CHAPTER 3

Making Choices in Risky Situations

Chapter Outline

3.1 Introduction

The first stage of the equilibrium perspective on asset pricing consists of developing an understanding of the determinants of the **demand** for securities of various risk classes. Individuals demand securities (in exchange for current purchasing power) in their attempt to redistribute income across time and states of nature. This is a reflection of the consumption-smoothing and risk-reallocation function central to financial markets.
Our endeavor requires an understanding of three building blocks:

1. How financial risk is defined and measured.
2. How an investor's attitude toward or tolerance for risk is to be conceptualized and then measured.
3. How investors' risk attitudes interact with the subjective uncertainties associated with the available assets to determine an investor's desired portfolio holdings (demands).

In this and the next chapter, we give a detailed overview of points 1 and 2; point 3 is treated in succeeding chapters.

Intermediate Financial Theory.

3.2 Choosing Among Risky Prospects: Preliminaries

When we think of the "risk" of an investment, we are typically thinking of uncertainty in the future cash-flow stream to which the investment represents title. Depending on the state of nature that may occur in the future, we may receive different payments and, in particular, much lower payments in some states than others. That is, we model an asset's associated cash flow in any future time period as a **random variable**.

Consider, for example, the investments listed in Table 3.1, each of which pays off next period in either of two equally likely states. We index these states by $\theta = 1,2$ with their respective probabilities labeled π_1 and π_2.

First, this comparison serves to introduce the important notion of **dominance**. Investment 3 clearly dominates both investments 1 and 2 in the sense that it pays as much in all states of nature and strictly more in at least one state. The **state-by-state dominance** illustrated here is the strongest possible form of dominance. Without any qualification, we will assume that all rational individuals would prefer investment 3 to either of the other two. Basically, this means that we are assuming the typical individual to be nonsatiated in consumption: she desires more rather than less of the consumption goods these payoffs allow her to buy.

In the case of dominance, the choice problem is trivial and, in some sense, the issue of defining risk is irrelevant. The ranking defined by the concept of dominance is, however, very incomplete. If we compare investments 1 and 2, we see that neither dominates the other. Although it performs better in state 2, investment 2 performs much worse in state 1. There is no ranking possible on the basis of the dominance criterion. The different prospects must be characterized from a different angle. The concept of risk enters necessarily.

On this score, we would probably all agree that investments 2 and 3 are comparatively riskier than investment 1. Of course, for investment 3, the dominance property means that the only risk is an upside risk. Yet, in line with the preference for smooth consumption discussed in Chapter 1, the large variation in date 1 payoffs associated with investment 3 is to be viewed as undesirable in itself. When comparing investments 1 and 2, the qualifier

Table 3.1: Asset payoffs ($)

		Value at $t = 1$	
		$\pi_1 = \pi_2 = 1/2$	
	Cost at $t = 0$	$\theta = 1$	$\theta = 2$
Investment 1	−1000	1050	1200
Investment 2	−1000	500	1600
Investment 3	−1000	1050	1600

"riskier" undoubtedly applies to the latter. In the worst state, the payoff associated with 2 is much lower; in the best state it is substantially higher.

These comparisons can alternatively, and often more conveniently, be represented if we describe investments in terms of their performance on a per dollar basis. We do this by computing the state-contingent rates of return (ROR) that we will typically associate with the symbol r. In the case of the above investments, we obtain the results given in Table 3.2.

One sees clearly that all rational individuals should prefer investment 3 to the other two and that this same dominance cannot be expressed when comparing 1 and 2.

The fact that investment 2 is riskier, however, does not mean that all rational risk-averse individuals would necessarily prefer 1. Risk is not the only consideration, and the ranking between the two projects is, in principle, preference dependent. This is more often the case than not; dominance usually provides a very incomplete way of ranking prospects. This fact suggests we must turn to a description of preferences, the main objective of this chapter.

The most well-known approach at this point consists of summarizing such investment return distributions (i.e., the random variables representing returns) by their mean (Er_i) and variance (σ_i^2), $i = 1,2,3$. The variance (or its square root, the standard deviation) of the rate of return is then naturally used as the measure of "risk" of the project (or the asset). For the three investments just listed, we have:

$$Er_1 = 12.5\%; \quad \sigma_1^2 = \frac{1}{2}(5-12.5)^2 + \frac{1}{2}(20-12.5)^2 = (7.5)^2, \quad \text{or} \quad \sigma_1 = 7.5\%$$

$$Er_2 = 5\%; \quad \sigma_2 = 55\% \ \text{(similar calculation)}$$

$$Er_3 = 32.5\%; \quad \sigma_3 = 27.5\%$$

If we decided to summarize these return distributions by their means and variances only, investment 1 would clearly appear more attractive than investment 2: it has both a higher mean return and a lower variance. In terms of the mean–variance criterion, investment 1 dominates investment 2; 1 is said to *mean–variance dominate* 2. Our previous discussion makes it clear that **mean–variance dominance** neither implies nor is implied by state-by-state dominance. Investment 3 mean–variance dominates 2 but not 1, although it dominates

Table 3.2: State-contingent ROR (r)

	$\theta = 1$	$\theta = 2$
Investment 1	5%	20%
Investment 2	−50%	60%
Investment 3	5%	60%

them both on a state-by-state basis! This is surprising and should lead us to be cautious when using any mean–variance return criterion. Later, we will detail circumstances where it is fully reliable. At this point, let us anticipate that it is not generally so and that restrictions will have to be imposed to legitimize its use.

The notion of mean–variance dominance, which plays a prominent role in modern portfolio theory, can be expressed in the form of a criterion for selecting investments of equal magnitude:

1. For investments of the same Er, choose the one with the lowest σ.
2. For investments of the same σ, choose the one with the greatest Er.

In the framework of modern portfolio theory, one could not understand a rational agent choosing investment 2 rather than investment 1.

We cannot limit our inquiry to the concept of dominance, however. Mean–variance dominance provides only an incomplete ranking among uncertain prospects, as Table 3.3 illustrates.

When we compare these two investments, we do not clearly see which is best; there is no dominance in either state-by-state or mean–variance terms. Investment 5 is expected to pay 1.25 times the expected return of investment 4, but, in terms of standard deviation, it is also 3 times riskier. The choice between 4 and 5, when restricted to mean–variance characterizations, would require specifying the terms at which the decision maker is willing to *substitute* expected return for a given risk reduction. In other words, what decrease in expected return is the decision maker willing to accept for a 1% decrease in the standard deviation of returns? Or, conversely, does the 1 percentage point additional expected return associated with investment 5 adequately compensate for the (3 times) larger risk? Responses to such questions are preference dependent (i.e., they vary from individual to individual).

Suppose, for a particular individual, the terms of the trade-off are well represented by the index E/σ. Since $(E/\sigma)_4 = 4$ while $(E/\sigma)_5 = 5/3$, investment 4 is better than investment 5 for that individual. Of course, another investor may be less risk averse; i.e., he may be willing to accept more extra risk for the same expected return. For example, his preferences may be

Table 3.3: State-contingent ROR (r)

	$\theta = 1$	$\theta = 2$
Investment 4	3%	5%
Investment 5	2%	8%
	$\pi_1 = \pi_2 = \frac{1}{2}$ $ER_4 = 4\%;\ \sigma_4 = 1\%$ $ER_5 = 5\%;\ \sigma_5 = 3\%$	

adequately represented by $(E - 1/3\sigma)$ in which case he would rank investment 5 (with an index value of 4) above investment 4 (with a value of $3\frac{2}{3}$).[1]

All these considerations strongly suggest that we have to adopt a more general viewpoint for comparing potential return distributions. This viewpoint is part of utility theory, to which we now turn after describing some of the problems associated with the empirical characterization of return distributions in Box 3.1.

BOX 3.1 Computing Means and Variances in Practice

Useful as it may be conceptually, calculations of distribution moments such as the mean and the standard deviation are difficult to implement in practice: we rarely know what the *future* states of nature are, let alone their probabilities. We also do not know the returns in each state. A frequently used proxy for a future return distribution is its historical distribution. This amounts to selecting a historical time period and a periodicity, say monthly prices for the past 60 months, and computing the historical (net) returns as follows:

a. Discrete compounding

$$r_j^e = \text{(net) return to stock ownership in month } j = ((q_j^e + d_j)/q_{j-1}^e) - 1$$

where q_j^e is the price of the stock in month j, and d_j its dividend, if any, that month; $1 + r_j^e$ is referred to as the gross return. We then summarize the past distribution of stock returns by the average historical return and the variance of the historical returns. By doing so, we, in effect, assign an equal probability of $\frac{1}{60}$ to each past observation or event.

b. Continuous compounding

To understand how to compute period-by-period returns "under continuous compounding," we must first explain what this convention entails. Conceptually, continuous compounding is the result of discrete compounding when the corresponding time interval becomes infinitesimally small. Suppose an investor's wealth is Y_0, which he invests at a rate r for one period (let us say a month as in (a)). If the rate r is continuously compounded over this single period, the cumulative wealth consequence is as follows:

$$Y_0 \mapsto Y_0 \lim_{n \mapsto \infty} \left(1 + \frac{r}{n}\right)^n = Y_0 e^r$$

(*Continued*)

[1] Observe that the proposed index is not immune to the criticism discussed above: investment 3 ($E/\sigma = 1.182$) is inferior to investment 1 ($E/\sigma = 1.667$). Yet, we know that three dominates one because it pays a higher return in every state. This problem is pervasive with the mean–variance investment criterion: whatever the terms of the trade-off between mean and variance or standard deviation, one can produce a paradox such as the one illustrated above. Accordingly, this criterion is not generally applicable without additional restrictions. The index E/σ resembles, but is not identical to, the Sharpe ratio, $(E\tilde{r} - r_f)/\sigma_{\tilde{r}}$, where r_f denotes the risk-free rate.

