (0.012)	-0.010 (0.015)	-0.010 (0.015)
Household Income	-0.000 (0.000)	-0.000 (0.000)	-0.003 (0.003)	-0.003 (0.003)	-0.009 (0.007)	-0.009 (0.007)	-0.000 (0.000)	-0.000 (0.000)	0.000* (0.000)	0.000* (0.000)
Married	-0.092 (0.068)	-0.091 (0.068)	-1.663** (0.816)	-1.593* (0.819)	-1.611 (1.877)	-1.484 (1.902)	-0.037 (0.029)	-0.035 (0.029)	-0.060** (0.024)	-0.061** (0.024)
Hours Worked	0.010 (0.007)	0.010 (0.007)	0.286*** (0.101)	0.280*** (0.100)	0.427** (0.213)	0.415** (0.211)	0.008** (0.004)	0.008** (0.004)	0.010** (0.004)	0.010*** (0.004)
Household Size	-0.009 (0.007)	-0.009 (0.007)	-0.081 (0.124)	-0.078 (0.124)	-0.008 (0.301)	-0.002 (0.300)	-0.003 (0.004)	-0.003 (0.004)	-0.000 (0.004)	-0.000 (0.004)
Observations	1613	1613	1613	1613	611	611	1613	1613	1610	1610
Change in Education's Coefficient		0.0%		0.9%		1.1%		0.0%		0.0%

Notes: Months Attended is the cumulative number of months attended to any school. The outcome variables are listed at the top of columns. Odd (even) numbered columns exclude (include) Health Knowledge (Smoking or Drinking). Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. OLS is employed on the first differenced data. Robust standard errors are in parentheses. *, **, and *** indicate significance at 10%, 5% and 1 % levels, respectively.

^a The sample include individuals who were smokers in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 3
Health Knowledge, Highest Grade Completed, Health Behaviors – First Differences

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.018** (0.007)	-0.018** (0.007)	-0.415*** (0.098)	-0.408*** (0.096)	-0.267 (0.171)	-0.250 (0.170)	-0.009*** (0.003)	-0.009*** (0.003)	-0.007* (0.004)	-0.007* (0.004)
Health Knowledge		-0.023 (0.049)		-1.399** (0.684)		-1.938 (1.663)		-0.034 (0.022)		-0.064** (0.031)
Age	0.027 (0.024)	0.027 (0.025)	0.230 (0.359)	0.258 (0.362)	0.381 (0.834)	0.378 (0.830)	-0.003 (0.011)	-0.002 (0.011)	-0.011 (0.015)	-0.011 (0.015)
Household Income	-0.000 (0.000)	-0.000 (0.000)	-0.004 (0.003)	-0.004 (0.003)	-0.009 (0.007)	-0.010 (0.007)	-0.000 (0.000)	-0.000 (0.000)	0.000* (0.000)	0.000* (0.000)
Married	-0.079 (0.067)	-0.078 (0.068)	-0.886 (0.789)	-0.816 (0.795)	-1.087 (1.842)	-0.961 (1.872)	-0.014 (0.028)	-0.012 (0.028)	-0.055** (0.024)	-0.056** (0.024)
Hours Worked	0.012* (0.007)	0.012* (0.007)	0.358*** (0.101)	0.350*** (0.100)	0.463** (0.213)	0.450** (0.210)	0.010*** (0.004)	0.010** (0.004)	0.010*** (0.004)	0.010*** (0.004)
Household Size	-0.009 (0.007)	-0.009 (0.007)	-0.078 (0.130)	-0.075 (0.130)	-0.052 (0.310)	-0.043 (0.308)	-0.003 (0.004)	-0.003 (0.004)	-0.000 (0.004)	-0.000 (0.004)
Observations	1613	1613	1613	1613	611	611	1613	1613	1610	1610
Change in Education's Coefficient	0.0%		1.7%		6.4%		0.0%		0.0%	

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Odd (even) numbered columns exclude (include) Health Knowledge (Smoking or Drinking). Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. OLS is employed on the first differenced data. Robust standard errors are in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively.

