Chapter 1: INTRODUCTION

"Markowitz' early works have suffered the fate of those of other pioneers: often cited, less often read (at least incompletely)."[1]

1.1 THE INVESTMENT DECISION PROBLEM

1.1.1 Portfolio theory

The field of portfolio theory deals with investment decision problems. In order to establish an intertemporal reallocation of consumption opportunities, economic subjects make investment decisions under uncertainty. In a simple two moment - single period setting, a subject may at time 0 choose to consume less than his or her wealth, or the income he or she[2] is originally endowed with at time 0. The complement, i.e. the savings, he can invest in a set of financial securities: a portfolio. Doing so, this investor hopes to increase his future original income (at time 1) by the additional receipts from the investment, so that he may obtain more future consumption. In this way, investment serves to trade present consumption against more future consumption.[3] The investment decision can be extended to a multi-period (dynamic) setting, where it serves to move from an **endowed** income pattern to a **desired** consumption pattern (or another cash outflow pattern, for example in the case of matching).

In either case, the individual is confronted with a variety of (financial) securities and faces the problem of selecting one portfolio out of the multiplicity of portfolios that can be formed. As the investment decision is prospective in nature and is made in an uncertain environment, the aspect of risk or uncertainty is inextricably bound up with it.

Since the seminal analyses of Markowitz [1950,1952,1959], portfolio theory represents a widely studied class of financial decision

---

[1] Sharpe [1989, p.535].
[2] For brevity and because of personal preference, we'll refer in the following to persons in the male form.
[3] See for example Hirshleifer [1969] and Sharpe [1970].

making problems.[4] Aside from the appealing analytical and practical nature of investment decision problems, results in portfolio theory are characterized by their general applicability to other fields of decision making.[5]

1.1.2 Context

The investment decision is studied in a model world that, by definition, represents a simplified world. This introduces the dilemma between keeping the grasp of reality and obtaining a sufficient degree of tractability. Most frequently (but at the same time mostly implicitly), a one-period time frame is assumed. For simplicity, any slates are wiped clean at the beginning of the period and the investment portfolio is built from scratch.

A time horizon can have different meanings (cf. Merton [1975, pp.662]). The 'trading horizon' refers to the minimum length of time between successive transactions that investors can make in the market. The 'decision horizon' is the length of time between which the investor makes successive decisions, and the 'planning horizon' is the maximum length of time for which the investor evaluates the consequences of his decision(s).[6] A one-period time frame is commonly understood as the situation in which the decision and planning horizons are of equal length and overlap. A multi-period time frame then indicates a planning horizon that captures multiple decision horizons. The choice between a one-period and a multi-period time frame marks the difference between portfolio selection and portfolio management. In the latter process, which is of a dynamic nature, the composition of the initial portfolio and the effect of transaction costs on revising that portfolio are important considerations. In the following, we will constrain ourselves to the static case of portfolio selection and assume a one-period time frame.

---

[4] We hesitate to use the term 'Modern Portfolio Theory' and even propose to reserve the acronym MPT for 'Markowitz Portfolio Theory'. Indicating some stage of development by 'modern' raises the problem whether to call next developments 'neo-modern' or 'post-modern', &c. Instead of these indiscriminate terms, the distinction between (conditional-) normative portfolio theory on one side and positive portfolio theory (capital market theory) on the other has content and is useful.

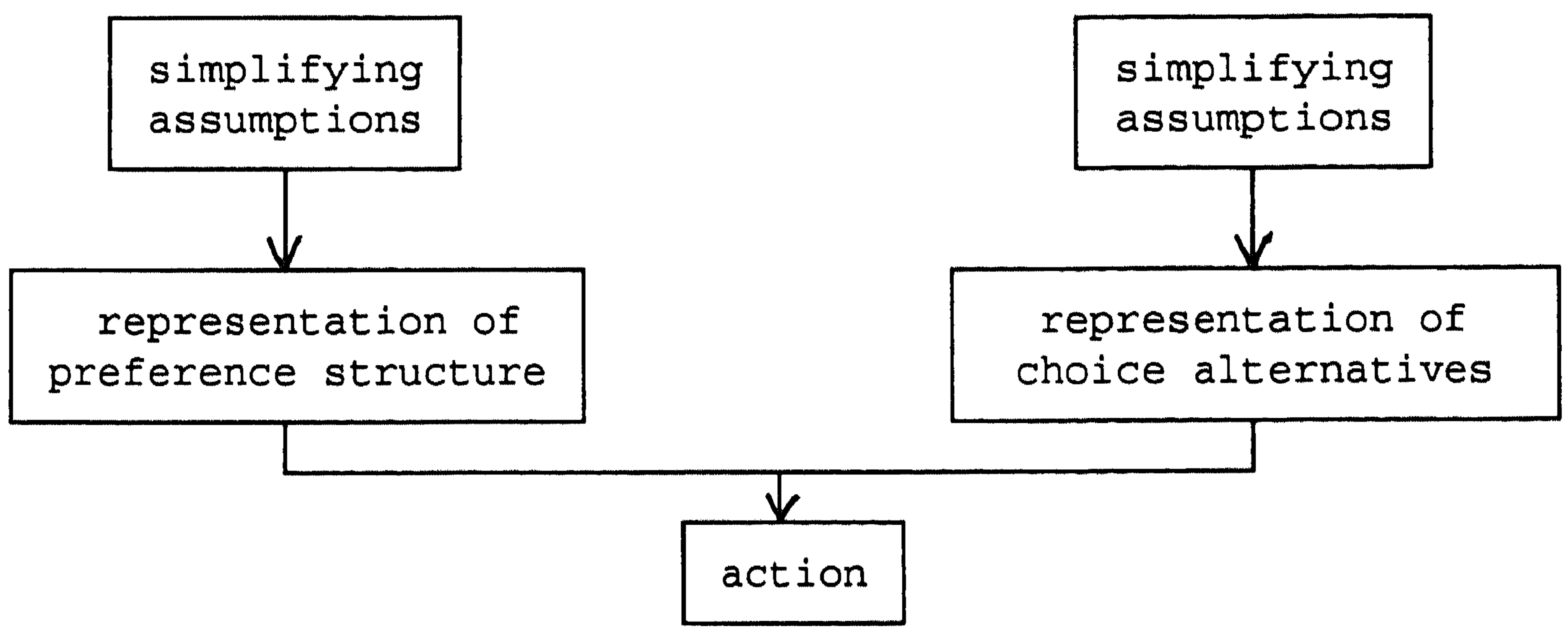
[5] Sharpe [1970, p.1] characterizes portfolio theory as "The Theory of Making Decisions Involving Interrelated Uncertain Outcomes". This stresses the **general** applicability of the attainments, as well as the **joint** nature of the risks.

[6] A fourth horizon is the 'observation horizon': the length of time between successive observations that are used in empirical work.

Conditional-normative framework

In contrast to positive capital asset pricing theories, portfolio selection models are of a (conditional-) normative nature. I.e., **if** the investor (the decision maker) satisfies the assumptions made with respect to goals, tastes and preferences, beliefs and restrictions, **then** the implied decision rule is optimal (cf. Keynes [1891, pp.34-35]). This implies that the validity of the underlying simplifying assumptions is crucial to its adequate applicability. The underlying assumptions relate to the delimitation of the inputs to the decision process and to the combination of the inputs in order to derive a decision rule.

Figure 1.1: General ingredients of a decision making process.



