1. Risk and uncertainty

It is hard to think of decisions where the outcome can be predicted with certainty. For example, the decision to buy a house involves several elements of uncertainty: Will house prices increase or decrease in the near future? Will the house require expensive repairs? Will my job be stable enough that I will be able to live in this area for a sufficiently long time? Will I have a good relationship with the neighbors? And so on.

Whenever the outcome of a decision involves future states of the world, uncertainty is unavoidable. Thus one source of uncertainty lies in our inability to predict the future: at most we can formulate educated guesses. The price of a commodity one year from now is an example of this type of uncertainty: the relevant facts are not settled yet and thus cannot be known. Another type of uncertainty concerns facts whose truth is already settled but unknown to us. An example of this is the uncertainty whether a second-hand car we are considering buying was involved in a serious accident in the past. The seller is likely to know, but it will be in his interest to hide or misrepresent the truth.

In his seminal book *Risk, uncertainty, and profit*, first published in 1921, Frank Knight established the distinction between situations involving risk and situations involving uncertainty. A common factor in both is the ability, in general, to list (at least some of) the possible outcomes associated with a particular decision. What distinguishes them is that in the case of risk one can associate “objective” probabilities to the possible outcomes, while in the case of uncertainty one cannot. An example of a decision involving risk is the decision of an insurance company to insure the owner of a car against theft. The insurance company can make use of statistical
information about past car thefts in the particular area in which the customer lives to calculate
the probability that the customer’s car will be stolen within the period of time specified in the
contract. An example of a decision involving uncertainty is the decision of an airline to purchase
fuel on futures markets. Since the future spot price of fuel will be affected by a variety of hard-
to-predict factors (such as the political situation in the Middle East, the future demand for oil,
etc.) it is impossible to assign an objective probability to the various future spot prices. In a
situation of uncertainty the decision maker might still rely on probabilistic estimates, but such
probabilities are called subjective since they are merely an expression of that particular decision
maker’s beliefs.

2. Attitudes to risk and insurance

Suppose that a decision-maker has to choose one among several available actions \( a_1, \ldots, a_r \) \((r \geq 2)\). For example, the decision-maker could be a driver who has to decide whether to
remain uninsured (action \( a_1 \)) or purchase a particular collision-insurance contract (action \( a_2 \)).
With every action the decision-maker can associate a list of possible outcomes with
 corresponding probabilities (which could be objective or subjective probabilities). If the possible
outcomes are denoted by \( x_1, \ldots, x_n \) and the corresponding probabilities by \( q_1, \ldots, q_n \) (thus, \( q_i \geq 0, \)
for every \( i = 1, \ldots, n \), and \( q_1 + \ldots + q_n = 1 \)), the list
\[
\begin{pmatrix}
  x_1 & \cdots & x_n \\
  q_1 & & q_n
\end{pmatrix}
\]
\( \) is called the lottery
corresponding to that action. Choosing among actions with uncertain outcomes can thus be
viewed as choosing among lotteries. When the possible outcomes are numbers, typically
representing sums of money, we will denote them by \( m_1, \ldots, m_n \) and call the lottery a money-
lottery. With a money-lottery \( L = \begin{pmatrix} m_1 & \cdots & m_n \\ q_1 & & q_n \end{pmatrix} \) one can associate the number
\( EV_L = q_1 m_1 + \ldots + q_n m_n \), called the expected value of \( L \). For example, if \( L = \begin{pmatrix} 5 & 2 & 4 \\ \frac{2}{8} & \frac{5}{8} & \frac{1}{8} \end{pmatrix} \) then

\[
EV_L = \frac{2}{8} \cdot 5 + \frac{5}{8} \cdot 2 + \frac{1}{8} \cdot 4 = 3. \]

An individual is defined to be risk-averse if, when given a choice between a money-lottery \( L = \begin{pmatrix} m_1 & \ldots & m_n \\ q_1 & \ldots & q_n \end{pmatrix} \) and its expected value for sure [that is, the (trivial) lottery \( \begin{pmatrix} EV_L \\ 1 \end{pmatrix} \)], she would strictly prefer the latter. If the individual is indifferent between the lottery and its expected value, she is said to be risk-neutral and if she prefers the lottery to the expected value she is said to be risk-loving. For example, given a choice between being given $50 for sure and tossing a fair coin and being given $100 if the coin lands Heads and nothing if the coin lands Tails, a risk-averse person would choose $50 for sure, a risk-loving person would choose to toss the coin and a risk-neutral person would be indifferent between the two options.

Most individuals display risk aversion when faced with important decisions, that is, decisions that involve substantial sums of money. We now show that risk-aversion is what makes insurance markets profitable.

---

\(^2\) The expected value of lottery \( L \) is the amount of money that one would get on average if one were to play the lottery a large number of times. For example, consider the lottery \( L = \begin{pmatrix} \$5 & \$2 & \$4 \\ \frac{2}{8} & \frac{5}{8} & \frac{1}{8} \end{pmatrix} \) whose expected value is $3. Suppose that this lottery is played \( N \) times. Let \( N_5 \) be the number of times that the outcome of the lottery turns out to be $5 and similarly for \( N_2 \) and \( N_4 \) (thus \( N_5 + N_2 + N_4 = N \)). By the Law of Large Numbers in probability theory, if \( N \) is large then the frequency of the outcome $5, that is, the ratio \( \frac{N_5}{N} \) will be approximately equal to the probability of that outcome, namely \( \frac{2}{8} \), and, similarly, \( \frac{N_2}{N} \) will be approximately equal to \( \frac{5}{8} \) and \( \frac{N_4}{N} \) will be approximately equal to \( \frac{1}{8} \). The total amount the individual will get is \( 5 N_5 + 2 N_2 + 4 N_4 \) and the average amount (that is, the amount per trial) will be \( \frac{5 N_5 + 2 N_2 + 4 N_4}{N} = \frac{N_5}{N} \cdot 5 + \frac{N_2}{N} \cdot 2 + \frac{N_4}{N} \cdot 4 \) which is approximately equal to \( \frac{2}{8} \cdot 5 + \frac{5}{8} \cdot 2 + \frac{1}{8} \cdot 4 = 3 \), the expected value of \( L \).
We shall consider the simple case where the insurance industry is a monopoly (that is, it consists of a single firm) and all individuals are identical, in the sense that they have the same initial wealth (denoted by $W$) and face the same potential loss (denoted by $\ell$) and the same probability of loss (denoted by $q$). Suppose that the potential loss $\ell$ (with $0 < \ell < W$) is incurred if there is a fire. Thus there are two possible future “states of the world”: the good state, where there is no fire, and the bad state, where there is a fire. Suppose that the probability that there will be a fire within the period under consideration (say, a year) is $q$ (with $0 < q < 1$). Each individual has the option of remaining uninsured, which corresponds to the lottery $\left( \frac{W}{1-q}, \frac{W-\ell}{q} \right)$.

