Economic & Financial Decisions under Risk (Chapters 1&2)
Eeckhoudt, Gollier & Schlesinger (Princeton Univ Press 2005)

1
Risk Aversion

This chapter looks at a basic concept behind modeling individual preferences in the face of risk. As with any social science, we of course are fallible and susceptible to second-guessing in our theories. It is nearly impossible to model many natural human tendencies such as "playing a hunch" or "being superstitious." However, we can develop a systematic way to view choices made under uncertainty. Hopefully, our models can capture the basic human tendencies enough to be useful in understanding market behavior towards risk. In other words, even if we are not correct in predicting behavior under risk for every individual in every circumstance, we can still make general claims about such behavior and can still make market predictions, which after all are based on the "marginal consumer."

To use (vaguely) mathematical language, the understanding of this chapter is a *necessary* but not *sufficient* condition to go further into the analysis. Because of the importance of risk aversion in decision making under uncertainty, it is worthwhile to first take an "historical" perspective about its development and to indicate how economists and decision scientists progressively have elaborated upon the tools and concepts we now use to analyze risky choices. In addition, this "history" has some surprising aspects that are interesting in themselves. To this end, our first section in this chapter broadly covers these retrospective topics. Subsequent sections are more "modern" and they represent an intuitive introduction to the central contribution to our field, that of Pratt (1964).

1.1 An Historical Perspective on Risk Aversion

As it is now widely acknowledged, an important breakthrough in the analysis of decisions under risk was achieved when Daniel Bernoulli, a distinguished Swiss mathematician, wrote in St Petersburg in 1738 a paper in Latin entitled: "*Specimen theoriae novae de mensura sortis*," or "Exposition of a new theory on the measurement of risk." Bernoulli's paper, translated into English in Bernoulli (1954), is essentially nontechnical. Its main purpose is to show that two people facing the same lottery may value it differently because of a difference in their psychology. This idea was quite novel at the time, since famous scientists before Bernoulli (among them

Pascal and Fermat) had argued that the value of a lottery should be equal to its mathematical expectation and hence identical for all people, independent of their risk attitude.

In order to justify his ideas, Bernoulli uses three examples. One of them, the "St Petersburg paradox" is quite famous and it is still debated today in scientific circles. It is described in most recent texts of finance and microeconomics and for this reason we do not discuss it in detail here. Peter tosses a fair coin repetitively until the coin lands head for the first time. Peter agrees to give to Paul 1 ducat if head appears on the first toss, 2 ducats if head appears only on the second toss, 4 ducats if head appears for the first time on the third toss, and so on, in order to double the reward to Paul for each additional toss necessary to see the head for the first time. The question raised by Bernoulli is how much Paul would be ready to pay to Peter to accept to play this game.

Unfortunately, the celebrity of the paradox has overshadowed the other two examples given by Bernoulli that show that, most of the time, the value of a lottery is not equal to its mathematical expectation. One of these two examples, which presents the case of an individual named "Sempronius," wonderfully anticipates the central contributions that would be made to risk theory about 230 years later by Arrow, Pratt and others.

Let us quote Bernoulli:[1]

> Sempronius owns goods at home worth a total of 4000 ducats and in addition possesses 8000 ducats worth of commodities in foreign countries from where they can only be transported by sea. However, our daily experience teaches us that of [two] ships one perishes.

In modern-day language, we would say that Sempronius faces a risk on his wealth. This wealth may represented by a lottery $\tilde{x}$, which takes on a value of 4000 ducats with probability $\frac{1}{2}$ (if his ship is sunk), or 12 000 ducats with probability $\frac{1}{2}$. We will denote such a lottery $\tilde{x}$ as being distributed as $(4000, \frac{1}{2}; 12\,000, \frac{1}{2})$. Its mathematical expectation is given by:

$$E\tilde{x} \equiv \tfrac{1}{2}4000 + \tfrac{1}{2}12\,000 = 8000 \text{ ducats.}$$

Now Sempronius has an ingenious idea. Instead of "trusting all his 8000 ducats of goods to one ship," he now "trusts equal portions of these commodities to two ships." Assuming that the ships follow independent but equally dangerous routes, Sempronius now faces a more diversified lottery $\tilde{y}$ distributed as

$$(4000, \tfrac{1}{4}; 8000, \tfrac{1}{2}; 12\,000, \tfrac{1}{4}).$$

[1] We altered Bernoulli's probabilities to simplify the computations. In particular, Bernoulli's original example had one ship in ten perish.

Indeed, if both ships perish, he would end up with his sure wealth of 4000 ducats. Because the two risks are independent, the probability of these joint events equals the product of the individual events, i.e. $(\frac{1}{2})^2 = \frac{1}{4}$. Similarly, both ships will succeed with probability $\frac{1}{4}$, in which case his final wealth amounts to 12 000 ducats. Finally, there is the possibility that only one ship succeeds in downloading the commodities safely, in which case only half of the profit is obtained. The final wealth of Sempronius would then just amount to 8000 ducats. The probability of this event is $\frac{1}{2}$ because it is the complement of the other two events which have each a probability of $\frac{1}{4}$.

Since common wisdom suggests that diversification is a good idea, we would expect that the value attached to $\tilde{y}$ exceeds that attributed to $\tilde{x}$. However, if we compute the expected profit, we obtain that

$$E\tilde{y} = \tfrac{1}{4}4000 + \tfrac{1}{2}8000 + \tfrac{1}{4}12\,000 = 8000 \text{ ducats},$$

the same value as for $E\tilde{x}$! If Sempronius would measure his well-being *ex ante* by his expected future wealth, he should be indifferent about whether to diversify or not. In Bernoulli's example, we obtain the same expected future wealth for both lotteries, even though most people would find $\tilde{y}$ more attractive than $\tilde{x}$. Hence, according to Bernoulli and to modern risk theory, the mathematical expectation of a lottery is not an adequate measure of its value. Bernoulli suggests a way to express the fact that most people prefer $\tilde{y}$ to $\tilde{x}$: a lottery should be valued according to the "expected utility" that it provides. Instead of computing the expectation of the monetary outcomes, we should use the expectation of the utility of the wealth. Notice that most human beings do not extract utility from wealth. Rather, they extract utility from consuming goods that can be purchased with this wealth. The main insight of Bernoulli is to suggest that there is a nonlinear relationship between wealth and the utility of consuming this wealth.

What ultimately matters for the decision maker *ex post* is how much satisfaction he or she can achieve with the monetary outcome, rather than the monetary outcome itself. Of course, there must be a relationship between the monetary outcome and the degree of satisfaction. This relationship is characterized by a utility function u, which for every wealth level x tells us the level of "satisfaction" or "utility" $u(x)$ attained by the agent with this wealth. Of course, this level of satisfaction derives from the goods and services that the decision maker can purchase with a wealth level x. While the outcomes themselves are "objective," their utility is "subjective" and specific to each decision maker, depending upon his or her tastes and preferences. Although the function u transforms the objective result x into a perception $u(x)$ by the individual, this transformation is assumed to exhibit some basic properties of rational behavior. For example, a higher level of x (more wealth) should induce a higher level of utility: the function should be increasing in x. Even for someone who is very altruistic, a higher x will allow them to be more philanthropic. Readers familiar with indirect utility functions from microeconomics (essentially utility over budget sets, rather than over bundles of goods and services) can think of $u(x)$ as essentially an indirect utility of wealth, where we assume that prices for goods and services are fixed. In other words, we may think of $u(x)$ as the highest achievable level of utility from bundles of goods that are affordable when our income is x.

Bernoulli argues that if the utility u is not only increasing but also concave in the outcome x, then the lottery $\tilde{y}$ will have a higher value than the lottery $\tilde{x}$, in accordance with intuition. A twice-differentiable function u is concave if and only if its second derivative is negative, i.e. if the marginal utility $u'(x)$ is decreasing in x.[2] In order to illustrate this point, let us consider a specific example of a utility function, such as $u(x) = \sqrt{x}$, which is an increasing and concave function of x. Using these preferences in Sempronius's problem, we can determine the expectation of $u(x)$:

$$Eu(\tilde{x}) = \tfrac{1}{2}\sqrt{4000} + \tfrac{1}{2}\sqrt{12\,000} = 86.4$$

$$Eu(\tilde{y}) = \tfrac{1}{4}\sqrt{4000} + \tfrac{1}{2}\sqrt{8000} + \tfrac{1}{4}\sqrt{12\,000} = 87.9.$$

Because lottery $\tilde{y}$ generates a larger expected utility than lottery $\tilde{x}$, the former is preferred by Sempronius. The reader can try using concave utility functions other than the square-root function to obtain the same type of result. In the next section, we formalize this result.

Notice that the concavity of the relationship between wealth x and satisfaction/utility u is quite a natural assumption. It simply implies that the marginal utility of wealth is decreasing with wealth: one values a one-ducat increase in wealth more when one is poorer than when one is richer. Observe that, in Bernoulli's example, diversification generates a mean-preserving transfer of wealth from the extreme events to the mean. Transferring some probability weight from $x = 4000$ to $x = 8000$ increases expected utility. Each probability unit transferred yields an increase in expected utility equaling $u(8000) - u(4000)$. On the contrary, transferring some probability weight from $x = 12\,000$ to $x = 8000$ reduces expected utility. Each probability unit transferred yields a reduction in expected utility equaling $u(12\,000) - u(8000)$. But the concavity of u implies that

$$u(8000) - u(4000) > u(12\,000) - u(8000), \tag{1.1}$$

i.e. that the positive effect of these combined transfers must dominate the negative effect. This is why all investors with a concave utility would support Sempronius's strategy to diversify risks.

[2]For simplicity, we maintain the assumption that u is twice differentiable throughout the book. However, a function need not be differentiable to be concave. More generally, a function u is concave if and only if $\lambda u(a) + (1-\lambda)u(b)$ is smaller than $u(\lambda a + (1-\lambda)b)$ for all (a, b) in the domain of u and all scalars λ in [0, 1]. A function must, however, be continuous to be concave.

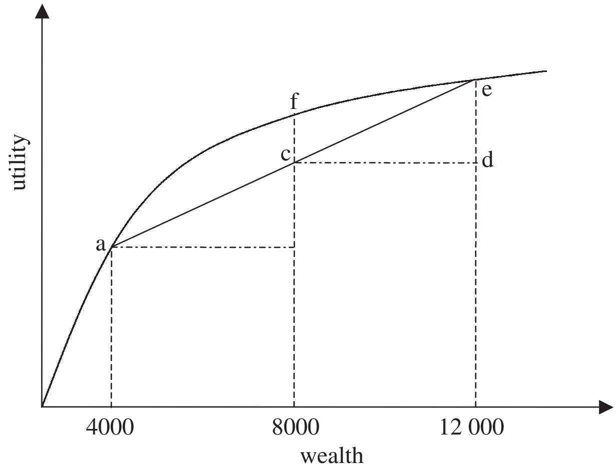

Figure 1.1. Measuring the expecting utility of final wealth (4000, $\frac{1}{2}$; 12000, $\frac{1}{2}$).

1.2 Definition and Characterization of Risk Aversion

We assume that the decision maker lives for only one period, which implies that he immediately uses all his final wealth to purchase and to consume goods and services. Later in this book, we will disentangle wealth and consumption by allowing the agent to live for more than one period. Final wealth comes from initial wealth w plus the outcome of any risk borne during the period.

Definition 1.1. An agent is risk-averse if, at any wealth level w, he or she dislikes every lottery with an expected payoff of zero: $\forall w$, $\forall \tilde{z}$ with $E\tilde{z} = 0$, $Eu(w + \tilde{z}) \leqslant u(w)$.