The inputs and their delimitation

To gain a better insight in the nature of investment decisions, the structure of the decision making problem must be analyzed. In most general terms, Figure 1.1 highlights the ingredients of a decision making process.[7] These inputs can be related to the standard investment decision in the following way:

- a set of alternative actions that can be taken:
  the investor can buy and/or sell financial securities in order to arrive at a portfolio with the desired composition (this can be the composition of a new portfolio or the revision of an existing portfolio). The set of potential actions that an investor can take is obvious, but of an unmanageable size. Therefore the set of all

[7] See for example Winkler [1979] for a more extensive list of inputs. Keeney [1982] provides a global overview of decision analysis and further references.

possible financial securities must somehow be restricted to come to an **opportunity set** N that can be handled. The decision then concentrates on buying or selling a specific (money-) amount of security i, with $i \in N$;

- a representation of the choice alternatives:
  This entails the specification of a set of **attributes** that adequately characterizes the securities in the opportunity set. In portfolio theory, it is generally assumed that only financial attributes of securities are relevant. Because "investment is merely a means to the end of future consumption" (Sharpe [1970, pp.19-20]), the future price performance of a financial security or a portfolio is thus crucial to the investor. This price performance can be reduced to a single dimensional rate of return on the investment portfolio that relates current wealth to future (next period's) wealth. The issue becomes more complicated if also non-monetary characteristics are to be considered, such as attributes of an ethical, esthetical or even emotional nature.[8] At this time, however, we will focus on monetary characteristics and specifically on rates of return.

  Obviously, in a world with perfect certainty, there exists no real portfolio selection problem.[9] The actual investment decision, however, is made in a risky environment. This risk can be represented in a formal manner by means of a probability distribution, defined over the possible outcomes of the investment decision (i.e., the future returns). Considering Knight's [1921, p.233] distinction between 'uncertainty' at one side and 'risk' as "measurable uncertainty" on the other, one could argue that the investment decision is more characterized by uncertainty than by risk. Indeed, the objective (i.e. true underlying) probability distributions that generate investment returns are unknown. However, it can be assumed that investors form **probability beliefs** regarding the future and use these subjective or personal probability distributions as inputs for the decision process.[10]

---

[8] Ethical aspects can relate to investments in arms industries, in 'political incorrect' countries (like South-Africa for a long time) or relate to environmental considerations (like tropical hardwood plantations in rain forests), for example. Esthetical and emotional aspects can be important for investments in jewelry, art or oldtimers and the like.

[9] In fact, all financial claims are then identical. If, under perfect certainty, greedy investors operate on a competitive and frictionless capital market, every investment must yield the risk free rate in order to avoid profitable arbitrage opportunities. The same applies to an uncertain environment where all investors are risk-neutral.

[10] Cf. Savage [1971; 1972, Chs.3&4]. For a recent discussion and further references, see Machina & Schmeidler [1992]. From the beginning, Markowitz [1950, p.326; 1952, p.81 fn.8; 1959, pp.257ff] assumes subjective probability.

So, we have a situation of 'risk' "when an individual is willing to base his actions on probability distributions" (Sharpe [1970, p.25]) but the probability distributions are allowed to be of a subjective nature. It is assumed that financial assets can fully be characterized by the probability distribution of their returns; two assets with the same probability distribution will in this view be perfect substitutes.

- the investor's **preference structure**:
  usually a (class of) **utility function**(s) is assumed which is meant to **represent** the preference structure of the particular investor. It serves to incorporate his tastes and preferences and risk aversion in the decision making process in order to make the decision consistent with them. By means of the utility function, the relevant security attributes are made **commensurate** so that their desirability can be expressed as a single number. In this way, trade offs between the values of the various attributes can be handled and the **attitude towards risk** is reflected.

1.1.3 Deriving decision criteria

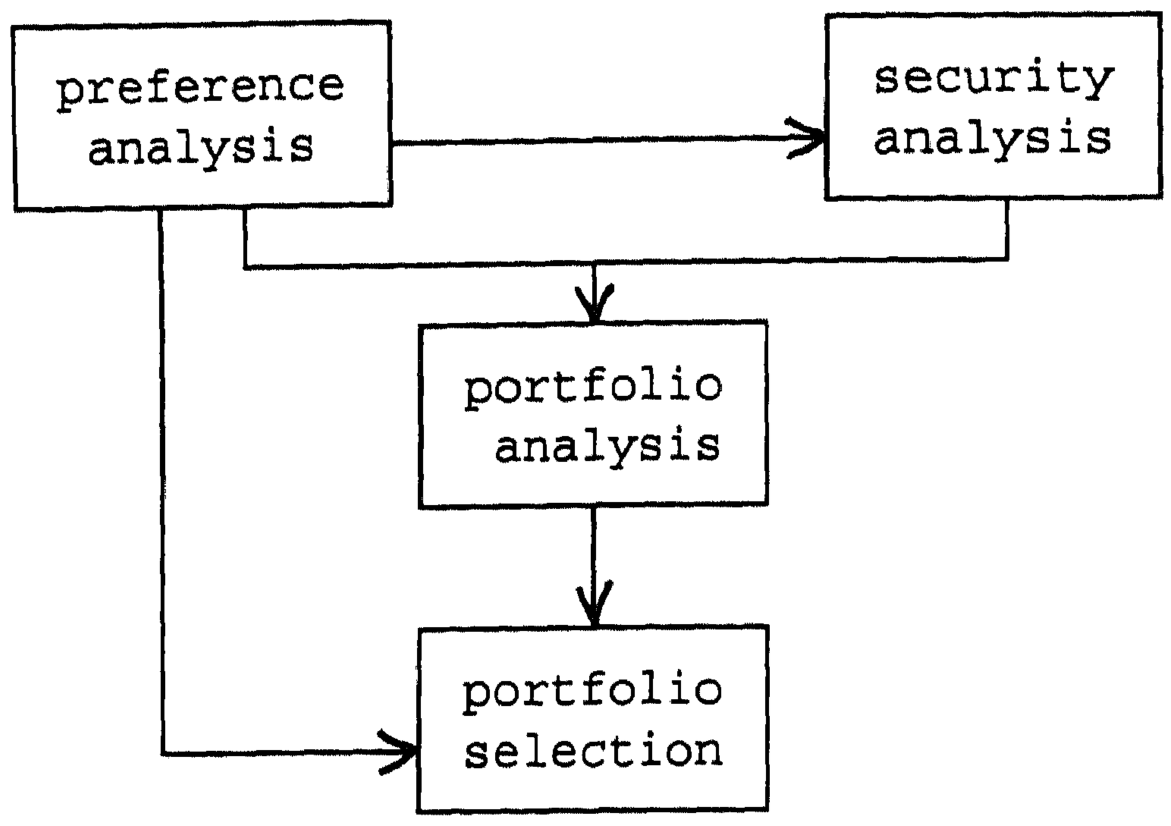
Stages in the decision process

In order to arrive at a decision, given a choice for each of the inputs as indicated above, the investor needs a criterion that specifies how the inputs are to be combined. In general, the investment decision process is decomposed into three stages:

- **security analysis:** the relevant characteristics of the investment opportunities are determined;
- **portfolio analysis:** the set of non-dominated or 'efficient' portfolios is determined;
- **portfolio choice:** the final choice of an optimal portfolio from the efficient set is made.

It is important to add a fourth stage to the process, i.e. the preference analysis. Crucial is the confrontation of the investor's preferences and probability beliefs. For the security analysis, preference information is needed to delimitate the full set of relevant attributes; distribution information is needed for determining the sub-set of relevant risk characteristics. Given the output from the security analysis, general characteristics of the preference structure combined with an optimization or combination procedure will allow portfolio analysis. Detailed knowledge of the specific characteristics of the preference structure finally allows the choice of an optimal portfolio. This is depicted in Figure 1.2.

Figure 1.2: Stages of the investment process.



Decision rules under conditions of risk

As noted before, uncertainty resolves to risk when it is assumed that the future (end-of-period) return $\underline{r}$ on the (financial) investment can be represented in a formal manner by means of a probability distribution. Decision rules under conditions of risk can then be summarized as:

$$\max Z(\underline{r}) \tag{1.1}$$

i.e., choose that action out of the set of possible alternatives which leads to a distribution of returns that maximizes the value of the preference functional $Z(\cdot)$. The expected utility criterion is (still) the most popular approach to formulating a preference functional for the evaluation of future prospects (although its primacy is fading).

