

# Basic Utility Theory for Portfolio Selection

In economics and finance, the most popular approach to the problem of choice under uncertainty is the *expected utility (EU) hypothesis*.

The reason for this to be the preferred paradigm is that, as a general approach to decision making under risk, it has a sound theoretical basis.\*

Suppose an individual at time 0 has to decide about the composition of her portfolio to be held until period 1, and that there are  $N$  assets which can be purchased, with (random) returns  $R_i$ ,  $i = 1, \dots, N$ .

If the initial wealth to be invested is  $W_0$ , she will have wealth

$$W_1 = \left( 1 + \sum_{i=1}^N x_i R_i \right) W_0 = (1 + r_p) W_0$$

\*For an introduction to utility theory, see Eeckhoudt, Gollier, and Schlesinger (2005): *Economic and Financial Decisions under Risk*. Princeton University Press, Princeton.

in period 1, where  $r_p = \sum_i x_i R_i$  is the *portfolio return*, and the  $x_i$ ,  $i = 1, \dots, N$ , satisfying  $\sum_i x_i = 1$ , are the *portfolio weights*.

In period 1, the individual will extract utility from consuming goods that can be purchased with this wealth.

The relationship between wealth and the utility of consuming this wealth is described by a utility function,  $U(\cdot)$ . In general, each investor will have a different  $U(\cdot)$ .

The expected utility hypothesis states that the individual will choose the portfolio weights such that the expected value of utility is maximized, i.e., the portfolio problem is

$$\max_{x_1, \dots, x_N} E[U(W_1)] = E \left\{ U \left[ \left( 1 + \sum_{i=1}^N x_i R_i \right) W_0 \right] \right\}$$

subject to  $\sum_i x_i = 1$ .

$U(\cdot)$  is also called *expected utility function*, or *von Neumann–Morgenstern utility function*.

## Properties of the Expected Utility Function

- An important property of an expected utility function is that it is unique up to affine transformations. That is, if  $U(\cdot)$  describes the preferences of an investor, then so does  $U^*(\cdot) = c_1 U(\cdot) + c_2$ , where  $c_1 > 0$ .

While the above is a purely mathematical result, we can restrict the range of reasonable utility functions by economic reasoning.

- Positive marginal utility. That is,  $U'(W) > 0$  for all  $W$ .
- Risk Aversion.

**Definition 1** An agent is risk-averse if, at any wealth level  $W$ , he or she dislikes every lottery with an expected payoff of zero, i.e.,

$$E[U(W + \epsilon)] < U(W)$$

for all  $W$  and every zero-mean random variable  $\epsilon$ .

Note that any random outcome  $Z$  can be written as  $Z = E(Z) + \epsilon$ , where  $\epsilon = Z - E(Z)$  is a zero-mean random variable.

Thus, a risk-averse agent always prefers receiving the expected outcome of a lottery with certainty, rather than the lottery itself.

Using Jensen's inequality, it can readily be shown that a necessary and sufficient condition for risk aversion is that the expected utility function is concave, i.e.,  $U''(W) < 0$  for all  $W$ .

In the context of applications to portfolio choice, it is important to note that risk aversion is closely related to *portfolio diversification*.

As a simple example, consider the case of two assets, the returns of which are identically and independently distributed.

Then the problem is to solve

$$\max_x E\{U[(1 + xR_1 + (1 - x)R_2)W_0]\}.$$

The first-order condition is

$$E\{U'[(1 + xR_1 + (1 - x)R_2)W_0](R_1 - R_2)\} = 0. \quad (1)$$

By the assumptions about the joint distribution of  $R_1$  and  $R_2$ , (1) will hold exactly if  $x = 1/2$ .

The second-order condition,

$E\{U''[(1 + xR_1 + (1 - x)R_2)W_0](R_1 - R_2)^2\} < 0$ ,  
is satisfied for a risk-averse investor.

# Measuring and Comparing Risk Aversion

## The Risk Premium and the Arrow–Pratt Measure

Risk averters dislike zero–mean risks.

Thus, a natural way to measure risk aversion is to ask how much an investor is ready to pay to get rid of a zero–mean risk  $\epsilon$ .

This is called the *risk premium*,  $\pi$ , and is defined implicitly by

$$E[U(W + \epsilon)] = U(W - \pi). \quad (2)$$

In general, the risk premium is a complex function of the distribution of  $\epsilon$ , initial wealth  $W$ , and  $U(\cdot)$ .

However, let us consider a small risk.

