Chapter 2
Insuring Health Capital

2.1 The Economic Value of Health

2.1.1 Consuming and Investing in Health

Economics is concerned with the behavior of individuals as consumers. Individuals make choices about their consumption every day. They choose what to eat for lunch, whether to walk or drive to their next meeting, and whether or not to go to the gym after work. How individuals make these consumption choices is one of the most important forms of behavior that economists study. A primary assumption of economics is that consumers have preferences over different types of consumption, and that they make choices in order to satisfy their preferences. In economic terms, individuals make the consumption choices that "maximize their utility." For example, a consumer may choose to drink coffee instead of water because coffee delivers more utility than water.[1]

The economic view of health is that it is a good that individuals consume. Individuals derive utility from good health in the same way that they derive utility from drinking coffee. The "consumption of health" refers to the same phenomenon: people experience a positive benefit from good health in the present and in the future. Over a given time period, say a particular day, part of the utility that individuals derive comes from the consumption of health. If individuals are relatively healthy, then they enjoy a large amount of utility from their health. If individuals are in relatively poor health, then they derive much less utility from their health. That people prefer being healthy is the economic model for a common sense idea.

Health is distinct from many other goods in that it has both an intrinsic and an instrumental value. Being in good health is a state that gives rise to positive

[1]A full description of the economics of utility maximization is beyond the scope of this book. An excellent treatment of utility maximization can be found in Chap. 7 of Varian's "Microeconomic Analysis" (Varian 1992).

R.D. Lieberthal, *What Is Health Insurance (Good) For?*,
DOI 10.1007/978-3-319-43796-5_2

utility—people enjoy being in good health. Good health also allows a person to work to earn money, to take care of themselves and others, and to engage in other activities. Feeling well is the intrinsic value of health—it directly increases individual utility. Using a state of good health to earn a living is an instrumental value of health—it indirectly increases individual utility by providing monetary wages that people can spend on utility inducing forms of consumption. Using good health to engage in activities of daily living is also an instrumental value of health—it indirectly increases individual utility by providing people with the opportunity to engage in activities they enjoy, such as spending time with friends and family.

Health is also like other forms of consumption in that it is subject to a degree of individual choice. The economics of choice relate to overall consumption decisions, including how much to consume of various goods today, and how much to put off (save) until tomorrow (Mas-Colell et al. 1995, this chapter). The economics of choice apply to the consumption of health in the sense that people can, to an extent, choose how much health to consume. Individuals can decide how to enjoy the direct benefits of their health today, and can engage in health promoting or health destroying activities. They can also decide how much of the indirect benefits of their health they would like to use for market activities, specifically working to earn a living.

The economics of choice also applies to health in the sense that individuals have some choice about how much health to consume. Economics emphasizes the role of "nonsatiation" of goods as explaining the choice of more consumption over less. If one cookie is good, then people will be happier with two cookies than one, and even happier with three cookies than two, and so on. Since health is good, then people will want to consume more and more of it. Economics also emphasizes the role of declining marginal benefit as explaining the choice to limit consumption of a particular good. People may derive more utility from the first cookie than the second cookie, and more utility from the second cookie than the third cookie, and so on. While people will prefer good health to poor health, they might not want to spend all their time and resources trying to achieve and maintain perfect health.

Maximizing utility with respect to health also requires balancing consumption today with consumption in the future. This is referred to in economics as the problem of maximizing "intertemporal utility" (Mas-Colell et al. 1995, pp. 733–736). We have seen that there may be a declining marginal benefit of health consumption. We know that future health is both uncertain and determined, in part, by health and non-health decisions made today. We can best examine this balancing through the health-related choices that individuals make. The amount spent on healthcare demonstrates a strong desire by people to remain healthy in the future even if it means spending a significant amount of time and money today.

Declining marginal benefit of health and the fact that health deteriorates over time also implies that individuals are worse off than they would be if they could somehow "transfer" some of their good health from the present to the future. Most forms of consumption provide greater utility when spread over time rather than being consumed all at once. A person who receives their monthly paycheck is generally better off spending the money over the course of the month rather than

spending all the money at once. In finance, money and banking technologies allow a person to "transfer" their earnings between periods. If individuals had a similar way to transfer health across periods, they would be better off since they would be able to choose when to consume their health. However, this is not a choice that people can make—it is impossible to "save" current health in the same way that a person can save a dollar or a physical asset like a car. Based on currently available choices, the main alternative is to invest in health to improve future consumption.

2.1.2 Health Investments

While people are unable to save their health, they are able to invest in their health. In economic terms, investment involves a cost paid today that generates returns in the future. Costs and returns for investments are generally expressed in both monetary and nonmonetary terms. For example, a person starting a new business must invest both start-up capital (money) as well as their own time. The returns would include both the income from the business and any positive feelings arising from being one's own boss. Health investments could have similar monetary and nonmonetary costs and returns. The cost of investment in health might include money spent on a doctor's visit and the time spent in the doctor's office. The returns could include a feeling of good health in the future and the improved ability to work to earn a living. This ability to invest in health implies that investment opportunities allow individuals to achieve a "payoff" in health consumption in the future by expending money, effort, and other resources today.

Two particularly important types of health investment are prevention and health promotion. The purpose of prevention is to raise health consumption in the future, thereby raising future utility. In that sense, preventative spending can be considered as a form of investment. The same is true of different forms of health promotion. Exercise, good diet, and proper sleep could be considered costly activities today: exercise requires time and effort, healthy food may be expensive, and time spent on sleep cannot be spent on other activities. To the extent that these activities improve health in the future, they are forms of investment in health.[2]

While many of the general principles of investment apply to health, one specific difficulty in applying investment models to health is that there is no market for good health. Health is not a fungible commodity—it cannot be transacted directly between individuals or bought and sold in a market. Financial securities like stocks and bonds can be bought and sold. The same is true in markets for physical capital. In contrast, investments in health are one way investments—money spent on health investment cannot be converted back into money or other financial goods. In this

[2]Note that it is not necessarily the case that a person will be healthier tomorrow than today as a result of these activities. Rather, they will be healthier tomorrow than they would have been in the absence of such activities. Economics measures the future relative to the counterfactual, i.e., what possible future states does a person face?

sense, the technologies for improving health today or transferring health across time periods are limited and costly. In order to understand why that fact matters for the formation of health insurance, we next formalize these aspects of health through the introduction and explanation of the health capital model.