BOX 3.1 Computing Means and Variances in Practice (Continued)

which exceeds the cumulative effect under discrete compounding ($Y_0 e^r > Y_0(1+r)$ if $r > 0$). It also follows from this identification that if a one period rate r is continuously compounded for a succession of J periods, the cumulative wealth effect will be

$$\begin{array}{lcl} t & & t+J \\ Y_0 & \longmapsto & Y_0 \lim_{n \mapsto \infty} \underbrace{\left\{ \left(1+\frac{r}{n}\right)^n \left(1+\frac{r}{n}\right)^n \cdots \left(1+\frac{r}{n}\right)^n \right\}}_{J \text{ terms}} = Y_0 e^{Jr} \end{array}$$

Lastly, if wealth Y_0 is invested successively at the continuously compounded discrete rates r_1 and r_2, the cumulative effect will be

$$t \xrightarrow{r_1} t+1 \xrightarrow{r_2} t+2$$

$$Y_0 \longmapsto Y_0 \lim_{n \mapsto \infty} \left\{ \left(1+\frac{r_1}{n}\right)^n \left(1+\frac{r_2}{n}\right)^n \right\} = Y_0 e^{r_1+r_2};$$

i.e., under continuous compounding rates of return may simply be added to give their cumulative effect. As we will see in subsequent chapters, this additive feature will allow various calculations to be simplified if the continuous compounding convention is assumed. Note also that our notation assigns the "age interpretation" to time periods: period 10, say, corresponds to the end of the 10th time interval just as a child's 10th birthday is celebrated at the conclusion of his 10th year of life.

Given this setting, how are returns computed from discrete data under the continuous compounding convention? Again, let us assume the following data for some stock is presented to us:

$$\begin{array}{cc} t+j-1 & t+j \\ q^e_{j-1} & q^e_j + \mathrm{div}_j \end{array}$$

(*Continued*)

BOX 3.1 Computing Means and Variances in Practice (Continued)

We may now ask the question: what discrete rate of return $r_j^{e,\text{cont.}}$, *when continuously compounded*, was earned by this stock in the course of period j? Equivalently, what rate of return $r_j^{e,\text{cont.}}$ satisfies:

$$q_{j-1}^e \, e^{r_j^{e,\text{ cont.}}} = q_j^e + \text{div}_j?$$

Thus,

$$r_j^{e,\text{ cont.}} = \ell n \left(\frac{q_j^e + \text{div}_j}{q_{j-1}^e} \right) = \ln(1 + r_j^e)$$

In what follows in this book, the unspoken assumption is that reported return data is computed under the continuous compounding convention as per the above calculation. Accordingly, we generally do not employ the "cont." superscript.[2]

Using the pattern of historical returns (however measured) to infer properties of the future return distribution makes sense if we think the "mechanism" generating these returns is "stationary": that the future will in some sense closely resemble the past. In practice, this hypothesis is rarely fully verified and, at the minimum, it requires careful checking.[3]

[2] When returns are small, $r_j \approx \ln(1 + r_j)$. This is a standard approximation. Note that continuously compounded returns are the natural logarithm of discrete gross returns and, in this sense, are the continuously compounded cointegral to discrete gross returns.

[3] The accuracy of the mean and variance estimates from historical data as stand-ins to the underlying return distribution's true (future) mean and variance is a topic of great significance and to which we will return (c.f. Chapter 7).

3.3 A Prerequisite: Choice Theory Under Certainty

A good deal of financial economics is concerned with how people make choices. The objective is to understand the systematic part of individual behavior and to be able to predict (at least in a loose way) how an individual will react under specific economic circumstances. Economic theory describes individual behavior as the result of a process of optimization under constraints, the objective to be reached being determined by individual preferences, and the constraints being a function of the person's income or wealth level and of market prices. This approach, which defines the *homo economicus* and the notion of **economic rationality**, is justified by the fact that individual behavior is predictable only to the extent that it is systematic, which must mean that there is an attempt to achieve a well-defined objective. It is not to be taken literally or normatively.[4]

[4] By this we mean that economic science does not *prescribe* that individuals maximize, optimize, or simply behave as if they were doing so. It just finds it productive to summarize the systematic behavior of economic agents with such tools.

To develop this sense of rationality systematically, we begin by summarizing the objectives of investors in the most basic way: we postulate the existence of a preference relation, represented by the symbol $\succeq$, describing investors' ability to compare various bundles of goods and services. For two bundles a and b, the expression

$$a \succeq b$$

is to be read as follows: For the investor in question, bundle a is either strictly preferred to bundle b, or he is indifferent between them. Pure indifference is denoted by $a \sim b$, strict preference by $a \succ b$.

The notion of economic rationality can then be summarized by the following assumptions:

A.1 Every investor possesses such a preference relation and it is *complete*, meaning that he is able to decide whether he prefers a to b, b to a, or both, in which case he is indifferent with respect to the two bundles. That is, for any two bundles a and b, either $a \succeq b$ or $b \succeq a$, or both. If both hold, we say that the investor is indifferent with respect to the bundles and write $a \sim b$.

A.2 This preference relation satisfies the fundamental property of transitivity: For any bundles a, b, and c, if $a \succeq b$ and $b \succeq c$, then $a \succeq c$.

A further requirement is also necessary for technical reasons:

A.3 The preference relation $\succeq$ is continuous in the following sense: Let $\{x_n\}$ and $\{y_n\}$ be two sequences of consumption bundles such that $x_n \mapsto x$ and $y_n \mapsto y$.[5] If $x_n \succeq y_n$ for all n, then the same relationship is preserved in the limit $x \succeq y$.

A key result can now be expressed in the following proposition.

Theorem 3.1 Assumptions A.1 through A.3 are sufficient to guarantee the existence of a continuous, time-invariant, real-valued utility function[6] u, such that for any two objects of choice (consumption bundles of goods and services),

$$a \succeq b \quad \text{if and only if} \quad u(a) \geq u(b).$$

Proof See, for example, Mas-Colell et al. (1995), Proposition 3.C.1.

This result asserts that to endow decision makers with a utility function (which they are assumed to maximize) is, in reality, no different than to assume their preferences among objects of choice define a relation possessing the (weak) properties summarized in A.1 through A.3.

5 We use the standard sense of (normed) convergence in R^N.
6 In other words, u: $R^N \to R^+$.

Note that Theorem 3.1 implies that if $u(\ )$ is a valid representation of an individual's preferences, any increasing transformation of $u(\ )$ is also valid since such a transformation, by definition, will preserve the ordering induced by $u(\ )$. Note also that the notion of a consumption bundle is, formally, very general. Different elements in a bundle may represent the consumption of the same good or service in different time periods. One element might represent a vacation trip in the Bahamas this year; another may represent exactly the same vacation next year. We can further expand our notion of different goods to include the same good consumed in mutually exclusive states of the world. Our preference for hot soup, for example, may be very different if the day is warm rather than cold. These thoughts suggest that Theorem 3.1 is really quite general and can, formally at least, be extended to accommodate uncertainty. Under uncertainty, however, ranking bundles of goods (or vectors of monetary payoffs, see below) involves more than pure elements of taste or preferences. In the hot soup example, it is natural to suppose that our preferences for hot soup are affected by the probability we attribute to the day being hot or cold. Disentangling pure preferences from probability assessments is the subject to which we now turn.

3.4 Choice Theory Under Uncertainty: An Introduction

Under certainty, the choice is among consumption baskets with known characteristics. Under uncertainty, however, our emphasis changes. The objects of choice are typically no longer consumption bundles but vectors of state-contingent money payoffs (we will reintroduce consumption in Chapter 5). Such vectors are formally what we mean by an investment or an *asset* available for purchase. When we purchase a share of a stock, for example, we know that its sale price in one year will differ depending on what events transpire within the firm and in the world economy. Under financial uncertainty, therefore, the choice is among alternative investments leading to different possible income levels and, hence, ultimately different consumption possibilities. As before, we observe that people do make investment choices, and if we are to make sense of these choices, there must be a stable underlying order of preference defined over different alternative investments. The spirit of Theorem 3.1 will still apply. With appropriate restrictions, these preferences can be represented by a utility index defined on investment possibilities, but obviously something deeper is at work. It is natural to assume that individuals have no intrinsic taste for the assets themselves (IBM stock as opposed to Royal Dutch Petroleum stock (hereafter RDS), for example). Rather, they are interested to know what payoffs these assets will yield and with what likelihood (see Box 3.2, however).

BOX 3.2 Investing Close to Home

Although the assumption that investors only care for the final payoff of their investment without any trace of "romanticism" is standard in financial economics, there is some evidence to the contrary and, in particular, for the assertion that many investors, at the margin at least, prefer to purchase the claims of firms whose products or services are familiar to them.
In particular, Huberman (2001) examines the stock ownership records of the seven regional Bell operating companies (RBOCs) (due to a series of mergers, these seven firms have combined into presently two entities, Verizon and AT&T). He discovered that, with the exception of residents of Montana, Americans were more likely to invest in their local RBOC than in any other. When they did, their holdings averaged $14,400. For those who ventured farther from home and hold stocks of the RBOC of a region other than their own, the average holding is only $8246. Considering that every local RBOC cannot be a better investment choice than all of the other six, Huberman interprets his findings as suggesting investors' psychological need to feel comfortable with where they put their money.

