The Puzzling Impact of Schooling on Health in Later Life: A Comparative Analysis of Common Chronic Illnesses

Sven E. Wilson

Brigham Young University

June, 2001

PRELIMINARY

Contact Information:

Departments of Political Science and Economics,

Brigham Young University

732 SWKT, Provo, UT 84602

(801)378-9018, sven_wilson@byu.edu

[0]Research support has been provided by grants from Brigham Young University and from the National Institute on Aging (AG10120). I thank John Bound, Robert Fogel, Jay Goodli¤e, John Kim, Diane Lauderdale, David Meltzer, Linda Martin, Tomas Philipson, and Tyler Shumway Ben Howell, Shawn Waddoups and, especially, Lauren Cundick provided excellent research assistance.

Abstract

This study uses the Health and Retirement study to investigate the impact of schooling on the incidence of common chronic illnesses in later life. Respondents aged 51-61 are tracked from 1992 to 1998, and the incidence of disease is recorded at two-year intervals. Schooling is shown to be highly associated with all of the eight diseases studied except cancer, with particularly high correlation for stroke, diabetes, lung disease and psychological disorders. Arthritis and hypertension are also strongly related to schooling, but to a lesser extent. However, after controlling for early-life characteristics, wealth, and post-schooling behaviors (estimated at the 1992 baseline), the schooling-health correlation is virtually non-existent for stroke, diabetes, hypertension and psychological disorders and is largely eliminated for lung disease and arthritis. For heart disease, schooling has a positive impact on disease incidence. Although the disease-speci…c di¤erences cannot be fully reconciled here, the analysis suggests that the schooling-health correlation results largely due to the behavioral pathway—the direct and indirect impacts of schooling on healthy behaviors such as smoking and diet—rather than from the direct e¤ects of early-life characteristics or wealth (which have generally small and statistically insigni…cant impact on disease incidence).

Key Words: Health production; human capital

1. Introduction

For those interested in improving the public health, simple descriptive statistics such as those presented in Table 1 are, at the least, intriguing. In Table 1, years of formal schooling are shown to be strongly associated with the prevalence of eight common chronic conditions: arthritis, cancer, diabetes, heart disease, hypertension, lung disease, psychological disorders and stroke. Similar …gures could be and have been produced for other measures of health, including global health status, disability and mortality. Indeed, a legion of studies going back to the 1960s have repeatedly shown that formal education is the most important correlate of good health.[1]

Given the robustness of the schooling-health correlation, it is not surprising that it has been extensively studied by scholars across the academic disciplines that have an interest in health. Clearly a prime motivating force for these studies is to determine whether additional investments in education will reap returns in public health. And even if the schooling-health correlation is only incidental, uncovering the roots of the correlation may point toward other policy levers that will improve health, such as reducing income inequality or increasing the ‡ow of information regarding healthy life-styles to particular sub-groups in the population. Though the simple correlations are compelling, the reasons to doubt the frequently hypothesized linkages between educational level and health are also quite plausible. Empirical regularities exist, but no consensus view has yet emerged.

Numerous complexities exist that make an understanding of the relationship between health and schooling very elusive. For instance, most studies have relied on cross-sectional data and even longitudinal data sets cover only a small portion of the life course and cannot adequately capture the e¤ect of variables that may take many years (or generations, for that matter) to develop.

[1] This is asserted by Grossman and Kaestner (1997) in their exhaustive review of the existing literature. The review of literature in the next section was signi…cantly informed by their excellent discussion.

Additionally, health itself is a multi-faceted concept that is conceived of and measured in myriad ways. The e¤ect of schooling has been studied in all commonly used health measures.

This study attempts to shed some light on the issue by examining the e¤ect of schooling on the onset of chronic illnesses in later life. Information on the eight chronic diseases shown in Table 1 is taken from the Health and Retirement Study (HRS), and the incidence of each disease is tracked over four waves of the study, from 1992 to 1998. Using duration analysis, the incidence of disease is explained by the baseline (1992) values for survey respondents. These include early-life characteristics, basic demographics, wealth and a variety of health-related behaviors.

While narrow in scope, this approach has several appealing characteristics. First, the longitudinal nature of the data reduces (though does not eliminate) problems with the potential endogeneity of regressors. Furthermore, the use of speci…c diseases, rather than more general health measures, also aids in the interpretation of the empirical results. Because diseases have di¤erent risk factors, convincing explanations for the schooling-health correlation must incorporate the unique characteristics of each disease. Finally, a focus on physician-diagnosed disease also reduces the subjectivity associated with other self-reported health measures, since the respondent merely states whether she has been diagnosed and need not interpret the illness as it presents itself in her individual case.[2]

2. Foundations: Health Production over the Life Course

2.1. A Pseudo-Recursive Model

The determination of health over the life cycle is an extremely complex process. Any honest analysis, therefore, necessities a clear delineation of the assumptions to be made. In order to …x ideas, this section presents a heuristic model of the disease prevention process that will serve two functions. First it will be used to indicate key assumptions that are made in the empirical sections

[2] This is not to say, of course, that diseases are reported without error.

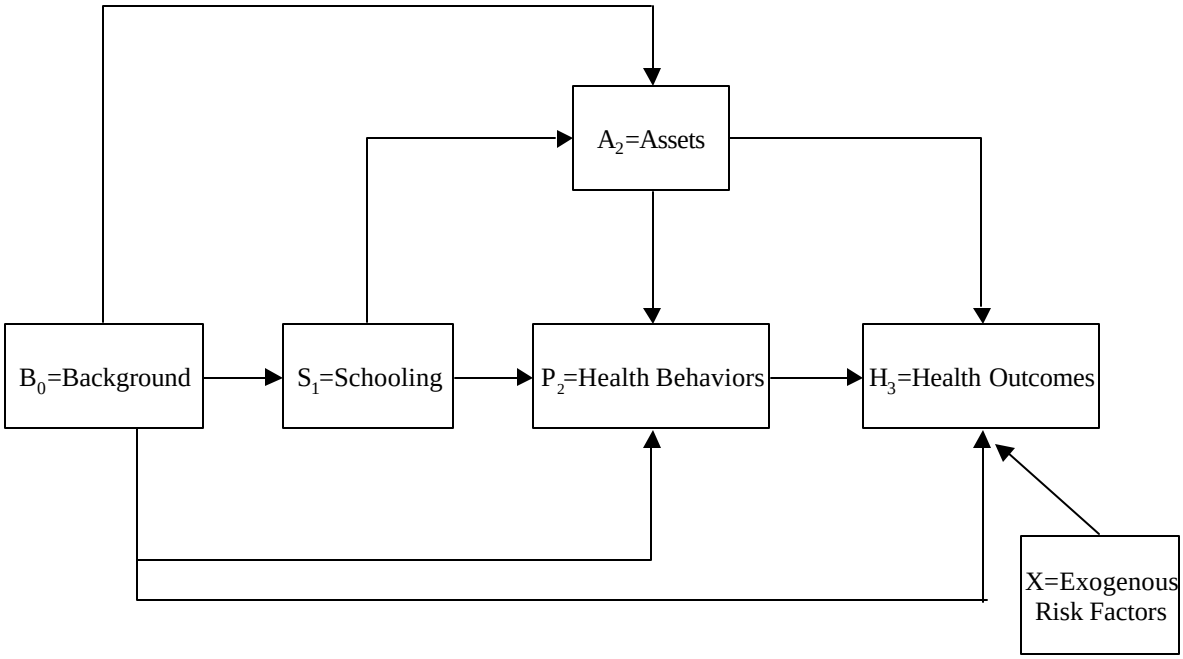

Figure 2.1: Health Production over the Life Course

to follow. And second, it will provide a starting point for discussing the plethora of theories that have been advanced to explain the robust correlation between schooling and health.

Figure 1 below represents a simple schema that is designed to highlight the central features of health production over the life course. Terminology will focus on the prevention of chronic disease, rather than on other important aspects of health, such as a general sense of well-being, disability, or life span. Furthermore, most dimensions of chronic disease will be ignored, including severity, duration, resulting disability, and medical treatment of the disease. Empirically the key variables under study will be the age of onset for each of the chronic disease groups explored.

The driving force in this model is B_0, the agent's endowment. For the purpose at hand it helps to think of this endowment in a broader sense than simply H_0, which is the initial capital stock in Grossman-type models. To wit, the endowment includes at least four different important components. First, G_0 is the genetic endowment, which is an immutable contributor to health over

the life course and to the individual's intellectual abilities. Second, P_0 is the health investment made by parents and other family measures. This may include a variety of investments in child health including nutrition, immunizations, and cigarette smoking in the household. It also includes decisions made by the parents before the child is born, such as smoking, alcohol consumption or contraction of HIV or other diseases that can be transferred to the fetus. Third, A_0 represents parental assets devoted to child welfare. And fourth, θ_0 is the set of parameters characterizing the preferences of the child. Whether these preferences are innate or in some way transferred to the child from the parents or other factors in the childhood environment is a question that is not addressed in this model.

