1
Choice under Uncertainty

ASSET PRICING THEORY aims to describe the equilibrium in financial markets, where economic agents interact to trade claims to uncertain future payoffs. Both adjectives, "uncertain" and "future," are important—as suggested by the title of Christian Gollier's book *The Economics of Risk and Time* (2001)—but in this chapter we review the basic theory of choice under uncertainty, ignoring time by assuming that all uncertainty is resolved at a single future date. The chapter draws on both Gollier (2001) and Ingersoll (1987).

Section 1.1 begins by briefly reviewing the axiomatic foundations of expected utility theory. Section 1.2 applies expected utility theory to the measurement of risk aversion and the comparison of risk aversion across agents. Section 1.3 discusses the hyperbolic absolute risk averse (HARA) class of utility functions, which are widely used because they are so tractable in applications. Section 1.4 discusses critiques of expected utility theory, including the Allais (1953) paradox and the Rabin (2000) critique. Section 1.5 shows how to compare the riskiness of different distributions.

1.1 Expected Utility

Standard microeconomics represents preferences using ordinal utility functions. An ordinal utility function $\Upsilon(.)$ tells you that an agent is indifferent between x and y if $\Upsilon(x) = \Upsilon(y)$ and prefers x to y if $\Upsilon(x) > \Upsilon(y)$. Any strictly increasing function of $\Upsilon(.)$ will have the same properties, so the preferences expressed by $\Upsilon(.)$ are the same as those expressed by $\Theta(\Upsilon(.))$ for any strictly increasing Θ. In other words, ordinal utility is invariant to monotonically increasing transformations. It defines indifference curves, but there is no way to label the curves so that they have meaningful values.

A cardinal utility function $\Psi(.)$ is invariant to positive affine (increasing linear) transformations but not to nonlinear transformations. The preferences expressed by $\Psi(.)$ are the same as those expressed by $a + b\Psi(.)$ for any $b > 0$. In other words, cardinal utility has no natural units, but given a choice of units, the rate at which cardinal utility increases is meaningful.

Asset pricing theory relies heavily on von Neumann-Morgenstern utility theory, which says that choice over lotteries, satisfying certain axioms, implies maximization of the expectation of a cardinal utility function, defined over outcomes.

1.1.1 Sketch of von Neumann-Morgenstern Theory

The content of von Neumann-Morgenstern utility theory is easiest to understand in a discrete-state example. Define states $s = 1 \ldots S$, each of which is associated with an outcome x_s in a set X. Probabilities p_s of the different outcomes then define lotteries. When $S = 3$, we can draw probabilities in two dimensions (since $p_3 = 1 - p_1 - p_2$). We get the so-called Machina triangle (Machina 1982), illustrated in Figure 1.1.

We define a compound lottery as one that determines which primitive lottery we are given. For example, a compound lottery L might give us lottery L^a with probability α and lottery L^b with probability $(1 - \alpha)$. Then L has the same probabilities over the outcomes as $\alpha L^a + (1 - \alpha) L^b$.

We define a preference ordering $\succeq$ over lotteries. A person is indifferent between lotteries L^a and L^b, $L^a \sim L^b$, if and only if $L^a \succeq L^b$ and $L^b \succeq L^a$.

Next we apply two axioms of choice over lotteries.

Continuity axiom: For all L^a, L^b, L^c s.t. $L^a \succeq L^b \succeq L^c$, there exists a scalar $\alpha \in [0, 1]$ s.t.

$$L^b \sim \alpha L^a + (1 - \alpha) L^c. \tag{1.1}$$

This axiom says that if three lotteries are (weakly) ranked in order of preference, it is always possible to find a compound lottery that mixes the highest-ranked and lowest-ranked lotteries in such a way that the economic agent is indifferent between this compound lottery and the middle-ranked lottery. The axiom implies the existence of a preference functional defined over lotteries, that is, an ordinal utility function for lotteries that enables us to draw indifference curves on the Machina triangle.

Independence axiom:

$$L^a \succeq L^b \Rightarrow \alpha L^a + (1 - \alpha) L^c \succeq \alpha L^b + (1 - \alpha) L^c \tag{1.2}$$

for all possible lotteries L^c.

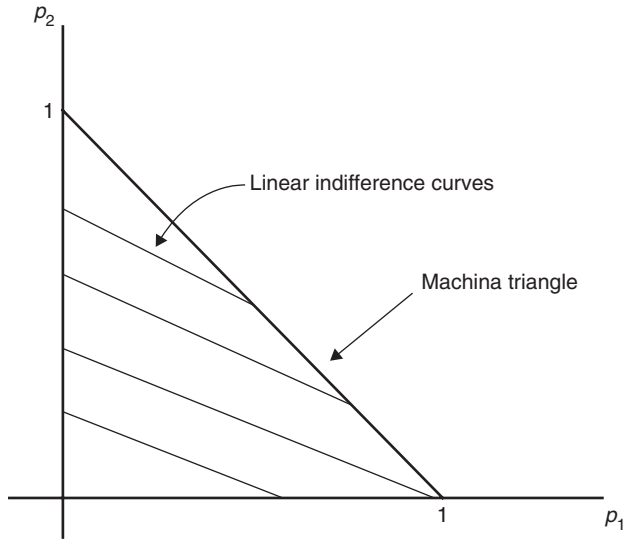

Figure 1.1. *Machina Triangle*

This axiom says that if two lotteries are ranked in order of preference, then the same rank order applies to two compound lotteries, each of which combines one of the original two lotteries with an arbitrary third lottery, using the same mixing weights in each case.

The independence axiom implies that the preference functional is linear in probabilities. In the Machina triangle, the indifference curves are straight lines, as illustrated in Figure 1.1. This means that a given increase in one probability, say p_1, requires the same change in another probability, say p_3, to leave the agent indifferent regardless of the initial levels of p_1 and p_3.

Then we can define a scalar u_s for each outcome x_s s.t.

$$L^a \succeq L^b \Rightarrow \sum_{s=1}^{S} p_s^a u_s \geq \sum_{s=1}^{S} p_s^b u_s. \tag{1.3}$$

The scalars u_s define the slopes of the linear indifference curves in the Machina triangle. Since probabilities sum to one and a constant can be added to all u_s without changing preferences, two scalars can be normalized (say the lowest to zero and the highest to one).

Equation (1.3) shows that a lottery is valued by the probability-weighted average of the scalars u_s associated with each outcome x_s. Call these scalars "utilities." A probability-weighted average of utilities u_s in each state s is the mathematical expectation of the random variable "utility" that takes the value u_s in state s. Hence, we have implicitly defined a cardinal utility function $u(x_s)$, defined over outcomes, such that the agent prefers the lottery that delivers a higher expectation of this function. The free normalization of lowest and highest utility corresponds to the two arbitrary parameters a and b that define the units in which cardinal utility is measured.

This construction can be generalized to handle continuous states. Strictly speaking, the resulting utility function must be bounded above and below, but this requirement is routinely ignored in modern applications of utility theory.

