Myocardial Infarction, Health Behavior, and the Grossman Model

Emma Marie Mehlhop

Professor Frank A. Sloan PhD., Faculty Advisor
Professor Charles M. Becker PhD., Faculty Advisor
Professor M. Kate Bundorf PhD., Faculty Advisor
Professor Grace Kim PhD., Seminar Instructor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University
Duke University
Durham, North Carolina
2021

Acknowledgements

A sincere thank you to my wonderful advisors, classmates, friends and family for your continual support throughout this process. This thesis would not be possible without the hours upon hours of Zoom guidance and encouragement from Dr. Frank Sloan, Dr. Charlie Becker, and Dr. Kate Bundorf. Thank you to Dr. Grace Kim for offering your expertise as I navigated this process, and for the collaboration and feedback from my honors seminar classmates. I would also like to thank the SSRI, Alan Jaske, and the Masters Student TAs for all of your help as I completed this project. This could not have happened without your continued support!

Abstract

This paper contributes an empirical test of Michael Grossman's model of the demand for health and a novel application of the model to myocardial infarction (MI) incidence. Using data from the University of Michigan's Health and Retirement Study (HRS), I test Grossman's assumptions regarding the effects of hourly wage, sex, educational attainment, and age on health demand along with the effects of new variables describing health behaviors, whether or not a respondent is insured, and whether or not they are allowed sufficient paid sick leave. I use logistic regression to estimate health demand schedules using five different health demand indicators: exercise, doctor visits, drinking, smoking, and high BMI. I apply the Cox Proportional Hazard model to examine two equations for the marginal product of health investment both in terms of propensity to prevent death and to prevent MI, one of the leading causes of mortality in the United States. This study considers the effects of the aforementioned health demand indicators, among other factors, on the marginal product of health investment for the prevention of death compared to the prevention of MI. Additionally, there is significant evidence of a negative effect of health insurance on likelihood of exercising regularly, implying some effect of moral hazard on the health demand schedule.

JEL Codes: I10, I12, I19
Keywords: Health economics, Myocardial infarction, Health behavior

I. Introduction

The present analysis applies the Grossman model, a prominent model of health investment and consumption in the field of health economics, to the probability of survival and the probability of the incidence of heart attack. The dual purpose of this thesis is to test the assumptions of the Grossman model generally and to determine whether or not it can be used as a valid predictor of a specific health condition as opposed to general health, as it has been used historically.

In 2019, almost 50% of US residents were affected by Cardiovascular disease (CVD) (AHA, 2019). Myocardial infarction (MI), colloquially known as heart attack, is one of the most common manifestations of CVD. MI and other complications of CVD have been the leading causes of mortality in the US since 1900 (Hamil-Luker and O'Rand, 2007). The enormous burden of cardiovascular disease caused by MI alone warrants further research into the risks, treatments, and causes of MI specifically. Further, the substantial portion of US morbidity and mortality that can be attributed to MI indicates that MI may serve as a meaningful proxy of general population health.

One of the seminal models in health economics is the Grossman model, which was proposed by Michael Grossman in his 1972 dissertation *The Demand for Health: A Theoretical and Empirical Investigation.* This model can be used to better understand how our decisions to "invest in", "consume", or "demand" health can change over the course of the lifetime and can differ across individuals. In his model, Grossman asserts that people can effectively determine their own lifespan by choosing the degree to which they "demand" health through either medical care or other health-promoting measures. Grossman proposes that this demand for health is influenced by factors such as age, wage, and educational attainment. While the original Grossman model does not consider health behaviors explicitly, several past empirical reformulations of the Grossman model have included such factors. In order to apply the Grossman model to MI, it will be useful to use similar methods by assessing the effects of health behaviors alongside factors such as age, wage, and educational attainment.

MI risk factors such as high blood pressure, diabetes, and smoking are considered avoidable or "modifiable" as a healthy lifestyle and careful diet can significantly reduce one's risk of MI. My empirical analysis will apply Grossman's theory to the study of MI risk and prevention. To this end, this thesis also contributes the inclusion of a Cox proportional hazard model in order to identify the effects of a series of health investment indicators on an individual's risk of both death and MI risk. Additionally, to test the validity of the model more generally, I will reformulate Grossman's health demand equations using a logistic regression model with binary indicators of various types of health demand. In this thesis, I will use behavioral, demographic, and health information from the Health and Retirement Study (HRS), a large longitudinal panel study of American adults conducted from 1992 through 2018, to carry out my analyses.

The primary risk factors of MI are hypertension (high blood pressure), high cholesterol and dyslipidemia (abnormal lipid levels in the blood), diabetes, obesity and inactivity, smoking, and family history of heart disease (NHS, 2018). A majority of patients presenting with their first MI in a 2011 study by Canto et. al exhibited at least one of the aforementioned risk factors of CVD. Among patients presenting with their first MI, 52.3% had hypertension and 31.3% had some history of smoking. It was rare, however, for patients to present with obesity as a lone risk factor. Additionally, Canto et al found statistically significant evidence that patients who were obese were likely to have more co-risk factors than those presenting without obesity. While this thesis will not focus on patient mortality post-MI, as cause of death is masked in the data, it is useful to note that after controlling for age and other clinical risk factors, as the primary risk factors for CVD increased, the number of in-hospital deaths following a patient's first MI decreased. The authors postulate that the aforementioned inverse relationship between number of risk factors and mortality may be a result of better health management or a higher propensity to consult health care providers among those who are conscious of the fact that they have a high number of MI risk factors (Canto et al., 2011).

There is presently a wealth of literature on how we can prevent MI which cites health behaviors that also show up in empirical analyses of the Grossman model. As of now, however, I am unaware of any academic papers that have applied the Grossman model to CVD or MI specifically. Rather than focusing on a particular disease, past empirical applications of the Grossman model, which will be discussed below, have been used to predict overall health. Due to the prevalence of MI in the US, however, modeling MI risk may be a viable proxy for overall health. Additionally, there is a vast body of evidence that behaviors such as exercising or smoking can have significant and observable impacts on our overall health and risk of MI. This case study in MI and health behavior is an ideal candidate for the exploration of the degree to which we have real agency in determining our own health through an empirical application of the Grossman model.

II. Literature Review

The Grossman model defines health as a durable capital good which depreciates over time. In this model, people produce health through a health production function in which the input is some type of health investment. Health, as a capital good, then "produces" both consumption and investment benefits. Consumption benefits describe "direct increases to utility," such as when some ailment is resolved and one feels generally healthier. Investment benefits describe the "increased healthy time" one will enjoy over the course of their life, allowing them to work and earn wages, as a result of investment in their own health stock

(Muurinen, 1982). The depreciation rate of health capital (δ_i) can be defined exogenously as a product of an individual's age or endogenously as a function of both age and the "intensity of use" of one's health. Intensity of use describes the degree to which a person maintains their health stock by either practicing health-improving or health-diminishing behaviors. The effect of the gradual increase in depreciation rate over the course of a lifetime can be mitigated by "investing" in or "consuming" health (Grossman, 1972). When the depreciation rate changes, the marginal efficiency of health capital (MEC) in the health production function changes accordingly.

The concept of intensity of use is of particular interest regarding MI, as a person at risk of MI may be using their health stock more intensely if they are overweight, have high blood pressure, and are not actively seeking opportunities to invest in health. Health behavior and lifestyle choices are thus incorporated into the model through this concept of "intensity of use" or "use-related deterioration." For the purposes of my analysis, therefore, depreciation rate is defined, at least in part, endogenously through the incorporation of these health behaviors. The depreciation rate itself will not be estimated in this paper, rather it will be implicitly represented in the model through changes in health stock due to changes in behavior.

We can define current health stock (H_i) as an amalgamation of all aspects of health which an individual has thus far attained. In other words, H_i is made up of one's heart health, lung health, mental health, digestive health, and so on. Each of these aspects of H_i have a certain "age" which refers to how long it has been since the individual invested in that aspect of their health. By combining H_i and δ_i, we can measure how quickly a person will lose their health. For instance, consider a young person who trains for a marathon, thereby improving their heart health, but when they finish their race they immediately begin eating fast food each day. The age of their heart health investment as a component of H_i will be low, but their intensity of use will be high. Since they are young, they may be able to get away with this. However, if they continue with their unhealthy behavior, the age of H_i will increase along with their own age. If they maintain their high intensity of use without investing further in heart health, this aspect of their health stock will diminish increasingly rapidly. If they instead choose to eat healthier and increase their heart health periodically through exercise or healthy diet, they will of course benefit from a lower intensity of use (Muurinen, 1982).

In Grossman's 1972 manuscript, he specifies three major findings. First, the depreciation rate of health stock increases and the demand for health stock is decreases as we age. Second, the demand for medical care increases as hourly wages increase. Third, education increases the efficiency with which we produce health, allowing us to optimize at higher levels of health stock. Following his original analysis, many other health economists developed their own empirical reformulations of the Grossman model. Many past empirical reformulations of the Grossman model generally support the prediction that health is increasing in income/wages as well as education. Health is decreasing, however, in age, the cost of medical care, and stressful work environments or psychological stressors. These past empirical studies have also found that unmarried/single people and females tend to have worse health than married people and males. Further, health is increasing in exercise, healthy eating, and good sleeping habits, and is lower for those who are overweight or smoke cigarettes (Galama, 2012). Some of the most notable empirical reformulations were those conducted by Maureen Cropper in 1981 and by Adam Wagstaff in 1986.

Past empirical findings are influenced, of course, by how health is measured in each study. Grossman uses a self-reported assessment of health measured on a scale of 1 to 5 (1 being poor and 5 being excellent) to estimate health stock (*H*). Health stock itself, however, is intangible and inherently difficult to measure. For this reason, Grossman also considers the production of “healthy time” such that health stock serves as a latent variable within the model. “Healthy time” is defined as the proportion of days in which an individual did *not* experience restricted activity or inability to work due to illness or injury. Grossman uses two-stage least squares to measure health using a system of equations in which health stock (*H*), work loss days (*WLD*), and restricted activity days (*RAD*) serve as dependent variables. He also uses OLS to predict the demand for health using medical care (*M*) as the dependent variable (Grossman, 1972). In his 1986 empirical application of the Grossman model, Wagstaff also treats health capital as a latent variable in a system of three demand equations which predict the quantity of demand for general practitioner visits, hospital stays, and medicine use. Workplace characteristics which describe the environment of a respondent’s primary workplace as well as what is demanded of them on a daily basis is also included in the estimation of health capital. Wagstaff incorporates use-related depreciation of health stock in his equations by controlling for sex, family size, and a series of variables describing the respondent’s employment information (Wagstaff, 1986).

In Cropper’s empirical work, she estimates the monetary value of a decrease in air pollution by using the Grossman framework to estimate the effect of air pollution on health seeking behaviors including exercising, smoking, and sleeping habits. Like Grossman, Cropper measures health stock using each respondent's rating of their own health on a scale of 1 to 5 (1 being poor and 5 being excellent). In her estimation of the demand for health schedule, Cropper includes covariates for level of air pollution as well as a respondent’s race and parents’ income. Cropper also controls for chronic health conditions, education level, marital status, wage, and a measure of risk aversion (Cropper, 1981).