Observe that any lottery $\tilde{z}$ with a non-zero expected payoff can be decomposed into its expected payoff $E\tilde{z}$ and a zero-mean lottery $\tilde{z} - E\tilde{z}$. Thus, from our definition, a risk-averse agent always prefers receiving the expected outcome of a lottery with certainty, rather than the lottery itself. For an expected-utility maximizer with a utility function u, this implies that, for any lottery $\tilde{z}$ and for any initial wealth w,

$$Eu(w + \tilde{z}) \leqslant u(w + E\tilde{z}). \tag{1.2}$$

If we consider the simple example from Sempronius's problem, with only one ship the initial wealth w equals 4000, and the profit $\tilde{z}$ takes the value 8000 or 0 with equal probabilities. Because our intuition is that Sempronius must be risk averse, it must follow that

$$\tfrac{1}{2}u(12\,000) + \tfrac{1}{2}u(4000) \leqslant u(8000). \tag{1.3}$$

If Sempronius could find an insurance company that would offer full insurance at an actuarially fair price of $E\tilde{z} = 4000$ ducats, Sempronius would be better off by purchasing the insurance policy. We can check whether inequality (1.3) is verified in Figure 1.1. The right-hand side of the inequality is represented by point 'f' on the utility curve u. The left-hand side of the inequality is represented by the middle point on the arc 'ae', i.e. by point 'c'. This can immediately be checked by observing that the two triangles 'abc'and 'cde'are equivalent, since they have the same base and the same angles. We observe that 'f' is above 'c': *ex ante*, the welfare derived from lottery $\tilde{z}$ is smaller than the welfare obtained if one were to receive its expected payoff $E\tilde{z}$ with certainty. In short, Sempronius is risk-averse. From this figure, we see that this is true whenever the utility function is concave. The intuition of the result is very simple: if marginal utility is decreasing, then the potential loss of 4000 reduces utility more than the increase in utility generated by the potential gain of 4000. Seen *ex ante*, the expected utility is reduced by these equally weighted potential outcomes.

It is noteworthy that Equations (1.1) and (1.3) are exactly the same. The preference for diversification is intrinsically equivalent to risk aversion, at least under the Bernoullian expected-utility model.

Using exactly the opposite argument, it can easily be shown that, if u is convex, the inequality in (1.2) will be reversed. Therefore, the decision maker prefers the lottery to its mathematical expectation and he reveals in this way his inclination for taking risk. Such individual behavior will be referred to as risk loving. Finally, if u is linear, then the welfare Eu is linear in the expected payoff of lotteries. Indeed, if $u(x) = a + bx$ for all x, then we have

$$Eu(w + \tilde{z}) = E[a + b(w + \tilde{z})] = a + b(w + E\tilde{z}) = u(w + E\tilde{z}),$$

which implies that the decision maker ranks lotteries according to their expected outcome. The behavior of this individual is called risk-neutral.

In the next proposition, we formally prove that inequality (1.2) holds for any lottery $\tilde{z}$ and any initial wealth w if and only if u is concave.

Proposition 1.2. *A decision maker with utility function u is risk-averse, i.e. inequality (1.2) holds for all w and $\tilde{z}$, if and only if u is concave.*

Proof. The proof of sufficiency is based on a second-order Taylor expansion of $u(w + z)$ around $w + E\tilde{z}$. For any z, this yields

$$u(w + z) = u(w + E\tilde{z}) + (z - E\tilde{z})u'(w + E\tilde{z}) + \tfrac{1}{2}(z - E\tilde{z})^2 u''(\xi(z))$$

for some $\xi(z)$ in between z and $E\tilde{z}$. Because this must be true for all z, it follows that the expectation of $u(w + \tilde{z})$ is equal to

$$Eu(w + \tilde{z}) = u(w + E\tilde{z}) + u'(w + E\tilde{z})E(\tilde{z} - E\tilde{z}) + \tfrac{1}{2}E[(\tilde{z} - E\tilde{z})^2 u''(\xi(\tilde{z}))].$$

Observe now that the second term of the right-hand side above is zero, since $E(\tilde{z} - E\tilde{z}) = E\tilde{z} - E\tilde{z} = 0$. In addition, if u'' is uniformly negative, then the third term takes the expectation of a random variable $(\tilde{z} - E\tilde{z})^2 u''(\xi(\tilde{z}))$ that is always negative, as it is the product of a squared scalar and negative u''. Hence, the sum of these three terms is less than $u(w + E\tilde{z})$. This proves sufficiency.

Necessity is proven by contradiction. Suppose that u is not concave. Then, there must exist some w and some $\delta > 0$ for which $u''(x)$ is positive in the interval $[w - \delta, w + \delta]$. Now take a small zero-mean risk $\tilde{\varepsilon}$ such that the support of final wealth $w + \tilde{\varepsilon}$ is entirely contained in $(w - \delta, w + \delta)$. Using the same Taylor expansion as above yields

$$Eu(w + \tilde{\varepsilon}) = u(w) + \frac{1}{2}E[\tilde{\varepsilon}^2 u''(\xi(\tilde{\varepsilon}))].$$

Because $\xi(\tilde{\varepsilon})$ has a support that is contained in $[w - \delta, w + \delta]$ where u is locally convex, $u''(\xi(\tilde{\varepsilon}))$ is positive for all realizations of $\tilde{\varepsilon}$. Consequently, it follows that $E[\tilde{\varepsilon}^2 u''(\xi(\tilde{\varepsilon}))]$ is positive, and $Eu(w + \tilde{\varepsilon})$ is larger than $u(w)$. Thus, accepting the zero-mean lottery $\tilde{\varepsilon}$ raises welfare and the decision maker is not risk-averse. This is a contradiction. □

The above proposition is in fact nothing more than a rewriting of the famous Jensen inequality. Consider any real-valued function ϕ. Jensen's inequality states that $E\phi(\tilde{y})$ is smaller than $\phi(E\tilde{y})$ for any random variable $\tilde{y}$ if and only if ϕ is a concave function. It builds a bridge between two alternative definitions of the concavity of u: the negativity of u'' and the property that any arc linking two points on curve u must lie below this curve. Figure 1.1 illustrates this point. It is intuitive that decreasing marginal utility ($u'' < 0$) means risk aversion. In a certain world, decreasing marginal utility means that an increase in wealth by 100 dollars has a positive effect on utility that is smaller than the effect of a reduction in wealth by 100 dollars. Then, in an uncertain world, introducing the risk to gain or to lose 100 dollars with equal probability will have a negative net impact on expected utility. In expectation, the benefit of the prospect of gaining 100 dollars is overweighted by the cost of the prospect of losing 100 dollars with the same probability. Over the last two decades, many prominent researchers in the field have challenged the idea that risk aversion comes only from decreasing marginal utility. Some even challenged the idea itself, that there should be any link between the two.[3]

1.3 Risk Premium and Certainty Equivalent

A risk-averse agent is an agent who dislikes zero-mean risks. The qualifier "zero-mean" is very important. A risk-averse agent may like risky lotteries if the expected

---

[3]This question will be discussed in the last chapter of this book. Yaari (1987) provides a model that is dual to expected utility, where agents may be risk-averse in spite of the fact that their utility is linear in wealth.

payoffs that they yield are large enough. Risk-averse investors may want to purchase risky assets if their expected returns exceed the risk-free rate. Risk-averse agents may dislike purchasing insurance if it is too costly to acquire. In order to determine the optimal trade-off between the expected gain and the degree of risk, it is useful to quantify the effect of risk on welfare. This is particularly useful when the agent subrogates the risky decision to others, as is the case when we consider public safety policy or portfolio management by pension funds, for example. It is important to quantify the degree of risk aversion in order to help people to know themselves better, and to help them to make better decisions in the face of uncertainty. Most of this book is about precisely this problem. Clearly, people have different attitudes towards risks. Some are ready to spend more money than others to get rid of a specific risk. One way to measure the degree of risk aversion of an agent is to ask her how much she is ready to pay to get rid of a *zero-mean* risk $\tilde{z}$. The answer to this question will be referred to as the risk premium Π associated with that risk. For an agent with utility function u and initial wealth w, the risk premium must satisfy the following condition:

$$Eu(w + \tilde{z}) = u(w - \Pi). \tag{1.4}$$

The agent ends up with the same welfare either by accepting the risk or by paying the risk premium Π. When risk $\tilde{z}$ has an expectation that differs from zero, we usually use the concept of the certainty equivalent. The certainty equivalent e of risk $\tilde{z}$ is the sure increase in wealth that has the same effect on welfare as having to bear risk $\tilde{z}$, i.e.

$$Eu(w + \tilde{z}) = u(w + e). \tag{1.5}$$

When $\tilde{z}$ has a zero mean, comparing (1.4) and (1.5) implies that the certainty equivalent e of $\tilde{z}$ is equal to minus its its risk premium Π.

A direct consequence of Proposition 1.2 is that the risk premium Π is nonnegative when u is concave, i.e. when she is risk-averse. In Figure 1.2, we measure Π for the risk $(-4000, \frac{1}{2}; 4000, \frac{1}{2})$ for initial wealth $w = 8000$. Notice first that the risk premium is zero when u is linear, and it is nonpositive when u is convex.

One very convenient property of the risk premium is that it is measured in the same units as wealth, e.g. we can measure Sempronius's risk premium in ducats. Although the measure of satisfaction or utility is hard to compare between different individuals—what would it mean to say Sempronius was "happier" than Alexander?—the risk premium is not. We can easily determine whether Sempronius or Alexander is more affected by risk $\tilde{z}$ by comparing their two risk premia.

The risk premium is a complex function of the distribution of $\tilde{z}$, of initial wealth w and of the utility function u. We can estimate the amount that the agent is ready to pay for the elimination of this zero-mean risk by considering small risks. Assume

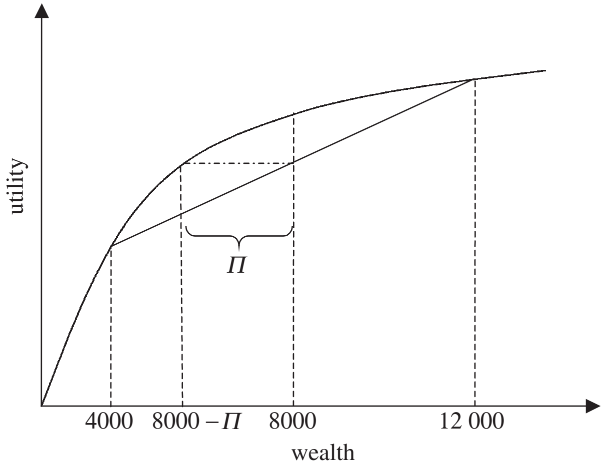

Figure 1.2. Measuring the risk premium P of risk $(-4000, \frac{1}{2}; 4000, \frac{1}{2})$ when initial wealth is $w = 8000$.

that $E\tilde{z} = 0$. Using a second-order and a first-order Taylor approximation for the left-hand side and the right-hand side of equation (1.4), respectively, we obtain that

$$u(w - \Pi) \simeq u(w) - \Pi u'(w)$$

and

$$\begin{aligned} Eu(w + \tilde{z}) &\simeq E[u(w) + zu'(w) + \tfrac{1}{2}\tilde{z}^2 u''(z)] \\ &= u(w) + u'(w)E\tilde{z} + \tfrac{1}{2}u''(w)E\tilde{z}^2 \\ &= u(w) + \tfrac{1}{2}\sigma^2 u''(w), \end{aligned}$$

where $E\tilde{z} = 0$ and $\sigma^2 = E\tilde{z}^2$ is the variance of the outcome of the lottery. Replacing these two approximations in equation (1.4) yields

$$\Pi \simeq \tfrac{1}{2}\sigma^2 A(w), \tag{1.6}$$

where the function A is defined as

$$A(w) = \frac{-u''(w)}{u'(w)}. \tag{1.7}$$

Under risk aversion, function A is positive. It would be zero or negative respectively for a risk-neutral or risk-loving agent. $A(\cdot)$ is hereafter referred to as the degree of absolute risk aversion of the agent. From (1.6), we see that the risk premium associated with risk $\tilde{\varepsilon}$ for an agent with wealth w is approximately equal to one-half the product of the variance of $\tilde{z}$ and the degree of absolute risk aversion of the agent evaluated at w. Equation (1.6) is known as the Arrow–Pratt approximation, as it was developed independently by Arrow (1963) and Pratt (1964).