Insurance contracts are normally specified in terms of two quantities: the premium (denoted by $p$) and the deductible (denoted by $d$). The premium is the price of the contract, that is, the amount of money that the insured person pays to the insurance company, irrespective of whether there is a fire or not. If a fire does not occur, then the insured receives no payment from the insurance company. If there is a fire then the insurance company reimburses the insured for an amount equal to the loss minus the deductible, that is, the insured receives a payment from the insurance company in the amount of $\ell - d$. Thus the decision to purchase contract $(p,d)$ corresponds to the lottery $\left( \frac{W-p}{1-q}, \frac{W-p-d}{q} \right)$. If $d = 0$ the contract is called a full-insurance contract, while if $d > 0$ the contract is called a partial-insurance contract. Let $L_{NI}$ denote the lottery corresponding to the decision not to insure ($NI$ stands for ‘No Insurance’); thus $L_{NI} = \left( \frac{W}{1-q}, \frac{W-\ell}{q} \right)$. The expected value of this lottery is $W - q\ell$, that is, initial wealth minus expected loss. Given our assumption that the individual is risk-averse, she will prefer $W - q\ell$ for sure to the lottery $L_{NI}$, that is, she would be better off (relative to not insuring) if she purchased a full-insurance contract with premium $p = q\ell$. Since such a contract makes her strictly better off, she will also be willing to buy a full-insurance contract with a slightly larger premium $p > q\ell$. If the insurance company
sells a large number of such contracts, its average profit, that is, its profit per contract will be \( \bar{p} - q\ell > 0 \). Hence the sale of insurance contracts would yield positive profits.

Would a profit-maximizing monopolist want to offer a full insurance contract or a partial insurance contract to its customers? A simple argument shows that the monopolist would in fact want to offer full insurance. Consider any partial insurance contract \((p_0, d_0)\) with premium \(p_0\) and deductible \(d_0 > 0\). Denote by \(\pi_0\) the average profit per customer given this contract. Then \(\pi_0 = p_0 - q(\ell - d_0) = p_0 + qd_0 - q\ell\). Consider the alternative full-insurance contract with premium \(\hat{p} = p_0 + qd_0\). The average profit per customer from the sale of this contract would be \(\hat{\pi} = \hat{p} - q\ell = p_0 + qd_0 - q\ell\). Thus \(\hat{\pi} = \pi_0\), so that the insurance company is indifferent between these two contracts. The customers, however, would strictly prefer the full-insurance contract with premium \(\hat{p} = p_0 + qd_0\). In fact, purchasing contract \((p_0, d_0)\) can be viewed as playing the lottery \(\left( \begin{array}{cc} W - p_0 & W - p_0 - d_0 \\ 1 - q & q \end{array} \right)\) whose expected value is \(W - p_0 - qd_0 = W - \hat{p}\); the full insurance contract guarantees this amount for sure and thus, by the assumed risk-aversion of the customer, makes her strictly better off. Hence the customer would be willing to purchase a full-insurance contract with a slightly higher premium \(\hat{p} > \hat{p}\); such a contract would yield an average profit of \(\hat{\pi} = \hat{p} - q\ell > \hat{p} - q\ell = \hat{\pi} = \pi_0\). Hence contract \((p_0, d_0)\) cannot be profit-maximizing.

In the above analysis it was assumed that the probability of loss \(q\) remained the same, no matter whether the individual was insured or not and no matter what insurance contract she bought. It is often the case, however, that the individual’s behavior has an effect on the chances that a loss will occur, in which case a situation of moral hazard is said to arise. Moral hazard refers to situations where the individual, by exerting some effort or incurring some expenses, has

\[\ldots\]

\[\begin{array}{c}
\text{3 By the argument of the previous note, the fraction of insured customers who would suffer a loss and submit a reimbursement claim would be approximately } q \text{ and thus the total profit would be approximately } \bar{p}N - qN\ell \text{ (where } N \text{ is the number of contracts sold) so that the profit per customer would be } \\
\frac{\bar{p}N - qN\ell}{N} = \bar{p} - q\ell.
\end{array}\]
some control over either the probability or magnitude of the loss, but these preventive measure are not observed by the insurer and hence the premium cannot be made a function of them. For example, the chances that a bicycle is stolen are lower if the owner is very careful and always locks the bicycle when she leaves it unattended. If the bicycle is not insured, the owner might be more conscientious about locking it, while if it is covered by full insurance she might at times not bother to lock it (after all, if the bicycle is stolen the insurance company will pay for a replacement). When the individual stands to lose if the loss occurs (e.g. if she is uninsured or if she incurs a high deductible) then she will have an incentive to try to reduce the chances of a loss. Hence in situations where moral hazard is present, the insurance company will prefer to offer partial insurance rather than full insurance.

So far we have focused on the case where the two parties to the insurance contract have the same information. Often, however, potential customers have more information than the insurance company, in which case we say that information is asymmetric. We now turn to the issues that arise when there is asymmetric information.

3. Asymmetric information

The expression ‘asymmetric information’ refers to situations where two parties to a potential transaction do not have the same information; in particular, one of the two parties has valuable information that is not available to the other party. Examples abound. The owner of a used durable good has had enough experience through use to know the true quality of the good he wants to sell; the potential buyer, on the other hand, cannot help but wondering if the seller is merely trying to get rid of a low-quality item he regretted buying. A loan applicant knows whether her intentions are to do her best to repay the loan and what the chances are that she will be able to repay it; the lender, on the other hand, will worry about the possibility that the borrower will simply “take the money and run”. The owner of a house has more information than the prospective buyer about matters that are important to the latter, such as the quality of the house (e.g. how many repairs were needed in the past), the neighborhood (e.g. whether the neighbors are noisy), the upkeep of the house, etc.

In such situations the uninformed party cannot simply rely on verbal assurances by the informed party, since the latter will have an incentive to lie or misrepresent the truth: after all,
talk is cheap! Thus the uninformed party will need to try to infer the relevant information from the actions of the other party or from other observable characteristics. This often leads to market failures where society gets stuck in a *Pareto inefficient* situation. A situation $X$ is defined to be *Pareto inefficient* if there is an alternative situation $Y$ which is feasible and such that everybody is at least as well off in situation $Y$ as in situation $X$ and some individuals strictly prefer $Y$ to $X$.\(^1\)

In the next two sections we discuss two important phenomena associated with asymmetric information: adverse selection and signaling. Both phenomena can give rise to Pareto inefficiencies. The seminal work in this area was laid out most notably by three economists, George Akerlof, Michael Spence and Joseph Stiglitz who shared the 2001 Nobel Memorial Prize in Economic Sciences “for their analyses of markets with asymmetric information”.

### 4. Adverse selection

George Akerlof’s seminal paper (1970) pointed out what is now known as the phenomenon of *adverse selection* (also called *hidden information*). Akerlof considers markets where buyers are unable to ascertain the quality of the good they intend to purchase, while sellers know the quality. He shows that this asymmetry of information may lead to the breakdown of the market. He illustrates this possibility by focusing on the market for used cars where buyers’ inability to determine the quality of the car they are considering buying makes them worried that they might end up with a ‘lemon’ (the American term for ‘bad car’). Since buyers cannot distinguish a good car from a bad car, all cars must sell at the same price. This fact would not create a problem if the average quality of cars in the market were given exogenously. However, because of the sellers’ knowledge, the average quality will in fact depend on the market price. The lower the price, the smaller the number of cars offered for sale and the lower the average quality. Realizing this, buyers will be willing to pay lower and lower prices, leading to a situation where only the lowest-quality cars are offered for sale: the bad quality cars drive the good quality cars out of the market. We will illustrate this with a simple example.