^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 4A
The Impact of Highest Grade Completed on Health behaviors, models with and without ASVAB – Cross Section

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.044*** (0.008)	-0.037*** (0.008)	-0.551*** (0.109)	-0.505*** (0.109)	-0.496** (0.199)	-0.478** (0.202)	-0.011*** (0.003)	-0.011*** (0.003)	-0.007 (0.004)	-0.006 (0.005)
Asvab		-0.002*** (0.0001)		-0.010*** (0.004)		-0.006 (0.012)		0.00001 (0.00010)		-0.0003 (0.0002)
Health Knowledge	0.023 (0.036)	0.043 (0.036)	-0.978* (0.592)	-0.848 (0.587)	-4.662* (2.578)	-4.593* (2.552)	-0.026 (0.017)	-0.026 (0.017)	-0.039 (0.024)	-0.033 (0.025)
Male	0.017 (0.016)	0.015 (0.016)	0.486*** (0.177)	0.469*** (0.177)	1.794*** (0.653)	1.790*** (0.653)	0.020*** (0.006)	0.020*** (0.006)	0.040*** (0.008)	0.039*** (0.008)
Black	-0.126*** (0.018)	-0.160*** (0.020)	-1.723*** (0.211)	-1.943*** (0.241)	-5.320*** (0.862)	-5.413*** (0.914)	-0.033*** (0.007)	-0.033*** (0.008)	-0.035*** (0.008)	-0.042*** (0.010)
Hispanic	-0.047** (0.024)	-0.068*** (0.025)	-1.357*** (0.290)	-1.490*** (0.307)	-4.364*** (1.013)	-4.399*** (1.033)	-0.033*** (0.008)	-0.033*** (0.008)	-0.007 (0.013)	-0.011 (0.013)
Age	0.034* (0.018)	0.031* (0.018)	0.285 (0.221)	0.267 (0.221)	0.076 (0.896)	0.082 (0.898)	0.000 (0.007)	0.000 (0.007)	0.006 (0.009)	0.005 (0.009)
Household Income	-0.000** (0.000)	-0.000 (0.000)	-0.005** (0.002)	-0.004 (0.003)	-0.014* (0.008)	-0.013* (0.008)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Married	-0.064 (0.075)	-0.073 (0.077)	-1.512** (0.696)	-1.568** (0.702)	-1.963 (1.759)	-1.946 (1.767)	-0.030 (0.025)	-0.030 (0.025)	-0.057** (0.026)	-0.059** (0.026)
Hours Worked	0.023*** (0.007)	0.023*** (0.007)	0.355*** (0.107)	0.351*** (0.107)	0.402* (0.232)	0.403* (0.232)	0.011** (0.004)	0.011** (0.004)	0.007* (0.004)	0.007* (0.004)
Household Size	-0.010** (0.005)	-0.012** (0.005)	-0.067 (0.088)	-0.078 (0.088)	0.193 (0.342)	0.188 (0.342)	-0.001 (0.002)	-0.001 (0.002)	-0.004 (0.002)	-0.004* (0.002)
Mother HS Grad	0.031 (0.022)	0.053** (0.022)	-0.184 (0.302)	-0.040 (0.293)	-1.517 (1.068)	-1.426 (1.057)	-0.004 (0.009)	-0.004 (0.009)	0.022** (0.010)	0.026** (0.010)
Observations	2735	2735	2735	2735	596	596	2735	2735	2733	2733

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Asvab is the individual's percentile score in the math and verbal sections of the ASVAB test. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. Results are obtained using OLS. Standard errors clustered at the individual level are presented in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively. ^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave. ^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 4B
The Impact of Highest Grade Completed on Health Behaviors, models with and without ASVAB – Models with Family Fixed effects