All of these methods are imperfect to some extent in terms of their ability to accurately represent one’s health stock. Wagstaff comments that a weakness of the latent variable method is that we cannot differentiate between patient-initiated and provider-initiated visits if this information is not explicitly provided in the data. While both patient and provider-initiated visits imply some level of demand for health, the extent to which a patient is actually demanding health may not be as observable when considering provider-initiated visits. A useful set of variables that are included in Wagstaff's analysis are the availability of general practitioners (GPs) or hospital beds near a given respondent's residence. This improves the accuracy of a respondent's predicted demand for health as their actual demand for health can be greatly influenced by their ability to pursue health given their location and resources (Wagstaff, 1986). Measuring health stock directly also has its flaws, as it relies on one's subjective view of their own health status. It appears as if some combination of the two methods is standard practice, as one method may fail to capture the effect of a given covariate on health stock or demand.

Now, we can consider how health stock, whether measured directly or as a latent variable, changes over the course of a lifetime. Grossman found a negative coefficient for age in the health capital stock (*H*) equation but a positive value in the demand for medical care equation. This implies that as people age, they have less health stock and seek more medical care. This supports Grossman's theory that the depreciation rate rises with age and the MEC curve is inelastic. This causes health capital to fall as one ages at a "continuously compounded rate".

Regarding wages, the Grossman model predicts that an increase in hourly wage rate will lead to an increase in the monetary return on health investment. In other words, while the opportunity cost of seeking medical care is now higher, so is the cost of inability to work due to illness. A higher wage rate is expected to increase one's health stock, number of healthy days, and demand for medical care. However, the wage elasticity of medical care is negative in Grossman's empirical specification, but not significant (Grossman, 1972). Cropper's results also show that the effect of wage on health stock is either insignificant or positively related to illness. Cropper suggests, as Grossman posits in his own analysis, that this unexpected result may be caused by an increase in glutinous consumption that follows an increase in wealth (Cropper, 1981).

Finally, regarding education, Grossman's education efficiency hypothesis states that a higher level of education causes one to be a more efficient producer of health. According to Grossman's theoretical specifications, therefore, an increase in education should increase the marginal product of the inputs to the health production function and shift the MEC curve to the right. Better educated people should also experience a lower marginal cost when investing in health. Grossman predicts that the demand for health will therefore be increasing in education. While this implies that education should be negatively correlated with medical expenditures, Grossman found a positive but insignificant coefficient for education in his estimated health investment equation (Grossman, 1972). In Wagstaff's empirical work, when controlling for covariates including health, age, and sex, he found that the better educated demand more health care overall. This may be due to the fact that better educated people see lower marginal returns to health so they invest in more. Alternatively, better educated people may have better relationships with their physicians and will therefore be incentivized to visit their doctor more frequently (Wagstaff, 1986).

One factor which is not explicitly included in Grossman's original specification is the availability of paid sick leave. This is relevant as the true price of health investment is made up of the actual monetary cost of a given health investment opportunity and the estimated price of the time that it takes the individual to seek that health investment. The time cost of seeking healthcare is therefore vastly reduced when the consumer does not have to forego wages. Cropper's 1981 study uses paid sick leave for this purpose, yet many people in her sample reported zero sick days per year. This is a potential downside of defining health capital in terms of healthy days, and Cropper accounts for this by constructing a variable describing the natural log of the number of sick days such that the number of sick days is continuous and can become arbitrarily small without ever actually approaching zero (Cropper, 1981).

As discussed above, empirical results often contradict Grossman's expectations. Wage has been shown to have an inverse relationship with health and education has been found to be an insignificant predictor of health demand. Wagstaff (1986) suggests that the presence of some of these "incorrect signs" in the health demand system of equations may be a product the fact that health demand is not just demand for formal medical care (in his case, through GP visits, hospital stays, and medication use). Health demand also consists of the demand for a healthy lifestyle through exercise and the avoidance of health-damaging behaviors such as smoking.

As stated above, exercise has been shown to increase health in the Grossman literature. In response to Wagstaff's comment, however, exercise may be a source of demand for health in and of itself. Higher levels of exercise or physical activity (PA) can result in a better overall health as well as a decreased risk of CVD. Until recently, however, there has been uncertainty regarding this relationship between PA and CVD at the highest levels of PA. A recent study used motion sensor data to assess the effect of moderate-intensity and vigorous-intensity physical exercise. The authors found that the inverse association between PA and CVD exists even at the highest levels of PA. In other words, the benefits of PA do not plateau at a certain intensity/frequency level, as the highest benefits of PA were seen at the highest activity levels. Further, the highest PA levels were associated with the lowest CVD risk in the study cohort (the UK Biobank) (Ramakrishnan et al, 2021). This indicates that exercise may be a significant measure of health investment both generally and for MI specifically in my empirical analysis.

If we can measure health stock as a latent variable using health-improving behaviors such as exercise, we may also be able to approximate health stock using health-diminishing behaviors such as smoking, drinking, and inactivity. According to Galama (2012), moderate alcohol consumption shows either a positive or insignificant effect on health. While this seems counterintuitive, this may simply be an endogeneity issue in determining the direct effect of drinking on health. For example, if a person becomes ill and is not able to drink as frequently, they will show a positive association between drinking and health though this is not necessarily causal behavior. For this reason, I expect insignificant or counterintuitive results regarding the effect of alcohol consumption.

Smoking kills over 8 million people a year and is one of the world's biggest public health problems. The Mayo Clinic reported that there is a causal relationship between smoking and CHD. Though the prevalence of cigarette smoking has declined in the US from approximately 40% to 13.7% in 2018 in the past 50 years due to a variety of public health initiatives, it remains a relevant concern regarding both general health and MI risk. Smoking cessation can be interpreted as a health-seeking behavior as, per the1990 US surgeon general report, "smoking cessation improves immediate and long-term health and increases longevity, even for those who already suffer from smoking-related illness." Further, smoking cessation can reduce the risk of MI and CVD-related mortality by over 50% among those with diagnosed heart problems (Lahiri and Li, 2020).

Finally, obesity is considered to be a major risk factor of MI, and it is expected that high BMI (which can also serve as a proxy for inactivity) will be associated in an increase of MI prevalence and lower survival rates. However, it is worth nothing that a 2011 study by Canto found that among those with diagnosed CHD, overweight/obese people show relatively lower long-term mortality rates. The study defined overweight/obesity using the following BMI categories: underweight: BMI < 18.5, normal: 18.5 < BMI < 25, overweight: 25 < BMI < 30, obese: 30 < BMI < 40; and morbidly obese: BMI > 40. Canto states that obesity has a strong and positive relationship with the presence of hypertension, dyslipidemia, and diabetes (three major risk factors of CVD). Interestingly, however, patients with preexisting CHD and a BMI > 30 showed lower mortality compared to those with BMI <30. This seemingly paradoxical relationship may result in a "protective effect" of these risk factors at first incidence of MI (Canto, 2011). This may explain any unusual findings regarding BMI in empirical analysis.

III. Theoretical Framework

In order to contextualize the Grossman model and the predictions of the aforementioned empirical formulations, it is necessary to develop an understanding of the theoretical framework presented by Grossman (1972) in his original manuscript. The Grossman model utilizes a lifetime budget constraint in which the "consumer" is a utility maximizer over a finite number of periods in their lifetime (*n*). As in classical economic theory, agents must optimize their utility by splitting their disposable income between investing in health capital (H_i) and all other consumption goods (represented by the composite consumption good Z_i). An individual's lifetime utility ($U_{lifetime}$) can be represented by the intertemporal utility function:

$$U_{lifetime} = U(\varphi_0 H_0, \dots, \varphi_i H_i, \dots, \varphi_n H_n, Z_0, \dots, Z_i, \dots, Z_n) \quad \textbf{[1]}$$

where H_0 represents the health stock that one is born with and H_i represents health stock in a given period (*i*). For each period, the value of each unit of health is represented by φ_i, the "service flow" or utility gained by each unit of H_i. An individual's utility (U_i) in a given period can be represented by:

$$U_i = U(H_i, \varphi_i, Z_i) \quad \textbf{[2]}$$

Finally, Z_i, the composite consumption good, can be represented in more detail with:

$$Z_i = Z(X_i, T_i; E_i) \qquad \textbf{[3]}$$

where X_i represents the goods (such as money or time) that one exchanges for some quantity of Z_i, T_i represents the time one gives up in order to attain a given amount of Z_i, and E_i represents human capital stock.

In a given period, an individual has a certain amount of monetary stock and human capital stock (E_i) which is made up of health stock and "skills and knowledge" stock. A change in human capital can affect one's efficiency of health production, thereby affecting the estimated monetary value of a "unit" of health, or the "shadow price" of health. In other words, gaining the same amount of health becomes costlier for those who have lower health or lower "skills and knowledge" to begin with. Rational choosers will use their human capital resources to invest in health and the composite consumption good (Z_i) in a proportion which maximizes their utility given their preferences, disposable income, and initial health stock.

In the Grossman model, the production of health occurs through investment in health stock which can be achieved through medical care or time spent doing any activity which improves one's health. For instance, when an individual makes a tradeoff between health and other goods such as going to the doctor or exercising, they are investing in their health. Health investment for a given period (I_i) can be represented as

$$I_i = I(M_i, TH_i; E_i) \qquad \textbf{[4]}$$

where investment is a function of medical care (M_i) and time spent seeking health (TH_i) given one's human capital stock (E_i). The marginal products of TH_i and M_i can be written as $\partial I_i/\partial TH_i$ and $\partial I_i/\partial M_i$ respectively. We can then use these to isolate the marginal product of time. According to Grossman, the health and consumption good production functions are both homogenous of degree 1, meaning that for an n-level increase in income, the demand for both TH_i and M_i will increase by a factor of n, implying parallel indifference curves. We can therefore simplify these marginal products by dividing all elements of I_i by M_i and expressing the gross investment production function as such:

$$I_i = M_i g(t_i; E_i) \qquad \textbf{[5]}$$

where time is represented by $t_i = TH_i/M_i$ and g is some function of time for which efficiency is determined, in part, by E_i. Therefore,

$$\partial I_i/\partial TH_i = \partial g_i/\partial t_i = g' \qquad \textbf{[6]}$$

$$\partial I_i/\partial M_i = g - t_i g' \qquad \textbf{[7]}$$

giving us the marginal product of time (g') and the marginal product of medical care ($g - t_i g'$) in the gross investment function. As noted previously, when an individual invests in their health, the amount of health stock that they gain is dependent on some health capital production function. This function is unique to the individual and can change over the course of their lifetime. We can refer to the MEC and its behavior when identifying a change in the health production function.

When an individual produces some level of health, they determine their health stock for the upcoming period (H_{i+1}). The net investment in the stock of health can be represented by

$$H_{i+1} - H_i = I_i - \delta_i H_i \qquad \textbf{[8]}$$

where an individual's net investment in health stock ($H_{i+1} - H_i$) is defined by their gross health investment (I_i) less the health stock lost to depreciation ($\delta_i H_i$).