The cost of risk, as measured by the risk premium, is approximately proportional to the variance of its payoffs. Thus, the variance might appear to be a good measure of the degree of riskiness of a lottery. This observation induced many authors to use a mean–variance decision criterion for modeling behavior under risk. In a mean–variance model, we assume that individual risk attitudes depend only upon the mean and the variance of the underlying risks. However, the validity of these models is dependent on the degree of accuracy of the approximation in (1.6), which can be considered accurate only when the risk is small or in very special cases. In such cases, the mean–variance approach for decisions under risk, which has historically played a very important role in the development of the theory of finance, can be seen as a special case of the expected-utility theory. In most cases however, the risk premium associated with any (large) risk will also depend upon the other moments of the distribution of the risk, not just its mean and variance. For example, it seems intuitive that whether or not $\tilde{x}$ is symmetrically distributed about its mean matters for determining the risk premium. The degree of skewness (i.e. third moment) might very well affect the desirability of a risk. Hence, two risks with the same mean and variance, but one with a distribution that is skewed to the right and the other with a distribution that is skewed to the left, should not be expected to necessarily have the same risk premium. A similar argument can be made about the kurtosis (fourth moment), which is linked to the probability mass in the tails of the distribution.

At this stage, it is worth noting that, at least for small risks, the risk premium increases with the size of the risk proportionately to the square of this size. To see this, let us assume that $\tilde{z} = k\tilde{\varepsilon}$, with $E\tilde{\varepsilon} = 0$. Parameter k can be interpreted as the size of the risk. When k tends to zero, the risk becomes very small. Of course, the risk premium is a function of the size of the risk. We may expect that this function $\Pi(k)$ is increasing in k. We are interested in describing the functional form linking the risk premium Π to the size k of the risk. Because the variance of $\tilde{z}$ equals k^2 times the variance of $\tilde{\varepsilon}$,[4] we obtain that

$$\Pi(k) \simeq \tfrac{1}{2}k^2\sigma_{\tilde{\varepsilon}}^2 A(w),$$

i.e. the risk premium is approximately proportional to the square of the size of the risk. From this observation, we can observe directly that, not only does $\Pi(k)$ approach zero as k approaches zero, but also $\Pi'(0) = 0$. This is an important property of expected-utility theory. At the margin, accepting a small zero-mean risk has *no* effect on the welfare of risk-averse agents! We say that risk aversion is a

[4]The general formula is

$$\operatorname{var}(a\tilde{x} + b\tilde{y}) = a^2 \operatorname{var}(\tilde{x}) + b^2 \operatorname{var}(\tilde{y}) + 2ab \operatorname{cov}(\tilde{x}, \tilde{y}).$$

second-order phenomenon.[5] "In the small," we—the expected-utility maximizers—are all risk neutral.

Proposition 1.3. *If the utility function is differentiable, the risk premium tends to zero as the square of the size of the risk.*

Proof. In the following, we prove formally that $\Pi'(0) = 0$, as suggested by the Arrow–Pratt approximation in our comments above. The relationship between Π and k can be obtained by fully differentiating the equation $Eu(w + k\tilde{\varepsilon}) = u(w - \Pi(k))$ with respect to k. This yields

$$\Pi'(k) = \frac{-E\tilde{\varepsilon}u'(w + k\tilde{\varepsilon})}{u'(w - \Pi(k))}. \tag{1.8}$$

We directly infer that $\Pi'(0) = 0$, since by assumption $E\tilde{\varepsilon} = 0$. □

1.4 Degree of Risk Aversion

Let us consider the following simple decision problem. An agent is offered a take-it-or-leave-it offer to accept lottery $\tilde{z}$ with mean μ and variance σ^2. Of course, the optimal decision is to accept the lottery if

$$Eu(w + \tilde{z}) \geqslant u(w), \tag{1.9}$$

or, equivalently, if the certainty equivalent e of $\tilde{z}$ is positive. In the following, we examine how this decision is affected by a change in the utility function.

Notice at this stage that an increasing linear transformation of u has no effect on the decision maker's choice, and on certainty equivalents. Indeed, consider a function $v(\cdot)$ such that $v(x) = a + bu(x)$ for all x, for some pair of scalars a and b, where $b > 0$. Then, obviously $Ev(w + \tilde{z}) \geqslant v(w)$ yields exactly the same restrictions on the distribution of $\tilde{z}$ as condition (1.9). The same analysis can be done on equation (1.5) defining certainty equivalents. The neutrality of certainty equivalents to linear transformations of the utility function can be verified in the case of small risks by using the Arrow–Pratt approximation. If $v \equiv a + bu$, it is obvious that

$$A(x) = \frac{-v''(x)}{v'(x)} = \frac{-bu''(x)}{bu'(x)} = \frac{-u''(x)}{u'(x)}$$

for all x. Thus, by (1.6), risk premia for small risks are not affected by the linear transformation. Because the certainty equivalent equals the mean payoff of the risk minus the risk premium, the same neutrality property holds for certainty equivalents.

[5]This property in general models, not restricted to expected utility, is called "second-order risk aversion." Within the expected-utility model, this property relies on the assumption that the utility function is differentiable.

Limiting the analysis to small risks, we see from this analysis that agents with a larger absolute risk aversion $A(w)$ will be more reluctant to accept small risks. The minimum expected payoff that makes the risk acceptable for them will be larger. This is why we say that A is a measure of the degree of risk aversion of the decision maker. From a more technical viewpoint, $A = -u''/u'$ is a measure of the degree of concavity of the utility function. It measures the speed at which marginal utility is decreasing.

We are now interested in extending these observations to any risk, not only small risks. We consider the following definition for comparative risk aversion.

Definition 1.4. Suppose that agents u and v have the same wealth w, which is arbitrary. An agent v is more risk-averse than another agent u with the same initial wealth if any risk that is undesirable for agent u is also undesirable for agent v. In other words, the risk premium of any risk is larger for agent v than for agent u.

This must be true independently of the common initial wealth level w of the two agents. If this definition were restricted to small risks, we know from the above analysis that this would be equivalent to requiring that

$$A_v(w) = \frac{-v''(w)}{v'(w)} \geqslant \frac{-u''(w)}{u'(w)} = A_u(w),$$

for all w. If limited to small risks, v is more risk-averse than u if function A_v is uniformly larger than A_u. We say in this case that v is more concave than u in the sense of Arrow–Pratt. It is important to observe that this is equivalent to the condition that v is a *concave* transformation of u, i.e. that there exists an increasing and concave function ϕ such that $v(w) = \phi(u(w))$ for all w. Indeed, we have that $v'(w) = \phi'(u(w))u'(w)$ and

$$v''(w) = \phi''(u(w))(u'(w))^2 + \phi'(u(w))u''(w),$$

which implies that

$$A_v(w) = A_u(w) + \frac{-\phi''(u(w))u'(w)}{\phi'(u(w))}.$$

Thus, A_v is uniformly larger than A_u if and only if ϕ is concave. This is equivalent to requiring that A_v be uniformly larger than A_u or that v be a concave transformation of u. It yields that agent v values small risks less than agent u. Do we need to impose more restrictions to guarantee that agent v values any risk less than agent u, i.e. that v is more risk-averse than u? The following proposition, which is due to Pratt (1964), indicates that no additional restriction is required.

Proposition 1.5. *The following three conditions are equivalent.*

(a) *Agent v is more risk-averse than agent u, i.e. the risk premium of any risk is larger for agent v than for agent u.*

(b) *For all* w, $A_v(w) \geqslant A_u(w)$.

(c) *Function* v *is a concave transformation of function* u : $\exists \phi(\cdot)$ *with* $\phi' > 0$ *and* $\phi'' \leqslant 0$ *such that* $v(w) = \phi(u(w))$ *for all* w.

Proof. We have already shown that (b) and (c) are equivalent. That (a) implies (b) follows directly from the Arrow–Pratt approximation. We now prove that (c) implies (a). Consider any lottery $\tilde{z}$. Let Π_u and Π_v denote the risk premium for zero-mean lottery $\tilde{z}$ of agent u and agent v, respectively. By definition, we have that

$$v(w - \Pi_v) = Ev(w + \tilde{z}) = E\phi(u(w + \tilde{z})).$$

Define random variable $\tilde{y}$ as $\tilde{y} = u(w + \tilde{z})$. Because ϕ is concave, $E\phi(\tilde{y})$ is smaller than $\phi(E\tilde{y})$ by Jensen's inequality. It thus follows that

$$v(w - \Pi_v) \leqslant \phi(Eu(w + \tilde{z})) = \phi(u(w - \Pi_u)) = v(w - \Pi_u).$$

Because v is increasing, this implies that Π_v is larger than Π_u. □

In the case of small risks, the only thing that we need to know to determine whether a risk is desirable is the degree of concavity of u locally at the current wealth level w. For larger risks, the proposition above shows that we need to know much more to take a decision. Namely, we need to know the degree of concavity of u at all wealth levels. The degree of concavity must be increased at all wealth levels to guarantee that a change in u makes the decision maker more reluctant to accept risks. If v is locally more concave at some wealth levels and is less concave at other wealth levels, the comparative analysis is intrinsically ambiguous.

To illustrate the proposition, let us go back to the example of Sempronius's single ship yielding outcome $\tilde{z} = (0, \frac{1}{2}; 8000, \frac{1}{2})$, with a initial wealth $w_0 = 4000$ ducats. If Sempronius's utility function is $u(w) = \sqrt{w}$, his certainty equivalent of $\tilde{z}$ equals $e_u = 3464.1$, since

$$\tfrac{1}{2}\sqrt{4000} + \tfrac{1}{2}\sqrt{12\,000} = 86.395 = \sqrt{7464.1}$$

Alternatively, suppose that Sempronius's utility function is $v(w) = \ln(w)$, which is also increasing and concave. It is easy to check that v is more concave than u in the sense of Arrow–Pratt. Indeed, these functions yield

$$A_v(w) = \frac{1}{w} \geqslant \frac{1}{2w} = A_u(w)$$

for all w. From the above proposition, this change in utility should reduce the certainty equivalent of any risk. In the case of $w_0 = 4000$ and $\tilde{z} \sim (0, \frac{1}{2}; 8000, \frac{1}{2})$, the certainty equivalent of $\tilde{z}$ under v equals $e_v = 2928.5$, since

$$\tfrac{1}{2}\ln(4000) + \tfrac{1}{2}\ln(12\,000) = 8.8434 = \ln(6928.5).$$

Thus, e_v is smaller than e_u. Notice that the risk premium $\Pi_v = 1071.5$ under v is approximately twice the risk premium $\Pi_u = 535.9$. This was predicted by the Arrow–Pratt approximation, since A_v is equal to $2A_u$.

1.5 Decreasing Absolute Risk Aversion and Prudence

We have seen that risk aversion is driven by the fact that one's marginal utility is decreasing with wealth. In this section, we examine another question related to increasing wealth. Namely, we are interested in determining how the risk premium for a given zero-mean risk $\tilde{z}$ is affected by a change in initial wealth w. Arrow argued that intuition implies that wealthier people are generally less willing to pay for the elimination of fixed risk. A lottery to gain or lose 100 with equal probability is potentially life-threatening for an agent with initial wealth $w = 101$, whereas it is essentially trivial for an agent with wealth $w = 1\,000\,000$. The former should be ready to pay more than the latter for the elimination of risk. We can check that this property holds for the square-root utility function, with $\Pi = 43.4$ when $w = 101$ and $\Pi = 0.0025$ when $w = 1\,000\,000$. If wealth is measured in euros, the individual would be willing to pay over 43 euros to avoid the risk when wealth is $w = 101$, whereas the same individual would not even pay one euro cent to get rid of this risk when wealth is one million euros! In the following, we characterize the set of utility functions that have this property.