The expected utility criterion

Von Neumann & Morgenstern [1947] proved that, if a subject's preference structure satisfies their basic axioms of rational choice in the presence of risk, then there exists a utility function $U(\cdot)$ which assigns numerical utilities to the possible outcomes, such that the preference ordering by this subject over a set of risky outcomes can be represented by the expected value of the utility.[11)]

[11)] In general, given a set of risky outcomes $\{\underline{x}\}$, an expected utility model can be defined as one which —under appropriate assumptions— predicts or prescribes that subjects maximize $\int F\{f(\underline{x})\}U(\underline{x})\,dx$. Within this general model, Schoemaker [1982, pp.530,538] distinguishes nine different variants, depending on *(i)* the transformation $U(\cdot)$ on the outcomes, *(ii)* the type of probability transformations $F(\cdot)$ that are

Assuming a one-period planning horizon, initial wealth W_0 and a probability density function $f(\cdot)$ of next-period wealth $\underline{W}_1$, expected utility EU becomes:

$$(1.2) \qquad EU(\underline{W}_1) = \int_{-\infty}^{+\infty} f(\underline{W}_1)U(\underline{W}_1)dW_1$$

where E denotes the expectations operator. So in the expected utility framework, investment decisions can be viewed as choices among alternative probability distributions of single dimensional portfolio returns[12], the optimal choice being determined by the maximization of the expected value of the investor's utility function.

There are some assumptions underlying the specification of (1.2), apart from Von Neumann & Morgenstern's [1947] basic choice axioms. In particular it is assumed that the stochastic end-of-period wealth is evaluated in terms of a continuous differentiable utility function $U(\cdot)$, which is invariant under positive affine transformations.[13]

For the inclusion of the utility function in the expected utility expression (1.2) to be valid, $U(\cdot)$ must be defined over whole range of $f(\cdot)$. The combination of a logarithmic utility function and a normal probability density function, for example, is not allowed. For the inclusion of $U(\cdot)$ in (1.2) to be meaningful, $U(\cdot)$ must be bounded. When $U(\cdot)$ is unbounded, one can always find a density function that implies an infinitely high expected utility.[14] The solution to this 'generalized St.Petersburg Paradox' entails a restriction, either on the utility function[15] or on the density function.[16]

Although we refer to 'the' utility function, we first remark that the individual's choice problem not only applies to the selection of an investment portfolio but is of a general nature. Yet, many analyses are cast in the form of returns $\underline{r}$, that specify the relationship between the initial proportion of wealth invested in (financial) securities and its

---

allowed, and *(iii)* the way in which the outcomes are measured. Although the general formulation includes Kahnemann & Tversky's [1979] Prospect Theory, it should be noted that the latter theory is more commensurable with human decision making in practice. In particular, prospect theory is able to explain choice behavior that is inconsistent with the expected utility model (in the sense that the behavior gives rise to paradoxes in the latter model).

12) Either objective or as a representation of the investor's subjective beliefs.

13) An affine transformation is obtained by composing any linear transformation with a translation of the origin.

14) Menger [1934]. See also Levy & Sarnat [1972, p.201].

15) I.e., assume a bounded utility function. Cf. Arrow [1970, p.92].

16) For concave utility functions, implying risk aversion, a finite level of expected future wealth is a sufficient condition for finite expected utility. Cf. Arrow [1974, p.137] and Huang & Litzenberger [1988, pp.13-14].

terminal value:

$$(1.3)\quad \underline{W}_1 \equiv W_0(1+\underline{r})$$

So, not only part of total initial wealth is considered, also only the relative change in wealth, irrespective of the level of initial wealth. The former limitation cannot easily be overcome, since it may be expected that there exist interdepencies between the portions of wealth allocated to various appropriations. Hence, detaching an individual's financial investment decision from his other decision problems may result in gross overall sub-optimality. The latter limitation, however, can be justified by assuming either logarithmic, exponential or power utility functions. These utility functions exhibit constant relative risk aversion in the Pratt [1964]-Arrow [1970] sense and hence display the 'separation property'.[17] According to this property, the optimal mix of the investments in risky assets does not depend on the scale of the total investment in risky assets.

Even when limiting ourselves to financial investment decision making, we encounter serious problems in extracting utility functions. Not only must the tastes and preferences of an investor (as of any other individual) be **explicitized**, also a utility function must be specified that adequately **represents** this preference structure in a single dimensional way.

Although we refer to 'the' probability density function of future wealth, it must be realized that it is not sufficient to know the probability distributions of the individual securities' future returns. As an investor will likely combine several securities into a portfolio, the distribution of this portfolio's return is relevant. Because security returns are not statistically independent, knowledge about their **joint** probability distribution is necessary in order to derive the distributions of portfolio returns. Even for a relatively small opportunity set of securities, the assessment and **specification** of the joint probability distribution implies a heavy information burden. An additional difficulty is the adequate **representation** of the distributions in order to incorporate them in some tractable way in the decision process.

Problems .....

In evaluating expected utility as a decision criterion, we are first confronted with various conflicts that exist between reasonable human

---

[17] See Pratt [1964, p.127], Mossin [1968, p.215] and Alexander & Francis [1986, pp.27-28].

behavior and 'maximum expected utility' behavior.[18] Even when we are willing to sidestep this observation and tacitly accept expression (1.2) as a valid decision criterion, we are left with two other major problems that immediately follow from (1.2). First, it appears that the expected utility model can only be a relevant decision making framework for investors who can explicitize in detail their preferences, represented by $U(\cdot)$, and their beliefs, described by the joint probability density function $f(\underline{r}_1,...,\underline{r}_N)$ of the N security returns in the opportunity set. This information problem is expressed in the apt and frequently cited phrase of Roy [1952, p.433]: "The man who seeks advice about his actions will not be grateful for the suggestion that he maximize expected utility." Second, there exists what we could call a combination problem in the confrontation of the investor's preferences and his probability beliefs. The (static) programming problem of direct EU-maximization is mathematically intractable unless both the utility function and the joint distribution of security returns are severely restricted.[19]

Summarizing, the assessment of the joint return distribution, the determination of the investor's utility function, as well as the combination of the two into a decision rule are indeed the central problems in (expected utility-based) portfolio theory. For this reason, simplifying assumptions are necessary in developing pragmatic investment decision criteria. At one side, one can restrict the class of utility functions to be considered. In this way, various classes of utility functions can be specified for each of which specific decision rules can be developed. Besides these **preference assumptions**, one can apply **distribution assumptions** by means of which the number of relevant characteristics of the (joint) probability distributions is reduced. Alternatively, one can **approximate** the preference functional by approximating either the utility function or the (assumed) joint return distribution by some relatively simple form.

..... and solutions and approximations

To gain insight in the way in which the utility function and the probability density function are interrelated in the expected utility

---

[18] These conflicts were recognized by Markowitz [1959, pp.218-228]. As there exist preference patterns that violate the axioms of expected utility theory, the expected utility framework (and thus any embedded decision criterion) is challenged itself. For a recent review and for prospects, we refer to Fishburn [1989] and Fishburn [1988], respectively.

[19] Cf. Ohlson & Ziemba [1976, p.57] and Dexter, Yu & Ziemba [1980, p.507]. The full expected utility maximization becomes computationally intractable for moderate sizes of the opportunity set, i.e. for a number of securities as small as 5!

framework, we can expand the utility function around the expected future wealth in a Taylor series. Taking the expectation then yields:[20]

$$\text{(1.4)}\quad EU[W_0(1+\underline{r})] = U[W_0(1+E\{\underline{r}\})] + \tfrac{1}{2}U''[W_0(1+E\{\underline{r}\})]W_0\mathrm{Var}(\underline{r}) + \sum_{k=3}^{\infty} \frac{U^{(k)}[W_0(1+E\{\underline{r}\})]}{k!} W_0 \cdot M_k$$

where the primes and superscripts between parentheses denote derivatives and where $M_k \equiv E\{[\underline{r}-E(\underline{r})]^k\}$ is the k-th central moment of the portfolio return distribution. Eq. (1.4) makes clear how in the expected utility framework, characteristics of the utility function are linked with the characteristics of the portfolio return distribution. The second and higher central moments M_k are pairwise combined with the corresponding derivatives $U^{(k)}(\cdot)$. Thus, for an n-th degree polynomial utility function, n moments of the probability distribution are needed.[21] If higher moments can be expressed in terms of lower moments, only these lower moments are **explicitly** needed.