Now, we can define the individual's budget constraint such that the total present value of all medical care and other consumption goods consumed over one's lifetime is equal to the total present value of lifetime earnings and initial assets (A_0). Therefore, the budget constraint given some interest rate (*r)* is

$$\sum \frac{P_i M_i + V_i X_i}{(1+r)^i} = \sum \frac{W_i TW_i}{(1+r)^i} + A_0 \quad \textbf{[9]}$$

where P_i is the price of M_i, V_i is the price of X_i, TW_i is time spent working, and W_i is the person's hourly wage. The time constraint (Ω) must account for all time spent working (TW_i), time spent investing in health (TH_i), time lost/spent unhealthy (TL_i), as well as all other possible uses of an individual's disposable time (T_i). We can assume, intuitively, that TH_i is directly related to both I_i and H_{i+1} and inversely related to TL_{i+1}. Similar to the concept of disposable income, we can choose how to allocate a portion of our time (Ω) according to our preferences, but time spent doing necessary tasks such as sleeping and eating is not taken into account in equation 10 (shown below). Therefore:

$$TW_i + TL_i + TH_i + T_i = \Omega \quad \textbf{[10]}$$

Now, we can define the equilibrium conditions which are as follows:

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \left[\frac{W_i G_i}{(1+r)^i} + \cdots + \frac{(1-\delta_i)\ldots(1-\delta_{n-1}) W_n G_n}{(1+r)^n}\right] + \left[\frac{Uh_i G_i}{\lambda} + \cdots + \frac{(1-\delta_i)\ldots(1-\delta_{n-1}) Uh_n G_n}{\lambda}\right] \quad \textbf{[11]}$$

$$\pi_{i-1} = \frac{P_{i-1}}{g - t_i g'} = \frac{W_{i-1}}{g'} \quad \textbf{[12]}$$

where we define the marginal product of health stock in producing healthy days as $G_i = \frac{\partial h_i}{\partial H_i}$ and the marginal utility of healthy days as $Uh_i = \frac{\partial U}{\partial h_i}$. $G_i W_i$ then describes the marginal benefit of health investment in terms of earning wages and $G_i Uh_i$ describes the marginal benefit of health investment in terms of producing health itself. Equation 11 therefore sets the present value of the marginal cost of gross health investment in the previous period (π_{i-1}) equal to the present value of the marginal benefits of health investment over the course of a lifetime. Equation 12 describes the minimum investment needed to produce the optimal benefits of health as described in equation 11. The marginal cost of gross investment in the previous period (π_{i-1}) must equal the benefits of medical care and labor. The benefit of M_i is defined by the price of M_i in the previous period (P_{i-1}) divided by the marginal product of M_i ($g - t_i g'$) and the marginal benefit of labor is a person's wage in the previous period (W_{i-1}) divided by the marginal product of time (g'). Similarly, the present value of the marginal benefits of health can be represented by

$$G_i \left[\frac{W_i}{(1+r)^i} + \frac{Uh_i}{\lambda} \right] \qquad \textbf{[13]}$$

where G_i is the marginal product of health capital, Uh_i is the marginal utility of healthy days ($h_i = \Omega - TL_i$), and λ is the marginal utility of wealth. This formulation of the benefits of health allows us to contextualize the aforementioned distinction between the investment and consumption benefits of health. In order to simplify the model and provide opportunity for comparisons with other capital goods, Grossman splits his theoretical framework into a pure investment and a pure consumption model. In the pure investment model, we consider only the monetary returns of health stock and ignore the utility of holding health stock. In the pure consumption model, we consider only the utility of holding health stock and ignore the monetary returns to health. Note that the term Uh_i/λ in equation 12 describes the marginal utility of a one-unit increase in healthy time. In the pure consumption model, this term can represent the utility of healthy days and the disutility of unhealthy days. The pure investment model, however, sets $Uh_i/\lambda = 0$ in order to isolate the investment benefits of health without considering the utility of health (Grossman, 1972). The majority of past empirical work, including Grossman's own formulations, has focused on the pure investment model. The present analysis will also focus on the pure investment model, as it will account for the monetary benefits of health and how they are affected by age, hourly wage, and human capital (which is represented by education). I will also briefly consider the pure consumption model in the consideration of the effect of household income on health demand. Through the inclusion of Uh_i/λ, the pure consumption model reflects the diminishing marginal utility of *Z*. Within this framework, higher income individuals may substitute away from *Z* and toward *H* due to the diminishing returns to investing in *Z.* Put plainly, a person only needs so many luxury cars or vacations, so the utility they gain from non-health consumption eventually levels off causing an increase in demand for health.

Grossman (1972) transforms equation 13 algebraically into the following:

$$G_i \left[W_i + \frac{Uh_i}{\lambda}(1+r)^i \right] = \pi_{i-1}(r - \tilde{\pi}_{i-1} + \delta_i) \qquad \textbf{[14]}$$

where π_{i-1} represents the marginal cost of gross health investment in the previous period and $\tilde{\pi}_{i-1}$ represents the percentage change in π_i between the previous and current periods. The expression $r - \tilde{\pi}_{i-1} + \delta_i$ is effectively the rental rate of health capital which is made up of the interest and depreciation rate as well as any capital gains derived from past health investment ($\tilde{\pi}_{i-1}$). This provides an optimality condition which asserts that the undiscounted marginal product of an individual's health stock must be equal to the cost of holding one unit of health capital. Continuing with the pure investment model, we can ignore the consumption benefits of health by setting $Uh_i/\lambda = 0$ which results in

$$\frac{G_i W_i}{\pi_{i-1}} = r - \tilde{\pi}_{i-1} + \delta_i = \gamma_i \qquad \textbf{[15]}$$

which implies that the monetary value of the marginal rate of return on an investment in health ($G_i W_i/\pi_{i-1}$) must be equal to the cost of holding one unit of health capital. This expression can be described as the marginal efficiency of health capital (MEC) which will be denoted by γ_i.

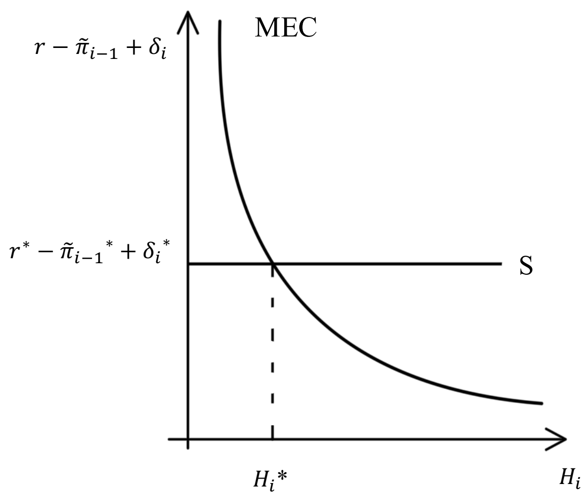


Figure 1
The Marginal Efficiency of H_i

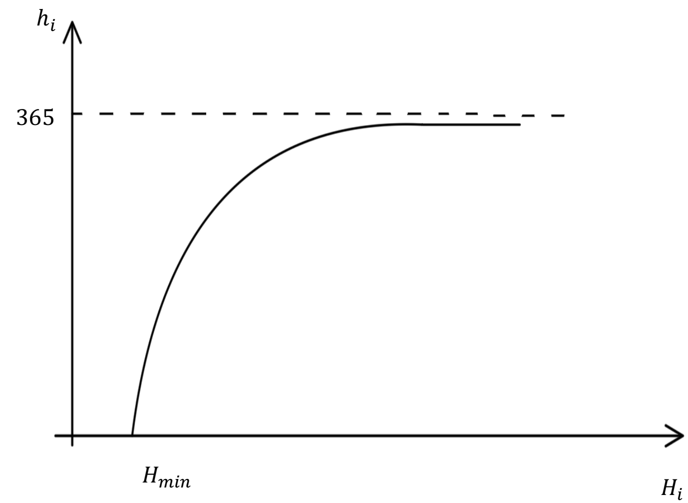


Figure 2
The diminishing marginal returns to H_i in producing h_i

In figure 1, the demand for health stock is represented by the MEC schedule where health stock is shown on the horizontal axis and the marginal return on health investment (γ_i) is shown on the vertical axis. The MEC is downward sloping as the marginal product of health capital (G_i) is diminishing. The infinitely elastic supply curve (S), on the other hand, is plotted such that health stock is shown on the horizontal axis and the rental rate of health stock is shown on the vertical axis. The supply curve is perfectly elastic, as Grossman assumes that this rental rate does not depend on the amount of health stock owned or the current period. The optimal health stock (H_i*) that a person will hold in a given period (i) must therefore occur when $\gamma_i = r - \tilde{\pi}_{i-1} + \delta_i$. Since the marginal cost of health investment does not change, we can conclude that $\gamma_i = r + \delta_i$.

Figure 2 provides further illustration of the diminishing marginal product of health. If we assume, as stated in the pure investment model, that the only benefit to health is increased healthy days (h_i), these benefits must be bounded by the period length which, in this case, we will assume is one year.

As we age, it becomes increasingly difficult to maintain our health and to generate new health stock. Past a certain point, age must be positively associated both with δ_i and the marginal cost of gross health investment (the shadow price of health). As age increases, so does $r + \delta_i$ which decreases one's optimal health stock (H_i*). Depreciation must increase with age otherwise a single value of H_i* would always satisfy the $\gamma_i = r - \tilde{\pi}_{i-1} + \delta_i$ condition. Since we can assume that people do not maintain one level of health stock over their lifetime, especially in the later years of life, the depreciation rate must be increasing in age. In order to describe this inverse relationship between level of health and age, we can take the derivative of $\gamma_i = r + \delta_i$ with respect to i. The percentage decrease in health stock in a given period ($\widetilde{H}_i$) is then

$$\widetilde{H}_i = -s_i \varepsilon_i \widetilde{\delta_\iota} \quad \textbf{[16]}$$

where the ratio of the depreciation rate to the total cost of health capital is given by $s_i = \frac{\delta_i}{r+\delta_i}$ and the elasticity of the MEC is given by ε_i. In the definition of ε_i, note that for a higher level of ε_i we expect a lower change in γ_i from a one unit of change in the health stock. For the purposes of this analysis, however, assume that ε_i is constant.

$$\varepsilon_i = \frac{-\partial \ln (H_i)}{\partial \ln (r+\delta_i)} = \frac{-\partial \ln (H_i)}{\partial \ln (\gamma_i)} = \frac{-\partial \ln (H_i)}{\partial \ln (G_i)} \quad \textbf{[17]}$$

As people age, they are likely to invest in relatively more health but receive relatively less healthy days implying a positive association between TL_i and both M_i and TH_i ceteris paribus. In order to prove this, given the relationship in equation 8, we can state that

$$\ln (I_i) = \ln (H_i) + \ln (\widetilde{H_\iota} + \delta_i) \qquad \textbf{[18]}$$