In this model, 3 periods are assumed beyond childhood. In early adulthood, period 1, the education level, S1, is determined. Following Figure 1, this speci…cation can be represented as:

$$S_1 = S(B_0) = S(G_0, P_0, A_0, \theta_0) \tag{2.1}$$

At this level of generality, the model presumes that educational levels are determined both by levels of the speci…c endowments and by individual choice. In most models of educational choice, the agent chooses the level of education that maximizes discounted lifetime utility or, more narrowly, the net present value of his or her future income stream. Such models typically minimize the role of parental choices other than through parental assets transferred to the child. But it seems highly reasonable to assume that the wishes of parents strongly in‡uence the educational choices made by their children. In a more general framework we might consider P_0 and θ_0 as jointly determined by optimizing parents.

In the second period, the agent earns a level of income that is a¤ected by education and the child endowment (ability, bequests, etc.). In this period he or she also picks a set of preventative behaviors, P_2 (this is clearly a vector, but will be treated here, for convenience, as a scalar). At this stage of the analysis, preventative behaviors should be conceived of as broadly as possible. They

include not only the obvious decisions such as diet, exercise, smoking, drinking and physical exams, but also such actions such as residential and occupational decisions and leisure time activities. In short, P_2 includes any action that may conceivably a¤ect health.

The model is intentionally vague about consumption and saving. Expenditures that are usually considered as consumption—food, leisure activities—are included here if they in‡uence health (i.e., if they are part of P_2) but are otherwise not modeled. The level of assets, A_2, is that portion of income and accrued wealth which can be exploited to produce P_2. Although not modeled explicitly here, it makes sense to think of P_2 as an output of the household production technology. Assets are used to buy inputs to the technology, and education a¤ects the e¢ciency of those inputs (Michael 1973). Given this framework we can write:

$$A_2 = A(B_0, S_1) \tag{2.2}$$

$$P_2 = P(B_0, S_1, A_2) \tag{2.3}$$

Thus the agent, given preferences, genes, native abilities, educational level and the level of assets available based on previous consumption decisions, chooses a set of behaviors in period 2 that determine health in period 3, H_3. The risk of acquiring particular chronic diseases is also dependent on three other factors: 1) the original endowments, particularly genetic factors; 2) exogenous environmental variables that are unknown to the agent, represented by X, which is left un-subscripted since these environmental forces may occur anywhere over the life course; and 3) assets in period 2. This last e¤ect captures the possibility that socioeconomic status has a direct e¤ect on health due to "hierarchy stress," (see below) rather than indirect e¤ects through P_2. These do not include medical care, which is a preventative behavior. We can thus represent the "structural" health status equation as:

$$H_3 = H(B_0, P_2, A_2, X) \tag{2.4}$$

Schooling in this framework has no direct effect on health. Although this may seem surprising at first, the next section will make clear that the hypotheses concerning the schooling-health relationship all depend on a causal variable, such as behavior, being correlated with schooling. The only direct effect that schooling might have on health in later life is if something happens at school that affects long-term health risks, such as exposure to infectious disease.[3]

Equations 2.1–2.4 form a strictly recursive model of health production over the life cycle. I refer to it here as "pseudo-recursive" to alert the reader that this formulization is highly stylized. Numerous other linkages between the variables in the model could easily be hypothesized, and several other variables (such as values for A and P for all the periods) could be implemented. Any of these changes would likely cause the recursive structure of the model to break down. Also, though this formulation is obviously not derived from a well-specified utility maximization problem, the spirit of such models is present in this analysis.[4]

In spite of the limited formality of the model, I argue that the simple recursive structure captures many of the prominent features of the disease prevention process. In brief, the model makes reasonable claims: Agents make decisions about their schooling (or have those decisions made for them) based on a variety of early life characteristics, including their native abilities and the values acquired in their early environment. The level of schooling then affects both their economic position in later life and their health related behavior. Finally, these factors, in combination with the long-term effects of their early environment and with exogenous factors, work together to impact disease risks in later life.

[3]Costa (2000) has shown with a cohort of U.S. Civil War veterans that exposure to disease in young adulthood has significant impacts on a variety of chronic diseases in later life

[4]Grossman (2000) provides a detailed survey of the health capital literature from his original seminar articles (Grossman 1972am 1972b) to the present.

2.2. Explaining Schooling-Health Correlations

At the outset of this paper, I presented strong evidence that the presence of chronic diseases is closely associated with educational level. Obviously the ultimate question is whether a causal relationship can be inferred from these statistics. In particular, does changing an exogenous parameter (such as the price of schooling) that induces individuals to get more schooling have long-term positive health consequences? In mathematical terms, the question at hand is to determine if $\frac{dH}{dz} > 0$, where z represents the exogenous parameter.

2.2.1. Health Capital Explanations: Direct and Indirect

Economists, though relatively latecomers to the debate, have conceived of health as a stock of human capital that can be exploited in the production of utility-generating household commodities and augmented, at will, by investment. (Grossman 1972a, 1972b). Health capital models lie at the center of the recursive model above. Two channels exist through which schooling a¤ects health in the health capital model. First, it increases the agent's market wage, which results in greater earning potential. Second, it is hypothesized (Grossman, 1975) that schooling increases the e¢ciency of health investment. These two e¤ects can be encompassed as:

$$\frac{dH}{dz} = (H_P [P_S + P_A A_S]) S_Z, \tag{2.5}$$

where subscripts refer to partial derivatives. Thus S, under the health capital model, can follow a direct behavioral path through P or an indirect path through higher earnings, A. Assuming that schooling will always increase wages ($A_S > 0$), empirical con…rmation of the health capital model requires either that higher income improves behavior ($P_A > 0$) or that schooling a¤ects behavior directly ($P_S > 0$). The H_P term links the direct and indirect pathways through health behavior. Obviously it is of central importance whether the pathway is through income or household production, especially in terms of drawing accurate policy implications.

Numerous studies have been conducted that link schooling to health via some variation of the behavioral pathway discussed above. Grossman (1975) was one of the early studies to perform such an analysis. Using the NBER-Thorndike sample of men originally sample in 1955, he found that 40% of schooling's impact on survival to 1969 can be explained by the wife's level of schooling, job satisfaction and excess weight (relative to ideal weight). Leigh (1983) …nds that most of the schooling e¤ect on health can be explained by smoking, exercise and occupational choice, though Kemna (1987), in a more detailed analysis, …nds only very small impacts of occupation.

The direct pathway necessitates that $P_S > 0$. Grossman and Kaestner (1997) distinguish between two closely related behavioral responses to health knowledge that are potentially a¤ected by schooling. The …rst is "allocative e¢ciency," which is that individuals reallocate their resources as new information becomes available, and schooling, under this hypothesis, improves the knowledge people have about the relationship between health behaviors and health outcomes. The second is "productive e¢ciency," which is the extent to which schooling raises the marginal productivity of endogenous health inputs (Michael 1973). Kenkel (1991) investigates the relationship between schooling and health knowledge. He …nds that even after controlling for age, family income, race, marital status, employment status, veteran status and self-reported stress, that schooling leads to decreased smoking, decreased alcohol abuse and increased exercise. But his results also cast doubt on the behavioral pathway between schooling and health knowledge. Schooling remains a signi…cant predictor of health and inclusion of the health knowledge variables only reduce the schooling coe¢cients between 5% and 20%.[5]

Recently published work by Ross and Mirowsky (1998, 1999) attributes the largest portion of the schooling e¤ect to health behaviors. Their work …nds that it is only after including the

[5]Berhman and Wolfe (1987, 1989) look at the e¤ect of woman's schooling on family nutrition in developing countries, but their results are mixed and little can be concluded about causal mechanisms. Behrman, Rozenzweig and Taubman (1994) also …nd con‡icting results in estimating the impact of schooling on body mass index.

health behaviors that the schooling effect becomes insignificant (though in the most recent study, the schooling effects are still relatively large in magnitude relative to the schooling effects without economic and behavioral controls). In their 1999 study, Ross and Mirowsky also reject the hypothesis of Collins (1979), who argues that schooling merely provides a "credential," which leads to higher status jobs and better living standards, in favor of the quantity argument (number of years of schooling). The credential hypothesis is essentially the extreme version of the human capital hypothesis described above where P_S = 0, so that schooling can only affect health through assets and economic status. The difference is that in the credential hypothesis, number of years has no impact on A; it is only the credential from obtaining a degree that matters.