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.024*	-0.023*	-0.468***	-0.467***	-0.420	-0.415	-0.010*	-0.010*	-0.008	-0.008
	(0.012)	(0.012)	(0.166)	(0.166)	(1.863)	(1.894)	(0.005)	(0.005)	(0.007)	(0.007)
Asvab		-0.004		-0.006		-0.012		0.0002		0.0003
		(0.003)		(0.010)		(0.164)		(0.0001)		(0.0002)
Health Knowledge	-0.068	-0.070	-2.066*	-2.069*	0.217	0.187	-0.056	-0.056	-0.062	-0.062
	(0.078)	(0.078)	(1.187)	(1.188)	(8.368)	(8.368)	(0.034)	(0.034)	(0.049)	(0.049)
Male	0.048	0.026	-0.164	-0.192			-0.008	-0.007	(0.049)	(0.049)
	(0.058)	(0.067)	(0.338)	(0.337)			(0.010)	(0.010)	-0.005	-0.003
Black	0.458***	0.435***	0.386	0.356			0.003	0.003	(0.010)	(0.010)
	(0.081)	(0.108)	(1.162)	(1.145)			(0.034)	(0.034)	0.012	0.014
Hispanic	-0.000	0.059	0.004	0.081			0.000	-0.002	(0.042)	(0.042)
	(0.002)	(0.042)	(0.023)	(0.138)			(0.001)	(0.002)	0.005	0.001
Age	0.036	0.036	0.241	0.241	-1.759	-1.768	-0.013	-0.013	(0.005)	(0.005)
	(0.042)	(0.042)	(0.637)	(0.637)	(6.228)	(6.318)	(0.020)	(0.020)	0.003	0.003
Household Income	-0.000	-0.000	-0.002	-0.002	0.022	0.022	-0.000	-0.000	(0.025)	(0.025)
	(0.000)	(0.000)	(0.005)	(0.005)	(0.057)	(0.058)	(0.000)	(0.000)	0.000	0.000
Married	-0.037	-0.036	-1.450	-1.450	-6.392	-6.413	-0.040	-0.040	(0.000)	(0.000)
	(0.129)	(0.129)	(1.043)	(1.043)	(6.400)	(6.545)	(0.037)	(0.037)	-0.060	-0.060
Hours Worked	0.013	0.013	0.353**	0.353**	-0.499	-0.503	0.010	0.010	(0.038)	(0.038)
	(0.011)	(0.011)	(0.162)	(0.162)	(1.466)	(1.495)	(0.006)	(0.006)	0.008	0.008
Household Size	-0.010	-0.010	-0.090	-0.090	-1.226	-1.228	-0.004	-0.004	(0.006)	(0.006)
	(0.013)	(0.013)	(0.235)	(0.235)	(1.420)	(1.429)	(0.007)	(0.007)	0.000	0.000
Mom HS Grad	-0.490***	-0.392***	-1.373	-1.244	-6.452	-6.437	-0.008	-0.011	(0.006)	(0.006)
	(0.075)	(0.130)	(0.852)	(0.866)	(19.686)	(19.819)	(0.028)	(0.028)	-0.044	-0.050
Observations	2735	2735	2735	2735	596	596	2735	2735	(0.034)	(0.036)

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Asvab is the individual's percentile score in the math and verbal sections of the ASVAB test. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. Results are obtained using OLS with family fixed effects included. Standard errors clustered at the individual level are presented in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively. ^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 5A
The Impact of Highest Grade Completed on Health Behaviors, models with and without PIAT – Cross Section

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.039*** (0.007)	-0.035*** (0.007)	-0.521*** (0.095)	-0.487*** (0.093)	-0.571*** (0.196)	-0.549*** (0.195)	-0.011*** (0.003)	-0.010*** (0.003)	-0.007* (0.004)	-0.006 (0.004)
Piat		-0.0012*** (0.0002)		-0.009*** (0.003)		-0.007 (0.008)		-0.0001 (0.0001)		-0.0001 (0.0001)
Health Knowledge	0.046 (0.033)	0.061* (0.033)	-0.448 (0.495)	-0.324 (0.491)	-2.908 (2.214)	-2.791 (2.200)	-0.007 (0.014)	-0.005 (0.014)	-0.030 (0.023)	-0.029 (0.023)
Male	0.023 (0.015)	0.025* (0.015)	0.500*** (0.162)	0.520*** (0.162)	1.552*** (0.554)	1.561*** (0.553)	0.017*** (0.005)	0.017*** (0.005)	0.038*** (0.007)	0.039*** (0.007)
Black	-0.134*** (0.017)	-0.158*** (0.018)	-1.759*** (0.190)	-1.949*** (0.207)	-4.775*** (0.726)	-4.872*** (0.740)	-0.033*** (0.006)	-0.035*** (0.007)	-0.036*** (0.008)	-0.038*** (0.008)
Hispanic	-0.067*** (0.022)	-0.080*** (0.022)	-1.521*** (0.262)	-1.625*** (0.271)	-4.410*** (0.896)	-4.432*** (0.903)	-0.030*** (0.008)	-0.032*** (0.008)	-0.004 (0.011)	-0.005 (0.012)
Age	0.043*** (0.016)	0.040** (0.016)	0.531*** (0.201)	0.506** (0.200)	0.936 (0.775)	0.930 (0.774)	0.008 (0.007)	0.007 (0.007)	0.006 (0.008)	0.005 (0.008)
Household Income	-0.000*** (0.000)	-0.000** (0.000)	-0.005** (0.002)	-0.004* (0.002)	-0.010 (0.007)	-0.009 (0.007)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Married	-0.058 (0.062)	-0.061 (0.062)	-0.650 (0.797)	-0.672 (0.797)	-0.020 (1.749)	-0.008 (1.750)	-0.007 (0.028)	-0.007 (0.028)	-0.054** (0.024)	-0.054** (0.024)
Hours Worked	0.018*** (0.006)	0.018*** (0.006)	0.335*** (0.100)	0.329*** (0.099)	0.430** (0.198)	0.433** (0.198)	0.009** (0.004)	0.009** (0.004)	0.009** (0.004)	0.009** (0.004)
Household Size	-0.011** (0.004)	-0.012*** (0.004)	-0.080 (0.075)	-0.087 (0.074)	0.039 (0.287)	0.036 (0.286)	-0.001 (0.002)	-0.001 (0.002)	-0.003 (0.002)	-0.003 (0.002)
Mother HS Grad	0.016 (0.020)	0.035* (0.020)	-0.211 (0.269)	-0.066 (0.265)	-1.256 (0.897)	-1.142 (0.886)	-0.004 (0.009)	-0.002 (0.009)	0.024*** (0.009)	0.025*** (0.009)
Observations	3224	3224	3224	3224	717	717	3224	3224	3220	3220