Suppose that there are two groups of individuals: the owners of cars and the potential buyers. The quality of each car is known to the seller (he used the car for a sufficiently long
time) but cannot be ascertained by a potential buyer (a buyer will discover the true quality of a
car only after owning it for a while). Denote the possible qualities by $A, B, \ldots F$ where $A$
represents the best quality, $B$ represents the second best quality and so on (thus $F$ is the lowest
quality). Quality could be measured in terms of durability or fuel efficiency or other
characteristics that all consumers rank in the same way. Suppose that, for each quality level, the
Corresponding car is valued less by its owner than by a potential buyer. Thus, in the absence of
asymmetric information, all cars would be traded (assuming a sufficiently large number of
potential buyers). Suppose also that some general information is available to everybody (e.g.
through consumer magazines) giving, for each quality, the proportion of all cars produced that
are of that quality. All this is shown in the following table, which we take to be common
knowledge among sellers and potential buyers.

| Quality | $A$ | $B$ | $C$ | $D$ | $E$ | $F$
|---------|----|----|----|----|----|----|
| Value to potential buyers | $6,000$ | $5,000$ | $4,000$ | $3,000$ | $2,000$ | $1,000$
| Value to current owners | $5,400$ | $4,500$ | $3,600$ | $2,700$ | $1,800$ | $900$
| Proportion | $p_A = \frac{1}{12}$ | $p_B = \frac{2}{12}$ | $p_C = \frac{1}{12}$ | $p_D = \frac{4}{12}$ | $p_E = \frac{3}{12}$ | $p_F = \frac{1}{12}$

**Table 4.1**

Information which is common knowledge among sellers and potential buyers.

According to Table 4.1, for each possible quality, the seller’s valuation of the car is 10% less
than a potential buyer’s valuation. Since buyers cannot determine the quality of any particular
car prior to purchase, all cars must sell for the same price, denoted it by $P$. For what values of $P$
can there be trade in this market? Suppose that all potential buyers are risk-neutral, so that they
view a money lottery as equivalent to its expected value. A potential buyer might reason as
follows: “buying a car can be viewed as playing the following lottery
whose expected value is \[ \frac{1}{12}(6,000 - P) + \frac{2}{12}(5,000 - P) + \ldots + \frac{1}{12}(1,000 - P) = 3,250 - P; \] thus, as long as \( P < 3,250 \) I would gain from buying a second-hand car.” This reasoning, however, is naïve in that it assumes that the average quality of the cars offered for sale is independent of \( P \). A sophisticated buyer, on the other hand would realize that if, say, \( P = 3,100 \) then buying a car would not yield an expected gain of $150 (= 3,250 – 3,100), because only the owners of cars of qualities \( D, E \) and \( F \) would be willing to sell at that price: the higher qualities \( A, B \) and \( C \) would not be offered for sale. Thus buying a car at price \( P = 3,100 \) would really correspond to playing the following lottery:\footnote{The probabilities are obtained by conditioning on the set of qualities \{\( D,E,F \). While the probability of picking a car of quality \( D \) from the entire population of cars is \( \frac{4}{12} \), the probability of picking a car of quality \( D \) from the subpopulation containing only cars of qualities \( D, E \) and \( F \) is obtained, by using what is known as Bayes’ rule, as the ratio \( \frac{P_D}{P_D + P_E + P_F} = \frac{\frac{4}{12}}{\frac{1}{12} + \frac{1}{12} + \frac{1}{12}} = \frac{4}{8} \). This becomes clear if one considers the case where the number of cars of each quality is as follows: \( \begin{pmatrix} A & B & C & D & E & F \\ 1 & 2 & 1 & 4 & 3 & 1 \end{pmatrix} \). Then within the subpopulation of cars of qualities \( D, E \) and \( F \) (a total of 8) there are 4 that are of quality \( D \), so that the fraction of cars of quality \( D \) among those of qualities \( D, E \) and \( F \) is \( \frac{4}{8} \). The conditional probabilities for the cars of qualities \( E \) and \( F \) are obtained similarly.} \[
\begin{pmatrix}
 3,000 - 3,100 & 2,000 - 3,100 & 1,000 - 3,100 \\
 4 & 3 & 1 \\
 8 & 8 & 8
\end{pmatrix}
\]
which has an expected value of \(- 725 (= 2,375 – 3,100)\); thus buying a car at price \( P = 3,100 \) is equivalent to losing $725! Should then the buyer be willing to pay some price lower than $2,375, say, \( P = 2,100 \)? The answer is No! When the price is $2,100, only cars of qualities \( E \)
and $F$ are offered for sale and thus a buyer faces the lottery \[
\begin{pmatrix}
2,000 & 1,000 \\
3 & 1 \\
4 & 4
\end{pmatrix}
\]
whose expected value $-350 = (1,750 - 2,100)$, so that buying a car at price $P = $2,100 is equivalent to losing $350. Trading in the market is only possible at some price $P$ between $900$ and $1,000$, in which case only cars of quality $F$ are traded. This is an extremely inefficient outcome, as compared to what would happen if both buyers and sellers had the same information: in such a case there would be a different price for each quality and all cars would be traded.\(^2\)

Second-hand durable goods, such as cars, provide just one example of markets where the phenomenon of adverse selection occurs. Another example is health insurance. Individuals differ in their likelihood of needing medical care; for example, some are genetically predisposed to certain diseases while others are not, some lead healthier lifestyles than others (e.g. make better nutritional choices and/or exercise more often), some are more prone to engage in risky and dangerous activities while others are more cautious and less likely to meet with accidents, etc. If the relevant characteristics are known to the potential customer but not observable by the provider of health insurance, we have a situation of asymmetric information. Demand for insurance will be a decreasing function of the insurance premium, but high-risk individuals will be willing to pay higher premia than low-risk individuals. Thus an increase in the insurance premium will adversely affect not only the size but also the composition of the pool of applicants: a larger proportion will consist of high-risk individuals who are more costly to insure, since they are more likely to submit claims. If health costs increase, the insurance company will need to increase the premium in order to cover its costs, but a premium increase will lead to low-cost, low-risk individuals dropping out of the market and thus to a worse pool of customers and higher costs for the insurance provider. This phenomenon of adverse selection can lead to spiraling increases in costs and potentially to market failure.

