# Economic Decisions under Uncertainty

by Hans-Werner Sinn

North Holland: Amsterdam, New York and Oxford 1983

Chapter 4: Multiple Risk

# Chapter Four
# Multiple Risks

Up to now, the analysis has been limited to a single period of time where a single risk project has to be chosen. This chapter provides a twofold generalization. On the one hand we discuss the possibility that the risk projects referred to previously come about by summing up the incomes of various *non-rival* subprojects that can be carried out at the *same time*. On the other hand, as promised earlier, we proceed to the analysis of *repeated* choïce under risk. If, in the second case, the optimal sequence of risk projects had to be determined before the sequence starts[1], then, except for the interest problem, it would not differ from the first. However if we assume realistically that, at the beginning of each period, the decision maker again may choose between the risk projects of that period knowing the outcomes of all previous decisions, then there is a decision problem of a new kind which requires a separate analysis. Moreover, in the multiperiod case, it seems that we can no longer factor out the problem of a simultaneous optimization of the consumption decision which was briefly considered in chapter one.

In what follows we shall be concerned particularly with searching for decision rules that permit single risks to be evaluated without having to worry about all the other choices at the same time. In the classical literature, the most important example of such a rule is to base the choice of a single risk project on the mean-value criterion. This criterion is founded in the Law of Large Numbers which implies that, in the case of multiple risks, the risk aspect of the decision problem may vanish. In the case of multiple risks occurring simultaneously or sequentially, other rules will be met.

[1] This assumption is made by KRELLE (1969, pp. 95-97, 100; 1968, pp. 172-174) and SCHNEEWEISS (1967a, pp. 173-183; 1968) when analyzing repeated risk situations.

# Section A
# Simultaneous Risks

## 1. *The Law of Large Numbers and the Mean-Value Criterion*

The *Law of Large Numbers*, called also 'Bernoulli's Theorem'[1] after its originator, implies that, when a game is played a number of times, the average gain converges stochastically towards the expected gain from a single performance as the number of performances approaches infinity. Let $\Theta$ be an arbitrary random variable with finite variance and $\alpha$ an arbitrary number $>0$. Suppose a game is played $m$ times bringing about the (gross, or balance sheet) random gain $X_i$, $i=1,\dots,m$. Then Chebyshev's inequality[2] says

$$(1) \qquad W[\,|\Theta - E(\Theta)| \geq \alpha] \leq \left(\frac{\sigma(\Theta)}{\alpha}\right)^2, \; \alpha > 0.$$

Because of

$$E\left(\frac{\Sigma X_i}{m}\right) = E(X_i) \equiv E(X) \quad \forall i$$

and[3]

$$\sigma^2\left(\frac{\Sigma X_i}{m}\right) = \frac{1}{m^2}\sigma^2(\Sigma X_i) = \frac{\sigma^2(X_i)}{m} \equiv \frac{\sigma^2(X)}{m} \quad \forall i,$$

we thus have

$$(2) \qquad W\left(\left|\frac{\Sigma X_i}{m} - E(X)\right| \geq \alpha\right) \leq \frac{1}{m}\left(\frac{\sigma(X)}{\alpha}\right)^2.$$

By taking the limit, this expression gives the Law of Large Numbers:

$$(3) \qquad \lim_{m \to \infty} W\left(\left|\frac{\Sigma X_i}{m} - E(X)\right| \geq \alpha\right) = 0.$$

This is the usual formulation.

---

[1] Jacob Bernoulli, 1654–1705, uncle of Daniel Bernoulli. A presentation of the original version can be found in TODHUNTER (1865, pp. 71–73).

[2] Cf. equation (II B 3).

[3] It is assumed that the single performances of the game are stochastically independent of one another.

Another one, that for our purposes is even clearer, can be given by comparing two gambles with the stochastic prizes or gains $X$ and $X'$. Both gambles are carried out $m$ times. It is assumed that $E(X) > E(X')$. If we set $\Theta = \Sigma(X_i - X_i')$ and $\alpha = E[\Sigma(X_i - X_i')]$ ($>0$), then Chebyshev's inequality (1) becomes

$$(4) \qquad W\{|\Sigma(X_i - X_i') - E[\Sigma(X_i - X_i')]| \geq E[\Sigma(X_i - X_i')]\} \leq \left[\frac{\sigma[\Sigma(X_i - X_i')]}{E[\Sigma(X_i - X_i')]}\right]^2.$$

*A fortiori* this implies

$$(5) \qquad W\{-[\Sigma(X_i - X_i')] - E[\Sigma(X_i - X_i')] \geq E[\Sigma(X_i - X_i')]\} \leq \left[\frac{\sigma[\Sigma(X_i - X_i')]}{E[\Sigma(X_i - X_i')]}\right]^2.$$

Assuming stochastic independence between $X_i$ and $X_j$ on the one hand and $X_i'$ and $X_j'$ on the other, $i \neq j$, then, because of

$$\sigma^2[\Sigma(X_i - X_i')] = m\sigma^2(X - X') \quad \text{and}$$

$$E[\Sigma(X_i - X_i')] = mE(X - X')$$

we find the expression

$$(6) \qquad W(\Sigma X_i \leq \Sigma X_i') \leq \frac{1}{m}\left[\frac{\sigma(X - X')}{E(X - X')}\right]^2$$

and its limit

$$(7) \qquad \lim_{m \to \infty} W(\Sigma X_i \leq \Sigma X_i') = 0.$$

Since, by assumption, $E(X) > E(X')$, expression (7) says that the probability that the game with the higher expected gain will bring about a higher sum of gains approaches certainty as the number of times the games are played approaches infinity.

This formulation seems to suggest that the mathematical expectation can be taken to be the preference functional provided that the assumptions underlying the above reasoning are at least approximately satisfied in practical decision making. But this hardly seems to be the case. The condition that, as a rule, is violated most severely is that $m$ is sufficiently large. Apart from that, however, there is an important effect that, even when the games are played an infinite number of times, may thwart the Law of Large Numbers. This effect is studied in the next section.

## 2. *The Correlation of Risks*

A condition for the Law of Large Numbers, that is often not satisfied in practical decision problems, is the stochastic independence of the single games. How the Law of Large Numbers is modified when interdependencies are taken into account can be well demonstrated by reference to an insurance example which can easily by reinterpreted for other decision problems.

Consider an insurance company that has to decide which of two competitive insurance markets $K$ and $H$ it should operate in. In both markets, there are completely homogeneous risks and given insurance premiums. A single contract sold in market $K$ brings a net gain $X^K$, and a single contract sold in market $H$ brings a net gain $X^H$. Assume that $E(X^K)>E(X^H)$. Suppose the company decides to sell $m$ policies either in market $K$ or in market $H$. Can it be almost certain that operating in market $K$ will bring about a higher level of profit than operating in $H$ when $m$ is chosen sufficiently large?

In the case of stochastic independence, the Law of Large Numbers ensures that it can. This is easily seen by substituting $X_i \equiv X_i^K$ and $X_i' \equiv X_i^H$ in (7). The situation is different, though, if the risks are correlated. The step from (5) to (6) is no longer possible. Instead of (6) the more general formula

$$(8) \qquad W[\Sigma X_i^K \leq \Sigma X_i^H] \leq \frac{\sum\limits_i \sum\limits_j \varrho_{ij}\sigma(X_i^K - X_i^H)\sigma(X_j^K - X_j^H)}{[mE(X^K - X^H)]^2}$$

with[4]

$$(9) \qquad \varrho_{ij} \equiv \frac{\operatorname{cov}(X_i^K - X_i^H, X_j^K - X_j^H)}{\sigma(X_i^K - X_i^H)\sigma(X_j^K - X_j^H)}, \quad \varrho_{ij} = 1 \Leftrightarrow i = j,$$

[4] The covariance between two random variables $A$ and $B$, $\operatorname{cov}(A,B) \equiv E\{[A-E(A)][B-E(B)]\}$, is a measure of the strength of the linear correlation between the variables. If $A$ and $B$ are stochastically independent from one another we have $\operatorname{cov}(A,B)=0$. The general formula for the variance of a sum of random variables $Z_i$ is

$$\sigma^2[\sum(Z_i)] = \sum_i \sum_j \operatorname{cov}(Z_i, Z_j) \text{ where } \operatorname{cov}(Z_i, Z_i) = \sigma^2(Z_i)$$

$$= \sum_i \sum_j \frac{\operatorname{cov}(Z_i, Z_j)}{\sigma(Z_i)\sigma(Z_j)}\sigma(Z_i)\sigma(Z_j) = \sum_i \sum_j \varrho_{ij}\sigma(Z_i)\sigma(Z_j).$$

Cf. footnote 35 in chapter III D.

is obtained. Here $\varrho_{ij}$ is the coefficient of correlation between the difference in the profits obtained from the $i$th $K$ and $H$ contracts and the difference in the profits obtained from the $j$th $K$ and $H$ contracts. Fortunately, this coefficient of correlation can be reduced to three other coefficients of correlation that can be interpreted in a more plausible way: $\varrho_{KK}$ and $\varrho_{HH}$ for the correlation between the losses[5] of two arbitrary, but different, $K$ or $H$ contracts respectively, and $\varrho_{KH}$ for the correlation between an arbitrary $K$ and an arbitrary $H$ contract.

With

$$\sigma^2(X_i^N)=\sigma^2(X_j^N) \quad \text{and} \quad E(X_i^N)=E(X_j^N) \;\forall i,j, (N=K,H)$$

(9) may then be transformed to

$$\text{(10)} \qquad \varrho_{ij}=\frac{\varrho_{KK}\sigma^2(X^K)+\varrho_{HH}\sigma^2(X^H)+2\varrho_{KH}\sigma(X^K)\sigma(X^H)}{\sigma^2(X^K)+\sigma^2(X^H)+2\varrho_{KH}\sigma(X^K)\sigma(X^H)}, \; i\neq j,$$

a step that can easily be verified by splitting (9) and (10) according to the rules of expectation algebra. Since $\varrho_{ij}\equiv\varrho=\text{const.}$ if $i\neq j$ and since $\sigma(X_i^K-X_i^H)=\sigma(X_j^K-X_j^H)$, expression (8) can be written as

$$\text{(11)} \qquad W[\Sigma X_i^K\leq\Sigma X_i^H]\leq\frac{m+\varrho(m^2-m)}{m^2}\left[\frac{\sigma(X^K-X^H)}{E(X^K-X^H)}\right]^2.$$

(Here the superfluous indices have been dropped.) Hence, the upper limit of the probability that operating in the market with the higher expected profit per contract brings about a lower level of aggregate profits is given by

$$\text{(12)} \qquad \lim_{m\to\infty} W[\Sigma X_i^K\leq\Sigma X_i^H]\leq\varrho\left[\frac{\sigma(X^K-X^H)}{E(X^K-X^H)}\right]^2.$$

Obviously, when $\varrho>0$, this upper limit does not take on a value of zero.

For random variables that are correlated in the way described above, the Law of Large Numbers has no longer any significant influence. The reason is that the netting out of dispersions takes place only for those parts of the variance that cannot be explained by mutual regressions. The part of the variance which is brought about by factors common to

[5] With non-random insurance premiums the coefficient of correlation between the profits from two contracts equals that between the corresponding losses the company underwrites. We forgo the proof of this simple fact.

all single risks cannot be eliminated by increasing the number of contracts pooled by the company[6].

### 3. *Weber's Relativity Law as the Proper Basis of the Mean-Value Criterion in the Case of Large Numbers*

The doubts concerning the practical relevance of the Law of Large Numbers discussed so far were related to the possibly unrealistic conditions for this law, but they were not related to the law itself. The law is, however, fundamentally called into question by SCHNEEWEISS (1967a, pp. 173–183 and 1968)[7]. He doubts whether it is suitable for legitimating the mean-value criterion even if all its conditions are met. The probability that a multiple performance of one project brings about a better result than a multiple performance of another does not, he argues, indicate reliably which distribution of the sum of profits is characterized by the higher level of expected utility. For example a project, which is very likely to perform better than some other project, may produce only a slight comparative gain in utility if it performs better, but may bring about a dramatic comparative loss in utility, if it performs worse. Hence, despite the operation of the Law of Large Numbers, this project might have a lower level of expected utility than the other.

Schneeweiß indicates a utility function which, for the class of normal distributions, has the property that the '$\mu$ criterion for multiple risks', as he calls it, does not hold. It is Freund's function of constant absolute risk aversion $U(v) = -e^{-\beta v}$ that was criticized above[8].

Even when the problem of a wealth independence of risk aversion is neglected, this function is not strictly applicable, since it neglects the BLOOS rule[9] which implies that $U(v) = U(0)$ for $v \leq 0$. Assume, however, that the analysis is confined to projects with $E(X) > 0$. Then the coefficient of variation of the end-of-period wealth distribution approaches zero as the number of projects simultaneously carried out nears infinity, that is,

---

[6] This objection was made to Knight (for example by NIEHANS (1948)) who, with reference to the Law of Large Numbers, had contended that economic risks can normally be consolidated by forming large groups. See KNIGHT (1921, chapter VIII, esp. pp. 213 and 238 f.).

[7] A similar point was made by SAMUELSON (1963) who proved the following theorem: 'If at each income or wealth level within a range, the expected utility of a certain investment or bet is worse than abstention, then no sequence of such independent ventures (that leaves one within the specified range of income) can have a favorable expected utility.'

[8] Cf. chapter III A 2.3.2.

[9] Cf. chapter III B.

$$(13)\qquad \lim_{m\to\infty}\frac{\sigma(aq+\sum_{i=1}^{m}X_i)}{E(aq+\sum_{i=1}^{m}X_i)}=\frac{\sqrt{m}\,\sigma(X)}{aq+mE(X)}=0$$

($aq$ = interest-augmented initial wealth).

Because of Chebyshev's inequality this property ensures that the probability of negative gross wealth vanishes. In other words, the contribution to expected utility of the modified branch of the utility function which everywhere takes on finite values becomes insignificant [10]. Thus, since Schneeweiß is only interested in the limiting case $m\to\infty$, it turns out that under the condition $E(X)>0$ it is permissible to use the function $U(v)=-e^{-\beta v}$ right from the beginning. We shall therefore do this when considering the argument that Schneeweiss (1967a, p. 178), in a slightly different form, presented in favor of his paradox.