The risk premium $\Pi = \pi(w)$ as a function of initial wealth w can be evaluated by solving

$$Eu(w + \tilde{z}) = u(w - \pi(w)) \tag{1.10}$$

for all w. Fully differentiating (1.10) with respect to w yields

$$Eu'(w + \tilde{z}) = (1 - \pi'(w))u'(w - \pi),$$

or, equivalently,

$$\pi'(w) = \frac{u'(w - \pi) - Eu'(w + \tilde{z})}{u'(w - \pi)}. \tag{1.11}$$

Thus, the risk premium is decreasing with wealth if and only if

$$Ev(w + \tilde{z}) \leqslant v(w - \pi(w)), \tag{1.12}$$

where function $v \equiv -u'$ is defined as minus the derivative of function u. Because the function v is increasing, we can also interpret it as *another* utility function. Condition (1.12) then just states that the risk premium of agent v is larger than the risk premium π of agent u. From Proposition 1.5, this is true if and only if v is more concave than u in the sense of Arrow–Pratt, that is, if $-u'$ is a concave transformation of u. For this utility v, the measure of absolute risk aversion is $A_v = A_{-u'} = -u'''/u''$. This measure has several uses, which will be made clearer later in this book. For this reason, without justifying the terminology at this stage, we will define $P(w) = -u'''(w)/u''(w)$ as the degree of absolute prudence of the agent with utility u. It follows from (1.12) that $-u'$ is more concave than u if and only if

$$P(w) \geqslant A(w)$$

for all w. We conclude that condition $P \geqslant A$ uniformly is necessary and sufficient to guarantee that an increase in wealth reduces risk premia. Because

$$A'(w) = A(w)[A(w) - P(w)],$$

condition $P \geqslant A$ is equivalent to the condition $A' \leqslant 0$. We obtain the following proposition.

Proposition 1.6. *The risk premium associated to any risk $\tilde{z}$ is decreasing in wealth if and only if absolute risk aversion is decreasing; or equivalently if and only if prudence is uniformly larger than absolute risk aversion.*

Observe that the utility function $u(w) = \sqrt{w}$ satisfies this condition. Indeed, we have $A_u(w) = \frac{1}{2}w^{-1}$, which is decreasing. This can alternatively checked by observing that $v(w) = -\frac{1}{2}w^{-1/2}$ and $A_v(w) = P_u(w) = 1.5w^{-1}$, which is uniformly larger than $A_u(w)$. Notice that Decreasing Absolute Risk Aversion (DARA) requires that the third derivative of the utility function be positive. Otherwise, prudence would be negative, which would imply that $P < A$: a condition that implies that absolute risk aversion would be increasing in wealth. Thus, DARA, a very intuitive condition, requires the necessary (but not sufficient) condition that u''' be positive, or that marginal utility be convex.

1.6 Relative Risk Aversion

Absolute risk aversion is the rate of decay for marginal utility. More particularly, absolute risk aversion measures the rate at which marginal utility decreases when wealth is increased *by one euro*.[6] If the monetary unit were the dollar, absolute risk aversion would be a different number. In other words, the index of absolute risk aversion is not unit free, as it is measured per euro (per dollar, or per yen).

Economists often prefer unit-free measurements of sensitivity. To this end, define the index of *relative* risk aversion R as the rate at which marginal utility decreases

---

[6] In general, the *growth rate* for a function $f(x)$ is defined as

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} \cdot \frac{1}{f(x)}.$$

Since marginal utility $u'(x)$ declines in wealth, its growth rate is negative. The absolute value of this negative growth rate, which is the measure of absolute risk aversion, is called the *decay rate*.

when wealth is increased *by one percent*. In terms of standard economic theory, this measure is simply the wealth-elasticity of marginal utility. It can be computed as

$$R(w) = -\frac{\mathrm{d}u'(w)/u'(w)}{\mathrm{d}w/w} = \frac{-wu''(w)}{u'(w)} = wA(w). \tag{1.13}$$

Note that the measure of relative risk aversion is simply the product of wealth and absolute risk aversion.

The (absolute) risk premium and the index of absolute risk aversion are linked by the Arrow–Pratt approximation and by Propositions 1.5 and 1.6. We can develop analogous kinds of results for relative risk aversion. Suppose that your initial wealth w is invested in a portfolio whose return $\tilde{z}$ over the period is uncertain. Let us assume that $E\tilde{z} = 0$. Which share of your initial wealth are you ready to pay to get rid of this proportional risk? The solution to this problem is referred to as the relative risk premium $\hat{\Pi}$. This measure also is a unit-free measure, unlike the absolute risk premium, which is measured in euros. It is defined implicitly via the following equation:

$$Eu(w(1+\tilde{z})) = u(w(1-\hat{\Pi})). \tag{1.14}$$

Obviously, the relative risk premium and the absolute risk premium are equal if we normalize initial wealth to unity. More generally, the relative risk premium for proportional risk $\tilde{z}$ equals the absolute risk premium for absolute risk $w\tilde{z}$, divided by initial wealth w: $\hat{\Pi}(\tilde{z}) = \Pi(w\tilde{z})/w$. From this observation, we obtain the fact that, if agent v is more risk-averse than agent u with the same initial wealth, then agent v will be ready to pay a larger share of his wealth than agent u to insure against a given proportional risk $\tilde{z}$. Moreover, if σ^2 denotes the variance of $\tilde{z}$, then the variance of $w\tilde{z}$ equals $w^2\sigma^2$. Using the Arrow–Pratt approximation thus yields

$$\hat{\Pi}(\tilde{z}) = \frac{\Pi(w\tilde{z})}{w} \simeq \frac{\frac{1}{2}w^2\sigma^2 A(w)}{w} = \tfrac{1}{2}\sigma^2 R(w). \tag{1.15}$$

The relative risk premium is approximately equal to half of the variance of the proportional risk times the index of relative risk aversion. This can be used to establish a range for acceptable degrees of risk aversion. Suppose that one's wealth is subject to a risk of a gain or loss of 20% with equal probability. What is the range that one would find reasonable for the share of wealth Π that one would be ready to pay to get rid of this zero-mean risk? From our various experiments in class, we found that most people would be ready to pay between 2% and 8% of their wealth. Because risk $\tilde{z}$ in this experiment has a variance of $0.5(0.2)^2 + 0.5(-0.2)^2 = 0.04$, using approximation (1.15) yields a range for relative risk aversion between 1 and 4. This information will be useful later in this book.

There is no definitive argument for or against decreasing relative risk aversion. Arrow originally conjectured that relative risk aversion is likely to be constant, or perhaps increasing, although he stated that the intuition was not as clear as was the intuition for decreasing absolute risk aversion. Since then, numerous empirical studies have offered conflicting results. We might also try to examine this question by introspection. If your wealth would increase, would you want to devote a larger or a smaller share of your wealth to get rid of a given zero-mean proportional risk? For example, what would you pay to avoid the risk of gaining or losing 20% of your wealth, each with an equal probability? If the share is decreasing with wealth, you have decreasing relative risk aversion. There are two contradictory effects here that need to be considered. On the one hand, under the intuitive DARA assumption, becoming wealthier also means becoming less risk-averse. This effect tends to reduce Π. But, on the other hand, becoming wealthier also means facing a larger absolute risk $w\tilde{z}$. This effect tends to raise Π. There is no clear intuition as to whether the first effect or the second effect will dominate. For example, many of the classic models in macroeconomics are based on relative risk aversion being constant over all wealth levels, which is implicitly assuming that our two effects exactly cancel each other out. Of course, there also is no *a priori* reason to believe that the dominant effect will not change over various wealth levels. For instance, some recent empirical evidence indicates a possible "U-shape" for relative risk aversion, with R decreasing at low wealth levels, then leveling off somewhat before increasing at higher wealth levels.

1.7 Some Classical Utility Functions

As already noted above, expected-utility (EU) theory has many proponents and many detractors. In Chapter 13, we examine some generalizations of the EU criterion that satisfy those who find expected utility too restrictive. But researchers in both economics and finance have long considered—and most of them still do—EU theory as an acceptable paradigm for decision making under uncertainty. Indeed, EU theory has a long and prominent place in the development of decision making under uncertainty. Even detractors of the theory use EU as a standard by which to compare alternative theories. Moreover, many of the models in which EU theory has been applied can be modified, often yielding better results.

Whereas the current trend is to generalize the EU model, researchers often *restrict* EU criterion by considering a specific subset of utility functions. This is done to obtain tractable solutions to many problems. It is important to note the implications that derive from the choice of a particular utility function. Some results in the literature may be robust enough to apply for all risk-averse preferences, while others might be restricted to applying only for a narrow class of preferences. In this section, we examine several particular types of utility functions that are often encountered in the economics and the finance literature. Remember that utility is unique only up to a linear transformation.

Historically, much of the theory of finance was developed during the 1960s by considering the subset of utility functions that are quadratic of the form

$$u(w) = aw - \tfrac{1}{2}w^2, \quad \text{for } w \leqslant a.$$

Note that the domain of wealth on which u is defined comes from the necessary requirement that u be nondecreasing, which is true only if w is smaller than a. This set of functions is useful because the EU generated by any distribution of final wealth is a function of only the first two moments of this distribution:

$$Eu(\tilde{w}) = aE\tilde{w} - \tfrac{1}{2}E\tilde{w}^2.$$

Therefore, in this case, the EU theory simplifies to a mean–variance approach to decision making under uncertainty. However, as already discussed, it is very hard to believe that preferences among different lotteries be determined only by the mean and variance of these lotteries.

Above wealth level a, marginal utility becomes negative. Since quadratic utility is decreasing in wealth for $w > a$, many people might feel this is not appropriate as a utility function. However, it is important to remember that we are trying to model human behavior with mathematical models. For example, if the quadratic utility function models your behavior quite well with $a = 100$ million euros, is it really a problem that this function declines for higher wealth levels? The point is that the quadratic utility might work well for more realistic wealth levels, and if it does, we should not be overly concerned about its properties at unrealistically high wealth levels. However, the quadratic utility function has another property that is more problematic. Namely, the quadratic utility functions exhibit increasing absolute risk aversion:

$$A(w) = \frac{1}{a - w} \Rightarrow A'(w) = \frac{1}{(a - w)^2} > 0.$$

For this reason, quadratic utility functions are not as in fashion anymore.

A second set of classical utility functions is the set of so-called constant-absolute-risk-aversion (CARA) utility functions, which are exponential functions characterized by

$$u(w) = -\frac{\exp(-aw)}{a},$$

where a is some positive scalar. The domain of these functions is the real line. The distinguishing feature of these utility functions is that they exhibit constant absolute risk aversion, with $A(w) = a$ for all w. It can be shown that the Arrow–Pratt approximation is exact when u is exponential and $\tilde{w}$ is normally distributed with mean μ and variance σ^2. Indeed, we can take expectations to see that

$$\begin{aligned} Eu(\tilde{w}) \\ &= \frac{-1}{\sigma a\sqrt{2\pi}} \int \exp(-aw) \exp\left(- \frac{(w-\mu)^2}{2\sigma^2} \right) \mathrm{d}w \\ &= -\frac{1}{a} \exp(-a(\mu - \tfrac{1}{2}a\sigma^2)) \left[\frac{1}{\sigma\sqrt{2\pi}} \int \exp\left(- \frac{(w - (\mu - \frac{1}{2}a\sigma^2))^2}{2\sigma^2} \right) \mathrm{d}w \right] \\ &= -\frac{1}{a} \exp(-a(\mu - \tfrac{1}{2}a\sigma^2)) = u(\mu - \tfrac{1}{2}a\sigma^2). \end{aligned} \tag{1.16}$$

The third equality comes from the fact that the bracketed term is the integral of the density of the normal distribution $N(\mu - \frac{1}{2}a\sigma^2, \sigma)$, which must be equal to unity. Thus, the risk premium is indeed equal to $\frac{1}{2}\sigma^2 A(w)$. In this very specific case, we obtain that the Arrow–Pratt approximation is exact. The fact that risk aversion is constant is often useful in analyzing choices among several alternatives. As we will see later, this assumption eliminates the income effect when dealing with decisions to be made about a risk whose size is invariant to changes in wealth. However, this is often also the main criticism of the CARA utility, since absolute risk aversion is constant rather than decreasing.