A second-order and a first-order Taylor approximation of the left-hand and the right-hand side of (2) gives

$$\begin{aligned} \mathbb{E}[U(W + \epsilon)] &\simeq \mathbb{E}[U(W) + \epsilon U'(W) + \frac{\epsilon^2}{2} U''(W)] \\ &= U(W) + \frac{\sigma_\epsilon^2}{2} U''(W), \end{aligned}$$

and

$$U(W - \pi) \simeq U(W) - \pi U'(W),$$

respectively, where  $\sigma_\epsilon^2 := \mathbb{E}(\epsilon^2)$  is the variance of  $\epsilon$ .

Substituting back into (2), we get

$$\pi \simeq \frac{\sigma_\epsilon^2}{2} A(W),$$

where

$$A(W) := -\frac{U''(W)}{U'(W)} \tag{3}$$

is the *Arrow–Pratt* measure of absolute risk aversion, which can be viewed as a measure of the degree of concavity of the utility function.

The division by  $U'(W)$  can be interpreted in the sense that it makes  $A(W)$  independent of affine transformations of  $U(\cdot)$ , which do not alter the preference ordering.

In view of the above, it is reasonable to say that Agent 1 is *locally* more risk-averse than Agent 2 if both have the same initial wealth  $W$  and  $\pi_1 > \pi_2$ , or, equivalently,  $A_1(W) > A_2(W)$ .

We can say more, however. Namely, **Pratt's theorem**<sup>†</sup> on *global* risk aversion states that the following three conditions are equivalent:

- 1)  $\pi_1 > \pi_2$  for any zero-mean risk and any wealth level  $W$ .
- 2)  $A_1(W) > A_2(W)$  for all  $W$ .
- 3)  $U_1(W) = G(U_2(W))$  for some increasing strictly concave function  $G$ , where  $U_1$  and  $U_2$  are the utility functions of Agents 1 and 2, respectively.

<sup>†</sup>Pratt (1964): Risk Aversion in the Small and in the Large. *Econometrica*, 32: 122-136.



A related question is what happens with the degree of risk aversion as a function of  $W$ .

Since Arrow (1971)<sup>‡</sup>, it is usually argued that absolute risk aversion should be a *decreasing* function of wealth.

For example, a lottery to gain or loose 100 is potentially life-threatening for an agent with initial wealth  $W = 101$ , whereas it is negligible for an agent with wealth  $W = 1000000$ . The former person should be willing to pay more than the latter for the elimination of such a risk.

Thus, we may require that the risk premium associated with any risk is decreasing in wealth.

It can be shown that this holds if and only if the Arrow–Pratt measure of absolute risk aversion is decreasing in wealth, i.e.,

$$\frac{d\pi}{dW} < 0 \Leftrightarrow \frac{dA(W)}{dW} < 0.$$

<sup>‡</sup>*Essays in the Theory of Risk-Bearing.* Markham, Chicago.

Thus, we can equivalently require that  $A'(W) < 0$  for all  $W$ .

Note that this requirement means

$$A'(W) = -\frac{U'''(W)U'(W) - U''(W)^2}{U'(W)^2} < 0.$$

A necessary condition for this to hold is  $U'''(W) > 0$ , so that we can also sign the third derivative of the utility function.

## Mean–Variance Analysis and Expected Utility Theory

- Mean–variance portfolio theory (or  $\mu$ – $\sigma$  analysis) assumes that investor's preferences can be described by a preference function,  $V(\mu, \sigma)$ , over the mean ( $\mu$ ) and the standard deviation ( $\sigma$ ) of the portfolio return.
- Standard assumptions are

$$V_\mu := \frac{\partial V(\mu, \sigma)}{\partial \mu} > 0,$$

and

$$V_\sigma := \frac{\partial V(\mu, \sigma)}{\partial \sigma} < 0,$$

which may be interpreted as “risk aversion”, if the variance is an appropriate measure of risk.

- The existence of such a preference function would greatly simplify things, because, then, the class of potentially optimal portfolios are those with the greatest expected return for a given level of variance and, simultaneously, the smallest variance for a given expected return.
- Moreover, portfolio means and variances are easily computed once the mean vector and the covariance matrix of the asset returns are given, and “efficient sets” can be worked out straightforwardly.
- However, in general, mean–variance analysis and the expected utility approach are not necessarily equivalent.
- Thus, because the expected utility paradigm, as a general approach to decision making under risk, has a sound theoretical basis (in contrast to the mean–variance criterion), we are interested in situations where both approaches *are* equivalent.