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Piat is the individual's percentile score in the math section of the PIAT test. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. Results are obtained using OLS. Standard errors clustered at the individual level are presented in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively. ^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 5B
The Impact of Highest Grade Completed on Health Behaviors, models with and without PIAT – Models with Family Fixed effects

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.017 (0.011)	-0.016 (0.011)	-0.392*** (0.142)	-0.395*** (0.143)	0.001 (1.368)	0.004 (1.374)	-0.009* (0.005)	-0.009* (0.005)	-0.008 (0.006)	-0.008 (0.006)
Piat		-0.002 (0.002)		0.015 (0.010)		-0.013 (0.133)		0.0004** (0.0002)		0.0004 (0.0003)
Health Knowledge	-0.029 (0.072)	-0.029 (0.072)	-1.390 (1.084)	-1.388 (1.084)	1.787 (8.648)	1.795 (8.718)	-0.028 (0.032)	-0.028 (0.032)	-0.054 (0.046)	-0.054 (0.046)
Male	0.057 (0.074)	0.050 (0.074)	0.254 (0.373)	0.324 (0.386)			-0.003 (0.010)	-0.001 (0.010)	0.002 (0.012)	0.004 (0.011)
Black	0.027 (0.121)	0.086 (0.176)	-1.498 (0.977)	-2.080** (0.958)			-0.023 (0.026)	-0.037 (0.026)	-0.016 (0.038)	-0.031 (0.037)
Hispanic	0.001 (0.002)	0.004 (0.005)	0.012 (0.022)	-0.018 (0.030)			0.001 (0.001)	-0.000 (0.001)	0.004 (0.005)	0.003 (0.005)
Age	0.034 (0.037)	0.034 (0.037)	0.344 (0.534)	0.347 (0.535)	-1.912 (3.401)	-1.917 (3.422)	-0.004 (0.017)	-0.004 (0.017)	-0.007 (0.022)	-0.007 (0.022)
Household Income	-0.000 (0.000)	-0.000 (0.000)	-0.004 (0.004)	-0.004 (0.004)	0.015 (0.046)	0.015 (0.046)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Married	-0.068 (0.099)	-0.068 (0.099)	-0.697 (1.173)	-0.701 (1.173)	-1.089 (11.714)	-1.097 (11.776)	-0.011 (0.041)	-0.011 (0.041)	-0.055 (0.035)	-0.055 (0.035)
Hours Worked	0.011 (0.010)	0.011 (0.010)	0.351** (0.149)	0.351** (0.149)	0.021 (1.098)	0.018 (1.110)	0.009 (0.006)	0.009 (0.006)	0.010* (0.005)	0.010* (0.005)
Household Size	-0.009 (0.011)	-0.009 (0.011)	-0.070 (0.196)	-0.070 (0.196)	-1.159 (1.193)	-1.160 (1.198)	-0.003 (0.006)	-0.003 (0.006)	-0.000 (0.006)	-0.000 (0.006)
Mom HS Grad	-0.210 (0.176)	-0.222 (0.181)	0.034 (1.048)	0.145 (1.049)	0.368 (13.940)	-0.244 (15.941)	-0.001 (0.021)	0.002 (0.018)	-0.038 (0.028)	-0.036 (0.026)
Observations	3224	3224	3224	3224	717	717	3224	3224	3220	3220