1.2 Risk Aversion

We now assume the existence of a cardinal utility function and ask what it means to say that the agent whose preferences are represented by that utility function is risk averse. We also discuss the quantitative measurement of risk aversion.

To bring out the main ideas as simply as possible, we assume that the argument of the utility function is wealth. This is equivalent to working with a single consumption good in a static two-period model where all wealth is liquidated and consumed in the second period, after uncertainty is resolved. Later in the book we discuss richer models in which consumption takes place in many periods, and also some models with multiple consumption goods.

For simplicity we also work with weak inequalities and weak preference orderings throughout. The extension to strict inequalities and strong preference orderings is straightforward.

1.2.1 Jensen's Inequality and Risk Aversion

An important mathematical result, Jensen's Inequality, can be used to link the concept of risk aversion to the concavity of the utility function. We start by defining concavity for a function f.

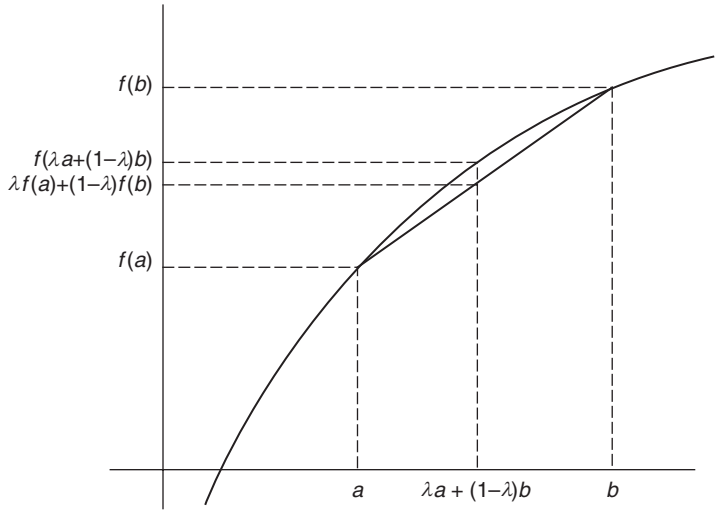

Figure 1.2. *Concave Function*

Definition. f is *concave* if and only if, for all $\lambda \in [0,1]$ and values a, b,

$$\lambda f(a) + (1-\lambda) f(b) \leq f(\lambda a + (1-\lambda) b). \tag{1.4}$$

If f is twice differentiable, then concavity implies that $f'' \leq 0$. Figure 1.2 illustrates a concave function.

Note that because the inequality is weak in the above definition, a linear function is concave. Strict concavity uses a strong inequality and excludes linear functions, but we proceed with the weak concept of concavity.

Now consider a random variable $\tilde{z}$. *Jensen's Inequality* states that

$$\mathrm{E} f(\tilde{z}) \leq f(\mathrm{E}\tilde{z}) \tag{1.5}$$

for all possible $\tilde{z}$ if and only if f is concave.

This result, due to the Danish mathematician and telephone engineer Johan Jensen, is so useful in finance that the field might almost be caricatured as "the economics of Jensen's Inequality." As a first application, we can use it to establish the equivalence of risk aversion and concavity of the utility function.

Definition. An agent is *risk averse* if she (weakly) dislikes all zero-mean risk at all levels of wealth. That is, for all initial wealth levels W_0 and risk $\tilde{x}$ with $\mathrm{E}\tilde{x} = 0$,

$$\mathrm{E} u(W_0 + \tilde{x}) \leq u(W_0). \tag{1.6}$$

To show that risk aversion is equivalent to concavity of the utility function, we simply rewrite the definition of risk aversion as

$$\mathrm{E} u(\tilde{z}) \leq u(\mathrm{E}\tilde{z}), \tag{1.7}$$

where $\tilde{z} = W_0 + \tilde{x}$, and apply Jensen's Inequality.

Since risk aversion is concavity, and concavity restricts the sign of the second derivative of the utility function (assuming that derivative exists), it is natural to construct a quantitative measure of risk aversion using the second derivative u'', scaled to avoid dependence on the units of measurement for utility. The *coefficient of absolute risk aversion* $A(W_0)$ is defined by

$$A(W_0) = \frac{-u''(W_0)}{u'(W_0)}. \tag{1.8}$$

As the notation makes clear, in general this is a function of the initial level of wealth.

1.2.2 Comparing Risk Aversion

Let two agents with utility functions u_1 and u_2 have the same initial wealth. An agent rejects a lottery if taking it lowers expected utility, that is, if the expected utility of initial wealth plus the lottery payout is lower than the utility of initial wealth. Continuing with our use of weak inequalities, we will also say that the agent rejects the lottery if it gives her the same expected utility as the utility of initial wealth.

Definition. u_1 is *more risk-averse* than u_2 if u_1 rejects all lotteries that u_2 rejects, regardless of the common initial wealth level.

Many utility functions cannot be ranked in this way. It is quite possible for agents to disagree about lotteries at a given initial wealth level (with the first agent accepting some that the second agent rejects and vice versa). It is also quite possible for the initial wealth level to matter, so that the first agent rejects all lotteries that the second agent rejects at a low level of initial wealth, but the second agent rejects all lotteries that the first agent rejects at a higher level of initial wealth.

What else is true if u_1 is more risk-averse than u_2? To answer this question, we first define a function

$$\phi(x) = u_1(u_2^{-1}(x)). \tag{1.9}$$

This function has three important properties:

(a) $u_1(z) = \phi(u_2(z))$, so $\phi(.)$ turns u_2 into u_1.
(b) $u_1'(z) = \phi'(u_2(z))\,u_2'(z)$, so $\phi' = u_1'/u_2' > 0$.
(c) $u_1''(z) = \phi'(u_2(z))\,u_2''(z) + \phi''(u_2(z))\,u_2'(z)^2$, so

$$\phi'' = \frac{u_1'' - \phi' u_2''}{{u_2'}^2} = \frac{u_1'}{{u_2'}^2}(A_2 - A_1). \tag{1.10}$$

The second of these properties is obtained by differentiating the first, and the third by differentiating the second. This trick (repeated differentiation to obtain restrictions on derivatives) often comes in handy in this field.

The third property is important because it shows that concavity of the function $\phi(x)$, $\phi'' \leq 0$, is equivalent to higher absolute risk aversion for agent 1, $A_1 \geq A_2$.