Finally, one set of preferences that has been by far the most used in the literature is the set of power utility functions. Researchers in finance and in macroeconomics are so accustomed to this restriction that many of them do not even mention it anymore when they present their results. Suppose that

$$u(w) = \frac{w^{1-\gamma}}{1-\gamma} \quad \text{for } w > 0.$$

The scalar γ is chosen so that $\gamma > 0$, $\gamma \neq 1$. It is easy to show that γ equals the degree of relative risk aversion, since $A(w) = \gamma/w$ and $R(w) = \gamma$ for all w. Thus, this set exhibits decreasing absolute risk aversion and constant relative risk aversion, which are two reasonable assumptions. For this reason, these utility functions are called the constant-relative-risk-aversion (CRRA) class of preferences. Notice that our definition does not allow for $\gamma = 1$. However, it is straightforward to show that function $u(w) = \ln(w)$ satisfies the property that $R(w) = 1$ for all w. Thus, the set of all CRRA utility functions is completely defined by[7]

$$u(w) = \begin{cases} \dfrac{w^{1-\gamma}}{1-\gamma} & \text{for } \gamma \geqslant 0,\ \gamma \neq 1, \\ \ln(w) & \text{for } \gamma = 1. \end{cases} \tag{1.17}$$

[7] We can also show that $u(w) = \ln(w)$ as a limiting case of the power utility function. To this end, rewrite the power utility function, using a linear transformation, as

$$u(w) = \frac{1}{1-\gamma}(w^{1-\gamma} - 1).$$

As we will see later in this book, this class of utility functions eliminates any income effects when making decisions about risks whose size is proportional to one's level of wealth. For example, the relative risk premium $\hat{\Pi}$ defined by equation (1.14) is independent of wealth w in this case. The assumption that relative risk aversion is constant enormously simplifies many of the problems often encountered in macroeconomics and finance.

1.8 Bibliographical References, Extensions and Exercises

The contribution by Pratt (1964) basically opened and closed the field covered in this chapter. It is, however, fair to mention that the measure of absolute risk aversion has been discovered independently by Arrow (1963) and de Finetti (1952). The paper by de Finetti was written in Italian and even today is not given the attention it deserves. The paper by Pratt is by far the most advanced in defining the notions of an increase in risk aversion and of decreasing absolute risk aversion. The orders of risk aversion are introduced by Segal and Spivak (1990).

Ross (1981) challenged the idea that $A = -u''/u'$ is a good measure of the degree of risk aversion of an agent. Kihlstrom, Romer and Williams (1981) and Nachman (1982) showed that if initial wealth is uncertain, it is not true that an agent v, who is more risk-averse than another agent u in the sense of Arrow–Pratt, will be ready to pay more to get rid of another risk. Ross (1981) characterized the conditions on u and v that imply that $\Pi_v \geqslant \Pi_u$ even when initial wealth is uncertain and potentially correlated with the risk under scrutiny. These conditions are of course stronger than $A_v \geqslant A_u$.

There is much contradictory empirical evidence on the shape of relative risk aversion as a function of wealth. Many authors have empirically estimated R, assuming that we have CRRA. Fewer authors have examined whether R might be increasing or decreasing in wealth. A good summary of many of these results appears in Ait-Sahalia and Lo (2000).

Chapter Bibliography

Ait-Sahalia, Y. and A. W. Lo. 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94:9–51.

Arrow, K. J. 1963. Liquidity preference. Lecture VI in *Lecture Notes for Economics 285, The Economics of Uncertainty*, pp. 33-53, undated, Stanford University.

———. 1965. *Yrjo Jahnsson lecture notes*, Helsinki. (Reprinted in Arrow 1971).

———. 1971. *Essays in the theory of risk bearing*. Chicago: Markham Publishing Co.

Taking the limit as $\gamma \to 1$ and applying L'Hôpital's rule, we obtain

$$\lim_{\gamma \to 1} u(w) = \lim_{\gamma \to 1} \frac{-(w^{1-\gamma})\ln(w)}{-1} = \ln(w).$$

Bernoulli, D. 1954. Exposition of a new theory on the measurement of risk. (English Transl. by Louise Sommer.) *Econometrica* 22: 23–36.

Bernstein, P. L. 1998. *Against the Gods*. Wiley.

de Finetti, B. 1952. Sulla preferibilita. *Giornale Degli Economisti E Annali Di Economia* 11:685–709.

Kihlstrom, R., D. Romer, and S. Williams. 1981. Risk aversion with random initial wealth. *Econometrica* 49:911–920.

Nachman, D. C. 1982. Preservation of 'more risk averse' under expectations. *Journal of Economic Theory* 28:361–368.

Pratt, J. 1964. Risk aversion in the small and in the large. *Econometrica* 32:122–136.

Ross, S. A. 1981. Some stronger measures of risk aversion in the small and in the large with applications. *Econometrica* 3:621–638.

Segal, U. and A. Spivak. 1990. First order versus second order risk aversion. *Journal of Economic Theory* 51:111–125.

Yaari, M. E. 1987. The dual theory of choice under risk. *Econometrica* 55:95–115.

Exercises

(1.1) An individual has the following utility function:

$$u(w) = w^{1/2}.$$

Her initial wealth is 10 and she faces the lottery $\tilde{X} : (-6, \frac{1}{2}; +6, \frac{1}{2})$.

(a) Compute the exact value of the certainty equivalent and of the risk premium.

(b) Apply Pratt's formula to obtain an approximation of the risk premium.

(c) Show that with such a utility function absolute risk aversion is decreasing in wealth while relative risk aversion is constant.

(d) If the utility function becomes

$$v(w) = w^{1/4},$$

answer again part (a). Are you surprised by the changes in the certainty equivalent and in the risk premium? Relate this change to the notion of 'more risk averse' (i.e. express $v(w)$ as a concave transformation of $u(w)$).

(e) If the risk becomes $\tilde{Y} : (-3, \frac{1}{2}; +3, \frac{1}{2})$, compute the new risk premium as approximated by Pratt's formula. Why is the approximated risk premium four times smaller than the risk premium for $\tilde{X}$?

(1.2) As in the previous exercise, consider an initial wealth of 10 and the lottery $\tilde{X}$. Assume now that the utility is:

$$u = \begin{cases} w & \text{for } w \leqslant 10, \\ \frac{1}{2}w + 5 & \text{for } w \geqslant 10. \end{cases} \tag{1.18}$$

(a) Draw the utility function. Is it globally concave?

(b) Compute the certainty equivalent and the risk premium attached to $\tilde{X}$.

(c) Can you apply the Arrow–Pratt approximation? Why?

(d) Consider now the lottery $\tilde{Y}$ defined in exercise 1.1. Compute the risk premium attached to $\tilde{Y}$. Is it smaller than for $\tilde{X}$? Why?

(e) Answer (b) and (d) above if the individual has an initial wealth of 20. How do the risk premia for $\tilde{X}$ and $\tilde{Y}$ compare?

(1.3) Let $u = w^2$ for $w \geqslant 0$.

(a) Compute the exact risk premium if initial wealth is 4 and if a decision maker faces the lottery $(-2, \frac{1}{2}; +2, \frac{1}{2})$. Explain why the risk premium is negative.

(b) If the utility function becomes $v = w^4$, what happens to the risk premium? Show that v is a convex transformation of u.

(1.4) Let $u = \ln w$.

(a) Does this utility function exhibit the DARA property?

(b) Compute $-u'''/u''$ and compare it with $-u''/u'$.

(e) Prove that $-u'(w)$ is a concave transformation of $u(w)$ (hint: use Pratt's theorem).

(1.5) Consider the family of exponential utility functions

$$u = \frac{1 - \exp(-aw)}{a}.$$

(a) Show that a is the degree of absolute risk aversion.

(b) Show that u becomes linear in w when a tends to zero (hint: use L'Hôpital's rule).

(c) Consider lottery $\tilde{x}$ with positive and negative payoffs. Determine the value of $Eu(\tilde{x})$ when a tends to infinity.

2
The Measures of Risk

In Chapter 1, we defined the concept of risk aversion by considering the effect of the introduction of a zero-mean risk on welfare. That is, we assumed that the initial environment of the consumer was risk free. Our only conclusion from risk aversion was that the individual preferred no risk to a zero-mean risk. But what about choices among different zero-mean risks? For example, recall from the previous chapter that we might interpret Sempronius's situation as facing the risk $(-4000, \frac{1}{2}; 4000, \frac{1}{2})$ with initial wealth $w = 8000$. The alternative, using two separate ships, can be thought of as facing the risk $(-4000, \frac{1}{4}; 0, \frac{1}{2}; +4000, \frac{1}{4})$ from the same initial wealth. We argued that the second alternative seemed more valuable, in some sense. In this chapter, we will examine this question more closely and consider the comparison of such competing risks.

If one could know the utility function of the agent, ranking lotteries would be easy. For example, let us compare two different wealth prospects (i.e. two distributions of final wealth) $\tilde{w}_1$ and $\tilde{w}_2$. The first is preferred to the second by an agent with utility function u if $Eu(\tilde{w}_1) \geqslant Eu(\tilde{w}_2)$. This preference order, which is specific to a single utility function u, is complete in the sense that, for any pair $(\tilde{w}_1, \tilde{w}_2)$, $\tilde{w}_1$ is preferred to $\tilde{w}_2$, or $\tilde{w}_2$ is preferred to $\tilde{w}_1$, or we are indifferent to both wealth prospects. In this chapter, we consider several relatively weak restrictions on preferences; for example, that "agents are risk-averse" or that "agents are prudent." We want to find restrictions on the change in risk from $\tilde{w}_1$ to $\tilde{w}_2$ that are unanimously disliked by the group of agents under scrutiny. In that case, we say that $\tilde{w}_1$ dominates $\tilde{w}_2$ for this class of utility functions. As soon as the class is not limited to a single utility function, these preference orders are incomplete in the sense that it is not true that, for any pair of lotteries, one must necessarily dominate the other. Some people in the group may prefer the first, whereas other members in the group may prefer the second. Imposing unanimity in the group is a very strong constraint on the change in risk. Considering a larger group makes the constraint of unanimity stronger. On the other hand, if among our group we do find unanimity that $\tilde{w}_1$ dominates $\tilde{w}_2$, it allows us to choose $\tilde{w}_1$ over $\tilde{w}_2$ when these are our only two options.

The theory of stochastic dominance looks at certain statistical properties of the distributions of $\tilde{w}_1$ and $\tilde{w}_2$, which allow us to infer unanimous agreement for certain classes of preferences. This is important not only for our understanding of individual behavior, but also for decision making designed to benefit a group, such as a corporate manager making decisions on behalf of the company's shareholders. In this book, we will consider basically three stochastic orders. In Section 2.1, we first consider the most natural set of utility functions from what we have seen in Chapter 1. We consider in that section the set of all risk-averse agents. It generates the concept of an increase in risk that was first examined by Rothschild and Stiglitz (1970). In Section 2.2, we focus on the set of prudent agents, which yields the concept of an increase in downside risk introduced by Menezes, Geiss and Tressler (1980). Finally, in Section 2.3, we assume only that agents have an increasing utility function. The corresponding stochastic order is called "first-order stochastic dominance."

2.1 Increases in Risk

In this section, we characterize the changes in risk that make all risk-averse agents worse off. We focus the analysis on changes in risk which preserve the expected outcome, i.e. mean-preserving changes in risk. These changes are called "increases in risk." There are at least three equivalent ways to define them.

2.1.1 Adding Noise

Consider the binary lottery faced by Sempronius using a single ship. This may be written as $\tilde{w}_1 \sim (4000, \frac{1}{2}; 12\,000, \frac{1}{2})$. In the low wealth state, the ship is lost, whereas in the high wealth state, the ship succeeds in bringing the spices safely to the harbor. Let us assume that the ship contains 8000 pounds of spices, which will be sold at a unit price of one ducat. This environment generates a distribution $\tilde{w}_1$ for Sempronius's final wealth.