Both these studies point towards two additional contributing factors–a sense of personal control and the presence of social support.[6] They provide theoretical reasons why these factors are influenced by schooling, and in their 1999 study, both these measures are statistically significant after controlling for both economic, demographic and behavioral variables.[7] Personal control provides a link between human capital and behavior in two related ways: first, it gives individuals a variety of skills and habits that can potentially improve decision making (as other human capital theorists have argued); and second, it instills people with the belief that outcomes are contingent on one's choices and actions. In some ways, the idea of personal control and the economists' notion of productive efficiency are closely related. Both notions capture the idea that schooling gives on the ability to actually implement health information in the decision making process.

[6] A large literature linking social support to health outcomes exists. See Umberson (1987); House, Landis, and Umberson (1988); Ross, Mirowsky and Goldsteen (1990); Lack of social support also increases depression, anxiety and psychological distress (Aneshensel, Frerichs, and Huba 1984; Kessler and Mcleod 1985; Bruce and Leaf 1989;), which may lead to a deterioration in health.

[7] A central problem with these studies is the cross-sectional nature of the data, which comes from a 1995 telephone survey, "Aging, Status, and the Sense of Control Survey." It is highly feasible that the causal direction runs the other direction: good health leads to a sense of personal control and higher quality personal relationships.

The indirect e¤ect of schooling functions through income, given the well-established a¤ect of schooling on earnings. Those with greater assets then have more resources to devote towards health maintenance and disease prevention. While most studies show an e¤ect of income, wealth and other economic variables on health, these studies are in wide agreement that the education has large e¤ects on health that are not mediated by income,[8] though some studies dissent from this generalization (Duleep 1986; Behrman, Sickles, Taubman, and Yazbeek 1991).

2.2.2. Preferences-Based Explanations

A frequently cited criticism of the health capital explanation is that put forward by Fuchs. He has argued (1982) that the schooling-health correlation is not due to anything gleaned from the education experience. Instead, people who invest in education are similar to those who invest in health, particularly in terms of their rate of time preference. In terms of the nomenclature above, Fuchs' explanation of the health-schooling correlation is that P_S = 0 and $P_S S_A$ = 0 but that $S_B > 0$ and $H_P P_B > 0$. The latter two assumptions result in a positive correlation even though $\frac{dH}{dz}$ = 0.

The immediate problem with the Fuchs hypothesis is that time preference is not directly observable and is very di¢cult to infer from available data In his 1982 study that attempts to measure time preference, his time preference measure is not statistically signi…cant and is dominated by the schooling e¤ect. Farrell and Fuchs (1982) do, however, present evidence that schooling does not have a direct e¤ect on smoking behavior but that a third unobservable variable (which they conjecture is time preference) is related to both smoking and education. In the extreme, Fuch's notion seems implausible, since it is hard to envision a PhD having no inherent advantage over an illiterate person in terms of acquiring and exploiting health information. But it is also the case that health information is widely available and many common sense disease prevention techniques

[8] See, for instance, Rosen and Taubman (1982) and Ross and Mirowsky (1999),

require little formal education to undertake. Subsequent work by Leigh (1985), Berger and Leigh (1989), Leigh (1990) have also failed to support the Fuch's hypothesis. Furthermore, the work of Sander (1995a, 1995b) contradicts the Farrell and Fuchs results on smoking behavior.

A recent wrinkle in this debate has been added by Becker and Mulligan (1997), who claim that schooling actually shapes preferences, including the rate of time preference. In their model individuals realize that the present value of lifetime utility rises if the time preference for the future rises, so it is e¢cient to make investments that lower the preference for the present. They show that such investment has greater return as income or wealth rises, which implies that the income-enhancing e¤ect of education will lower the rate of time preference. Additionally, they argue that schooling may entail activities which focus the attention of young people on the future, particularly the challenges of adult life. This, in turn, will lead to greater investment in schooling. The Becker-Mulligan hypothesis cannot be accommodated in the simple model here, but note that empirically it is di¢cult to distinguish this argument from that made by Fuchs, since both follow the same pathway between B and H. To empirically di¤erentiate these two ideas would require measurement of time preference before and after the schooling experience occurs.

2.2.3. Non-Behavioral Explanations

The important elements of commonality in all the propositions discussed above are that prevention matters ($H_P > 0$) and that people make rational investments in their health. Even though Fuchs argues against a direct e¤ect of schooling, he still assumes that investment in health capital occurs; he simply attributes it to preferences rather than increased productivity. Though such assumptions appeal strongly to economists, there are other plausible explanations for why health and schooling are correlated that do no rely on these assumptions.

First, some have argued that low status in a social hierarchy is inherently stressful, so much so

that health su¤ers. This idea is central to the recent controversial work of Wilkinson (1996, 1997), which claims that it is social inequality (income inequality, in particular) that drives socioeconomic di¤erentials in health and mortality, rather than the actual levels of socioeconomic variables, though these …ndings have been challenged recently by Link and Phelan (2000), Mellor and Milyo (2001) and Milyo and Mellor (2001). Under the "hierarchy stress" hypothesis, we can de…ne

$$\frac{dH}{dz} = H_A A_S S_z. \tag{2.6}$$

Of course hierarchy stress could still lead to the health-schooling correlation if schooling were to have no direct impact on assets ($A_S = 0$). This would occur if $A_B > 0$ and $S_B > 0$.

It should be noted, however, that it is inaccurate to attribute hierarchy stress simply to assets, since the theories' proponents would include other measures of socioeconomic status as stressors, including occupation, and formal education. A broader measure of social class does not …t well into the context of the recursive model being employed here. It may be the case that schooling e¤ects not attributable to income, behavior or background can be attributed to hierarchy stress, though this implication would be di¢cult to test with existing data. The concept of personal control, discussed above, is likely a correlate of hierarchy stress.

In the discussion above, the schooling-health correlation often occurs simply because the schooling decision is a function of B_0, which a¤ects a variety of health-related variables other than schooling. The hierarchy stress hypothesis links the endowment to assets (more accurately, social status in general), and the Fuchs' hypothesis connects the endowment (particularly time preference) to behaviors. Of course other in‡uences can follow these same pathways. To be more concrete, assume that $P_S = 0$ and $H_S > 0$, meaning that schooling has no direct e¤ect on behaviors, but behavior still in‡uences health. For instance, A may a¤ect behavior if costly medical care is vital for preventing disease. Also, direct e¤ects on behavior ($P_B > 0$)[9] can occur through any of

[9] This notation is very loose here, given that B is actually a vector, as de…ned above. Here P_B should be interpreted

the elements of B_0 as de…ned above. Surely many of the key health-enhancing behaviors people take into adulthood (diet, exercise, smoking, etc.) are a direct result of the home and community environment of the child—independent of any interaction with educational decisions.

Another alternative explanation is that health is correlated with schooling due to reverse causal relationship from health to schooling. Since this study focuses on health in later life, this relationship will only be important if factors such as G_0 and P_0, which determine child health, lead directly to health in later life ($H_B > 0$) and if health in childhood reduces the level of schooling ($S_B > 0$). Studies linking schooling to child health include Edwards and Grossman (1979), Shakotko, Edwards and Grossman (1981), Perri (1984), Wolfe (1985), and Chaikind and Corman (1991)

Two …nal potential explanations deserve note. First, it is obvious that B_0 can have a direct e¤ect on health, since genetics clearly play an important role for many health conditions. Parental decisions and parental assets may also have a direct e¤ect if the health and nutritional experience of childhood have e¤ects in later life. The work of Barker (1998) implies that even the pre-natal environment of an individual may have an impact on health in later life. If the elements of B_0 also a¤ect the level of education then health and schooling will be correlated. Second, the social, economic and political forces that in‡uence X will also be correlated with H. For example, a community that subsidizes higher education may also be more likely to tightly regulate the air pollution that leads to respiratory disease or be more likely to o¤er free blood pressure tests.

2.3. Implications for the Study of Chronic Illness

Numerous studies in the biomedical literature include schooling as a risk factor in the study of disease epidemiology. But this literature does not frequently attempt to uncover the mechanisms by which schooling impacts disease risk. Furthermore, this study is novel in its comparative analysis

as the partial derivitive of an element P with respect to any of the elements of B.

of speci…c conditions within the context of a common data set and a common methodology. Pincus, Callahan and Berkhauser (1987) are a previous example of this approach. Theirs is a cross-sectional analysis of 37 chronic conditions reported in the 1978 Survey of Disability and Work. Of the 23 conditions reported by at least 1% of the sample, they …nd signi…cant correlation with schooling choice for 19, though they employ a very limited set of additional controls (age, sex, race and smoking). Wilson (1997) employs a similar approach with 14 disease groups using cross-sectional data from the New Jersey Demographics of Disability Survey conducted in 1992.