Now consider a risk $\tilde{x}$ that is rejected by u_2, that is, a risk s.t. $\mathrm{E}u_2(W_0 + \tilde{x}) \leq u_2(W_0)$. If u_1 is more risk-averse than u_2, we must also have $\mathrm{E}u_1(W_0 + \tilde{x}) \leq u_1(W_0)$. Using the function $\phi(.)$,

$$\mathrm{E}u_1(W_0 + \tilde{x}) = \mathrm{E}\phi(u_2(W_0 + \tilde{x})), \tag{1.11}$$

while

$$u_1(W_0) = \phi(u_2(W_0)) \geq \phi(\mathrm{E}u_2(W_0 + \tilde{x})) \tag{1.12}$$

since $\phi' > 0$. So for u_1 to be more risk-averse than u_2, we need

$$\mathrm{E}\phi(u_2(W_0 + \tilde{x})) \leq \phi(\mathrm{E}u_2(W_0 + \tilde{x})) \tag{1.13}$$

for all $\tilde{x}$. By Jensen's Inequality, this is equivalent to the concavity of the function $\phi(x)$, $\phi'' \leq 0$.

Putting these results together, we have shown that if one agent is more risk-averse than another, then the more risk-averse utility function is a concave transformation of the less risk-averse utility function and has a higher coefficient of absolute risk aversion at all levels of initial wealth. We have also shown the converse of these statements.

These concepts can be related to the amounts of wealth that agents are prepared to pay to avoid a zero-mean risk.

Definition. The *risk premium* $\pi(W_0, u, \tilde{x})$ is the greatest amount an agent with initial wealth W_0 and utility function u is willing to pay to avoid a risk $\tilde{x}$, assumed to have zero mean. Suppressing the arguments for notational simplicity, π is found by solving

$$\mathrm{E}u(W_0 + \tilde{x}) = u(W_0 - \pi). \tag{1.14}$$

Defining $z = W_0 - \pi$ and $\tilde{y} = \pi + \tilde{x}$, this can be rewritten as

$$\mathrm{E}u(z + \tilde{y}) = u(z). \tag{1.15}$$

Now define π_2 as the risk premium for agent 2, and define z_2 and $\tilde{y}_2$ accordingly. We have

$$\mathrm{E}u_2(z_2 + \tilde{y}_2) = u_2(z_2). \tag{1.16}$$

If u_1 is more risk-averse than u_2, then

$$\mathrm{E}u_1(z_2 + \tilde{y}_2) \leq u_1(z_2), \tag{1.17}$$

which implies $\pi_1 \geq \pi_2$. The same argument applies in reverse, so $\pi_1 \geq \pi_2$ implies that u_1 is more risk-averse than u_2.

We can extend the above analysis to consider a risk that may have a nonzero mean μ. It pays $\mu + \tilde{x}$ where $\tilde{x}$ has zero mean.

Definition. The *certainty equivalent* C^e satisfies

$$\mathrm{E}u(W_0 + \mu + \tilde{x}) = u(W_0 + C^e). \tag{1.18}$$

This implies that

$$C^e(W_0, u, \mu + \tilde{x}) = \mu - \pi(W_0 + \mu, u, \tilde{x}). \tag{1.19}$$

Thus if u_1 is more risk-averse than u_2, then $C_1^e \leq C_2^e$. Again, the reverse implication also holds.

In summary, the following statements are equivalent:

- u_1 is more risk-averse than u_2.
- u_1 is a concave transformation of u_2 at all initial wealth levels.
- $A_1 \geq A_2$ at all initial wealth levels.
- $\pi_1 \geq \pi_2$ at all initial wealth levels.
- $C_1^e \leq C_2^e$ at all initial wealth levels.

It is also possible to use the above ideas to ask how risk aversion for a single agent changes with the agent's level of wealth. It is natural to think that a richer person will care less about a given absolute risk than a poorer person, and will pay less to avoid it; in other words, that the risk premium for any risk should decline with initial wealth W_0. One can show that the following conditions are equivalent:

- π is decreasing in W_0.
- $A(W_0)$ is decreasing in W_0.
- $-u'$ is a concave transformation of u, so $-u'''/u'' \geq -u''/u'$ everywhere. The ratio $-u'''/u'' = P$ has been called *absolute prudence* by Kimball (1990), who relates it to the theory of precautionary saving.

Decreasing absolute risk aversion (DARA) is intuitively appealing. Certainly we should be uncomfortable with increasing absolute risk aversion.

1.2.3 The Arrow-Pratt Approximation

In the previous section, we defined the risk premium and certainty equivalent implicitly, as the solutions to equations (1.14) and (1.18). A famous analysis due to Arrow (1971) and Pratt (1964) shows that when risk is small, it is possible to derive approximate closed-form solutions to these equations.

Consider a zero-mean risk $\tilde{y} = k\tilde{x}$, where k is a scale factor. Write the risk premium as a function of k, $g(k) = \pi(W_0, u, k\tilde{x})$. From the definition of the risk premium, we have

$$\mathrm{E}u(W_0 + k\tilde{x}) = u(W_0 - g(k)). \tag{1.20}$$

Note that $g(0) = 0$, because you would pay nothing to avoid a risk with zero variability.

We now use the trick of repeated differentiation, in this case with respect to k, that was introduced in the previous subsection. Differentiating (1.20), we have

$$\mathrm{E}[\tilde{x}u'(W_0 + k\tilde{x})] = -g'(k)\,u'(W_0 - g(k)). \tag{1.21}$$

At $k = 0$, the left-hand side of (1.21) becomes $\mathrm{E}[\tilde{x}u'(W_0)] = \mathrm{E}[\tilde{x}]u'(W_0)$, where we can bring $u'(W_0)$ outside the expectations operator because it is deterministic. Since $\mathrm{E}[\tilde{x}] = 0$, the left-hand side of (1.21) is zero when $k = 0$, so the right-hand side must also be zero, which implies that $g'(0) = 0$.

We now differentiate with respect to k a second time to get

$$\mathrm{E}\tilde{x}^2 u''(w_o + k\tilde{x}) = g'(k)^2 u''(W_0 - g(k)) - g''(k)\,u'(W_0 - g(k)), \tag{1.22}$$

which implies that

$$g''(0) = \frac{-u''(W_0)}{u'(W_0)}\mathrm{E}\tilde{x}^2 = A(W_0)\mathrm{E}\tilde{x}^2. \tag{1.23}$$

Now take a Taylor approximation of $g(k)$ around the point of zero variability, $k = 0$:

$$g(k) \approx g(0) + kg'(0) + \frac{1}{2}k^2 g''(0). \tag{1.24}$$

Substituting in the previously obtained values for the derivatives, we get

$$\pi \approx \frac{1}{2}A(W_0)k^2\mathrm{E}\left[\tilde{x}^2\right] = \frac{1}{2}A(W_0)\mathrm{E}\left[\tilde{y}^2\right]. \tag{1.25}$$

The risk premium is proportional to the *square* of the risk. This property of differentiable utility is known as *second-order risk aversion*. It implies that people are approximately risk-neutral with respect to a single small risk (and more generally to small risks that are independent of other risks they face). The coefficient of proportionality is one-half the coefficient of absolute risk aversion, so we have a quantitative prediction linking the risk premium to the scale of risk and the level of risk aversion. This result is the basis for much modern quantitative research.

