Financial Economics - Asset Pricing and Portfolio Selection
Lecture 1

Hongtao Zhou *

1 Utility Functions and Risk Aversion Coefficients

The first part of this book addresses the decision problem of an investor in a one-period framework. We suppose the investor makes certain decisions at the beginning of the period (how much of his wealth to spend and how to invest what he does not spend) and the generally random results of the investments then determine his wealth at the end of the period. The investor has preferences for spending at the beginning of the period and wealth at the end of the period, and these preferences, in conjunction with the available investment opportunities, determine his choices. This is a simplification in the sense that we are ignoring the allocation decision for beginning-of-period spending and end-of-period wealth across the different consumption goods available. Likewise, we are ignoring the prices of different consumption goods at the beginning and end of the period. This is the common practice in finance, and it will be followed throughout this book. To put this formally, we can say we are assuming there is only a single consumption good and we are using it as the numeraire (so its price is always one). The returns on investments will therefore be "real returns," i.e., normalized in terms of buying power for the single consumption good.

We are also simplifying here in assuming the investor consumes all of his end-of-period wealth. We will relax this assumption later and study multi-period models. This will typically introduce "state-dependence" in the investor's preferences for end-of-period wealth. Specifically, the investor will care about the investment opportunities available at the end of the period as well as his wealth, because it is the combination of wealth and investment opportunities that determine the possibilities for future wealth (and consumption). State dependence is possible even in a single-period model; for example, the investor's health may be regarded as random, in which case he faces joint gambles over wealth and health, and his preferences over wealth gambles will likely depend on how they correlate with his health. We are assuming this is not the case: the investor's preferences for wealth gambles depend only on the (marginal) probability distribution of wealth.

---

*These notes are based on *Asset Pricing and Portfolio Choice Theory* by Kerry Back. We thank Professor Back for his kindly and enormous support.

We will assume in most of the book that each investor satisfies certain axioms of rationality, which imply that his choices will be those that maximize the expected value of a utility function. Specifically, letting c_0 denote beginning-of-period consumption and $\tilde{c}_1$ denote end-of-period consumption (which equals end-of-period wealth), we assume there is a function v such that the investor maximizes the expected value of $v(c_0, \tilde{c}_1)$. In many parts of the book - in particular, in this chapter and the next - the probabilities with respect to which expected values are computed can be "subjective probabilities;" that is, we do not need to assume the investor knows the "true probabilities" of the outcomes implied by his choices. In the discussion of market equilibrium, it will be assumed that all investors agree on the probabilities and hence that the probabilities are "objective."

In this chapter and in many other places in the book, we will simplify our model even further and focus on end-of-period wealth. We can do this by assuming c_0 is optimally chosen and considering the derived utility function $w \mapsto v(c^*, w)$ where c^* denotes the optimal beginning-of-period consumption. Denoting this function of w by $u(w)$, we assume the investor chooses his investments to maximize $E[u(\tilde{w})]$.

1.1 Uniqueness of Utility Functions

A utility function u is said to represent preferences over wealth gambles if for any random $\tilde{w}_1$ and $\tilde{w}_2$ such that $\tilde{w}_1$ is at least as preferred as $\tilde{w}_2$ we have $E[u(\tilde{w}_1)] \geq E[u(\tilde{w}_2)]$. When making decisions under certainty, the utility function representing preferences is unique only up to monotone transforms. However, for decisions under uncertainty, utility functions are unique up to monotone affine transforms: if two utility functions u and f represent the same preferences over all wealth gambles, then u must be a monotone affine transform of f; i.e., there exists a constant a and a constant $b > 0$ such that $u(w) = a + bf(w)$ for every w.