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Piat is the individual's percentile score in the math section of the PIAT test. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. Results are obtained using OLS with family fixed effects included. Standard errors clustered at the individual level are presented in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively. ^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 6
The Impact of Highest Grade Completed on Health Behaviors, models with and without Learning Disability – Cross Section

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.042*** (0.007)	-0.041*** (0.007)	-0.532*** (0.094)	-0.525*** (0.094)	-0.472*** (0.175)	-0.468*** (0.176)	-0.010*** (0.003)	-0.010*** (0.003)	-0.007* (0.004)	-0.008* (0.004)
Learning Disability		0.059* (0.032)		0.469 (0.374)		0.400 (0.923)		0.003 (0.012)		-0.012 (0.014)
Health Knowledge	0.062* (0.034)	0.069** (0.034)	-0.546 (0.539)	-0.489 (0.543)	-3.802 (2.472)	-3.717 (2.517)	-0.017 (0.015)	-0.016 (0.015)	-0.049** (0.023)	-0.051** (0.023)
Male	0.020 (0.016)	0.016 (0.016)	0.557*** (0.173)	0.528*** (0.174)	1.857*** (0.588)	1.824*** (0.595)	0.018*** (0.006)	0.018*** (0.006)	0.042*** (0.007)	0.042*** (0.007)
Black	-0.136*** (0.018)	-0.134*** (0.018)	-1.857*** (0.205)	-1.846*** (0.206)	-5.128*** (0.760)	-5.119*** (0.763)	-0.035*** (0.007)	-0.035*** (0.007)	-0.039*** (0.008)	-0.039*** (0.008)
Hispanic	-0.090*** (0.022)	-0.086*** (0.022)	-1.633*** (0.283)	-1.605*** (0.284)	-4.353*** (0.971)	-4.319*** (0.974)	-0.032*** (0.008)	-0.032*** (0.008)	-0.011 (0.012)	-0.012 (0.012)
Age	0.043** (0.017)	0.042** (0.017)	0.543** (0.213)	0.541** (0.214)	0.841 (0.825)	0.834 (0.828)	0.009 (0.007)	0.009 (0.007)	0.000 (0.008)	0.000 (0.008)
Household Income	-0.001*** (0.000)	-0.001*** (0.000)	-0.007*** (0.002)	-0.007*** (0.002)	-0.015* (0.008)	-0.015* (0.008)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Married	-0.052 (0.067)	-0.048 (0.067)	-0.587 (0.886)	-0.561 (0.888)	-0.143 (1.845)	-0.111 (1.851)	-0.006 (0.032)	-0.005 (0.032)	-0.051* (0.026)	-0.051* (0.026)
Hours Worked	0.018*** (0.007)	0.018*** (0.007)	0.284*** (0.100)	0.288*** (0.100)	0.293 (0.206)	0.293 (0.206)	0.007* (0.004)	0.007* (0.004)	0.009** (0.004)	0.009** (0.004)
Household Size	-0.009* (0.005)	-0.009* (0.005)	-0.047 (0.084)	-0.045 (0.084)	0.150 (0.302)	0.154 (0.302)	-0.000 (0.002)	-0.000 (0.002)	-0.003 (0.002)	-0.003 (0.002)
Mom HS Grad	0.017 (0.021)	0.017 (0.021)	-0.211 (0.295)	-0.208 (0.295)	-1.384 (0.954)	-1.375 (0.955)	-0.003 (0.009)	-0.003 (0.009)	0.020** (0.009)	0.020** (0.009)
Observations	3042	3042	3042	3042	686	686	3042	3042	3040	3040

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Learning Disability is an indicator that takes the value of one if the parent of the individual reported that the individual has a learning disability such as dyslexia or attention deficit disorder. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. Results are obtained using OLS. Standard errors clustered at the individual level are presented in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively. ^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave. ^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 7A
The Impact of School Attendance on Health Behaviors, models with and without Learning Disability – First Differences