A similar analysis can be performed for the certainty equivalent. The result is that

$$C^e \approx k\mu - \frac{1}{2}A(W_0)k^2\mathrm{E}\left[\tilde{x}^2\right]. \tag{1.26}$$

This shows that the mean has a dominant effect on the certainty equivalent for small risks.

In finance, risks are often multiplicative rather than additive. That is, as the level of wealth invested increases, the absolute scale of the risk increases in proportion. The above theory can easily be modified to handle this case. Define a multiplicative risk by $\tilde{W} = W_0(1 + k\tilde{x}) = W_0(1 + \tilde{y})$. Define $\widehat{\pi}$ as the share of one's wealth one would pay to avoid this risk:

$$\widehat{\pi} = \frac{\pi\left(W_0, u, W_0 k\tilde{x}\right)}{W_0}. \tag{1.27}$$

Then

$$\widehat{\pi} \approx \frac{1}{2}W_0 A(W_0)k^2\mathrm{E}\tilde{x}^2 = \frac{1}{2}R(W_0)\mathrm{E}\tilde{y}^2, \tag{1.28}$$

where $R(W_0) = W_0 A(W_0)$ is the *coefficient of relative risk aversion*.

1.3 Tractable Utility Functions

Almost all applied theory and empirical work in finance uses some member of the class of utility functions known as linear risk tolerance (LRT) or hyperbolic absolute risk aversion (HARA). Continuing to use wealth as the argument of the utility function, the HARA class of utility functions can be written as

$$u(W) = a + b\left(\eta + \frac{W}{\gamma}\right)^{1-\gamma}, \tag{1.29}$$

defined for levels of wealth W such that $\eta + W/\gamma > 0$. The parameter a and the magnitude of the parameter b do not affect choices but can be set freely to deliver convenient representations of utility in special cases.

For these utility functions, risk tolerance—the reciprocal of absolute risk aversion—is given by

$$T(W) = \frac{1}{A(W)} = \eta + \frac{W}{\gamma}, \tag{1.30}$$

which is linear in W. Absolute risk aversion itself is then hyperbolic in W:

$$A(W) = \left(\eta + \frac{W}{\gamma}\right)^{-1}. \tag{1.31}$$

Relative risk aversion is, of course,

$$R(W) = W\left(\eta + \frac{W}{\gamma}\right)^{-1}. \tag{1.32}$$

There are several important special cases of HARA utility.

Quadratic utility has $\gamma = -1$. This implies that risk tolerance declines in wealth from (1.30), and absolute risk aversion increases in wealth from (1.31). In addition, the quadratic utility function has a "bliss point" at which $u' = 0$. These are important disadvantages, although quadratic utility is tractable in models with additive risk and has even been used in macroeconomic models with growth, where trending preference parameters are used to keep the bliss point well above levels of wealth or consumption observed in the data.

Exponential or constant absolute risk averse (CARA) utility is the limit as $\gamma \to -\infty$. To obtain constant absolute risk aversion A, we need

$$-u''(W) = Au'(W) \tag{1.33}$$

for all $W > 0$. Solving this differential equation, we get

$$u(W) = \frac{-\exp(-AW)}{A}, \tag{1.34}$$

where $A = 1/\eta$. This utility function does not have a bliss point, but it is bounded above; utility approaches zero as wealth increases. Exponential utility is tractable with normally distributed risks because then utility is lognormally distributed. In addition, as we will see in the next chapter, it implies that wealth has no effect on the demand for risky assets, which makes it relatively easy to calculate an equilibrium because one does not have to keep track of the wealth distribution.

Power or constant relative risk averse (CRRA) utility has $\eta = 0$ and $\gamma > 0$. Absolute risk aversion is declining in wealth — a desirable property — while relative risk aversion $R(W) = \gamma$, a constant. For $\gamma \neq 1$, a and b in equation (1.29) can be chosen to write utility as

$$u(W) = \frac{W^{1-\gamma} - 1}{1 - \gamma}. \tag{1.35}$$

For $\gamma = 1$, we use L'Hôpital's rule to take the limit of equation (1.35) as γ approaches one. The result is

$$u(W) = \log(W). \tag{1.36}$$

Power utility is appealing because it implies stationary risk premia and interest rates even in the presence of long-run economic growth. Also it is tractable in the presence of multiplicative lognormally distributed risks. For these reasons it is a workhorse model in the asset pricing and macroeconomics literatures and will be used intensively in this book. The special case of log utility has even more convenient properties, but relative risk aversion as low as one is hard to reconcile with the substantial risk premia observed in financial markets as we discuss in Chapter 6.

Subsistence level. A negative η represents a subsistence level, a minimum level of consumption that is required for utility to be defined. Litzenberger and Rubinstein (1976) argued for a model with log utility of wealth above the subsistence level, which they called the Generalized Log Utility Model. The proposal did not gain traction, perhaps in part because economic growth renders any fixed subsistence level irrelevant in the long run.[1] Models of habit formation, discussed in Chapter 6, have time-varying subsistence levels that can grow with the economy.

1.4 Critiques of Expected Utility Theory

1.4.1 Allais Paradox

This famous paradox, due to Allais (1953), challenges the von Neumann-Morgenstern framework. Consider a set of lotteries, each of which involves drawing one ball from an urn containing 100 balls, labeled 0–99. Table 1.1 shows the monetary prizes that will be awarded for drawing each ball, in four different lotteries L^a, L^b, M^a, and M^b.

Lottery L^a offers \$50 with certainty, while lottery L^b offers an 89% chance of \$50, a 10% chance of \$250, and a 1% chance of receiving nothing. Many people, confronted with this choice, prefer L^a to L^b even though the expected winnings are higher for lottery L^b. Lottery M^a offers an 11% chance of winning \$50 and an 89% chance of receiving nothing, while lottery M^b offers a 10% chance of winning \$250 and a 90% chance of receiving nothing. Many people, confronted with this choice, prefer M^b to M^a.

The challenge to utility theory is that choosing L^a over L^b, while also choosing M^b over M^a, violates the independence axiom. As the structure of the table makes clear, the only difference between L^a and L^b is in the balls labeled 0–10; the balls labeled 11–99 are identical in these two lotteries. This is also true for the pair M^a and M^b. According to the independence axiom, the rewards for drawing balls 11–99 should then be irrelevant

Table 1.1. *Allais Paradox*

	0	1–10	11–99
L^a	50	50	50
L^b	0	250	50
M^a	50	50	0
M^b	0	250	0

[1]The model's gloomy acronym may also have hurt its prospects. Possibly only Deaton and Muellbauer (1980) were less fortunate in this respect.

to the choices between L^a and L^b, and M^b and M^a. But if this is the case, then the two choices are the same because if one considers only balls 0–10, L^a has the same rewards as M^a, and L^b has the same rewards as M^b.

There is a longstanding debate over the significance of this paradox. Either people are easily misled (but can be educated) or the independence axiom needs to be abandoned. Relaxing this axiom must be done carefully to avoid creating further paradoxes (Chew 1983, Dekel 1986, Gul 1991).[2] Recent models of dynamic decision making, notably the Epstein and Zin (1989, 1991) preferences discussed in section 6.4, also relax the independence axiom in an intertemporal context, taking care to do so in a way that preserves time consistent decision making.