### Example<sup>§</sup>

Consider random variable  $X$  with probability function

$$p(x) = \begin{cases} 0.8 & \text{if } x = 1 \\ 0.2 & \text{if } x = 100 \end{cases}$$

with  $E(X) = 20.8$  and  $\text{Var}(X) = 1568.16$ .

Let random variable  $Y$  have probability function

$$p(y) = \begin{cases} 0.99 & \text{if } y = 10 \\ 0.01 & \text{if } x = 1000 \end{cases}$$

with  $E(Y) = 19.9$  and  $\text{Var}(Y) = 9702.99$ .

Thus  $E(Y) < E(X)$  and  $\text{Var}(Y) > \text{Var}(X)$ .

Let  $U(W) = \log W$ . Then,

$$\begin{aligned} E[U(X)] &= 0.8 \log 1 + 0.2 \log 100 = 0.9210 \\ &< 2.3486 = 0.99 \log 10 + 0.01 \log 1000 \\ &= E[U(Y)]. \end{aligned}$$

<sup>§</sup>Taken from Hanoch and Levy (1969): The Efficiency Analysis of Choices Involving Risk. *Review of Economic Studies*, 36: 335-346.

To make  $\mu - \sigma$  analysis reconcilable with expected utility theory, we have to make assumptions about either 1) the utility function of the decision maker, or 2) the return distribution.

## 1) Restricting the Utility Function: Quadratic Utility

Assume that the investor's expected utility function is given by

$$U(W) = W - \frac{b}{2}W^2, \quad b > 0.$$

Note that this is the most general form of quadratic utility, because expected utility functions are unique only up to an affine transformation.

Marginal utility is  $U'(W) = 1 - bW > 0$  for  $W < 1/b$ .

As all concave quadratic functions are decreasing after a certain point, care must be taken to make sure that the outcome remains in the lower, relevant range of utility.

Furthermore,  $U''(W) = -b < 0$ , so  $b > 0$  guarantees risk aversion.

Expected utility is

$$\begin{aligned} E[U(W)] &= E(W) - \frac{b}{2}E(W^2) \\ &= E(W) - \frac{b}{2}[\text{Var}(W) + E^2(W)] \\ &= \mu - \frac{b}{2}(\sigma^2 + \mu^2) \\ &=: V(\mu, \sigma), \end{aligned}$$

a function of  $\mu$  and  $\sigma$ .

$V(\mu, \sigma)$  is called a  $\mu - \sigma$  preference function.

We have

$$\begin{aligned} V_\mu : &= \frac{\partial V}{\partial \mu} = 1 - b\mu > 0, \quad \text{as } \max\{W\} < 1/b \\ V_\sigma : &= \frac{\partial V}{\partial \sigma} = -b\sigma < 0. \end{aligned}$$

In this framework, variance is the appropriate measure of risk.

For the analysis of portfolio selection, it will be useful to introduce the concept of an *indifference curve* in  $(\sigma, \mu)$ -space.

**Definition 2** The indifference curve in  $(\sigma, \mu)$ -space, relative to a given utility level  $\bar{V}$ , is the locus of points  $(\sigma, \mu)$  along which expected utility is constant, i.e., equal to  $\bar{V}$ .

For quadratic utility, indifference curves can be derived from

$$\mu - \frac{b}{2}(\sigma^2 + \mu^2) \equiv \bar{V}.$$

Multiply both sides by  $-2/b$  and add  $1/b^2$ , to get

$$\sigma^2 + \mu^2 - \frac{2}{b}\mu + \frac{1}{b^2} \equiv \frac{2}{b} \left( \frac{1}{2b} - \bar{V} \right),$$

or

$$\sigma^2 + \left( \mu - \frac{1}{b} \right)^2 \equiv \text{const.}$$

Note that  $1/(2b) - \bar{V} > 0$ , as  $\max\{U\} = 1/(2b)$ .

Thus, indifference curves are semicircles centered at  $(0, 1/b)$ .