For the approximate case, truncating the Taylor series expansion clearly suggests that the approximation of either the utility function in terms of relevant derivatives or the density function in terms of relevant moments can simplify the decision problem enormously.[22]

The use of (a limited number of) moments to summarize (approximate) probabilistic information was already suggested by Hicks [1934]. We will postpone a discussion of the issues connected to moments to chapter 4, section 4.4, where multi-moment portfolio selection is treated.[23]

Concerning the derivatives of the utility function, two most commonly made assumptions are $U'(\cdot) > 0$, implying strictly positive marginal utility or non-satiation, and $U''(\cdot) < 0$, implying decreasing marginal utility or risk aversion, so that the investor will reject

---

20) Assuming, however, that the interchange of the summation and the (expectations) integral is allowed; cf. Kendall & Stuart [1969, p.17], Borch [1973, pp.335ff] and Loistl [1976]. Summarized, there must be uniform convergence of the individual terms under the integral sign; the derivatives of $U(\cdot)$ must exist (see the assumption made before); the moments of $f(\cdot)$ must exist; $f(\cdot)$ must generate wealth realizations that fall within the region of convergence, which in turn depends on $U(\cdot)$. In addition, the moments of $f(\cdot)$ must also determine $f(\cdot)$ completely and uniquely, so that they provide the same information in the Taylor expansion (1.4) as the complete density does in the expected utility integral (1.2). Cf. footnote 30 below.

21) This can readily be seen from eq.(1.4); cf. Richter [1960, p.154].

22) However, neglecting higher moments in the expansion can actually improve the approximation; cf. Loistl [1976, p.909] and Hasset, Sears & Trennepohl [1985, pp.38-41].

23) These issues concern *inter alia* whether moments exist, whether they provide any information about the underlying distribution and, if so, what information they can provide.

actuarially fair bets. Risk aversion can be summarized by the Pratt [1964]-Arrow [1970] risk aversion coefficients. The measure of absolute risk aversion is $-U''(\cdot)/U'(\cdot)$, which is insensitive to the (positive) affine transformations which are allowed for the utility function. Commonly, non-increasing absolute risk aversion is assumed, implying $U^{(3)}(\cdot) > 0$, i.e. preference for the third central moment or 'skewness preference' (note that the reverse is not necessarily true). The measure of relative risk aversion is $-W_0 \cdot U''(\cdot)/U'(\cdot)$, which is commonly assumed to be constant or increasing.

1.2 THE REPRESENTATION OF INVESTMENT ALTERNATIVES

In this study, we decribe a general framework that can serve as an aid in making investment decisions. In connection with the central problems as marked out in the previous section, any decision approach must deal with the closely related issues concerning the distributions of the security returns and the preferences of the investor. Before giving a further exposition of our general approach, we briefly review the way in which the 'distribution' and 'preference' issues are handled in the literature. We start describing the characteristics of one of the most frequently used frameworks in which the investment decision is cast: the mean-variance framework. In addition, we will touch upon some alternative approaches.

1.2.1 The mean-variance framework

Undoubtly the most popular approach to the investment decision problem is the mean-variance- or (E,σ^2)- framework, developed by Markowitz [1952, 1959] and Tobin [1958]. Within that framework, the investor is assumed to be risk averse; he considers an investment in terms of the probability distribution of its return over a fixed holding period (planning horizon) and makes his decision solely on basis of the portfolio's expected return and its variance.

From a historical perspective, the major contribution of Markowitz and Tobin is that they **formalize** the formerly undefined notion of risk by identifying risk with variability of returns in a portfolio context, and **operationalize** the concept by means of the (co-) variance or standard deviation.[24] In this way, risk can be explicitly dealt with. Furthermore, the fundamental concept of risk reduction through diversification emerges.

The stage of security analysis comprises the estimation of the

[24] Note that the variance is a specific way of measuring and representing the more general notion of 'variability'.

expected returns and the variance-covariance matrix. Then the mean-variance efficient set of portfolios is computed; it consists of all portfolios that have a minimum variance for a given value of the expected return and a maximum expected return for a given value of the variance. Now the efficient portfolios are separated from the dominated portfolios, the final choice of an optimal portfolio depends on the specific form of the investor's utility function, as before. This procedure can be simplified greatly when there exist unrestricted possibilities to borrow at and lend against the risk free rate. With that additional assumption the so-called **two fund separation theorem** applies: every portfolio of the efficient set is a combination of the risk free investment and just one portfolio of risky securities.

Basically, there are three ways to 'defend' the use of the (E,σ^2)-decision criterion: as a two-parameter substitutive criterion, as an exact expected utility criterion, or as an approximate expected utility criterion.

(E,σ^2) as a two-parameter substitutive criterion

In his seminal paper, Markowitz [1952, p.82] simply assumes that the investor desires to act according to the mean-variance 'maxim'. He recommends the (E,σ^2)-rule because of its implications: it implies not only diversification, but also the 'right' kind of diversification for the 'right' reason. This 'right' diversification is now commonly referred to as 'Markowitz diversification'; the resulting risk reduction comes from carefully exploiting the non-perfect correlations between the investment returns. This type of diversification can be contrasted with naive diversification that, given a sufficient degree of independence between the returns, follows from the law of large numbers.

However, there is a much longer history to the use of mean and variance.[25)] Edgeworth [1888], for example, invokes the central limit theorem and suggests to employ the normal distribution to describe a banker's solvency risk. This automatically implies the use of mean and standard deviation, but at the same time we admit that this case is not so interesting since it entails the standard use of the 'error distribution'. Fisher [1906, pp.281,409], however, explicitly suggests the use of the standard deviation in addition to the mean in the context of evaluating risks. Marschak [1938, p.320] suggests to cast the utility function not in terms of future returns, but in "parameters (e.g., moments and joint moments) of the joint-frequency distribution" of returns and argues that "[i]t is sufficiently realistic, however, to confine ourselves (...) to two parameters only: the mathematical expectation ("lucrativity") and the coefficient of variation ("risk")." Hicks [1939, pp.125-126] also adheres to a two-parameter framework in

---

25) This material is not covered by the 'historical note' in Markowitz [1987, pp.36-40].

suggesting the use of the dispersion around the expected value in order to make an "allowance for risk". Finally, Hicks [1934, p.195] and Markowitz [1950, p.326] condense probability information in terms of moments, and although realizing that more moments may be relevant, they limit themselves to the mean and variance.

(E,σ^2) as an expected utility decision criterion

As Markowitz [1959, pp.209-210] remarks: "It is logically possible to accept the expected utility maxim and either accept or reject the use of mean and variance as criteria of portfolio selection. Conversely, it is logically possible to accept the use of mean and variance and either accept or reject the expected utility maxim." The latter case was discussed above. Accepting the expected utility maxim, an investor following the mean-variance rule maximizes his expected utility precisely if and only if either an additional preference assumption is made, or an additional distribution assumption. The issue is either whether there exist Von Neumann-Morgenstern utility functions $U(\cdot)$ that generate an expected utility function $EU(E,\sigma^2)$ that is solely a function of mean E and variance σ^2, irrespective the density function $f(\cdot)$, or whether there exist densities $f(\cdot)$ that generate a $EU(E,\sigma^2)$ irrespective the utility function $U(\cdot)$. In both cases, the choice according to the mean-variance decision rule must be consistent with the optimal choice.

The necessary preference assumption is that the investor's utility function should be restricted to a quadratic function (in wealth or returns).[26] In the Taylor expansion eq. (1.4), the third and higher derivatives of $U(\cdot)$ are then zero and the corresponding terms can be ignored. Quadratic utility functions, however, exhibit some undesirable properties. Assuming non-satiation and risk aversion, only a limited range of a quadratic function can serve as a utility function. For some level of wealth (or return) the function reaches its maximum, whereafter negative marginal utility is displayed. But this implies that the assumption of quadratic preferences is **not sufficient** to defend the mean-variance decision criterion as an expected utility criterion: also the return (or wealth) distribution must be restricted in order to prevent drawings beyond the maximum of the utility parabola. As a solution to this problem, Alexander & Francis [1986, p.39] suggest that "the utility function should be scaled so that the maximum (...) is sufficiently high whereby returns greater than this amount are extremely unlikely." However, their statement rests on faulty reasoning, since it can easily be checked that a positive affine transformation of the utility function does **not** affect the position of the maximum.