One may further hypothesize that investor preferences are indeed very simple after uncertainty is resolved: They prefer a higher monetary payoff to a lower one or, equivalently, to earn a higher return rather than a lower one. Of course they do not know *ex ante* (i.e., before the state of nature is revealed) which asset will yield the higher payoff. They have to choose among prospects, or probability distributions representing these payoffs. And, as we saw in Section 3.2, typically, no one investment prospect will strictly dominate the others. Investors will be able to imagine different possible scenarios, some of which will result in a higher return for one asset, with other scenarios favoring other assets. For instance, let us go back to our favorite situation where there are only two states of nature; in other words, two conceivable scenarios and two assets, as seen in Table 3.4.

There are two key ingredients in the choice between these two alternatives. The first is the probability of the two states. All other things being the same, the more likely is state 1, the more attractive IBM stock will appear to prospective investors. The second is the ex post (once the state of nature is known) level of utility provided by the investment. In Table 3.4, IBM yields $100 in state 1 and is thus preferred to RDS, which yields $90 if this scenario is realized. RDS, however, provides $160 rather than $150 in state 2. Obviously, with

Table 3.4: Forecasted price per share in one period

	State 1	State 2
IBM	$100	$150
RDS	$90	$160
	Current price of both assets is $100.	

unchanged state probabilities, things would look different if the difference in payoffs were increased in one state as in Table 3.5.

Here even if state 1 is slightly more likely, the superiority of RDS in state 2 makes it look more attractive. A more refined perspective is introduced if we go back to our first scenario but now introduce a third contender, Sony, with payoffs of $90 and $150, as seen in Table 3.6.

Sony is dominated by both IBM and RDS. But the choice between the latter two can now be described in terms of an improvement of $10 over the Sony payoff, either in state 1 or in state 2. Which is better? The relevant feature is that IBM adds $10 when the payoff is low ($90), while RDS adds the same amount when the payoff is high ($150). Most people would think IBM more desirable, and with equal state probabilities, would prefer IBM. Once again this is an illustration of the preference for smooth consumption (smoother income allows for smoother consumption).[7] In the present context, one may equivalently speak of risk aversion or of the well-known microeconomic assumption of decreasing marginal utility (the incremental utility steadily declines when adding ever more consumption or income).

The expected utility theorem provides a set of hypotheses under which an investor's preference ranking over investments with uncertain money payoffs may be represented by a utility index combining, in the most elementary way (i.e., linearly), the two ingredients just

Table 3.5: Forecasted price per share in one period

	State 1	State 2
IBM	$100	$150
RDS	$90	$200
	Current price of both assets is $100.	

Table 3.6: Forecasted price per share in one period

	State 1	State 2
IBM	$100	$150
RDS	$90	$160
Sony	$90	$150
	Current price of all assets is $100.	

[7] Of course, for the sake of our reasoning, one must assume that nothing else important is going on simultaneously in the background, and that other things, such as income from other sources, if any, and the prices of the consumption goods to be purchased with the assets' payoffs, are unchanged irrespective of what the payoffs actually are.

discussed—the preference ordering on the ex post money payoffs and the respective probabilities of these payoffs.

We first illustrate this notion in the context of the two assets considered earlier. Let the respective probability distributions on the price per share of IBM and RDS be described, respectively, by $\tilde{p}_{\mathrm{IBM}} = p_{\mathrm{IBM}}(\theta_i)$ and $\tilde{p}_{\mathrm{RDS}} = p_{\mathrm{RDS}}(\theta_i)$ together with the probability π_i that the state of nature θ_i will be realized. In a two-state context, the expected utility theorem provides sufficient conditions on an agent's preferences over uncertain asset payoffs, denoted $\succeq$, such that there exists a function $\mathbb{U}(\ )$, defined over uncertain asset payoffs, and an associated utility-of-money function $U(\ )$ such that

i. $\tilde{p}_{\mathrm{IBM}} \succeq \tilde{p}_{\mathrm{RDS}}$ if and only if $\mathbb{U}(\tilde{p}_{\mathrm{IBM}}) \geq \mathbb{U}(\tilde{p}_{\mathrm{RDS}})$ where

ii. $\mathbb{U}(\tilde{p}_{\mathrm{IBM}}) = EU(\tilde{p}_{\mathrm{IBM}}) = \pi_1 U(p_{\mathrm{IBM}}(\theta_1)) + \pi_2 U(p_{\mathrm{IBM}}(\theta_2))$

$$> \pi_1 U(p_{\mathrm{RDS}}(\theta_1)) + \pi_2 U(p_{\mathrm{RDS}}(\theta_2)) = EU(\tilde{p}_{\mathrm{RDS}}) = \mathbb{U}(\tilde{p}_{\mathrm{RDS}})$$

More generally, for these preferences, the utility of any asset A with payoffs $p_A(\theta_1)$, $p_A(\theta_2), \ldots, p_A(\theta_N)$ in the N possible states of nature with probabilities $\pi_1, \pi_2, \ldots, \pi_N$ can be represented as

$$\mathbb{U}(\tilde{p}_A) = EU(p_A(\theta_i)) = \sum_{i=1}^{N} \pi_i U(p_A(\theta_i))$$

In other words, by the weighted mean of ex post utilities using the state probabilities as weights. $\mathbb{U}(\tilde{p}_A)$ is a real number. Its precise numerical value, however, has no more meaning than if you are told that the temperature is 40° when you do not know if the scale being used is Celsius or Fahrenheit. It is useful, however, for comparison purposes. By analogy, if it is 40° today, but it will be 45° tomorrow, you at least know it will be warmer tomorrow than it is today. Similarly, the expected utility number is useful because it permits attaching a number to a probability distribution and this number is, under appropriate hypotheses, a good representation of the relative ranking of a particular member of a family of probability distributions (assets under consideration).

3.5 The Expected Utility Theorem

We elect to discuss this theorem in the simple context where objects of choice take the form of simple lotteries. A generic lottery will be denoted (x,y,π); it offers payoff (consequence) x with probability π and payoff (consequence) y with probability $1-\pi$. This notion of a lottery is actually very general and encompasses a huge variety of possible payoff structures. For example, x and y may represent specific monetary payoffs as in Figure 3.1, or x may be a payment while y is a lottery as in Figure 3.2, or even x and y may both be lotteries as in Figure 3.3. Extending these possibilities, some or all of the x_i's

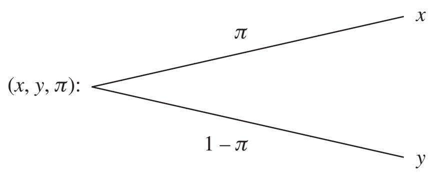


Figure 3.1
A simple lottery (*x*, *y* are monetary payoffs).

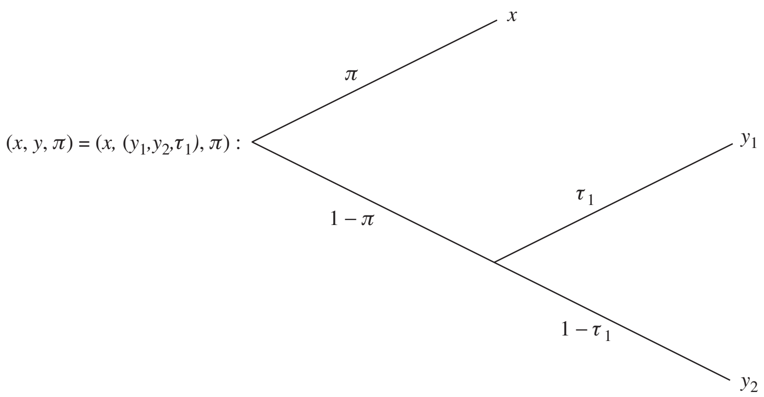


Figure 3.2
A compound lottery (*y* is itself a lottery).

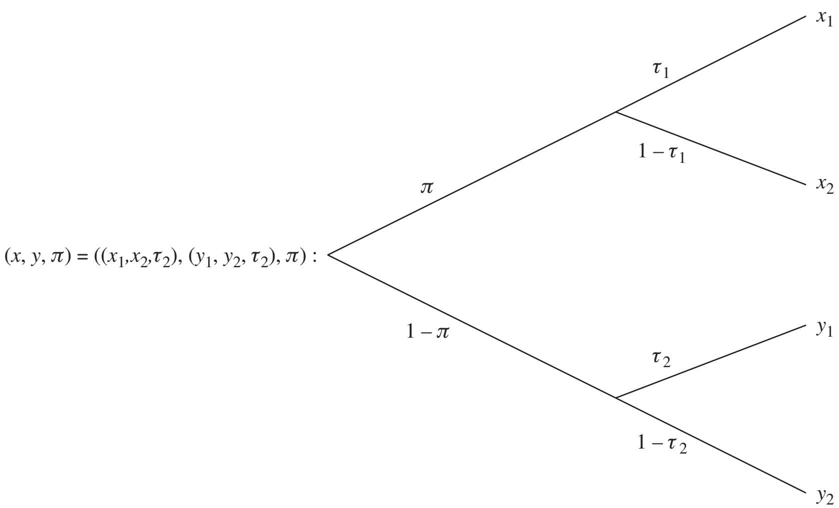


Figure 3.3
A compound lottery (both *x* and *y* are themselves lotteries).

and y_i's may themselves be lotteries, and so on. We also extend our choice domain to include individual payments, lotteries where there is one, certain, monetary payoff; for instance,

$$(x, y, \pi) = x \text{ if (and only if) } \pi = 1 \text{ (see axiom C.1)}$$

Moreover, the theorem holds as well for assets paying a continuum of possible payoffs, but our restriction to discrete payoffs makes the necessary assumptions and justifying arguments easily accessible. Our objective is conceptual transparency rather than absolute generality. All the results extend to much more general settings.