As shown by Stiglitz and Weiss (1981), adverse selection can also explain credit rationing, that is, the situation where lenders limit the supply of additional credit to borrowers who demand funds, even if the latter are willing to pay higher interest rates. In other words, at the prevailing market interest rate, demand exceeds supply but lenders are not willing to lend more funds and do not find it profitable to raise the interest rate that they charge. Stiglitz and
Weiss consider the case where lenders are faced with borrowers characterized by different risk levels: some want to borrow in order to finance low-risk projects, while others need funds for high-risk investments. For low-risk borrowers the chances of default on the loan are low, but so are the potential returns on the investment; thus low-risk borrowers are not willing to borrow if the interest rate is high. High-risk borrowers, on the other hand, have higher chances of default as well as higher potential returns and are thus willing to borrow at high interest rates. If each borrower knows his own risk-level while lenders cannot distinguish between high-risk and low-risk applicants, we have a situation of asymmetric information. The lenders are thus in the same situation as the buyers of cars of unknown quality in Akerlof’s model. Since safe borrowers are not willing to apply for a loan when the interest rate is high, while risky borrowers are, a situation of adverse selection arises: an increase in the interest rate leads to a worse pool of applicants and thus to lower expected profits for the lender (because of the higher proportion of borrowers who are likely to default).

5. Signaling

The analysis of signaling was initiated by Spence’s path-breaking book *Market signaling*, published in 1974. Spence considers informational asymmetries between two parties to a potential transaction: some relevant information is available to only one party, while the other party tries to infer that information from some observable characteristics.

5.1 Signals and indices

There are many characteristics that can be associated with a particular individual. Some of these (such as race, gender, educational certificates, previous work experience) are either directly observable by - or can be credibly and verifiably communicated to - the other party of a potential transaction (e.g. a prospective employer). Other characteristics (e.g. character traits such as honesty, dependability, conscientiousness, punctuality) are known to the individual but cannot be credibly communicated at the time of the transaction. The observable characteristics are often used by the individual to attempt to convey information about the unobservable attributes. For example, an entrepreneur might know her own ability to bring an investment to profitable fruition but might be unable to credibly convey this information to a potential investor;
on the other hand, the amount of the entrepreneur’s personal wealth invested in the project can be verifiably communicated. Thus the entrepreneur might decide to devote a large fraction of her wealth to the project in order to convince a potential investor that the project is worth financing.

Spence distinguishes between those observable characteristics that can be manipulated by - and are potentially available to - all individuals (e.g. the way one dresses or the number of years of schooling) and those attributes that are fixed and cannot be changed (e.g. one’s race or gender or the blushing associated with shyness). Spence calls the former characteristics *signals* and the latter *indices*. If employers offer higher salaries to those who have a college degree relative to those who have only a high school diploma, then everybody can (but might choose not to) acquire a college degree and avail oneself of the higher salary. Thus a college degree is a signal. On the other hand, if employers offer higher salaries to white applicants than to black applicants, then the higher salaries are not accessible to blacks because race cannot be changed (it is an index). We begin our analysis by restricting attention to signals.

### 5.2 Signaling equilibria

The situation considered by Spence, typified by the job market, is one where some relevant information is available to only one side of the market (the prospective employees), while the other side of the market (the employer) has to try to infer this information from some observable characteristics, which are potentially available to all applicants (and thus are signals). The situation can be illustrated in the following example.

Let \( y \) denote the amount of time spent in school (measured in years). Education certificates are obtained after several years of schooling as shown in Table 5.1.
Suppose that employers believe that education affects productivity; more specifically, suppose that employers believe that somebody with an elementary school certificate has a productivity of $6,000 and that a high school diploma adds $14,000 to a person’s productivity, a college degree adds another $5,000, a Master’s degree adds another $5,000 and a PhD degree adds another $2,000. The employer’s beliefs are reflected in the following wage schedule, which is known to the potential applicants.

<table>
<thead>
<tr>
<th>Educational certificate</th>
<th>Offered wage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary school</td>
<td>$6,000</td>
</tr>
<tr>
<td>High-school diploma</td>
<td>$20,000</td>
</tr>
<tr>
<td>College degree</td>
<td>$25,000</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>$30,000</td>
</tr>
<tr>
<td>PhD degree</td>
<td>$32,000</td>
</tr>
</tbody>
</table>

Table 5.2
The wage schedule reflecting the employers’ beliefs concerning the relationship between education and productivity

To make the example more striking, suppose that, as a matter of fact, education does not affect productivity at all: there are two types of individuals, those (Type L) with productivity $20,000 and those (Type H) with productivity $30,000. The proportion of individuals of Type L
in the population is \( q_L \) with \( 0 < q_L < 1 \) (and the proportion of Type \( H \) is \( 1 - q_L \)). If the two types of individuals are otherwise identical, they will make the same choices. Suppose, however, that they face different costs of acquiring education. The cost could be measured in terms of effort (for example, productivity could be correlated with intelligence and thus Type \( H \) individuals find it easier to progress through the schooling system). For simplicity we will take costs to be monetary costs (e.g. the monetary equivalent of expended effort). Suppose that Type \( L \) individuals have the following cost of acquiring \( y \) years of education (for \( y \geq 6 \)):

\[
C_L(y) = 2,000(y - 6),
\]

while Type \( H \) individuals have the following cost of acquiring education:

\[
C_H(y) = 1,000(y - 6).
\]

Thus every extra year of schooling costs $2,000 to a Type \( L \) but only $1,000 to a Type \( H \).

Given the wage schedule offered by employers, every individual will only consider values of \( y \) from the set \{6, 12, 16, 18, 21\}, since every other level of education will imply the same wage as some value in the set \{6, 12, 16, 18, 21\} but higher costs. Each individual will make her choice on the basis of a cost-benefit analysis as shown in the Tables 5.3 and 5.4.

<table>
<thead>
<tr>
<th>( y ) (years of schooling)</th>
<th>( w ) (gross wage)</th>
<th>( C_L ) (cost)</th>
<th>wage net of cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6,000</td>
<td>0</td>
<td>6,000</td>
</tr>
<tr>
<td><strong>12</strong></td>
<td><strong>20,000</strong></td>
<td><strong>12,000</strong></td>
<td><strong>8,000</strong></td>
</tr>
<tr>
<td>16</td>
<td>25,000</td>
<td>20,000</td>
<td>5,000</td>
</tr>
<tr>
<td>18</td>
<td>30,000</td>
<td>24,000</td>
<td>6,000</td>
</tr>
<tr>
<td>21</td>
<td>32,000</td>
<td>30,000</td>
<td>2,000</td>
</tr>
</tbody>
</table>

**Table 5.3**
The cost-benefit analysis for a Type \( L \) individual
Thus Type $L$ individuals will choose to obtain a high-school diploma and will be employed at a salary of $20,000$, which happens to be exactly their true productivity. Type $H$ individuals, on the other hand, will choose to obtain a Master’s degree and will be employed at a salary of $30,000$, which happens to be their true productivity. This is what Spence called a *signaling equilibrium*. Such an equilibrium has the following features. First of all, different types of individuals make different choices, in particular Type $H$ individuals signal their higher productivity by acquiring more education. The higher education certificate (and corresponding higher salary) is available also to Type $L$ individuals, but they choose not to avail themselves of this signal, since the cost to them is too high as compared to the benefit in terms of a higher wage. Secondly, the employers’ beliefs are confirmed. Initially they offer different wages on the basis of the signal produced by the applicant (the education certificate); later - after observing the employees and learning their true productivity - the employers find that what they believed to be the case turns out to be true: applicants with a higher investment in education are indeed more productive. Thus they are confirmed in their (objectively wrong) beliefs that more education causes higher productivity and have no reasons to change those beliefs: the employers’ beliefs become *self-fulfilling*.