[26] Tobin [1958, p.67], Markowitz [1959, p.286] and Richter [1960, pp.153-154].

Furthermore, as the third derivative is zero, quadratic utility functions exhibit increasing absolute risk aversion in the Pratt-Arrow sense over their whole domain.[27] Hence, quadratic preferences imply that risky assets are inferior goods.

The necessary distribution assumption is that the security returns are (jointly) elliptically (or spherically) distributed.[28] The elliptical family of distributions can fully be characterized by two parameters. Long time, attention concentrated on a subset of the elliptical family: normal distributions (Tobin [1958, p.75]). Here, the first two moments are the mean and variance; higher odd moments are zero and higher even moments are functions of the variance.[29] Although the elliptical distributions form a wider class than the normal, the argument cannot be extended to any two-parameter family.[30]

We finally conclude that, due to the limited range over which a quadratic function can serve as a utility function, *(i)* the preference and distribution assumptions as mentioned above cannot be invoked simultaneously, and *(ii)* the quadratic preferences assumption is not sufficient and must be supplemented with a distribution assumption that restricts the domain of the utility function.[31]

---

27) Pratt [1964, p.132], Arrow [1970, p.97].

28) See Chamberlain [1983] and Owen & Rabinovitch [1983]. However, Meyer [1987] presents a sufficiency condition in the form of distributions that can be described by only a location and a scale parameter. Bigelow [1993] presents a complete (i.e. necessary and sufficient) characterization of distributions.

29) For the incorporation in the Taylor series expansion, see Rubinstein [1973b, pp.613ff]. Unlike the case of quadratic preferences, the Taylor series expansion cannot be truncated beyond the second term since higher order terms are relevant when higher moments are (exact) functions of mean and/or variance.

30) Although this was initially suggested by Tobin [1958, pp.74-76]). Two-parameter distributions are only fully characterized by their parameters when the parameters do not depend on one another. Lognormal distributions, for example, do not meet this requirement. Cf. Feldstein [1969] and chapter three, section 3.3.3. Furthermore, individual securities can be combined into portfolios and various portfolios can in turn be grouped into other portfolios. It is then required that the (assumed) return distributions are closed under addition. Otherwise, the two-parameter distributions of the component securities or portfolios do not translate into two-parameter distributions of the aggregate portfolios.

31) From another point of view, the mean-variance decision rule is sufficient but not necessary when considering restricted quadratic utility functions. Hanoch & Levy [1970, pp.181-184] show that the efficient set, based on the mean-variance rule, can be reduced without additional preference information. This contrasts Bawa [1975, p.96]. Under the normal distribution assumption (plus non-satiation and risk aversion), however, the mean-variance rule is necessary **and** sufficient (cf. Hanoch & Levy [1969, p.343] and Baron [1977, p.1694]).

(E,σ^2) as an approximation to expected utility maximization

The restrictive distribution or preference assumptions seriously weaken the generality of the (E,σ^2)-approach. But, as Ross [1982, p.53] remarks: "Yet the mean-variance analysis has both a simplicity and a heurism to recommend it, and there does not appear to be any ready substitute. If exact validity is too much to hope for, perhaps there is an approximate sense in which the results remain true." The question is here how adequate the true expected utility function is approximated by a truncated Taylor series expansion, i.e. when non-zero terms behind the summation sign in eq. (1.4) are ignored. As an approximated expected utility function still involves a combination of utility and density, the approximations can be categorized into preference approximations and distribution approximations.

Although Markowitz [1959, p.286] presents quadratic utility as the assumption behind the validity of the mean-variance rule as an exact expected utility decision criterion, the rationale he offers for the mean-variance rule is the **approximation** of a utility function in a certain neighborhood by a quadratic function (Markowitz [1959, pp.121-125,282-286]). The approximation of a utility function by means of a Maclaurin series expansion or a Taylor series expansion around the mean return will only be satisfactory for some region around the spanning point. As the ultimate goal is to approximate expected utility, the approximated utility function must be combined with the return distribution in order to judge its adequacy. The observation that an approximated utility function closely fits the true utility function is no guarantee that the preference ordering according to approximated expected utility is close to that of true expected utility. Levy & Markowitz [1979] and Kroll, Levy & Markowitz [1984], for example, empirically demonstrate the validity of quadratic approximations using historical security return distributions for realistic holding periods. For more information, we refer to Markowitz [1987, pp.63-68] and Simaan [1993].

On the side of distribution assumptions, we have Samuelson [1970], who shows that mean-variance analysis is asymptotically valid in situations involving less and less risk. For these 'compact' distributions, the mean-variance solution to the investor's portfolio problem is asymptotically correct as the return's standard deviation per unit of time approaches zero.[32)] In this category of asymptotic validity, we also have Ohlson [1975] who considers the situation in which the length of the trading interval goes to zero, and Pulley [1981] who considers short holding periods. Whereas Samuelson [1970] considers a situation in which risk is limited in an absolute sense, Tsiang [1972]

[32)] Cf. Pratt [1964, p.125]. However, Samuelson [1970, p.542] notes that the cubic (three moment) solution is a **closer** approximation to the solution for any arbitrarily short, finite time interval. This is a motivation for portfolio analysis in three moments (cf. chapter three, section 3.3.3).

defends mean-variance analysis when the risk is small relative to expected future wealth.

Reviewing the results as presented above, a critical remark is justified. The underlying argument for defending mean-variance analysis as an approximation is that an approximate result or guideline is better than none. We agree that some analytical rigor and academic exactness must be sacrificed in favor of increased practical relevance and enhanced applicability. It must be realized, however, that although the analyses above intend to increase the validity of mean-variance analysis, the underlying assumptions at the same time weaken their practical relevance. When the risks involved in a decision problem under uncertainty are being limited or when the proportion of wealth at stake is being restrained, this implies that at the same time the relevance of the decision problem is being reduced.[33)]

Of a more general character is the analysis of Ross [1982] who shows that, given a very large number of securities, neither the assumption of normally distributed returns nor the assumption of quadratic utility is critical to the validity of mean-variance analysis. In this case, the return distributions are restricted by assuming that security returns are generated by a (single or multi-) factor model. This restriction on the return generating process, together with the process of naive diversification of non-factor risks (from the law of large numbers) ensures that the optimal solution to the portfolio problem converges to the mean-variance solution when the number of securities increases.

1.2.2 Some alternative approaches

Limitations of (E,σ^2)-analysis

The former section briefly discussed the conditions under which mean-variance analysis can be defended, either in an exact or an approximate way. It is important to note that Tobin [1958] and Markowitz [1959, pp.286-288] have recognized and stressed the limitations of (E,σ^2)-analysis. This, however, cannot be said of all of their followers. A very striking example, indeed, is Sennetti [1976, p.962] who even wonders "who needs a utility function to solve the St. Petersburg Paradox?" and uses 'instead' (E,σ^2)-analysis to show why rational men refuse to play this game.[34)] In this respect, we dearly miss a detailed review of the mean-variance discussion (that spans over three decades), more recent than that of Baron [1977] and more detailed than that of

[33)] Asymptotic arguments of the Samuelson [1970]-type for the relevance of only mean and variance can be defended from the viewpoint of continuous time (Brownian motion) processes, but "the concept of a compact family is an independent and completely general one" (Samuelson [1970, p.539]).

[34)] He is corrected by Epps [1978].

Markowitz [1987, pp.52-70]. This is not only interesting from a historical point of view, or as a means to gain easy access to and insight in a huge and scattered literature, or for discovering the links with other (competing) approaches. It is especially relevant since the mean-variance approach is one of the corner stones of financial theory. As new insights and analytical refinements —though they may be marginal— are still published (cf. Bigelow [1993]), mean-variance analysis is not just a 'resolved case'.