A focus on illnesses of later life has several advantages when investigating the impact of schooling on health, not the least of which is that disease-speci…c e¤ects are understudied. Because diseases have di¤erent risk factors, any theory of schooling impacts—particularly one that tries to identify the behavioral pathway from schooling to health outcomes—needs to be consistent with all types of diseases. For instance, if schooling improves health because it augments health-enhancing behaviors, then diseases that have few behavioral risk factors should have smaller schooling e¤ects than diseases with strong behavioral risks. In sum, a comparative analysis of di¤erent diseases increases the number and variety of cases that can be explained by a given theory. Theories that are successful in explaining disease-speci…c impacts are more convincing than theories that only explain general health measures, such as mortality, disability or general health status.

3. Methods

3.1. Data

The data used come from the Health and Retirement Study (HRS). The HRS is a nationally representative sample of the U.S. population begun in 1992. Face to face interviews were completed with all respondents and their partners in 1992, and respondents were re-interviewed by phone every

two years since that time. The analysis here, therefore, covers the period from 1992 to 1998.[10] Sample respondents are between the ages of 51-61 at the time of interview, though partners outside of that age range were also interviewed. To insure representativeness, only age-eligible respondents are included in this analysis.

A detailed set of health questions is asked during each interview, including whether the respondent had ever been diagnosed with particular diseases categories. The eight diseases used in this analysis were selected because they were consistently queried across the four waves of the survey. Some of the diseases are very speci…c, such as hypertension; others are extremely broad, such as lung disease, which of course covers numerous speci…c diagnoses. Others are narrow in scope, but encompass more than one speci…c diagnostic category. Arthritis, for instance, is not di¤erentiated between osteoarthritis and rheumatoid arthritis.

Several covariates are available for estimating the health equations. In addition to years of schooling, demographic variables include sex, age, race, parental characteristics, and marital status. Extensive income, wealth and employment information is available in the HRS. The intent here is to summarize the economic welfare (A) of the respondent during adulthood. Therefore, I use net household wealth at the baseline of the survey as a measure of economic status. I also include a dummy variable indicating whether or not the respondent has health insurance. Insurance could be interpreted as either a measure of …nancial well-being or as a behavioral choice. Given the age group used in this survey, I interpret it as a measure of …nancial well-being.

Health related behaviors, P, include measures of smoking, drinking, exercise and diet. Smoking is measured by the number of cigarettes smoked both for current smokers and previous smokers. Daily alcohol consumption, as measured by number of drinks, is included, as is a measure of exercise,

[10] The data from wave 4 (1998) is still in preliminary form and may change as HRS researchers prepare the data for o¢cial public release.

which incorporates both the frequency and intensity of physical activity. Diet is proxied with body mass index (BMI). BMI, which is weight (in kilos) divided by the square of height (in kilos), is a widely used measure of obesity. Obviously it contains a large amount of noise as a proxy for diet, since weight is a function of food intake, physical activity and basic metabolism, which has a strong genetic component.

Wave 4 (1998) of the HRS contains a set of questions that ask the respondents about their early life conditions, which are used to capture some of the potential variables in the B vector. These include health status, socioeconomic status, whether the father was frequently unemployed, and whether the family moved frequently. Using this set of variables is problematic since they were not collected at baseline and the data is not collected for those who died or dropped out of the survey before the …nal wave of data was collected. These issues will be discussed further in the next section.

3.2. Estimation

Ideally, estimation of health models would yield coe¢cient estimates that are consistent, e¢cient and unambiguously interpretable. In practice, unrealistic (but necessary) assumptions must be made, samples are created with unavoidable and often unknowable biases, and numerous important variables are either missing or poorly measured. Health econometrics is particularly challenging because health outcomes are the result of complex interactions of numerous variables over numerous decades. Section 2 pointed to the importance of many assumptions, and this section moves from that general theoretical discussion to a discussion of more speci…c empirical issues.

Although still a relatively "young" data set, the HRS has accumulated enough waves of data to initiate studies on the dynamic aspects of health production. The analysis here does not pretend to uncover all such dynamic structures, but it is possible to mitigate several of the problems associated

with the cross-sectional results. In a longitudinal context it is possible to estimate the marginal impact of schooling on disease incidence by imposing controls for health at the baseline period, K. Furthermore, the direction of causality between covariates and the dependent variable is easier to infer in the longitudinal context.

The longitudinal estimates will be obtained in the context of a Cox proportional hazard model. In this common speci…cation, the hazard rate (the probability of acquiring the disease in period t conditional on not acquiring it before period t) is speci…ed as $h^j(t)$. Of central concern here is estimation of the disease equations. The model of the previous section suggests the following speci…cation of the Cox model:

$$h_i^j(t) = h_0^j(t)\exp(\alpha_0^j + \alpha_1 X_i + \alpha_2^j S_i + \alpha_3^j B_i + \alpha_4^j A_i + \alpha_5^j P_i + \alpha_6^j K_i + e_i^j). \tag{3.1}$$

where $h_0^j(t)$ represents the baseline hazard that is shifted in a proportional fashion by the covariates. Individuals are indexed by i,diseases by j. The disease variables are binary indicators representing whether or not the respondent reports that a physician has ever diagnosed her with the disease. The α coe¢cients (or vectors of coe¢cients) are assumed to vary across diseases (since each disease has di¤erent risk factors) but are constant across individuals within a disease category. Given the recursive structure of the model, the covariates in 3.1 are treated as pre-determined variables, even though they may be jointly determined, as the case with A and P. The coe¢cients can be consistently estimated as long as the e_i^j are uncorrelated with the covariates of the model. All independent variables are measured at wave 1 values. Naturally, cases where the disease already exists at wave 1 are excluded from analysis. The wave 1 health variables include both indicators for the other seven wave 1 indicators of disease as well as self-assessed general health status (SAGHS).[11]

[11] Where repondent's rate their own health as excellent, very good, good, fair or poor. Dummy variables for each value are used in the regression.

Substituting 2.2 and 2.3 into 3.1 obtains the following reduced-form equation[12]:

$$h^j(t) = h_0^j(t)\exp(\pi_0^j + \pi_1 X_i + \pi_2^j S_i + \pi_3^j B_i + u_i^j). \quad (3.2)$$

If a complete set of variables exists in X and B then this equation can be consistently estimated. Estimates of these reduced form equations will be provided as part of the analysis on the causal pathway between health and schooling, though it is still likely that important variables are omitted from B that will cause S and u to be correlated.

The data in this study are measured at 2-year intervals for 3 periods following the baseline interview. These characteristics do not make for ideal analysis of Cox models, given the high number of ties that exist and the short time horizon. The Cox model, however, is advantageous because it can account for censoring of the data due to the competing risks of death and loss of follow-up. An additional complication is that the hazard for mortality is not independent of disease hazard. The analysis to be reported here has been performed, for comparative purposes, with simple probit analysis where the dependent variable is the occurrence of disease anytime during the six-year period and the cases are restricted to those who lived and were not lost to follow-up during the six year period. The results of this analysis, not shown here, are very similar to the Cox results reported in the next section.

3.2.1. Uncovering Health Pathways

The discussion above points to two estimating equations for each type of analysis: a "complete" structural equation and a reduced-form equation. The reduced-form equation represents the total impact of schooling, while the structural equation represents the impact of schooling after assets

[12] Technically, the recursive model implies that S should not be in the reduced-form equation. However, assuming that S is completely determined by equation (X), this speci…cation has the properties of a reduced-form equation and captures the "total" e¤ect of S on disease prevalence.

and behavior–the primary hypothesized pathways–have been added to the model. In the simplest version of the health capital model, the coe¢cients on S, B, and X, will all be zero once the appropriate set of behaviors, P, is included. Non-zero coe¢cients represent either non-behavioral pathways or missing data.

In addition to these two equations, several other "sub-models" will be estimated to illustrate the e¤ect of adding additional covariates to the speci…cation. In total, seven longitudinal models will be estimated. This step-wise approach for uncovering mediating and confounding variables is very common and has been recently been exploited in the schooling-health debate in the work of Ross and Mirowsky (1999). Similar analysis was followed in Grossman (1975) and Leigh (1987). It should be noted, however, that this approach cannot be construed to make de…nitive comparisons across the di¤erent covariates in the model. The impact of any one variable depends critically on the order in which it is added to the model, and (if one believes that the structural model is correctly speci…ed), the sub-models will yield biased and inconsistent estimates due to omitted variables. Still, if the variables are added in a common sensical manner, such as the order in which the present themselves over the life course of an individual, they can be informative about the underlying causal mechanisms even if they can never be de…nitive.