As shown by Freund (1956), a combination of the utility function $U(v)=-e^{-\beta v}$ and the normal distribution which has the density [11]

$$\frac{1}{\sqrt{2\pi}}e^{-(1/2)[(v-\mu)/\sigma]^2}$$

implies that the aim of maximizing expected utility,

$$(14)\qquad \max\int\frac{1}{\sqrt{2\pi}}e^{-(1/2)[(v-\mu)/\sigma]^2}(-e^{-\beta v})dv,$$

is equivalent to

$$(15)\qquad \max\left\{-e^{(\beta^2\sigma^2/2)-\beta\mu}\int\frac{1}{\sqrt{2\pi}}e^{-(1/2)[(v-\mu+\beta\sigma^2)/\sigma]^2}\right\}dv.$$

This expression reduces to

$$(16)\qquad \max\left(\mu-\frac{\beta}{2}\sigma^2\right),$$

---

[10] Schneeweiss (1967a, p. 182) demonstrates this for normal distributions by allowing for almost arbitrary modifications of the utility function over the negative half of the *income* axis.

[11] Cf. footnote 22 in chapter II A.

since $\beta$ = const. and the integral in (15) takes on the constant value 1 independently of $\mu$ and $\sigma$.

Assume now that all single projects $X_i$ are characterized by a normal distribution so that $\Sigma X_i$ is normal, too. Then, because of (16), the expected-utility maximizer whose preferences are described by Freund's function has to choose the projects so as to

$$(17) \qquad \max\{E(aq+\sum_{i=1}^{m} X_i)-\frac{\beta}{2}\sigma^2(aq+\sum_{i=1}^{m} X_i)\}.$$

Because $aq$ = const. and because of the independence of the $X_i$, this postulate is equivalent to

$$(18) \qquad \max\sum_{i=1}^{m}[E(X_i)-\frac{\beta}{2}\sigma^2(X_i)]$$

$$(19) \qquad =\sum_{i=1}^{m}\max[E(X_i)-\frac{\beta}{2}\sigma^2(X_i)].$$

The crucial point in (19) is that the best project can be determined *independently* of the frequency with which it is carried out. Thus we indeed have an isolated decision rule of the type we sought. Paradoxically[12] the rule contradicts the mean-value criterion $\max E(X_i)$ even in the case where the number of performances approaches infinity, i.e., even where, according to the Law of Large Numbers, the project with the highest mathematical expectation almost certainly brings about a higher sum of gains than any other project, including the project that has the highest value for $E(X)-\beta\sigma^2(X)/2$.

Unfortunately the relevance of this result is unnecessarily limited since the $X_i$ were assumed to be normally distributed[13]. However, the same result may be obtained even without the assumption of a particular distribution class[14]. Only the finiteness of expected utility has to be assumed, a condition that, for concave utility functions, is definitely satisfied if the first moment exists and the distribution under consideration is bounded to the left.

---

[12] In fact, of course, there is only an apparent paradox.

[13] SCHNEEWEISS (1968c, p. 100) formulates a general distribution-free criterion for examining whether a utility function implies the mean-value criterion for multiple risks. When verifying the function $-e^{-\beta v}$, however, he again assumes a normal distribution.

[14] The argument is similar to the one that SAMUELSON (1971) raised against the growth-optimum portfolio model of LATANÉ (1959).

According to the expected-utility rule, with $m$ performances, the best project is characterized such that the postulate

$$(20) \qquad \max E[U(aq + \sum_{i=1}^{m} X_i)]$$

or, for $U(v) = -e^{-\beta v}$, the postulate

$$(21) \qquad \max E[-e^{-\beta(aq + \sum_{i=1}^{m} X_i)}]$$

is satisfied. The latter is equivalent to

$$(22) \qquad \max E[-\prod_{i=1}^{m} e^{-\beta X_i}],$$

so that, because of the assumed stochastic independence of the $X_i$, we finally have [15]

$$(23) \qquad \max \prod_{i=1}^{m} E(-e^{-\beta X_i}) = \prod_{i=1}^{m} \max E(-e^{-\beta X_i}).$$

This expression confirms the above result for nearly arbitrary distribution classes. In the case of constant absolute risk aversion, the best project can be chosen independently of the number of performances by reference to the expected utilities of the single projects.

It must be stressed that the assumptions used above are, in fact, more general than those of Schneeweiß although he explicitly (1967a, p. 174) claims that, for his argument, there are no constraints on the distribution class [16]. Schneeweiß argues that, with an increase in the number of performances, the sum distribution in any case converges towards the normal distribution so that the use of equations (14)–(19) is possible even though the $X_i$ are not distributed normally. It is true that, according to the Central Limit Theorem, the *standardized* form of the sum distribution converges towards the standard normal distribution. However, does this really imply, as would be necessary, that the expected utility of the sum distribution converges towards the expected utility of

[15] In the case of stochastic independence between two random variables $Z$ and $X$ we have $\text{cov}(X, Z) = 0$ and hence $E(ZX) = E(X)E(Z)$. This allows for a stepwise transformation of (22) into (23) since independence between $X_i$ and $X_j$ of course also means independence between $-e^{-\beta X_i}$ and $-e^{-\beta X_j}$, $i \neq j$.

[16] Schneeweiß assumes that the mathematical expectation if finite.

a normal distribution with equal mean and standard deviation? The proof has not been given and cannot be given. Schneeweiß contends that, for almost arbitrary distribution classes, the approaches (17) and (20) lead to the same choice of project, i.e.,

$$(24) \qquad \lim_{m\to\infty} \max \left\{ E(aq + \sum_{i=1}^{m} X_i) - \frac{\beta}{2}\sigma^2(aq + \sum_{i=1}^{m} X_i) \right\} \sim \lim_{m\to\infty} \max E[U(aq + \sum_{i=1}^{m} X_i)].$$

If this contention were correct then, as a comparison between (19) and (23) shows, the following equivalence would have to hold on the level of single projects:

$$(25) \qquad \max\,[E(X_i) - \frac{\beta}{2}\sigma^2(X_i)] \sim \max E(-e^{-\beta X_i}) \quad \forall i.$$

This, however, is not true for general distribution classes [17] but only for normal distributions, a fact which, incidentally, was shown by SCHNEEWEISS himself (1967a, pp. 89–98, 146–148). Thus, for the reasoning given by Schneeweiß, the restrictive assumption that, for each single project, the distribution of gains has to be normal cannot be avoided.

Nevertheless we can conclude that, under constant absolute risk aversion, the mean-value criterion for multiple risks cannot be justified [18]. When confronted with practical decision making under uncertainty, this result seems highly implausible as 'rational' as it might be and as 'reasonable' as SCHNEEWEISS (1967a, p. 175) thinks the utility function $-e^{-\beta v}$ is. KRELLE (1968, fn. p. 174) therefore attempts to correct the implausible result by truncating the tails of the probability distributions, a procedure he justifies by appealing to the frequently observable neglect of small probabilities. This procedure, however, has an *ad hoc* character and cannot be accepted for a normative analysis [19].

The apparent paradox can be solved in a quite natural way if it is

---

[17] Cf. the discussion about the rationality of the $(\mu, \sigma)$ criterion in chapter II D 2.1.3, esp. footnote 14.

[18] Provided that the decision maker is risk averse. In the case of a linear utility function there is, of course, risk neutrality that prevails independently of wealth and hence the mean-value criterion is appropriate.

[19] KRELLE (1961, p. 588) argued that the mean-value criterion is a 'zwingendes Gebot der Rationalität' (imperative dictated of rationality). In KRELLE (1968), there is no corresponding remark; here he merely seems to be interested in a positive description of behavior.

assumed realistically that the preference structure of the decision maker obeys Weber's law[20]. According to chaper III[21] this implies that there is an indifference-curve system in the $(\mu, \sigma)$ diagram where the certainty equivalent is a function of the type

$$(26) \qquad S(V) = \Omega\left(\frac{\sigma(V)}{E(V)}, Z\right) E(V).$$

$Z$ is defined as a description of the linear distribution class to which $V$ belongs and it holds that

$$(27) \qquad \lim_{b \to 0} \Omega(b, Z) = 1,$$

$$(28) \qquad \Omega(b, Z) < 1, \quad \text{if} \quad b > 0,$$

provided that at least one of the following constellations prevails:

$$(a) \qquad \upsilon > 0, \qquad\qquad 0 < \varepsilon < \infty.$$

$$(b) \qquad \left.\begin{cases} -\infty \leq \upsilon \leq +\infty, \\ \dfrac{\sigma(V)}{E(V)} \text{ is small,} \\ \text{for } \upsilon \to -\infty \text{ the density converges at least} \\ \text{as fast as with a normal distribution} \end{cases}\right\} 0 < \varepsilon < 1.$$

Here, as usual, $\varepsilon$ is the measure of relative risk aversion and $E(V) < \infty$ is assumed to be self-evident.

Under these conditions, the best project can be found according to the maxim

$$(29) \qquad \max S(V) = \max\left\{\Omega\left[\frac{\sigma(aq + \sum_{i=1}^{m} X_i)}{E(aq + \sum_{i=1}^{m} X_i)}, Z\right] E(aq + \sum_{i=1}^{m} X_i)\right\}$$

$$= \max\left\{\Omega\left[\frac{\sqrt{m}\,\sigma(X)}{aq + mE(X)}, Z\right] [aq + mE(X)]\right\},$$

[20] Maintaining the assumption of normal distributions SCHNEEWEISS (1967a, pp. 179–181) finds a similar solution for the utility functions following from Weber's law. He however needs versions of these utility functions that have already been replaced on the positive half of the wealth axis (identifying wealth with what he calls 'income') by suitable other functions and thus are not strictly compatible with our preference hypothesis.

[21] Cf. chapter III A 2.2, A 2.3, B 1.2, and B 2.

which because of $\lim_{m\to\infty}\Omega(.,Z)=1=\text{const.}$ leads to the mean-value criterion:

$$\lim_{m\to\infty} \max S(V) \sim \max E(X). \tag{30}$$

It should be noted that this result is valid although of course, when $m$ approaches infinity, the linear distribution class $Z$ changes continuously. Since the limit (27) is independent of $Z$, this change is irrelevant. After all, albeit under somewhat more restrictive conditions, we could also make use of the method of local approximation where $\Omega$ is, in any case, independent of $Z$.

For a proper interpretation of (30) it is useful to clarify the meaning of conditions (a) and (b) for the case of single projects. When $m\to\infty$, condition (a) is only satisfied if $x>0$ and, since $\lim_{m\to\infty}\sigma(V)/E(V)=0$, condition (b) is already satisfied if $-\infty<\underline{x}<x<\bar{x}<+\infty$. The boundedness of $x$ satisfies condition (b) since it ensures that the sum distribution $aq+\Sigma x_i$ is bounded for finite $m$ and approaches the normal distribution as $m\to\infty$ so that the convergence condition concerning the density of the wealth distribution is always met. But of course condition (b) would also be satisfied if $X$ were normally distributed. It is worth noting that the mean-value criterion is not appropriate if $x$ can take on negative values and $\varepsilon\geq 1$. Whatever the size of the initial wealth, in this case, for $m$ sufficiently large, $v$ may become negative so that the lexicographic critical wealth level at $v=\bar{v}=0$ requires the maximization of the, always strictly positive, survival probability.

Although the mean-value criterion has now received its proper basis from the expected-utility theory, the arguments concerning the correlation of risks and the 'smallness' of numbers, that were raised above against the Law of Large Numbers, of course remain valid. Both these arguments imply that the coefficient of variation of the wealth distribution never approaches zero so that its size has to be taken into account when a choice is made among the different projects. Concerning the smallness of numbers we can surely dispense with an explanation. Concerning the correlation the reader is advised to verify for himself that

$$\frac{\sigma(V)}{E(V)}=\frac{\sigma(aq+\Sigma X_i)}{E(aq+\Sigma X_i)}=\frac{\sqrt{\sum\limits_i\sum\limits_j \varrho_{ij}\sigma(X_i)\sigma(X_j)}}{aq+\Sigma E(X_i)}, \tag{31}$$

which with $\varrho_{ii}=1$ and $\varrho_{ij}\equiv\varrho=\text{const.}$, $i\neq j$, similarly to (11) implies that

$$\frac{\sigma(V)}{E(V)}=\frac{\sqrt{m+(m^2-m)\sigma^2(X)}}{aq+mE(X)}=\frac{\sqrt{1/m+\varrho(1-1/m)\sigma^2(X)}}{\dfrac{aq}{m}+E(X)}, \tag{32}$$

$$\lim_{m\to\infty}\frac{\sigma(V)}{E(V)}=\sqrt{\varrho}\frac{\sigma(X)}{E(X)}.$$

4. *Conclusions*

According to the Law of Large Numbers, when two risk projects are considered, the one that has the higher expected gain with a probability approximating certainty brings about a higher utility *ex post* if the frequency of performances is sufficiently high. Thus it appears that, independently of subjective preferences, the expected gain may be taken to be the preference functional. This appearance is not deceptive, but the reason given is wrong. The project that, with a probability approximating certainty, brings about a higher utility will not necessarily be characterized by the higher level of expected utility *ex ante*; whether or not both advantages coincide is a matter of the utility function. If the utility function is $U(v) = -e^{-\beta v}$, so that the decision maker's preferences exhibit constant absolute risk aversion, then a choice among single projects can be made, irrespective of how often they are carried out. The Law of Large Numbers is irrelevant in this case. If, however, the utility function obeys Weber's law, then the Law of Large Numbers is reinstated. For strictly positive distributions of gains and/or weak risk aversion ($\varepsilon < 1$), a sufficiently high frequency of performances ensures that the project with the highest expected gain brings about the highest level of expected utility. Thus both of the above-mentioned advantages coincide. However, for distributions that extend partly over the negative half of the wealth axis the Law of Large Numbers continues to be ineffective, if the degree of risk aversion is sufficiently high ($\varepsilon \geq 1$).

Apart from these fundamental aspects of the Law of Large Numbers in practical decision making, it seems that the problems of 'small numbers' and of a correlation of risks are severe obstacles to its effectiveness. Therefore it seems wise to base decisions on the mean-value criterion in rare cases only.

## Section B
## Sequential Risks

With the analysis of sequential risks, time becomes a new dimension in the theory of decision making under uncertainty. Even after the early studies of von Böhm-Bawerk (1884, 1888) and Fisher (1906, 1930) this dimension remained foreign to economic theory for a long time, indeed, there are areas like, for example, the theory of general equilibrium where it did not appear until quite recently.