1.2 Concavity and Risk Aversion

An investor is said to be (weakly) risk averse if, for any random $\tilde{w}$ with mean $\bar{w}$, we have

$$u(\bar{w}) \geq E[u(\tilde{w})] \tag{1.1}$$

An equivalent definition is that a risk-averse investor would prefer to avoid a fair bet, meaning that if $\tilde{\epsilon}$ is a zero-mean random variable and a is a constant, then

$$u(a) \geq E[u(a + \tilde{\epsilon})] \tag{1.2}$$

Risk aversion is equivalent to concavity of the utility function. The fact that concavity of u implies (1.1) is known as Jensen's inequality. Note that concavity is preserved by monotone affine transforms (though not by general monotone transforms). Strict concavity is equivalent to strict risk aversion, meaning strict preference for a sure thing over a gamble with the same mean (strict inequality in (1.1) and (1.2), unless $\tilde{w} = \bar{w}$ or $\tilde{e} = 0$ with probability one).

For a differentiable function u, concavity is equivalent to nonincreasing marginal utility ($u'(w_1) \leq u'(w_0)$ if $w_1 > w_0$) and strict concavity is equivalent to decreasing marginal utility ($u'(w_1) < u'(w_0)$ if $w_1 > w_0$). For a twice differentiable function u, concavity is equivalent to $u''(w) \leq 0$ for all w and strict concavity is implied by $u''(w) < 0$ for all w.

1.3 Coefficients of Risk Aversion

The coefficient of absolute risk aversion at a wealth level w is defined as

$$\alpha(w) = -\frac{u''(w)}{u'(w)},$$

where the primes denote derivatives. The second derivative of the utility function measures its concavity; dividing by the first derivative eliminates the dependence on the arbitrary scaling of the utility - i.e., the coefficient of absolute risk aversion is unaffected by a monotone affine transform of the utility function. Hence, it depends on the preferences, not on the particular utility function chosen to represent the preferences. Note that $\alpha(w) \geq 0$ for any risk averse investor, because concavity implies $u'' \leq 0$. Clearly, a high value of α indicates a high curvature of the utility function; moreover, as will be explained in the next section, this implies high aversion to risk.

The coefficient of relative risk aversion is defined as

$$\rho(w) = w\alpha(w) = -\frac{wu''(w)}{u'(w)}.$$

The coefficient of risk tolerance is defined as

$$\tau(w) = \frac{1}{\alpha(w)} = -\frac{u'(w)}{u''(w)}.$$

Because risk can be shared among investors, the aggregate risk tolerance in the economy is frequently important. If there are H investors with coefficients of absolute risk aversion α_h and coefficients of risk tolerance $\tau_h = 1/\alpha_h$, then the aggregate risk tolerance is $\sum_{h=1}^{H} \tau_h$. We can define the aggregate absolute risk aversion as the reciprocal of the aggregate risk tolerance:

$$\frac{1}{\sum_{h=1}^{H} 1/\alpha_h}.$$

This is equal to the harmonic mean of the absolute risk aversion coefficients divided by H.

In the next section, we will describe the sense in which $\alpha(w)$ and $\rho(w)$ measure risk aversion. We will also see why $\alpha(w)$ is called the coefficient of *absolute* risk aversion and why $\rho(w)$ is called the coefficient of *relative* risk aversion.

1.4 Risk Aversion Coefficients and Certainty Equivalents

Given a wealth level w and a zero-mean random variable $\tilde{\epsilon}$, define $w - \pi$ to be the certainty equivalent of $w + \tilde{\epsilon}$ if

$$u(w - \pi) \ = E[u(w + \tilde{\epsilon})]$$

In other words, starting at wealth w, π is the largest amount the investor would pay to avoid the gamble $\tilde{\epsilon}$. We can show that, for "small gambles" (and assuming u is twice continuously differentiable and the gamble is a bounded random variable)

$$\pi \approx \frac{1}{2}\sigma^2 \alpha(w) \tag{1.3}$$

where σ^2 is the variance of $\tilde{\epsilon}$. Thus, the amount one would pay to avoid the gamble is proportional to the coefficient of absolute risk aversion. Eq. (1.3) is derived, and the meaning of the approximation explained, at the end of this section.