Finally, uncovering the pathway between schooling and health also necessitates the estimation of particular links between key variables in the model. All such relationships will not be estimated here, but two important ones will be. First, a regression of years of schooling on background variables will be used to ascertain the usefulness of the variables in B as a proxy for the background characteristics of the agent. If such a link exists, then a variety of possible explanations for "incidental" correlation between health and schooling exist, since these all rely on $S_B > 0$, as discussed earlier. This regression is stated simply as:

$$S = \lambda_0 + \lambda_1 B_i + \omega_i \tag{3.3}$$

Second, the relationship between schooling and health behaviors is fundamental to any behavioral pathway. Particularly important is the effect of education on behavior after controlling for assets. This implies a set of regressions for each health behavior. If we index these by k, we can define the following set of regressions:

$$P_i^k = \delta_0 + \delta_1 X_i + \delta_2^k S_i + \delta_3^k B_i + \delta_4^k A_i + \mu_i^k. \quad (3.4)$$

Regressions for BMI will be estimated by OLS, while those for drinking, current smoking and exercise will be estimated by ordered probit.

4. Results

The equations discussed in the previous section generate a very large number of regression equations–far too many to discuss in detail here. The complete set of regressions is included in the appendix to this article, and selected coefficients are presented and/or discussed here, with the primary attention focused on the schooling coefficients. All estimation is performed with the STATA 7 statistical package. All test-statistics are calculated using robust (heteroskedasticity-consistent) standard errors.

4.1. Estimates of Schooling Impacts

Table 2 presents the schooling coefficients for the proportional hazard analysis discussed above. In each case a series of sub-models is estimated. Briefly, these are:

- M1: Education alone
- M2: M1 + Age & sex
- M3: M2 + Race/ethnicity

- M4: M3 + Early life characteristics (health and socioeconomic status as children, parental education)
- M5: M4 + Assets (household wealth in 1991 and health insurance dummy)
- M6: M5 + Health behaviors (BMI, smoking, drinking, exercise and marital status)
- M7: M6 + Baseline Health (self-assessed general health status and baseline disease indicators)

The speci…cation of these models follows the model of section 2, though clearly other reasonable approaches could be taken. Of particular note is the treatment of race/ethnicity and marital status. Although race and ethnicity are clearly an early-life characteristic, it is not clear at what point in the life-cycle they a¤ect health. Race has been identi…ed as a risk factor for particular diseases, such as hypertension, but it is unclear whether the racial di¤erences are due to di¤erent environmental and behavioral risk factors across the life cycle or whether there are genetic di¤erences between groups that a¤ect disease risk. Marital status is another demographic variable that has shown to be correlated with health in a variety of studies. Although marriage often occurs early in life and, therefore, should possibly be included in M3, changes in marital status also represent forces at work across long periods of time. Since the marriage variable is measured at the baseline of the study, it is considered here as a behavioral variable. In separate analysis (not shown) the inclusion of the marital status variable, on its own, has very trivial impacts the schooling coe¢cients

Because the early life variables were not collected until the fourth wave, individuals who died or were lost to follow-up are dropped from the regressions associated with models M4-M7. Models M4a-M7a are analogous to M4-M7 except for the fact that they exclude the variables that are only present in wave 4. Thus these alternative models more accurately capture the impact of censoring due to death or exit from the survey since individuals are not constrained to remain in the sample until wave 4.

Several important patterns can be drawn out of the Table 2. First, the unconditional correlation (M1) of schooling with disease incidence is both strikingly large and highly signi…cant in a statistical sense. For example, just one additional year of school reduces the hazard of lung disease by 8.4%. Since these estimates do not capture the cumulative declines in health over the course of several decades, but instead re‡ect the marginal declines over a relatively short time period, their size of the estimates is particularly surprising. The only exception to this pattern is cancer, which has virtually no association with the level of schooling. Furthermore, if it is the case that educated people are more likely to detect the presence of a disease (from, say, more frequent visits to the doctor), then the simple negative association between schooling and disease is even more striking.

Second, the magnitude of the education e¤ect varies signi…cantly across diseases. We expect diseases that have signi…cant behavioral risk factors to have the strongest education e¤ects. This is certainly the case for lung disease, stroke and diabetes. But no disease is more frequently studied and discussed than heart disease, and the simple correlation of schooling with heart disease is dramatically smaller than for these other diseases. After controlling for intervening variables, the e¤ect of schooling on heart disease actually becomes positive. Equally surprising is the very high correlation of schooling with psychological disorders, though Pincus, Callahan and Burkhauser (1987) also …nd a very strong correlation between education and mental health. Arthritis and hypertension are also signi…cantly correlated with schooling, but to a lesser extent. Finally, schooling seems to have no impact in cancer, even though a wide variety of dietary and environmental risk factors for cancer have been proposed in the popular media. In fact, the impact of schooling on heart disease is modestly positive, after controlling for other covariates in the model.

Third, background is an important determinant of the schooling-health correlation for some, but not all, of the diseases analyzed. Roughly half the e¤ect of schooling is explained by the background variables for stroke, diabetes and hypertension, and modest e¤ects are seen for psychological

disorders, arthritis, while virtually no e¤ect is seen for heart and lung disease.[13] These …ndings suggest that importance of background as an important factor in explaining the health-schooling correlation,[14] but further work needs to be done to account for disease-speci…c e¤ects.

Fourth, the schooling e¤ects remain strong even after controlling for background characteristics such as race and childhood socioeconomic status. This is seen in Model M4, which is the reduced-form equation 3.2 discussed previously. Again heart disease and cancer are the exceptions.

Fifth, the inclusion of both assets and health behaviors reduce the magnitude of the schooling e¤ects. The inclusion of asset values changes the schooling e¤ect by a notable degree for all diseases except arthritis, and health behaviors have even further impacts. Further discussion of the role of assets and health behaviors will be taken up shortly.

Finally, the extent to which the schooling coe¢cient can be reduced by the inclusion of variables that generally occur after the schooling period (asset accumulation and health behaviors, (models M5-M7) is quite remarkable. The e¤ects of schooling on stroke, diabetes and hypertension are completely eliminated,[15] and the e¤ects on lung disease and psychological disorders are mostly eliminated when controls are included for baseline health. (the remaining e¤ects are statistically insigni…cant and in the range of 2-3%). To the extent that schooling has long-term bene…cial health e¤ects, these e¤ects work primarily through the post-schooling variables in the model, a …nding even more remarkable in light of the relatively short list of mediating variables included in the analysis.

[13] Lung disease is anamalous in that inclusion of the race variable in M3 raises (in absolute value) the e¤ect of schooling. But once the othe early life charactersitics are included, the background variables together change the schooling e¤ect by only .1 percentage points. (from -8.4% to -8.3%)

[14] Recall that the "incidental" correlation between health and schooling depends on S_B.

[15] Even in M6, which doesn't control for baseline health, the schooling coe¢cients for these three diseases are very small

4.2. Correlates of Chronic Disease Incidence

The complete regressions tables for all the sub-models are included in the appendix to this document. Table 3 is a brief summary of M7 for each of the eight diseases, where the coe¢cients are summarized by a + or - , indicating the direction of the e¤ect, with + sign indicating that the covariate increases the hazard of getting the disease. As noted above, the schooling e¤ects are largely eliminated after additional controls are put in place.

Age, sex and race are signi…cant determinants of disease in later life, which is hardly surprising, though the strength of these relationships after controlling for baseline health and for health behaviors remains somewhat a puzzle. Early childhood variables, on the other hand, have little long-term e¤ects on disease. The notable exception to this is are psychological disorders, which are also increased by having an absent father in childhood. Some signi…cant long-term e¤ects of childhood health are also found for arhtritis, though the absence of these e¤ects at the poorest levels of health suggest that the other e¤ects are anomalous. It is notable that, in general, the signs of the childhood health e¤ects are what we would expect—poor health in childhood increaes the incidence of disease in later life, though the e¤ects are neither statistically signi…cant nor substantial.

Economic variables, whether in childhood or at baseline are largely unimportant. Thus the e¤ect of schooling on income and wealth seems to be transferred primarily through health behaviors. Again there are some anomalous …ndings for stroke and arthritis for the childhood SES variables, but the pattern of the e¤ects is not consistent. Parental education variables have no long-term e¤ects on health in this analysis

In terms of health behaviors, the dominant story is cigarette smoking. Srong and signi…cant e¤ects exist for all the diseases except arthritis and hypertension. The lack of smoking e¤ects for arthritis is to be expected, since previous studies have not established a role for smoking. The e¤ects of previous smoking are much smaller, but are generally of the right sign. The other health

behaviors with relatively strong e¤ects is BMI. However, the impact is restricted to the cases of arthritis, hypertension and diabetes.