From a general perspective, the variance as a risk measure may miss its link with an investor's preference structure or with the distributions of security and portfolio returns. Information concerning mean and variance then is not sufficient to adequately discriminate between return distributions.

Full domain measures: multi-moment portfolio analysis

For example, when an investor is interested in characteristics of distributions over the whole range of the returns ('full domain'), return distributions may exhibit asymmetry or other form characteristics different from normal distributions. In that case, the investor may consider third and higher moments relevant. The relevance of skewness (the normalized third moment) in addition to mean and variance, for example, can be defended by skewed security return distributions[35] or by cubic utility functions.[36] Also higher moments may be considered, in which case we arrive in the general setting of multi-moment portfolio analysis. In this setting, the investor can either specify preferences for the moments, or use the moment information to reconstruct the return distributions (see chapter 3, section 3.3.3). The latter point brings us to the situation in which an investor considers distributions only over some part of the range of returns ('partial domain').

Partial domain measures: downside risk

Partial domain measures for a distribution provide information for some distribution over some part of its domain. Of special relevance is

---

[35] Cf. the empirical results of Sears & Trennepohl [1986], Singleton & Wingender [1986] and Lau & Wingender [1989]. Although they concentrate on pricing, see also Sears & Wei [1988] and Tan [1991].

[36] See for example Friedman & Savage [1948], Levy [1969] and Hanoch & Levy [1970] for a third degree polynomial utility function. However, like a quadratic, a cubic function cannot be a valid utility function for all wealth levels. In general, a polynomial utility function cannot reflect preference assumptions (non-satiation and decreasing absolute risk aversion) on all wealth levels at the same time (cf. Borch [1969, p.2] and Tsiang [1972, pp.355-356]). In an approximate sense, a third order Taylor series approximation is not necessarily better than a second order approximation; see footnote 22 above.

'downside risk', which comes in many forms. Probability of loss is connected to the probability of returns falling below some critical level. An early example is Fisher [1906, p.409], who considered "the chance of earnings falling below the interest-paying line". As he assumed normal distributions, this information can easily be obtained from mean and variance. For other types of distributions, this information may be difficult to obtain, however, and Chebychev-type inequalities only provide crude upper bounds. Markowitz [1959, pp.287-297] discussed various downside risk measures (and the form of the associated utility function), like probability of loss, expected value of loss and maximum loss. In this context, many 'safety first' decision rules have been developed.[37]

Another approach to downside risk specifies risk in terms of probability-weighted functions of deviations below some target return.[38] Markowitz [1959, pp.188-201], for example, suggested the semi-variance. Semi-variance and the more general 'mean lower partial moments' form the partial domain analogons of variance and higher moments.[39] Though intuitively very appealing, these downside risk measures are computationally much more complex than their full domain equivalents and very problematic to use in a portfolio context.[40] Since the (lower) partial moments of individual securities cannot be aggregated in some way into portfolio partial moments (unlike full domain moments), their computation requires knowledge of the **entire** joint distribution of security returns.

The downside risk concept enjoys an increasing popularity and portfolio models based on this principle are even termed by the generic name 'post-modern portfolio theory'. Considering the historic roots and analyses of downside risk measures, this seems to be a little exaggerated (see our footnote 4 above). Still, some thirty years ago, Lorie [1966, p.108] remarked: "I believe that we will ultimately find an objective measure of sensitivity to decline [in price] which avoids the inherent absurdity of calling a stock risky because in the past it has gone up much faster than the market in some years and only as fast in others whereas we call a security which never varies in price not risky at all." Although objectivity may be utopian, theory has gone a long way since then in shaping this notion of relevant risk. At the same time,

---

37) Pyle & Turnovsky [1970] provide alternative specifications of safety first rules and the corresponding seminal references, and Arzac [1974,1977] discusses chance constrained portfolio selection. Both discuss the link with expected utility theory.

38) Cf. Fishburn's [1977] α-t model. Holthausen [1981] extends this approach by considering 'return' as a probability-weighted function of deviations above the target rate of return t. Thus, downside risk is supplemented with 'upside potential'.

39) See Harlow & Rao [1989] for recent references in this area.

40) When the distribution is **given**, they can easily be computed or approximated. Choobineh & Branting [1986], for example, provide a simple approximation for semivariance.

Lorie [1966, p.109] seems convinced of the importance of "measurements of the relationship between the stock market and other things in the economy such as the money supply, interest rates, industrial production, etc." This directly leads us to factor models (see section 1.4.1 below). The use of factor models —even in a mean-variance context— permits replacing return variance as a uni-dimensional risk measure by multi-dimensional risk measures, i.e. the return exposures to changes in the return generating factors. In addition to this decision-theoretic argument we have a statistical argument, in that the use of factor models for simplifying the representation of joint return distributions is indispensable for enhancing the computational tractability and practical applicability of (downside) risk measures.[41]

1.3 THE INVESTMENT DECISION: A CLOSER LOOK

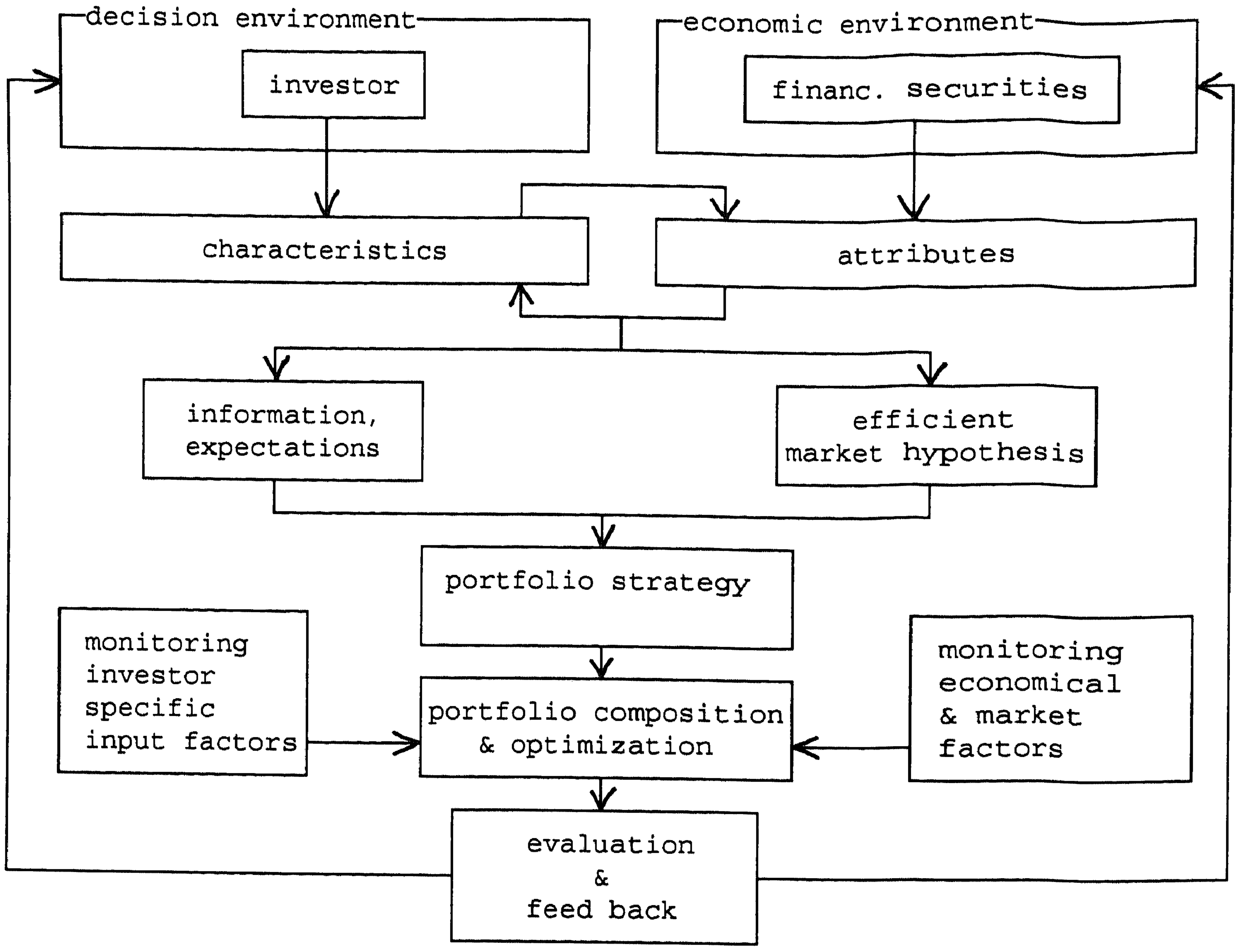
For the construction of a conditional-normative portfolio theory, the real world is replaced by a simplified model-world. As is clear from the former section, the mean-variance framework places quite restrictive assumptions on the preferences of the investor and/or the representation of investment alternatives. In many applications of the framework, the choice for mean-variance analysis seems almost natural and taken for granted, the restrictive nature of the underlying assumptions being unsufficiently recognized. When is opted for one out of many alternative approaches, it may likewise be questioned whether the particular choice is well-founded. Although some simplifying assumptions cannot be avoided, we stress that sensible decision rules can only be obtained when both the desires and preferences of the investor and the characteristics of the investment opportunities are adequately understood.