The neglect of time is not necessarily a short-coming. Certainly there are people with myopic behavior which SAVAGE (1954, p. 16) characterized by the precept 'You can cross the bridge when you come to it.' Thus, for the sake of positive analysis, we could possibly be content with the theory presented up to now, but, at least from a normative point of view, the observation of myopic behavior is no excuse for ignoring time. It is certainly unreasonable to close one's eyes to the future. But, even for describing man's real behavior, myopia seems a slender reed. There are good reasons why the Keynesian myopic consumer was dethroned in FRIEDMAN's (1957) Nobel Prize winning study *A Theory of the Consumption Function*. With this study, which became famous primarily because of its empirical results, a process of reconsideration of preference theory began that did not put the theory of economic decision making under uncertainty to one side.

In this section it is assumed that *homo oeconomicus* optimizes over time. We shall see to what extent, if at all, the results derived in the previous two chapters have to be modified. Surprisingly, this can be revealed already, Savage's precept will be shown to be wiser than it seems at first sight.

The problem of optimal multiperiod planning is approached in two steps. First, the one-period approach discussed above is generalized to the case of sequential decision making under the assumption that no consumption occurs until the planning horizon. After this, the assumption is removed and the task becomes to find an intertemporal decision strategy that maximizes a preference functional over probability distributions of consumption paths which has yet to be specified. The strategy involves sequential replanning where, at each point in time, a simultaneous choice of the optimal risk project for the current period and of the optimal level of period consumption is made.

In contrast to the one-period approach, in the multiperiod analysis it is necessary to consider the decision maker's opportunity set of risk projects or probability distributions in more detail. In the one-period analysis there was nothing wrong with assuming this opportunity set to be exogenous. In the multiperiod case this assumption no longer makes sense, since, in general, the set of alternatives available depends on the decision maker's wealth which, in turn, depends on the variates of previously chosen risk projects. This is very clear for the problem of portfolio holding, where risk taking involves the lending of capital, or for the problem of the investment decisions of a firm which is partly able to finance risky and profitable projects by credit but where the credit itself is limited by the amount of equity. Moreover, in insurance the decision maker's wealth is usually very closely connected with the size of his insurable risk.

A simple way of modelling the relationship between risk and wealth is to interpret the amount of wealth $a_t$ net of period consumption that is available at a point in time $t$ as a scale factor which, after a multiplication with a stochastic *standard income factor* $R_{t+1}$, determines the period income $Y_{t+1}$ available at point in time $t+1$:

$$Y_{t+1} = R_{t+1} a_t. \tag{1}$$

Up to now the prevalence of a non-random interest income had been assumed. Since it, too, is proportional to initial wealth, this formulation includes it as a special case. Given the period income as formulated in (1) the level of wealth available at point in time $t+1$ is

$$V_{t+1} = Q_{t+1} a_t, \tag{2}$$

where $Q_{t+1} = 1 + R_{t+1}$. With probability distribution $Q_{t+1}$ (respectively $R_{t+1}$) one out of a set of alternative *standard risk projects* is defined. If this set is independent of $a_t$ then we have what ARROW (1965, p. 37) calls *stochastic constant returns to scale.* Because of

$$E(V_{t+1}) = a_t E(Q_{t+1}) \tag{3}$$

and

$$\sigma(V_{t+1}) = a_t \sigma(Q_{t+1}) \tag{4}$$

stochastic constant returns to scale bring about alternative opportunity sets in the $(\mu, \sigma)$ diagram which, as shown in Figure 1, may be derived from one another by a projection through the origin.

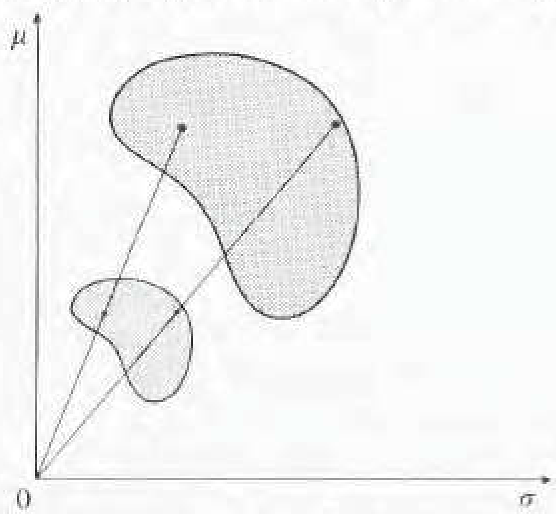


*Figure 1*

For the decision problem of a portfolio optimizer, the assumption of stochastic constant returns to scale is immediately plausible: if Peter buys a packet of shares for \$ 1000 and Paul buys one for \$ 2000 and if, furthermore, Paul buys exactly twice as many of each kind of share, after one period Paul's packet will be precisely twice as valuable as Peter's, whatever happens to the share prices. This and further examples of exact, or at least approximately, stochastic constant returns to scale are discussed in chapter V which is concerned with applications. Here it is sufficient to have a rough idea of one application.

In addition to the basic assumption of stochastic constant returns to scale the following assumptions, that are helpful for finding a solution to the multiperiod problem, are made.

(1) The set of standard risk projects contains a non-empty subset of those projects for which $q_t > q_{\min} > 0$, where $q_t$ is a variate of the random variable $Q_t$ and $q_{\min}$ is an arbitrarily choosable small number. This assumption implies that in the opportunity set there is at least one choice possibility that with certainty prevents the decision maker from losing his wealth. Whether this choice will then be made is another question.

(2) The opportunity sets of standard risk projects $Q_t$ and $Q_{t^*}$, $t \neq t^*$, are stochastically independent. The justification of this assumption will be examined in chapter V case by case.

(3) The Bloos rule holds[1]. It implies

1. that the gross distributions that were referred to in the above remarks are associated with net distributions which are constraint to positive values and
2. that for these net distributions stochastic constant returns to scale (including auxiliary assumptions (1) and (2)) prevail also.

If we distinguish between a gross risk project $Q_t^g$ and its net counterpart $Q_t^n$ then both aspects follow from relationship (III B 1) which implies

$$Q_t^n = \begin{cases} 0 & \text{if } Q_t^g \leq 0, \\ Q_t^g & \text{if } Q_t^g \geq 0. \end{cases}$$

The calculations of this chapter always refer to net distributions. To abbreviate the notation, however, the superscript '$n$' is dropped.

(4) The decision maker either has no human capital or, if he has, it comes from an exogenous and non-random flow of labor income that is discounted at a non-random rate of interest. This assumption once again makes use of the wealth concept introduced at the begin-

[1] Cf. chapter III B.

ning of chapter III. Other non-random and exogenous flows of income, from whatever source, may occur in a similar way.

With these assumptions and specifications we are well prepared to attempt the first step into the multiperiod world.

## 1. *Optimal Multiperiod Planning of a Pure Investment Program under Uncertainty*

Consider an investor who at point in time zero has a level of wealth $a_0$ and plans to invest this wealth until a point in time $T$, $1 \leq T < \infty$. Before $T$ he does not withdraw any capital for consumption purposes but at $T$ he wants to have his potlatch, the great feast where everything is used up. This investor's aim is

$$(5) \qquad \max E[U(V_T)]$$

where $U(.)$ is one of the Weber functions described in (III A 34). If we set $a_t = v_t$ to indicate that the total wealth available at point in time $t$ is immediately reinvested then

$$(6) \qquad V_T = a_0 \prod_{t=1}^{T} Q_t,$$

so that (5) becomes

$$(7) \qquad \max_{Q} E[U(a_0 \prod_{t=1}^{T} Q_t)].$$

The maximization operator in this formula refers to the choice of a standard risk project for each period. The standard risk projects do not have to be identical nor do they have to be already determined at point in time 0. It is assumed that they have to be determined no earlier than one period before their outcomes are revealed.

### 1.1. *The Growth Optimum Model*

On the basis of the approach of Kelly (1956), the following solution was suggested by Latané (1959, cf. esp. the footnote on p. 151). Independently of the decision maker's personal preferences, the risk projects should be chosen so as to maximize the expected logarithms of the end-of-period wealth distribution[2]:

$$(8) \qquad \max_{Q_{t+1}} E[\ln(a_t Q_{t+1})] \quad \forall t.$$

This is the classical decision rule proposed by Bernoulli. The reason for it is, however, completely different.

Like the argument in favor of the mean-value criterion for multiple risks, it is based on the Law of Large Numbers. Suppose that in a particular period the risk projects $Q$ and $Q'$ are available with $E(\ln Q) > E(\ln Q')$. Then, an immediate application of the above formula (A 6) yields

$$(9) \qquad W[U(a_0 \prod_{t=1}^{T} Q_t) \leq U(a_0 \prod_{t=1}^{T} Q'_t]$$
$$= W[\prod_{t=1}^{T} Q_t \leq \prod_{t=1}^{T} Q'_t]$$
$$= W[\ln \prod_{t=1}^{T} Q_t \leq \ln \prod_{t=1}^{T} Q'_t]$$
$$= W[\sum_{t=1}^{T} \ln Q_t \leq \sum_{t=1}^{T} \ln Q'_t]$$
$$\leq \frac{1}{T}\left[\frac{\sigma(\ln Q - \ln Q')}{E(\ln Q - \ln Q')}\right]$$

and, because of (A 7),

$$(10) \qquad \lim_{T\to\infty} W[U(a_0 \prod_{t=1}^{T} Q_t) \leq U(a_0 \prod_{t=1}^{T} Q'_t)] = 0.$$

Thus, since $\max E[\ln Q_{t+1}] \sim \max E[\ln(a_t Q_{t+1})]$, in the long run the choice of the project for which the expected logarithm of end-of-period wealth is maximal leads, with a probability approximating certainty, to a higher level of final wealth and hence to a higher level of utility.

The result can be criticized in that it was derived on the basis of an *ex ante* decision over all projects rather than from a sequential decision

[2] Cf. also Latané and Tuttle (1967), Breimann (1960), and Thorp (1971). On a theoretical basis Hakansson (1971) compares the $(\mu, \sigma)$ with the growth optimum approach, and Roll (1973) draws the corresponding empirical comparison which, however, cannot, in principle, have a discriminatory power. Cf. footnote 8 below.

making process[3]. Moreover, Latané's conjecture that the result implies the inequality

$$\text{(11)} \qquad \lim_{T\to\infty} E[U(a_0 \prod_{t=1}^{T} Q_t)] > \lim_{T\to\infty} E[U(a_0 \prod_{t=1}^{T} Q_t')]$$

is wrong for a reason similar to that which prevents the Law of Large Numbers, by itself, from legitimating the mean-value criterion. To promise a higher utility with a probability that approximates certainty does not necessarily mean to indicate a higher level of expected utility.

### 1.2. *The Solution by Means of Stochastic Dynamic Optimization*

The true solution of the optimization problem (14) is provided by MOSSIN (1968a)[4]. It is based on *Bellman's Principle of Optimality* which says that[5,6]

$$\text{(12)} \qquad \max\{E[U(V_T)]\,|_{a_0} = E[U(a_0 \prod_{t=1}^{T} Q_t)]\}$$

implies

$$\max\{E[U(V_T)]\,|_{a_{t^*}} = E[U(a_{t^*} \prod_{t=t^*+1}^{T} Q_t)]\} \quad \forall t^*, 0 \le t^* \le T-1.$$

This implies that, at point in time zero, the optimal decision must be sought under the constraint that the wealth realized at point in time 1 is reinvested in the best possible way as seen from that point in time, that, similarly, the wealth available at point in time 2 is optimally reinvested as seen from point in time 2, and so on until finally the chain of reinvestment is interrupted at point in time $T$ by consuming the then available wealth.

From a long-run point of view the optimal decision at point in time zero can be determined by a process of period-by-period recursive

---

[3] For the same reason, we do not consider the multiperiod approach of TOBIN (1965) which is, quite correctly, criticized by MOSSIN (1968, pp. 217 f.).

[4] SAMUELSON (1971) discusses the growth optimum model and achieves the correct solution. His procedure, however, is slightly unsatisfactory since he determines the optimal sequence of risk projects in a one step decision at the beginning of the sequence.

[5] BELLMAN (1957, p. 83) described that principle as follows: 'An optimal policy has the property that, whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision.'

[6] In words, $\max E[U(V_T)]\,|_{a_0}$ means: 'Maximize the expected utility of wealth available at the planning horizon $T$ subject to the constraint that the initial wealth at point in time zero, $a_0$, is given.'

optimization where, at each point in time $t^*$, $0 \le t^* < T-1$, one of the postulates

$$(13)\quad \frac{\begin{Bmatrix}\max\\ \min\end{Bmatrix}\{E(V_T^{1-\varepsilon'})|_{a_{t^*}} = E[(a_{t^*}\prod_{t=t^*+1}^{T} Q_t)^{1-\varepsilon'}]\} \equiv z_{t^*}, \varepsilon'\{\lessgtr\}1,}{\max\{E(\ln V_T)|_{a_{t^*}} = E[\ln(a_{t^*}\prod_{t=t^*+1}^{T} Q_t)]\} \equiv z_{t^*},\ \ \varepsilon' = 1,}$$

has to be satisfied. Here, $U(v_T)$ is specified by the Weber functions $(1-\varepsilon')v^{1-\varepsilon'}$, $0<\varepsilon'\neq 1$, and $\ln v$, $\varepsilon'=1$. The measure of relative risk aversion $\varepsilon'$ has, in principle, the same role as the parameter $\varepsilon$ used up to now. The reason for chosing another symbol is that $\varepsilon$ is to be kept for the measure of relative risk aversion relevant for the decision in the *current* period. The constant factor $(1-\varepsilon')$, which does not affect the optimization task and which, in any case has the sole function of determining the sign, was dropped so that for $\varepsilon'>1$ a *minimization* of the expected value has to be undertaken. In what follows, it is not mentioned when a maximization and when a minimization is appropriate. We agree that only in the case $\varepsilon'>1$ a minimization is carried out. In order to avoid unnecessary repetitions, all functions are discussed simultaneously. The case of the power function is always written first and the case of the logarithmic function always comes second.