Suppose alternatively that the price at which the spices will be sold at the harbor is unknown at the time that the ship leaves the East Indies. More precisely, let us suppose that the unit price will be either 0.5 ducats or 1.5 ducats with equal probabilities. In this alternative environment, the final wealth is still 4000 in the case of the ship being sunk. But, conditional on the ship's arriving safely in Europe, Sempronius's final wealth will be either 8000 or 16 000 with equal probabilities, or $12\,000 + \tilde{\varepsilon}$, with $\tilde{\varepsilon} \sim (-4000, \frac{1}{2}, 4000, \frac{1}{2})$. Because $E\tilde{\varepsilon} = 0$, the price uncertainty adds a zero-mean noise to Sempronius's final wealth, conditional upon the no-loss state. *Ex ante*, the agent faces an uncertain wealth distributed as $\tilde{w}_2 \sim (4000, \frac{1}{2}; 12\,000 + \tilde{\varepsilon}, \frac{1}{2})$. This situation describes what is called a compound lottery, i.e. a lottery for which some of the outcomes are themselves lotteries.

Intuition suggests that Sempronius should dislike this additional uncertainty. The reader can check that this is indeed the case. For example, if Sempronius's utility

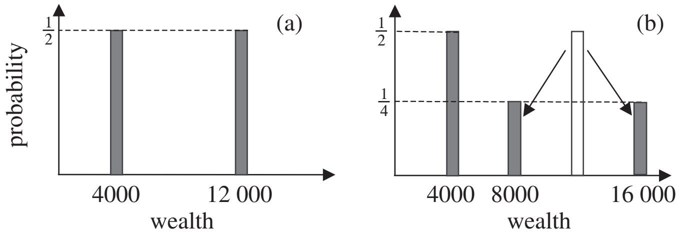

Figure 2.1. Transfer of probability mass to describe an increase in risk.

function is a square root, we have that

$$Eu(\tilde{w}_1) = \tfrac{1}{2}\sqrt{4000} + \tfrac{1}{2}\sqrt{12\,000} = 86.395,$$

whereas

$$Eu(\tilde{w}_2) = \tfrac{1}{2}\sqrt{4000} + \tfrac{1}{2}\Big[\tfrac{1}{2}\sqrt{8000} + \tfrac{1}{2}\sqrt{16\,000}\Big] = 85.606 < 86.395.$$

The same qualitative result would hold if Sempronius would have another concave utility function. In fact, adding a zero-mean noise conditional upon some specific state always reduces the EU of risk-averse agents, as we now show.

To keep the presentation relatively simple, suppose that $\tilde{w}_1$ can take n different possible values $\omega_1, \omega_2, \dots, \omega_n$. Let p_s denote the probability that $\tilde{w}_1$ takes value ω_s. Suppose that the alternative wealth distribution $\tilde{w}_2$ be obtained by compounding $\tilde{w}_1$ with zero-mean noises $\tilde{\varepsilon}_s$ for the different outcomes ω_s of $\tilde{w}_1$. This means that each outcome ω_s of $\tilde{w}_1$ is replaced by $\omega_s + \tilde{\varepsilon}_s$ with $E\tilde{\varepsilon}_s = 0$:

$$\tilde{w}_2 = \tilde{w}_1 + \tilde{\varepsilon} \quad \text{with } E[\tilde{\varepsilon} \mid \tilde{w}_1 = \omega_s] = E[\tilde{\varepsilon}_s] = 0.$$

The price uncertainty presented above is an example of this technique of adding noises to each possible outcome of the initial distribution of wealth.

With this notation, it is easy to show that any such alternative lottery $\tilde{w}_2$ makes all risk-averse agents worse off. Because $E\tilde{\varepsilon}_s$ is zero, risk aversion implies that $Eu(\omega_s + \tilde{\varepsilon}_s) \leqslant u(\omega_s)$. It follows that

$$Eu(\tilde{w}_2) = \sum_{s=1}^{n} p_s Eu(\omega_s + \tilde{\varepsilon}_s) \leqslant \sum_{s=1}^{n} p_s u(\omega_s) = Eu(\tilde{w}_1).$$

All risk-averse agents dislike adding zero-mean noises to the possible outcomes of their wealth.

2.1.2 Mean-Preserving Spreads in Probability

The existence of price uncertainty in the situation faced by Sempronius can alternatively be seen as transferring probability masses. In Figure 2.1(a), we represent the probability distribution in the absence of price uncertainty. Figure 2.1(b) describes the probability distribution when the price uncertainty is taken into account. We see that adding noise $\tilde{\varepsilon} \sim (-4000, \frac{1}{2}, 4000, \frac{1}{2})$ is equivalent to transfering half of the $\frac{1}{2}$-probability mass at 12 000 to 8000, and the remaining of the probability mass at 12 000 to 16 000. By doing this, we do not modify the center of gravity of the probability distribution, i.e. we preserve the mean. In short, we construct what is called a "mean-preserving spread" of the probability distribution.

Let $f_i(w)$ denote the probability mass of $\tilde{w}_i$ at w. In the case of a continuous distribution, $f_i(\cdot)$ is the probability density of $\tilde{w}_i$. The following definition formalizes the concept of a mean-preserving spread, which is an operation consisting of the partial removal of probability mass from some interval I in order to transfer it outside this interval.

Definition 2.1. $\tilde{w}_2$ is a mean-preserving spread (MPS) of $\tilde{w}_1$ if

1. $E\tilde{w}_2 = E\tilde{w}_1$, and
2. there exists an interval I such that $f_2(w) \leqslant f_1(w)$ for all w in I, and $f_2(w) \geqslant f_1(w)$ for all w outside I.

Adding noise or constructing a sequence of MPS's are obviously two equivalent ways to increase risk. In some circumstances, it is easier to use one representation than the other. For example, let us compare distribution $\tilde{w}_1 \sim (4000, \frac{1}{2}; 12\,000, \frac{1}{2})$ to distribution $\tilde{w}_2 \sim (2000, \frac{1}{2}; 14\,000, \frac{1}{2})$. Obviously, the two distributions have the same mean, and the second is obtained by transferring some probability mass from interval $I = [4000, 12\,000]$ outside I. Thus, $\tilde{w}_2$ is an increase in risk of $\tilde{w}_1$. It must be the case that $\tilde{w}_2$ is obtained from $\tilde{w}_1$ by adding some noise $\tilde{\varepsilon}_1$ to outcome 4000 and another noise $\tilde{\varepsilon}_2$ to outcome 12 000 of $\tilde{w}_1$. The reader may easily verify that indeed defining

$$\tilde{\varepsilon}_1 \sim (-2000, \tfrac{5}{6}; 10\,000, \tfrac{1}{6}) \quad \text{and} \quad \tilde{\varepsilon}_2 \sim (2000, \tfrac{5}{6}; -10\,000, \tfrac{1}{6})$$

does the job.

It is often useful to translate the definition of a mean-preserving spread into a condition on the *cumulative* distribution functions of $\tilde{w}_1$ and $\tilde{w}_2$. Let $F_i(w)$ denote the probability that $\tilde{w}_i$ be no greater than w. That is, define

$$F_i(w) = \sum_{s \mid \omega_s \leqslant w} f_i(s)$$

in the discrete case, and

$$F_i(w) = \int^{w} f_i(s)\, \mathrm{d}s$$

in the continuous case. In the latter case, the density function f_i is simply the derivative of F_i. To keep the level of technicality at a minimum, let us assume that

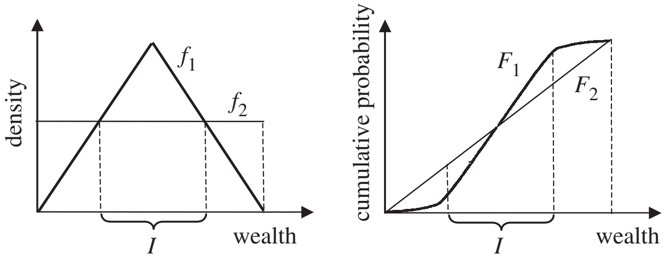

Figure 2.2. Example of a mean-preserving spread for a continuous distribution.

all possible final wealth levels are in the interval $(a, b]$. Suppose that $\tilde{w}_2$ is an MPS of $\tilde{w}_1$. Integrating by parts, preservation of the mean implies that

$$\int_a^b [F_2(s) - F_1(s)]\,\mathrm{d}s = -\int_a^b s[f_2(s) - f_1(s)]\,\mathrm{d}s = E\tilde{w}_2 - E\tilde{w}_1 = 0.$$

The fact that the expectation is preserved means that the area between F_1 and F_2 (counted as positive when F_2 is above F_1 and counted as negative otherwise) must sum up to zero. Also, by definition of an MPS, the derivative of F_2 is smaller (resp. larger) than the derivative of F_1 within the interval I (resp. outside I). Thus, F_2 must be larger than F_1 to the left of some threshold $\hat{w}$ and F_2 must be smaller than F_1 to its right. We illustrate this property in Figure 2.2 in the continuous case, and in Figure 2.3 in the discrete case considered in the previous paragraph.

This so-called "single-crossing" property of MPS implies in particular that

$$S(w) = \int_a^w [F_2(s) - F_1(s)]\,\mathrm{d}s \geqslant 0 \tag{2.1}$$

for all w, with an equality when w equals b. This integral condition is examined in more detail in the next section.

2.1.3 The Integral Condition and Risk-Averse Preferences

We now examine the problem of characterizing changes in risk that reduce the EU of all risk-averse agents. By integrating by parts, we obtain

$$Eu(\tilde{w}_i) = \int_a^b u(\omega) f_i(\omega)\,\mathrm{d}\omega = u(\omega)F_i(\omega)|_{\omega=a}^{\omega=b} - \int_a^b u'(\omega)F_i(\omega)\,\mathrm{d}\omega,$$

or, equivalently, that

$$Eu(\tilde{w}_i) = u(b) - \int_a^b u'(\omega)F_i(\omega)\,\mathrm{d}\omega.$$

It follows that

$$Eu(\tilde{w}_2) - Eu(\tilde{w}_1) = \int_a^b u'(\omega)[F_1(\omega) - F_2(\omega)]\,\mathrm{d}\omega. \tag{2.2}$$

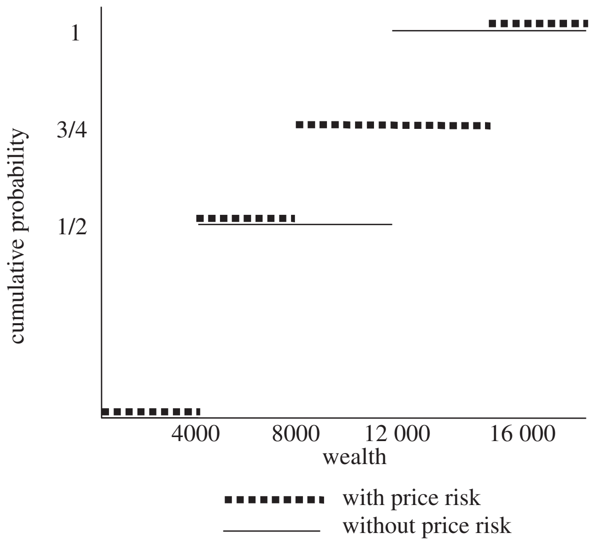

Figure 2.3. Cumulative distribution functions of Sempronius's final wealth with one ship, and with or without price uncertainty.

The difference in EU in transforming $\tilde{w}_1$ into $\tilde{w}_2$ is equal to the areas between F_1 and F_2 ("+" if F_1 is above F_2, and "−" otherwise) weighted by the marginal value of wealth. Integrating by parts once again yields

$$Eu(\tilde{w}_2) - Eu(\tilde{w}_1) = -u'(\omega)S(\omega)|_{\omega=a}^{\omega=b} + \int_a^b u''(\omega)S(\omega)\,\mathrm{d}\omega,$$

where function S is defined by equation (2.1) and is such that $S'(w) = F_2(w) - F_1(w)$. Because we focused the analysis on changes in risk that preserve the mean, we have that $S(a) = S(b) = 0$. The above equation thus simplifies to

$$Eu(\tilde{w}_2) - Eu(\tilde{w}_1) = \int_a^b u''(\omega)S(\omega)\,\mathrm{d}\omega. \tag{2.3}$$

Equation (2.3) implies that all risk-averse agents dislike mean-preserving increases in risk, that is changes in risk for which the condition $S(w) \geqslant 0$ is satisfied for all w. This condition would indeed imply that the integrand in (2.3) is uniformly negative. Its integral in $[a, b]$ should therefore be negative. The condition that $S(w) \geqslant 0$ for all w is also necessary to guarantee that every risk averter would unanimously prefer $\tilde{w}_1$ over $\tilde{w}_2$. Indeed, suppose by contradiction that S is positive in some interval $J \subseteq [a, b]$. Let us consider the concave utility function u that is linear outside J, and which is strictly concave in J. Then, from equation (2.3), agent u increases her EU by transforming $\tilde{w}_1$ into $\tilde{w}_2$. The integrand $u''S$ is zero for w outside J and is positive for w in J.