After differentiating [18] with respect to age (i), and assuming that $\tilde{\delta}_i$ and ε_i are constant, Grossman finds that the gross investment of health over a lifetime ($\widetilde{I_\iota}$) can be represented by

$$\widetilde{I_\iota} = \frac{\widetilde{H_\iota}^2 + \delta_i \widetilde{H_\iota} + \frac{\partial \widetilde{H_\iota}}{\partial i} + \delta_i \widetilde{\delta}}{\widetilde{H_\iota} + \delta_i} \quad \textbf{[19]}$$

Given [16], we find that

$$\frac{\partial \widetilde{H_\iota}}{\partial i} = -s_i(1 - s_i)\varepsilon\, \delta^2 \qquad \textbf{[20]}$$

and since $\frac{\partial \widetilde{H_\iota}}{\partial i} < 0$ by construction, we can rewrite [19] as

$$\widetilde{I_\iota} = \frac{\widetilde{\delta}(1-s_i\varepsilon)(\delta_i - s_i\varepsilon\widetilde{\delta}) + s_i^2\varepsilon\widetilde{\delta}^2}{\delta_i - s_i\varepsilon\widetilde{\delta}}. \qquad \textbf{[21]}$$

Given that $\tilde{\delta}_i$ and ε_i are constant, gross health investment over the lifetime ($\widetilde{I_\iota}$) must be nonnegative, which is logical as one cannot sell health stock. If we assume that $\widetilde{I_\iota} \neq 0$, we know that the MEC must be inelastic (ε_i<1) over the lifetime in order to satisfy the condition that $\widetilde{I_\iota}$ is positive. This inelasticity causes health capital to fall as one ages at a "continuously compounded rate". This implies a positive association between gross health investment and δ_i as well as a negative association between gross health investment and H_i over the course of a lifetime. We can interpret the former by concluding that, in order to maintain one's desired level of health capital, individuals must marginally increase their health investment over the course of the lifetime in order to account for the latter (the rising depreciation rate). Age is therefore associated with an increase in demand for medical care and own time health inputs but a decrease in health capital. Maintaining a desired level of health becomes more and more difficult such that people will gradually invest less and less in health until they "choose" death by optimizing utility at zero healthy days ($h_i = \Omega - TL_i = 0$) where $H_i{*} = H_{min}$ as shown in figure 2.

Now that we have established the association between age and depreciation rate, we can consider the variables that affect the MEC: wage and human capital. Beginning with hourly wage, we can assume that a higher wage implies higher benefits associated with healthy time (h_i) since TL_i implies an inability to work. While this incentivizes an increase in demand for health, it is important to consider the effect of the opportunity cost of time. The Grossman model defines health investment input as a combination of time (TH) and medical care (M). If TH accounts for x% of health inputs, a 1% increase in wage would cause an x% increase in the marginal cost of health investment (π).

The value of π therefore reflects an individual's hourly wage. An increase in wage, for this reason, is associated with decrease health investment demand due to a higher marginal cost. However, wage also increases the returns to health investment. Therefore, if these returns outweigh the costs, the increase in the MEC will cause a net increase in wage by (1-*x*) % given that x ≠ 100 as shown below in figure 3. If the returns to investment from an increase in wage do not outweigh the added cost, the effect of wage on health demand will be negative. Therefore, an increase in wage *may* cause a net increase in demand for medical care, but not in all cases.

Regarding human capital, Grossman asserts that we can assess the effect of human capital (E_i) on the MEC by measuring changes in educational attainment. The non-monetary benefits to education can be shown by their contributions to human capital (skills and knowledge stocks).

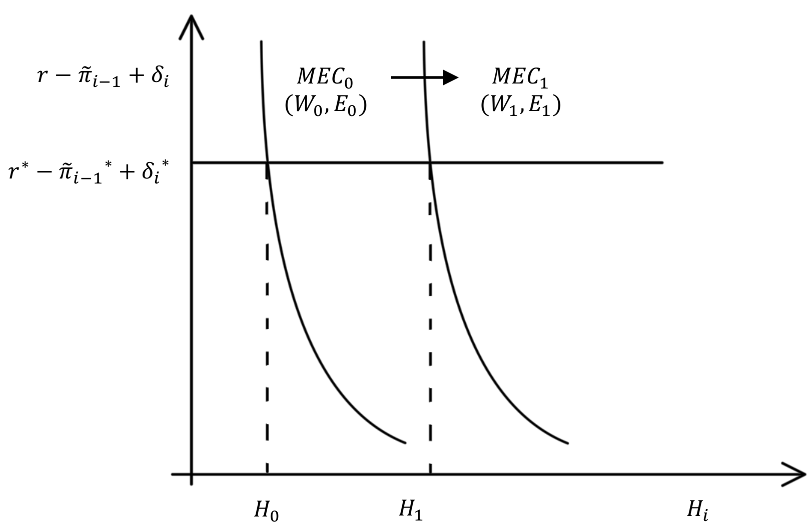


Figure 3
The effect of an increase in W_i or E_i on the MEC

Given equations 4-7, the marginal product of E_i can be represented by the following:

$$\frac{\partial I}{\partial E} = M\frac{\partial(g - tg')}{\partial E} + TH\frac{\partial g'}{\partial E}. \quad \textbf{[22]}$$

This implies that the weighted average of the percentage changes in the marginal products of *M* and *TH* is increasing in education. With all else held constant, an increase in marginal product implies less *M* and *TH* needed to produce a unit of health stock. It follows, therefore, that an increase in *E* causes a decrease in the marginal cost of health and an increase in the MEC (which will shift to the right as shown in figure 3). Given [22], and assuming that an increase in *E* causes a same percentage increase in the marginal product of time (g') and the marginal product of medical care ($g - t_i g'$), we can prove this by showing that

$$r_H = \frac{\partial I}{\partial E}\frac{1}{I} = \left(\frac{M(g-tg')}{I}\right)\left(\frac{g\hat{g}-tg'\widehat{g'}}{g-tg'}\right) + \left(\frac{THg'}{I}\right)(\hat{g}) = \hat{g} = \widehat{g'} \qquad \textbf{[23]}$$

where $\hat{g}$ represents the percentage increase in g and r_H represents the percentage increase in gross investment each caused by a 1% increase in *E*. People will therefore be able to optimize at a higher level of demand for H_i. The percentage increase in demand for health that is caused by this increase in education ($\widehat{H}$) can be represented by $\widehat{H} = r_H\varepsilon$. Since we have established that the MEC is inelastic, we know that $\widehat{H} \neq r_H$. Using [22], [23], and some algebra, $\widehat{M} = T\widehat{H} = r_H(\varepsilon - 1)$. This tells us that an increase in education is associated with a relatively lower demand for medical care likely due to an increased incentive to substitute away from monetary health investment through *M* following an increase in education in order to maintain one's initial utility level (Grossman, 1972).

Table 1: Overview of key variables and their predicted effects

Variable	Grossman's Assumptions	Expected empirical result: Demand curves	Grossman's Assumptions	Expected empirical result: MEC curves
Age	Increase in the demand for M_i Increase in the demand for TH_i	Increase in the demand for doctor visits and exercise Decrease in the demand for inactivity (BMI≥ 30), smoking, and drinking	MEC shifts left	Increase in the risk of MI or death
Wage	Increase in the demand for M_i Decrease in the demand for TH_i	Increase in the demand for doctor visits and inactivity (BMI ≥ 30) Decrease in the demand for exercise, smoking, and drinking	MEC shifts right	Decrease in the risk of MI or death

Education	Decrease in the demand for M_i Increase in the demand for TH_i	Increase in the demand for exercise Decrease in the demand for doctor visits, inactivity (BMI ≥ 30), smoking, and drinking	MEC shifts right	Decrease in the risk of MI or death *(Grossman, 1972)*

IV. Empirical Methods

In my empirical specification, I will use a Cox Proportional Hazard Model to estimate the stocks and flows of health capital (*H)*. The Grossman Model posits that individuals can choose their lifespans by investing in health. In this application of Grossman's framework, I hypothesize that individuals can also effectively increase the time until their first heart attack (or avoid one entirely) by making decisions related to their heart health. These decisions can serve either as a health investment or as a health divestment. In this analysis, I will control for the effects of the primary risk factors for MI including high blood pressure, angina, congestive heart failure, diabetes, obesity, smoking, and alcohol use. My sample is limited to those who have not had a heart attack upon entry to the study and those who are observed in at least two waves of the study. Each respondent will be evaluated upon entry (in their baseline year). The baseline information will be used to predict the number of days until the failure event: MI or death. By predicting a hazard ratio for each covariate, I will estimate each variable's effect on number of healthy days (until either MI or death). By using both death and MI as outcomes, I will be able to compare my findings on the validity of the application of the Grossman model to MI to its original use: general health. In this equation, health stock will serve as a latent variable while healthy days will represent the flow of health capital.

The following explanation describes the Cox proportional hazard model in the context of MI. The interpretation is the same for survival if we substitute risk of death for risk of MI. Risk can be measured by the hazard rate which, in this case, is the instantaneous probability of having a heart attack at a given time *t* conditional on *not* having had a heart attack yet. The risk of having an MI at a given time (*h(t)*) can be represented by

$$h(t) = h_0(t) \cdot \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i)$$

where $h_0(t)$ represents the risk of having a heart attack when all covariates are equal to zero. In other words, this represents the risk when an individual in my model does *not* have high blood pressure, or diabetes, does not smoke, and so on. When these covariates are non-zero, $\beta_i x_i$ represents the effect of a given covariate on the risk of heart attack. The hazard ratio ($\exp(\beta_i)$) represents the proportion of increase or decrease in heart attack risk which we can expect given a change in the corresponding covariate x_i. An increase in this covariate by one unit leads to an approximately $[\beta_i \cdot 100]$ % higher chance of heart attack at any given time (Cox Proportional-Hazards Model, n.d). In the context of the model, we can interpret as follows:

$\beta_i < 0,\ \exp(\beta_i) < 1$ indicates that risk is decreasing in the covariate x_i

$\beta_i = 0,\ \exp(\beta_i) = 1$ indicates that risk is constant in the covariate x_i

$\beta_i > 0,\ \exp(\beta_i) > 1$ indicates that risk is increasing in the covariate x_i

The Cox model assumes the presence of “proportional hazards” which means that the ratio of the risk of MI for one individual to another is constant over the course of the study. I tested this assumption using a chi square test for proportional hazards, and found a p-value < 15%. This is a limitation of my study, as the significance level is high. However, it would be difficult to ensure constant risk over the course of a 28 year-long study, so I will continue with my analysis. We can further test the proportional hazards assumption by visualizing the Kaplan-Meier survival estimate graphs. If the lines cross at any point or if one line goes to 0 while the other line flattens out, this condition is not satisfied. The appendix displays the Kaplan-Meier curves for the variable for regular exercise in figures 6 and 7, indicating that this condition is reasonably satisfied. Next, we must address the linearity assumption which requires that each covariate makes “a linear contribution to the model”. In order to test this assumption, we can use a residual plot which is shown in figures 4 and 5 in the appendix. As we can see, the predicted hazard ratios appear to have a fairly random distribution with no discernable pattern across observations. We can therefore assume that the linearity assumption is reasonably satisfied. Next, in order to ensure that the data does not include any potentially disruptive extreme values or outliers, I tabulated and generated summary statistics on each covariate in my analysis. During the data cleaning process, I eliminated several extreme values that were either errors or coding conventions (ex: 997 = Don’t know/Refused). Summary statistics and tabulations are displayed in the appendix. Finally, we must consider the independence assumption which prohibits relationships between survey responses that are not accounted for in the data. Since the HRS surveys such a wide population both in-person and via phone in order to promote accessibility and a wide reach, I will assume that these data are independent.