2.1.3 The Health Capital Model

The Grossman health capital model was among the first to recognize the capital-like nature of health. When the model was first published in 1972, Grossman said that he aimed "…to construct a model of the demand for the commodity 'good health'" (Grossman 1972). The model is based on the proposition that individuals can invest in their health through both health promoting activities and healthcare (which Grossman labels "medical care" in his paper), in the way described in this chapter. Investment in future health "produces" health in the same way that investment in a factory may lead to the production of cars. However, Grossman contrasts the production of health capital with physical capital and human capital as established by Gary Becker, among others, as a person's productive capacity which can be used to earn a living (Becker 1967). Comparatively, health capital is a nonmarket good—it cannot be bought or sold, and it cannot be used to obtain money to buy other commodities.

The main questions the health capital model raises and that Grossman aims to answer are: (1) to what extent will people invest in their health capital? and (2) how do variables, such as age, alter health capital investment decisions? The answer to the first question depends largely on the price of investment in health capital. If medical care or other forms of health capital protection are less costly (has a lower price), people will tend to consume more of them. The answer to the second question depends on the elasticity of demand for health, meaning the response of individuals to changes in their health. Generally, as people age (or get less healthy for any reason), their demand for medical care rises as a direct result of their demand for health.

In terms of the health capital model, a person with an injury might possibly return to a pre-injury state of health. A person who breaks his wrist from falling off a ladder might heal fully through rest and physical therapy and return to the health state that he enjoyed before the injury. In this case, his loss would be the cost that he paid in time and money to recover from the injury. The price of healthcare to recover from the injury might determine whether a person recovers fully. If the price of returning to full health in terms of money and time is too high, then individuals may choose to seek a lower level of health than that the one they enjoyed before their injury. These decisions will vary based on personal characteristics such as age, and the same person may make different decisions at age 25 than at age 50. The question of how the value of health capital changes over time, and how people respond to risks related to their health capital is the subject of the next section.

2.2 Risk Aversion and Health Risks

2.2.1 Changes in Health Capital

One of the most important elements of the health capital model is that the value of health capital changes over time. That health tends to decline over time is a common trend for a capital asset. "Depreciation" is the common terminology for these changes in the value of real assets. Even with regular maintenance, safe behavior, and good care, a car tends to become less functional and less valuable over time. Eventually, a car will stop working no matter what is done to protect it. Similarly, health capital tends naturally to decline over time in that no one lives forever. More specifically, at some point in a person's lifetime, their health begins to decline, and they are faced with choices about how to address that decline in health.

In addition to depreciation, health can also change in a discrete fashion. It is possible for health capital to "jump" up or down in value similar to the way that the value of a financial or physical asset can suddenly become more or less valuable. Many of these sharp changes in the value of health—also known as health capital "shocks"—are negative. For example, illness and injury both result in sudden decreases in the value of a person's health capital. A person that slips on the ice, falls, and breaks an ankle is said to have suffered a negative health shock. The injured person clearly enjoys less health as a result of the injury. We can relate this to experiences in terms of financial or physical capital. The stock in an oil company may drop suddenly if the company is sued for illegal business practices. A car that is in an accident instantaneously becomes much less valuable.

It is important to note that there is no clear distinction between many types of shocks and depreciation. Instead, there is a large amount of gray area in terms of determining whether people have experienced a health shock or depreciation in their health capital. For example, a people who develop lung cancer have clearly suffered a reduction in the value of their health. The cancer grew over a relatively long period of time, so the diagnosis could be considered as either a negative health shock or the end of a long process of health depreciation. While injuries are generally negative health shocks, most illnesses and some injuries will occur over a longer period of time. The time it takes to incur a health shock relates to the insurability of such declines in health capital, as we will see in later chapters.

It is also the case that there are degrees of predictability in terms of changes in health. Certain changes in health can be predicted with a high degree of certainty. Depreciation in health, as a long-term process, is easier to foresee and plan for than health shocks. Health shocks may be more difficult to predict, especially if they are low probability events. The purpose of predicting such changes is to try to prevent them or to mitigate them, i.e., engage in investment in health. Rational individuals examining health investments subject them to an economic "cost/benefit" test: do the benefits outweigh the costs?

The challenges related to how health changes over time also relate to the time horizon associated with health capital. Many financial assets have a fixed time horizon, such as a one-year bond. Other financial and physical assets have a stochastic time horizon, meaning that the asset does not have a fixed date where it has zero value. For example, stocks do not have a fixed "end date." Physical assets, such as cars, have an expected time of use, but no fixed date when the car is worthless. Most physical assets do have a finite expected time horizon, meaning they will eventually have a value of zero. Health capital is like physical capital in this sense—eventually, health capital will be worth zero. The date when that occurs is not fixed and is, in part, within the control of the owner.

2.2.2 *Valuing Health Capital*

If health is a form of capital, then one implication is that it has a value just like any other financial or physical asset. Assets can appreciate (go up in value) or depreciate (go down in value), but once an asset reaches a value of zero, it can have no further changes in its value. The expected value of health in the future is the health capital today, minus the cost of any depreciation and negative shocks, plus the value of any appreciation and positive shocks. These changes in value then impact the overall wealth that an individual holds in his or her portfolio, meaning the total amount of wealth owned. It may be natural to think of wealth in terms of financial wealth, such as money in bank accounts, or a financial portfolio of stocks and bonds. It is also common to think of wealth in terms of physical wealth, such as a house or a car. While these forms of financial and physical wealth are important parts of a person's overall portfolio of assets, health capital can be considered as equally if not more important in terms of its overall value (Murphy and Topel 2006).

Health capital is like a stock or a bond in that it can be conceptualized as a "complex security," or a "… bundle of state contingent claims" (Huang and Litzenberger 1988). The cost of each contingent claim would be the expected payoff to that claim—meaning the probability of the event occurring times the payout. For example, the value of a ten-year bond with semiannual coupon payments could be thought of as being comprised of the twenty individual payments that make up the bond. The value of a stock that pays off quarterly dividends could be thought of as being comprised of all the individual dividend payments over time. Health also makes many small "payments" to the individual in terms of the amount of health consumed each day. Valuing these smaller aspects of overall health is at the heart of the value of life literature, which is concerned with examining the small changes in health behavior that may increase the risk of mortality and morbidity in the future while delivering additional benefits today in the form of additional pay or utility benefits.