Under these representations, we will adopt the following **axioms and conventions**:

C.1. a. $(x,y,1) = x$

b. $(x,y,\pi) = (y,x, 1 - \pi)$

c. $(x,z,\pi) = (x,y,\pi + (1 - \pi)\tau)$ if $z = (x,y,\tau)$

C.1c informs us that agents are concerned with the net cumulative probability of each outcome. Indirectly, it further accommodates lotteries with multiple outcomes; see Figure 3.4, for an example with lotteries (x, y, π'), and $(z, w, \hat{\pi})$, where $\pi_1 = \pi'$, $\pi = \pi_1 + \pi_2$, etc.

C.2. There exists a preference relation $\succeq$, defined on lotteries, which is complete and transitive.

C.3. The preference relation is continuous in the sense of A.3 in Section 3.3.

By C.2 and C.3 alone, we know (Theorem 3.1) that there exists a utility function, which we will denote by $\mathbb{U}(\ )$, defined both on lotteries and on specific payments

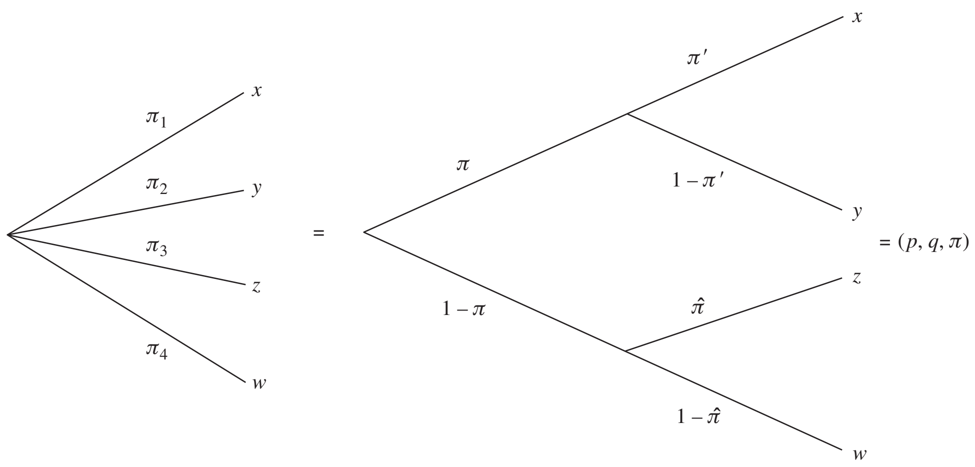


Figure 3.4
A lottery with multiple outcomes reinterpreted as a compound lottery.

since, by assumption C.1a, a payment may be viewed as a (degenerate) lottery. For any payment x, we identify

$$U(x) = \mathbb{U}((x, y, 1)) \tag{3.1}$$

Our remaining assumptions are thus necessary only to guarantee that $\mathbb{U}$ assumes the expected utility form.

C.4. Independence of irrelevant alternatives. Let (x, y, π) and (x, z, π) be any two lotteries; then, $y \succeq z$ if and only if $(x, y, \pi) \succeq (x, z, \pi)$.

C.5. For simplicity, we also assume that there exists a best (i.e., most preferred lottery), b, as well as a worst, least desirable, lottery w.

In our argument to follow (which is constructive, i.e., we explicitly exhibit the expected utility function), it is convenient to use relationships that follow directly from these latter two assumptions. In particular, we will use C.6 and C.7:

C.6. Let x, k, z be consequences or payoffs for which $x > k > z$. Then there exists a π such that $(x, z, \pi) \sim k$.

C.7. Let $x \succ y$. Then $(x, y, \pi_1) \succeq (x, y, \pi_2)$ if and only if $\pi_1 > \pi_2$. This follows directly from C.4.

Theorem 3.2 Consider a preference ordering, defined on the space of lotteries, that satisfies axioms C.1 to C.7. Then there exists a utility function $\mathbb{U}$ defined on the lottery space, with associated utility-of-money function $U(\ )$, such that:

$$\mathbb{U}((x, y, \pi) = \pi U(x) + (1 - \pi)U(y) \tag{3.2}$$

Proof We outline the proof in a number of steps:

By Theorem 3.1 we know that $\mathbb{U}(\ )$ exists with associated $U(\ )$ as its restriction to certain monetary payments, defined as per Eq. (3.1). We must now show that $\mathbb{U}(\ )$ and $U(\ )$ are related by Eq. (3.2).

1. Without loss of generality, we may normalize $\mathbb{U}(\ )$ so that $\mathbb{U}(b) = 1$, $\mathbb{U}(w) = 0$.
2. For all other lotteries z, define $\mathbb{U}(z) = \pi_z$ where π_z satisfies $(b, w, \pi_z) \sim z$

   Constructed in this way $\mathbb{U}(z)$ is well defined since

   a. by C.6, $\mathbb{U}(z) = \pi_z$ exists, and
   b. by C.7, $\mathbb{U}(z)$ is unique. To see this latter implication, assume, to the contrary, that $\mathbb{U}(z) = \pi_z$ and also $\mathbb{U}(z) = \pi'_z$ where $\pi_z > \pi'_z$. By assumption C.4,

   $$z \sim (b, w, \pi_z) \succ (b, w, \pi'_z) \sim z, \quad \text{a contradiction}$$

3. It follows also from C.7 that if $m \succ n$, $\mathbb{U}(m) = \pi_m > \pi_n = \mathbb{U}(n)$. Thus, $\mathbb{U}(\ )$ has the property of a utility function.

4. Lastly, we want to show that $\mathbb{U}(\ )$ has the required property. Let x, y be monetary payments, π a probability. By C.1a, $U(x)$, $U(y)$ are well-defined real numbers. By C.6,

$$\begin{aligned}(x, y, \pi) &\sim ((b, w, \pi_x), (b, w, \pi_y)), \pi) \\ &\sim (b, w, \pi\pi_x + (1-\pi)\pi_y), \text{by C.1c.}\end{aligned}$$

Thus, by definition of $\mathbb{U}(\ )$,

$$\mathbb{U}((x, y, \pi)) = \pi\pi_x + (1-\pi)\pi_y = \pi U(x) + (1-\pi)U(y)$$

Although we have chosen x, y as monetary payments, the same conclusion holds if they are lotteries.

Before going on to refine our understanding of the expected utility theorem, it is important to be absolutely clear on terminology: First, the overall $\mathbb{U}(\ )$ is defined over lotteries. It is referred to as the von Neumann−Morgenstern (VNM) utility function, so named after the originators of the theory, the justly celebrated mathematicians John von Neumann and Oskar Morgenstern. In the construction of a VNM utility function, it is customary first to specify its restriction to certainty monetary payments, the so-called utility-of-money function $U(\ )$ or simply (and hereafter) the *utility function*. Note that the VNM utility function and its associated utility function are not the same. The VNM utility function is defined over uncertain asset payoff structures, while its associated utility function is defined over individual money payments.

The key identifier of the "expected utility" construct is that these two concepts are linearly related: either the utility function is linearly related to the VNM function via the probability weights or the VNM function is linearly related to the state probabilities, with weights being the state-by-state utility-of-money values. The second interpretation leads to the common expression that VNM-expected utility preferences are "linear in the probabilities."

Given the objective specification of probabilities (thus far assumed), it is the utility function that uniquely characterizes an investor. As we will see shortly, different additional assumptions on $U(\ )$ will identify an investor's tolerance for risk. We do, however, impose the maintained requirement that $U(\ )$ be increasing for all candidate utility functions (more money is preferred to less). Note also that the expected utility theorem confirms that investors are concerned only with an asset's final payoffs and the cumulative probabilities of achieving them. For expected utility investors, the structure of uncertainty resolution is thus irrelevant (Axiom C.1c).[8]

Although the introduction to this chapter concentrates on comparing rates of return distributions, our expected utility theorem in fact gives us a tool for comparing different asset

[8] See Section 5.7.1 for a generalization on this score.

payoff distributions. Without further analysis, it does not make sense to think of the utility function as being defined over a rate of return. This is true for a number of reasons. First, returns are expressed on a per unit (per US$, Swiss CHF, etc.) basis and do not identify the magnitude of the initial investment to which these rates are to be applied. We thus have no way to assess the implications of a return distribution for an investor's wealth position. It could, in principle, be anything. Second, the notion of a rate of return implicitly suggests a time interval: the payout is received after the asset is purchased. So far we have only considered the atemporal evaluation of uncertain investment payoffs. In Chapter 4, we generalize the VNM representation to preferences defined over rates of returns.

As in the case of a general order of preferences over bundles of commodities, the VNM-expected utility representation is preserved under linear transformations. If $\mathbb{U}(\ )$ is a von Neuman–Morgenstern utility function, then $\mathbb{V}(\ ) = a\mathbb{U}(\ ) + b$, where $a > 0$, is also such a function. To verify this assertion, let (x, y, π) be some uncertain payoff, and let $U(\ )$ be the utility-of-money function associated with $\mathbb{U}$.

$$\begin{aligned}\mathbb{V}((x, y, \pi)) &= a\mathbb{U}((x, y, \pi) + b = a[\pi U(x) + (1-\pi)U(y)] + b \\ &= \pi[a\mathbb{U}(x) + b] + (1-\pi)[a\mathbb{U}(y) + b] = \pi\mathbb{V}(x) + (1-\pi)\mathbb{V}(y)\end{aligned}$$

Every linear transformation of an expected utility function is thus also an expected utility function. The utility-of-money function associated with $\mathbb{V}$ is $[aU(\ ) + b]$; $\mathbb{V}(\ )$ represents the same preference ordering over uncertain payoffs as $\mathbb{U}(\ )$. On the other hand, a nonlinear transformation does not always respect the preference ordering. It is in that sense that utility is said to be **cardinal**.