The distinction between absolute and relative risk aversion can be seen by contrasting (1.3) with the following: let $w - w\pi$ be the certainty equivalent of $w + w\tilde{\epsilon}$, where w is a constant and $\tilde{\epsilon}$ is a zero-mean random variable with variance σ^2. Then

$$\pi \approx \frac{1}{2}\sigma^2 \rho(w) \tag{1.4}$$

Thus, the proportion π of initial wealth w that one would pay to avoid a gamble equal to the proportion $\tilde{\epsilon}$ of initial wealth depends on *relative* risk aversion and the variance of $\tilde{\epsilon}$. The result (1.4) follows immediately from (1.3): Define $w\tilde{\epsilon}$ to be the gamble we considered when discussing absolute risk aversion; then the variance of the gamble is $w^2\sigma^2$; thus, one would pay approximately

$$\frac{1}{2}w^2\sigma^2\alpha(w) = \frac{1}{2}w\sigma^2\rho(w)$$

to avoid it; this means that one would pay $w\pi$ to avoid it, where π satisfies (1.4).

To make (1.3) more concrete, consider flipping a coin for \$1. In other words, take $\tilde{\epsilon} = \pm 1$ with equal probabilities. The standard deviation of this gamble is 1, and so the variance is 1 also. Condition (1.3) says that one would pay approximately \$1 $\times \alpha(w)/2$ to avoid it. If one would pay 10 cents to avoid it, then $\alpha(w) \approx 0.2$.

To make (1.4) more concrete, let w be your wealth and consider flipping a coin where you win 10% of w if it comes up heads and lose 10% of w if it comes up tails. This is a large gamble, so the approximation in (1.4) may not be very good. Nevertheless, it can help us interpret (1.4). The standard deviation of the random variable $\tilde{\epsilon}$ defined as $\tilde{\epsilon} = \pm 0.1$ with equal probabilities is 0.1, and its variance is $0.01 = 1\%$. According to (1.4), one would pay approximately

$$\frac{1}{2}\rho(w) \times 1\%$$

of one's wealth to avoid the gamble. If one would pay exactly 2% of one's wealth to avoid this 10% gamble, then (1.4) says that $\rho(w) \approx 4$.

We will end this section with a proof of (1.3) - and hence also of (1.4).

Let u be twice continuously differentiable, and let w be in the interior of the domain of u with $u'(w) \neq 0$. Take $\tilde{\epsilon}$ to be a bounded zero-mean random variable with unit variance and define $\tilde{\epsilon}_n = \sigma_n \tilde{\epsilon}$ for a sequence of positive numbers σ_n converging to zero. For sufficiently large n, $w + \tilde{\epsilon}_n$ is in the domain of u with probability one. Moreover, the variance of $\tilde{\epsilon}_n$ is σ_n^2.

Let $w - \pi_n$ be the certainty equivalent of $w + \tilde{\epsilon}_n$. We will show that

$$\frac{\pi_n}{\sigma_n^2} \to \frac{1}{2}\alpha(w). \tag{1.5}$$

This is the meaning of the approximation (1.3).

To establish (1.5), take Taylor series expansions of $u(w - \pi_n)$ and $u(w + \tilde{\epsilon}_n)$. We have

$$u(w - \pi_n) = u(w) - u'(x_n)\pi_n$$

for some numbers x_n between w and $w - \pi_n$. Likewise,

$$u(w + \tilde{\epsilon}_n) = u(w) + u'(w)\tilde{\epsilon}_n + \frac{1}{2}u''(\tilde{y}_n)\tilde{\epsilon}_n^2$$

for some random numbers $\tilde{y}_n$ between w and $w + \tilde{\epsilon}_n$. Using the fact that $w - \pi_n$ is the certainty equivalent of $w + \tilde{\epsilon}_n$ and the fact that $\tilde{\epsilon}_n$ has zero mean, we have

$$\begin{aligned} u(w) - u'(x_n)\pi_n &= E[u(w) + u'(w)\tilde{\epsilon}_n + \frac{1}{2}u''(\tilde{y}_n)\tilde{\epsilon}_n^2] \\ &= u(w) + \frac{1}{2}E[u''(\tilde{y}_n)\tilde{\epsilon}_n^2]. \end{aligned}$$