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Months Attended	-0.003** (0.001)	-0.003** (0.001)	-0.119*** (0.020)	-0.117*** (0.020)	-0.089** (0.036)	-0.087** (0.036)	-0.003*** (0.001)	-0.003*** (0.001)	-0.001* (0.001)	-0.001* (0.001)
Attend × Disability	-0.000 (0.001)	-0.000 (0.001)	0.005 (0.011)	0.005 (0.011)	-0.002 (0.025)	-0.002 (0.025)	0.001 (0.000)	0.001 (0.000)	-0.000 (0.000)	-0.000 (0.001)
Health Knowledge		-0.032 (0.053)		-1.291* (0.744)		-2.267 (1.909)		-0.035 (0.023)		-0.074** (0.033)
Age	0.014 (0.027)	0.015 (0.027)	0.203 (0.409)	0.231 (0.413)	0.373 (0.936)	0.389 (0.932)	-0.003 (0.013)	-0.002 (0.013)	-0.018 (0.016)	-0.018 (0.016)
Household Income	-0.000 (0.000)	-0.000 (0.000)	-0.004 (0.003)	-0.004 (0.003)	-0.013 (0.008)	-0.013 (0.008)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Married	-0.062 (0.071)	-0.061 (0.071)	-1.473 (0.935)	-1.407 (0.938)	-1.032 (2.050)	-0.856 (2.083)	-0.032 (0.033)	-0.031 (0.033)	-0.061** (0.028)	-0.062** (0.027)
Hours Worked	0.009 (0.007)	0.009 (0.007)	0.222** (0.106)	0.215** (0.105)	0.278 (0.224)	0.268 (0.222)	0.005 (0.004)	0.005 (0.004)	0.010*** (0.004)	0.010*** (0.004)
Household Size	-0.007 (0.008)	-0.007 (0.008)	-0.061 (0.141)	-0.059 (0.140)	0.056 (0.325)	0.063 (0.324)	-0.004 (0.004)	-0.004 (0.004)	-0.001 (0.004)	-0.000 (0.004)
Observations	1442	1442	1442	1442	554	554	1442	1442	1440	1440

Notes: Months Attended is the cumulative number of months the individual has attended any school. The outcome variables are listed at the top of columns. Odd (even) numbered columns exclude (include) Health Knowledge (Smoking or Drinking). Learning Disability is an indicator that takes the value of one if the parent of the individual reported that the individual has a learning disability such as dyslexia or attention deficit disorder. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. OLS is applied to the first differenced data. Robust standard errors are in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively.

^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

Table 7B
The Impact of Highest Grade Completed on Health Behaviors, models with and without Learning Disability – First Differences

	Smoker		Cigarettes/day		Cigarettes/day among smokers ^a		One Pack/day		Heavy Drinker ^b	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Gr. Completed	-0.019** (0.008)	-0.019** (0.008)	-0.429*** (0.104)	-0.422*** (0.103)	-0.198 (0.176)	-0.180 (0.174)	-0.010*** (0.003)	-0.009*** (0.003)	-0.009* (0.005)	-0.009* (0.005)
Highest Gr.Com × Disability	0.001 (0.009)	0.001 (0.009)	0.072 (0.104)	0.072 (0.104)	0.063 (0.236)	0.064 (0.238)	0.006 (0.004)	0.006 (0.004)	-0.000 (0.005)	-0.001 (0.005)
Health Knowledge		-0.033 (0.052)		-1.413* (0.756)		-2.349 (1.932)		-0.039* (0.023)		-0.072** (0.033)
Age	0.013 (0.026)	0.014 (0.027)	0.103 (0.407)	0.135 (0.411)	0.249 (0.925)	0.263 (0.921)	-0.006 (0.013)	-0.006 (0.013)	-0.018 (0.016)	-0.018 (0.016)
Household Income	-0.000 (0.000)	-0.000 (0.000)	-0.005 (0.003)	-0.005 (0.003)	-0.013 (0.008)	-0.014* (0.008)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Married	-0.050 (0.071)	-0.048 (0.071)	-0.669 (0.906)	-0.603 (0.913)	-0.414 (1.997)	-0.235 (2.038)	-0.008 (0.032)	-0.006 (0.032)	-0.056** (0.027)	-0.057** (0.027)
Hours Worked	0.011 (0.007)	0.011 (0.007)	0.301*** (0.105)	0.293*** (0.103)	0.305 (0.222)	0.293 (0.220)	0.007* (0.004)	0.007* (0.004)	0.010*** (0.004)	0.011*** (0.004)
Household Size	-0.008 (0.008)	-0.008 (0.008)	-0.081 (0.146)	-0.079 (0.145)	0.017 (0.335)	0.027 (0.333)	-0.004 (0.004)	-0.004 (0.004)	-0.001 (0.004)	-0.001 (0.004)
Observations	1442	1442	1442	1442	554	554	1442	1442	1440	1440