Lastly, we need to clarify the direct connection between $U(\ )$ and $u(\ )$ (c.f., Theorem 3.1). Economic science recognizes that money has no value *per se*; its significance lies in the consumption goods that may be purchased with it. Accordingly, consider some financial asset (portfolio) that pays $(Y(\theta_1),\ldots, Y(\theta_N))$, with $Y(\theta_i)$ denoting its money payoff in state θ_i, $i = 1, 2,\ldots, N$. Suppose also that the investor who acquires this asset has available to him J distinct consumption goods in each of the states. The proportions of the consumption goods the investor elects to consume may differ across states reflecting potentially different state-contingent prices as denoted by $(P_1(\theta_i), P_2(\theta_i),\ldots, P_J(\theta_i))$, where $P_j(\theta_i)$ is the price of good j in state θ_i.

Presuming the investor wishes to spend his money as wisely as possible irrespective of what state may be realized, we can define the utility-of-money function $U(Y(\theta_i))$ as the maximum level of consumption utility he may achieve in state i given his income $Y(\theta_i)$ and the consumption goods prices noted above; in effect, we define $U(Y(\theta_i))$ by

$$U(Y(\theta_i)) \equiv_{def} \max_{\{c_1(\theta_i),\ldots,c_J(\theta_i)\}} u(c_1(\theta_i), \ldots, c_J(\theta_i)) \tag{3.3}$$

$$\text{s.t. } c_1(\theta_i)P_1(\theta_i) + \cdots + c_J(\theta_i)P_J(\theta_i) \leq Y(\theta_i)$$

where $c_j(\theta_i)$ is the consumption of good j in state θ_i. The constraint is referred to as the investor's budget constraint. Note that $U(Y(\theta_i))$ subsumes three important quantities: the investor's relative preference for the different goods available in state θ_i (as per $u(\ )$), the relative prices of these goods, and the investor's state θ_i income, $Y(\theta_i)$.

A fuller treatment of this identification would also acknowledge that investors typically save some portion of their income. This consideration requires a multiperiod setting and comes to the fore beginning in Chapter 4.

3.6 How Restrictive Is Expected Utility Theory? The Allais Paradox

Although apparently innocuous, the above set of axioms has been hotly contested as representative of rationality. In particular, it is not difficult to find situations in which investor preferences violate the independence axiom. Consider the following four possible asset payoffs (lotteries):

$$L^1 = (10{,}000, 0, 0.1) \quad L^2 = (15{,}000, 0, 0.09)$$
$$L^3 = (10{,}000, 0, 1) \quad L^4 = (15{,}000, 0, 0.9)$$

When investors are asked to rank these payoffs, the following ranking is frequently observed:

$$L^2 \succ L^1$$

(presumably because L^2's positive payoff in the favorable state is much greater than L^1's while the likelihood of receiving it is only slightly smaller) and

$$L^3 \succ L^4$$

(Here it appears that the certain prospect of receiving 10,000 is worth more than the potential of an additional 5000 at the risk of receiving nothing.)

By the structure of compound lotteries, however, it is easy to see that:

$$L^1 = (L^3, L^0, 0.1)$$
$$L^2 = (L^4, L^0, 0.1) \quad \text{where } L^0 = (0, 0, 1)$$

By the independence axiom, the ranking between L^1 and L^2, on the one hand, and L^3 and L^4, on the other, should thus be identical!

This is the Allais paradox.[9] There are a number of possible reactions to it.

1. Yes, my choices were inconsistent; let me think again and revise them.
2. No, I'll stick to my choices. The following kinds of things are missing from the theory of choice expressed solely in terms of asset payoffs:
   - the pleasure of gambling, and/or
   - the notion of regret.

The idea of regret is especially relevant to the Allais paradox, and its application in the prior example would go something like this. L^3 is preferred to L^4 because of the regret involved in receiving nothing if L^4 were chosen and the bad state ensued. We would, at that point, regret not having chosen L^3, the certain payment. The expected regret is high because of the nontrivial probability (0.10) of receiving nothing under L^4. On the other hand, the expected regret of choosing L^2 over L^1 is much smaller (the probability of the bad state is only 0.01 greater under L^2, and in either case the probability of success is small), and insufficient to offset the greater expected payoff. Thus L^2 is preferred to L^1.

The Allais paradox is but the first of many phenomena that appear to be inconsistent with standard preference theory. Another prominent example is the general pervasiveness of *preference reversals*, events that may approximately be described as follows. Individuals participating in controlled experiments were asked to choose between two lotteries, (4, 0, 0.9) and (40, 0, 0.1). More than 70% typically chose (4, 0, 0.9). When asked at what price they would be willing to sell the lotteries if they were to own them, however, a similar percentage demanded the higher price for (40, 0, 0.1). At first appearances, these choices would seem to violate transitivity. Let *x, y* be, respectively, the sale prices of (4, 0, 0.9) and (40, 0, 0.10). Then this phenomenon implies

$$x \sim (4, 0, 0.9) \succ (40, 0, 0.1) \sim y, \quad \text{yet} \quad y > x$$

Alternatively, it may reflect a violation of the assumed principle of procedure invariance, which is the idea that investors' preference for different objects should be indifferent to the manner by which their preference is elicited. Surprisingly, more narrowly focused experiments, which were designed to force a subject with expected utility preferences to behave consistently, gave rise to the same reversals. The preference reversal phenomenon could thus, in principle, be due either to preference intransitivity or to a violation of the independence axiom, or of procedure invariance.

Through a series of carefully constructed experiments, some researchers have attempted to assign responsibility for preference reversals to procedure invariance violations. But this is a particularly alarming conclusion as Thaler (1992) notes. It suggests that "*the context and*

[9] Named after the Nobel Prize-winner Maurice Allais who was the first to uncover the phenomenon. See Allais (1964).

procedures involved in making choices or judgements influence the preferences that are implied by the elicited responses. In practical terms this implies that (economic) behavior is likely to vary across situations which economists (would otherwise) consider identical." This is tantamount to the assertion that the notion of a preference ordering is not well defined. While investors may be able to express a consistent (and thus mathematically representable) preference ordering across television sets with different features (e.g., size of the screen and quality of the sound), this may not be possible with lotteries or consumption baskets containing widely diverse goods.

Grether and Plott (1979) summarize this conflict in the starkest possible terms:

> *Taken at face value, the data demonstrating preference reversals are simply inconsistent with preference theory and have broad implications about research priorities within economics. The inconsistency is deeper than the mere lack of transitivity or even stochastic transitivity. It suggests that no optimization principles of any sort lie behind the simplest of human choices and that the uniformities in human choice behavior which lie behind market behavior result from principles which are of a completely different sort from those generally accepted.*

At this point it is useful to remember, however, that the ultimate goal of financial economics is not to describe individual, but rather market, behavior. There is a real possibility that occurrences of seeming individual irrationality essentially "wash out" when aggregated at the market level. On this score, the proof of the pudding is in the eating and we have little alternative but to see the extent to which the basic theory of choice we are using is able to illuminate financial phenomena of interest. All the while, the discussion above should make us alert to the possibility that unusual phenomena might be the outcome of deviations from the generally accepted preference theory articulated above. While there is, to date, no preference ordering that accommodates preference reversals—and it is not clear there will ever be one—more general constructs than expected utility have been formulated to admit other, seemingly contradictory, phenomena. Further complications arise under collective choice; see Box 3.3.

BOX 3.3 On the Rationality of Collective Decision Making

Although the discussion in the text pertains to the rationality of individual choices, it is a fact that many important decisions are the result of collective decision making. The limitations to such a process are important and, in fact, better understood than those arising at the individual level. It is easy to imagine situations in which transitivity is violated once choices result from some sort of aggregation over more basic preferences.

Consider three portfolio managers who decide which stocks to add to the portfolios they manage by majority voting. The stocks currently under consideration are General Electric

(*Continued*)

BOX 3.3 On the Rationality of Collective Decision Making (Continued)

(GE), Daimler (DAI), and Sony (S). Based on his fundamental research and assumptions, each manager has rational (i.e., transitive) preferences over the three possibilities:

Manager 1: GE $\succsim_1$ DAI $\succsim_1$ S
Manager 2: S $\succsim_2$ GE $\succsim_2$ DAI
Manager 3: DAI $\succsim_3$ S $\succsim_3$ GE

If they were to vote all at once, they know each stock would receive one vote (each stock has its advocate). So they decide to vote on pairwise choices: (GE versus DAI), (DAI versus S), and (S versus GE). The results of this voting (GE dominates DAI, DAI dominates S, and S dominates GE) suggest an intransitivity in the aggregate ordering. Although an intransitivity, it is one that arises from the operation of a collective choice mechanism (voting) rather than being present in the individual orders of preference of the participating agents. There is a large literature on this subject that is closely identified with Arrow's "Impossibility Theorem". See Arrow (1963) for a more exhaustive discussion.

3.7 Behavioral Finance

"The political man of the Greeks, the religious man of the Hebrews and Christians, the enlightened economic man of eighteenth century Europe (the original of that mythical present day character the 'good European') [have] been superseded by a new model for the conduct of life. Psychological man is... more native to American culture than the Puritan sources of that culture would indicate."[10]

The notion of a "rational investor" underlies most of financial theory and, indeed, most of what is presented in this book. A rational investor, very simply, is one with two essential attributes: (i) his preferences over random money payoffs are VNM-expected utility, as just described, and (ii) the probabilities he assigns to these payoffs are objective in that they incorporate all past and present information available to the investor in a manner that respects correct statistical procedure.[11] Unless specified otherwise, VNM-expected utility will represent our default context going forward. But despite the elegant and straightforward nature of the rational investor construct, there remain numerous empirical choice phenomena, which it cannot rationalize. How are they to be understood? One answer to this question lies in the domain of "behavioral finance".