In order to value health, we apply expected value calculations to changes in, and decisions about, the "flow" of good health that is "paid off" by health capital. The

Table 2.1 Possible flu season contingencies

State of the world	Probability	Payoff
No flu	p_n	V_n
Mild flu	p_m	V_m
Severe flu and hospitalization	p_s	V_s
Fatal flu	$P_f = 1 - p_n - p_m - p_s$	$V_f = 0$

expected value of future health is the value of each state multiplied by the probability that each state will occur. For example, one small part of a person's overall health could relate to the possibility of contracting the flu. In this example, the value of health accounting for the flu is comprised of four different possible future health contingencies, or states of the world (see Table 2.1). In each contingency, the expected value of health is the value of health in that state of the world multiplied by the probability that the specified state of the world obtains. Summing all of the possibilities leads to the expected value of health with regards to the flu. With regards to all possible diseases, summing up all of the possibilities could be considered a thought experiment that leads to an overall value of health. This calculation, which is also known as the expected value method, can only be performed to some approximation in reality as shown in Eq. 2.1.

$$E[V] = \sum_{i=\{n,m,s,f\}} p_i V_i \tag{2.1}$$

While the calculation above gives the average expected utility connected to the flu season, it is also possible to look at the variance in payoffs. Looking at the variance involves examining the number of different possible outcomes. The more outcomes, the more variability an individual would face in terms of payoffs. Looking at the variance would also involve looking at the spread between the sizes of the different payoffs. In a variance calculation, larger differences between the probability or payoff to each possible outcome (i.e., no flu, mild flu, severe flu, and death) lead to greater the variance over the entire scenario. This variance calculation can also be formalized mathematically as the variance of payoffs as shown in Eq. 2.2.

$$Var[V] = \left(\sum_{i=\{n,m,s,f\}} p_i V_i^2 \right) - E[V]^2 \tag{2.2}$$

The point of this calculation is to look at how much risk a person in this flu season scenario faces. The reason relates to how people relate to their health capital—they want both good health and certainty (predictability) about their health over the short and long term.

2.2.3 Health Risk Aversion

The health capital model captures the riskiness of health through the evolution of the value of health capital over time. Economic risk means that there is more than one possible state of the world that could obtain (occur) in the future in terms of the value of an individual's wealth (Gollier 2001). Risk is important both for the management of health capital and the intertemporal substitution problem. Risk is one of the motivations for investment in health capital because people would like to avoid both deprecation of health capital and negative health shocks. However, risk is also a barrier to investment in health because it is a reason that people may prefer current consumption over future consumption. There is risk associated with putting off the utility associated with consumption to the future because of the uncertainty of the length of life. Present consumption may be preferred over investment in health capital because future consumption is uncertain.

There is an element of risk to any capital good because of the contingent nature of assets and economic returns. The definition of an asset is that the owner of an asset will receive a good only if a certain state occurs (Mas-Colell et al. 1995, p. 699). The owner of bonds may have a fairly low risk asset, in that bonds are paid out a fixed, predefined return as long as the bond issuer is able to repay them. An asset, such as a risky stock, has many potential payoffs, each of which occurs with a much smaller probability. In that sense, a stock is a much higher risk asset. Capital goods can also be subject to risk management. Individuals can undertake costly investments to reduce their risk. They can also respond to changes in the risk of their assets by investing in different amounts of protection.

Health capital can be considered as a more or less risky asset depending upon the individual. Health capital contains a great many payoffs, since there are so many different states of health and so many ways that health can change. People have risks arising because of the health of different systems in the body, and because of the many different types of possible injuries and illnesses they can suffer. They also face a large number of choices in terms of investment in health. There is a wide array of preventative and precautionary health activities. People can also choose from a wide range of healthcare services. Therefore, economists need some way to explain how people make choices related to the management of their health capital.

Risk aversion is the economic concept designed to explain small choices about wealth. Risk aversion can be used to explain and optimize portfolio allocation choices. People who have a particular level of wealth have a wide array of choices in terms of how to invest and manage that wealth. They could invest their money in real assets, like houses and cars. They can also invest in financial assets like stocks and bonds. Certain assets may have a higher expected payoff, and certain assets may have a higher degree of risk (more variance in terms of payoffs). Risk aversion explains why one person would invest more of their wealth in stocks, whereas another person with the same amount of wealth would invest more of their money

in bonds—the latter individual is more risk averse.[3] We can use the same principles of risk aversion that apply to any choices about wealth and apply them specifically to health capital.

Risk aversion is generally measured with respect to marginal choices that individuals make that trade off the expected return of their portfolio with the riskiness of their portfolio. In the example of the flu from above, one small choice that people could make with respect to their health capital is to use the flu vaccine. The use of the flu vaccine is considered an investment because it costs money, time, and discomfort in order to deliver a future gain of reduced risk of the flu. The flu vaccine is also a marginal choice because it only changes the probability of getting the flu in the future, which is a small part of a person's overall health capital. The flu vaccine is also a good choice for measuring risk aversion because of the possibility that the expected return on the flu vaccine is negative—it could be that the cost of the flu vaccine exceeds the expected return in terms of the expected increased value of health capital. The flu vaccine might still be worthwhile because it reduces the risk associated with health capital—it reduces the number of future possible health states, thus reducing the variance in future health.

Risk aversion can be compared with risk neutrality in terms of looking at how people make investment decisions in their health capital. A risk neutral person would be willing to undertake any investment with a positive expected return. This is a common sense description of the tendency for everyone to prefer higher returns over lower returns. In the case of health capital, the expected benefit of any intervention or financial product would have to exceed the expected costs. However, the risk neutral person would be unconcerned about the distribution of outcomes. In other words, the person would be indifferent between the safe bond and the risky stock if their expected return were the same—all they care about is the expected value of the return. Similarly, in the case of a flu vaccine, a risk neutral person would get the flu shot only if the expected benefits outweighed the expected costs.