Thus,

$$-u'(x_n)\pi_n = \frac{1}{2}E[u''(\tilde{y}_n)\tilde{\epsilon}_n^2],$$

which implies

$$\frac{\pi_n}{\sigma_n^2} = -\frac{1}{2}\frac{E[u''(\tilde{y}_n)\tilde{\epsilon}^2]}{u'(x_n)}.$$

The random variables $u''(\tilde{y}_n)\tilde{\epsilon}^2$ are bounded (because u'' is continuous and hence bounded on bounded sets and because the $\tilde{y}_n$ are bounded) and converge to $u''(w)\tilde{\epsilon}^2$. Hence,

$$E[u''(\tilde{y}_n)\tilde{\epsilon}^2] \to E[u''(w)\tilde{\epsilon}^2] = u''(w)E[\tilde{\epsilon}^2] = u''(w),$$

using the fact that $\tilde{\epsilon}$ has zero mean and unit variance for the last equality. Moreover,

$u'(x_n) \to u'(w)$. Therefore,

$$\frac{\pi_n}{\sigma_n^2} = -\frac{1}{2}\frac{E[u''(\tilde{y}_n)\tilde{\epsilon}^2]}{u'(x_n)} \to -\frac{1}{2}\frac{u''(w)}{u'(w)} = \frac{1}{2}\alpha(w).$$

1.5 Constant Absolute Risk Aversion

If absolute risk aversion is the same at every wealth level, then one says that the investor has CARA (Constant Absolute Risk Aversion) utility. It is left as an exercise (Prob. 1.5) to demonstrate that every CARA utility function is a monotone affine transform of the the utility function

$$u(w) = -e^{-\alpha w},$$

where α is a constant and equal to the absolute risk aversion. This is called "negative exponential utility" (or sometimes just "exponential utility").

CARA utility is characterized by the "absence of wealth effects." This "absence" applies to the risk premium discussed in the previous section and also to portfolio choice. For the risk premium, note that

$$u(w - \pi) = -e^{-\alpha w}e^{\alpha \pi},$$

and

$$u(w + \tilde{\epsilon}) = -e^{-\alpha w} - e^{\alpha \tilde{\epsilon}},$$

so the condition

$$u(w - \pi) = E[u(w + \tilde{\epsilon})]$$

is equivalent to

$$e^{\alpha \pi} = E[e^{-\alpha \tilde{\epsilon}}],$$

implying

$$\pi = \frac{1}{\alpha}\log E[e^{-\alpha \tilde{\epsilon}}] \tag{1.6}$$

which is independent of w. Thus, an individual with CARA utility will pay the same to avoid a fair gamble no matter what his initial wealth might be. This seems somewhat unreasonable, as will be discussed further below.

If the gamble $\tilde{\epsilon}$ is normally distributed, then the risk premium (1.6) can be calculated more explicitly. We use the fact, which has many applications in finance, that if $\tilde{x}$ is normally distributed with mean μ and variance σ^2, then

$$E[e^{\tilde{x}}] = e^{\mu + \frac{1}{2}\sigma^2} \tag{1.7}$$

In the case at hand, $\tilde{x} = -\alpha\tilde{\epsilon}$, which has mean zero and variance $\alpha^2\sigma^2$. Thus

$$E[e^{-\alpha \tilde{\epsilon}}] = e^{\frac{1}{2}\alpha^2\sigma^2},$$

and (1.6) implies

$$\pi = \frac{1}{2}\alpha\sigma^2. \tag{1.8}$$

This shows that the approximate formula (1.3) is exact when absolute risk aversion is constant and the gamble is normally distributed.