Baseline health variables are also highly signi…cant. Since these variables capture the e¤ect of health across the life course, it is possible that the e¤ect of other covariates actually appear through the health variables. Other chronic diseases at baseline also have sign…cant e¤ects. Some are anticipated, such as the impact of hypertension on heart disease and stroke, while others have no apparent explanation, such as the impact of arthritis on lung disease. Even though baseline health is important, most of the reduction in schooling e¤ects, as noted above, occurs without the inclusion of the baseline health variables in M7. The signi…cant health variables point towards an accelerating rate of decline in health, in that those who are already in the worst health are more likely to su¤er additional declines in health. However, this accelerated rate of decline does not seem to be a¤ected by schooling, other than through the behavior covariates in the model.

Estimates of equations 3.1 and 3.2 are given in Table 4 and 5, respectively. These tables are designed to estimate two important branches of the model picture in …gure 2.1. First, Table 4 provides an esimate of the impact of family background variables on schooling. The R^2 on this regression is .30, which is remarkably high given the limited number of variables and the number of observations. It shows that childhood charachteristics (including race and sex) are highly correlated with the schooling levels of the surviving members of this cohort. Being in poor health and in poor socioeconomic status as a child reduces years of schooling by a combined 2.55 years, which is close to a full standard devation (S.D.=3.26). Thus all the "alternative" explanations for the schooling-health correlation that relied on the correlation between background and schooling are de…nitely plausible, given the results in Table 4.

Finally, Table 5 shows the impact of schooling, other early life behaviors and wealth on the four health behaviors included in the analysis. BMI is estimated with OLS, and the other three

are estimated by ordered probit. Positive coe¢cients represent increased body mass, increased (current) smoking, increased drinking, and increased exercise. For each behavior, the coe¢cients for both schooling and wealth are highly signi…cant in each case. It should be noted, however, that the estimated e¤ects of these variables are small and the models as a whole explain only about 3% of the variation in the behavioral variables. Thus a variety of forces other than either schooling or wealth are determining the behavioral variables.

5. Conclusions

Which of the variety of theories discussed in Section 2 …nd empirical support from the analysis here, and which are rendered suspect? If anything, the results here seem to imply that almost everything matters, but no one explanation dominates. I …nd two of the results here particularly striking. First, the simple correlation between years of schooling and the incidence rate of disease in later life are very high for six of the eight diseases studied. It is not surprising that health status at a point in time is correlated with schooling, but that the incidence rates of disease (the rate of decline, in other words) should be so strongly associated with an early life variable such as disease is not necessarily expected. Second, this correlation is completely eliminated or mostly eliminated by including the post-schooling variables for all the diseases, even though the variables included are far from what an exhaustive list might contain.

Evidence for direct e¤ects, through behavior, on health is compelling in this analysis. Each piece of the puzzle …ts together. Education is shown to impact behavior, and behaviors, particularly smoking and diet, in‡ucence health. However, it is also the case that wealth in‡uences health behaviors. Thus both the direct and indirect components of the health capital pathway in equation 2.5 can be supported. Furthermore, since background variables have a strong impact on both schooling and income, a strong case can be made that the background variables—genetics, partental

investments in health, parental socioeconomic status and early-life preference formation—are very important in explaining the schooling-health correlation. These variables on their own explain a substantial chunk of the simple correlation for most diseases.

It does not appear to be the case, however, that the background variables have a signi…cant direct impact on disease incidence. Wealth also does not have a direct impact. Thus the schooling-health correlation may be largely "incidental," but the pathway is, nonetheless, through the post-schooling behavior of the agent. This implies that no role is found for the hierarchy stress hypothesis that has received so much attention recently. The Fuch's hypothesis, on the hand, remains highly plausible, especially if one believes that the childhood variables that are shown to have such strong impacts on schooling are also found to be correlated with the rate of time preference.

Finally, if the diseases under analysis were limited to stroke, diabetes, hypertension and psychological disorders, it would seem very straightforward that schooling a¤ects disease (either directly or indirectly through income) by alterning post-schooling behavior. The addition of lung disease and arthritis, would pose a bit of a puzzle, since reasonably large schooling e¤ects remain even after controlling for other covariates. But the inclusion of heart disease and cancer throw a considerable wrench in works. Particularly troubling is heart disease, which has strong behavioral risk factors. While others have found a similar relationship between schooling and heart disease, none of the theories advanced can readily explain why heart disease should di¤er so markedly from other diseases such as lung disease or stroke. These …ndings point to the importance of pushing our theories to explain not only health in general, but also to account for the extensive variation in the types of health conditions that people face in later life.

References

Aneschensel, C.S., R. Frerichs, and G. Huba. 1984. "Depression and Physical Illness: A Multiwave, Nonrecursive Causal Model." *Journal of Health and Social Behavior* 25:350-71.

Barker, D.J.P. 1998. *Mothers, babies, and health in later life*. 2d ed. Edinburgh; New York: Churchill Livingstone.

Becker, Gary S. and Casey B. Mulligan. 1997. The Endogenous Determination of Time Preference. *Quarterly Journal of Economics* 112(3), August 1997, pp. 729-58.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102(6):1131-1174.

Behrman, Jere R., Robin Sickles, Paul Taubman, and Abdo Yazbeck. 1991. "Black-White Mortality Inequalities." *Journal of Econometrics* 50:183-203.

Behrman, Jere R. and Barbara L. Wolfe. 1987. "How does Mother's Schooling Affect Family Health, Nutrition, Medical Care Usage, and Household Saniation?" *Journal of Econometrics* 36:185-204

Behrman, Jere R., and Barbara L. Wolfe. 1989. "Does More Schooling Make Women Better Nourished and Healthier? Adult Sibling Random and Fixed Effects Estimates for Nicaragua." *Journal of Human Resources* 24(4)644-663.

Berger, M.C. and J.P. Leigh. 1988. "Schooling, Self-Selection, and Health." *Journal of Human Resources* 24(3):434-454.

Bruce, M.L. and P.J. Leaf. 1989. "Psychiatric Disorders and 15-Month Mortality in a Community Sample of Older Adults." *American Journal of Public Health* 79:727-30.

Chaikind, Stephen, and Hope Corman. 1991. "The Impact of Low Birthweight on Special Education Costs." *Journal of Health Economics* 10:291-311.

Costa, Dora L. 2000. "Understanding the Twentieth-Century decline in Chronic Conditions Among Older Men." *Demography* 37(1): 53-72.

Duleep, H.O. 1986. "Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data." *Journal of Human Resources* 21:238-51.

Edwards, L.N., and M. Grossman. 1979. "The Relationship Between Chidren's Health and Intellectual Development." In *Health: What Is It Worth*?, ed. S.J. Muschkin and D.W. Dunlop. Elmsford:Pergamon Press, pp. 273-314

Farrell, P. and V.R. Fuchs. 1982. "Schooling and Health: The Cigarette Connection." *Journal of Health Economics* 1:217-30.

Fuchs, Victor R. Time Preference and Health: An Exploratory Study. In *Economic Aspects of Health*, ed. Victor R. Fuchs. Chicago: University of Chicago Press for NBER), pp. 93-120.

Grossman, Michael. 1975. "The Correlation between Health and Schooling." In *Household Production and Consumption*, ed. N.E. Terleckyj, Studies in Income and Wealth, Vol. 40, Conference on Research in Income and Wealth. New York: Columbia University Press for the National Bureau of Economic Research. Pp. 147-211.

Grossman, Michael. 2000. "The Human Capital Model." In *Handbook of Health Economics, Vol.1A*, ed., Anthony J. Culyer and Joseph P. Newhouse. Amsterdam, The Netherlands: Elsevier Science B.V., pp. 347-405.

Grossman, Michael. 1972a. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80:223-55.

Grossman, Michael. 1972b. *The Demand for Health: A Theoretical and Empirical Investigation*. New York: Columbia University Press for the National Bureau of Economic Research.

Grossman, Michael and Robert Kaestner. 1997. Effects of Education on Health. In *The Social Benefits of Education*, ed., Jere R. Behrman, David. L. Crawford, and Nevzer Stacey. Ann Arbor: University of Michigan Press, 1997, pages 247-53.

House, James S., Karl R. Loandis, and Debra Umberson. 1988. "Social Relationships and Health." *Science* 241:540-45.