Two different people can make different choices based on their different degrees of risk aversion. A risk averse person has less of an appetite for risk than a risk neutral person. There are also degrees of risk aversion where more risk averse people have a lower appetite for risk—they prefer certainty about their health capital to investments with higher payoffs. A person with a greater appetite for risk may make the more risky choice. There are formal mathematical methods to define and measure risk aversion, which in a sense boil down to how big of a premium a person is willing to pay in order to avoid a risk. This is also important for insurance because it captures the heterogeneity in the population. Risk aversion is not one size fits all—there are degrees of difference within risk aversion. A person could be almost risk neutral, somewhat risk averse, or very risk averse.

[3]This assumes that, on average, bonds are safer than stocks.

2.2.4 Health Insurance Choices

One reason to measure risk aversion in this way is that it explains why people buy insurance. When people purchase insurance, they generally give up high returns. Insurance arrangements are often "actuarially unfair," meaning that the individual expects to pay more in premiums than they will receive in future claims. That makes insurance a low return investment (generally negative). In the canonical models of risk aversion, Arrow and Pratt actually measure the value of insurance in this way—how much are people willing to forgo in current goods in order to achieve future certainty? (Arrow 1963a; Pratt 1964) That makes health insurance distinct from many forms of health prevention, promotion, and healthcare in the sense that many of these activities are chosen because they have a positive expected return on investment. Insurance, however, does not.[4]

Risk aversion also implies a complex set of calculations in terms of the management of health capital. The risk averse person must also examine the spread of outcomes (standard deviation); the possibility that the majority of outcomes will be positive or negative (skew); and the possibility of low probability events (kurtosis). Thus, a risk averse person has a seemingly endless number of considerations of risk.[5] Applying this decision-making framework to health capital, the risk averse person will look at every decision with respect to future health in terms of both expected value calculations and the different states of the world. The risk neutral person does not have the need for such a complex set of calculations. In mathematical terms, they can translate all of these considerations into a single number, the expected value, which is a form of average (mean). Applying this decision-making framework to health capital, the risk neutral person will look at every decision with respect to future health in terms of a pure expected value calculation.

A risk averse person looking at all the risks to their health capital might like to purchase insurance against all possible outcomes. If insurance markets allowed for the purchase of a contingent claim that pays out the monetary value of contracting the flu on a particular day, then that would allow an individual to buy insurance against that contingency. Individuals could calculate their total cost of the flu in terms of the reduction in their human capital, and then buy enough contingent claims to cover the entire loss. The same individuals would likely want coverage for all of the risks that they face, and so would buy protection against the flu, an accident where they slip and fall, and any other foreseeable injury and accident. In this way, risk averse individuals would alter their portfolios to include protection against insurable events through their holdings of various microinsurance contracts.

[4]In fact, any time that a person chooses a health-related activity where the expected return is negative; they are in some sense "buying" insurance.

[5]Risk averse people are concerned with reducing variance, while reducing skew and kurtosis (prudence) are further refinements of the totality of risk aversion.

Originally, their wealth consisted of the value of their health capital, money, and other non-health assets. After the purchases, their wealth consists of their health capital, the contingent claims, and less money.

In practice, it is more efficient to buy a single insurance policy that covers all claims. The transaction costs associated with buying so many diverse contingent claims against so many risks would be staggering. In addition, it is not clear that there is any market where such claims are bought and sold. This is also known as the incomplete markets problem in finance, and in health insurance in particular (Arrow 1963b; Doherty and Schlesinger 1983). Insurance policies are generally "bundled" in that they pay for a large number of contingencies, such as all the various types of illness and injury, over a long period of time.

There are degrees of risk aversion, and these different degrees become important when considering the cost of health insurance. When insurance is actuarially fair, then risk neutral people are indifferent between their current portfolio and one that includes less money but more insurance. Conversely, risk averse people strongly prefer the portfolio that includes less money but full insurance at an actuarially fair rate. When insurance is actuarially unfair, different individuals will demand different amounts of insurance. A moderately risk averse person might prefer a portfolio with insurance against only the most catastrophic hazards, such as life-threatening illnesses. A more risk averse person would prefer a portfolio with insurance against both catastrophic and less severe hazards to health capital, even though that extra insurance would be costly in terms of the decreased expected return. The optimal scope of health insurance is the subject of Chap. 3.

There is also an in-between model, which applies under Medicare and Medicaid, in which insurers bear some risk but it is limited by government risk corridors and subsidies. Medicare and Medicaid are better described as forms of social insurance than as private insurance that individuals obtain in order to preserve their health capital. The published literature on social insurance is vast and this book considers only the literature on social health insurance. Regarding Medicare, Medicaid, and similar programs, the key considerations are how well they preserve health capital and how the cost of such insurance is shared between the individual and the government. Medicare and Medicaid may be chosen as forms of health insurance to the extent that they allow individuals to manage health capital risks. They may also be "actuarially favorable" forms of insurance in that individuals pay less in premiums than they receive in expected benefits (taxpayers pay the difference).

One other important consideration in terms of the scope of health insurance is how it defines and delineates what is and is not paid for by the insurance contract. In a traditional economic model where risk averse individuals face risks to their portfolio of wealth, they consider the total level of risk that they are willing to accept. Then, they purchase insurance that covers the entire portfolio in order to achieve their desired level of risk given the price of insurance. What consumers would really prefer is a "super insurance" policy that covers all hazards—e.g.,

health, life, homeowners—rather than individual policies for each separate risk that they face.[6]

Health insurance is particularly challenging to individuals because of the scope of its benefits. Health insurers tend to subdivide health insurance into benefit categories like "inpatient services," "outpatient services," "prescription drugs," and "mental health services," making the insurance less comprehensive from the consumer point of view—consumers want to insure their health without regard to how that health loss is recovered based on the insurer's definition of healthcare services. The main concern of individuals is to minimize the monetary and time costs of maintaining their health, as well as pain and discomfort associated with a given course of treatment. Models that assume that people can simply adjust the degree of insurance to suit their risk aversion level gloss over the fact that individuals do not simply wish to minimize their financial outlays, but rather have preferences over the type of care they receive. This book will address the scope of health insurance—what it should cover and whether it should be full or partial—as the subject of the next chapter. Before doing so, it is important to describe how health insurance actually pays for healthcare. That is the subject of the final section of this chapter.