In order to elaborate the relationship between the decision context of the investor and the economic environment of the securities, Figure 1.3 presents a global scheme of the investment process and thus extends Figure 1.2 of section 1.1.3.

In the economic environment, we have the securities in the opportunity set. These securities possess characteristics or attributes, and these attributes identify various dimensions in which securities are likely to differ: expected return, 'risk', maturity, income component, liquidity, manageability, taxability &c. When securities are issued by a firm, like common stocks or corporate bonds, also characteristics of the corresponding firm can be linked to the securities. This view on securities is objective in the sense that it is an 'outsider's view'. Many of these attributes are enumerated in investment text books (see

[41] This is further elaborated in chapter three.

Figure 1.3: Global scheme of the investment process.



also section 1.4).

In the decision environment, the investor's profile is described. This profile reflects the decision context and comprises the investment objectives that the investor wishes to attain, the restrictions he thereby faces, and his tastes and preferences. In more formal terms: the investor's objective function may be multifarious and complex, and may be subject to constraints. The investor's view on the economic environment in general and on the security characteristics in particular is of a subjective nature.

First of all, the investor's profile determines which of the securities' attributes are relevant in the decision-making process. For example, the investor may have a reference portfolio (for example a liability portfolio in case of a pension fund) which calls for an evaluation of security attributes relative to this portfolio. In other cases, there may be restrictions on foreign investments. The investor's profile also determines the degree of relevance of various attributes, as perceived by the investor. This interdependence between investor's and securities' characteristics, and especially the evaluation of security characteristics relative to an investor's own unique circumstances, is stressed by Smith [1974, p.53].

Secondly, it is important to realize that an investor's evaluation of security attributes is subject to his 'bounded rationality'. It would be unrealistic not to recognize the limitations of the human mind in both observing data and processing these data into information, and next in translating this information to an investment decision. The investor's perspective on and perception of the multitude of aspects characterizing a decision situation is not only subjective but (partly as a result) also limited. The investor will not possess perfect insight in the real, 'objective' world and, hence, does not possess 'perfect' information. There may be simply to many context variables and choice alternatives to monitor, and an investor is likely to use any circumstantial evidence to form a picture of the world.

We conclude that there exists an interdependence between investor's and securities' characteristics, and it is important that this interrelationship between the decision context and the economic environment is explicitly recognized in designing decision rules. It is precisely this notion that underlies the approach we propose in this study.

We can imagine that an investor compares his subjective information about the investment opportunities and his subjective expectations with respect to the future to a more objective set of information and expectations. The latter set is more accurately identified as an interpersonal-subjective set and comprises the information and expectations of market participants. The degree of market efficiency determines to what extent this information is embodied in the market prices of securities. For an investor, the information in the market can serve as a touchstone against which he can evaluate his own information and expectations. Furthermore, the degree to which information already is incorporated in market prices determines the potential success of active investment strategies as opposed to more passive strategies. This issue is discussed in more detail in section 1.6.

The investor's subjective information serves as input for the investment decision. Given the relevant securities' attributes as perceived by the investor, the portfolio composition stage comprises the combination of securities into a portfolio that exhibits a constellation of attributes according to the investor's feelings or preferences. Here again, bounded rationality will leave its traces. Given investor's limitations as indicated before, it would be utopian to suggest that the investor can list all available alternatives, compare them and chose the best or optimal alternative. The decision process will instead be characterized by a step-by-step search for an alternative that satisfies his requirements. Optimizing behavior is then replaced by satisficing behavior. Contributions to investment decision problems then do not entail the specification of 'optimal' decision rules, but the design of systematic search procedures that help the investor scan the feasible choice alternatives.

As portfolio investment is an ongoing process, the investor's profile as well as the securities profiles need to be monitored continuously. Any relevant change is incorporated in the portfolio composition process. In addition, information about the performance of the investment portfolio is fed back and the investment cycle starts again.

In this study, we adopt a one-period planning and decision horizon. This implies that that we will not pursue these issues further than the stage of portfolio selection.

Section 1.4 is devoted to a discussion of security attributes that may be relevant for an investor. In formulating guidelines for investment decisions under conditions of risk, especially the ambiguity of the aspect of perceived risk is a major problem. Indeed, the very lack about adequate definitions of risk, as appeared from section 1.2, may be symptomatic for the multi-dimensional nature of risk. Coupled with the notion that shaping the investment decision to a large extent boils down to shaping the risk characteristics of an investment portfolio, this study mainly focusses on aspects of multi-dimensional risk.

1.4 REFINING THE REPRESENTATION OF INVESTMENT ALTERNATIVES

By relating the investor's decision environment and the securities' economic environment, a set of relevant attributes can be specified. This set can be decomposed into two sub-sets: direct return related attributes and indirect return related attributes. This section discusses both categories of attributes, with special attention for the latter category. The end of this section will show that both categories can be linked together. Hence, the categorization does not imply a separation of various attributes but entails a means to distinguish between the attributes.

1.4.1 Direct return related attributes

The characterization of securities in terms of direct return related attributes rests on the description of securities in terms of their joint probability distribution. In many academic work, it is even assumed that securities can fully be characterized by the distribution of their returns. Two assets whose return distributions take the same position in the joint return distribution of the opportunity set will in this view be perfect substitutes. In a more general view, additional attributes may be required to adequately characterize a security. These other attributes will be treated below. We here constrain ourselves to direct return related attributes.

Any probability distribution can fully be described by means of its locus and its shape. An obvious first attribute candidate then is the 'expected return' on a security, which refers to the location parameter of the return distribution (not necessarily the expected value).

The risk attached to a security's return, in turn, is directly related to the shape of its distribution. The shape of a distribution can be described by its shape parameters. But then one must explicitly assume that the security returns are generated by some specific distribution. Furthermore, as individual securities will be aggregated into a portfolio, it is required that the distribution belongs to the stable class, i.e. the set of distributions that are closed under addition. Otherwise, portfolio returns would obey distributions different from those of the securities and hence possess different shape characteristics. Alternatively, one could try to describe the distributions' shapes by means of their moments. Unfortunately, there is no one-to-one relationship between the shape of a distribution and its moments.[42)]

When relying on either shape parameters or statistical moments as risk attributes, one faces three problems, each related to one of the stages in the investment decision process as described in section 1.1.3. In order to incorporate the probabilistic information in the decision process (i.e. to confront it with the preference structure), one must specify both the investor's tastes with respect to each of these attributes and their relative importance. This gives rise to the 'criteria problem' in specifying the investor's preference functional $Z(\cdot)$ and hence implies a heavy burden on the stage of preference analysis. Furthermore, from the view of portfolio formation, not the (marginal) distributions of the securities must be represented, but their joint distribution. This gives rise to two additional problems. In the stage of security analysis, an 'information problem' arises because all of the interactions between the security returns on the level of the relevant parameters or moments have to be accounted for. In the stage of portfolio analysis, a 'combination problem' arises because the relevant security attributes must be processed in order to obtain portfolio attributes. It follows that the risk dimension is truly problematic in its complexity.