*Point in Time $t^* = T-1$*

Suppose there is only one period left before the potlatch and, by chance, a wealth of size $a_{T-1}$ has been accumulated. Then, (13) implies the usual one-period decision problem which is to maximize the expected utility of end-of-period wealth $V_T$ given the initial wealth $a_{T-1}$:

$$\begin{Bmatrix}\max\\ \min\end{Bmatrix}\{E(V_T^{1-\varepsilon'})|_{a_{T-1}} = E[(a_{T-1}Q_T)^{1-\varepsilon'}] = a_{T-1}^{1-\varepsilon'}E(Q_T^{1-\varepsilon'})\}$$
$$= a_{T-1}^{1-\varepsilon'}\begin{Bmatrix}\max\\ \min\end{Bmatrix}E(Q_T^{1-\varepsilon'})$$

$$(14)\quad \frac{\equiv a_{T-1}^{1-\varepsilon'} j_T \equiv z_{T-1}.}{\begin{aligned}\max\{E(\ln V_T)|_{a_{T-1}} &= E[\ln(a_{t-1}Q_T)] = \ln a_{t-1} + E(\ln Q_t)\}\\ &= \ln a_{T-1} + \max E(\ln Q_T)\\ &\equiv \ln a_{t-1} + j_T \equiv z_{T-1}.\end{aligned}}$$

It is worth noting that obviously the choice of one project out of the set of available standard risk projects $Q_T$ can be made independently of the available investment capital $a_{T-1}$. This is a particular property of the Weber functions that is usually called *separation property*[7].

*Point in Time $t=T-2$*

Two periods before the planning horizon, the level of wealth $V_{T-1}=A_{T-1}$ available at point in time $T-1$ is a random variable, while the current level of wealth $a_{T-2}$ is given historically. From (13) we then have

$$z_{T-2}\equiv\begin{Bmatrix}\max\\\min\end{Bmatrix}\{E[V_T^{1-\varepsilon'}]|_{a_{T-2}}=E[(a_{T-2}Q_{T-1}Q_T)^{1-\varepsilon'}]\}$$

(15)

$$z_{T-2}\equiv\max\{E[\ln V_T]|_{a_{T-2}}=E[\ln(a_{T-2}Q_{T-1}Q_T)]\}$$

and

$$\begin{aligned}z_{T-2}&=a_{T-2}^{1-\varepsilon'}\begin{Bmatrix}\max\\\min\end{Bmatrix}E(Q_{T-1}^{1-\varepsilon'})\begin{Bmatrix}\max\\\min\end{Bmatrix}E(Q_T^{1-\varepsilon'})\\&=a_{T-2}^{1-\varepsilon'}j_T\begin{Bmatrix}\max\\\min\end{Bmatrix}E(Q_{T-1}^{1-\varepsilon'})\\&\equiv a_{T-2}^{1-\varepsilon'}j_{T-1}.\end{aligned}$$

(16)

$$\begin{aligned}z_{T-2}&=\ln a_{T-2}+\max E(\ln Q_{T-1})+\max E(\ln Q_T)\\&=\ln a_{T-2}+j_T+\max E(\ln Q_{T-1})\\&\equiv\ln a_{T-2}+j_{T-1}.\end{aligned}$$

In the case $\varepsilon\neq 1$, the step from (15) to (16) is possible, since $a_{T-2}=\text{const.}$ and $Q_T$ is independent of $Q_{T-1}$ so that

$$\begin{aligned}E[(a_{T-2}Q_{T-1}Q_T)^{1-\varepsilon'}]&=a_{T-2}^{1-\varepsilon'}E[Q_{T-1}^{1-\varepsilon'}Q_T^{1-\varepsilon'}]\\&=a_{T-2}^{1-\varepsilon'}E(Q_{T-1}^{1-\varepsilon'})E(Q_T^{1-\varepsilon'}).\end{aligned}$$

There are two reasons for independence. One is the assumption of stochastically independent opportunity sets of standard risk projects that was made at the beginning of section B. The other is the separation property. It is worth noting that, in the case $\varepsilon'=1$, the assumption of independence is not needed because

$$\begin{aligned}E[\ln(a_{T-2}Q_{T-1}Q_T)]&=\ln a_{T-2}+E[\ln Q_{T-1}+\ln Q_T]\\&=a_{T-2}+E[\ln Q_{T-1}]+E[\ln Q_T]\end{aligned}$$

holds even if $Q_{T-1}$ and $Q_T$ are correlated.

[7] Pye (1967) and, less explicitly, also Arrow (1965, pp. 28–44) have detected this implication of constant relative risk aversion.

*Point in Time $t^* = T - \tau$*

We can now continue by using a procedure similar to that used for point in time $T-2$. For point in time $T-\tau$, including the case $\tau = T$ which characterizes the ultimately interesting point in time 0 where the current decision has to be made, we have

$$(17)\quad \frac{\begin{aligned} &z_{T-\tau} = a_{T-\tau}^{1-\varepsilon'} j_{T-\tau+2} \begin{Bmatrix}\max\\ \min\end{Bmatrix} E(Q_{T-\tau+1}^{1-\varepsilon'}) \\ &\text{with}\quad j_{T-\tau+2} \equiv \prod_{t=T-\tau+2}^{T} \begin{Bmatrix}\max\\ \min\end{Bmatrix} E(Q_t^{1-\varepsilon'}),\end{aligned}}{\begin{aligned} &z_{T-\tau} = \ln a_{T-\tau} + j_{T-\tau+2} + \max E(\ln Q_{T-\tau+1}) \\ &\text{with}\quad j_{T-\tau+2} \equiv \sum_{t=T-\tau+2}^{T} \max E(\ln Q_t),\end{aligned}}$$

so that the decision at $t^* = T-\tau$ is optimal from a long run perspective if, and only if, it satisfies

$$(18)\quad \left.\frac{\begin{Bmatrix}\max\\ \min\end{Bmatrix} E(Q_{t^*+1}^{1-\varepsilon'}) \sim \begin{Bmatrix}\max\\ \min\end{Bmatrix} E[(V_{t^*+1})^{1-\varepsilon'}]}{\max E(\ln Q_{t^*+1}) \sim \max E[\ln(V_{t^*+1})]}\right\} \sim \max E(U(V_{t^*+1})].$$

This is an interesting result, for its says that, if the decision maker evaluates wealth available at the planning horizon in a way consistent with Weber's law, he should have a derived preference ordering over end-of-period wealth that also is consistent with Weber's law. It is worth noting, too, that the risk aversion parameter $\varepsilon$, characterizing the derived utility function for end-of-period wealth, is identical with the corresponding parameter $\varepsilon'$ of the utility function for evaluating wealth available at the planning horizon. We can therefore use the same symbol ($U$) for characterizing both functions. Thus the stupid, myopic man, who merely calculates expected utility of end-of-period wealth, reaches the same decision as the smart man, who optimizes his current decision with respect to all future alternatives, provided that both are equally risk averse.

Some words must be added concerning the existence of solution (18). They also apply to the other approach to the intertemporal optimization problem that will be discussed below, but they will not be repeated there. If all distributions $Q_{t+1}$ of the opportunity sets available at points in time $t = 0, 1, \ldots, T-1$ have the property $Q_{t+1} > q_{\min} > 0$ then the existence is ensured since

$$U(q_{\min}) = \begin{cases} (1-\varepsilon')q_{min}^{1-\varepsilon'}, & \varepsilon' \neq 1, \\ \ln q_{\min}, & \varepsilon' = 1, \end{cases}$$

is finite and hence also $j_t$ for $t=1, \ldots, T$. Because of the lower bound of $U(Q_t)=(1-\varepsilon')Q_t^{1-\varepsilon'}$, $\varepsilon'<1$, the existence is even ensured for distributions with $q_{\min}=0$ if risk aversion is weak. What changes if, with $\varepsilon' \geq 1$, strong risk aversion prevails and also $q_{\min}=0$? In this case the opportunity sets may contain risk projects with

$$E[U(Q_t)] = \begin{Bmatrix} E[(1-\varepsilon')Q_t^{1-\varepsilon'}] \\ E(\ln Q_t) \end{Bmatrix} = -\infty$$

so that $j_t$ becomes infinite if it comprises at least one such project. In fact, however, the latter will not occur. According to the definition of $j$ given in (17) only the maxima of $E(\ln Q_t)$ or the minima of $E(Q_t^{1-\varepsilon'})$, $\varepsilon'>1$, enter the formula for $j$ and these are finite, since, by assumption, each opportunity set contains at least one risk project with $Q_t > q_{\min} > 0$ that clearly dominates those which render $j$ infinite. Thus an optimal solution exists even in the case $\varepsilon' \geq 1$ and $q_{\min}=0$.

A more important implication of the preceding considerations is that the multiperiod approach which was first formulated for actual or net wealth can, by using the Bloos rule, easily be generalized to an evaluation of balance sheet or gross distributions incorporating the possibility of negative variates. The implications of the Bloos rule, for example the indifference curves in the $(\mu, \sigma)$ diagram for linear distribution classes, can be almost completely maintained. The only new constraint is that, in the case $\varepsilon' \geq 1$, the opportunity set must contain projects with $Q_t > q_{\min} > 0$. For all the examples of application studied in chapter V, this constraint will not be binding.

The result (18) obviously implies that Latané's argument for the short-run rule $\max E(\ln V)$ cannot be valid. It is true, this rule is included in (18) as a special case[8], but the reason has nothing to do with the Law of Large Numbers. It is simply that Weber's law shows up in the form of a logarithmic utility function. In the presence of a power function which is also compatible with Weber's law the rule $\max E(\ln V)$ definitely is suboptimal.

The result of this section may be summarized as follows. Suppose, in a multiperiod investment program, the decision maker attempts to maximize the expected utility of wealth at the planning horizon, his preferences are compatible with Weber's relativity law, and stochastic constant returns to scale, including auxiliary conditions, prevail. Then the current decision is optimal, if the expected utility of net wealth at the

[8] Thus an empirical discrimination is impossible, and the empirical evidence that Roll (1973) was able to find in favor of the growth optimum model supports equally well the hypothesis that people choose risk projects in line with an intertemporal optimization approach based on Weber's law.

end of the current period is maximized by using the utility function for net wealth at the planning horizon. Because of the BLOOS rule, used here with a minor constraint, this result applies analogously to gross distributions that partly extend over the negative half of the wealth axis. Thus, for example, the indifference curve systems of Figures 10 and 12 in chapter III can still be used for finding the optimal distribution out of a set of wealth distributions that all belong to the same linear class.

## 2. *Optimal Multiperiod Planning of a Consumption-Investment Program*

The aim of reinvesting all returns and consuming the total stock of capital available at the planning horizon $T$, which was assumed above, does not seem to be realistic for most investors. It is more realistic to assume that the funds available at the beginning of each period can, in principle, also be used for consumption (in the model of the firm: for paying dividends). To what extent they actually will be used for current consumption has to be determined by an optimization calculus[9].

Of crucial importance for such an approach is the question of how to model the preference structure of the decision maker. This question is studied in the following section B 2.1 and, as a result, a hypothesis concerning the shape of the multiperiod preference functional is formulated. The implications of this preference functional are derived in section B 2.2 and interpreted in section B 2.3.

### 2.1. *The Multiperiod Preference Functional*

Assume, as before, that there is a given planning horizon $T$. Then the decision maker's task is

$$(19) \qquad \max R(C_0, \ldots, C_{T-1}, V_T),$$

where $R(.)$ is the preference functional yet to be specified, $(C_0, \ldots, C_{T-1})$ a stochastic consumption path, and $V_T$ a stochastic level of wealth available at the planning horizon. In the model of the firm, $V_T$ may be interpreted as the final stock of equity capital available for further

[9] PHELPS (1962) considered this problem in a stochastic multiperiod model, but excluded the choice of the optimal risk project by assuming that only one such project is available to the decison maker. The simultaneous analysis of the twofold problem of how much to invest in which project was studied in two-period models by SANDMO (1968, 1969) and TOBIN (1968). Multiperiod models, developed subsequently, that offer simultaneous solutions to both problems are mentioned in the text.

investment and, in the model of the household, it may be interpreted as the legacy at the end of the decision maker's life.

Concerning the properties of $R(.)$, there are in principle two problems. On the one hand, what $R(.)$ looks like in the special case of a non-random consumption path has to be clarified[10] and, on the other, the particular aspects that emerge in the case of uncertainty have to be determined. Only the second problem is considered in detail here. Concerning the first, use is made of the available literature.

The preference functional for planning problems under certainty is assumed to be

$$(20) \qquad \Sigma = \sum_{t=0}^{T-1} \lambda_t u(c_t) + \lambda_T u(v_T); \quad \lambda_t, \lambda_T > 0,\ u'(.) > 0,\ u''(.) < 0.$$

While $\Sigma$ is an ordinal function defined up to a strictly increasing *monotonic* transformation, the period-utility function $u(.)$ is defined up to a strictly increasing *linear* transformation. The parameter $\lambda_t$ is a period-specific weight factor. The preference functional was introduced in a rudimentary form by RAMSEY (1928) and supplemented with discount factors by SAMUELSON (1936/37). Its axiomatic basis was established by KOOPMANS (1960), and it has been used by STROTZ (1955/56), MODIGLIANI/BRUMBERG (1955), and many others.

Despite its axiomatic foundation by Koopmans, (20) is not a generally accepted maxim for wise intertemporal planning. The reason is primarily the intertemporal separability implied by the additivity of the functional[11]. The separability implies that the level of consumption in one period has no influence on the preference ordering over alternative consumption paths during the periods remaining. Strictly speaking, such an implication is not realistic. However, there are reasons for expecting that the complementarities or substitutabilities between the levels of consumption of two different points in time are weaker the further these points are from one another so that, simply by increasing the length of the periods, the disturbing influences can be reduced[12]. Thus, among the conceivable preference functionals which are compromises between simplicity and realism, (20) seems to be an attractive candidate.

Despite its shortcomings, it must be admitted that (20) has properties that a multiperiod preference functional should, in general, have. An

---

[10] For the sake of brevity the term 'consumption path' is here used to characterize the whole sequence $(c_1, \ldots, c_{T-1}, v_T)$ including the level of wealth $v_T$ remaining at the planning horizon.

[11] Cf. KOOPMANS's (1960) postulate 3.