Notes: Highest Grade Completed is the number of years of schooling the individual has completed. The outcome variables are listed at the top of columns. Odd (even) numbered columns exclude (include) Health Knowledge (Smoking or Drinking). Learning Disability is an indicator that takes the value of one if the parent of the individual reported that the individual has a learning disability such as dyslexia or attention deficit disorder. Health Knowledge is measured as the share of the correct responses individual provided to the questions related to potential health risks of smoking or of heavy alcohol consumption. OLS is applied to the first differenced data. Robust standard errors are in parentheses. *, **, and *** indicate significance at 10%, 5% and 1% levels, respectively.

^a The sample include individuals who were smoker in the 1997 wave or in the 2002 wave.

^b Indicator for whether individual drinks more than 2 alcoholic drinks every day for a month as defined by Dawson (1995).

References

- Altindag, D. T., C. Cannonier and N. Mocan. 2011. "The Impact of Education on Health Knowledge." *Economics of Education Review* 30(5), pp.792-812.
- Auld, M.C., and N. Sidhu. 2005. "Schooling, Cognitive Ability and Health." *Health Economics* 14, pp. 1019–1034.
- Card, D. 2000. "The Causal Effect of Education on Earnings," in The Handbook of Labor Economics, Vol. 3, Orley Ashenfelter and David Card, eds. New York: North Holland.
- Cunha, Flavio and James J. Heckman, 2007. The Technology of Skill Formation. *American Economic Review* 97 (2), pp. 31-47.
- Cutler, D. and A. Lleras-Muney. 2010. "Understanding Differences in Health Behaviors by Education." *Journal of Health Economics* 29, pp.1-28.
- Currie, J. 2011. "Inequality at Birth: Some Causes and Consequences." NBER WP No.16798.
- Currie, J and E. Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings" *The Quarterly Journal of Economics* 118(4), pp. 1495-1532.
- Dawson, D.A., B.F. Grant, and P.S. Chou.1995. "Gender Differences in Alcohol Intake." In: Hunt, W.A., and S. Zakhari, eds. Stress, Gender, and Alcohol-Seeking Behavior. National Institute on Alcohol Abuse and Alcoholism Research Monograph No. 29. NIH Pub. No. 95–3893. Bethesda, MD: the Institute, 1995. pp. 1–21.
- Grossman, M. 2008. "The Relationship between Health and Schooling." *Eastern Economic Journal* 34, pp. 281-92.
- Grossman, M. 2006. "Education and Non-Market Outcomes." In: Hanushek, E., and F. Welch, eds. Handbook of Economics of Education, volume 1, Amsterdam: Elsevier, pp. 577-633.
- Grossman, M. 1975. "The Correlation Between Health and Schooling." In: Terleckyj, N. ed. Household Production and Consumption, volume 40 of Studies in Income and Wealth, pp. 147–211. New York: Columbia University Press for the National Bureau of Economic Research.

- Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80(2), pp. 223-255.
- Hansen K., J. Heckman and K. Mullen. 2004. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121, pp. 39-98.
- Heckman J, J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3), pp. 411-482.
- Kenkel D.S. 1991. "Health Behavior, Health Knowledge and Schooling." *Journal of Political Economy* 99(2), pp. 287-305.
- Pastor P.N., and Reuben C.A. 2008. "Diagnosed Attention Deficit Hyperactivity Disorder and Learning Disability: United States, 2004–2006." National Center for Health Statistics. *Vital Health Statistics* 10(237).
- Rosenzweig, M. R., and T.P. Schultz. 1982. "The Behavior of Mothers as Inputs to Child Health: The Determinants of Birth Weight, Gestation, and Rate of Fetal Growth." In: Fuchs, V. R., eds. Economic Aspects of Health. Chicago: University of Chicago Press, pp. 53–92.
- Sacerdote, B. 2002. "The Nature and Nurture of Economic Outcomes." *American Economic Review* 92(2). pp.344-48.