2.3 Financial Intermediation of Healthcare Spending

2.3.1 Insuring Health Indirectly

We have seen so far that it is convenient to conceptualize of health as a form of capital and health behaviors as investments. However, directly transferring health using savings or investment is not possible. As Phelps (2003) puts it: "Some people complain sardonically that 'health insurance' is nothing of the sort, because it doesn't insure our health. Of course, this is correct, but meaningless. Our society simply does not possess the technology to insure health. We must accept the second best alternative of insurance against the financial risks associated with buying medical care" (Phelps 2003, p. 320). The point is that there is an element of this criticism that is superficial, but there is also a more substantive element to this criticism. People have a direct demand for health but an indirect demand for healthcare. Similarly, they have a direct demand for less risk but an indirect demand for health insurance.

It is common for health insurance policies to pay for care on a nonmonetary basis rather than a monetary basis. Chapter 1 examined how it became common for health insurers to pay on a "benefits" basis rather than an "indemnity" basis.

[6]In addition, people may be rightly concerned about the "border issues" between insurance policies where insurers will try to evade liability through putting it on another insurance company. If a person is in an automobile accident, the automobile insurance company and health insurance company may each claim that the other insurer is responsible for the loss, leading to uncertainty for the insured individual.

However, one problem that this raises in the context of health insurance is that individuals have, in some sense, a "doubly" indirect demand for health insurance. Individuals use health insurance to finance healthcare, and then healthcare is used to obtain health. In this form of financial intermediation, health insurance provides individuals an important service, in that it allows people to manage risks related to the value of health and the cost of healthcare.

The financial risk associated with buying medical care is the problem of exceeding one's budget. Healthcare to address a particular injury or illness may be unaffordable in the economic sense that a person could not purchase it at all. It could also exceed what economists refer to as a person's "liquidity constraints"—in theory, it might be a sizeable but affordable share of a person's lifetime income, but in practice, that person could never borrow the money in order to pay for care (Deaton 1991). It is worth noting that this is a financial risk that is related to, but somewhat distinct from, the risk related to health capital.

While the financial risk related to healthcare spending is large, it is also distributed unevenly across the population. Healthcare spending, on average, consumes a sizeable share of individual, organizational, and governmental budgets, as we have seen. This average masks enormous variation across individuals both because so much of the spending is indirect (spent by insurers on behalf of individuals), and because of the "skew" in healthcare spending whereby the majority of healthcare spending is for a small share of the population (Berk and Monheit 2001). Given that health insurance decisions are often made at the individual level, this heterogeneity could make the purchase of insurance relatively unattractive. For the risk neutral person, this possibility of catastrophe may still not be enough to get them to purchase insurance if the insurance is actuarially unfair. For the risk averse person, the financial risk associated with a large, uncertain healthcare bill is the key motivator to purchase insurance in an economic model. Then, their decision becomes dependent upon the probability and severity of large losses.

Many forms of healthcare that people purchase are fairly predictable and have a relatively low cost. Individuals may prefer to finance these low risk purchases with health insurance in order to smooth out their healthcare spending over time. Healthcare purchases that are discrete and episodic make an individual budget more difficult to manage. This is a way to use health insurance to address the intertemporal substitution problem. Contrast healthcare with the payments on a fixed-rate, 30-year mortgage. The consumption of housing through a mortgage is very smooth, in the sense that the payments are fixed over a long time horizon. Healthcare spending is often "lumpy" in that it is concentrated in specific time periods. We have previously seen that distaste for lumpy consumption is a key consequence of decreasing marginal benefit of the consumption of any good. As a result, health insurance serves a financial purpose in smoothing out consumption. This financial management is distinct from the management of larger risks provided through health insurance. Even when consumption is certain, people will have a taste for consumption smoothing over time, and health insurance could serve this function (Eeckhoudt et al. 2005, p. 92).

2.3.2 Monetizing the Value of Health

Since health insurance is designed to pay for healthcare, it is important to understand how people would value the healthcare services paid for by health insurance in the absence of such insurance. Economists often define the value of through the amount people would be willing to pay for a given form of care. "Benefit is defined as *the individual's maximum willingness to pay for the program* when supplied with information as complete as it can be, given the scientific knowledge available at the time" (Pauly 1995, p. 102). When considering the value of any healthcare intervention, the value is based on the individual's willingness to pay. The reason that they are willing to pay for a health intervention is to protect or restore their health capital. This exercise is designed to reflect, to the greatest extent possible, the actual effect on health: Given a negative health shock, what would an individual be willing to pay to avoid or reverse that shock if the direct purchase of health were possible?

The maximum willingness to pay is an important determinant of the monetary price of healthcare for two reasons. The first is that, in almost any situation, a person will end up paying a price that is less than the maximum. Thus, the economic value that a consumer gets for a health intervention is the net benefit—their maximum willingness to pay minus the amount they actually had to pay. The second reason that this is important is that the maximum willingness to pay for a health intervention for a risk averse individual will typically exceed the expected benefit because of the value that a risk averse individual places on certainty. This will be important for health insurance because the maximum willingness to pay for insurance will typically allow for insurers to charge actuarially unfair rates—that is, premiums that exceed the expected financial benefits from the insurance contract.

One way to put a financial value in these small changes to health capital brought about through the use of contingent claims is the "value of life" literature. The value of life literature determines "…how much people are willing to pay for a small reduction in the probability of their dying (Viscusi 1978)" (Phelps 2003, p. 496; Viscusi 1978). That allows an insurance company to place a financial value on health capital. The value of life approach also addresses the lack of market for health by looking at health-related choices as a substitute for such markets.

Health insurance functions by quantifying the size of health shocks and health depreciation in monetary terms. In other words, the financial intermediary role of health insurance is to take the valuations that people put on their own life and turn them into a dollar value. The economic rationale for putting health shocks in dollar terms relates to the actual goods that individuals demand, which is health capital preservation and reduction in health risks. The financial rationale is to actually implement a reimbursement system for care that achieves the purpose of health insurance, which is managing the financial risks related to the cost of healthcare. In order to help individuals preserve health capital and manage health risks, an insurer that was acting as a "perfect agent" would start with wiliness to pay as a foundation

and determine the valuation that people place on their own life.[7] Such an exercise would form the basis for the total value of health capital, and then an individual could decide how much of that value to insure given the price of insurance. However, while this is a theoretically useful way to determine the value life for health insurance, it is also highly impractical considering the great number of treatments available. As a result, insurers implement the practical way to value life, which is to measure the price of healthcare and to pay for it as long as it is within the scope of covered benefits within an insurance contract.