Kenkel, Donald S. 1991. "Health Behavior, Health Knowledge, and Schooling." *Journal of Political Economy* 99(2):287-305.

Kemna, Harrie J.M.I. 1987. "Working Conditions and the Relationship Between Schooling and Health." *Journal of Health Economics* 6:189-210.

Kessler, R.C. and J.D. McLeod. 1985. "Social Support and Mental Health in Community Samples." *In Social Support and Health*, ed., S. Cohen and S.L. Syme. New York:Academic, pp. 219-40.

Leigh, J. Paul 1983. "Direct and Indirect Effects of Education on Health." *Social Science Medicine* 17(4):227-234.

Leigh, J. Paul. 1985. "An Empirical Analysis of Self-reported, Work-Limiting Disability." *Medical Care* 23:310-19.

Leigh J. Paul. 1990. "Schooling and Seat Belt Use." *Southern Economic Journal* 57:195-207.

Link, Bruce G., and Jo C. Phelan. 2000. "Evaluating the Fundamental Cause Explanation for Social Disparities in Health." In *Handbook of Medical Sociology, 5th Ed*., ed., Chloe E. Bird, Peter Conrad, and Allen N. Fremont. Upper Saddle River, NJ: Prentice Hall, pp.33-46.

Mellor, Jennifer M. and Jeffrey Milyo. 2001. "Re-Examining the Evidence of an Ecological Association Between Income Inequality and Health." *Irving B. Harris Graduate School of Public Policy Studies, Working Paper Series: 00.17.*

Milyo, Jeffrey and Jennifer M. Mellor. 2001. "Is Inequality Bad for Your Health?" *Irving B. Harris Graduate School of Public Policy Studies, Working Paper Series: 00.16.*

Perri, T.J. 1984. “Health Status and Schooling Decisions of Young Men.” *Economics of Education Review* 3:207-13.

Pincus, Theodore, Leigh F. Callahan, and Richard V. Burkhauser. 1986. “Most Chronic Diseases Are Reported More Frequently By Individuals with Fewer Than 12 Years of Formal Education I the Age 18-64 United States Population.” *Journal of Chronic Diseases* 40(9):865-874.

Rosen, Sherwin and Paul Taubman. 1982. “Some Socioeconomic Determinants of Mortality.” In *Economics of Health Care*, ed. J. van der Gaag, W.B. Neenan, and T. Tsukahara Jr. New York: Praeger, pp. 255-271.

Ross, Catherine E., and John Mirowsky. 1998. “Education, Personal Control, Lifestyle and Health: A Human Capital Hypothesis.” *Research on Aging* 20(4):415-449.

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

Ross, Catherine E., J. Mirowsky, and K. Goldsteen. 1990. “The Impact of the Family on Health: The Decade in Review.” *Journal of Marriage and the Family* 52:1059-78.

Sander, W. 1995a. “Schooling and Quitting Smoking.” *Review of Economics and Statistics* 77:191-99.

Sander, W. 1995b. “Schooling and Smoking.” *Economics of Education Review* 14:23-33.

Shakotko, R.A., L.N. Edwards, and M. Grossman. 1981. “An Exploration of the Dynamic Relationship Between Health and Cognitive Development in Adolescence.” In *Contributions to Economic Analysis: Health, Economics, and Health Economics*, ed. J. van der Gaag and M. Perlman. Amsterdam:North-Holland Publishing Company, pp.305-25.

Umberson, Debra. 1987. “Family Status and Health Behaviors: Social Control as a Dimension of Social Integration.” *Journal of Health and Social Behavior* 28:306-19.

Wilkinson, Richard G. 1996. *Unhealthy Societies: The Afflictions of Inequality*. London: Routledge.

Wilkinson, Richard G. 1997. “Health Inequalities: Relative or Absolute Standards?” *British Medical Journal* 413:591-95.

Wolfe, B.L. 1985. “The Influence of Health on School Outcomes: A Multivariate Approach.” *Medical Care* 23:1127-38.

Table 1: Disease Prevalence by Years of Schooling

	Years of Schooling				
Disease	0-11	12	13-15	16	17+
Arthritis	46.6%	38.8%	34.6%	28.5%	26.9%
Hypertension	42.4%	37.7%	36.9%	34.7%	33.4%
Psych. Disorders	16.3%	10.3%	9.5%	8.9%	6.8%
Heart Disease	16.0%	12.5%	11.8%	11.0%	10.0%
Diabetes	14.2%	9.2%	8.6%	7.4%	6.9%
Lung Disease	12.8%	7.9%	7.3%	3.5%	4.8%
Cancer	6.0%	5.7%	5.6%	5.7%	4.4%
Stroke	3.9%	2.3%	2.4%	1.4%	2.0%

N=9825

Notes: Data are from 1992 Health and Retirement Survey. Includes men and women aged 51-61.

Table 2: Schooling Effects: Proportional Hazard Results

Percentage Change in the Hazard Rate Attributable to One Additional Year of School:

Model	Stroke	Diabetes	Psych. Disorders	Lung Disease	Hypertension	Arthritis	Heart Disease	Cancer
M1: Schooling Alone	-10.0%	-9.1%	-9.1%	-8.4%	-5.1%	-4.9%	-2.6%	-0.2%
	(-6.26)	*(-8.55)*	*(-8.84)*	*(-7.47)*	*(-5.77)*	*(-7.50)*	*(-2.54)*	*(-0.14)*
M2: M1 + Age + Sex	-9.5%	-8.8%	-9.3%	-8.2%	-5.1%	-4.7%	-2.4%	0.1%
	(-5.92)	*(-8.32)*	*(-8.92)*	*(-7.26)*	*(-5.80)*	*(-7.03)*	*(-2.41)*	*-0.06*
M3: M2 + Race	-8.2%	-6.5%	-9.2%	-11.4%	-3.9%	-5.0%	-3.2%	-0.9%
	(-4.07)	*(-5.15)*	*(-7.59)*	*(-8.15)*	*(-4.05)*	*(-6.56)*	*(-2.74)*	*(-0.53)*
M4: M3 + Early Life Variables	*-5.8%*	*-5.6%*	*-7.6%*	*-8.3%*	*-2.7%*	*-3.9%*	*-2.0%*	*1.8%*
(Reduced Form)	*(-2.36)*	*(-3.61)*	*(-5.30)*	*(-4.64)*	*(-2.26)*	*(-4.29)*	*(-1.33)*	*-0.83*
M5: M4 + Assets	-3.6%	-4.3%	-6.0%	-6.2%	-2.1%	-3.5%	-0.6%	3.1%
	(-1.41)	*(-2.62)*	*(-3.96)*	*(-3.25)*	*(-1.71)*	*(-3.67)*	*(-0.41)*	*(-1.33)*
M5a: M5 w/o Early Life Vars.	-4.4%	-5.4%	-6.3%	-7.6%	-2.3%	-3.9%	-0.2%	0.0%
	(-1.89)	*(-3.67)*	*(-4.57)*	*(-4.65)*	*(-2.06)*	*(-4.41)*	*(-0.14)*	*(-0.02)*
M6: M5 + Health Behaviors	-1.5%	-1.6%	-5.2%	-4.0%	-1.4%	-3.1%	1.2%	3.6%
	(-0.57)	*(-0.92)*	*(-3.29)*	*(-1.92)*	*(-1.10)*	*(-3.09)*	*(-0.73)*	*(-1.56)*
M6a: M6 w/o Early Life Vars.	-2.7%	-2.8%	-5.3%	-5.4%	-1.7%	-3.5%	1.8%	0.6%
	(-1.09)	*(-1.81)*	*(-3.59)*	*(-3.05)*	*(-1.49)*	*(-3.83)*	*-1.27*	*-0.29*
M7: M6 + Baseline Health	0.7%	-0.4%	-3.1%	-2.0%	-0.7%	-2.3%	2.9%	4.3%
	(-0.24)	*(-0.21)*	*(-1.85)*	*(-0.95)*	*(-0.54)*	*(-2.36)*	*(-1.75)*	*(-1.80)*
M7a: M7 w/o Early Life Vars.	0.0%	-1.2%	-2.7%	-3.1%	-1.0%	-2.6%	4.2%	1.2%
	(-0.01)	*(-0.77)*	*(-1.78)*	*(-1.69)*	*(-0.86)*	*(-2.83)*	*(-2.83)*	*(-0.60)*
Cumulative Incidence Rate	2.4%	6.7%	7.5%	5.0%	18.1%	29.0%	9.4%	4.7%