The Cox proportional hazard model is semi-parametric, as the model estimates parameters (β_i) though the estimation of these parameters themselves is not conducted parametrically. The Cox model is particularly useful in this case, compared to other models, as it allows for unbalanced panel data, as members of the study dropped out due to death or attrition in varying years (The Analysis Factor, 2020).

This empirical analysis will also estimate the health demand equations proposed in the Grossman framework using logistic regression. I have selected logit over probit in order to account for extreme values and in order to produce more robust estimates. The independence, linearity, and extreme value assumptions are met and discussed above. Health demand will be measured by the likelihood that an individual will visit the doctor at least twice per year or exercise regularly (at least 2 or 3 times per week). Negative health demand will be measured by the likelihood that an individual will be overweight (with a body mass index of at least 30), a current smoker, or a heavy/binge drinker. In my results section, I will report the marginal effect of each explanatory variable on each health demand indicator. This analysis will differ from past reformulations of the model as I will use binary dependent variables in order to avoid attributing meaning to marginal changes in number of doctor visits, for instance. While the impact of a marginal change from 0 to 1 doctor visit per year may be meaningful, we cannot necessarily say the same regarding the difference between 100 and 101 doctor visits per year.

Data

I will be using the University of Michigan's Health and Retirement Study (HRS) data to conduct my analysis. The HRS is a longitudinal panel study composed of in-depth interviews of American adults over 26 years. Interviews were conducted every other year beginning in 1992 and ending, for the time being, in 2018 resulting in 14 waves of data. I will be using the free and publicly available variables provided by the HRS and by the RAND corporation's HRS Fat Files and HRS Longitudinal Files. The HRS is the "largest and most comprehensive" survey of American adults. It is supported by the National Institute on Aging (NIA) and funded by the Social Security Administration in order to provide researchers with the first ever longitudinal dataset which includes both health and economic information about respondents. Interviews were conducted either in person or by phone and generally lasted between 1.5 and 3 hours. The HRS includes respondent, spouse (or other proxy), and household level data. I will be primarily using the respondent level data in order to compare my findings with the predictions of the Grossman model which predicts health outcomes based on individual, rather than household level, health investment. In the event that a participant dies, the HRS collects "Exit interview" data which includes medical costs, social interactions, and other relevant information in the time leading up to the person's death. This information is obtained from a close family member or spouse and was used to construct a variable specifying when a respondent died.

Within each interview wave, participants were asked questions about their demographic information, health and healthcare utilization, employment history, and many other topics. Each topic is contained in its own data file, so I merged the topic files that are relevant to my study by year. I then cleaned the data within these sections by constructing and labeling relevant variables and appended the data across each year/wave of interviews. I then cleaned the data across waves by interpolating missing values when appropriate and turning continuous variables (such as doctor visits or BMI) into binary variables. I then constructed variables indicating when each respondent entered and exited the study. For the purposes of the MI Cox model, I used participants' responses in each wave to generate a new variable describing the number of waves until a person reported their first MI as well as a binary variable describing whether they reported an MI at all. For the survival Cox model, I did the same using indicators for death and waves until death. Once I determined when, in the course of the study, each respondent either had a failure event or dropped out, I limited my observations to each respondent's baseline year responses leaving me with 27,417 total observations.

Table #2: Key Variables

Variable	Type	Description	Expected Effect on Health Demand	Expected Effect on Health Stock
Age	Explanatory (Quantitative, discrete)	Respondent age in baseline year	Positive	Negative
Natural Log of Hourly Wage	Explanatory (Quantitative, continuous)	Hourly wage (0 if inapplicable)	Positive	Positive
Natural Log of Household Income	Explanatory (Quantitative, continuous)	Household income	Positive	Positive
Doctor visits (≥ 2 per year)	Explanatory in health production function, Response in health demand function (Binary)	Respondent reports at least 2 doctor visits each year (not including overnight hospital or nursing home stays)	---	Positive
Educational Attainment	Explanatory (Categorical, ordinal)	Years of formal education ranging from 1 to 17 where 17 indicates any post-undergraduate education	Positive	Positive

Male	Explanatory (Binary)	Respondent is male (otherwise female)		
Residence type	Explanatory (Categorical, nominal)	Respondent lives on a farm (1), in a mobile home (2), does not live in a farm or mobile home but owns their property (3), does not live in a farm or mobile home but rents their property (4)	Farm (-) Mobile (-) Own (+) Rent (+)	Farm (-) Mobile (-) Own (+) Rent (+)
Government or Private Health Insurance	Explanatory (Binary)	Respondent has indicated that they hold either a government-provided or private health insurance policy (0 if they have not reported any type of health insurance provided in questionnaire)	Positive	Positive
Exercise (≥ 3 times/week)	Explanatory in health production function, Response in health demand function (Binary)	Respondent exercises at least 2 or 3 times per week	---	Positive
BMI ≥ 30	Explanatory in health production function, Response in health demand function (Binary)	Calculated BMI exceeds 30, placing respondent in "overweight" category	---	Negative
Current Smoker	Explanatory in health production function, Response in health demand function (Binary)	Respondent reports that they currently smoke cigarettes	---	Negative
Heavy of Binge Drinker	Explanatory in health production function, Response in health demand function (Binary)	Respondent is a heavy drinker (has ≥ 4 drinks on days that they drink) or binge drinker (has ≥ 14 drinks per week)	---	Negative
Paid Sick Leave (≥ 2 days/year)	Explanatory (Binary)	Respondent has at least 2 days of paid sick leave per year if employed.	Positive	Positive

The baseline age variable was constructed using the respondent's birth year and first year in study. Doctor visits were reported continuously across two years, so I converted the responses to a binary in which those who visited the doctor at least twice, on average, in a given year (or four times over two years) received a 1 and all others received a 0. I constructed the residence type variable using a series of variables provided in the data on housing information to assign values to the residence type variable as shown above. Regarding health insurance, since there was no question that asked whether or not a respondent was insured, I created a binary variable for insurance by assigning a 0 to anyone who did not report any type of health insurance and 1 otherwise. Since the interview question on exercise changed in several waves of the study, my exercise variable was constructed by standardizing the responses as much as possible such that respondents who reported exercising at least 2 or 3 times per week receive a 1 and 0 otherwise. I constructed the variable for body mass index (BMI) using height and weight information and converted this into a binary variable where respondents who reported BMI ≥ 30 received a 1 and 0 otherwise. The heavy/binge drinking variable was constructed based on the number of days that respondents reported drinking per week and the number of drinks that a respondent consumes on the days on which they drink. Finally, the number of days of paid sick leave was reported continuously in the HRS, so I constructed a binary variable indicating whether or not a respondent had at least 2 paid sick days per year (1) or less than 2 (0). This of course depends on whether or not the respondent is employed. All variables not discussed were provided in the data and did not have to be constructed.

A limitation that I have encountered in the HRS is the masking of respondent-identifying variables. I am unable to control for race or predict the effect of a person's travel time to their nearest clinic or their region of residence. What this dataset gains in specificity through the provision of personal medical details, mental health information, and self-perception, it lacks in specificity for the sake of protecting respondent privacy.

V. Results

Overall, in my empirical findings, education and age predict MI and survival outcomes in support of the Grossman model, while the results for wage and income do not support Grossman's framework. Those who are more educated and those who exercise regularly are at a significantly lower risk of MI, yet the significance of each effect is largely reduced with the inclusion of controls for self-reported health status and a series of heart health-related comorbidities. An increase in age is a highly significant contributor to increased MI risk. Men, people who have a BMI over 30, and people who smoke cigarettes all show a significantly higher risk of MI. Those who are heavy or binge drinkers show a significant decrease in MI risk. Hourly wage appears to indicate a decreased risk of MI while household income indicates an increased risk of MI, though neither of these effects are significant.

Those who visit the doctor more than twice per year have a higher risk of MI, a result which is significant even when controlling for health status and comorbidities. Similarly, those who reported some type of government-provided or private health insurance plan had an overall higher risk of having an MI. The only significant effect of residence type on MI risk was a relatively higher risk of MI for those who live in mobile homes relative to those who live on farms. Finally, an improvement in self-reported health showed a highly significant decrease in MI risk and each heart health-related comorbidity shows an expected increase in MI risk.

Table 3: Risk of Myocardial Infarction Given Health Inputs

Health Input	Hazard Ratio	p-value	[95% Conf	Interval]
Age	1.034***	0.000	1.029	1.039
Natural Log of Hourly Wage	0.956	0.223	0.890	1.028
Natural Log of HH. Income	1.008	0.764	0.956	1.063
Doctor visits (≥ 2 per year)	1.418***	0.000	1.300	1.548
Educational Attainment	0.971***	0.000	0.960	0.983
Male	1.614***	0.000	1.486	1.753
Baseline Residence: Farm				
Mobile Home	1.244*	0.076	0.977	1.585
(Neither) Own	0.888	0.241	0.728	1.083
(Neither) Rent	1.026	0.816	0.826	1.274
Health Insurance	1.129**	0.010	1.029	1.238
Paid Sick Leave (≥ 2 days/year)	0.850**	0.015	0.745	0.969
Exercise (≥ 2 or 3 times/week)	0.858***	0.001	0.782	0.941

BMI ≥ 30	1.428***	0.000	1.299	1.570
Current Smoker	1.793***	0.000	1.633	1.969
Heavy/Binge Drinker	0.889**	0.043	0.793	0.996

**** p<.01, ** p<.05, * p<.1*

Table 4: Risk of Myocardial Infarction Given Health Inputs (with additional controls)

Health Input	Hazard Ratio	p-value	[95% Conf	Interval]
Age	1.031***	0.000	1.026	1.036
Natural Log of Hourly Wage	0.948	0.143	0.882	1.018
Natural Log of HH. Income	1.010	0.703	0.959	1.065
Doctor visits (≥ 2 per year)	1.111**	0.026	1.012	1.219
Educational Attainment	0.992	0.248	0.980	1.005
Male	1.626***	0.000	1.497	1.767
Baseline Residence: Farm				
Mobile Home	1.142	0.281	0.897	1.455
(Neither) Own	0.867	0.159	0.710	1.057
(Neither) Rent	0.915	0.424	0.736	1.137
Health Insurance	1.182***	0.000	1.077	1.296
Paid Sick Leave (≥ 2 days/year)	0.865**	0.030	0.759	0.986
Exercise (≥ 2 or 3 times/week)	0.938	0.180	0.855	1.030
BMI ≥ 30	1.267***	0.000	1.150	1.395
Current Smoker	1.78***	0.000	1.620	1.957
Heavy/Binge Drinker	0.889**	0.044	0.793	0.997
Baseline Health: Poor				
Fair	0.839**	0.034	0.714	0.987
Good	0.749***	0.000	0.637	0.880
Very Good	0.619***	0.000	0.520	0.737
Excellent	0.495***	0.000	0.406	0.603
High Blood Pressure	1.326***	0.000	1.214	1.447
Diabetes	1.515***	0.000	1.353	1.697
Cancer	1.145*	0.090	0.979	1.340
Angina	1.848***	0.000	1.525	2.239
Congestive Heart Failure	1.836***	0.000	1.362	2.475
Stroke	1.203*	0.066	0.988	1.465

**** p<.01, ** p<.05, * p<.1*

Regarding the Cox Proportional Hazard Model for survival (Tables 5 and 6), education and regular exercise show significant evidence of a decrease in risk of death. In this case, the addition of health ratings and comorbidities reduce the significance of the effect of education, yet the effect of exercise remains highly significant even after the addition of controls in table 6. Men and people who smoke cigarettes both show a significantly higher risk of death. Those who are heavy or binge drinkers are also at a higher risk, yet this effect is not significant. Higher hourly wage and household income are associated with a decrease in mortality, yet these effects are not significant. Visiting the doctor more than twice per year appears to significantly increase risk of death in the restricted model (table 5), and while the addition of health controls appears to reverse this effect, the result in table 6 is not significant. Respondents who reported some type of government-provided or private health insurance plan showed higher risks, yet this result was only significant in table 6 when controlling for health variables. Those who own their homes but do not live on a farm or mobile home have a significantly lower risk of mortality compared to those who live on a farm, in a mobile home, or rent their property. As expected, older individuals, those with lower self-reported health, and those with any of the listed comorbidities showed relatively higher risk of death.