2.3.3 *Challenges in Monetizing Health Capital*

Using a more pragmatic value of life approach to health insurance is problematic because of the substantial heterogeneity in terms of how life is valued. "Even within a single type of study (such as the labor-market studies by Moore and Viscusi 1988a, b), wide variability exists in the estimated results" (Moore and Viscusi 1988a, b; Phelps 2003, pp. 496–497). The value of health can differ from person to person due to differing preferences in terms of the valuations placed on particular health outcomes. Recall that the overall value of health is based on summing up all of the possibilities for health over time. Each contingency has a value of the payoff that is received from each potential state of health, and the value of each healthcare good is the maximum amount that a person is willing to pay. Diversity with respect to these valuations implies different valuations of the same amount of health capital by different individuals. This difference across people helps to explain different levels of willingness to pay for health insurance due to different valuation of health capital.

Another form of diversity in terms of the valuation of risk over health outcomes is variation in risk aversion. A person who is more risk averse has a higher willingness to pay for a preventative program than a less risk averse person, all else being equal. Thus, even for two people with the same amount of health capital and the same valuation of the monetary value of health capital, one might have a greater taste for protection from health risks. This difference across people is important for explaining why people might demand different willingness to pay for health insurance due to different levels of risk aversion.

Individuals who are receiving healthcare through health insurance must also face the nonfinancial risks associated with such an arrangement. Because of how health insurance pays for healthcare to protect health, an individual must define the current, "whole," state to which insurance, in general, would help them return after harm or illness. The insurer, in a sense, defines the adequate compensation for an

[7]Principle-agent considerations are at the core of the health insurance industry, as the insurer is in some sense "representing" the patient in their dealings with providers. More on this issue can be found in the health economics and health policy literature, such as Smith et al. (1997).

illness or injury by its price list (reimbursements) for all of the forms of healthcare that it pays for under specified contingencies. For example, if a person with an unrecognized bad heart valve has open heart surgery for another heart issue, that person is not now "good as new," but rather repaired with a bad heart valve that is diagnosed but untouched. The fact that health capital can be valued but not bought or sold directly means that, to some extent, health is an irreplaceable commodity that health insurance cannot fully replace. The literature on insurance of irreplaceable commodities (e.g., Cook and Graham 1977) is extensive, but for this book, it suffices to say that such risks exist.

Basis risk is the financial term for any situation where the payouts for a financial instrument differ from the consumer's ideal. For example, if an investor wants insurance against potential changes in the value of their stock portfolio, but can only obtain a contract based on the performance of the S&P500 stock index, they accept some "basis risk" related to the difference between their portfolio and the market portfolio (Heckinger et al. 2014). In the case of health insurance, people demand health, and yet health insurance pays for healthcare. Basis risk is important within health insurance because it limits the extent to which health insurance can fully manage financial risks. Health insurance cannot protect health capital to the extent that money cannot replace health either because of technological limitations (a cure does not exist yet) or more philosophical limitations (a person who has been cured of a disease still feels in poorer health than one who never had the disease). While basis risk exists in all types of insurance—a homeowners insurance policy may not be able to truly restore a home to its condition before a fire—it is more acute in health because of the personal valuation people put on their health and the lack of a "replacement market" for human beings or human health.

While basis risk is a problem across the population, the extent of basis risk for a particular individual depends on personal characteristics, the diseases a person has, and the state of healthcare technology. In other words, the more idiosyncratic a person's health risks or health capital is, the more likely they are to face a basis risk with respect to their health. The reason is that the health insurance contract written for the "average" or "representative" member of a particular health insurance plan is likely to be a poor fit for the health risks of outliers, or individuals with unusual health risks. The same is true at the disease level—a person who faces a rare disease will have a harder time insuring this risk than a person with a common disease, all else being equal. The same is true of healthcare technology, since more advanced healthcare technology in a certain area makes it easier to define the scope of healthcare, and thus health insurance, and lowers basis risk. The reason that the magnitude of basis risk is important is that it ties directly into the demand for insurance—the greater the basis risk, the lower the demand for health insurance by an individual. Each form of diversity with respect to willingness to pay for health insurance is important because of the need for health insurers to apply a common set of benefits across a population as described in the next section.

2.3.4 Health Insurance Pricing

One pragmatic task for any health insurer is to apply a standardized methodology for monetizing health risks to individuals and populations. This methodology for determining the payoffs for healthcare under different contingencies is known the healthcare reimbursement system. The more general actuarial terminology for the design, application, and estimation of spending under such a system is the actuarial term "loss modeling." The key to the implementation of this method from the point of view of the insurer is that it must be standardized, because that allows the insurer to write insurance contracts that cover more than one person. Then, the insurer applies that method to determining the cost of insurance for individuals and populations.

Determining individual expected cost also determines the actuarially fair premium for a particular person covered by a given insurance contract. The population's expected cost is an average based on what an insurer is likely to pay out for a group of similarly situated individuals. This group of individuals grouped together for the purpose of writing insurance is also called a "risk pool." The insurer forms a risk pool so that it can calculate a premium that will allow it to pay claims, pay the cost of administering the insurance, and give the insurer a margin for error, also known as a reserve. Many insurers also budget a profit that they aim to make on a given population.

A health insurer assesses the cost of claims for insuring a population using a similar expected value calculation to the one used by individuals to value their health under different contingencies. The basic insurance pricing formula is shown in Eq. 2.3.

$$E[\text{Loss}] = E[\text{Frequency}] * E[\text{Severity}] \tag{2.3}$$

The insurer is most interested in the expected loss, also known as the claims cost. Expected loss can be assessed at the level of a population, an individual, or even a specific intervention. For example, in the U.S. population, the average spending on healthcare is equal to the expected frequency with which a person uses healthcare during the year (utilization) multiplied by the cost of care for an average healthcare encounter (price).