To sum up, the condition that $S(w) = \int^{w}(F_2(s) - F_1(s))\,\mathrm{d}s$ be nonnegative is both a necessary and a sufficient condition for mean-preserving changes in risk to reduce the EU of all risk-averse agents. It was examined by Rothschild and Stiglitz (1970). Notice from equation (2.2) that the condition $S(w) \geqslant 0$ implies that agents with the concave utility functions $u_w(x) = \min(x, w)\ \forall w \in [a, b]$ all prefer risk $\tilde{w}_1$ to risk $\tilde{w}_2$. We hope that this observation makes this integral condition less artificial.

There is a clear link between the integral condition $S \geqslant 0$ and the notion of a mean-preserving spread. It has been partly derived at the end of the previous section, where we have shown that a mean-preserving spread *implies* that S is nonnegative. Rothschild and Stiglitz (1970) have shown that the integral condition is *equivalent to* a sequence of mean-preserving spreads. In fact, they have proved the following proposition, showing how several interpretations of a mean-preserving increase in risk are all the same.

Proposition 2.2. *Consider two random variables $\tilde{w}_1$ and $\tilde{w}_2$ with the same mean. The following four conditions are equivalent.*

(a) *All risk-averse agents prefer $\tilde{w}_1$ to $\tilde{w}_2$: $Eu(\tilde{w}_2) \leqslant Eu(\tilde{w}_1)$ for all concave functions u.*

(b) *$\tilde{w}_2$ is obtained from $\tilde{w}_1$ by adding zero-mean noise terms to the possible outcomes of $\tilde{w}_1$:*

$$\tilde{w}_2 \overset{d}{=} \tilde{w}_1 + \tilde{\varepsilon},$$

with $E[\tilde{\varepsilon} \mid \tilde{w}_1 = \omega] = 0$ for all ω, where "$\overset{d}{=}$" means "equal in distribution."

(c) *$\tilde{w}_2$ is obtained from $\tilde{w}_1$ by a sequence of mean-preserving spreads.*

(d) *$S(w) \equiv \int^{w}(F_2(w) - F_1(w))\,\mathrm{d}w \geqslant 0$ for all w.*

Any one of these four equivalent conditions may define what we call an increase in risk from $\tilde{w}_1$ to $\tilde{w}_2$. Correspondingly, a change from $\tilde{w}_2$ to $\tilde{w}_1$ is labeled a reduction in risk.

2.1.4 Preference for Diversification

In Chapter 1, we have shown that Sempronius prefers to transfer the spices from the colonies by two ships rather than by only one. By doing so, he diversifies the risk. Let us now formalize this example by defining the random variable $\tilde{x}_i$ which takes value 0 if ship i is sunk (probability $\frac{1}{2}$), and which takes value 1 otherwise. In short, $\tilde{x}_i$ is distributed as $(0, \frac{1}{2}; 1, \frac{1}{2})$. By assumption, the risks faced by the two ships are independent. If Sempronius puts his 8000 pounds of spice in ship 1, his final wealth would equal

$$\tilde{w}_2 = w + 8000\tilde{x}_1.$$

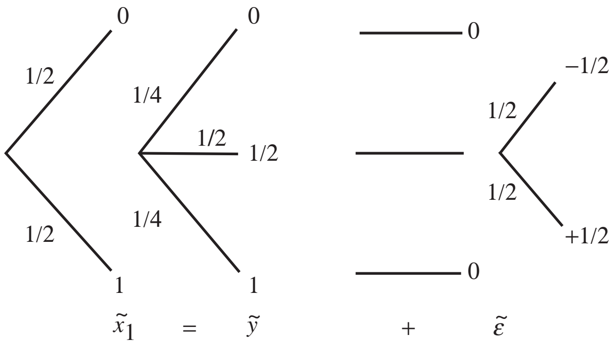

Figure 2.4. Diversification and reduction of risk.

We assume here that the price of spice is risk free and normalized to unity. If he splits the goods in two equal parts to be brought to London in ships 1 and 2, his final wealth equals

$$\tilde{w}_1 = w + 8000\left(\frac{\tilde{x}_1 + \tilde{x}_2}{2}\right) = w + 8000\tilde{y},$$

where $\tilde{y} = \frac{1}{2}(\tilde{x}_1 + \tilde{x}_2)$ can be interpreted as the rate of success. The rate of success is distributed as $(0, \frac{1}{4}; \frac{1}{2}, \frac{1}{2}; 1, \frac{1}{4})$. It is easy to check that $\tilde{x}_1$ can be obtained from $\tilde{y}$ by adding the noise $\tilde{\varepsilon} \sim (-\frac{1}{2}, \frac{1}{2}; +\frac{1}{2}, \frac{1}{2})$ conditional upon $\tilde{y} = \frac{1}{2}$, as seen in Figure 2.4. It then follows from Proposition 2.2 that, independent of the utility function of Sempronius, he must prefer two ships to one ship as soon as this function is concave. Diversifying the transfer of spice to two ships is a way to diversify the risk faced by Sempronius. Not "putting all the eggs in one basket" is a rational behavior for risk-averse agents.

More generally, one can verify that, if $\tilde{x}_1$ and $\tilde{x}_2$ are two independent and identically distributed (i.i.d.) random variables, then $\tilde{y} = \frac{1}{2}(\tilde{x}_1 + \tilde{x}_2)$ is a reduction of risk with respect to $\tilde{x}_1$. We have that

$$\tilde{x}_1 = \tilde{y} + \tilde{\varepsilon} \quad \text{with } \tilde{\varepsilon} = \frac{\tilde{x}_1 - \tilde{x}_2}{2}$$

and

$$\begin{aligned} E[\tilde{\varepsilon} \mid \tilde{y} = y] &= E\left[\frac{\tilde{x}_1 - \tilde{x}_2}{2} \,\middle|\, \frac{\tilde{x}_1 + \tilde{x}_2}{2} = y\right] \\ &= E\left[\tilde{x}_1 \,\middle|\, \frac{\tilde{x}_1 - \tilde{x}_2}{2}\right] - E\left[\tilde{x}_2 \,\middle|\, \frac{\tilde{x}_1 - \tilde{x}_2}{2}\right], \end{aligned}$$

which must be equal to zero by symmetry. Thus, $\tilde{x}_1$ is riskier than $\tilde{y}$. In other words, diversification is a risk-reduction device in the sense of Rothschild and Stiglitz. All risk-averse agents should diversify their risks when possible. This guideline does not rely on any preference restrictions other than risk aversion.

2.1.5 And the Variance?

The risk premium is the amount of money that the agent is ready to pay to eliminate the (zero-mean) risk. Facing the risk $\tilde{w}_i$ or receiving its certainty equivalent $E\tilde{w}_i - \Pi_i$ generates the same EU. Consider two risky wealth prospects $\tilde{w}_1$ and $\tilde{w}_2$ with equal means. It is clear that $\tilde{w}_1$ is preferred to $\tilde{w}_2$ if and only if Π_2 is larger than Π_1. Proposition 2.2 states the conditions on $\tilde{w}_1$ and $\tilde{w}_2$ that guarantee that Π_2 is larger than Π_1. Notice that, for small risks, we can use the Arrow–Pratt approximation

$$\Pi_i \simeq \tfrac{1}{2}\sigma_i^2 A$$

to claim that $\tilde{w}_1$ is preferred to $\tilde{w}_2$ if and only if the variance of $\tilde{w}_2$ is larger than the variance of $\tilde{w}_1 : \sigma_2^2 \geqslant \sigma_1^2$. Should not it also be the case that this holds for larger risks as well? That is to say, could we not add another equivalent statement (e) to Proposition 2.2 that would be written as follows:

(e) *the variance of* $\tilde{w}_2$ *is larger than the variance of* $\tilde{w}_1$?

The answer is definitely no! In general, statistical moments of orders higher than 2 will matter when comparing two random variables. By limiting the development of Taylor of the utility function to the second order, the Arrow–Pratt approximation is of no value for larger risks. The only exception is when the utility function is quadratic. A correct statement would then be that all risk-averse agents *with a quadratic utility function* prefer $\tilde{w}_1$ to $\tilde{w}_2$ if and only if the variance of the second is larger than the variance of the first.[1]

However, it is easy to check that an increase in variance is a necessary, but not sufficient, condition for an increase in risk. Indeed, it is a necessary condition for those agents with a quadratic concave utility function to dislike this change in risk. Thus, it is a necessary for *all* agents with a concave function to dislike it.

It is noteworthy that the necessary condition for an increase in the variance can be can be written as

$$\begin{aligned}\sigma_2^2 - \sigma_1^2 &= \int_a^b w^2(f_2(w) - f_1(w))\,\mathrm{d}w \\ &= 2\int_a^b S(w)\,\mathrm{d}w \geqslant 0.\end{aligned}$$

This equation is a direct consequence of equation (2.3) applied for $u(w) = w^2$. We see that the increase in variance just means that the integral of function S must be nonnegative. The Rothschild–Stiglitz increase in risk is a much stronger requirement that S be uniformly nonnegative.

[1] Another strategy is to limit the set of random variables to those that can be parametrized by their mean and variance only, as the set of normal distributions.

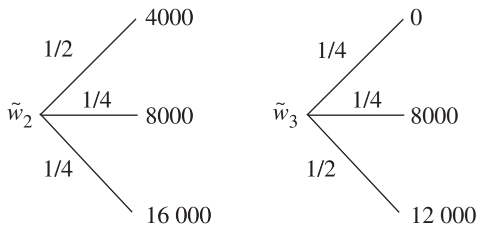

Figure 2.5. Increase in downside risk.

2.2 Aversion to Downside Risk

In this short section, we explore another set of changes in risk which are referred to as increases in downside risk. These changes have the property that they preserve both the mean and the variance of final wealth. To illustrate, consider again the distribution of final wealth $\tilde{w}_2 \sim (4000, \frac{1}{2}; 12\,000 + \tilde{\varepsilon}, \frac{1}{2})$, with $\tilde{\varepsilon} \sim (-4000, \frac{1}{2}, 4000, \frac{1}{2})$. Sempronius faces the price risk $\tilde{\varepsilon}$ in addition to the risk of losing his single ship. Observe that this is a situation where the additional zero-mean risk $\tilde{\varepsilon}$ is borne only in the good state, i.e. when the ship arrives safely at the harbor of London. Consider an alternative situation where this zero-mean risk would be borne in the bad state. It would thus yield a final wealth distributed as $\tilde{w}_3 \sim (4000 + \tilde{\varepsilon}, \frac{1}{2}; 12\,000, \frac{1}{2})$. Which of these two distributions of final wealth do you think Sempronius would prefer? To assist you in your choice, we represent these two distribution in Figure 2.5.

Observe that the means are the same: $E\tilde{w}_2 = E\tilde{w}_3 = 8000$. The variance is also unchanged by the change in distribution: $\sigma_2^2 = \sigma_3^2 = 24 \times 10^6$. In fact, the reader can check that function S alternates in sign in the interval of final wealth levels $[0, 16\,000]$. Thus, this cannot be an increase in risk. Hence, by Proposition 2.2, some risk-averse agents will like this change, whereas others will dislike it. However, experiments have shown that most people in the real world prefer $\tilde{w}_2$ to $\tilde{w}_3$. That is to say, they prefer to bear a zero-mean risk in the wealthier state. In other words, they dislike transferring a zero-mean risk from a richer to a poorer state. In this case, we say that they are averse to downside risk.