Consider flipping a fair coin for \$1,000. The formula (1.6) says that the amount an individual with CARA utility would pay to avoid the gamble is the same whether he starts with wealth of \$1,000 or wealth of \$1,000,000,000. One might think that in the latter case the gamble would seem much more trivial, and, since it is a fair gamble, the individual would pay very little to avoid it. On the other hand, one might pay a significant amount to avoid gambling all of one's wealth. If so - i.e., if one would pay less with an initial wealth of \$1,000,000,000 than with an initial wealth of \$1,000 to avoid a given gamble - then one has decreasing absolute risk aversion, meaning that absolute risk aversion is smaller when initial wealth is higher.

1.6 Constant Relative Risk Aversion

One says that an individual has CRRA (Constant Relative Risk Aversion) utility if the relative risk aversion is the same at all wealth levels. Note that any constant relative risk aversion utility has decreasing absolute risk aversion, because $\alpha(w) = \rho(w)/w$.

Any concave CRRA utility function is a monotone affine transform of one of the following functions (see Prob. 1.5): (i) $u(w) = \log w$, where log is the natural logarithm, (ii) $u(w)$ equals a positive power, less than one, of w, or (iii) $u(w)$ equals minus a negative power of w. The last two cases (power utility) can be consolidated by writing

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

where $\gamma < 1$. A slightly more convenient formulation, which we will adopt, is to write

$$u(w) = \frac{1}{1-\rho}w^{1-\rho} \tag{1.9}$$

where $\rho = 1 - \gamma$ is a positive constant different from 1. One can easily check that ρ is the coefficient of relative risk aversion of the utility function (1.9). Logarithmic utility has constant relative risk aversion equal to 1, and an investor with power utility (1.9) is said to be more risk averse than a log utility investor if $\rho > 1$ and to be less risk averse than a log utility investor if $\rho < 1$.

The fraction of wealth an individual with CRRA utility would pay to avoid a gamble that is proportional to initial wealth is independent of the individual's wealth. To see this, let $\tilde{\epsilon}$ be a zero-mean gamble. An individual will pay πw to avoid the gamble $\tilde{\epsilon}w$ if

$$u((1-\pi)w) = E[u((1+\tilde{\epsilon})w)]. \tag{1.10}$$

One can confirm (see Prob. 1.2) that π is independent of w for CRRA utility by using the facts that $\log(xy) = \log x + \log y$ and $(xy)^\gamma = x^\gamma y^\gamma$.

Logarithmic utility is a limiting case of power utility obtained by taking $\rho \to 1$, in the sense that a monotone affine transform of power utility converges to the natural logarithm function as $\rho \to 1$. Specifically,

$$\frac{1}{1-\rho}w^{1-\rho} - \frac{1}{1-\rho} \to \log w$$

as $\rho \to 1$ for each $w > 0$ (by l'Hopital's rule).

1.7 Linear Risk Tolerance

Many finance papers use one or more of the following special utility functions, the first three of which we have already introduced. All of these are concave functions. The risk tolerance formulas below are all straightforward calculations.

Negative Exponential For every real number w,

$$u(w) = -e^{-\alpha w},$$

for a constant $\alpha > 0$, where e is the natural exponential. The risk tolerance is

$$\tau(w) = \frac{1}{\alpha}.$$

Logarithmic For every $w > 0$,

$$u(w) = \log w,$$

where log is the natural logarithm function. The risk tolerance is

$$\tau(w) = w.$$

Power For a constant ρ with $\rho > 0$ and $\rho \neq 1$ and for every $w > 0$ if $\rho > 1$ and every $w \geq 0$ if $\rho < 1$,