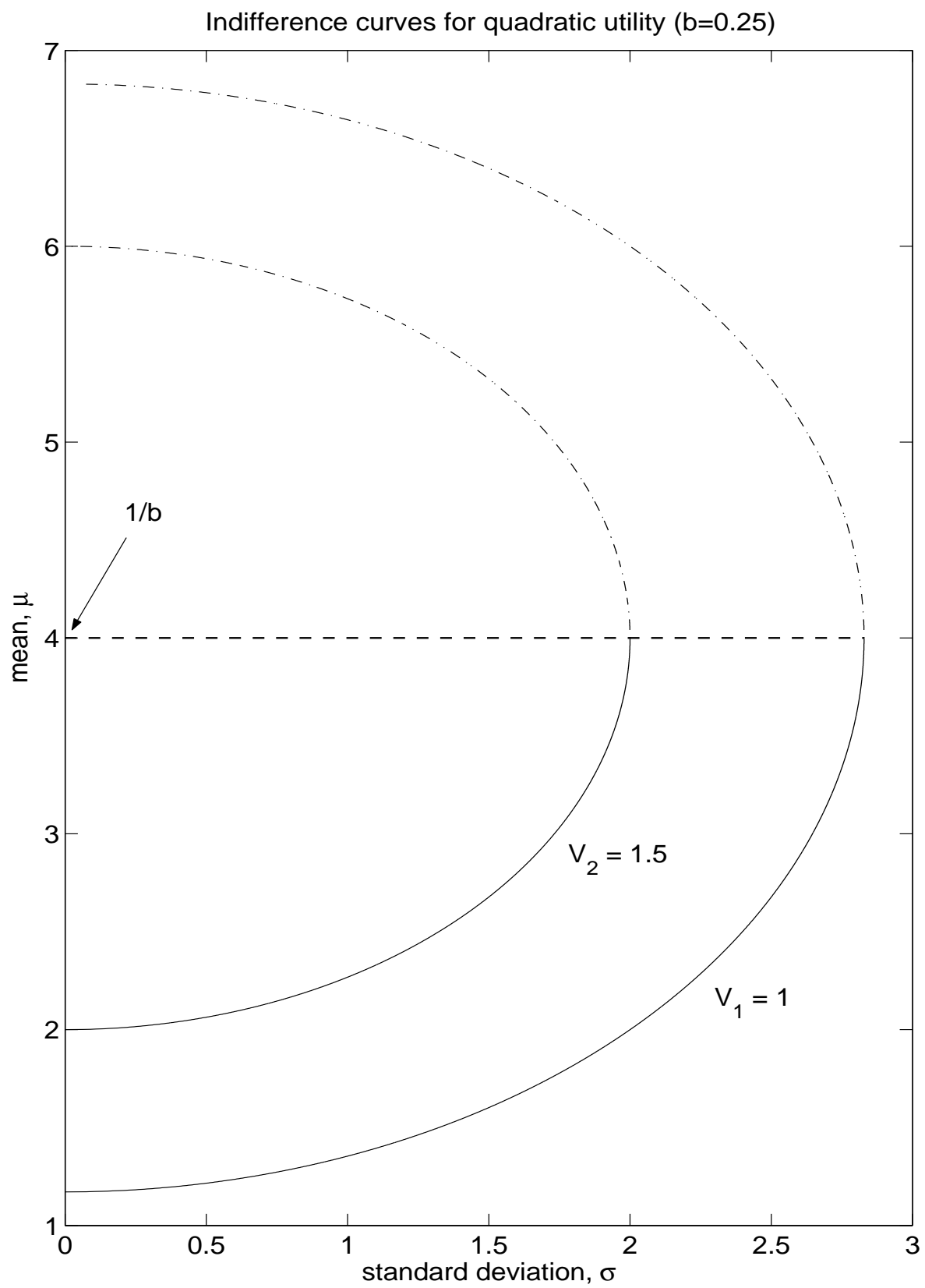
In particular, in the economically relevant range ( $\mu < 1/b$ ), they are convex in  $(\sigma, \mu)$ -space, i.e.,

$$\left. \frac{d^2 \mu}{d\sigma^2} \right|_{V(\mu, \sigma) = \bar{V}} > 0.$$

Moving in a westward or northern direction in (the economically relevant part of)  $(\sigma, \mu)$ -space, the indifference curves will correspond to higher levels of expected utility,  $\bar{V}$ , as the same mean is associated with a lower standard deviation, or the same standard deviation is associated with a higher mean.

### Example

Let  $b = 0.25$ , and let  $\bar{V}_1 = 1$ , and  $\bar{V}_2 = 1.5$ , so that the indifference curve associated with  $\bar{V}_2$  reflects a higher expected utility level than that associated with  $\bar{V}_1$ .