Appendix

Health Knowledge Questions in the NLSY 97 and the Correct Answers

1. Does smoking one or more packs of cigarettes per day, INCREASE THE RISK (chance) of getting heart disease?

Correct Answer: Yes

Sources:

- 1990 Surgeon General Report¹¹
- American Heart Association, <http://www.americanheart.org/presenter.jhtml?identifier=4545> (Accessed December 30, 2009)

2. Does having 5 or more drinks of alcohol once or twice each week, INCREASE THE RISK (chance) of damaging the liver?

Correct Answer: Yes

Sources:

- 1988 Surgeon General Report¹². “Excessive use of alcohol is also associated with liver disease...”
- American Liver Foundation, <http://www.liverfoundation.org/education/info/alcohol/> (Accessed December 30, 2009)

3. Does having 5 or more drinks of alcohol once or twice each week, INCREASE THE RISK (chance) of getting heart disease?

Correct Answer: Yes

Sources:

- 1988 Surgeon General Report¹³
- American Heart Association, <http://www.americanheart.org/presenter.jhtml?identifier=4488> (Accessed December 30, 2009); Cardiovascular Institute of the South, <http://www.medhelp.org/general/alcohol.HTM> (Accessed December 30, 2009)

¹¹ US Department of Health and Human Services. 1990. *The Health Benefits of Smoking Cessation: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. http://profiles.nlm.nih.gov/NN/B/B/C/T/_/nnbbct.pdf (accessed on May 19, 2010)

¹² US Department of Health and Human Services. 1988. *The Surgeon General's Report on Nutrition and Health*. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. http://profiles.nlm.nih.gov/NN/B/C/Q/G/_/nnbcqg.pdf (accessed on May 19, 2010)

¹³ US Department of Health and Human Services. 1988. *The Surgeon General's Report on Nutrition and Health*. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. http://profiles.nlm.nih.gov/NN/B/C/Q/G/_/nnbcqg.pdf (accessed on May 19, 2010)

4. Does having 5 or more drinks of alcohol once or twice each week, INCREASE THE RISK (chance) of getting arthritis?

Correct Answer: No

Sources:

- Voight, L, et al. (1994) find that “Post menopausal women who averaged more than 14 alcoholic drinks per week had a reduced risk of rheumatoid arthritis.” (p. 525)¹⁴
- Science Daily, <http://www.sciencedaily.com/releases/2007/06/070615110218.htm> (Accessed December 30, 2009)

5. Does having 5 or more drinks of alcohol once or twice each week, INCREASE THE RISK (chance) of becoming addicted to alcohol?

Correct Answer: Yes.

Sources:

- 1988 Surgeon General Report¹⁵
- American Heart Association, <http://www.americanheart.org/presenter.jhtml?identifier=4488>, <http://www.americanheart.org/presenter.jhtml?identifier=4422> (Accessed December 30, 2009)

6. Does having 5 or more drinks of alcohol once or twice each week, INCREASE THE RISK (chance) of harming an unborn child?

Correct Answer: Yes

Sources:

- 1988 Surgeon General Report¹⁶
- American Heart Association, <http://www.americanheart.org/presenter.jhtml?identifier=3017032> (Accessed December 30, 2009)

¹⁴ Voight, Lynda F, Thomas D.Koepsell, J. Lee Nelson, Carin E. Dugowson and Janet R. Daling. 1994. “Smoking, Obesity, Alcohol Consumption, and the Risk of Rheumatoid Arthritis.” *Epidemiology*, volume 88, pp. 525-532.

¹⁵ US Department of Health and Human Services. 1988. *The Surgeon General’s Report on Nutrition and Health*. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. http://profiles.nlm.nih.gov/NN/B/C/Q/G/_/nnbcqg.pdf (accessed on May 19, 2010)

¹⁶ US Department of Health and Human Services. 1988. *The Surgeon General’s Report on Nutrition and Health*. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control. http://profiles.nlm.nih.gov/NN/B/C/Q/G/_/nnbcqg.pdf (accessed on May 19, 2010)