Yet another feature of a signaling equilibrium is that it can be *Pareto inferior* to a situation where signaling is not available. A situation $X$ is said to be *Pareto inferior* to an

<table>
<thead>
<tr>
<th>$y$</th>
<th>$w$</th>
<th>$C_H$</th>
<th>wage net of cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6,000</td>
<td>0</td>
<td>6,000</td>
</tr>
<tr>
<td>12</td>
<td>20,000</td>
<td>6,000</td>
<td>14,000</td>
</tr>
<tr>
<td>16</td>
<td>25,000</td>
<td>10,000</td>
<td>15,000</td>
</tr>
<tr>
<td>18</td>
<td>30,000</td>
<td>12,000</td>
<td>18,000</td>
</tr>
<tr>
<td>21</td>
<td>32,000</td>
<td>15,000</td>
<td>17,000</td>
</tr>
</tbody>
</table>

Table 5.4
The cost-benefit analysis for a Type $H$ individual
alternative situation $Y$ if some individuals are strictly better off in situation $Y$ as compared to situation $X$ and everybody else is at least as well off in situation $Y$ as in situation $X$. To continue our example, suppose that the higher-education signal is not available (for example, because the government shuts down all institutions of higher education). In such a hypothetical situation, everybody will choose to obtain a high school diploma (this can be seen from the cost-benefit analysis of Tables 5.3 and 5.4 truncated at the value $y = 12$). Hence all individuals would look identical to the employers and the employers would have to offer the same salary to every applicant, namely a salary equal to the average productivity which is

$$20,000q_L + 30,000(1-q_L) = 30,000 - 10,000q_L.$$ \textsuperscript{5}

Who would be better off in this hypothetical situation? Type $L$ individuals for sure, since $30,000 - 10,000q_L - 12,000 > 8,000$ (this is true for every $q_L$ such that $0 < q_L < 1$). Type $H$ individuals are better off if and only if

$$30,000 - 10,000q_L - 6,000 > 18,000,$$ that is, if and only if $q_L < \frac{3}{5}$. Hence if less than 60% of the population is of Type $L$, shutting down all institutions of higher education would make everybody better off! \textsuperscript{6}

---

\textsuperscript{5} In the situation described, hiring a new employee can be viewed as playing a lottery where, with probability $q_L$, the employee is worth $20,000$ to the employer and, with probability $(1-q_L)$, she is worth $30,000$. Assuming risk neutrality on the part of the employers, such a lottery is equivalent to its expected value which is $\$(30,000 - 10,000q_L)$.

\textsuperscript{6} The employers – if risk neutral – will be indifferent between the two situations. In the signaling equilibrium the probability that an applicant will be of type $L$ and will thus produce a high school certificate and receive a salary of $20,000$ is $q_L$ and the probability that the applicant will be of type $H$, will produce a Master’s degree certificate and obtain a salary of $30,000$ is $1-q_L$. Thus, on average, employers will pay each worker $\$(30,000 - 10,000q_L)$, which is the same salary offered to each applicant in the alternative situation where signaling is not available.
5.3 The interaction of indices and signals

The example of the previous section highlights the possibility of objectively wrong beliefs (by employers) that give rise to choices (by prospective employees) that confirm those beliefs. The example involves the use of a signal (the level of education) by one group of individuals (the $H$ types) to separate themselves from another group (the $L$ types). Could the same phenomenon happen when the false beliefs involve an index? For example, suppose that there is no objective difference between men and women concerning productivity and yet employers incorrectly believe that women are less productive than men. Wouldn’t employers at some stage discover that their beliefs were wrong (by observing the equal productivity of men and women) and thus be forced to abandon those beliefs? Spence showed that, in the presence of signals, false beliefs concerning indices might also be self-fulfilling. We illustrate this possibility by expanding on the example of the previous section. Suppose that, as before, there are two types of individuals: Type $L$ with productivity $20,000$ and Type $H$ with productivity $30,000$. Within each type there are both men and women, who are identical in terms of productivity, that is, a woman of Type $L$ has a productivity of $20,000$, just like a man of Type $L$, and similarly for Type $H$. Thus, within each type, productivity is independent of both the level of education and of gender. Suppose, as before, that employers wrongly believe that education increases productivity; we now add a second false belief, namely that women tend to be less productive than men. Suppose that, according to the employers’ wrong beliefs, for men the relationship between education and productivity is as in the previous section; women, on the other hand, progress in the same way as men up to high school, but gain less than men from attending institutions of higher education. The employers’ beliefs are reflected in the offer of different wages to men and women, as detailed in the following table.
In our example, within each type, men and women do not differ in any relevant respect. In particular the costs associated with education are the same as in the previous section. The cost benefit analysis for men of the two types are as shown in Tables 5.3 and 5.4. The cost benefit analysis for women is shown in Tables 5.6 and 5.7.
Table 5.7
The cost-benefit analysis for a woman of Type $H$

<table>
<thead>
<tr>
<th>$y$ (years of schooling)</th>
<th>$w$ (wage)</th>
<th>$C_H$ (cost)</th>
<th>wage net of cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6,000</td>
<td>0</td>
<td>6,000</td>
</tr>
<tr>
<td>12</td>
<td>20,000</td>
<td>6,000</td>
<td>14,000</td>
</tr>
<tr>
<td>16</td>
<td>23,000</td>
<td>10,000</td>
<td>13,000</td>
</tr>
<tr>
<td>18</td>
<td>26,000</td>
<td>12,000</td>
<td>14,000</td>
</tr>
<tr>
<td>21</td>
<td>30,000</td>
<td>15,000</td>
<td>15,000</td>
</tr>
</tbody>
</table>

Thus both men and women of Type $L$ choose to obtain a high-school diploma and are hired at a salary of 20,000 (which corresponds to their true productivity), while Type $H$ individuals make different choices depending on their gender: men obtain a Master’s degree while women obtain a PhD degree; both are paid $30,000, which corresponds to their true productivity. Once again, the employers’ false beliefs are confirmed: women are “slower learners”, that is, they need to invest more in education than men in order to achieve the productivity level of $30,000.

6. Screening and separating equilibria

We saw in Section 4 that situations of asymmetric information characterized by adverse selection can lead to inefficiencies and market failure. Are there ways in which the uninformed party can alleviate this problem? The answer is affirmative and draws from the insights gained from the phenomenon of signaling discussed in Section 5. In a signaling equilibrium, some types of individuals can take actions (invest in a signal such as education) which enables them to separate themselves from the other types. In this section we show that the uninformed party in an adverse selection situation can bring about an outcome where different types make different choices and thereby reveal their types. This phenomenon is called screening and was first pointed out by Rothschild and Stiglitz (1976) in the context of competitive insurance markets. We will illustrate it in the case where the insurance industry is a monopoly, elaborating on the example of Section 4. The main idea is that the insurance company, instead of offering just one
insurance contract, can offer a *menu* of contracts and let potential customers choose which kind of contract to purchase. The menu can be designed in such a way that different types of consumers will choose different contracts, thereby revealing their types.