We are interested in determining a condition on the utility function that guarantees that an agent is averse to downside risk. Suppose that the agent is initially facing a risk that is characterized by $\tilde{w} \sim (z_1, 1/n; z_2, 1/n; \ldots; z_n, 1/n)$. We assume for simplicity that these n states have the same probability of occurrence. Consider an additional risk $\tilde{\varepsilon}$ with a zero mean. The EU of final wealth depends upon the state i to which this additional risk is imposed. We denote

$$V^{\mathrm{i}} = \frac{1}{n} E u(z_i + \tilde{\varepsilon}) + \sum_{j \neq i} \frac{1}{n} u(z_j)$$

for the EU when $\tilde{\varepsilon}$ is borne in state i. Observe that

$$n(V^{\mathrm{i}} - V_j) = [Eu(z_i + \tilde{\varepsilon}) - u(z_i)] - [Eu(z_j + \tilde{\varepsilon}) - u(z_j)]$$
$$= \int_{z_j}^{z_i} [Eu'(\omega + \tilde{\varepsilon}) - u'(\omega)]\, \mathrm{d}\omega.$$

Although intuition suggests the empirical observation that $V^{\mathrm{i}} > V_j$ when $z_i > z_j$, we see from the above equation that this is true if and only if

$$Eu'(\omega + \tilde{\varepsilon}) \geqslant u'(\omega)$$

for all ω. Because $\tilde{\varepsilon}$ is constrained only to have a zero mean, this condition is satisfied if and only if $u'(\cdot)$ is itself convex, i.e. if the agent is prudent. This is a direct application of Jensen's inequality. We thus obtain the following result.

Proposition 2.3. *An agent dislikes any increase in downside risk if and only if he is prudent.*

Prudence and aversion to downside risk are two equivalent concepts.

2.3 First-Degree Stochastic Dominance

Up to now, we have focused the analysis to changes in risk that preserve the mean. This is a strong requirement. For example, two portfolios with different proportions invested in stocks typically have different expected returns. Or, purchasing more insurance typically induces a reduction in expected wealth, since the insurance premium likely contains a loading. More generally, most decision making under uncertainty yields a trade-off between risk and (expected) return. In this section, we explore an important stochastic order named "First-degree Stochastic Dominance" (FSD), in which changes in mean are required.

There is often a discrepancy in the common use of the wording of "an increase in risk" between economists and the rest of the world. In common language, one often says that the risk is increased when the probability of an accident is increased. However, taken to the extreme this would imply that someone who always has an accident with the highest possible loss is the most "risky," whereas in a technical sense, there is no risk at all involved here: the lowest wealth value is realized with certainty! Of course, if the probability of an accident is increased, the expected final wealth of the risk bearer is reduced, which implies that this change in risk cannot be an increase in risk in the sense of Rothschild and Stiglitz. Economists say that the risk undergoes a dominated shift in the sense of FSD. More generally, any change in risk that is generated by a transfer of probability mass from high wealth states to low wealth states is said to be FSD-deteriorating. Such transfers of probability obviously raise $F(w)$, the probability that final wealth be no greater than than w, for all w.

Definition 2.4. $\tilde{w}_2$ is dominated by $\tilde{w}_1$ in the sense of the first-degree stochastic dominance order if $F_2(w) \geqslant F_1(w)$ for all w.

It is obvious that all consumers in the real world dislike FSD-dominated shifts in the distribution of final wealth. Rewriting condition (2.2) as

$$Eu(\tilde{w}_2) - Eu(\tilde{w}_1) = -\int_a^b u'(\omega)[F_2(\omega) - F_1(\omega)]\,\mathrm{d}\omega, \tag{2.4}$$

we see that $Eu(\tilde{w}_2)$ is smaller than $Eu(\tilde{w}_1)$ if $\tilde{w}_2$ is dominated by $\tilde{w}_1$ in the sense of FSD and if u' is positive. These two conditions indeed imply that the integrand of the above equation is always positive. Suppose that the only restriction that we impose on the utility function is that it be nondecreasing: more wealth is preferred to less. This means that we allow for both risk aversion and risk-loving behavior. Then, equation (2.4) tells us that $F_2 - F_1$ nonnegative is a necessary and sufficient condition for $Eu(\tilde{w}_2)$ to be smaller than $Eu(\tilde{w}_1)$. To prove this suppose by contradiction that $F_2 - F_1$ is negative in the neighborhood of some ω_0. Then, consider the nondecreasing utility function that is flat everywhere except in this neighborhood of ω_0. For this specific utility function, the integrand of (2.4) is zero everywhere except in the neighborhood of ω_0, where it is negative. Thus, the integral is negative, and this agent prefers $\tilde{w}_2$ to $\tilde{w}_1$, a contradiction. We have thus just proven the equivalence of (a) and (b) in the following proposition. The equivalence of (c) has been shown by several authors.

Proposition 2.5. *The following conditions are equivalent.*

(a) *All agents with a nondecreasing utility function prefer $\tilde{w}_1$ to $\tilde{w}_2$: $Eu(\tilde{w}_2) \leqslant Eu(\tilde{w}_1)$ for all nondecreasing functions u.*

(b) *$\tilde{w}_2$ is dominated by $\tilde{w}_1$ in the sense of FSD: $\tilde{w}_2$ is obtained from $\tilde{w}_1$ by a transfer of probability mass from the high wealth states to lower wealth states, or $F_2(\omega) \geqslant F_1(\omega)$ for all ω.*

(c) *$\tilde{w}_1$ is obtained from $\tilde{w}_2$ by adding nonnegative noise terms to the possible outcomes of $\tilde{w}_2$: $\tilde{w}_1 \overset{d}{=} \tilde{w}_2 + \tilde{\varepsilon}$, where $\tilde{\varepsilon} \geqslant 0$ with probability one.*

Without surprise, this type of change in risk, where the probability of lower wealth states is increased, is disliked by a very wide set of agents.

Of course, we can combine various changes in risk. For example, combining any FSD-dominated shift in distribution with any increase in risk yields what is called a (SSD) shift in distribution. Obviously, SSD shifts are disliked by the set of agents with a nondecreasing and concave utility function. Combining any SSD shift with any increase in downside risk yields a Third-degree Stochastically Dominated (TSD) shift in distribution. They are disliked by all prudent agents with a nondecreasing and concave utility function.

Of these three stochastic dominance orders, FSD is the most demanding one. For many applications however, it is considered to be too broad to yield unambiguous comparative static results. In the principal–agent literature, for example, one usually uses the more restrictive concept of the Monotone Likelihood Ratio (MLR) order. We say that $\tilde{w}_2$ is dominated by $\tilde{w}_1$ in the sense of MLR if $f_2(\omega)/f_1(\omega)$ is nonincreasing in ω. One can check that MLR is a special case of FSD.

2.4 Bibliographical References, Extensions and Exercises

The origin of the concepts developed in the literature on stochastic dominance can be found in an old book by famous mathematicians Hardy, Littlewood and Polya (1934). Its revival in the late 1960s is due to Hadar and Russell (1969) and Hanoch and Levy (1969) for the concepts of first-degree and second-degree stochastic dominance, while Rothschild and Stiglitz (1970) discussed mean-preserving increases in risk, a special case of SSD. Whitmore (1971) and Menezes, Geiss and Tressler (1980) were interested in third-degree stochastic dominance. The proof that diversification is liked by all risk-averse agents can be found in Samuelson (1967) and Rothschild and Stiglitz (1971).

Chapter Bibliography

Geiss, C., C. Menezes, and J. Tressler. 1980. Increasing downside risk. *American Economic Review* 70(5):921–931.

Hadar, J. and W. R. Russell. 1969. Rules for ordering uncertain prospects. *American Economic Review* 59:25–34.

Hanoch, G. and H. Levy. 1969. Efficiency analysis of choices involving risk. *Review of Economic Studies* 36:335–346.

Hardy, G. H., J. E. Littlewood and G. Polya. 1934. *Inequalities*. (Reprinted in 1997 by Cambridge University Press.)

Rothschild, M. and J. Stiglitz. 1970. Increasing risk. I. A definition. *Journal of Economic Theory* 2:225–243.

———. 1971. Increasing risk. II. Its economic consequences. *Journal of Economic Theory* 3:66–84.

Samuelson, P. A. 1967. General proof that diversification pays. *Journal of Financial and Quantitative Analysis*. 2(2):1–13.

Whitmore, G. A. 1970. Third-degree stochastic dominance. *American Economic Review* 60:457–459.

Exercises

(2.1) Consider the following two random variables: $\tilde{X}$ has a (continuous) uniform density on $[-1, +1]$, while $\tilde{Y}$ is a discrete random variable defined by $(-1, \frac{1}{2}; +1, \frac{1}{2})$.

(a) Do $\tilde{X}$ and $\tilde{Y}$ have the same mean?

(b) Compute their variances.

(c) Draw their cumulative distributions.

(d) Which random variable is riskier? Apply the 'integral condition' and also ask yourself which random variable has more weight in the center.

(e) Find the distributions of the 'white noise' that must be added to the less risky lottery to obtain the riskier one.

(2.2) Let $\tilde{X}$ be a binomial random variable with $n = 2$ and $p = \frac{1}{2}$, while $\tilde{Y}$ is also binomial but with $n = 3$ and $p = \frac{1}{3}$.

(a) Draw their cumulative distribution functions. Is $\tilde{Y}$ riskier than $\tilde{X}$ in the sense of Rothschild and Stiglitz?

(b) Compute σ^2 for $\tilde{X}$ and $\tilde{Y}$. (You should of course obtain that $\sigma^2_{\tilde{Y}}$ exceeds $\sigma^2_{\tilde{X}}$.)

(2.3) Besides a certain wealth of 100, Ms. A owns one house, the value of which is 80. The probability of full loss (due to fire) for this house is equal to 0.10 for a given time period and Ms. A has no access to an insurance market. In the absence of fire, the value of the house remains equal to its initial value. Mr. B has the same initial wealth but owns two houses valued at 40 for each. The probability of full loss for each house is 0.10 and the fires are assumed to be independent random variables (e.g. because one house is in Toulouse (France) and the other one in Mons (Belgium)).

(a) Draw the cumulative distribution functions of final wealth for Ms. A and Mr. B and compute the expected final wealth for each of them.

(b) Show that Ms. A has a riskier portfolio of houses.

(c) Select three (or more) concave utility functions and compute the expected utility for Ms. A and Mr. B. If you do not make mistakes, then the expected utility of Mr. B must be systematically higher than that of Ms. A for each utility curve you have selected.

(2.4) Consider lottery $\tilde{X}$ distributed as $(-10, \frac{1}{3}; 0, \frac{1}{3}; +10, \frac{1}{3})$ and a 'white noise' $\tilde{\varepsilon}$ distributed as $(-5, \frac{1}{2}; +5, \frac{1}{2})$. First generate lottery $\tilde{Y}$ by attaching the white noise to the worst outcome of $\tilde{X}$ (i.e. -10) and then generate lottery $\tilde{Z}$ by attaching the white noise instead to the best outcome of $\tilde{X}$ (i.e. $+10$).

(a) Compute $E(\tilde{Y})$, $E(\tilde{Z})$, $\mathrm{var}(\tilde{Y})$, $\mathrm{var}(\tilde{Z})$.

(b) Draw the cumulative distributions of $\tilde{Y}$ and $\tilde{Z}$. Can you say that one is riskier than the other? Why or why not?

(c) If a decision maker has a quadratic utility such as $u(w) = w - 0.01w^2$, compute $E[u(\tilde{Y})]$ and $E[u(\tilde{Z})]$. Are you surprised by the fact that $E[u(\tilde{Y})] = E[u(\tilde{Z})]$?

(d) Choose a utility function such that $u''' > 0$ and then show that

$$E[u(\tilde{Y})] < E[u(\tilde{Z})].$$

(2.5) A corporation must decide between two mutually exclusive projects. Both projects require an initial outlay of 100 million euro, and they generate cash flows that are independent of the growth of the economy. Project A has an equal probability of four gross payoffs: 80 million euro, 100 million euro, 120 million euro or 140 million euro. Project B has a 50:50 chance of paying either 90 million euro or 130 million euro. Assuming that shareholders are all risk averse, show that they unanimously prefer Project B to Project A.