[12] This was stressed by ARROW and KURZ (1970, pp. 11 f.). Cf. however STROTZ (1957 and 1959).

important aspect, implied by $u'(c)>0$ and $u''(c)<0$, is that the marginal rate of substitution between the consumption levels of two points in time $t$ and $t^*$, $t \neq t^*$, is negative,

$$\left.\frac{dc_{t^*}}{dc_t}\right|_{\Sigma} = -\frac{u'(c_t)\lambda_t}{u'(c_{t^*})\lambda_{t^*}} \leq 0, \tag{21}$$

and in addition diminishing in absolute terms:

$$\left.\frac{d^2c_{t^*}}{dc_t^2}\right|_{\Sigma} = \frac{-\lambda_{t^*}^2 u'(c_{t^*})^2 \lambda_t u''(c_t) - \lambda_t^2 u'(c_t)^2 \lambda_{t^*} u''(c_{t^*})}{u'(c_{t^*})^3} > 0. \tag{22}$$

Another aspect is represented by the weight factors $\lambda_t$. Because of the assumption of constant period utility $u(c_t)$ we have $-dc_{t^*}/dc_t|_{\Sigma}=1$ if $\lambda_{t^*}=\lambda_t$ and $c_{t^*}=c_t$, $t\neq t^*$. Thus, with $\lambda_t=\text{const.}>0$ for all $t$, a rate of time preference $-dc_{t^*}/dc_t|_{\Sigma}-1$ different from zero can only occur if $c_{t^*}\neq c_t$: only VON BÖHM-BAWERK's (1888, pp. 328–331) first reason for time preference, namely, the *Verschiedenheit des Verhältnisses von Bedarf und Deckung* (difference in the relationships between needs and funds) is captured. Von Böhm-Bawerk's second reason (pp. 332–338), the *Minderschätzung zukünftiger Bedürfnisse* (underestimation of future wants), which occurs because the future seems less important when looked at from the perspective of the present, is taken into account by the subjective discount factors $\lambda_t$. In dynamic optimization models it is usually assumed for these that[13]

$$\lambda_{t-1} > \lambda_t > \lambda_{t+1} \quad \forall t < T-1 \tag{23}$$

so that, even in the case $c_{t^*}=c_t$, there is a positive rate of time preference: $-dc_{t^*}/dc_t|_{\Sigma}-1>0$. For the sake of illustration, Figure 2 shows the indifference curves for both the case without time preference and the case where $\lambda_{t^*}<\lambda_t$ for $t^*>t$ in Fisher's two-period diagram.

Although there can be no doubt that people do see their future wants in a diminished perspective, it is by no means clear if account should be taken of this fact in multiperiod dynamic optimization. Actually, von Böhm-Bawerk's argument is based on the *irrationality* of man: *ex post*

[13] This formulation includes the case where, with the passage of time $t$, the relative weights $\lambda_{t+\tau}/\lambda_t$, $\tau\geq 1$, of the future periods, as seen from point $t$, are changing. If we follow STROTZ (1955/56) and postulate that $\lambda_{t+\tau}/\lambda_t$ is independent of $t$, then $\lambda_t=e^{-\varrho t}$ emerges, where $\varrho$ is the rate of discount in the sense of von Böhm-Bawerk's second reason. But why should $\lambda_{t+\tau}/\lambda_t$ be independent of $t$? There seems to be hardly any justification for discrediting as irrational the possibility that the current rate of time preference changes with age.

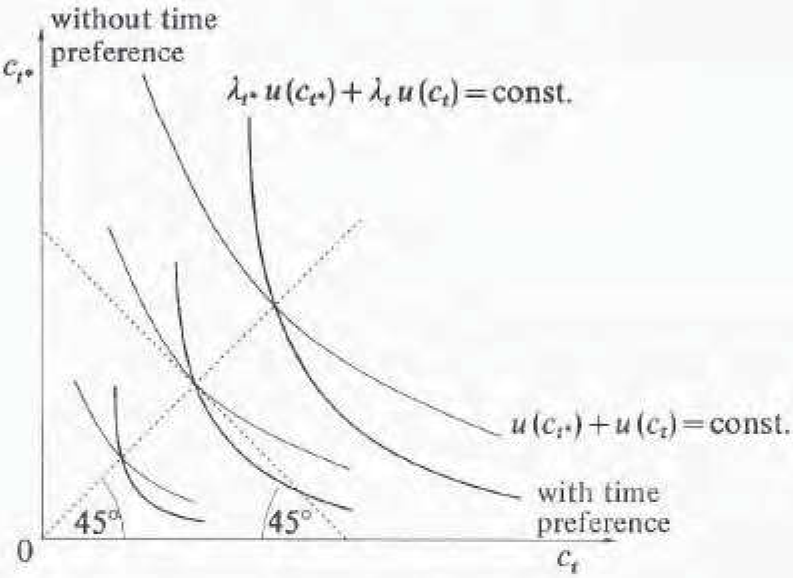


*Figure 2*

he regrets having underestimated his wants[14]. Irrationality, however, is not particularly appropriate in a dynamic optimization model, where a very high level of rationality is otherwise assumed. But at least, if the approach is interpreted in the sense of an individual life-cycle model, the reason for time preference, instead of being an underestimation of futures wants, may be a lower valuation of these wants which people stick to even *ex post*. The sentiment 'just as well I had a good time while I was still young enough to enjoy it' is a pertinent example of this kind of thinking. Thus we maintain (23) as characterizing the normal case without, however, excluding the possibility $\lambda_t = 1 \; \forall t < T$.

The rule (23) does not cover the factor $\lambda_T$. The reason is that this factor not only has a discounting function but also estimates the importance of the legacy left to the heirs. If the model is interpreted from the viewpoint of a firm, then $\lambda_T$ is a measure of the importance of the equity capital available at the planning horizon.

2.1.1. Specific Risk Preference in Multiperiod Planning

How can the preference functional (20) be generalized to the case of stochastic consumption paths? One possibility is simply to put the expectation operator in front, i.e., to set

$$(24) \qquad R(.) = E[\sum_{t=0}^{T-1} \lambda_t u(C_t) + \lambda_T u(V_T)].$$

[14] A similar point of view is expressed by Pigou (1932, p. 25): '... people distribute their resources between the present, the near future and the remote future on the basis of a wholly irrational preference.' Cf. also Strotz (1955/56, esp. p. 178).

This is what HAKANSSON (1969 and 1970a), SAMUELSON (1969), MERTON[15] (1969), LELAND (1974), and many others have done without, however, their making any attempt to justify the procedure.

The approach, unfortunately, is by no means satisfactory, since it does not allow for a *specific risk preference*. Suppose there are two people who, for all conceivable multiperiod decision problems under certainty, reach the same decisions. Is there any reason to assume that these people will also act in the same way when making decisions under uncertainty? Surely an answer in the affirmative cannot be given, for they may well differ in their risk preferences. Specific differences in risk preferences are clearly excluded by the preference functional (24). People who act alike under certainty have, up to a unique positive linear transformation, the same $u$'s and $\lambda$'s and hence are forced to behave alike under risk as well.

This by no means implies that (24) does not bring about risk averse behavior. As the cited approaches show, the concavity of period utility in fact ensures that, for the choice in each period, risk aversion prevails. The only problem is that this risk aversion is simply a byproduct of the concavity of $u(.)$, which in the non-stochastic model has the task of providing for a diminishing marginal rate of substitution between the consumption levels at two points in time (cf. (22)) and hence ensures that the decision maker does not concentrate his consumption in a single period. What is missing in (24) is an additional tool by which the decision maker's risk preference can be manipulated without at the same time altering the preferences relating to the time profiles of consumption in a world of certainty. Fortunately, a suitable tool seems to be available.

By its very nature, the deterministic preference functional $\Sigma$ from (20) is defined up to a strictly positive monotonic transformation. This means that, without any behavioral implications in the case of certainty, we can replace $\Sigma$ by $\Psi(S)$ where $\Psi(.)$ is an arbitrarily choosable, strictly increasing, function. Of course, this does not affect the marginal rate of substitution under certainty

$$(25) \qquad \left.\frac{dc_{t^*}}{dc_t}\right|_{\Psi(\Sigma)} = -\frac{\dfrac{\partial \Psi(.)}{\partial \Sigma}\,\dfrac{\partial \Sigma}{\partial c_t}}{\dfrac{\partial \Psi(.)}{\partial \Sigma}\,\dfrac{\partial \Sigma}{\partial c_{t^*}}} = -\frac{u'(c_t)\lambda_t}{u'(c_{t^*})\lambda_{t^*}}$$

(cf. (21)). However, if we follow the procedure of Hakansson, Samuelson, and others and construct the preference functional for the

[15] Merton uses a model with continuous time.

case of uncertainty by simply putting the expectation operator in front,

$$(26) \qquad R(.)=E\{\Psi[\sum_{t=0}^{T-1}\lambda_t u(C_t)+\lambda_T u(V_T)]\},$$

then, in general, the shape of the function $\Psi(.)$ takes on an important role. While positive linear transformations of $\Psi(.)$ are irrelevant, changes in its *curvature*, like changes in the curvature of the one-period utility function, do have a significant influence on the optimal decision if probability distributions of consumption paths and hence probability distributions of $\Sigma$ are to be evaluated. Hakansson *et al.* assumed $\Psi(.)$ to be linear. This is a special assumption that is possible but arbitrary. *A priori* it is no more and no less arbitrary than any other special assumption for $\Psi(.)$. Rather than arbitrarily assuming a particular shape of $\Psi(.)$, it is therefore tempting to use this function as a *specific risk preference function*[16], that is, to use it as the additional tool we sought.

Provided the deterministic preference functional is of the type (20), the use of the specific risk preference function $\Psi(.)$ is not only plausible, but is close to being a cogent rule of logic, as cogent, at any rate, as the expected-utility rule in the one-period case. This can easily be shown with the use of four axioms, the first three of which are known in principle from the one-period analysis (cf. chapter II C 2.1).

(1) Axiom or Ordering: *The decision maker has a complete weak ordering over probability distributions C of consumption paths*

$$c=(c_0, c_1, \ldots, c_{T-1}, v_T).$$

(2) Axiom of Strong Independence: *Suppose that, comparing the probability distributions $C^1$ and $C^2$, the decision maker reveals the preference $C^1\{\precsim\}C^2$. Then, combining these distributions with an arbitrarily given third distribution $C^3$, it holds that*

$$\begin{pmatrix} w & 1-w \\ C^1 & C^3 \end{pmatrix}\{\precsim\}\begin{pmatrix} w & 1-w \\ C^2 & C^3 \end{pmatrix} \quad \textit{if } 0<w\leq 1.$$

(3) Archimedes Axiom: *Consider three deterministic consumption paths $c^1$, $c^2$, and $c^3$ with $c^1 \prec c^2 \prec c^3$. For these paths there is one, and only*

[16] For the one-period case, such a function was postulated by KRELLE (1968, pp. 144–147) in order to adapt a preference structure over non-random outcomes to the case of uncertainty. Related ideas seem to underlie the approaches by DIAMOND/STIGLITZ (1974) and KIHLSTROM/MIRMAN (1974).

*one, probability $w$, $0<w<1$, such that*

$$c^2 \sim \begin{pmatrix} w & 1-w \\ c^1 & c^3 \end{pmatrix}.$$

(4) Koopmans Axiom: *The decision maker's preference functional for multiperiod planning problems under certainty is*

$$\Sigma = \sum_{t=0}^{T-1} \lambda_t u(c_t) + \lambda_T u(v_T).$$

Axioms (1), (2), and (3) are very similar to the corresponding axioms for the one-period case and hence do not need any further explanation. Axiom (4) saves us from searching for an appropriate preference functional for planning problems under certainty. For a possible axiom system producing the non-stochastic preference functional $\Sigma$ the reader is referred to KOOPMANS (1960). In the following, the proof showing why axioms (1)–(4) imply the preference functional (26) for the stochastic case is sketched.

As we know, the additive preference functional required by axiom (4) can partly be altered without affecting its behavioral implications in a world of certainty. Suppose, however, that $\Sigma$ has been given a special functional form which is compatible with the decision maker's preference over non-random consumption paths, so that there is a unique mapping from the set of these paths to the set of real numbers.

Consider now the Axiom of Strong Independence. This axiom includes the limiting case where $C^3$ is a degenerated 'probability distribution' which is a particular non-random consumption path. It therefore says that, in a probability distribution of consumption paths, those paths with the same $\Sigma$ can be interchanged without altering the evaluation of the whole distribution [17]. This is an important implication, for it

[17] A criticism against this property of the Axiom of Strong Independence was raised by DIAMOND (1967) in a comment on HARSANYI's (1955) famous derivation of an additive social welfare function based on the von Neumann–Morgenstern axioms. Diamond referred to consumption paths over different generations rather than over different periods in an individual life span. Consider two non-random consumption paths that are equivalent from a welfare point of view but which favor different generations. Is it permissible to interchange these paths in a probability distribution of paths, without asking which generations are favored by the other variates in the probability distribution? Diamond answered this question in the negative, arguing that there should be an element of justice in the social welfare function, in the sense that chances should be evenly distributed among the generations. This criticism clearly suggests that we must be careful when using the preference functional (26) for intergenerational welfare comparisons. In the present case, however, the criticism does not apply. There is no point in postulating 'justice' between the years in a single person's life span.

means that, in evaluating probability distributions of consumption paths, the decision maker is only interested in the corresponding probability distributions of $\Sigma$. The Axiom of Independence therefore allows the multidimensional risk problem to be reduced to a one-dimensional problem.

The reduction in dimension permits axioms (1)–(3) to be transformed in a way that elucidates the similarities between the one-period and the multiperiod choice problems under uncertainty. Axiom (1) implies that the decision maker has a complete preference ordering over probability distributions of $\Sigma$. Axiom (2) can be transformed into an analogous postulate referring to three distributions of $\Sigma$ rather than of consumption paths $C$. And axiom (3) can be reinterpreted in such a way as to require the existence and uniqueness of a probability in the open unit interval that renders the decision maker indifferent between a binary distribution of $\Sigma$ and a non-random value of $\Sigma$ which is placed strictly between the variates of this binary distribution. In short: the reduction to one dimension implies that axioms (1)–(3) bring about postulates that are identical to the corresponding one-period axioms of chapter II C 2.1, except that they refer to values of $\Sigma$ rather than to wealth.

Since the Non-Saturation Axiom assumed in the one-period analysis is satisfied for $\Sigma$ by its very definition, we may now combine the postulates just derived in the same way as shown in chapter II C 2.2 for the one-period von Neumann–Morgenstern axioms. If account is taken of axiom (4) the result is the preference functional (26) for an evaluation of stochastic consumption paths.

### 2.1.2. A Preference Functional According to Fechner's Law

In the previous section rational behavior for multiperiod planning under uncertainty was studied. Analogously to the one-period case, the task is now to put some life into the preference functional (26) by adding a special hypothesis on the preference structure of man. For this hypothesis we again make use of Weber's law, but, in addition, we also employ Fechner's law that, up to now, has only served as a means of interpretation.