As in Section 2, consider the case where the insurance industry is a monopoly facing a large number of potential customers who are risk-averse and identical in terms of initial wealth (denoted by $W$) and potential loss (denoted by $\ell$). While in Section 2 it was assumed that all consumers had the same probability of loss, here we consider the case where each consumer belongs to one of two groups: the high-risk group, which we call Group $H$, and the low-risk group, which we call Group $L$. The probability of loss for Group $L$ is $q_L$ and the probability of loss for Group $H$ is $q_H$, with $0 < q_L < q_H < 1$. The proportion of low-risk individuals in the population is $\lambda$ (with $0 < \lambda < 1$) and thus the proportion of high-risk individuals is $1 - \lambda$. While each individual knows whether she is high-risk or low-risk, the insurance company cannot tell whether an applicant belongs to Group $L$ or to Group $H$. We saw in Section 4 that when all individuals are identical, the monopolist would offer a full-insurance contract at a premium that exceeds the expected loss. In the two-type case considered here, if the monopolist were to offer the same full-insurance contract at premium $p$ to everybody then one of three situations would arise: (1) $p$ is so high that nobody applies (and thus the monopolist makes zero profits), (2) $p$ is sufficiently low for Group $H$ but too high for Group $L$, so that only high-risk individuals apply, and (3) $p$ is sufficiently low for everybody to apply. Case 2 arises when the premium $p$ is such that Type $L$ individuals prefer not to insure, that is, they prefer the lottery $\left( \begin{array}{c} W & W - \ell \\ 1 - q_L & q_L \end{array} \right)$ to the sure outcome $(W - p)$, while Type $H$ individuals prefer $(W - p)$ to the no-insurance lottery $\left( \begin{array}{c} W & W - \ell \\ 1 - q_H & q_H \end{array} \right)$; in this case the monopolist’s profit per customer is $\pi_2 = p - q_H \ell$. Case 3
arises when the premium $p$ is such that full insurance is preferred to no insurance by everybody; in this case the monopolist’s profit per customer is $\pi = p - [(1 - \lambda)q_H + \lambda q_L] \ell$. \(^7\)

A graphic illustration will be useful.

Figure 6.1
Graphical representation of insurance contracts, zero-profit line and indifference curve

Let us start with a risk-averse consumer with initial wealth $W$, potential loss $\ell$ and probability of loss $q$. In Figure 6.1 we denote by $W_1$ the individual’s wealth in the bad state (e.g. if a fire occurs) and measure it on the horizontal axis, while her wealth in the good state (no fire) is measured on the vertical axis and denoted by $W_2$. Point $NI$ represent the no insurance state,

\(^7\) The insurance company will collect the premium of $p$ from every consumer and pay $\ell$ to a Type $H$ with probability $q_H$ and $\ell$ to a Type $L$ with probability $q_L$. Since the proportion of Type $L$ in the population is $\lambda$,
corresponding to the lottery \( \begin{pmatrix} W & W - \ell \\ 1 - q & q \end{pmatrix} \). Points in the shaded triangle represent possible insurance contracts (the shaded triangle is the set of points \((W_1, W_2)\) satisfying the following constraints: (1) \(W_1 \geq W - \ell\), (2) \(W_2 \leq W\) and (3) \(W_1 \leq W_2\). Relative to no insurance, an insurance contract increases wealth in the bad state and reduces wealth in the good state. The portion of the shaded triangle which lies on the 45° line out of the origin represents full-insurance contracts, which equalize wealth in the two states (the 45° line is the set of points \((W_1, W_2)\) such that \(W_1 = W_2\)). A point in the shaded triangle, such as point \(A = (W_1^A, W_2^A)\) in Figure 8.1, can be viewed as the insurance contract with premium \(p_A = W - W_2^A\) and deductible \(d_A = W_2^A - W_1^A\). In fact, wealth in the good state is equal to initial wealth minus the premium: \(W_2^A = W - p_A\), and wealth in the bad state is equal to initial wealth minus the premium minus the deductible: \(W_1^A = W - p_A - d_A = W - (W - W_2^A) - d_A = W_2^A - d_A\).

The straight line in Figure 6.1 that goes through points NI and B is the zero-expected-profit line. In fact, point NI can be thought of as the trivial insurance contract with zero premium and deductible equal to the full loss; point B is the full-insurance contract with premium equal to expected loss, that is, \(p_B = q \ell\). (Points below the zero-profit line correspond to contracts that yield positive expected profits and points above the line correspond to contracts that yield negative expected profits, that is, a loss.)

Since the consumer is risk-averse, she will strictly prefer contract B to no insurance. The thick curve in Figure 6.1 is the indifference curve for the consumer that goes through the no-insurance point. For every contract there will be an indifference curve that goes through it, joining all the contracts that the consumer considers just as good as the contract under consideration. 8 Thus the average probability of making a payment to an insured customer is \((1 - \lambda)q_H + \lambda q_L\) and thus the average payment per customer is \((1 - \lambda)q_H + \lambda q_L\). 8

8 If the individual prefers more money to less (an assumption which we implicitly made) then each indifference curve must be downward-sloping. In fact, given two different points \((W_1, W_2)\) and \((W_1', W_2')\) with
curve in Figure 6.1 shows all the contracts that the consumer considers just as good as the option of not insuring. A contract below the indifference curve would be worse than not insuring and thus will be rejected by the consumer. Contracts above the indifference curve are considered better than not insuring and thus will be accepted by the consumer. For example, in Figure 6.1: contract $A$ is worse than not insuring, while contract $B$ is better than no insurance (indeed we know this from the definition of risk aversion).

Now let us return to the two-type case. Figure 6.2 shows two indifference curves, one for each type, that go through the no-insurance point $NI$.

$W' \geq W_1$ and $W'_2 \geq W_2$, the individual would strictly prefer the contract corresponding to $(W'_1, W'_2)$ to the contract corresponding to $(W_1, W_2)$, because the former would guarantee a higher wealth in at least one state (and the same or higher wealth in the other state) than the latter. Hence, in order for two points, say $A$ and $B$, to lie on the same indifference curve it must be that wealth in one state is higher in $A$ than in $B$ and wealth in the other state is lower in $A$ than in $B$. 

Figure 6.2
Two indifference curves, one for each type.
Given any point in the \((W_1,W_2)\) plane, there will be an indifference curve through it for the high-risk individuals (assuming that they all have the same preferences over insurance contracts) and a different indifference curve for the low-risk individuals (assuming that they, too, share the same preferences). In Figure 6.2 the indifference curve for the \(L\)-type is less steep than the indifference curve for the \(H\)-type.