At an individual level, expected healthcare spending also equals the probability that a person will utilize healthcare multiplied by the expected cost of that care. The individual cost will differ from the population cost based on the particular individual characteristics. Those in better health or who are younger have a lower average cost than those who are in poorer health or older. Other variables have a "nonmonotonic" association with spending—healthcare spending is higher for women than men when women are in their childbearing years, and then "crosses over" such that healthcare spending is higher for older men than for older women of the same age.

The same calculation can be applied at the intervention level, meaning the calculation of the average spending associated with a particular healthcare service. The cost of an intervention relates to the probability that a population or a particular individual will use that intervention, and how costly it is on average. As a result, expected losses for a population can be calculated in either a "top down" or "bottom up" fashion. We could look at a model for the "total claims attributable to an individual in a year" (the "density" approach), or model the cost of each intervention and then sum that into a total cost as the "sum of a number of composing pieces" (the "build up" approach) (Bluhm 2007, pp. 129–131).

The expected loss can differ across individuals and populations due to differences in either the expected probability or the expected severity associated with health shocks or depreciation. An example of differences in probabilities of illness across populations is the flu, which is more likely to affect the very young, the elderly, and those in poor health. The flu is also an example of the possibility of different expected severities with respect to the same health shock. For example, a relatively healthy person who contracts the flu is likely to have a relatively milder case than a relatively unhealthy person who contracts the flu. The same considerations may be true overall, leading to the cost of healthcare, and thus insurance, being higher for those in poorer health both because they are more likely to get sick, and because they are more likely to have higher costs when they do get sick. Higher spending on average need not be true for every type of illness or injury—a person who is healthy may be more likely to suffer a car accident because he or she is out of the house more often than a person who living is in a long-term care facility. However, for the less healthy person, the same accident is likely to have more severe consequences.

While a health insurer estimates the full cost of healthcare, it may only pay a portion of the cost of health shocks. The proportion of costs paid depends on financial features of the health insurance contract, including any deductibles, copayments, and coinsurance. For example, in the example of the flu, a health insurer may pay only a portion of the cost of receiving the vaccine. A person who is hospitalized with the flu may also have to pay a portion of the bill. These two examples highlight the two purposes of insurance—the cost of a vaccine will be manageable for most people, making it more an issue of consumption smoothing, whereas a hospitalization may actually exceed people's ability to pay, making it more an issue of a severe cost.

These financial payments by the individual factor into the expected loss in two ways: first, they directly reduce the amount the insurer expects to pay. A 20 % coinsurance that a person is responsible for lowers the health insurer's cost by 20 %. Second, such costs indirectly lower the expected loss through the channels of probability and severity. A person who knows he or she is responsible for part of the bill may be less likely to seek care and more likely to choose less intense interventions when possible. Also known as moral hazard, this is a critical factor for determining the expected cost of claims. Chapter 7 contains more detail on these financial aspects of health insurance whereby the insurer and the individual share responsibility for the cost of care.

Health insurance has costs that go beyond the cost of claims that must be factored into the premium for health insurance. Traditional health insurers assess the cost of providing insurance as the loading factor. As has already been noted, the loading factor includes the cost of administration (expenses), the profit margin for those insurers that must make a profit, and the cost of holding a reserve or a buffer against greater-than-expected claims (Bluhm 2007, pp. 146–148). The reason that administration is costly is because estimating the cost of care for interventions, individuals, and population is costly—it requires time and effort by health insurers, large databases and computers, and other resources. Third-party payment also requires an insurer to verify a claim and to pay it on behalf of an individual. Conversely, individuals without insurance would, in a sense, have to estimate their own healthcare costs and pay those costs out of their own pocket. They would save the cost of employing someone else at a health insurer to complete these tasks on their behalf. An individual without insurance would not need a profit margin. However, they might want reserves in the form of prudent savings against healthcare spending—those savings can also smooth consumption. Individuals considering the insurance premiums would treat these loading costs as the price for insurance when deciding whether to obtain health insurance.

One other consideration in terms of health insurance pricing is that the institution administering the bill may be the same or different from the one that actually pays it. Crucially, the third-party payment function of health insurance may be taken on by more than one organization. A health insurance company may be responsible for estimating the loss for a population and administering the payment of claims while another organization, such as an employer or a government, is responsible for the cost of the claims themselves. Alternatively, a government may take responsibility for determining what care will cost, while having an insurer determine which claims are within the scope of the insurance arrangement.

This split responsibility relates both to the cost of health insurance and to how it is administered. Split responsibility could decrease the cost of insurance if different organizations have different specializations in terms of health insurance services. In other words, health insurance companies could be relatively more efficient at administering claims while employers could be relatively better at forming insurance risk pools. The converse is also possible. Split responsibility is also important administratively because it determines who makes what choices with respect to health insurance. The organization of health insurance markets in terms of how health insurance is supplied is explored in Chap. 5.

2.3.5 Scale and Scope in Third-Party Payment

In order to determine how to minimize the cost of health insurance, it is important to determine whether third-party payment contains economies of scale and scope. There are reasons to believe that strong economies of scale in insurance would reduce the cost of insurance. Health insurance is based on medical underwriting. In

theory, an insurer that can find and bring together groups of people with similar loss characteristics would reduce the risk spreading aspect of insurance. The reason for this relates to a statistical property of groups known as the law of large numbers. The law of large numbers states that, for a given probability distribution, increasing the number of draws from that distribution increases the certainty over the distribution of draws. Formally, "(the) law of large numbers, in statistics, is the theorem that, as the number of identically distributed, randomly generated variables increases, their sample mean (average) approaches their theoretical mean" (Routledge 2016).[8] For example, if one person has a 20 % chance of getting the flu, and an insurer covers only that person, then the insurer is taking on a substantial risk—if that person gets sick, the insurer will pay out much more than the expected loss, whereas if the person stays healthy, the insurer pays out much less than the expected loss. If the insurer takes on 1000 such people, then it can be much more certain that the number of people that get the flu will be either 200 or very close to 200.

The key to capturing economies of scale in insurance through risk pooling is that the 1000 people need to be more "similar" with respect to their contracting the flu. Similar is a relatively intangible concept that relates to the probable occurrence of the flu, or any health shocks. Ideally, everyone in a risk pool would have the same probability distribution overall health events. Realistically, there is some variation within any risk pool because no two people are identical. Bringing similar but not identical individuals together in a risk pool reduces the cost of insurance because the reserve per person that an insurer has for 1000 people is much less than each person would hold alone. As a result, the financial savings that people would have held as a prudent reserve can be invested in other economically valuable investments. The point here is that mixing the healthy with the unhealthy will not allow the law of large numbers to apply, and thus bringing together larger groups of heterogeneous people will not lead to economies of scale in risk management.