Weber's law can be used by replacing end-of-period wealth from the previous analysis with a factor $x>0$ which measures the level of a consumption path $(xc_0^*, \ldots, xc_{T-1}^*, xv_T^*)$ that arose from a time-profile preserving multiplication with $x$ of an arbitrarily chosen basic path $(c_0^*, \ldots, c_{T-1}^*, v_T^*)$. This procedure is fully compatible with the definition of wealth given at the beginning of chapter III. According to the Weak

Relativity Axiom it is then required that

$$(27)\qquad \Psi[\sum_{t=0}^{T-1}\lambda_t u(xc_t^*)+\lambda_T u(xv_T^*)]$$

$$=\left\{\begin{array}{ll} x^{1-\varepsilon'}\Psi[\sum_{t=0}^{T-1}\lambda_t u(c_t^*)+\lambda_T u(v_T^*)], & \varepsilon'\neq 1 \\ \ln x+\Psi[\sum_{t=0}^{T-1}\lambda_t u(c_t^*)+\lambda_T u(v_T^*)], & \varepsilon'=1 \end{array}\right\}.$$

Here, $\varepsilon'$ is the measure of relative risk aversion for the evaluation of a probability distribution $X$ with variates $x$. We shall see later how $\varepsilon'$ is related to the risk aversion parameter $\varepsilon$ characterizing the current choice among end-of-period wealth distributions.

Unfortunately the information contained in (27) is not sufficient to determine $\Psi(.)$ *and* $u(.)$ since it merely indicates the over-all concavity of the preference functional, that is, a sum effect of both functions. Thus, additional information on $\Psi(.)$ *or* $u(.)$ is needed. For $u(.)$, it seems that such information can be obtained from the psychophysical relativity laws discussed in chapter III.

Suppose we carry out a number-matching experiment for the consumption level of a single period, keeping all other consumption levels constant. Suppose, further, that people evaluate this level of consumption independently of the levels in other periods, as was suggested by the intertemporal separability of the additive preference functional (20). Then it seems likely that a strict proportionality between the money value of period consumption and the numbers that people choose to signify its magnitude can be found. Assume this in the case or assume there is, in general, at least a relationship between the number and the consumption continua that can be described by a power function. Then period consumption belongs to the large family of Stevens's continua that include stimuli like loudness, brightness, length, weight, area, and numbers and that have been shown, in hundreds of cross-modality experiments, to be related to one another by power functions. Now, the interval experiments reported in chapter III A 1.3.4 clearly indicated that, for *all* Stevens's continua, equal relative changes in the objective intensity of a stimulus bring about equal absolute changes in the subjective intensity of its sensation or, in other words, that for *all* continua there are logarithmic sensation functions. This suggests that, if we wish to formulate a simple hypothesis regarding the shape of the period utility function, we should assume that this function is logarithmic. The hypothesis is taken up in the following axiom.

Strong Relatively Axiom: *Equal relative changes in period consumption bring about equal absolute changes in period utility.*

The preference functional for deterministic planning problems therefore is

$$\Sigma = \sum_{t=0}^{T-1} \lambda_t (a + b \ln c_t) + \lambda_T (a + b \ln v_T)$$

or, with some normalizations that can be carried out without any loss of generality[18],

$$(28) \qquad \Sigma = \sum_{t=0}^{T-1} \lambda_t \ln c_t + \lambda_T \ln v_T, \quad \sum_{t=0}^{T-1} \lambda_t = 1.$$

It is confirming to note that MODIGLIANI and BRUMBERG (1955, p. 396, fn. 15) considered a preference function of this type, referring explicitly to the results of psychophysics[19].

It is not difficult to find the specific risk preference function $\Psi(.)$ compatible with the two pieces of information given in (27) and (28). Obviously we have[20]

$$(29) \qquad \Psi(\Sigma) = \begin{cases} (1-\varepsilon')e^{(1-\varepsilon')\Sigma}, & \varepsilon' \neq 1, \\ \Sigma, & \varepsilon' = 1. \end{cases}$$

Accordingly, the possible versions of the multiperiod preference functional $R(C_1, \dots, C_{T-1}, V_T)$, that we have been looking for, are

$$(30) \qquad R(.) = E\left\{(1-\varepsilon')e^{(1-\varepsilon')[\Sigma_{t=0}^{T-1} \lambda_t \ln C_t + \lambda_T \ln V_T]}\right\}$$
$$= \mathrm{E}\left\{(1-\varepsilon') \prod_{t=0}^{T-1} C_t^{(1-\varepsilon')\lambda_t} V_T^{(1-\varepsilon')\lambda_T}\right\}, \quad \text{if} \quad \varepsilon' \neq 1,$$

$$(31) \qquad R(.) = E\left(\sum_{t=0}^{T-1} \lambda_t \ln C_t + \lambda_T \ln V_T\right), \quad \text{if} \quad \varepsilon' = 1.$$

The result shows that, for decision problems under uncertainty, the additive preference functional is only maintained in the case $\varepsilon' = 1$. In all

[18] The normalizations do not affect the classes of possible preference functionals given by equations (30) and (31).

[19] The formulation also corresponds to HELSON's (1947 and 1964) formula for the adaptation level. See equation (III A 29).

[20] Cf. equations (III A 38) and (III A 39).

other cases there is instead a multiplicative preference functional, for $\varepsilon < 1$ in a Cobb-Douglas version. Some of the implications of such a type of preference functional have been drawn out by PYE (1972) for the special case $\lambda_t = \lambda^t$, $t = 0, \ldots, T$, $0 < \lambda < 1$. Pye, however, did not try to give a legitimation of the preference functional, let alone a legitimation that resembles the one provided here. He merely made a favorable assessment of its behavioral implications. We shall see what these implications are, although without confining our attention to the case $\lambda_t = \lambda^t \forall t$.

## 2.2. *The Recursive Solution*

Utilizing the preference functionals (30) and (31) and assuming stochastic constant returns to scale, including auxiliary conditions, we are now going to solve the problem

$$(32) \qquad \max R(C_0, \ldots, C_{T-1}, V_T)\,|_{v_0}.$$

At each point in time $t$, the decision maker has control over the standard risk project $Q_{t+1}$, whose outcome is revealed one period later, and over the consumption-wealth ratio $\alpha_t \equiv c_t/v_t$. The variable $v_t$ denotes wealth available at point in time $t$, before consumption for the subsequent period is subtracted. Wealth available for investment in the standard risk project is $a_t = v_t(1 - \alpha_t)$. As usual, we write random variables with capital letters and denote non-random variables, i.e., those variables known by the decision maker, by lower-case letters[21].

If we again formulate a minimization problem in the case $\varepsilon' > 1$ then, according to Bellman's Principle of Optimality[22], at each point in time $t^*$, $0 \le t^* \le T-1$, the problem

$$(33) \qquad \frac{\begin{Bmatrix}\max\\ \min\end{Bmatrix} E\Big[\prod_{t=t^*}^{T-1} C_t^{(1-\varepsilon')\lambda_t} V_T^{(1-\varepsilon')\lambda_T}\Big]\Big|_{v_{t^*}}, \quad \text{if} \quad \begin{Bmatrix}\varepsilon' < 1\\ \varepsilon' > 1\end{Bmatrix},}{\max \; E\Big[\sum_{t=t^*}^{T-1} \lambda_t \ln C_t + \lambda_T \ln V_T\Big]\Big|_{v_{t^*}}, \quad \text{if} \quad \varepsilon' = 1,}$$

has to be solved. Expression (33) characterizes the optimal consumption-investment policy as seen from the initial point in time 0 in such a way that, at every subsequent point in time, on the basis of the currently available wealth a policy is chosen which is optimal at that point in time. This rule will allow the properties of the optimal plan to be discovered.

[21] Indices and parameters are excepted.

[22] Cf. expression (12) above.

*Point in Time $t^* = T-1$*

First, consider the last situation that obtains before the end of the planning period. With a historically given level of wealth $v_t$, whatever this level happens to be, the problem is to satisfy the following postulate[23] by a suitable simultaneous choice of the consumption-wealth ratio $\alpha_{T-1}$ and the standard risk project $Q_T$:

$$(34)\qquad \frac{\begin{Bmatrix}\max\\ \min\end{Bmatrix} E(c_{T-1}^{(1-\varepsilon')\lambda_{T-1}} V_T^{(1-\varepsilon')\lambda_T})\,|_{v_{T-1}} \equiv z_{T-1}.}{\max\ E(\lambda_{T-1}\ln c_{T-1} + \lambda_T \ln V_T)\,|_{v_{T-1}} \equiv z_{T-1}.}$$

Because of $V_T = (1-\alpha_{T-1})Q_T v_{T-1}$ and $c_{T-1} = \alpha_{T-1} v_{T-1}$, this postulate can be converted to the following expression, which can immediately be simplified by transferring constants forward:

$$(35)\qquad \frac{\begin{aligned} z_{T-1} &= \begin{Bmatrix}\max\\ \min\end{Bmatrix} E[(\alpha_{T-1}v_{T-1})^{(1-\varepsilon')\lambda_{T-1}} ((1-\alpha_{T-1})Q_T v_{T-1})^{(1-\varepsilon')\lambda_T}] \\ &= v_{T-1}^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)} \begin{Bmatrix}\max\\ \min\end{Bmatrix} [\alpha_{T-1}^{(1-\varepsilon')\lambda_{T-1}} (1-\alpha_{T-1})^{(1-\varepsilon')\lambda_T} E(Q_T^{(1-\varepsilon')\lambda_T})]. \end{aligned}}{\begin{aligned} z_{T-1} &= \max E[\lambda_{T-1}\ln(\alpha_{T-1}v_{T-1}) + \lambda_T \ln((1-\alpha_{T-1})v_{T-1}Q_T)] \\ &= (\lambda_{T-1}+\lambda_T)\ln v_{T-1} \\ &\quad + \max[\lambda_{T-1}\ln\alpha_{T-1} + \lambda_T\ln(1-\alpha_{T-1}) + \lambda_T E(\ln Q_T)]. \end{aligned}}$$

It is worth noting that here, as in the pure accumulation approach, the optimization problem can be solved independently of the size of the available wealth $v_{T-1}$ and hence independently of the previously realized outcome $q_{T-1}$ of the standard risk project $Q_{T-1}$. The reason is that, behind the max/min operators in the second lines of each of the equations, expressions with $v_{T-1}$ do not appear and the opportunity set for $Q_T$, by assumption, is not disturbed by autocorrelation. We do not attempt to find the solution for an optimal policy at this stage. At the moment, it is sufficient to state that, independently of the particular variate $v_{T-1}$, there are well-determined optimal values for $\alpha_{T-1}$ and $E(Q_T^{(1-\varepsilon')\lambda_T})$ or $E(\ln Q_T)$, respectively.

---

[23] As in the pure accumulation approach, the formula above the horizontal line refers to $\varepsilon' \neq 1$, the one below to $\varepsilon' = 1$.

*Point in Time $t^* = T-2$*

The decision maker's task now is

$$\begin{Bmatrix}\max\\ \min\end{Bmatrix}\{E(c_{T-2}^{(1-\varepsilon')\lambda_{T-2}} C_{T-1}^{(1-\varepsilon')\lambda_{T-1}} V_T^{(1-\varepsilon')\lambda_T}) \mid_{v_{T-2}}$$
$$= E[(v_{T-2}\alpha_{T-2})^{(1-\varepsilon')\lambda_{T-2}}$$
$$(v_{T-2}Q_{T-1}(1-\alpha_{T-2})\alpha_{T-1})^{(1-\varepsilon')\lambda_{T-1}}$$
$$(v_{T-2}Q_{T-1}Q_T(1-\alpha_{T-2})(1-\alpha_{T-1}))^{(1-\varepsilon')\lambda_T}]\}$$
$$\equiv z_{T-2}.$$

(36)

$$\max\{E(\lambda_{T-2}\ln c_{T-2} + \lambda_{T-1}\ln C_{T-1} + \lambda_T \ln V_T) \mid_{v_{T-2}}$$
$$= E[\lambda_{T-2}\ln(v_{T-2}\alpha_{T-2})$$
$$+ \lambda_{T-1}\ln(v_{T-2}Q_{T-1}(1-\alpha_{T-2})\alpha_{T-1})$$
$$+ \lambda_T \ln(v_{T-2}Q_{T-1}Q_T(1-\alpha_{T-2})(1-\alpha_{T-1}))]\}$$
$$\equiv z_{T-2}.$$

Transferring constants to the front and putting equal factors together, we have:

$$z_{T-2} = v_{T-2}^{(1-\varepsilon')(\lambda_{T-2}+\lambda_{T-1}+\lambda_T)}$$
$$\begin{Bmatrix}\max\\ \min\end{Bmatrix}[\alpha_{T-1}^{(1-\varepsilon')\lambda_{T-1}}(1-\alpha_{T-1})^{(1-\varepsilon')\lambda_T}$$
$$\alpha_{T-2}^{(1-\varepsilon')\lambda_{T-2}}(1-\alpha_{T-2})^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)}$$
$$E(Q_{T-1}^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)} Q_T^{(1-\varepsilon')\lambda_T})].$$

(37)

$$z_{T-2} = (\lambda_{T-2}+\lambda_{T-1}+\lambda_T)\ln v_{T-2}$$
$$+ \max[\lambda_{T-1}\ln\alpha_{T-1} + \lambda_T\ln(1-\alpha_{T-1})$$
$$+ \lambda_{T-2}\ln\alpha_{T-2} + (\lambda_{T-1}+\lambda_T)\ln(1-\alpha_{T-2})$$
$$+ E((\lambda_{T-1}+\lambda_T)\ln Q_{T-1} + \lambda_T \ln Q_T)].$$

Because of the assumption that $Q_{T-1}$ and $Q_T$ as well as any given functions of these random variables, are stochastically independent, (37) can also be expressed in the following way[24]:

$$z_{T-2} = v_{T-2}^{(1-\varepsilon')(\lambda_{T-2}+\lambda_{T-1}+\lambda_T)}$$
$$\begin{Bmatrix}\max\\ \min\end{Bmatrix}[\alpha_{T-1}^{(1-\varepsilon')\lambda_{T-1}}(1-\alpha_{T-1})^{(1-\varepsilon')\lambda_T}$$
$$E(Q_T^{(1-\varepsilon')\lambda_T})$$
$$\alpha_{T-2}^{(1-\varepsilon')\lambda_{T-2}}(1-\alpha_{T-2})^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)}$$
$$E(Q_{T-1}^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)})].$$

(38)

$$z_{T-2} = (\lambda_{T-2}+\lambda_{T-1}+\lambda_T)\ln v_{T-2}$$
$$\max[\lambda_{T-1}\ln\alpha_{T-1} + \lambda_T\ln(1-\alpha_{T-1})$$
$$+ \lambda_T E(\ln Q_T)$$
$$+ \lambda_{T-2}\ln\alpha_{T-2} + (\lambda_{T-1}+\lambda_T)\ln(1-\alpha_{T-2})$$
$$+ (\lambda_{T-1}+\lambda_T)E(\ln Q_{T-1})].$$

---

[24] Cf. footnote 4 in section A.