[10] Reiff (2006, page 48). Reiff (2006) is speaking of social trends, but it is not surprising that economic science would be similarly swept along.

[11] In equilibrium contexts, we will go one step further and strengthen the latter requirement to one of "rational expectations." This means that the probabilities objectively computed for the random payoffs in fact coincide with the payoffs' true probability distribution.

Behavioral finance is a theory-in-progress which seeks to fill this gap by departing from the rational investor assumptions in ways that are thought to better reflect various findings in experimental psychology. Most of the assumptions in behavioral finance have not been axiomatized in the context of choices-over-lotteries. Rather, they are supported by circumstantial empirical evidence. We give illustrations of several behavioral notions below. The difficulty in evaluating these concepts in the present context lies precisely in the absence of a formal axiomatic basis underlying them. In that sense we do not yet fully grasp "what they imply."

3.7.1 Framing

"Framing" is simply the notion that individuals' choices may be substantially influenced by the context in which they are presented. As a very simple illustration, would your decision to purchase a steak versus fish for dinner be different if the steak is advertised as:

"90% fat free," or
"10% fat content"?

While *ex post* recognizing that these alternatives convey the same information, it seems apparent that the former description is more likely to elicit a positive "steak" decision for the majority of shoppers.[12]

The same phenomena appear to be present in investment choices. In a classic study, Kahneman and Tversky (1979) explore individual choices across the following lotteries:

i. In the context of first being given $1000, participants were asked to choose between the following lotteries A and B:
   A: ($1000, 0, .5)
   B: ($500, 0, 1)
ii. In a context of first being given $2000, these same participants were asked to choose between
   C: (−$1000, 0, .5)
   D: (−$500, 0, 1).[13]

These lotteries are summarized in Figure 3.5.

[12] Procedure invariance (prior section) and "framing" are not the same notion. Procedure invariance requires that an investor's preference over lotteries be the same irrespective of whether they are directly compared or their certainty equivalents compared; i.e., the ranking is preserved under any methodology as to how the comparison is to be rendered (informally, irrespective of "how the problem is to be solved"). Framing concerns the context of the comparison, once a method of comparison has been chosen.

[13] Kahneman and Tversky (1979) actually conduct their study with payoffs denominated in terms of Israeli currency (lira). At the time, the average monthly income was approximately 3000 lira.

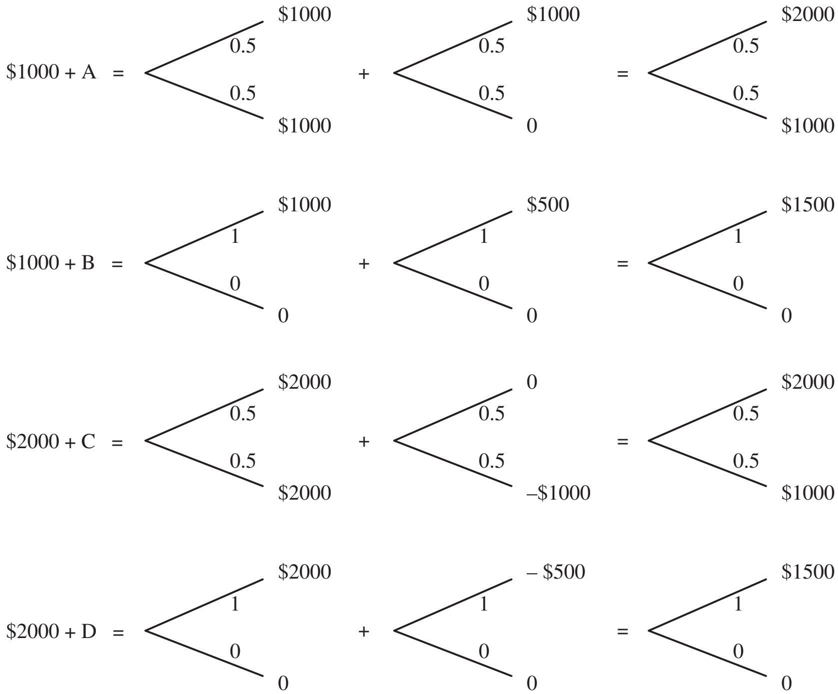


Figure 3.5
Four lotteries with preceding initial payments.

For a majority of those participating in the experiment, B $\succ$ A and C $\succ$ D despite the fact that A and C are equivalent as are B and D when taking full account of the differing initial payments.

How do we interpret these (inconsistent) choices? Apparently, it mattered to most participants that the choice between lotteries A and B was presented as a choice of *gains relative to $1000* while in the second case the choices were presented as *losses relative to $2000*. This distinction is viewed as a manifestation of the phenomenon of framing. Note that under VNM-expected utility, framing, as illustrated above, is irrelevant since only total wealth payoffs matter.

Framing is often cited as one factor potentially contributing to the failure of investors to diversify; i.e., to invest their wealth in portfolios of more than a few (one) assets. It is the idea that when investors consider the acquisition of various assets, they frame the decision on the basis of bilateral comparisons alone, without considering the interaction of multiple asset return patterns and the benefits that may follow. To illustrate this notion in the simple context of three assets and two states of nature θ_1 and θ_2, consider the assets in Figure 3.6.

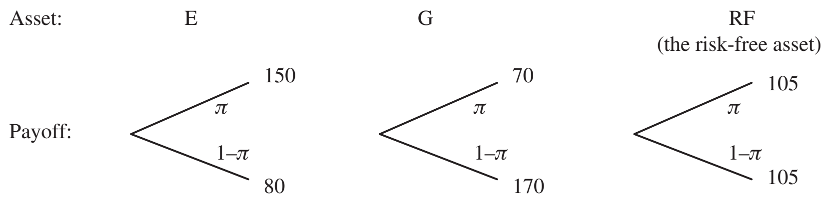


Figure 3.6
Three candidate assets.

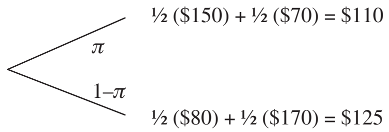


Figure 3.7
Payoff outcomes, portfolio {½E, ½G}.

Assume $\pi = ½$, and let the prices of the assets be $q_E = q_G = q_{RF} = \$100$. Under narrow framing, an investor may individually compare E and G to RF, find each individually less desirable (both E and G have large payoff variances and a substantial probability of significant loss), and end up with a portfolio composed exclusively of asset RF. Nevertheless, we see from Figure 3.7 that RF is clearly state by state *dominated* by the portfolio {½E, ½G}.

It is unclear what utility specification would eliminate all framing phenomena.

3.7.2 Prospect Theory

At present, Kahneman and Tversky's (1979, 1992) Prospect Theory is the most highly developed behavioral theory of choice.[14] It rests on a number of experimental observations:

i Consider the random payoff ($110, − $100, ½). In experimental settings, a majority of participants declined to accept this lottery, irrespective of their level of personal wealth, even if it was offered to them at zero cost.

ii Consider the four basic lotteries from Section 3.7.1

A: ($1000, 0, .5)
B: ($500, 0, 1)
C: (−$1000, 0, .5)
D: (−$500, 0, 1).

[14] Barberis (2013) provides a detailed overview of the theory and applications that have followed from it. This section owes much to him.

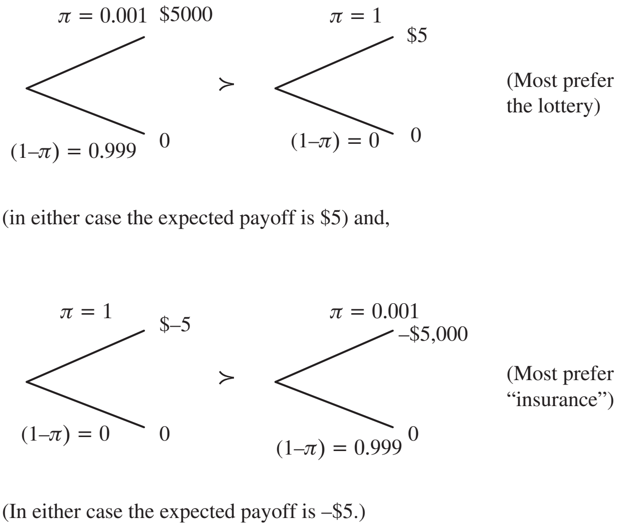


Figure 3.8
Preferences for lotteries and insurance.

When offered for comparison without any prior wealth distinctions (no issues of framing), a majority of participants displayed the following preference:

$B \succ A$ and $C \succ D$.

iii Participants generally displayed a preference for both lotteries and insurance when offered in closely related choice settings; in particular, see Figure 3.8.

Building on these and other observations Kahneman and Tversky (1979) propose a theory of choice under uncertainty (Prospect Theory) with four principal ingredients.