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

The risk tolerance is

$$\tau(w) = \frac{w}{\rho}.$$

Shifted Logarithmic For some constant ζ and every $w > \zeta$,

$$u(w) = \log(w - \zeta),$$

where log is the natural logarithm function. The risk tolerance is

$$\tau(w) = w - \zeta.$$

Shifted Power For a constant ζ and a constant ρ with $\rho \neq 0$ and $\rho \neq 1$ and for every $w > \zeta$ if $\rho > 1$ and every $w \geq \zeta$ if $0 < \rho < 1$ and every real w if $\rho < 0$,

$$u(w) = \frac{\rho}{1-\rho}(w-\zeta)^{1-\rho},$$

The risk tolerance is

$$\tau(w) = \frac{w-\zeta}{\rho}.$$

Obviously, the shifted log utility function includes logarithmic utility as a special case ($\zeta = 0$). Also, the shifted power utility function includes power utility as a special case (note that the numerator in the coefficient $\rho/(1-\rho)$ is irrelevant when $\rho > 0$, as was assumed for power utility). For the shifted utility functions, one can interpret the constant ζ as a "subsistence level of consumption" and interpret the utility as the utility of consumption in excess of the subsistence level. This interpretation probably makes more sense when $\zeta > 0$, but we do not require $\zeta > 0$ to use the utility functions.

Each of these special utility functions has linear risk tolerance, meaning that

$$\tau(w) = A + Bw \tag{1.11}$$

for some constants A and B. We also say that these utility functions have hyperbolic absolute risk aversion, meaning that

$$\alpha(w) = \frac{1}{A+Bw}.$$

The graph of the function $w \mapsto 1/(A+Bw)$ is a hyperbola, whence the name. One frequently sees the statement that these utility functions belong to the LRT (Linear Risk Tolerance) or HARA (Hyperbolic Absolute Risk Aversion) class of utilities. It can be shown in fact (see Prob. 1.5) that any utility function with linear risk tolerance (or hyperbolic absolute risk aversion) is a monotone affine transform of one of these functions.

The coefficient in the shifted power case has been defined as $\rho/(1-\rho)$ instead of $1/(1-\rho)$ to accommodate (ensure concavity in) the case $\rho < 0$. There are three different cases for the shifted power utility function, the first two of which parallel the cases for power utility.
(i) $\rho > 1$. The utility is defined for $w > \zeta$ (it converges to $-\infty$ as $w \downarrow \zeta$) and it is negative and converges up to zero as $w \to \infty$.
(ii) $0 < \rho < 1$. The utility is defined for $w \geq \zeta$. It is zero at ζ, positive for $w > \zeta$, and converges up to infinity as $w \to \infty$.
(iii) $\rho < 0$. The utility is defined for all w. It is zero at $w = \zeta$ and negative for $w \neq \zeta$. It is

monotone increasing up to zero for $w < \zeta$ and monotone decreasing down to $-\infty$ for $w > \zeta$.

A special case of category (iii) is $\rho = -1$, in which case the utility is

$$-\frac{1}{2}(w-\zeta)^2 = -\frac{1}{2}\zeta^2 + \zeta w - \frac{1}{2}w^2.$$

This is the case of quadratic utility, which has a special importance in finance theory, because it implies mean-variance preferences. Specifically, the investor's expected utility is, ignoring the additive constant $-\zeta^2/2$,

$$\zeta E[\tilde{w}] - \frac{1}{2}E[\tilde{w}^2] = \zeta E[\tilde{w}] - \frac{1}{2}E[\tilde{w}]^2 - \frac{1}{2}\text{var}(\tilde{w}),$$

where $\text{var}(\tilde{w})$ denotes the variance of $\tilde{w}$. Thus, preferences over gambles depend only on their means and variances when an investor has quadratic utility. As already noted, quadratic utility is decreasing in wealth for $w > \zeta$, which is an unreasonable feature if wealth greater than ζ is feasible.

Any utility function with linear risk tolerance $\tau(w) = A + Bw$ with $B > 0$ has decreasing absolute risk aversion. On the other hand, quadratic utility, in addition to being a decreasing function of wealth for $w > \zeta$, also has the unattractive property of increasing absolute risk aversion, even for $w < \zeta$. This property of *increasing* absolute risk aversion (decreasing risk tolerance) is shared by every shifted power utility function with $\rho < 0$.