Table 5: Risk of Death Given Health Inputs

Health Input	Hazard Ratio	p-value	[95% Conf	Interval]
Age	1.087***	0.000	1.084	1.090
Natural Log of Hourly Wage	0.991	0.648	0.952	1.031
Natural Log of HH. Income	0.985	0.320	0.957	1.014
Doctor visits (≥ 2 per year)	1.312***	0.000	1.251	1.377
Educational Attainment	0.975***	0.000	0.969	0.981
Male	1.375***	0.000	1.315	1.439
Baseline Residence: Farm				
Mobile Home	1.369***	0.000	1.187	1.578
(Neither) Own	1.016	0.795	0.901	1.146
(Neither) Rent	1.257***	0.001	1.104	1.430
Health Insurance	1.019	0.501	0.964	1.077
Paid Sick Leave (≥ 2 days/year)	0.863***	0.001	0.792	0.941
Exercise (≥ 2 or 3 times/week)	0.837***	0.000	0.796	0.879
BMI ≥ 30	1.025	0.301	0.978	1.075

Current Smoker	2.183***	0.000	2.071	2.302
Heavy/Binge Drinker	1.028	0.380	0.967	1.092

**** p<.01, ** p<.05, * p<.1*

Table 6: Risk of Death Given Health Inputs (with additional controls)

Health Input	Hazard Ratio	p-value	[95% Conf	Interval]
Age	1.086***	0.000	1.083	1.089
Natural Log of Hourly Wage	0.985	0.456	0.946	1.025
Natural Log of HH. Income	0.988	0.412	0.960	1.017
Doctor visits (≥ 2 per year)	0.996	0.277	0.989	1.003
Educational Attainment	1.050*	0.063	0.997	1.105
Male	1.376***	0.000	1.315	1.440
Baseline Residence: Farm				
Mobile Home	1.274***	0.001	1.104	1.469
(Neither) Own	0.993	0.906	0.880	1.120
(Neither) Rent	1.143**	0.043	1.004	1.301
Health Insurance	1.063**	0.032	1.005	1.123
Paid Sick Leave (≥ 2 days/year)	0.883***	0.004	0.810	0.962
Exercise (≥ 2 or 3 times/week)	0.910***	0.000	0.866	0.957
BMI ≥ 30	0.942**	0.015	0.897	0.988
Current Smoker	2.167***	0.000	2.054	2.286
Heavy/Binge Drinker	1.039	0.216	0.978	1.104
Baseline Health: Poor				
Fair	0.785***	0.000	0.719	0.857
Good	0.658***	0.000	0.603	0.717
Very Good	0.554***	0.000	0.504	0.608
Excellent	0.476***	0.000	0.429	0.530
High Blood Pressure	1.201***	0.000	1.145	1.259
Diabetes	1.617***	0.000	1.518	1.722
Cancer	1.233***	0.000	1.142	1.331
Angina	1.098	0.115	0.978	1.234
Congestive Heart Failure	1.853***	0.000	1.558	2.204
Stroke	1.341***	0.000	1.215	1.480

**** p<.01, ** p<.05, * p<.1*

In the health demand functions, age has a significant positive effect on the demand for doctor visits and a significant negative effect on the demand for inactivity/poor diet (for which body mass index ≥ 30 serves as a proxy), smoking, and heavy or binge drinking. Age has no significant effect, however, on demand for regular exercise in the absence of health status and comorbidity controls. With the inclusion of these controls, an increase in age indicates a small yet significant increase in demand for exercise. Those with a higher education show significant evidence of an increase in demand for doctor visits, exercise, and heavy/binge alcohol consumption. The better educated, however, have a lower demand for inactivity/poor diet (high BMI) and cigarette smoking. Men show significant evidence of having less demand for medical care (doctor visits), but higher demand for both exercise and inactivity/poor diet as well as cigarettes and alcohol. In the absence of the control variables included in table 8, it appears as if those own their houses have higher demand for medical care through doctor visits compared to those who rent their property, followed by those who live in mobile homes, and by those who live on farms. Once I controlled for health status, however, the only significant effect that remained was the relatively higher demand for medical care among those who own their houses. Those who rent or live in mobile homes showed decreased yet insignificant demand for doctor visits.

There is significant evidence of an overall higher demand for regular exercise among those who live on a farm relative to those who do not. The highest demand for cigarettes and heavy/binge alcohol consumption exists among those who live in mobile homes and the lowest exists among those who live on farms. These results are significant in the demand for smoking, yet the results for alcohol consumption are overall less significant and all results show a decrease in significance after the addition of controls. There is no significant evidence of a difference in demand for inactivity/poor diet (high BMI) across residence types. Across the board, higher self-reported health indicates a significant decrease in demand for doctor visits and cigarettes as well as a significant increase in demand for exercise. Finally, there is a significantly lower demand for exercise among those with high blood pressure, diabetes, congestive heart failure, and a history of stroke.

Table 7: Marginal Effects of Health Inputs in Health Demand Functions

	Doctor Visits (≥ 2 per year)	Exercise (≥ 2 or 3 times/week)	BMI ≥ 30	Current Smoker	Heavy/Binge Drinker
Age	0.007***	-0.000	-0.002***	-0.007***	-0.002***
Natural Log of Wage	0.000	0.003	0.008*	-0.000	0.019***
Natural Log of HH. Income	0.002	0.003	-0.004	0.001	0.001
Educational Attainment	0.009***	0.011***	-0.012***	-0.010***	0.002**
Male	-0.101***	0.051***	0.098***	0.039***	0.075***
Residence Baseline: Farm					
Mobile Home	0.048***	-0.107***	0.016	0.152***	0.034**
(Neither) Own	0.053***	-0.075***	0.018	0.021**	0.012
(Neither) Rent	0.053***	-0.134***	0.018	0.140***	0.031**
Health Insurance	-0.003	-0.105***	-0.062***	0.047***	-0.272***
Paid Sick Leave (≥ 2 days/year)	0.078***	-0.010	0.022***	-0.045***	-0.010

*** *p<.01, *** *p<.05, ** *p<.1*

Table 8: Marginal Effects of Health Inputs in Health Demand Functions (with additional controls)

	Doctor Visits (≥ 2 per year)	Exercise (≥ 2 or 3 times/week)	BMI ≥ 30	Current Smoker	Heavy/Binge Drinker
Age	0.005***	0.000	-0.004***	-0.007***	-0.002***
Natural Log of Wage	-0.004	0.007	0.006	-0.001	0.018***
Natural Log of HH. Income	0.003	0.003	-0.003	0.001	0.002
Educational Attainment	0.018***	0.003***	-0.008***	-0.007***	0.002***
Male	-0.093***	0.048***	0.096***	0.042***	0.076***
Residence Baseline: Farm					
Mobile Home	-0.004	-0.077***	-0.003	0.143***	0.034**
(Neither) Own	0.033**	-0.067***	0.013	0.022**	0.013
(Neither) Rent	-0.013	-0.098***	-0.011	0.130***	0.029**
Health Insurance	0.015**	-0.122***	-0.052***	0.054***	-0.270***
Paid Sick Leave (≥ 2 days/year)	0.072***	-0.010	0.017**	-0.041***	-0.009
Baseline Health: Poor					
Fair	-0.113***	0.101***	0.049***	-0.032***	0.050***
Good	-0.193***	0.158***	0.061***	-0.070***	0.031***
Very Good	-0.261***	0.222***	0.029**	-0.111***	0.020**
Excellent	-0.349***	0.327***	-0.049***	-0.153***	0.004
High Blood Pressure	0.143***	-0.024***	0.152***	-0.039***	0.004
Diabetes	0.140***	-0.022***	0.144***	-0.056***	-0.011
Cancer	0.167***	0.020**	-0.013	-0.004	0.005
Angina	0.220***	0.007	0.025	-0.025*	-0.017
Congestive Heart Failure	0.193***	-0.118***	0.085***	-0.004	-0.012
Stroke	0.074***	-0.048***	-0.030*	0.030**	-0.003

*** *p<.01, *** *p<.05, ** *p<.1*

VI. Discussion and Conclusion

The Cox models discussed above describe the changes in supply of health capital caused by a set of health inputs. The health capital flows are measured in terms of potential for MI prevention and in terms of the prevention of death. The results above show strong support for Grossman's theory that health stock is depreciating in age and increasing in education. The demand functions shown in tables 7 and 8 indicate that an increase in age causes an increase in demand for medical care through regular doctor visits. Exercise appears to be a meaningful indicator of health investment excluding the prediction of the effects of age. Finally, as age increases, we see decreases in demand for all three health divestment functions (demand for high BMI, smoking, and heavy/binge drinking).

The results for wage and income are overall insignificant. Grossman used wage in his original empirical formulation of the pure investment model as the effect of an increase in the opportunity cost of time can be reflected through a change in hourly wage. Grossman also used household income in his formulation of the pure consumption model due to the effects of diminishing returns to investment in Z on health demand. Overall, the results of this analysis support past empirical findings, including Grossman's own, which disprove his expectations of the effects of wage and income. While hourly wage increases are associated with a decrease in MI risk, increases in household income lead to increases in MI risk, yet both of these results are insignificant. Both wage and income appear to mitigate risk of death, yet these results are not significant either. Regarding demand for health, an increase in wealth (through wages or, in this case, income) is expected to lead to an increase in demand for medical care but a possible decrease in demand for time-consuming activities such as exercise. The demand functions provide inconclusive results regarding the relationship between wealth and medical care demand. The marginal effect of wage on demand for doctor visits is negative only when controlling for health status and the listed comorbidities, which is expected, yet this result is not significant. Income, however, does not have the predicted effect and is not significant. The demand for exercise should reflect the predicted decrease in demand for time-consuming health investment activities, yet the marginal effect of wage is positive and insignificant. These results are likely influenced by the fact that nearly 50% of respondents included in the final sample did not report an hourly wage. This reflects a shortcoming of the Grossman model due to its lack of ability to represent changes in opportunity cost of time without the presence of hourly wages.