1.4.2 Rabin Critique

Matthew Rabin (2000) has criticized utility theory on the ground that it cannot explain observed aversion to small gambles without implying ridiculous aversion to large gambles. This follows from the fact that differentiable utility has second-order risk aversion.

To understand Rabin's critique, consider a gamble that wins \$11 with probability 1/2 and loses \$10 with probability 1/2. With diminishing marginal utility, the utility of the win is at least $11u'(W_0 + 11)$. The utility cost of the loss is at most $10u'(W_0 - 10)$. Thus if a person turns down this gamble, we must have $10u'(W_0 - 10) > 11u'(W_0 + 11)$, which implies

$$\frac{u'(W_0 + 11)}{u'(W_0 - 10)} < \frac{10}{11}.$$

Now suppose the person turns down the same gamble at an initial wealth level of $W_0 + 21$. Then

$$\frac{u'(W_0 + 21 + 11)}{u'(W_0 + 21 - 10)} = \frac{u'(W_0 + 32)}{u'(W_0 + 11)} < \frac{10}{11}.$$

Combining these two inequalities,

$$\frac{u'(W_0 + 32)}{u'(W_0 - 10)} < \left(\frac{10}{11}\right)^2 = \frac{100}{121}.$$

If this iteration can be repeated, it implies extremely small marginal utility at high wealth levels, which would induce people to turn down apparently extremely attractive gambles.

Table 1.2 is an extract from Rabin (2000), Table I. The original caption reads "If averse to 50-50 lose \$100/gain g bets for all wealth levels, will turn down 50-50 lose L/gain G bets; G's entered in table." Values g are entered in the column headings, and values L are entered in the row labels, while the cells of the table report G. In other words, as one moves to the right, each column corresponds to an agent who is turning down more and

[2]For example, suppose that $L^a \succ L^b$ and $L^a \succ L^c$ but contrary to the independence axiom $L^d = 0.5L^b + 0.5L^c \succ L^a$. Then you would pay to switch from L^a to L^d, but once the uncertainty in the compound lottery L^d is resolved, you would pay again to switch back to L^a. This is sometimes called the "Dutch book" problem. It can be avoided by imposing Chew's (1983) property of "betweenness," that a convex combination of two lotteries (L^b and L^c in the example above) cannot be preferred to the more preferred of the two, and the less preferred of the two cannot be preferred to the convex combination.

Table 1.2. *Extract from Rabin (2000), Table I*

L/g	\$101	\$105	\$110	\$125
\$400	400	420	550	1,250
\$1,000	1,010	1,570	∞	∞
\$4,000	5,750	∞	∞	∞
\$10,000	∞	∞	∞	∞

Table 1.3. *Extract from Rabin (2000), Table II*

L/g	\$101	\$105	\$110	\$125
\$400	400	420	550	1,250
\$1,000	1,010	1,570	718,190	160 billion
\$4,000	5,750	635,670	60.5 million	9.4 trillion
\$10,000	27,780	5.5 million	160 billion	5.4 sextillion

more favorable small gambles. As one moves down the table, each row corresponds to a larger possible loss, and the table entries show the winnings that are required to induce the agent to take the bet. An entry of ∞ implies that the agent will turn down the bet for any finite upside, no matter how large.

A first obvious question is how is it possible for an agent to be unresponsive to arbitrarily large winnings, refusing to risk a finite loss. To promote careful thought, this question is posed as an informal problem and is answered at the end of the chapter. As a clue, Table 1.3 is an extract from Rabin (2000), Table II. The only difference between this and the previous table is that the numbers here are conditional on a specific initial wealth level (\$290,000), and the aversion to 50-50 lose \$100/gain g bets is known to hold only for wealth levels up to \$300,000.

1.4.3 First-Order Risk Aversion and Prospect Theory

Rabin's critique shows that the standard theory of expected utility cannot explain risk aversion with respect to small gambles over a significant range of wealth levels. At any one level of wealth, one can increase aversion to small gambles within the standard theory by relaxing the assumption that utility is twice differentiable, allowing a kink in the utility function that invalidates the standard formula for the risk premium given in (1.25). Such a kink makes risk aversion locally infinite and implies that the risk premium for a small gamble is proportional to its standard deviation rather than its variance; this is called "first-order" risk aversion by contrast with the "second-order" risk aversion implied by twice differentiable utility (Segal and Spivak 1990). However, this approach only increases aversion to small gambles at a single point, and Rabin's argument (which does not assume twice differentiability of the utility function) still applies if an agent is averse to small gambles over a range of wealth levels.

In response to this, economists and psychologists have explored models with reference points, in which utility results from gains or losses relative to a reference point that

is often set equal to current wealth. This has the effect of moving the kink in the utility function so that it is always relevant and induces first-order risk aversion at arbitrary levels of initial wealth.

The most famous example is Kahneman and Tversky's (1979) prospect theory, which has not only a kink at the reference point but also two other features designed to fit experimental evidence on risk attitudes: a preference function that is concave in the domain of gains and convex (risk-seeking) in the domain of losses, and subjective probabilities that are larger than objective probabilities when those probabilities are small. A standard parameterization of the prospect-theory preference function is

$$u(x) = x^{\beta} \text{ for } x \geq 0,$$
$$u(x) = -\lambda |x|^{\beta} \text{ for } x \leq 0, \tag{1.37}$$

where $x = W - W_{REF}$, the difference between wealth and the reference level of wealth. We assume $0 < \beta < 1$ to get concavity for gains and convexity for losses, and $\lambda > 1$ to deliver a kink at the reference point. Gul's (1991) disappointment averse preferences also have a kink at a reference point set equal to the endogenous certainty equivalent of a gamble (Backus, Routledge, and Zin 2004).

Barberis, Huang, and Thaler (2006) point out that even these preferences cannot generate substantial aversion to small delayed gambles. During the time between the decision to take a gamble and the resolution of uncertainty, the agent will be exposed to other risks and will merge these with the gamble under consideration. If the gamble is uncorrelated with the other risks, it is diversifying. In effect the agent will have second-order risk aversion with respect to delayed gambles. To deal with this problem, Barberis et al. argue that people treat gambles in isolation, that is, they use "narrow framing."

In this book, we will continue to work primarily with standard utility functions despite their inability to explain aversion to small risks. This reflects my belief that the theory is useful for asset pricing problems, consistent with Rabin's acknowledgement that it "may well be a useful model of the taste for very-large-scale insurance" (Rabin 2000). One might make an analogy with physics, where the force of gravity is dominant at cosmological scales even though it becomes negligible at subatomic scales where other forces are far more important.