1. Investors ultimately derive utility not from their absolute wealth levels (as in the VNM-expected utility case) but from gains or losses relative to some reference or benchmark value. This (critical) element in their theory is suggested by observation (i): since at very high wealth levels, the acceptance or rejection of ($110, $−100, ½) is really of little consumption consequence, its rejection at all wealth levels suggests that it is the gains or losses themselves, possibly relative to some preconceived benchmark, that really matter to investors. They thus propose a utility-of-money function of the form $U(Y-\overline{Y})$ where $\overline{Y}$ is the benchmark. The benchmark can be thought of as either a minimally acceptable wealth level or, under the proper transformations, a cutoff rate of return. It can be changing through time reflecting prior experience. Unfortunately, Prospect Theory does not offer a general guide as to how the benchmark should be selected in any specific choice setting.
2. Since ($110, $−100, ½) has a positive expected value, its rejection at all wealth levels also suggests that agents feel losses more acutely than gains (of greater comparative

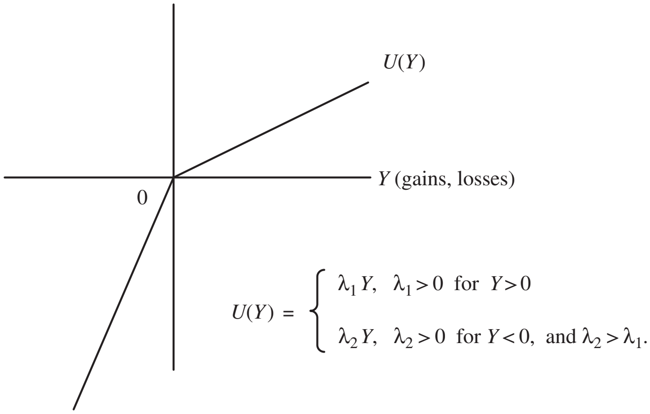


Figure 3.9
Loss-averse utility function.

magnitude). This is the sense of loss-averse preferences or simply "loss aversion." The simplest illustration of a loss-averse utility-of-money function is seen in Figure 3.9, where $\overline{Y} = 0$, a zero benchmark.

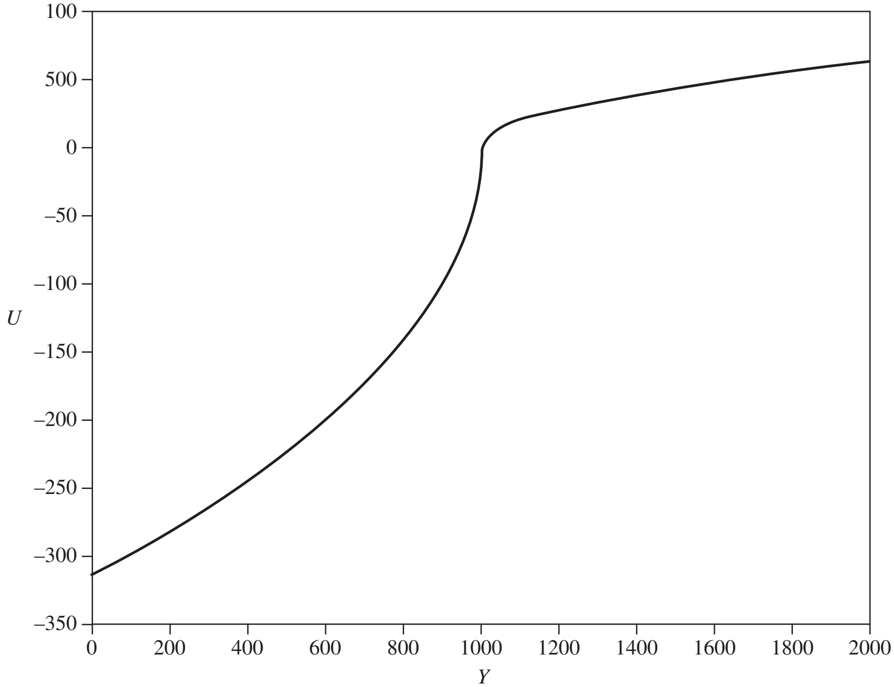
3. As a further refinement, consider the choices in observation (ii). The fact that most participants prefer lottery B to lottery A suggests dislike of risk over positive gain lotteries. The preference for lottery C over lottery D, however, suggests "risk loving" behavior over losses. As we will see in the next chapter these features imply that $U(Y-\overline{Y})$ is concave for $Y > \overline{Y}$ but convex for $Y < \overline{Y}$.

   An illustration of a utility representation satisfying (1)−(3) is as follows: Let $\overline{Y}$ denote the benchmark payoff and define the investor's utility-of-money function $U(Y)$ by

$$U(Y) = \begin{cases} \dfrac{(|Y-\overline{Y}|)^{1-\gamma_1}}{1-\gamma_1}, & \text{if } \; Y \geq \overline{Y} \\ \dfrac{-\lambda(|Y-\overline{Y}|)^{1-\gamma_2}}{1-\gamma_2}, & \text{if } \; Y \leq \overline{Y} \end{cases}$$

   where $\lambda > 1$ captures the extent of the investor's aversion to "losses" relative to the benchmark, and $\gamma_1 > 0$ and $\gamma_2 > 0$ need not coincide (but $\gamma_1 \neq 1$, $\gamma_2 \neq 1$). In other words, the curvature of the function may differ for deviations above or below the benchmark. See Figure 3.10 for an illustration. Clearly, all three features can have a large impact on the relative ranking of uncertain lottery payoffs.

4. There is a fourth attribute of Prospect Theory that also has its origins in observation. (iii) One interpretation of the choices found in that observation is that investors overweight low probability tail events, both favorable and unfavorable. Following on



Parameter values: $\bar{Y} = 1000; \gamma_1 = \gamma_2 = 0.5, \lambda = 5$

Figure 3.10
Utility function for Prospect Theory.

this possibility, Kahneman and Tversky (1979) elect to weigh the utilities of the various outcomes using a nonlinear function of the true probabilities which is asymmetric in a manner that gives a high weighting to tail events. These weightings are not necessarily to be interpreted as erroneous probability estimates, but as perhaps reflecting relative welfare consequences of the outcomes for the investor that are not observable. See Tversky and Kahneman (1992) for a full discussion. A sample weighting function taken from Tversky and Kahneman (1992) is found in Figure 3.11. Note that this weighting function resembles a probability distribution function where there is a large likelihood of "extreme" events.

As a paradigm for rationalizing laboratory observations (i)−(iv), Prospect Theory has no equal at the moment. But does it have any direct advantage over VNM-expected utility in explaining equilibrium market phenomena? There are at least two relevant works in this regard, Barberis and Huang (2008) and Benartzi and Thaler (1995). In the first paper, the authors show that the skewness in the return distribution of a common stock can have important pricing implications when investors' loss-averse preferences are defined over

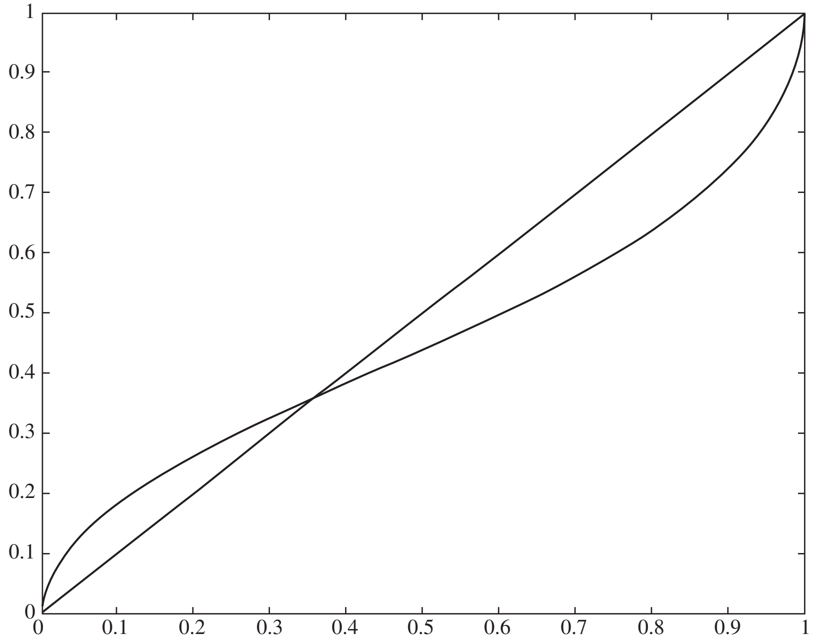


Figure 3.11
The probability weighting function. *Source: This figure is taken from Tversky and Kahneman (1992) as presented in Barberis (2013).*

changes in the value of their portfolios. In particular, securities with positively skewed return distributions are priced higher (and have lower average returns) and negatively skewed securities lower (and have higher average returns) than would be the case in a VNM-expected utility environment, a consequence that follows from the asymmetric weighting function fundamental to Prospect Theory. It is a feature that appears to be present in the cross section of security returns (see, for example, Boyer et al. (2010) and Conrad et al. (2013).

The basic intuition in the Benartzi and Thaler (1995) paper is that loss-averse investors will dislike equity securities because stock market returns are much more highly dispersed relative to bond returns. This fact should lead, they argue, to a higher equity premium in an equilibrium setting where investors have loss-averse preferences than would be possible in the same environment where investors are VNM-expected utility. In a formal dynamic equilibrium quantitative model where investors have loss-averse preferences, Barberis and Huang (2001) go on to confirm the assertions of Benartzi and Thaler (1995), subject to qualifications.

There are many other utility forms that have been proposed over the past 30 years which are linked in some way to Prospect Theory. Some are widely employed in the research literature. We conclude this section by enumerating a few of them below.