The shaded area in Figure 6.2 represents all the contracts that are preferred to no insurance by the \(H\)-type but are worse than no insurance for the \(L\)-type. The existence of such contracts points to the possibility, for the monopolist, of offering - instead of a single contract - a pair of contracts, say \(A\) and \(B\), designed in such a way that the \(H\)-type consumers will prefer contract \(A\) to contract \(B\) while the \(L\)-type will prefer contract \(B\) to contract \(A\). Such a pair of contracts is said to induce separation of types. If it is profit-maximizing for the monopolist to offer a pair of contracts that induces separation of types, the corresponding outcome is called a separating equilibrium. It can be shown that the monopolist will indeed choose to induce separation of types (by offering a pair of contracts) whenever the proportion of \(H\)-types in the population is not too large.

A pair of contracts that induces separation of types is shown in Figure 6.3.
Consider first Type $H$ individuals. The two thick indifference curves belong to the $H$-type, one goes through the no-insurance point and the other through contract $A$. Since contract $A$ is above the $H$-type indifference curve that goes through point $NI$, the $H$-types are better off purchasing contract $A$ than not insuring; furthermore, since contract $B$ lies below the $H$-type indifference curve that goes through contract $A$, the $H$-types prefer contract $A$ to contract $B$. Thus, of the three options: (1) do not insure, (2) purchase contract $A$ and (3) purchase contract $B$, the $H$-types will choose Option 2 (they rank $A$ above $B$ and $B$ above $NI$).

Consider now the $L$-type. Contract $A$ is below the thin indifference curve, which is the $L$-type indifference curve that goes through the no-insurance point; thus the $L$-type prefer to remain...
uninsured rather than purchase contract A. On the other hand, contract B is above the L-type indifference curve that goes through the no-insurance point and, therefore, the L-types prefer contract B to no insurance. Thus, of the three options: (1) do not insure, (2) purchase contract A and (3) purchase contract B, the L-types will choose Option 3 (they rank B above NI and NI above A). Hence everybody will buy insurance, but the H-types will choose a different contract than the L-types: the different types will reveal themselves by making different choices. This phenomenon is reminiscent of the signaling equilibria discussed in Section 5, where different types of workers made different educational choices.

The pair of contracts shown in Figure 6.3 does not maximize the monopolist’s profits. A profit-maximizing monopolist would not choose a partial-insurance contract - such as contract A in Figure 6.3 - as the contract targeted to the H-types: by the argument used in Section 2, the monopolist could increase its profits by replacing contract A with an appropriate contract on the 45° line, that is, a full insurance contract. On the other hand, the contract targeted to the L-types has to be a partial-insurance contract, since it cannot lie above the H-type indifference curve that goes through the full-insurance contract targeted to the H-types (otherwise both types would choose this other contract and there would be no separation). Hence in a separating equilibrium the low-risk individuals distinguish themselves from the high-risk individuals by choosing a contract with positive deductible, that is, a partial-insurance contract.

7. Optimal risk-sharing

There are several contractual situations where one party, whom we call the Principal, hires another party, the Agent, to perform a task whose future outcome is uncertain at the time of contracting. The uncertainty is due to factors that cannot be predicted with certainty. We call these factors external states. For example, the Principal could be the owner of a firm and the Agent the manager, hired to run the firm; the outcome is the firm’s profit, which will be affected by a number of external states, such as the state of the economy, the intensity of competition, input costs, etc. Another example is a contract between a land-owner (the Principal) and the farmer (the Agent), where the outcome is the quantity of, say, fruit that will be produced next year; in this case the external states are the weather, the availability of water for irrigation, the presence of a sufficiently large number of pollinating insects, etc. Yet another example is where
the Agent is a lawyer and the Principal is her client and the outcome is, say, the amount of damages that will be awarded to the client; in this case external states that will affect the outcome include the composition of the jury, whether some witnesses will be able or willing to testify, etc. In all these examples there is typically another factor that will influence the outcome, namely the level of effort that the Agent exerts in the enterprise. If the Agent’s effort cannot be observed by the Principal then we have a situation of *moral hazard*. For example, the manager could devote most of his time playing video games and then claim that the low profits were due an unusually low demand, the lawyer could be devoting his time and energy to another case and then claim that the disappointing outcome was due to bad luck, etc.

In this section we assume that the Agent’s effort is not an issue, for example because it is observable by the Principal and verifiable by a court of law so that it can be specified in the contract. Thus we concentrate on the pure element of uncertainty due to the possibility of different future external states.

The main issue that arises in these situations is what type of payment to the Agent should be agreed upon by the two parties. One possibility is a contract that establishes a fixed payment, that is, a payment which is the same in every state. With such a contract the Principal’s income varies with the state and thus the risk is entirely borne by the Principal. Another possible contract is one that guarantees a fixed income to the Principal, thereby leaving the Agent to bear all the risk. A third possibility is a risk-sharing contract where the Agent is assigned a fraction of the surplus (and the Principal the remaining fraction). Is there a sense in which some of these contracts are unambiguously better than others? This is the issue addressed by the theory of optimal risk-sharing.

We shall illustrate the notion of optimal risk-sharing in the case where the Principal is the owner of a firm, the Agent is the manager, the outcome is the profit of the firm and there are only two external states: a good (e.g. high-demand) state where the profit is $\pi_g$ and a bad (e.g. low-demand) state where the profit is $\pi_b$, with $0 < \pi_b < \pi_g$. We denote by $p_g$ (with $0 < p_g < 1$) the probability of the good state (so that the probability of the bad state is $1-p_g$). The set of possible contracts can be represented by what is known as an *Edgeworth box*, illustrated in Figure 7.1.
The length of the long side of the rectangle is equal to $\pi_G$ and the length of the short side is $\pi_B$. The lower left-hand corner of the box is viewed as the origin (denoted by $0_P$) of a two-dimensional diagram measuring, on the horizontal axis, the amount of money received by the Principal in the good state (denoted by $x_1$) and, on the vertical axis, the amount of money received by the Principal in the bad state (denoted by $x_2$). The upper right-hand corner of the box is viewed as the origin (denoted by $0_A$) of a diagram - rotated by 180° - measuring, on the horizontal axis, the amount of money received by the Agent in the good state (denoted by $w_1$) and, on the vertical axis, the amount of money received by the Agent in the bad state (denoted by $w_2$). Thus, since $x_1 + w_1 = \pi_G$ and $x_2 + w_2 = \pi_B$, any point in the box represents a possible distribution - between the Principal and the Agent - of the profit of the firm in each state. For example, point $A$ in Figure 9.1 represents a contract according to which, if the good state occurs, the Agent receives $sw_1^A$ (and the Principal collects the residual amount $x_1^A = \pi_G - w_1^A$) and if the
bad state occurs the Agent receives \( w_2^A \) (and the Principal collects the residual amount \( x_2^A = \pi_B - w_2^A \)).