Finally, it is worth noting that expected utility theory can be enriched to generate differences in aversion to medium-scale and large-scale risks. Notably, Chetty and Szeidl (2007) show that "consumption commitments" (fixed costs to adjust a portion of consumption) raise risk aversion over medium-sized gambles, relative to risk aversion over large gambles where extreme outcomes would justify paying the cost to adjust all consumption.

1.5 Comparing Risks

Earlier in this chapter we discussed the comparison of utility functions, concentrating on cases where two utility functions can be ranked in their risk aversion, with one turning down all lotteries that the other one turns down, regardless of the distribution of the risks. Now we perform a symmetric analysis, comparing the riskiness of two different distributions without making any assumptions on utility functions other than concavity.

1.5.1 Comparing Risks with the Same Mean

In this subsection we consider two distributions that have the same mean. Informally, there are three natural ways to define the notion that one of these distributions is riskier than the other:

(1) All increasing and concave utility functions dislike the riskier distribution relative to the safer distribution.
(2) The riskier distribution has more weight in the tails than the safer distribution.
(3) The riskier distribution can be obtained from the safer distribution by adding noise to it.

The classic analysis of Rothschild and Stiglitz (1970) shows that these are all equivalent. Consider random variables $\widetilde{X}$ and $\widetilde{Y}$, which have the same expectation.

(1) $\widetilde{X}$ is weakly less risky than $\widetilde{Y}$ if no individual with an increasing concave utility function prefers $\widetilde{Y}$ to $\widetilde{X}$:

$$E[u(\widetilde{X})] \geq E[u(\widetilde{Y})] \tag{1.38}$$

for all increasing concave u (.). $\widetilde{X}$ is less risky than $\widetilde{Y}$ (without qualification) if it is weakly less risky than $\widetilde{Y}$ and there is some increasing concave u(.) which strictly prefers $\widetilde{X}$ to $\widetilde{Y}$.

Note that this is a partial ordering. It is not the case that for any $\widetilde{X}$ and $\widetilde{Y}$, either $\widetilde{X}$ is weakly less risky than $\widetilde{Y}$ or $\widetilde{Y}$ is weakly less risky than $\widetilde{X}$. We can get a complete ordering if we restrict attention to a smaller class of utility functions than the concave, such as the quadratic.

(2) $\widetilde{X}$ is less risky than $\widetilde{Y}$ if the density function of $\widetilde{Y}$ can be obtained from that of $\widetilde{X}$ by applying a *mean-preserving spread* (MPS). An MPS $s(x)$ is defined by

$$s(x) = \begin{pmatrix} \alpha \text{ for } c < x < c+t \\ -\alpha \text{ for } c' < x < c'+t \\ -\beta \text{ for } d < x < d+t \\ \beta \text{ for } d' < x < d'+t \\ 0 \text{ elsewhere} \end{pmatrix}, \tag{1.39}$$

where $\alpha, \beta, t > 0$; $c + t < c' < c' + t < d < d + t < d'$; and $\alpha(c' - c) = \beta(d' - d)$; that is, "the more mass you move, the less far you can move it." This is illustrated in Figure 1.3.

An MPS is something you add to a density function $f(x)$. If $g(x) = f(x) + s(x)$, then (i) $g(x)$ is also a density function, and (ii) it has the same mean as $f(x)$.

(i) is obvious because $\int s(x)\,dx$ = area under $s(x) = 0$.
(ii) follows from the fact that the "mean" of $s(x)$, $\int xs(x)\,dx = 0$, which follows from $\alpha(c' - c) = \beta(d' - d)$. The algebra is

$$\begin{aligned} \int xs(x)\,dx &= \int_c^{c+t} x\alpha\,dx + \int_{c'}^{c'+t} x(-\alpha)\,dx + \int_d^{d+t} x(-\beta)\,dx + \int_{d'}^{d'+t} x\beta\,dx \\ &= \alpha\left[\frac{x^2}{2}\right]_c^{c+t} - \alpha\left[\frac{x^2}{2}\right]_{c'}^{c'+t} - \beta\left[\frac{x^2}{2}\right]_d^{d+t} + \beta\left[\frac{x^2}{2}\right]_{d'}^{d'+t} \\ &= t\left[\beta(d' - d) - \alpha(c' - c)\right] = 0. \end{aligned} \tag{1.40}$$

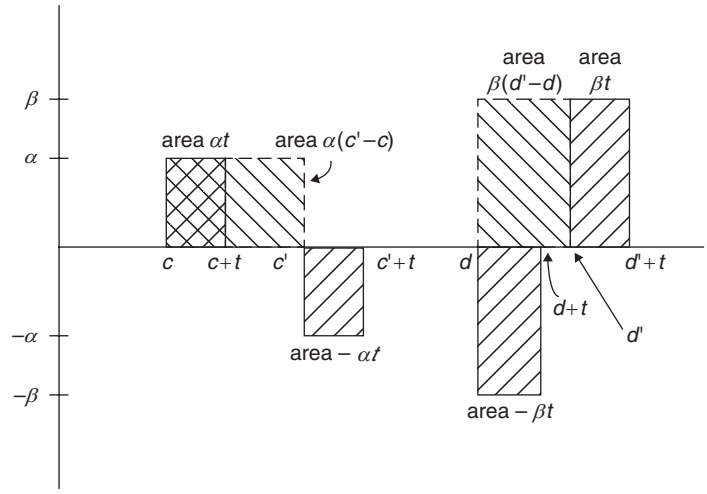

Figure 1.3. *Mean-Preserving Spread*

In what sense is an MPS a spread? It is obvious that if the mean of $f(x)$ is between $c' + t$ and d, then $g(x)$ has more weight in the tails. This is not so obvious when the mean of $f(x)$ is far to the left or the right in Figure 1.3. Nevertheless, we can show that $\widetilde{Y}$ with density g is riskier than $\widetilde{X}$ with density f in the sense of (1) above. In this sense the term "spread" is appropriate.

We calculate the expected utility difference between $\widetilde{X}$ and $\widetilde{Y}$ as

$$
\begin{aligned}
&E[u(\widetilde{X})] - E[u(\widetilde{Y})] \\
&\quad = \int u(z)\,[f(z) - g(z)]\,dz = -\int u(z)\,s(z)\,dz \qquad (1.41)\\
&\quad = -\alpha \int_c^{c+t} u(z)\,dz + \alpha \int_{c'}^{c'+t} u(z)\,dz + \beta \int_d^{d+t} u(z)\,dz - \beta \int_{d'}^{d'+t} u(z)\,dz \\
&\quad = -\alpha \int_c^{c+t} \left[u(z) - u(z + c' - c) - \frac{\beta}{\alpha}\{u(z + d - c) - u(z + d' - c)\} \right] dz.
\end{aligned}
$$

The definition of an MPS implies that $\beta/\alpha = (c' - c)/(d' - d)$. In addition, $u(z + h) - u(z) = u'(z^*)\,h$ for some z^* between z and $z + h$.