3.7.2.1 Preference Orderings with Connections to Prospect Theory

i. Survival benchmark: Investors evaluate lotteries according to $\mathbb{U}(\tilde{Y}) = EU(\tilde{Y} - \overline{Y})$ where $\overline{Y}$ represents a minimum level of income for a decent lifestyle.

ii. Habit formation preferences: Investors become accustomed to a particular income level and measure their utility by the extent to which their present income realization departs from it. For example,

$$\mathbb{U}(\tilde{Y}_t) = EU(\tilde{Y}_t - Y_{t-1})$$

where the habit is identified by the prior period's income (and associated consumption) level, thereby introducing a time dimension. See Constantinides (1990) and Sundaresan (1989).

iii. "Keeping up with the Joneses" or relative income status: Once again investors evaluate lotteries according to

$$\mathbb{U}(\tilde{Y}) = EU(\tilde{Y} - \overline{Y})$$

except that $\overline{Y}$ represents the average income level in the investor's reference community. See Abel (1990).[15]

iv. "Disappointment aversion"; Gul (1991): Here we will need to be a bit more detailed in our representation of the expectations operator E. A disappointment averse investor evaluates lotteries according to

$$\mathbb{U}(\tilde{Y}) = \int_{\underline{Y}}^{B} u(Y)\mathrm{d}F(Y) + AY\int_{B}^{\overline{Y}} u(Y)\mathrm{d}F(Y)$$

where $\underline{Y}$, $\overline{Y}$ denote, respectively, the minimum and maximum payoffs associated with $\tilde{Y}$, A is a number $0 < A < 1$, and B is a certain payment (a certainty equivalent) in exchange for which the investor would be willing to sell the lottery $\tilde{Y}$. With $A < 1$, the investor effectively weighs payments above this benchmark level less heavily than ones below it—a sort of indirect "loss" aversion. He is, essentially, more concerned with low-value outcomes, low in the sense of falling short of the amount for which the investor would have been willing to sell the lottery. Routledge and Zinn (2003) modify the original Gul (1991) representation in a way that endogenizes the construction of B so as to make low payoff realizations even more painful than in Gul (1991).

v. The notion of regret; Loomis and Sugden (1982): The idea here is that if an investor selects one lottery over another, then his utility benefit of a particular state's payment will be diminished had the payoff to the rejected asset in that same state been higher; i.e., given the realized state, the investor experiences regret for not having chosen the

[15] In all these three cases, the benchmark is calculated so that the utility function $U(\ )$ is always defined over a positive quantity.

other asset. More formally, in deciding which of lotteries $\tilde{Y}_1$ and $\tilde{Y}_2$ to choose, the investor computes (for $\tilde{Y}_1$; $\tilde{Y}_2$ is evaluated symmetrically):

$$\begin{aligned}\mathbb{U}(\tilde{Y}_1) &= EU(\tilde{Y}_1) + ER(\tilde{Y}_1 - \tilde{Y}_2)\\ &= \sum_{i=1}^{N}\pi(\theta_1)u(Y_1(\theta_i)) + \delta\sum_{i=1}^{N}\pi(\theta_i)R(Y_1(\theta_i) - Y_2(\theta_i))\end{aligned}$$

where i indexes the states, $\delta > 0$ and, like $U(\ )$, $R(\ )$, the regret function, is a monotone, strictly increasing function which satisfies $R(0) = 0$ and $-R(-\xi) = R(\xi)$ for any $\xi > 0$. $R(\ )$[16] diminishes expected utility in those states where $\tilde{Y}_2$ has the higher outcome. In this sense the alternative asset's payoff serves as the benchmark on a state-by-state basis. Loomes and Sugden (1982) demonstrate that regret preferences can rationalize many of the choice anomalies presented in Kahneman and Tversky (1979). As one might expect, however, regret preferences do not satisfy the VNM transitivity axiom which has the implication that, *per se*, they cannot necessarily be used to isolate the best (highest expected utility) of a collection of eligible lotteries.

Preference representations (i)−(v) are a representative sample of "what's out there". Except for disappointment aversion, all lack a full axiomatic basis grounded in choices-over-lotteries. It is also not clear what the corresponding "reverse engineered" preferences over consumption goods would look like, a problem that is typically sidestepped by assuming one composite consumption good with a normalized price of one so that income and consumption are always numerically equal. In all cases, however, the takeaway is the same: investors evaluate gains and losses relative to a benchmark and do so asymmetrically in a way that tends to "overweight losses" relative to expected utility.

3.7.3 Overconfidence

A variety of studies find evidence that suggests pervasive overconfidence among physicians, nurses, attorneys, engineers, and others.[17] Monitor (2007) finds, in a survey, that more than 70% of professional portfolio managers regard the service they provide as "above average." These various studies measure overconfidence in different ways, consistent with the difficulties inherent in making the notion precise in individual contexts. A more formal study again suggestive of overconfidence is Barber and Odean (2000); see also Odean (1998, 1999). Using data from 78,000 individually managed accounts at a large discount brokerage company, these authors find that these accounts substantially underperform various commonplace benchmarks largely because of the large transaction

[16] Alternatively, $R(\ )$ could assume the form $R(Y_1(\theta_i) - Y_2(\theta_i)) = \min\{0,(Y_1(\theta_i) - Y_2(\theta_i))\}$, etc.

[17] Some references are Baumann et al. (1991), Wagenaar and Kern (1986), Russo and Schoemaker (1992), and De Bont and Thaler (1990) for, respectively, nurses, attorneys, high-level managers, and portfolio managers.

costs attendant to frequent trading. They interpret these findings as suggesting that individual investors are overconfident in their ability to pick "winning stocks."

It is not clear how the notion of overconfidence can be modeled in a framework where agent preferences are represented by utility functions of some type. One possible approach is to borrow from Prospect Theory as regards subjective weightings on the various possible outcomes. In particular, we might expect that an overconfident investor is one who assigns lower weightings to negative outcomes than does the typical investor or who assumes the information he has collected regarding a stock's future return distribution is more precise than it actually is. This latter approach is taken by Daniel et al. (1998), who study the consequences of overconfidence in an equilibrium security pricing model. These authors find that overconfidence, modeled in this way, can lead to "momentum"-like effects: price increases today followed, on average, by further price increases tomorrow, and vice versa.

3.8 Conclusions

The expected utility theory is the workhorse of choice theory under uncertainty. It will be put to use systematically in this book, as it is in most of financial theory. We have argued in this chapter that the expected utility construct provides a straightforward, intuitive mechanism for comparing uncertain asset payoff structures. As such, it offers a well-defined procedure for ranking the assets themselves.

Two ingredients are necessary for this process:

1. An estimate of the probability distribution governing the asset's uncertain payments. While it is not trivial to estimate this quantity, it must also be estimated for the much simpler and less flexible mean/variance criterion.
2. An estimate of the agent's utility-of-money function; it is the latter that fully characterizes his preference ordering. How this can be identified is one of the topics of the next chapter.

Behavioral theories represent plausible, experimentally based deviations from the axioms underlying expected utility. They make us aware of departures from "rationality" that have potential implications for explaining phenomena that are difficult to account for in the standard VNM framework.

References

Abel, A., 1990. Asset pricing under habit formation and catching up with the Joneses. Am. Econ. Rev. Pap. Proc. 80, 38–42.

Allais, M., 1964. Le comportement de l'homme rationnel devant le risque: Critique des postulats de l'école Américaine. Econometrica. 21, 503–546.

Arrow, K.J., 1963. Social Choice and Individual Values. Yale University Press, New Haven, CT.

Barberis, N., 2013. Thirty years of prospect theory in economics: a review and assessment. J. Econ. Perspect. 27, 173–196.
Barberis, N., Huang, M., 2001. Mental accounting, loss aversion, and individual stock returns. J. Finan. 56, 1247–1292.
Barberis, N., Huang, M., 2008. Stocks as lotteries: the implications of probability weighting for security prices. Am. Econ. Rev. 98, 2066–2100.
Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: the common stock performance of individual investors. J. Finan. 55, 773–806.
Baumann, A., Deber, R., Thompson, G., 1991. Overconfidence among physicians and nurses: the macro-certainty, micro-uncertainty phenomenon. Soc. Sci. Med. 32, 167–174.
Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. Q. J. Econ. 110, 73–92.
Boyer, B., Mitton, T., Vorkink, K., 2010. Expected idiosyncratic skewness. Rev. Finan. Stud. 23, 169–202.
Conrad, J., Dittmar, R., Ghysels, E., 2013. Ex ante skewness and expected stock returns. J. Finan. 68, 85–124.
Constantinides, G., 1990. Habit formation: a resolution of the equity premium puzzle. J. Polit. Econ. 98, 519–543.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. J. Finan. 53, 1839–1885.
DeBont, W., Thaler, R., 1990. Do security analysts overreact? Am. Econ. Rev. 80, 52–57.
Grether, D., Plott, C., 1979. Economic theory of choice and the preference reversal phenomenon. Am. Econ. Rev. 75, 623–638.
Gul, F., 1991. A theory of disappointment aversion. Econometrica. 59, 667–686.
Huberman, G., 2001. Familiarity breeds investment. Rev. Financ. Stud. 14, 659–680.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica. 47, 263–291.
Loomes, G., Sugden, R., 1982. Regret theory: an alternative theory of rational choice under uncertainty. Econ. J. 92, 805–824.
Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford; New York: Oxford University Press.
Monitor, J., 2007. Behavioral Investing: A Practitioner's Guide to Applying Behavioral Finance. John Wiley & Sons, West Sussex.
Odean, T., 1998. Are investors reluctant to realize their losses? J. Finan. 53, 1775–1798.
Odean, T., 1999. Do investors trade too much? Am. Econ. Rev. 89, 1279–1298.
Reiff, P., 1966. The Triumph of the Therapeutic, New York: Harper and Row, 1966; quote taken from the reprinted edition, ISI Books: Wilmington, Delaware (2006).
Routledge, B., Zin, S., 2003. Generalized appointment aversion and asset prices, NBER Working Paper no. 10107.
Russo, J., Schoemaker, P., 1992. Managing overconfidence. Sloan Manage. Rev. 33, 7–17.
Sundaresan, S., 1989. Intertemporally dependent preferences and the volatility of consumption and wealth. Rev. Finan. Stud. 2, 73–89.
Thaler, R.H., 1992. The Winner's Curse. Princeton University Press, Princeton, NJ.
Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. 5, 297–323.
Wagenaar, W., Keren, G., 1986. Does the expect know? the reliability of predictions and confidence ratios of experts. In: Hollnagel, E., Mancini, G., Woods, D.D. (Eds.), Intelligent Decision Support in Process Environments. Springer, Berlin.