In this chapter, we discuss a fundamental problem in health insurance. We have seen how a consumer’s utility is affected by income fluctuations. Health insurance is a policy to guard against income fluctuations due to illnesses.

1 A Model of Illness and Treatment

1.1 Verifiable Health Status and Insurance

Illness is modeled by the uncertainty of losses due to illness. A consumer may be sick, and his illness may differ in severities. We will let the loss due to illness be measured in terms of money. For a simple analysis, we let there be six illness states: these states are like those that we have studied in the chapter on risks and preferences. These states are labelled \( \ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \) and \( \ell_6 \) and we will use the following numbers for these states:

\[
\begin{array}{cccccc}
\ell_1 & \ell_2 & \ell_3 & \ell_4 & \ell_5 & \ell_6 \\
0 & 30 & 60 & 90 & 120 & 150 \\
\end{array}
\]

The interpretation is that \( \ell_i, i = 1, 2, ..., 6 \) describes the income loss. We assume that each state occurs with probability \( \frac{1}{6} \).

An illness with a higher severity level leads to more income loss. Let \( W \) be the consumer’s wealth or income, and the strictly concave function \( U \) the risk-averse consumer’s utility function. In state \( i \), the consumer’s utility is \( U(W - \ell_i) \). Ex ante, before the illness happens, the consumer does not know the severity. Ex post, he learns that, and the income loss occurs. The consumer’s expected utility, ex ante, is

\[
\begin{align*}
E[U] &= \frac{1}{6} U(W - \ell_1) + \frac{1}{6} U(W - \ell_2) + \frac{1}{6} U(W - \ell_3) + \frac{1}{6} U(W - \ell_4) + \frac{1}{6} U(W - \ell_5) + \frac{1}{6} U(W - \ell_6) \\
&= \frac{1}{6} U(W - 0) + \frac{1}{6} U(W - 30) + \frac{1}{6} U(W - 60) + \frac{1}{6} U(W - 90) + \frac{1}{6} U(W - 120) + \frac{1}{6} U(W - 150).
\end{align*}
\]

Our discussion so far uses the general framework for risks and preferences. Clearly, if health states are verifiable, an insurance contract can be written for the consumers; the insurance contract will specify payments contingent on the health states \( \ell_i \)’s. The expected loss, the average of the \( \ell \)'s is simply the sum of their values multiplied by the probability \( 1/6 \):

\[
\text{Expected loss} = \frac{1}{6} [0 + 30 + 60 + 90 + 120 + 150] = 75.
\]
An insurance company, operating in a competitive environment, can offer the following contract: for a premium of 75, the consumer will be indemnified for all the losses, so that if $\ell_i$ occurs, the insurer will reimburse that loss. The consumer’s risk, due to fluctuations of income from illness will be eliminated. The premium is fair—it is equal to the expected loss—and the consumer has full insurance, so this is the best that can happen to the consumer. This kind of insurance of course critically relies on the assumption that the health losses are verifiable and hence contractible.

For now we continue with this assumption of verifiable health states. We next introduce a medical treatment option. This treatment costs 75, and it eliminates any illness loss. Clearly, this is a stylised assumption about medical treatment but it does capture the fact that medical treatment is costly, but can help alleviate the loss due to illness. In any case, the results are robust in many ways.

Facing with such an option, the consumer will reassess his situation. He knows that if he gets sick, there is always the option of spending 75 to get rid of the loss. Clearly, in states $\ell_4, \ell_5,$ and $\ell_6$ where the loss is more than 75, it is best to use the treatment. Conversely, in states $\ell_1, \ell_2,$ and $\ell_3$ where the loss is less than 75, it is best to forgo the treatment. Adopting the optimal treatment decision, the consumer’s expected utility is now

$$\frac{1}{6}U(W - 0) + \frac{1}{6}U(W - 30) + \frac{1}{6}U(W - 60) + \frac{1}{6}U(W - 75) + \frac{1}{6}U(W - 75) + \frac{1}{6}U(W - 75).$$

Medical treatment reduces income fluctuations. Whereas before the consumer’s income loss ranges between 0 and 150, now it is between 0 and 75. The medical treatment makes the consumer better off.