Thus

$$u(z) - u(z + c' - c) = -(c' - c)\,u'(z_1^*) \qquad (1.42)$$

for some z_1^* between z and $z + c' - c$, and

$$u(z + d - c) - u(z + d' - c) = -(d' - d)\,u'(z_2^*) \qquad (1.43)$$

for some z_2^* between $z + d - c$ and $z + d' - c$. Substituting into (1.41), we get

$$E[u(\widetilde{X})] - E[u(\widetilde{Y})] = \alpha(c' - c) \int_c^{c+t} \left[u'(z_1^*) - u'(z_2^*) \right] dz > 0, \qquad (1.44)$$

where the inequality follows because $z_1^* < z_2^*$ so $u'(z_1^*) > u'(z_2^*)$.

(3) A formal definition of "added noise" is that $\widetilde{X}$ is less risky than $\widetilde{Y}$ if $\widetilde{Y}$ has the same distribution as $\widetilde{X} + \widetilde{\varepsilon}$, where $E[\widetilde{\varepsilon}|X] = 0$ for all values of X. We say that $\widetilde{\varepsilon}$ is a "fair game" with respect to X.

The fair game condition is stronger than zero covariance, $\text{Cov}(\widetilde{\varepsilon}, \widetilde{X}) = 0$. It is weaker than independence, $\text{Cov}(f(\widetilde{\varepsilon}), g(\widetilde{X})) = 0$ for all functions f and g. It is equivalent to $\text{Cov}(\widetilde{\varepsilon}, g(\widetilde{X})) = 0$ for all functions g. To develop your understanding of this point, Problem 1.1 at the end of this chapter asks you to construct examples of random variables $\widetilde{X}$ and $\widetilde{\varepsilon}$ that have zero covariance but do not satisfy the fair game condition, or that satisfy the fair game condition but are not independent.

It is straightforward to show that added noise is sufficient for a concave utility function to dislike the resulting distribution, that is, (3) implies (1):

$$\begin{aligned} E[U(\widetilde{X} + \widetilde{\varepsilon})|X] &\leq U(E[\widetilde{X} + \widetilde{\varepsilon}|X]) = U(X) \\ &\Rightarrow E[U(\widetilde{X} + \widetilde{\varepsilon})] \leq E[U(\widetilde{X})] \\ &\Rightarrow E[U(\widetilde{Y})] \leq E[U(\widetilde{X})], \end{aligned} \tag{1.45}$$

because $\widetilde{Y}$ and $\widetilde{X} + \widetilde{\varepsilon}$ have the same distribution.

More generally, Rothschild and Stiglitz show that conditions (1), (2), and (3) are all equivalent. This is a powerful result because one or the other condition may be most useful in a particular application.

None of these conditions are equivalent to $\widetilde{Y}$ having greater variance than $\widetilde{X}$. It is obvious from (3) that if $\widetilde{Y}$ is riskier than $\widetilde{X}$ then $\widetilde{Y}$ has greater variance than $\widetilde{X}$. The problem is that the reverse is not true in general. Greater variance is necessary but not sufficient for increased risk. $\widetilde{Y}$ could have greater variance than $\widetilde{X}$ but still be preferred by some concave utility functions if it has more desirable higher-moment properties. This possibility can only be eliminated if we confine attention to a limited class of distributions such as the normal distribution.

1.5.2 Comparing Risks with Different Means

The Rothschild-Stiglitz conditions apply only to distributions that have the same mean. However, they extend straightforwardly to the case where a riskier distribution, in the Rothschild-Stiglitz sense, is shifted downward and therefore has a lower mean. Some brief definitions illustrate this point.

Definition. $\widetilde{X}$ *(first-order) dominates* $\widetilde{Y}$ if $\widetilde{Y} = \widetilde{X} + \widetilde{\xi}$, where $\widetilde{\xi} \leq 0$. In this case every outcome for $\widetilde{X}$ is at least as great as the corresponding outcome for $\widetilde{Y}$.

Definition. $\widetilde{X}$ *first-order stochastically dominates* $\widetilde{Y}$ if $\widetilde{Y}$ has the distribution of $\widetilde{X} + \widetilde{\xi}$, where $\widetilde{\xi} \leq 0$. Equivalently, if $F(.)$ is the cdf of $\widetilde{X}$ and $G(.)$ is the cdf of $\widetilde{Y}$, then $\widetilde{X}$ first-order stochastically dominates $\widetilde{Y}$ if $F(z) \leq G(z)$ for every z. In this case every quantile of the $\widetilde{X}$ distribution is at least as great as the corresponding quantile of the $\widetilde{Y}$ distribution, but a particular outcome for $\widetilde{Y}$ may exceed the corresponding outcome for $\widetilde{X}$. First-order stochastic dominance implies that every increasing utility function will prefer the distribution $\widetilde{X}$.

Definition. $\widetilde{X}$ *second-order stochastically dominates* $\widetilde{Y}$ if $\widetilde{Y}$ has the distribution of $\widetilde{X} + \widetilde{\xi} + \widetilde{\varepsilon}$, where $\widetilde{\xi} \leq 0$ and $\text{E}[\widetilde{\varepsilon}|X + \xi] = 0$. Second-order stochastic dominance (SOSD) implies that

every increasing, concave utility function will prefer the distribution $\widetilde{X}$. Increased risk is the special case of SOSD where $\widetilde{\xi} = 0$.

SOSD, based on the consistent preference of all risk-averse decision makers for one gamble over another, offers an uncontroversial comparison of risks. Unfortunately this also limits its applicability: SOSD is only a partial order of gambles; that is, many pairs of gambles cannot be ranked using SOSD. Specifically, when a riskier distribution, in the Rothschild-Stiglitz sense, is shifted upward—implying that it has a higher mean—then one cannot assert that any concave utility function will prefer the safer alternative. The choice will depend on the scale of the risk and the form of the utility function. This tradeoff is the subject of portfolio choice theory, which we explore in the next chapter.

It is possible to create a complete order, delivering a ranking of any two gambles, if one confines attention to a more specific set of decision makers (defined by their utility functions and wealth levels). A complete order can be used to create a riskiness index, that is, a summary statistic mapping a gamble to a real number that depends only on the attributes of the gamble itself. For example, Aumann and Serrano (2008) propose a riskiness index based on the preferences of agents with CARA utility, for whom wealth does not affect their attitudes toward gambles. The Aumann-Serrano index is the risk tolerance (the reciprocal of risk aversion) that makes a CARA agent indifferent to a gamble. Problem 1.2 invites you to explore this and another riskiness index proposed by Foster and Hart (2009). While riskiness indices lack the generality of SOSD and depend on the preferences considered, they can nonetheless be useful for descriptive and regulatory purposes.

1.5.3 The Principle of Diversification

We conclude this chapter by showing how the Rothschild-Stiglitz analysis can be used to prove the optimality of perfect diversification in a simple portfolio choice problem with identical risky assets.