## Drawbacks of quadratic utility

- Marginal utility becomes negative for  $W > 1/b$ .
- Quadratic utility implies globally *increasing* absolute risk aversion, given by

$$A(W) = -\frac{U''(W)}{U'(W)} = \frac{b}{1 - bW}.$$

This is increasing in  $b$ , which is reasonable.  
But,

$$A'(W) = \left( \frac{b}{1 - bW} \right)^2 > 0.$$

In a portfolio context, Arrow (1971) has shown that this implies that wealthier people invest *less* in risky assets, which contradicts both intuition and fact.

In view of its consequences, Arrow (1971) has characterized the quadratic utility assumption as “absurd”.

## 1) Restricting the Return Distribution: Normality

If the returns have a multivariate normal distribution, then the portfolio return (and wealth) will likewise be normal.

Thus, portfolio return (and wealth) distributions will differ only by means and variances.

In case of normally distributed wealth, we can write

$$\begin{aligned} E[U(W)] &= \int_{-\infty}^{\infty} U(W) \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(W - \mu)^2}{2\sigma^2}\right\} dW \\ &= \int_{-\infty}^{\infty} U(\sigma W + \mu) \phi(W) dW \\ &= V(\mu, \sigma), \end{aligned}$$

where  $\phi(\cdot)$  is the standard normal density.

We have

$$V_{\mu} = \int_{-\infty}^{\infty} U'(\sigma W + \mu) \phi(W) dW > 0$$

by the positivity of marginal utility, and

$$\begin{aligned}
 V_\sigma &= \int_{-\infty}^{\infty} W U'(\sigma W + \mu) \phi(W) dW \\
 &= \int_{-\infty}^0 W U'(\sigma W + \mu) \phi(W) dW \\
 &\quad + \int_0^{\infty} W U'(\sigma W + \mu) \phi(W) dW \\
 &= \int_0^{\infty} W [U'(\sigma W + \mu) - U'(-\sigma W + \mu)] \phi(W) dW,
 \end{aligned}$$

where the last line follows from the symmetry of  $\phi(W)$ .

By risk aversion, i.e.,  $U''(W) < 0$  for all  $W$ , we have  $U'(\sigma W + \mu) < U'(-\sigma W + \mu)$  for  $W > 0$ , thus

$$V_\sigma < 0,$$

i.e., investors like higher expected returns and dislike return variance.

To derive indifference curves, set

$$V(\mu, \sigma) \equiv \bar{V}.$$

Differentiating implicitly,

$$\left. \frac{d\mu}{d\sigma} \right|_{V(\mu, \sigma) = \bar{V}} = -\frac{V_\sigma}{V_\mu} > 0.$$

Not surprisingly, indifference curves are upward sloping in  $(\sigma, \mu)$ -space.

To figure out their shape, we follow Tobin (1958)<sup>¶</sup> and consider pairs  $(\mu, \sigma)$  and  $(\mu^*, \sigma^*)$  being on the same indifference locus.