Still the consumer faces income fluctuation, and has to bear some risks, those income variations when the treatment is not used. An insurance contract can improve on the situation. Remember that the consumer is risk averse, so any attempt to remove this income fluctuation is better for him. Furthermore, the availability of the medical treatment makes things easier. Previously, an insurance company operating in a competitive environment must charge a premium of 75 and indemnify all income losses. Now, with the medical treatment, an insurer can charge a premium, indemnify income losses in states $\ell_1, \ell_2,$ and $\ell_3,$ and then specify that treatment be used in states $\ell_4, \ell_5,$ and $\ell_6,$ but there will not be any payment to the consumer. This is Ideal Insurance.
Here, the treatment will be used when it is efficient, exactly the same as what the consumer would have chosen to do if he was bearing treatment costs. The consumer does not have to pay anything ex post; health states are verifiable, and the insurance contract simply specifies the treatment, instead of indemnify the losses in states $\ell_5$, and $\ell_6$. Now in states $\ell_1$, $\ell_2$, and $\ell_3$, the consumer does not receive treatment; again this is what he would have chosen to do if he was bearing treatment costs. His loss in these states will be reimbursed. The premium under Ideal Insurance is set at the expected cost. The insurer has to reimburse losses $\ell_1$, $\ell_2$, and $\ell_3$, as well as pay for treatment in states $\ell_4$, $\ell_5$, and $\ell_6$. The premium therefore is

$$\frac{1}{6} [0 + 30 + 60 + 75 + 75 + 75] = 52.50$$

Again the medical treatment makes it less costly to be sick! The premium has dropped from 75 to 52.50, the consumer is fully insured, and treatment will be used if and only if it is efficient. The insurance option has improved the consumer’s welfare.

1.2 Nonverifiable Health Status and Moral Hazard

The Ideal Insurance situation is the best. It relies critically on the health states being verifiable so that treatment decisions as well as monetary payments can be based on them. The fundamental problem in the health market is simply that health status is difficult to verify. First, in many situations, whether one is sick or not does not yield any easily detectable physical changes. We are all aware that headaches, back pains, and all sorts of physical discomfort do not always exhibit any visible signs. For our model, health states $\ell_i$ and $\ell_j$ may be physically indistinguishable. Second, even if one shows signs of being sick, the illness can exhibit various severity levels. It will be impossible to tell how sick one is. For our model, even if we can ascertain that someone is sick (so that state $\ell_1$ can be ruled out), we may not tell whether it is the loss is going to be high or low.

Lacking verifiable information, an insurer offering the contract under Ideal Insurance would face a problem. A patient, now completely indemnified for either loss or treatment cost, would always claim to be more seriously sick than he truly was. The insurance therefore would not be able to break even. As a result, the contract which offers full indemnity of loss and treatment cost will not be viable.
If nothing is verifiable, the insurance market will breakdown. The consumer will simply have to operate on his own, unable to smooth any of his income fluctuations through an insurance policy. The marketplace, however, comes up with a remedy, not perfect, but can solve some of the problems due to health status being nonverifiable. The key idea is that medical treatment itself is verifiable. A medical treatment is a readily observable physical event. It is afterall a form of transaction.

An insurance policy can be based on medical treatment. There can be a payment from the insurer to the consumer if some medical treatment is used. In other words, insurance can be in the form of subsidizing the cost of medical treatment. We can illustrate this by the following policy. In return for a subsidy of 10 whenever the medical treatment is used, the consumer pays a premium of 5. Now, under this subsidy, the cost of treatment is actually 65. The consumer retains the ability to make treatment decisions. Under this subsidy, the consumer continues to opt for treatment under states $\ell_4$, $\ell_5$, and $\ell_6$; in all these states, the loss is more than 65. In the other states, the loss is lower than 65, so the consumer does not use the treatment. The subsidy of 10 is paid in three states, so the total expected subsidy is $(10 + 10 + 10)/6 = 5$, which is the premium we have specified. The contract breaks even even for the insurer.

Now the consumer’s expected utility becomes

$$
\frac{1}{6}U(W - 5) + \frac{1}{6}U(W - 5 - 30) + \frac{1}{6}U(W - 5 - 60)
+ \frac{1}{6}U(W - 5 - 65) + \frac{1}{6}U(W - 5 - 65) + \frac{1}{6}U(W - 5 - 65)
= \frac{1}{6}U(W - 5) + \frac{1}{6}U(W - 35) + \frac{1}{6}U(W - 65) + \frac{1}{6}U(W - 70) + \frac{1}{6}U(W - 70) + \frac{1}{6}U(W - 70).
$$

The fluctuation of income now has a range between 5 and 70 under the insurance-subsidy contract. If there was no insurance contract, the range of fluctuation is between 0 and 75. Subsidizing 10 of the 75 treatment cost reduces the consumer’s risks. The consumer is allowed to trade between incomes when treatment is not used and incomes when treatment is used.