Educational attainment is predicted to cause a decrease in demand for health care through doctor visits yet an increase in demand for other methods of health investment such as exercise. My findings show that education increases the demand for both types of medical care, which partially supports Grossman's theory. Additionally, education is shown to significantly reduce risk of MI and death, supporting the claim that more educated individuals are more efficient producers of health and are able to optimize at relatively higher levels of health stock compared to the less-educated.

Doctor visits have a significant negative effect on health stocks and flows. Additionally, as a respondent's health improves in rating, they become less and less likely to demand doctor visits at least twice per year. These results are highly significant. According to Peter Zweifel (2012), the empirical finding that sicker people go to the doctor more implying an inverse relationship between health and doctor visits discredits a key assumption of Grossman's pure investment model. Grossman posits that those who demand more health will be healthier. Since people can demand health by "demanding" doctor visits, then surely those who demand more doctor visits will be healthier (Zweifel, 2012). Robert Kaestner (2012) writes a response to Zweifel stating that the Grossman model allows for a changing marginal cost of health. By utilizing a non-fixed ratio between health care expenditure and cost of health enhancing efforts, we can model changes in health demand not through changes in the demand function but through changes in investment cost. He further argues that due to the existing evidence of a positive association between good health behaviors (not just including doctor visits) and good health, Zweifel's argument of reverse causality does not invalidate Grossman's theory as his model allows for non-healthcare methods of health investment (Kaestner, 2012). If Grossman's predictions were correct, or supported by my analysis, however, any confounding effects of illness on number of doctor visits should be accounted for by the presence controls for health rating and comorbidities, yet evidence in support of Zweifel's argument remains even in the presence of these controls. Further, extreme values in number of doctor visits per year were masked through the use of the binary doctor visits variable, eliminating the possibility of attributing large changes in health to high numbers of doctor visits, as these are likely observed only in the case of chronic illness and not in the case of an extremely high demand for medical care. Zweifel also argues that there is no empirical evidence that more health care consumption is associated with higher levels of health. We can understand this argument intuitively as sick people need to use more health care. This reverse causality, according to Zweifel, invalidates Grossman's health investment production function (Zweifel, 2012).

According to Wagstaff (1993), this finding also contradicts the theory of derived demand. This framework assumes that when someone goes to the doctor, for instance, they are motivated by their demand for "good health" or "longevity" and not a demand for the healthcare service itself. Demand for goods and services such as healthcare (whether for preventative or treatment purposes), nutritious food, or gym access is then classified as "derived demand" as it serves as an "input" into the creation of health. When the shadow price of health changes, the demand for health as well as the derived demand for health inputs change (Wagstaff, 1993).

Grossman (1972) hypothesizes that females are less efficient producers of health, partially due to a higher intensity of use during childbirth. Past empirical studies have also found that females tend to have worse health than males (Galama, 2012). Men in my sample, however, had worse outcomes across the board. The lower life expectancy among men in the US as well as the fact that men are more likely to exhibit risk-seeking tendencies is a likely explanation for these findings. Men are significantly less likely to visit the doctor regularly and are more likely to pursue risky behaviors such as smoking and heavy/binge drinking.

I used residence type to proxy the effect of reduced access to medical care among people living in remote areas. My findings largely do not support the use of this proxy regarding the production of health stock apart from the finding that mortality risk is the highest among those who live in mobile homes. This does reflect the anticipated adverse effect of remote or low-income housing on one's ability to seek medical services even if they wanted to. Demand for medical care is significantly higher for those who do *not* live on a farm, further supporting the claim that remote housing may reduce ability to demand medical care in the first place. The results for exercise demand are interesting, however, as farm residents are significantly more likely to demand regular exercise compared to those who do not live on farms. This supports Wagstaff's aforementioned hypothesis that the demand for health cannot be measured solely or even predominantly through the use of medical care.

When predicting the demand for health using exercise as a dependent variable, I found a significant and positive association between level of self-reported health and likelihood of exercising at least 3 times per week. Those without at least two days of sick leave were significantly *less* likely to exercise frequently at a 5% significance level.

This model also shows significant evidence that participants who were insured were significantly less likely to exercise, implying some degree of moral hazard. Previous studies have shown a high likelihood that one's "anticipated behavioral response" to the purchase of an insurance contract will have an effect on their selection process (Finkelstein, 2014). This is not an effect that is directly considered in Grossman's framework, as insurance would only factor into the marginal cost of the demand for medical care. When we consider it as an explanatory variable, however, we can see that choosers may anticipate their own health-seeking and non-seeking behaviors making the purchase of insurance a predictor of health stock in and of itself. The Cox models show support for this hypothesis as those who are insured have significantly higher risk of MI and death. It appears as if insurance may have a significant effect on health and should be included in more Grossman model empirical works going forward.

Limitations and Critiques

Critics have remarked that the model does not properly account for SES and past health information, two factors which are critical in predicting MI outcomes especially for people on the margins. In his 1972 manuscript, Michael Grossman discusses the counterintuitive fact that the age-adjusted mortality rate is positively correlated with income in the US. Fluctuations in income, he theorizes, can "no longer be the major determinant of variations in mortality and morbidity" in a developed country such as the US in which income broadly "exceeds a subsistence level". Further, Grossman states that there is a positive correlation between income and the quantity and quality of medical care. He therefore proposes an economic model of the demand for health in which health itself is an endogenous variable. Additionally, SES should be accounted for through the inclusion of education and income variables. Support for the validity of these tools is lacking in this model, however.

The Grossman model and this paper do not take race into account as a predictor of health. This is certainly a limitation, as systemic medical racism in the US has led to worse outcomes for Black people, Indigenous people, and people of color (BIPOC) in comparison to white populations. A 2015 study found that the effect of race on MI mortality between Black and White patients is increasing in SES. While MI outcomes were similar for black and white patients of low SES, Black patients of higher SES had significantly lower life expectancies compared to their white counterparts. Even in the absence of medical racism, Black people are more likely than White people to suffer from CVD risk factors and to have low SES (Bucholz et al., 2015). SES is a risk factor in and of itself, as MI risk is two times as high for those of low SES compared to high SES. In fact, a study conducted by Hamad et al. found empirical evidence that CVD risk factors can predict *less* of the excess MI risk among low SES populations compared to factors such as poverty and education which are closely tied to SES (Hamad et al., 2020). These observations illuminate gaps in the argument that NCDs such as MI can be largely prevented by changing lifestyle choices. Dedicating one's time to exercise and a healthy diet may have lower returns for Black people or low SES populations and gaining access to high levels of education and high SES may have lower returns for Black people in terms of avoiding or recovering from MI.

The Grossman model is dynamic in its own right, as it has been widely discussed, refuted, and revised in the years since its inception. As in many economic models, the Grossman model comes with limitations. A particularly noteworthy weakness is its propensity to exclude information about high risk populations. This can be attributed to the fact that Grossman included only white Americans in his 1972 empirical analyses. While groups on the margins are consistently subjected to disparities in health outcomes and quality of care globally, MI is of particular concern due to its prevalence in the US as a whole and among marginalized populations.

My results show support for Grossman's model regarding age and education but not regarding wage and the positive effect of medical care. Going forward, it would be useful to consider more proxies of demand for health that do not rely on formal medicine and are not likely to be invalidated due to the association between illness and high numbers of doctor visits. It is also possible that the setting, the United States, creates substantial bias in the sense that healthcare is expensive and difficult to attain for many. Findings regarding health demand may be more meaningful in the presence of socialized medicine. Finally, the consideration of insurance status seems to play a large role in both health behaviors and outcomes. Further incorporation of this effect in future empirical tests of the Grossman model may help to explain how we make health-seeking and non-seeking decisions.

References

Alkhouli, M., MD, Alqahtani, F., MD, Holmes, D. R., MD, & Berzingi, C., MD. (2019, August 6). Racial Disparities in the Utilization and Outcomes of Structural Heart Disease Interventions in the United States. Retrieved October 19, 2020, from https://www.ahajournals.org/doi/epub/10.1161/JAHA.119.012125

American College of Cardiology. ASCVD Risk Estimator Plus. (n.d.). Retrieved October 19, 2020, from http://tools.acc.org/ASCVD-Risk-Estimator-Plus/

American Heart Association (AHA) News. (2019, January 31). Cardiovascular diseases affect nearly half of American adults, statistics show. Retrieved October 19, 2020, from https://www.heart.org/en/news/2019/01/31/cardiovascular-diseases-affect-nearly-half-of-american-adults-statistics-show

The Analysis Factor. (2020, August 26). Retrieved February 24, 2021, from https://www.theanalysisfactor.com/

Avis N. E., Smith K. W., and McKinlay J. B. (1989): Accuracy of perceptions of heart attack risk: what influences perceptions and can they be changed? American Journal of Public Health 79, 1608_1612, https://doi.org/10.2105/AJPH.79.12.1608

Bucholz, E. M., Ma, S., Normand, S. T., & Krumholz, H. M. (2015). Race, Socioeconomic Status, and Life Expectancy After Acute Myocardial Infarction. *Circulation, 132*(14), 1338-1346. doi:10.1161/circulationaha.115.017009

Canto JG, Kiefe CI, Rogers WJ, Peterson ED, Frederick PD, French WJ, Gibson CM,

Center for Health Informatics (CHI). (2019). Non-communicable diseases. Retrieved March 22, 2021, from https://www.nhp.gov.in/healthlyliving/ncd2019

Cropper, M. (1981). Measuring the Benefits from Reduced Morbidity. *The American Economic Review, 71*(2), 235-240. Retrieved March 21, 2021, from http://www.jstor.org/stable/1815724

Finkelstein, A. (2014). Moral hazard in health insurance. *Moral Hazard in Health Insurance,* 13-44. doi:10.7312/columbia/9780231163804.003.0002

Pollack CV Jr, Ornato JP, Zalenski RJ, Penney J, Tiefenbrunn AJ, Greenland P; NRMI Investigators. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. JAMA. 2011 Nov 16;306(19):2120-7. doi: 10.1001/jama.2011.1654. PMID: 22089719; PMCID: PMC4494683.