An insurer could also gather larger numbers of people within each risk pool in order to capitalize on economies of scope. Economy of scope differs from the economies of scale captured through the law of large numbers. Economy of scope in health insurance refers to the ability of the insurer to reduce the cost of administration of third-party payment. There are many fixed costs in insurance, such as the writing and administration of an insurance contract for a group of people and setting up a healthcare reimbursement system for making payments under the insurance contract. In this way, large insurers may be able to offer insurance more cheaply than small insurers. There may be some limit to this—beyond a certain point, more people may not bring down the cost of administration. What that limit is an empirical question, the answer to which could vary by the type of insurance and the population being insured.

[8]It is more accurate to say that there are a number of laws of large numbers including both a "weak" and "strong" form of the law. This book will treat the law of large numbers as a single law that specifies that the average of a number of random draws becomes more accurate with more and more draws under a number of specific conditions.

One challenge to economies of scope and scale in insurance is the issue of diversity across the population. An insurer could choose to offer many different types of insurance to appeal to people with different degrees of risk and risk aversion and different valuation of their health capital. Such an insurer could still attempt to capitalize on economies of scope in third-party payment by using the same underlying payment system to administer payments for each class of individuals. For example, the insurer may be able to use the same contract for a less healthy person and a person in better health, with the only difference being that the two individuals are in two different risk pools and thus pay different premiums. It may also be able to diversify its risks in terms of employing capital for reserves—it can "blend" greater-than-expected costs with one group and less-than-expected costs with another group to produce more stable financial results. However, an insurer that chooses to "slice and dice" the population in this way is forgoing potential economies of scale that it would obtain from having fewer health insurance contracts. In addition, there may be a limit to the extent to which a single reimbursement system (price list) can be used for a wide range of health insurance contracts. As a result, insurers are likely to make some trade-off between the personalization of health insurance and the price of such insurance.

While it is clear that health capital is valuable and worth insuring, it is not clear that everyone would want the same amount of protection. Indeed, the opposite is likely true given the diversity in the population we explored in this chapter. Next, we will extend beyond these general principles in order to focus on the specifics of optimal health insurance design. The question of how much insurance an individual person wants, how many risk pools there should be, and how tailored they should be, is answered by the optimal health insurance literature that we explore in the next chapter.

References

Arrow, K. J. (1963a). Liquidity preference, lecture VI. In *Lecture notes for economics 285, the economics of uncertainty* (pp. 33–53). Stanford, CA: Stanford University.

Arrow, K. J. (1963b). Uncertainty and the welfare economics of medical care. *The American Economic Review, 53*(5), 941–973.

Becker, G. S. (1967). *Human capital and the personal distribution of income: An analytical approach*. Ann Arbor, MI: Institute of Public Administration.

Berk, M. L., & Monheit, A. C. (2001). The concentration of health care expenditures, revisited. *Health Affairs (Project Hope), 20*(2), 9–18.

Bluhm, W. F. (2007). *Individual health insurance*. Winsted, CT: ACTEX Publications.

Cook, P. J., & Graham, D. A. (1977). The demand for insurance and protection: The case of irreplaceable commodities. *The Quarterly Journal of Economics, 91*(1), 143–156.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica, 59*(5), 1221–1248.

Doherty, N. A., & Schlesinger, H. (1983). Optimal insurance in incomplete markets. *The Journal of Political Economy, 91*(6), 1045–1054.

Eeckhoudt, L., Gollier, C., & Schlesinger, H. (2005). *Economic and financial decisions under risk* (1st ed.). Princeton, NJ: Princeton University Press.

Gollier, C. (2001). *The economics of risk and time* (1st ed.). Cambridge, MA: MIT Press.
Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy, 80*(2), 223–255.
Heckinger, R., Ruffini, I., & Wells, K. (2014). Over-the-counter derivatives. In R. Heckinger, D. Mengle, R. Steigerwald, I. Ruffini, & K. Wells (Eds.), *Understanding derivatives: Markets and infrastructure* (pp. 27–38). Chicago, IL: Federal Reserve Bank of Chicago.
Huang, C., & Litzenberger, R. H. (1988). *Foundations for financial economics*. Englewood Cliffs, N.J.: Prentice Hall.
Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory* (1st ed.). New York: Oxford University Press.
Moore, M. J., & Viscusi, W. K. (1988a). Doubling the estimated value of life: Results using new occupational fatality data. *Journal of Policy Analysis and Management, 7*(3), 476–490.
Moore, M. J., & Viscusi, W. K. (1988b). The quantity-adjusted value of life. *Economic Inquiry, 26* (3), 369–388.
Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. *The Journal of Political Economy, 114*(5), 871–904.
Pauly, M. V. (1995). Valuing health care benefits in money terms. In F. A. Sloan (Ed.), *Valuing health care: Costs, benefits, and effectiveness of pharmaceuticals and other medical technologies* (1st ed., pp. 99–124). Cambridge, UK: Cambridge University Press.
Phelps, C. E. (2003). *Health economics* (3rd ed.). Boston: Addison-Wesley.
Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica, 32*(1/2), 122–136.
Routledge, R. (2016). Law of large numbers. *Encyclopædia Britannica*. Retrieved from http://www.britannica.com/science/law-of-large-numbers
Smith, P. C., Stepan, A., Valdmanis, V., & Verheyen, P. (1997). Principal-agent problems in health care systems: An international perspective. *Health Policy, 41*(1), 37–60.
Varian, H. R. (1992). *Microeconomic analysis* (3rd ed.). New York: W.W. Norton and Company.
Viscusi, W. K. (1978). Wealth effects and earnings premiums for job hazards. *The Review of Economics and Statistics, 60*(3), 408–416.

http://www.springer.com/978-3-319-43795-8

What Is Health Insurance (Good) For?
An Examination of Who Gets It, Who Pays for It, and How to Improve It
Lieberthal, R.
2016, XIII, 271 p. 4 illus., Hardcover
ISBN: 978-3-319-43795-8