(Note that, in the case $\varepsilon'=1$, the step from (37) to (38) is possible even for stochastically correlated risks.)

The expressions in the second and third lines of (38) are known from the expressions following the max/min operators in (35). According to the Bellmann Principle of Optimality, we may therefore transform (38) into

$$
(39)\qquad
\begin{aligned}
z_{T-2} &= \upsilon_{T-2}^{(1-\varepsilon')(\lambda_{T-2}+\lambda_{T-1}+\lambda_T)} \\
&\quad \left\{\begin{matrix}\max\\ \min\end{matrix}\right\} [\alpha_{T-1}^{(1-\varepsilon')\lambda_{T-1}}(1-\alpha_{T-1})^{(1-\varepsilon')\lambda_T} E(Q_T^{(1-\varepsilon')\lambda_T})] \\
&\quad \left\{\begin{matrix}\max\\ \min\end{matrix}\right\} [\alpha_{T-2}^{(1-\varepsilon')\lambda_{T-2}}(1-\alpha_{T-2})^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)} \\
&\qquad E(Q_T^{(1-\varepsilon')(\lambda_{T-1}+\lambda_T)})].
\end{aligned}
$$

$$
\begin{aligned}
z_{T-2} &= (\lambda_{T-2}+\lambda_{T-1}+\lambda_T)\ln \upsilon_{T-2} \\
&\quad + \max[\lambda_{T-1}\ln\alpha_{T-1}+\lambda_T\ln(1-\alpha_{T-1})+\lambda_T E(\ln Q_T)] \\
&\quad + \max[\lambda_{T-2}\ln\alpha_{T-2}+(\lambda_{T-1}+\lambda_T)\ln(1-\alpha_{T-2}) \\
&\qquad + (\lambda_{T-1}+\lambda_T)E(\ln Q_{T-1})].
\end{aligned}
$$

*Point in Time $t^*=T-\tau$*

Proceeding in the manner described, we reach the conclusion that the maximized/minimized goal functions generally assume the following form

$$
(40)\qquad
\begin{aligned}
z_{T-\tau} &= \upsilon_{T-\tau}^{(1-\varepsilon')\sum_{i=T-\tau}^{T}\lambda_i} \prod_{t=T-\tau}^{T-1} \left\{\begin{matrix}\max\\ \min\end{matrix}\right\} [\alpha_t^{(1-\varepsilon')\lambda_t}(1-\alpha_t)^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i} \\
&\qquad E(Q_{t+1}^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i})].
\end{aligned}
$$

$$
\begin{aligned}
z_{T-\tau} &= \sum_{i=T-\tau}^{T}\lambda_i \ln \upsilon_{T-\tau} + \sum_{t=T-\tau}^{T-1}\max[\lambda_t\ln\alpha_t + \sum_{i=t+1}^{T}\lambda_i\ln(1-\alpha_t) \\
&\qquad + \sum_{i=t+1}^{T}\lambda_i E(\ln Q_{t+1})].
\end{aligned}
$$

Obviously, within (40), the expressions in the second lines have to be maximized/minimized, whatever value the consumption-wealth ratio $\alpha_t$ takes on. Thus we have

$$
(41)\qquad
\begin{aligned}
z_{T-\tau} &= \upsilon_{T-\tau}^{(1-\varepsilon')\sum_{i=T-\tau}^{T}\lambda_i} \prod_{t=T-\tau}^{T-1} \left\{\begin{matrix}\max\\ \min\end{matrix}\right\} [\alpha_t^{(1-\varepsilon')\lambda_t}(1-\alpha_t)^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i} \\
&\qquad \left\{\begin{matrix}\max\\ \min\end{matrix}\right\} E(Q_{t+1}^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i})].
\end{aligned}
$$

$$
\begin{aligned}
z_{T-\tau} &= \sum_{i=T-\tau}^{T}\lambda_i \ln \upsilon_{T-\tau} + \sum_{t=T-\tau}^{T-1}\max[\lambda_t\ln\alpha_t + \sum_{i=t+1}^{T}\lambda_i\ln(1-\alpha_t) \\
&\qquad \sum_{i=t+1}^{T}\lambda_i \max E(\ln Q_{t+1})].
\end{aligned}
$$

The last step is to put the expressions in the second lines in front of the first max/min operators, since these expressions are constants with respect to the optimization of the consumption-wealth ratio:

$$z_{T-\tau}=v_{T-\tau}^{(1-\varepsilon')\sum_{t=T-\tau}^{T}\lambda_t}\prod_{t=T-\tau}^{T-1}\Big[\begin{Bmatrix}\max\\ \min\end{Bmatrix}E\big(Q_{t+1}^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i}\big)$$

$$\begin{Bmatrix}\max\\ \min\end{Bmatrix}\big(\alpha_t^{(1-\varepsilon')\lambda_t}(1-\alpha_t)^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i}\big)\Big].$$

(42)

$$z_{T-1}=\sum_{t=T-\tau}^{T}\lambda_t\ln v_{T-\tau}+\sum_{t=T-\tau}^{T-1}\Big[\sum_{i=t+1}^{T}\lambda_i\max E(\ln Q_{t+1})$$

$$+\max\big(\lambda_t\ln\alpha_t+\sum_{i=t+1}^{T}\lambda_i\ln(1-\alpha_t)\big)\Big].$$

In what follows, these formulas will be interpreted with respect to their implications for the optimal behavioral strategy under multiperiod uncertainty.

### 2.3. *Interpretation of the Solution*

#### 2.3.1. The Rehabilitation of the One-Period Approach

For the optimal decision at point in time $t$, where $t=T-\tau=0$, and hence for the optimal decision at the beginning of the planning problem also, (42) brings about two significant results. The first is that the optimal standard risk project has to be chosen according to one of the following rules:

$$(43)\qquad \max\begin{Bmatrix}(1-\varepsilon')E\big(Q_{t+1}^{(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i}\big)\\ E(\ln \mathrm{Q}_{t+1})\end{Bmatrix},\quad \text{if}\quad \begin{Bmatrix}0<\varepsilon'\neq 1\\ \varepsilon'=1\end{Bmatrix}.$$

The second consists of a formula for the optimal consumption-wealth ratio $\alpha_t^*$, that is, for the decision maker's propensity to consume out of wealth:

$$(44)\qquad \alpha_t^*=\frac{1}{1+\frac{\sum_{i=t+1}^{T}\lambda_i}{\lambda_t}},\quad \varepsilon'>0.$$

This formula follows by differentiating (42) with respect to $\alpha$ and setting the derivative equal to zero.

In a way similar to the pure accumulation approach, (43) confirms the role of the Weber functions (III A 34) for a separate, isolated evaluation of the standard risk projects of each period. In the case of a power function, the exponent now takes on a time-dependent value but what matters is that it is constant at a given point in time. A new element is the propensity to consume out of wealth. With (44) we have a formula which shows that this propensity depends only on the time path of the weight factor $\lambda_t$; the propensity to consume is therefore determined by the decision maker's time preference, but not by his risk preference.

The results (43) and (44) give very simple behavioral rules for multiperiod planning. The most important aspect seems to be that the decision maker does not have to take into account the future investment opportunities either in choosing the current rate of consumption or in determining the optimal current risk project. Thus the level of information that is necessary for solving the intertemporal optimization problem is surprisingly low. All that the decision maker needs to know about the future are his own preferences. Anyone who up to now had the impression that the intertemporal optimization approach overburdened the decision maker, will be relieved to find that this is not so.

Another, more subtle, aspect is that the current decisions on the optimal rate of consumption and the optimal risk project are separable from each other. The decision maker can therefore proceed in two steps. First, independently of the available set of standard risk projects, he determines his optimal level of consumption and hence the level of wealth $a_t = v_t - c_t$ available for investment. Then, given the optimal value of $a_t$ and the corresponding set of end-of-period wealth distributions $V = a_t Q_{t+1}$, a choice is made according to the rule $\max E[U(V)]$ where $U(.)$ is one of the Weber functions.

*A priori*, the separability of the two choice problems could not have been expected. On the contrary, it would have seemed plausible for the profitability of the best available project to influence the amount of wealth $a_t$ maintained for reinvestment. The reason this conjecture is wrong is provided by the Strong Relativity Axiom. This axiom requires a logarithmic period utility function which, as is easily understandable, implies that the elasticity of substitution between the consumption levels at two points in time is unity. For such a value, the income and substitution effects of a change in a particular variate $q$ of the standard risk project $Q$ just offset each other. Thus the decision maker does not care whether he knows this change or the change in a whole probability distribution of such variates. In any case, he chooses the same level of consumption[25,26].

[25] MOSSIN (1969 and 1973, pp. 29–32) demonstrates that generally the choices of consumption and risk are intertwined in a difficult manner. Our result shows that we do not always have to 'cross our fingers', as he suggests, and proceed by neglecting the consump- →

The behavioral rules developed from the multiperiod approach imply a very surprising rehabilitation of the simple decision theoretic approach used in the previous chapters. It now turns out that we were indeed justified in abstracting from both the consumption decision and the time aspect of the decision problem[27].

The rehabilitation, however, goes even further. For the reason that was discussed above in connection with the pure accumulation approach, the BLOOS rule again enables us to evaluate even gross distributions that partly extend over the negative half of the wealth axis. The corresponding derived preference structure is the same as in the one-period case. Thus, for example, the kinked utility curve and the corresponding indifference-curve system in the $(\mu, \sigma)$ diagram that were derived in chapter III maintain their validity.

The rehabilitation of the one-period approach is certainly the most important outcome of the multiperiod approach, but it is not the only one. In the next two sections it will be shown how the passage of time affects the optimal decisions. This information, of course, could not be revealed by the one-period approach.

### 2.3.2. Time and Risk Aversion

The crucial change that (42) brings about, compared to the one-

tion decision. It is true that, despite Fechner's law, the terrors Mossin points out will persist in the case of pure 'income risks' that do not depend on wealth. In the present specification, however, where income is generated through wealth, they vanish. Let, in line with Mossin's 1973 terminology, $Y_2=(Y_1-c_1)(Q-1)$ denote the stochastic income at point in time 2 where $Y_1$ is the initial wealth, $c_1$ consumption in the first period, and $Q$ the standard risk project. Then Mossin's formula $\max E[u(c_1, Y-c_1)]$, where $Y \equiv Y_1+Y_2$ and $u(.)$ is a two-dimensional von Neumann-Morgenstern function, becomes $\max E\{u[c_1, Y_1+(Y_1-c_1)(Q-1)-c_1]\} = \max E\{u[c_1,(Y_1-c_1)Q]\}$. Concerning this expression, the analysis in the text gives the result

$$\max E\{u[(c_1,(Y_1-c_1)Q]\} \sim \max E\{U[(Y_1-c_1^*)Q]\}$$

where $c_1^*$ is the optimal level of consumption at point in time 1 that can be determined independently of the risk problem. By using a utility function $v(.)$ (Mossin's terminology) defined such that $v(x+c_1) \equiv U(x)$ the optimization problem therefore can also be expressed as

$$\max E\{v[(Y_1-c_1)Q+c_1^*]\} = \max E[v(Y)]$$

which is what Mossin would have liked to find. Cf. also SPENCE and ZECKHAUSER (1972).

[26] The result does not necessarily imply that consumption does not depend on a change in the non-random market rate of interest. In fact, METZLER's (1951) wealth effect is clearly present. For example, a rise in the non-random market rate of interest lowers the value of the decision maker's human capital as well as the capitalized value of other income streams and hence will reduce current consumption. This implication fits recent empirical results on the interest dependence of consumption. See BOSKIN (1978).

[27] It should be noted that this statement is contingent upon the assumption of stochastic constant returns to scale.

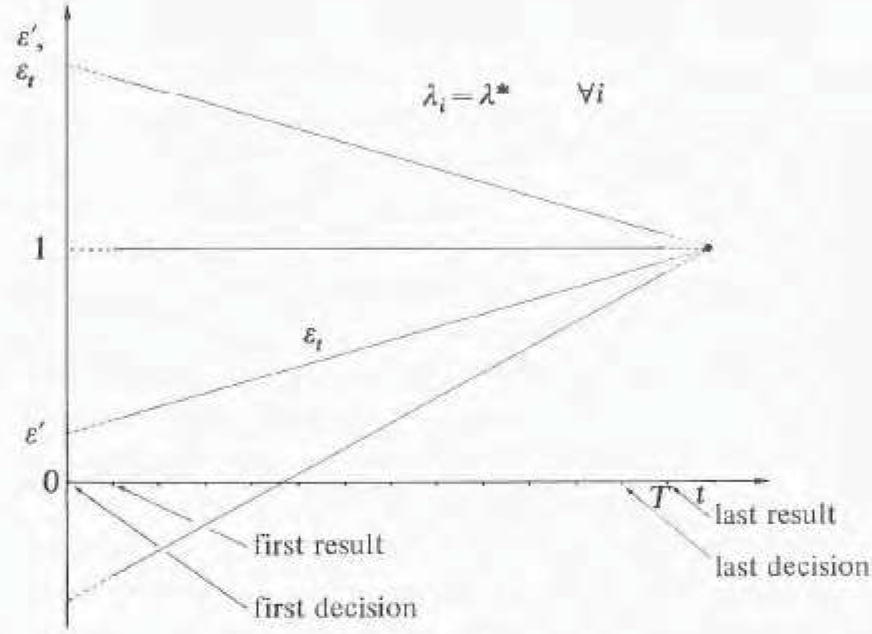


*Figure* 3

period approach, is the time dependence of risk aversion. Let $\varepsilon_{t+1}$ denote the measure of relative risk aversion[28] relevant for choosing at point in time $t$ among risk projects $Q_{t+1}$. Then (43) implies for both the cases $\varepsilon'=1$ and $\varepsilon'\neq 1$:

$$(45) \qquad \varepsilon_{t+1}=1-(1-\varepsilon')\sum_{i=t+1}^{T}\lambda_i.$$

The sum on the right-hand side of this equation diminishes over time as long as $\lambda_i>0$. Thus, with the passage of time, the current level of risk aversion approaches, either from above or below, the value of unity or maintains this value if it initially prevailed[29]. Figure 3 illustrates the possible time paths of risk aversion for the example $\lambda_i=\lambda^*=\text{const.}$ $\forall i\leq T$. Because of $\sum_{i=0}^{T}\lambda_i=1$, in this example we have $\lambda^*=1/(T+1)$, and (45) becomes

$$(46) \qquad \varepsilon_t=1-(1-\varepsilon')\frac{T-t+1}{T+1},$$

an expression that indicates a linear time path for the risk-aversion measure $\varepsilon$.