Table 3: Correlates of Disease--Proportional Hazard Regressions

Variable	Stroke	Diabetes	Psych. Disorders	Lung Disease	Arthritis	Heart Disease	Hyper-tension	Cancer
Education:	+	-	-*	-	-**	+*	-	+*
Age:	+***	+***	-***	+**	+***	+	-	+***
Sex: Female	-	-***	+***	+	+***	-***	+	-
Race:White					reference			
Race:Black	+	+***	-***	-	-	-**	+***	-
Race:Hispanic	+	+***	-	-***	-*	-***	+	-
Race:Other	-	+**	+	+	-	-	+	+
Childhood Health: Excellent					reference			
Childhood Health:Very Good	-	-**	+**	-	+**	-	+	+
Childhood Health:Good	-	+	+	+	+**	+	+	-
Childhood Health:Fair	-	+	+***	+	+	+	-*	+
Childhood Health:Poor	+**	+	+**	+	-	+	+*	-
Childhood SES: High					reference			
Childhood SES: Average	-***	-	-	+	-	-	-	-*
Childhood SES: Poor	-**	-	-	+	+	-	-	+
Childhood SES: Varied	+	-	-	+	+***	-	-	+
Childhood transience:	-	-	+	-	+	-	+	-**
Childhood financial help:	-**	+	+	+	-	+	-	+
Father Employed					reference			
Father Unemployed	+	+	+*	-	+*	+	+	-
Father not at Home	-	-	+***	-	-	+	+**	-***
Father's Education:	-	+	+	-	-	+	-	+
Mother's Education:	-	-	-	+	+	-	-	-
Log of Household Wealth	-*	-	-*	-	-	-	-***	-
Health Insurance: 1=yes, 0=no	-	-	-	+	+	+	+	-*
Never Smoked					reference			
Curr. Smoking: Light	+	+	+**	+*	+**	+	+	+
Curr. Smoking: Moderate	+	+**	+	+***	+*	+*	+	+
Curr. Smoking: Heavy	+***	+**	+***	+***	+	+***	-	+***
Curr. Smoking: Very Heavy	+***	+*	+	+***	+	+**	+	+**
Prev. Smoking: Light	-	+	+*	+	+	-	-*	+
Prev. Smoking: Moderate	+	+	+***	+	+**	+	+	+
Prev. Smoking: Heavy	+**	+	+	+	+	+	+	+
Prev. Smoking: Very Heavy	+*	+	+	+**	+*	+	-	-

Notes: +/- signs indicate sign of coefficients: p-values are * <.1; ** <.05; ***<.01

Table 3: Cont.

Variable	Stroke	Diabetes	Psych. Disorders	Lung Disease	Arthritis	Heart Disease	Hyper-tension	Cancer
Non-Drinker					reference			
Daily Drinks: <1	-	-***	-	-	+*	-	+	+*
Daily Drinks: 1-2	-	-	-	+	+***	-	+	+
Daily Drinks: 3-4	-	-	-	-	-	-	+	+
Daily Drinks: 5+	+	-	+	-	-	-	+	+
No Exercise					reference			
Exercise: Light	+	-**	-	-	-	-	-	-
Exercise: Moderate	-	+	-	+	-**	-	+	+
Exercise: Heavy	-	-***	-	-	+	-	-	-
Exercise: Very Heavy	-	-	-	-***	+*	-	-**	+
Body Mass Index	-	+***	+	+	+***	+	+***	+
Body Mass Index Squared	+	-***	-	-	-**	-	-***	+
Marital Status: Married					reference			
Marital Status:Cohabiting	+	-	+	-*	+	+	-	+
Marital Status:Separated	-	+	+	-	-	-	-	+
Marital Status:Divorced	-	+	+	+	-**	-	-	-
Marital Status:Widowed	-	-	-	+	-**	+	+	-*
Marital Status:Never Married	-	+	-	-	+	-	-	+
Baseline Health: Excellent					reference			
Baseline Health:Very Good	-	+***	+	+	+***	+**	+	-
Baseline Health:Good	+	+***	+**	+***	+***	+***	+	+
Baseline Health:Fair	+***	+***	+***	+***	+***	+***	+	-
Baseline Health: Poor	+***	+***	+***	+***	+***	+***	+**	-
Stroke		+	+	-**	+	-	+*	-
Diabetes	+***		+	-	+	+***	+	-
Psychological Disease	-	-		+***	+*	+	+	+
Lung Disease	+	+	+		+	+***	+*	+
Arthritis	+	-	+**	+***		+***	+	+***
Heart Disease	+	-	+***	+**	+		+	-
Hypertension	+**	+***	-	+	+	+***		+
Cancer	-	-	+	+	-	+*	+	

Notes: +/- signs indicate sign of coefficients: p-values are * <.1; ** <.05; ***<.01

Table 4: Early-Life Correlates of Schooling

Dependent Variable: Years of Schooling

Variable	*Coeff.*	*t-stat.*
Age:	-0.018	-1.83
Sex: Female=1	-0.260	-4.20
Race:White	reference	
Race:Black	-0.488	-5.00
Race:Hispanic	-2.268	-12.52
Race:Other	0.704	2.59
Childhood Health: Excellent	reference	
Childhood Health:Very Good	-0.386	-5.36
Childhood Health:Good	-0.769	-8.93
Childhood Health:Fair	-0.470	-2.92
Childhood Health:Poor	-1.386	-4.92
Childhood SES: High	reference	
Childhood SES: Average	-0.637	-5.14
Childhood SES: Poor	-1.164	-8.18
Childhood SES: Varied	-0.394	-1.14
Childhood transience:	0.000	0.00
Childhood financial help:	0.120	1.11
Father Employed	reference	
Father Unemployed	0.037	0.43
Father Not at Home	-0.394	-3.18
Father's Education:	0.177	15.56
Mother's Education:	0.193	14.55
Intercept	11.340	18.92
N=	7676	
R-squared	0.300	

Table 5: Correlates of Health Behaviors

	BMI		**Smoking**		**Drinking**		**Exercise**	
	OLS		Ordered Probit		Ordered Probit		Ordered Probit	
Variable	*Coeff.*	*t-stat.*	*Coeff.*	*t-stat.*	*Coeff.*	*t-stat.*	*Coeff.*	*t-stat.*
Education:	-0.088	-3.64	-0.050	-7.62	0.042	6.97	0.048	8.53
Age:	-0.039	-2.04	-0.029	-5.39	-0.003	-0.74	-0.004	-1.03
Sex: Female=1	-0.564	-4.84	-0.144	-4.37	-0.438	-15.77	-0.134	-4.90
Race:White								
Race:Black	1.580	7.64	-0.318	-7.03	-0.046	-1.03	-0.014	-0.34
Race:Hispanic	0.360	1.34	-0.552	-7.53	-0.006	-0.10	0.086	1.51
Race:Other	-0.844	-1.79	-0.176	-1.66	-0.442	-3.81	-0.077	-0.75
Childhood Health: Excellent								
Childhood Health:Very Good	-0.102	-0.71	-0.046	-1.16	-0.004	-0.11	-0.089	-2.70
Childhood Health:Good	-0.157	-0.92	-0.094	-2.02	-0.085	-2.18	-0.045	-1.20
Childhood Health:Fair	-0.165	-0.56	-0.103	-1.36	-0.048	-0.72	0.020	0.31
Childhood Health:Poor	-0.079	-0.15	0.108	0.88	0.065	0.44	-0.353	-3.14
Childhood SES: High								
Childhood SES: Average	0.300	1.15	-0.039	-0.52	-0.134	-2.21	-0.092	-1.42
Childhood SES: Poor	0.275	0.96	-0.015	-0.18	-0.191	-2.81	-0.066	-0.95
Childhood SES: Varied	0.745	1.00	-0.069	-0.38	0.034	0.23	-0.096	-0.63
Childhood transience:	-0.238	-1.35	0.085	1.81	-0.010	-0.24	-0.001	-0.02
Childhood financial help:	-0.043	-0.22	0.031	0.56	0.076	1.64	0.045	1.01
Father Employed								
Father Unemployed	0.274	1.57	0.000	0.00	0.044	1.15	0.000	-0.01
Father Not at Home	0.031	0.13	0.081	1.41	0.132	2.54	-0.136	-2.80
Father's Education:	-0.074	-3.51	-0.002	-0.28	0.015	2.87	0.019	3.84
Mother's Education:	-0.012	-0.49	0.010	1.50	0.014	2.47	-0.003	-0.53
Log of Household Wealth	-0.076	-3.00	-0.064	-11.92	0.024	3.86	0.040	7.92
Health Insurance	0.018	0.09	-0.245	-5.03	-0.116	-2.35	0.050	1.07
Intercept	31.769	26.68						
N=	7598		7598		7598		7598	
R-squared	0.033		0.032		0.038		0.024	