The consumer makes the same efficient treatment decisions whether the subsidy of 10 is available or not. This is not always the case. A subsidy of 20 certainly will make the consumer behave differently. In this case, the consumer’s net treatment cost becomes 55, down from 75. Here, the consumer will optimally choose to have treatment even in state $\ell_3$. This change in behavior due to subsidy is called “moral hazard.”
Under this premium, the subsidy of 20 will be paid in states \( \ell_3, \ell_4, \ell_5, \) and \( \ell_6 \). The total expected subsidy is \((20 + 20 + 20 + 20)/6 = 13\frac{1}{3} \), so the premium under this higher subsidy insurance contract will be \( 13\frac{1}{3} \).

The consumer’s expected utility is now

\[
\frac{1}{6} U(W - 13\frac{1}{3}) + \frac{1}{6} U(W - 13\frac{1}{3} - 30) + \frac{1}{6} U(W - 13\frac{1}{3} - 55) \\
+ \frac{1}{6} U(W - 13\frac{1}{3} - 55) + \frac{1}{6} U(W - 13\frac{1}{3} - 55) + \frac{1}{6} U(W - 13\frac{1}{3} - 55)
\]

\[
= \frac{1}{6} U(W - 13\frac{1}{3}) + \frac{1}{6} U(W - 43\frac{1}{3}) + \frac{1}{6} U(W - 68\frac{1}{3}) + \frac{1}{6} U(W - 68\frac{1}{3}) + \frac{1}{6} U(W - 68\frac{1}{3}) + \frac{1}{6} U(W - 68\frac{1}{3}).
\]

A higher subsidy of 20 further reduces the range of income fluctuation, which is now between \( 13\frac{1}{3} \) and \( 68\frac{1}{3} \), instead of the 5 and 70 range under the lower subsidy of 10. The key difference between the higher subsidy and the lower subsidy is moral hazard. Under the lower subsidy of 10, the consumer’s treatment decisions are all efficient. Under the higher subsidy of 20, the consumer treatment decision is inefficient at \( \ell_3 \).

In our discussions, we have two possibilities of offering partial insurance to the consumer. The low subsidy of 10 does not lead to inefficient treatment, while the high subsidy of 20 does. None of these subsidies can implement Ideal Insurance. The reason is that the consumer must still face income risks when no treatment is used. The insurance policy is written on partially covering treatment cost. When treatment is not used, as in states \( \ell_1 \) and \( \ell_2 \), there are still losses and income variations, but these cannot be insured—because health states are unverifiable.

We have illustrated the moral hazard versus risk sharing tradeoff. Medical expenses are uncertain, and impose financial risks on the consumer. The lower subsidy does not lead to inefficient treatment decisions at the expense of the consumer facing more income fluctuation. The higher subsidy does lead to inefficient decisions, but giving the benefit of the consumer facing less income fluctuation. Generally, it is seldom optimal to have medical treatment decisions always efficient. There is almost always some gain in getting the treatment decision to be a little inefficient. The inefficiency allows an improvement in risk sharing. (The formal, mathematica argument is that at the point of zero subsidy, raising the subsidy from zero induces a second-order loss due to inefficient treatment, but a first-order gain due to better risk sharing.) This is a fundamental point. Health insurance necessarily lead to inefficient treatments ex post; the benefit is a better risk sharing ex ante.
Sometimes moral hazard may be the only way to help with risk sharing. The following, somewhat extreme, example illustrates this. Suppose that there are only two possible losses, namely 0 and 100, and each happens with probability \( \frac{1}{2} \). There is a medical treatment that costs 110 for this, and again, this treatment eliminates the loss. Now without any insurance, the consumer does not use any treatment, and his expected utility is \( \frac{1}{2}U(W) + \frac{1}{2}U(W - 100) \). Now, suppose that an insurance policy covers the entire treatment cost. The consumer will use the treatment when the loss is 100. The premium therefore is \( 110 \times \frac{1}{2} = 55 \). Under this insurance policy, the consumer does not face any risk, but his utility is now \( U(W - 55) \). When \( U(W - 55) > \frac{1}{2}U(W) + \frac{1}{2}U(W - 100) \), then covering the cost of the inefficient treatment is the one and only way to reduce the consumer’s financial risks. The condition \( U(W - 55) > \frac{1}{2}U(W) + \frac{1}{2}U(W - 100) \) is the same as the consumer’s risk premium being at least 5. In other words, if the consumer is very risk averse, insurance on medical expenses, despite leading to inefficiency, can improve the consumer’s expected utility.