By concavity of the utility function,

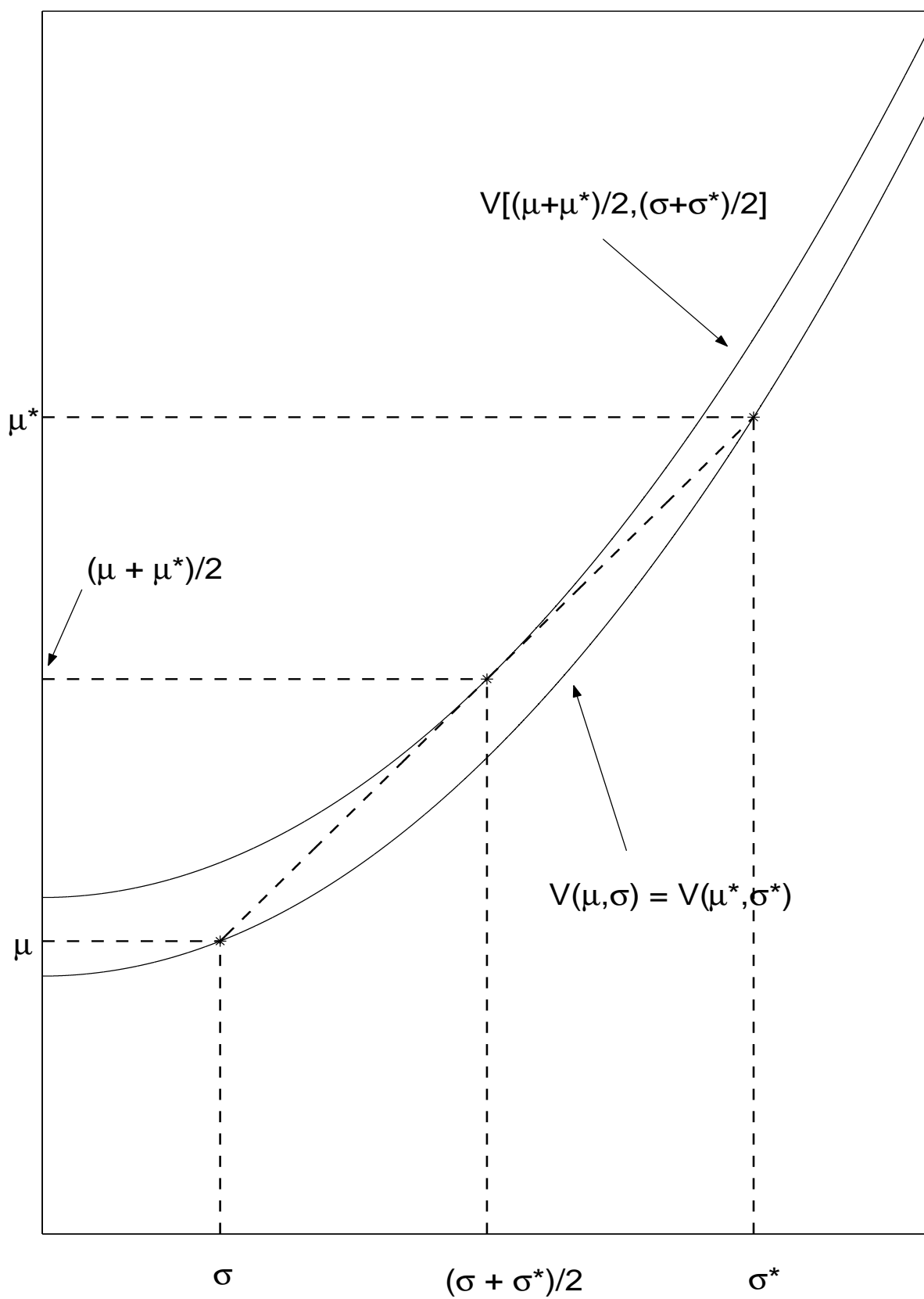
$$\begin{aligned} & \frac{1}{2}U(\sigma W + \mu) + \frac{1}{2}U(\sigma^* W + \mu^*) \\ & < U\left(\frac{\sigma + \sigma^*}{2}W + \frac{\mu + \mu^*}{2}\right). \end{aligned}$$

Taking expectations, we have

$$V\left(\frac{\mu + \mu^*}{2}, \frac{\sigma + \sigma^*}{2}\right) > V(\mu, \sigma) = V(\mu^*, \sigma^*),$$

so that the indifference curves are again convex in  $(\sigma, \mu)$ -space (see the picture).

<sup>¶</sup>Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies*, 25: 65-86.



## Example

When normality of returns is assumed, a convenient choice of  $U(\cdot)$  is the constant absolute risk aversion (CARA) utility function, given by

$$U(W) = -\exp\{-cW\}, \quad c > 0,$$

where  $c = A(W) = -U''(W)/U'(W)$  is the constant coefficient of absolute risk aversion.

Then,

$$\mathbb{E}[U(W)] = -\exp\left\{-c\mu + \frac{c^2}{2}\sigma^2\right\},$$

and the  $\mu - \sigma$  preference function may be written as

$$V(\mu, \sigma) = \mu - \frac{c}{2}\sigma^2.$$

Indifference curves relative to a preference level  $\bar{V}$  are defined by

$$\mu \equiv \bar{V} + \frac{c}{2}\sigma^2,$$

where  $\bar{V} = -\log\{-\mathbb{E}[U(W)]\}/c$ .



Note that  $\mu - \sigma$  analysis does not necessarily require normality of returns but holds within a more general class of distributions, namely, the *elliptical* distributions, see, e.g., Ingersoll (1987): *Theory of Financial Decision Making*. Roman and Littlefield Publishing, Savage.

Finally, even if  $\mu - \sigma$  analysis does not hold exactly, it will often serve as a reasonable approximation to the true, expected–utility maximizing solution, with the advantage of greatly simplifying the decision problem.

In fact, a number of studies (using real data and a range of utility functions) have shown that the exact solutions are often rather close to or even economically indistinguishable from  $\mu - \sigma$  *efficient portfolios* (to be defined below).