[28] See equation (III A 33).

[29] It is even possible that initial risk loving ($\varepsilon<0$) changes to risk aversion with the passage of time.

The paths in Figure 3 between points in time 0 and 1 are broken to indicate that point 1 is the first point in time where the outcome of a probability distribution is revealed. Despite this, it may be asked which degree of risk aversion would be appropriate if, at point in time zero, a choice between risk projects with immediate pay-offs were offered. Noting that $\sum_{i=0}^{T}\lambda_i = 1$, we find from (45) that $\varepsilon_0 = \varepsilon'$. This brings us back to the original interpretation of $\varepsilon'$ in (27). There, $\varepsilon'$ was assumed to indicate the evaluation of standardized consumption paths that differ only with respect to their levels $x$. Here, we find that obviously the initial level of wealth can be identified with the parameter $x$. This is plausible, but not self-evident, for, while we assumed in (27) that doubling $x$ implies a doubling of consumption at each point in time, nothing similar to this was assumed in the solution of the multiperiod optimization problem. Instead, the fact that in (44) the optimal consumption-wealth ratios are independent of wealth legitimates the previous assumption. Whatever the random variates of the standard risk projects, if the initial level of wealth is doubled, consumption at each point in time is twice as high as it otherwise would have been.

Figure 3 suggests that a path that deviates anywhere from the value of unity does so everywhere. This is a general rule to which there are no exceptions. Since $\varepsilon_t$ falls if it is larger than one and rises if it is smaller then one, it is possible to draw conclusions concerning the *size* of risk aversion from its *change*, this being easier to see than its size. With the model of optimal life-cycle planning in mind, the case of risk aversion increasing with age seems particularly realistic. Hence we must conclude that the Weber function $U(v) = v^{1-\varepsilon}$, $\varepsilon < 1$, defines the *standard type*.

This result restates a conclusion that was reached on different grounds in chapter III B. It was shown there that the neglect of large liability risks, which can often be observed in reality, can be explained by the BLOOS rule if, and only if, the utility function for net wealth is bounded from below. The latter implies that only the Weber function with $\varepsilon < 1$ is relevant.

For the sake of illustration, Figure 3 refers to linear time paths of the risk aversion measure $\varepsilon$. The property of linearity is not a general one but will only prevail if there is no time preference in the sense of von Böhm-Bawerk's second reason. How the appearance of time preference affects the time paths of $\varepsilon$ can easily be shown. Generally, an increase in time preference occurs if, for all $t = 1, \ldots, T$, the ratio $(\sum_{i=t}^{T}\lambda_i)/(\sum_{i=0}^{T}\lambda_i)$ is getting smaller which, because of $\sum_{i=0}^{T}\lambda_i = 1$, means that $\sum_{i=t}^{T}\lambda_i = 1$ is falling. In connection with (45), this implies that an increase in the rate of time preference makes $\varepsilon_t = 1 - (1 - \varepsilon')\sum_{i=t}^{T}\lambda_i$ approach the value of unity for all $t \geq 1$ if, with $\varepsilon_0 = \varepsilon'$, the origin of the path is maintained. Thus the following conclusions appear. Of two people, who have the same degree of risk aversion when they are young, the one with the higher rate of time preference will be nearer to the intermediate value $\varepsilon = 1$ when he is old and of two people, who exhibit the same degree of risk aversion in old age, the one with the higher rate of time preference had the greater tendency toward the extremes when he was young.

The reader may be tempted to doubt the time dependence of risk aversion when thinking of the phenomenon of 'rolling planning' observable in reality. Indeed, if, with the passage of time, the planning horizon is shifted ahead and, period by period, a new optimization problem relating to the corresponding new horizon is solved then risk aversion will be time-invariant[30]. In general, however, such a procedure is irrational, for it means optimizing the current decision on the assumption of a future behavioral strategy that, when the future arrives, is not actually carried out. A shift in the planning horizon is only innocuous if, for whatever reason, it does not affect the actual decision. Presumably this is the case when rolling planning is used in practice, for otherwise it would be hard to understand why a firm chooses, for example, to make rolling five-year plans if rolling six-year plans imply significant differences in current behavior.

Compared to other models of intertemporal planning under uncertainty, the present approach incorporates an unusual influence of time and time preference on risk aversion. This influence certainly deserves some explanation.

• At the beginning of his related article, SAMUELSON (1969) conjectured that risk aversion for the young might be smaller than for the old, since the young man may be able to recoup his losses. He was surprised when he found that, on the contrary, his own approach implied a time-invariance of risk aversion. At first glance, it is tempting to interpret our findings as confirming Samuelson's conjecture. In fact, however, this would not provide us with the true explanation of the time dependence of risk aversion. Contrary to the idea behind Samuelson's conjecture, decisions in youth do, in fact, have enormous implications for the future for they 'switch the points' and so influence the direction of later life. If, through a change in an initial decision, wealth changes by $x\%$ then, because of the ratio structure of preferences in connection with stochastic constant returns to scale, consumption in each future period also changes by $x\%$. There is no change of recouping. The true explanation for the time dependence of risk aversion is different.

The preference functional derived above has the property that a specific risk preference function $\Psi(.)$ is applied in order to evaluate dis-

[30] For arguments in favor of rolling plans see ROSENSTEIN-RODAN (1934, pp. 78–84).

persions of life-time utility $\Sigma$. This seems to be the clue to understanding what is going on in the model. A young man's decision alters consumption in a large number of future years and hence may significantly affect the dispersion of $\Sigma$. Risk aversion or risk loving on the subjective continuum ($\Sigma$), i.e., the concavity or convexity of $\Psi(.)$ will have large effects on a young man's evaluation of probability distributions. A young man's relative risk aversion parameter $\varepsilon$ may hence substantially deviate from the value of unity that would prevail if $\Psi(.)$ were linear. An old man's decision, on the other hand, can only affect comparatively fewer years and thus will not alter the dispersion of $\Sigma$ to any great extent. The curvature of $\Psi(.)$ is not very important in his decision making; the old man may decide roughly according to the mean-value criterion on the subjective continuum. This mean-value criterion is $\max E(\ln Q_t)$ and explains why the old man tends towards a risk aversion of $\varepsilon = 1$.

In the light of this interpretation, the role of time preference described above also becomes clear. A higher rate of time preference means that the weight of the later years is reduced in comparison to the earlier years. Hence, for an even stronger reason, the old man's decisions bring about comparatively smaller changes in dispersions of $\Sigma$ than the young man's decisions do. The obvious consequence is that the time change in the degree of relative risk aversion is reinforced.

### 2.3.3. The Optimal Consumption Strategy

We now study the properties of the optimal consumption strategy (or in the model of the firm: the optimal dividend policy) as implied by (44). It is immediately apparent from this equation that the propensity to consume out of wealth at the decision point $t$ is lower the more weight is attached to future consumption ($\sum_{i=t+1}^{T} \lambda_i$) compared to present consumption ($\lambda_t$). There are, however, a number of further implications of (44) that are worth investigating.

Suppose first that $\lambda = \lambda^* = \text{const.}\ \forall t$. Then (44) becomes

$$(47) \qquad \alpha_t^* = \frac{1}{T-t}.$$

This is a very simple rule for the intertemporal choice of consumption. At each point in time $t$, the then available wealth is divided up equally into one part for consumption in the current period and $T-t$ (putative) other parts. Of these, $T-t-1$ provide for later consumption and one provides for final wealth. The rule, of course, implies that the propensity to consume out of wealth is rising over time.

Next, consider a slightly more complicated, but also more realistic, case. It is assumed that the weight factors in (44) take on the form

$$\text{(48)} \qquad \begin{rcases} \lambda_t \equiv \zeta_t, & \text{if} \quad t < T, \\ \lambda_t \equiv \kappa \zeta_t, & \text{if} \quad t = T, \end{rcases} \quad \zeta_t \equiv \frac{\dfrac{1}{(1+d)^t}}{\sum_{t=0}^{T-1} \dfrac{1}{(1+d)^t} + \dfrac{\kappa}{(1+d)^t}},$$

$$d > 0, \ \kappa > 0.$$

With this assumption, a constant rate of time preference is postulated, but with the weight factor $\kappa$ final wealth can be excluded from the discounting rule. From (44) and (48) the following expression for an optimal consumption strategy can be calculated:

$$\text{(49)} \qquad \alpha_t^* = \frac{1}{1 + \dfrac{1}{d} - \dfrac{1}{(1+d)^{T-t}} \left[ \dfrac{1+d}{d} - \kappa \right]}.$$

According to this expression, a passage of time $t$ will only imply a rise in $\alpha_t^*$ if $\kappa < (1+d)/d$. In the case $\kappa > (1+d)/d$ the converse is true. With the passage of time, the propensity to consume out of wealth is falling, since a relative decline in the discount factor applied to final wealth makes saving more and more urgent.

Whatever the size of $\kappa$, the influence of this parameter is not significant if the time horizon is still in the far distance. In the limiting case of an infinite horizon we have

$$\text{(50)} \qquad \lim_{T \to \infty} \alpha_t^* = \frac{d}{1+d}.$$

Roughly speaking, for small $d$, the propensity to consume equals the rate of time preference.

To conclude the discussion, an interesting relationship between the consumption strategy and the degree of risk aversion should be mentioned. From (45) it is possible to calculate the expression

$$\frac{(1-\varepsilon_{t+1}) - (1-\varepsilon_t)}{(1-\varepsilon_t)} = -\frac{1}{1 + \dfrac{\sum_{i=t+1}^{T} \lambda_i}{\lambda_t}}.$$

Comparing it with (44), we find

$$\text{(51)} \qquad \alpha_t = -\frac{(1-\varepsilon_{t+1}) - (1-\varepsilon_t)}{1-\varepsilon_t}.$$

Hence the propensity to consume coincides with the 'shrinking rate' of the difference between the current level of risk aversion $\varepsilon$ and the value of unity. The more thrifty the decision maker is, the more slowly his degree of relative risk aversion approaches the value that characterizes a logarithmic utility function. This rule is independent of any particular aspects of the time path of the weight factors $\lambda_t$ and thus should open our preference hypothesis to empirical scrutiny.

### 2.4. *Result: The Surprising Simplicity of Multiperiod Planning*

The discussion in section B 2, which is about to be concluded, started from a criticism of the multiperiod preference functional normally used in the literature, which is constructed by simply putting the expectation operator in front of the additive preference functional developed by Ramsey, Samuelson, and Koopmans. It was shown that this procedure links risk aversion and intertemporal substitutability of consumption in a very special and arbitrary way and that some simple rationality axioms require the introduction of a specific risk preference function. The specific risk preference function models the decision maker's degree of risk aversion without at the same time affecting his behavior in a world of certainty.

The preference functional thus founded was then specified by using Weber's law and Fechner's law. It turned out that these laws allow for two possibilities: an additive logarithmic functional and a multiplicative power functional that, in the case of a relative risk aversion smaller than unity, takes on a Cobb-Douglas form.

Under the condition of stochastic constant returns to scale, these functionals imply very simple rules for current behavior compatible with intertemporal optimization. On the one hand, the decision maker does not need any particular information about the risk projects available in the future[31]. On the other, the decisions about the optimal level of consumption and the optimal risk project can be found independently of one another.

Since, in the evaluation of risk projects, one of the Weber functions has to be consulted and since, moreover, the Bloos rule can be applied in the usual way, the previously developed one-period approach is near-

[31] However, he has to know the basic assumptions of the model, namely, that stochastic constant returns to scale prevail and that there is no autocorrelation among the risk projects.

ly[32] perfectly rehabilitated. In particular, the shapes of the indifference curves in the $(\mu, \sigma)$ diagram that were derived in chapter III retain their validity. Savage's precept 'You can cross the bridge when you come to it', that was cited in the introduction to section B, therefore met the point fairly well.

Apart from a rehabilitation of the one-period approach, the class of preference functionals derived above has a number of further implications that are worth noting. The most important one is that, with the passage of time, the degree of relative risk aversion relevant for current decision making approaches the value of unity whatever its initial value. This suggests the interesting conclusion that people, who get more risk averse as they grow older, have a comparatively low level of risk aversion which is less than that implied by a logarithmic utility function. At each point in time these people's behavior is determined by one of the Weber functions that are bounded from below and hence exhibit a property that, in connection with the Bloos rule and the demand for liability insurance, was seen to be desirable in explaining observable behavior. A further implication is that the size of the propensity to consume out of wealth and the time change in the degree of risk aversion depend on the decision maker's time preference in exactly the same way. The propensity to consume just equals the rate at which the current degree of risk aversion approaches the degree characterizing a logarithmic utility function.

The highly surprising rehabilitation of the simplest risk-theoretic approach deserves some explanation. The explanation is *not* primarily that an attempt was made to find assumptions that bring about simple solutions. A simplifying assumption that is not immediately plausible, although, at the same time, it is not very restrictive either, is that of intertemporal separability. As Koopmans showed, this assumption implies, in connection with other plausible assumptions, an additive multiperiod preference functional for deterministic planning problems. But, apart from that, the results were derived by a straight-forward line of reasoning, primarily from cogent rationality postulates and the laws of psychophysics. The latter provide the true explanation of the simplicity of multiperiod planning.

Remember: according to Weber's law equal relative intensities of a stimulus are perceived as equally significant. The explanation for this

[32] An exception is that, in the case of strong ($\varepsilon \geq 1$) risk aversion, no comparison can be made between distributions that, with a probability greater than zero, bring about gross wealth levels equal to or less than zero. This does not appear to be a severe constraint, since these distributions are, in any case, suboptimal if at least one distribution is available that guarantees survival with certainty.

phenomenon seemed to be that, in his evolutionary development, man has adapted to the *ratio language* in which environmental signals are encoded. The parallel with our results seems obvious. With the assumption of stochastic constant returns to scale, we made *homo oeconomicus* act in a world in which the relevant information again is formulated in a ratio language. Is it then surprising that he succeeds in finding his way according to simple rules in this world, too?