Centers for Disease Control. NHIS - National Health Interview Survey. (2020, October 09). Retrieved October 19, 2020, from https://www.cdc.gov/nchs/nhis/index.htm

Charan, R. (2020, July 23). The Cox Proportional Hazards Model. Retrieved November 09, 2020, from https://towardsdatascience.com/the-cox-proportional-hazards-model-35e60e554d8f

Cox Proportional-Hazards Model. (n.d.). Retrieved November 09, 2020, from http://www.sthda.com/english/wiki/cox-proportional-hazards-model

Fox, A. N., Skrepnek,, G. H., Miller, J. L., Schwier, N. C., & Ripley, T. L. (2018, November 10). Optimal Medical Therapy Prescribing Patterns and Disparities Identified in Patients with Acute Coronary Syndromes at an Academic Medical Center in an Area with High Coronary Heart Disease-Related Mortality. Retrieved October 19, 2020, from https://link-springer-com.proxy.lib.duke.edu/article/10.1007/s40256-018-0308-x

Galama, T. J., Hullegie, P., Meijer, E., & Outcault, S. (2012). Is There Empirical Evidence For Decreasing Returns To Scale In A Health Capital Model? *Health Economics, 21*(9), 1080-1100. doi:10.1002/hec.2843

Galama, T., & Kippersluis, H. V. (2018). A Theory of Socioeconomic Disparities in Health Over the Life Cycle. doi:10.7249/wr773

Grossman, M. (1972). *The demand for health: A theoretical and empirical investigation*. New York: National Bureau of Economic Research.

Grossman, M. (1972a). 'On the concept of health capital and the demand for health', Journal of Political Economy, vol. 80(2), pp. 223–55.

Hamad R, Penko J, Kazi DS, et al. (2020). Association of Low Socioeconomic Status With Premature Coronary Heart Disease in US Adults. *JAMA Cardiol.* 2020;5(8):899–908. doi:10.1001/jamacardio.2020.1458

Hamil-Luker, J., & O'rand, A. M. (2007). Gender Differences in the Link Between Childhood Socioeconomic Conditions and Heart Attack Risk in Adulthood. *Demography, 44*(1), 137-158. doi:10.1353/dem.2007.0004

The Health and Retirement Study. (n.d). Retrieved November 09, 2020, from https://hrs.isr.umich.edu/about

Jones, A. M., Laporte, A., Rice, N., & Zucchelli, E. (2018). Dynamic panel data estimation of an integrated Grossman and Becker–Murphy model of health and addiction. *Empirical Economics, 56*(2), 703-733. doi:10.1007/s00181-017-1367-6

Kaestner, R. (2012). The Grossman model after 40 years: A reply to Peter Zweifel. *The European Journal of Health Economics, 14*(2), 357-360. doi:10.1007/s10198-012-0451-2

Kaiser Permanente. Atherosclerotic Cardiovascular Disease (ASCVD) Primary PreventionGuideline. (2018, April). Retrieved from https://wa.kaiserpermanente.org/static/pdf/public/guidelines/ascvd-primary.pdf

Lahiri, K., Li, X. Smoking Behavior of Older Adults: A Panel Data Analysis Using HRS. *J. Quant. Econ.* 18, 495–523 (2020). https://doi.org/10.1007/s40953-020-00196-x

Laporte, A. (2014). *Should the Grossman model retain its Iconic status in health economics?* (Doctoral dissertation, University of Toronto, November 24, 2014) (pp. 1-28). Canadian Centre for Health Economics.

Muurinen, J. (1982). Demand for health. *Journal of Health Economics, 1*(1), 5-28. doi:10.1016/0167-6296(82)90019-4

National Health Service (NHS) of the UK. (2018, September 17). Cardiovascular disease. Retrieved October 19, 2020, from https://www.nhs.uk/conditions/cardiovascular-disease/

Ramakrishnan R, Doherty A, Smith-Byrne K, Rahimi K, Bennett D, Woodward M, et al. (2021) Accelerometer measured physical activity and the incidence of cardiovascular disease: Evidence from the UK Biobank cohort study. PLoS Med 18(1): e1003487. https://doi.org/10.1371/journal.pmed.1003487

Ruhm, C.J. (2006). A healthy economy can break your heart. NBER working paper no. 12102.

Smith, J. P. (2007). The Impact of Socioeconomic Status on Health over the Life-Course. *Journal of Human Resources, XLII*(4), 739-764. doi:10.3368/jhr.xlii.4.739

Spertus, J. A. (2009). Factors Associated With Racial Differences in Myocardial Infarction Outcomes. *Annals of Internal Medicine, 150*(5), 314. doi:10.7326/0003-4819-150-5-200903030-00007

Svensson, M. (2007). Do not go breaking your heart: Do economic upturns really increase heart attack mortality? *Social Science & Medicine, 65*(4), 833-841. doi:10.1016/j.socscimed.2007.04.015

Wagstaff, A. (1986). The demand for health. *Journal of Health Economics: Some New Empirical Evidence, 5*(3), 195-233. doi:10.1016/0167-6296(86)90015-9

Wagstaff, A. (1993). The demand for health: An empirical reformulation of the grossman model. *Econometric Analysis of Health Data,* 13-23. doi:10.1002/0470846313.ch1

Zweifel, P. (2012). The Grossman model after 40 years. *The European Journal of Health Economics, 13*(6), 677-682. doi:10.1007/s10198-012-0420-9

VII. Appendix

Table 9: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	606246	59.719	11.767	18	106
Natural Log of Hourly Wage	606302	1.694	1.477	-4.894	10.086
Natural Log of HH Income	606302	10.205	2.007	-2.931	16.384
leave	42674	30.43	90.676	0	365

Table 10: Tabulation of Dr Visits (≤ 2x per year)

Value	Freq.	Percent	Cum.
0	188930	32.10	32.10
1	399584	67.90	100.00
Total	588514	100.00	

Table 11: Tabulation of Male

Value	Freq.	Percent	Cum.
0	339759	56.04	56.04
1	266515	43.96	100.00
Total	606274	100.00	

Table 12: Tabulation of Residence Type

Value	Freq.	Percent	Cum.
1	12938	2.37	2.37
2	33970	6.23	8.61
3	384746	70.61	79.22
4	113233	20.78	100.00
Total	544887	100.00	

Table 13: Tabulation of Govt. or Private Insurance

Value	Freq.	Percent	Cum.
0	408384	67.36	67.36
1	197918	32.64	100.00
Total	606302	100.00	

Table 14: Tabulation of Exercise (≥ 3x per week)

Value	Freq.	Percent	Cum.
0	456149	75.23	75.23
1	150153	24.77	100.00
Total	606302	100.00	

Table 15: Tabulation of BMI ≥ 30

Value	Freq.	Percent	Cum.
0	179997	32.87	32.87
1	367607	67.13	100.00
Total	547604	100.00	

Table 16: Tabulation of Current Smoker

Value	Freq.	Percent	Cum.
0	508957	84.01	84.01
1	96861	15.99	100.00
Total	605818	100.00	

Table 17: Tabulation of Heavy or Binge Drinker

Value	Freq.	Percent	Cum.
0	468931	77.36	77.36
.2	1	0.00	77.36
.2857143	1	0.00	77.36
1	137206	22.64	100.00
Total	606139	100.00	

Table 18: Tabulation of Educational Attainment

Value	Freq.	Percent	Cum.
0	4982	0.88	0.88
1	1535	0.27	1.15
2	2670	0.47	1.62
3	5753	1.02	2.64
4	4563	0.81	3.45
5	5506	0.97	4.42
6	12295	2.17	6.59
7	8316	1.47	8.06
8	24843	4.39	12.45
9	18808	3.32	15.77
10	27172	4.80	20.57
11	27368	4.83	25.41
12	177647	31.38	56.79
13	43193	7.63	64.42
14	58956	10.42	74.84
15	21897	3.87	78.70
16	63342	11.19	89.89
17	57204	10.11	100.00
Total	566050	100.00	

Table 19

Test of proportional-hazards assumption	chi2	df	Prob>chi2
	19.620	14	0.142

Table 20: Waves until respondent either left study or had an MI

	Respondent had an MI		
Waves	0	1	Total
1	5791	462	6253
2	2221	379	2600
3	2526	347	2873
4	5440	289	5729
5	1492	239	1731
6	1403	222	1625
7	2847	160	3007
8	1237	136	1373
9	1130	114	1244
10	2399	85	2484
11	738	62	800
12	930	49	979
13	3480	21	3501
Total	31634	2565	34199

Fig 4: Residual visualization for MI Cox

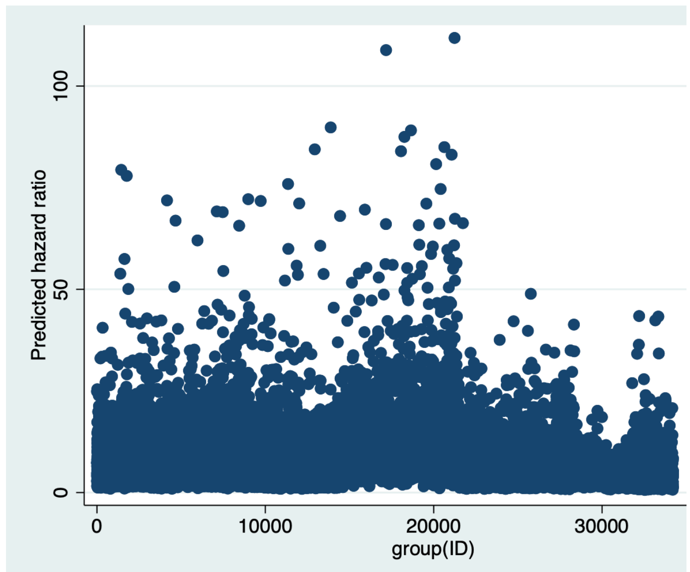
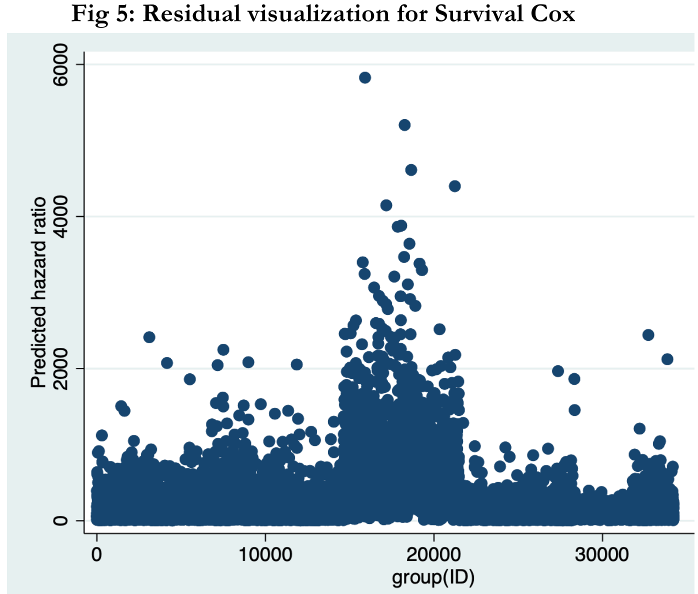


Table 21: Waves until respondent either left study or died

	Respondent died		
Waves	0	1	Total
1	4658	770	5428
2	1052	1152	2204
3	1473	1197	2670
4	4548	1182	5730
5	542	1049	1591
6	684	1021	1705
7	2242	898	3140
8	506	841	1347
9	682	783	1465
10	2228	567	2795
11	398	433	831
12	751	442	1193
13	3973	141	4114
Total	23737	10476	34213

Fig 5: Residual visualization for Survival Cox



Figures 6 & 7: Kaplan Meier Curves for Educational Attainment