Without any further information about the specific context of the situation, can we narrow down the set of contracts that could be agreed upon by Principal and Agent? It seems reasonable to expect that the two parties will not sign a contract \( A \) if there is an alternative contract \( B \) which is Pareto superior to \( A \), in the sense that both parties prefer \( B \) to \( A \). We define a contract \( C \) to be Pareto efficient if there is no other contract which is Pareto superior to it. Is it possible to identify the set of Pareto efficient contracts? We shall address this question under the assumption that each party only cares about how much money he ends up with and prefers more money to less. As we did in Section 6, we make use of indifference curves to elaborate on this point.

\[ x_2^A = \pi_B - w_2^A \]

\[ x_1^A = \pi_G - w_1^A \]

Figure 7.2

The shaded area represents contracts that are Pareto superior to contract \( A \)

Figure 7.2 shows two indifference curves through point \( A \): the straight line is the indifference curve of the Principal and the curved line is the indifference curve of the Agent. A
A straight-line indifference curve implies risk neutrality; thus in Figure 7.2 it is assumed that the Principal is risk-neutral. On the other hand, the fact that the indifference curve of the Agent is convex towards the Agent’s origin implies that the Agent is risk averse.

The straight-line indifference curve of the Principal that goes through point $A$ divides the box into three regions: the region above the line, consisting of contracts that the Principal prefers to contract $A$, the region below the line, consisting of contracts that the Principal finds worse than contract $A$, and the line itself, which consists of all the contracts that the Principal considers to be just as good as $A$.

Similarly, the indifference curve of the Agent that goes through point $A$ divides the box into three regions: the region between the curve and lower left-hand corner of the box, consisting of contracts that the Agent prefers to contract $A$ (such points are below the curve if we look at the curve from point $0_P$ but they are above the curve if we take the viewpoint of the Agent which is $0_A$), the region between the curve and the upper right-hand corner of the box, consisting of contracts that the Agent finds worse than contract $A$, and the curve itself, which consists of all the contracts that the Agent considers to be just as good as $A$.

Thus the shaded area between the two indifference curves represents contracts that are Pareto superior to $A$. Hence contract $A$ is not Pareto efficient.

---

9 Recall that an individual is said to be risk-neutral if he considers a money lottery to be just as good as the expected value of the lottery for sure. Let $A = (x_1^A, x_2^A)$ be a possible contract (a point in the Edgeworth box) and let $B = (x_1^B, x_2^B)$ be another contract which the Principal considers to be just as good as $A$ (that is, $A$ and $B$ lie on the same indifference curve for the Principal). Then, since $A$ corresponds to the lottery $\left(\begin{array}{cc} x_1^A & x_2^A \\ p_G & 1 - p_G \end{array} \right)$ and $B$ corresponds to the lottery $\left(\begin{array}{cc} x_1^B & x_2^B \\ p_G & 1 - p_G \end{array} \right)$, it must be that the expected value of $A$ is equal to the expected value of $B$, that is, $x_1^A p_G + x_2^A (1 - p_G) = x_1^B p_G + x_2^B (1 - p_G)$. Rearranging this equation we get $\frac{x_2^A - x_2^B}{x_1^A - x_1^B} = -\frac{p_G}{1 - p_G}$, whose left-hand side is the ratio of the change in the vertical coordinate to the change in the horizontal coordinate; thus this ratio is a negative constant, meaning that the indifference curve is a downward-sloping straight line.
In the case where the Principal is risk-neutral and the Agent is risk-averse the only Pareto efficient contracts (among those that involve positive payments to the Agent in both states) are the ones that lie on the $45^0$ line out of the origin for the Agent, that is contracts that guarantee the same income to the Agent, no matter what external state occurs.  

Whenever the indifference curves of Principal and Agent that go through a given contract cross, there will be an area between the two curves consisting of contracts that are Pareto superior to the contract under consideration. Thus at a Pareto efficient contract the two indifference curves cannot cross, that is, they must be tangent. Figure 7.3 shows a Pareto efficient contract, namely contract $C$ (where $w_1^C = w_2^C$), for the case where the Principal is risk-neutral and the Agent is risk averse.

---

10 This can be proved as follows, by showing that any contract not on the $45^0$ degree line for the Agent is Pareto inferior to some other contract. Fix a contract $A$ that involves a payment to the Agent of $W_1^A > 0$ in the good state and a payment of $W_2^A > 0$ in the bad state and suppose that $A$ is not on the $45^0$ degree line for the Agent so that $W_1^A \neq W_2^A$. From the point of view of the Agent this contract corresponds to the lottery $\left( \begin{array}{c} W_1^A \\ W_2^A \\ p_G \end{array} \right)$, whose expected value is $a = W_1^A p_G + W_2^A (1 - p_G)$. Since the Agent is risk-averse she will prefer contract $B$ defined by $W_1^B = W_2^B = a$, that is, the contract that guarantees a fixed income of $a$. How would the Principal rank $B$ versus $A$? From the point of view of the Principal, contract $A$ corresponds to the lottery $L_A = \left( \begin{array}{c} \pi_G - W_1^A \\ \pi_B - W_2^A \\ p_G \end{array} \right)$ and contract $B$ to the lottery $L_B = \left( \begin{array}{c} \pi_G - a \\ \pi_B - a \\ p_G \end{array} \right)$. The expected value of $L_A$ is $p_G(\pi_G - a) + (1 - p_G)(\pi_B - a) = p_G \pi_G + (1 - p_G) \pi_B - a$ and the expected value of $L_B$ is $p_G(\pi_G - a) + (1 - p_G)(\pi_B - a) = p_G \pi_G + (1 - p_G) \pi_B - [W_1^A p_G + W_2^A (1 - p_G)]$ which is equal to $p_G \pi_G + (1 - p_G) \pi_B - a$ (since $W_1^A p_G + W_2^A (1 - p_G) = a$). Thus the Principal, being risk-neutral, would be indifferent between contract $A$ and contract $B$. If the two parties switched to a contract like $B$ but with a slightly smaller payment $a' < a$, then the Agent would still be better off and the Principal would also be better off.
A similar analysis applies to the case where the Principal is risk-averse and the Agent is risk-neutral: in such a case Pareto efficiency requires that the Principal be guaranteed a fixed income (that is, the Pareto efficient contracts are the ones that lie on the 45° line out of the origin for the Principal).

Thus we have a general principle of optimal risk-sharing: *when one of the two parties to a contract is risk-neutral and the other is risk-averse, Pareto efficiency requires that the entire risk be borne by the risk-neutral party (so that the risk-averse party is guaranteed a fixed income).*

What if both Principal and Agent are risk-averse? In this case it is not possible to guarantee a fixed income to both individuals. It is still the case that Pareto efficiency of a contract, say $C$, requires that the indifference curves of both individuals that go through point $C$ not cross at $C$ (otherwise there would be an area between the two curves representing contracts that are Pareto superior to $C$). Thus the indifference curves must be tangent to each other at a Pareto efficient contract, as shown in Figure 7.4. Which individual comes closer to income certainty (that is, to which of the two 45° lines the contract is closer) depends on who is more risk averse.
Figure 7.4
A Pareto efficient contract for the case where both Principal and Agent are risk averse

References


Hey, John, Uncertainty in microeconomics, Martin Robertson, Oxford, 1981.