Consider n lotteries with payoffs $\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n$ that are independent and identically distributed (iid). You are asked to choose weights $\alpha_1, \alpha_2, \ldots, \alpha_n$ subject to the constraint that $\sum_i \alpha_i = 1$. It seems obvious that the best choice is a fully diversified, equally weighted portfolio with weights $\alpha_i = 1/n$ for all i. The payoff is then

$$\widetilde{z} = \frac{1}{n} \sum_{i=1}^{n} \widetilde{x}_i. \tag{1.46}$$

The Rothschild-Stiglitz analysis makes it very easy to prove that this is optimal. Just note that the payoff on any other strategy is

$$\sum_i \alpha_i \widetilde{x}_i = \widetilde{z} + \sum_i \left(\alpha_i - \frac{1}{n} \right) \widetilde{x}_i = \widetilde{z} + \widetilde{\varepsilon}, \tag{1.47}$$

and

$$\mathrm{E}[\widetilde{\varepsilon}|z] = \sum_i \left(\alpha_i - \frac{1}{n} \right) \mathrm{E}[\widetilde{x}_i|z] = k \sum_i \left(\alpha_i - \frac{1}{n} \right) = 0. \tag{1.48}$$

Thus, any other strategy has the payoff of the equally weighted portfolio, plus added noise (Rothschild-Stiglitz condition (3)). It follows that any concave utility function will prefer the equally weighted portfolio (Rothschild-Stiglitz condition (1)).

1.6 Solution and Further Problems

An informal problem posed in this chapter was how it is possible for an agent to turn down a 50-50 gamble with a fixed loss, regardless of the size of the potential winnings, as claimed in Rabin (2000), Table I. The answer is that if utility is bounded above, then the utility gain from a win converges to a finite limit even as the size of the win becomes arbitrarily large. Rabin's assumption in Table I—that an agent with expected utility turns down a given small gamble at all initial wealth levels—requires that absolute risk aversion is nondecreasing (because with decreasing absolute risk aversion, at some high enough level of wealth the agent will accept the small gamble). But the utility function with constant absolute risk aversion, the exponential utility function, is bounded above, and the same is true of all utility functions with increasing absolute risk aversion such as the quadratic utility function. This discussion suggests that Table II may be a more relevant critique of expected utility than Table I. Table II makes a weaker assumption about the range of wealth over which an agent turns down a given small gamble and is thus consistent with decreasing absolute risk aversion.

Problem 1.1 Fair Games

State whether each of the following statements is true or false. Provide a proof if the statement is true or a counterexample if the statement is false.

(a) If $\widetilde{X}$ is a fair game with respect to $\widetilde{Y}$, and $\widetilde{Y}$ is a fair game with respect to $\widetilde{X}$, then $\widetilde{X}$ and $\widetilde{Y}$ are independent.

(b) If $\widetilde{X}$ and $\widetilde{Y}$ have zero means and zero covariance, then $\widetilde{X}$ is a fair game with respect to $\widetilde{Y}$ and $\widetilde{Y}$ is a fair game with respect to $\widetilde{X}$.

(c) For jointly normally distributed random variables, zero covariance implies independence.

Problem 1.2 Riskiness Indices

This exercise explores the properties of two recently proposed riskiness indices: the Aumann and Serrano (AS 2008) index and the Foster and Hart (FH 2009) index.

A decision maker is characterized by an initial wealth level W_0 and von Neumann-Morgenstern utility function u over wealth with $u' > 0$ and $u'' < 0$. A gamble is represented by a real-valued random variable g representing the possible changes in wealth if the gamble is accepted by the decision maker. An investor (W_0, u) rejects a gamble g if $\mathrm{E}[u(W_0+g)] \leq u(W_0)$ and accepts g if $\mathrm{E}[u(W_0+g)] > u(W_0)$. We only consider gambles with $\mathrm{E}[g] > 0$ and $\Pr(g < 0) > 0$. For simplicity, we assume that gambles take finitely many values. Let $L_g \equiv \max(-g)$ and $M_g \equiv \max g$ denote the maximal loss and maximal gain of g, respectively.

For any gamble g, the AS riskiness index $R^{AS}(g)$ is given by the unique positive solution to the equation

$$E\left[\exp\left(-\frac{1}{R^{AS}(g)}g\right)\right] = 1. \tag{1.49}$$

For any gamble g, the FH riskiness $R^{FH}(g)$ index is given by the unique positive solution to the equation

$$\mathrm{E}\left[\log\left(1+\frac{1}{R^{FH}(g)}g\right)\right]=0. \tag{1.50}$$

(a) Show that the AS riskiness index equals the level of risk tolerance that makes a CARA investor indifferent between accepting and rejecting the gamble. That is, an investor with CARA utility $u(w)=-\exp(-Aw)$ will accept (reject) g if $A<1/R^{AS}(g)$ (if $A\geq 1/R^{AS}(g)$).

(b) Show that the FH riskiness index equals the level of wealth that would make a log utility investor indifferent between accepting and rejecting the gamble. That is, a log investor with wealth $W_0>R^{FH}(g)$ ($W_0\leq R^{FH}(g)$) will accept (reject) g.

(c) Consider binary gambles with a loss of L_g with probability p_L and a gain M_g with probability $1-p_L$. Calculate the values of the two indices for the binary gamble with $L_g=\$100$, $M_g=\$105$, and $p_L=1/2$ (Rabin 2000). Repeat for the binary gamble with $L_g=\$100$, $M_g=\$10,100$, and $p_L=1/2$. (The calculation is analytical for FH but numerical for AS.)

(d) Consider binary gambles with infinite gain, that is, M_g arbitrarily large. Derive explicit formulas for the two indices as a function of L_g and p_L at the limit $M_g\to+\infty$. Explain the intuition behind these formulas. Why do the indices assign nonzero riskiness to gambles with infinite expectation? What happens as $p_L\to 0$?

(e) The Sharpe ratio, defined as the ratio of the mean of a gamble to its standard deviation, $SR(g)\equiv \mathrm{E}[g]/\sqrt{\mathrm{Var}(g)}$,[3] is a widely used measure of risk-adjusted portfolio returns. We can interpret its reciprocal as a riskiness index.

(i) Show by example that the (inverse) Sharpe ratio violates first-order stochastic dominance (and hence second-order stochastic dominance). That is, if gamble h first-order stochastically dominates gamble g, then it is not always true that $SR(h)\geq SR(g)$.

(ii) AS (2008) propose a generalized version of the Sharpe ratio $GSR(g)\equiv \mathrm{E}[g]/R^{AS}(g)$, a measure of "riskiness-adjusted" expected returns. Argue that GSR respects second-order stochastic dominance (and hence first-order stochastic dominance).

(iii) Show that $GSR(g)$ is ordinally equivalent to $SR(g)$ when g is a normally distributed gamble.

Hint: use the probability density function of the normal distribution to show that $R^{AS}(g)=\mathrm{Var}(g)/(2\mathrm{E}[g])$.

[3]The definition of the Sharpe ratio in terms of asset returns is given in equation (2.37) of the next chapter.