1.8 Conditioning and Aversion to Noise

Given random variables $\tilde{x}$ and $\tilde{y}$, the conditional expectation $E[\tilde{x}|\tilde{y}]$ is defined in Appendix A. It depends on the realization of $\tilde{y}$ and hence is a random variable. Intuitively, one can think of it as the probability-weighted average value of $\tilde{x}$, given that one knows $\tilde{y}$. Observing $\tilde{y}$ will generally cause one to update the probabilities of various events, and this produces the dependence of $E[\tilde{x}|\tilde{y}]$ on $\tilde{y}$.

Some important facts about conditional expectations are:

(i) The "law of iterated expectations" states that the expectation of the conditional expectation is just the unconditional expectation; i.e.,

$$E[E[\tilde{x}|\tilde{y}]] = E[\tilde{x}].$$

(ii) If $\tilde{x}$ and $\tilde{y}$ are independent, then $E[\tilde{x}|\tilde{y}] = E[\tilde{x}]$. The interpretation is that knowing $\tilde{y}$ tells you nothing about the average value of $\tilde{x}$ when $\tilde{y}$ is independent of $\tilde{x}$.
(iii) If $\tilde{z}$ depends only on $\tilde{y}$ in the sense that $\tilde{z} = g(\tilde{y})$ for some function g, then $E[\tilde{z}\tilde{x}|\tilde{y}] = \tilde{z}E[\tilde{x}|\tilde{y}]$. The interpretation is that if $\tilde{y}$ is known, then $\tilde{z}$ is known, so it is like a constant, pulling out of the expectation.
(iv) Jensen's inequality applies to conditional expectations. Recall that Jensen's inequality

states that

$$E[u(\tilde{w})] \leq u(E[\tilde{w}])$$

for any concave function u. This generalizes to conditional expectations as

$$E[u(\tilde{w})|\tilde{y}] \leq u(E[\tilde{w}|\tilde{y}]).$$

Calling u a utility function, the left-hand side is the conditional expected utility and the right-hand side is the utility of the conditional expectation.

One concept of a risk $\tilde{\epsilon}$ being "noise" is that of mean independence. A random variable $\tilde{\epsilon}$ is said to be "mean independent" of another random variable $\tilde{y}$ if $E[\tilde{\epsilon}|\tilde{y}] = 0$. By the law of iterated expectations, $\tilde{\epsilon}$ must have a zero mean. Mean independence requires that $\tilde{\epsilon}$ has a zero mean even when we know the realization of $\tilde{y}$ and regardless of what realization of $\tilde{y}$ occurs. Mean independence is an intermediate concept between independence and zero correlation: If $\tilde{\epsilon}$ and $\tilde{y}$ are independent and $E[\tilde{\epsilon}] = 0$, then $\tilde{\epsilon}$ is mean independent of $\tilde{y}$, and if $\tilde{\epsilon}$ is mean independent of $\tilde{y}$, then $\text{cov}(\tilde{\epsilon}, \tilde{y}) = 0$. The latter fact is the subject of Prob. 1.4, and the former follows immediately from fact (ii) above.

Now we will show that risk-averse individuals dislike this type of noise. Suppose $\tilde{w} = \tilde{y} + \tilde{\epsilon}$ where $E[\tilde{\epsilon}|\tilde{y}] = 0$. Thus, $\tilde{w}$ equals $\tilde{y}$ plus noise. This implies $E[\tilde{w}|\tilde{y}] = \tilde{y}$, so $\tilde{y}$ is the conditional mean of $\tilde{w}$. Assuming a concave utility function u, Jensen's inequality states that

$$E[u(\tilde{w})|\tilde{y}] \leq u(\tilde{y}).$$

Taking expectations and applying the law of iterated expectations on the left-hand side, we have

$$E[u(\tilde{w})] \leq E[u(\tilde{y})]. \tag{1.12}$$

Thus, $\tilde{y}$ is preferred to $\tilde{y}$ plus noise. Other results of this type are described in the end-of-chapter notes.