In the light of the discussion in section 1.3, we opt for a (linear) multi-factor framework. We assume that security returns are generated from multiple sources, related to identifiable economic variables. Each of these variables or factors represents a dimension of the economic environment in which the security returns are generated. Depending on

---

42) See chapter three, section 3.3.3 for more details.

the specific circumstances (the profile) of the investor, a specific set of factors may be relevant. The relationship between a security's return and changes in these factors is described by a response coefficient or factor sensitivity. By means of these sensitivities, the joint distribution of security returns is linked to the joint distribution of factor changes. In this interpretation, the sensitivity coefficients can serve as risk measures. Considering factor sensitivities as relevant security attributes, we arrive at a multi-factor representation of security returns.

1.4.2 Indirect return related attributes

The step from a multi-factor representation of security returns to a multi-attribute representation of securities is made by allowing other attributes to appear in the decision process. The incorporation of additional attributes can be motivated from either the specific tastes and desires (goals) of the investor, from specific investment constraints he faces, or from distinctive characteristics of the investment alternatives. In short, the investor can simply indicate that there exist various other attributes with which he can discriminate between the attractiveness of various securities. For example, because of restrictions on foreign investments (as many institutional investors face), the investor may wish to distinguish between domestic and foreign investment opportunities. The same argument applies to the distinction between various asset classes, like stocks, bonds, real estate and cash holdings. Because of the investor's tax situation, the taxability of the portfolio components may be a relevant attribute, in which respect the portfolio's dividend yield may be important. In terms of 'liquidity' or the flexibility to revise the portfolio's composition, the marketability of the component securities may be relevant (and could be measured by their bid-ask spread, for example). Because of some method of performance measurement, the position with respect to some benchmark portfolio may be relevant, and so on. In addition, the investor may adher to the notion that not all future events can be reduced to probability distributions, not even when the latter are of a subjective nature. This also implies that attributes may be considered in addition to explicit elements of return and explicit components of risk. We must seriously consider the possibility that some of these 'other' attributes act in fact as proxies for (components of) expected return and risk.

In short, in practice there exist other elements or attributes that may be relevant, notwithstanding our own opinion regarding their relevance. Some of these attributes may be of a less tangible nature, in which case proxies must be specified. Having a set of attributes which adequately characterize securities in a portfolio context, multi-dimensional portfolio selection can be performed. This will be discussed in detail in chapter five.

'Demand pull' .....

So far, the discussion is of a fairly abstract nature. One way to concretisize the potential relevance of various stock attributes is to look at the various variables that appear in schemes for fundamental corporate firm analysis. One familiar example is the DuPont scheme[43] which allows a breakdown of a firm's return on equity and an analysis of changes in it. A more general example relates to the different aspects that are distinguished in company analyses and industry or sector analyses.[44]

A more direct way to detect various dimensions in which the appraisal of securities (stocks) may differ is to look at the security analyses as conducted by investors in practice. An early study by Baker & Haslem [1974], for example, presents results from a survey, conducted to gain insight in the decision variables employed by individual investors. They conclude that "the investor's investment analysis of common stock appears to be a multi-dimensional process" (p.1261). In particular, Baker & Haslem find that investors greatly differ in their perceptions of the importance of dividends (a.o. in the form of dividend growth, stability and dividend yield), 'future expectations' (towards growth in sales and in earnings per share, a.o.) and 'financial stability' (of earnings per share and stock price).

..... and 'supply push'

With some imagination, the attributes stemming from sources as mentioned above may be labelled 'demand pull'. The data are generally available or can be obtained without many efforts, and investors may use some or all of these data in some way or another. Although the specification of relevant attributes is on the discretion of the investor himself, we can draw a clearer picture of the importance of indirect return related attributes by referring to attributes whose relevance is acknowledged through empirical study. Many of these 'validated' attributes are used by professional investors and advertized in their publications. Hence, these attributes may be marked 'supply push'. Note that we do not suggest that for this reason any investor should consider these attributes. Rather, the implication applies in reverse: because investors in general show a preference (or an aversion) towards some attribute, this can have a negative (positive) effect on the return on a stock that 'has much' of this attribute, and vice versa.

In principle, there are two ways in which the relevance of attributes can be assessed. In what we label the 'transversal approach',

---

[43] See for example Reilly [1994, pp.338ff].

[44] See for example Shapiro [1991, Ch.23], Ross, Westerfield & Jaffe [1993, Ch.2] and Reilly [1994, Chs.17&18]. For an early overview and criticism of (cross-sectional) equity valuation models, based on corporate fundamental data, see Keenan [1970].

the cross-sectional relationships between the securities' scores on the attributes and their (mean) returns are analyzed. The estimated cross-sectional regression coefficient for an attribute then represents the premium or reward for an exposure to that attribute. In the 'longitudinal' approach, the securities are first ranked to their scores on an attribute. Next, securities with either very high or very low scores on (exposures to) some attribute are grouped into a portfolio, whereafter the (historical) return performance of this portfolio is studied. In this way, the relevance of an attribute is 'back tested'.[45] For illustrative purposes, we present below some examples of both approaches. This gives an impression of the range of attributes that can be considered.

Longitudinal approach

The 'multi-factor' model as developed and used by Goldman, Sachs & Co., New York, is an example of the latter approach.[46] By back testing, the following twelve attributes are considered important and included in the model:

- 'value factors':
  - the expected return from a three-stage Dividend Discount Model based on consensus forecasts;
  - private market value (i.e. value of controlling interest in a company);[47]
- 'yield factors':
  - cash flow to price ratio (cash flow yield);
  - consensus expected earnings to price ratio (earnings yield);
- 'momentum factors':
  - earnings estimate momentum (gauging the recent trend in earnings per share estimates);
  - earnings momentum (year change in earnings per share divided by price);
  - share price momentum (average historical price appreciation);

---

[45] Another design of the longitudinal approach is to use a time series regression of returns on attributes. One condition for applying this procedure is that the scores on the attributes vary over the observation intervals. Bauman & McLaren [1982], for example, estimate several time series regressions, relating the annual returns on the S&P 500 to various variables. They find that the return is significantly related to the earnings/price ratio, three-year average past returns and (actual) inflation rates. By including the inflation rate, their model is a mixed factor-attribute model.

[46] See Jones [1987, 1990] and Jones, Kohn & Melnikoff [1990]. Somewhat confusingly, the attributes are called 'factors'.

[47] This attribute replaces the formerly used 'K-ratio', which compared a company's fundamentals (intrinsic value) to its price/earnings ratio.

- 'growth factor':
  - sustainable growth (five year expected earnings growth rate);
- 'risk factors':
  - price risk (residual volatility from regressing return on market index);
  - earnings risk (dispersion of analysts' earnings estimates);
- 'liquidity factors':
  - market capitalization;
  - coverage (number of analysts providing earnings estimates; the opposite of 'neglect').

These variables are then combined (equally weighted) into a proprietary evaluation model.

Transversal approach

The transversal approach appears to be more popular. Sharpe [1982, p.7] selects "more or less *ex cathedra*" a list of attributes of securities. Five attributes are of a common nature, including dividend yield, size (measured by the logarithm of market capitalization), historic α and β values (from regressions on the excess return on the S&P 500), and the bond β (from a regression on the excess return of long term government bonds). Eight more attributes are dummy variables, representing economic sector membership. Sharpe then employs cross-sectional multiple regressions to fit '*ex post* security market hyperplanes' for each of the months from 1931 to 1979. By examining the resulting 'risk premia'[48], he concludes that all of the attributes are relevant for explaining cross sectional returns over time.

Aside from historical α's and β's and size (log market cap), Dowen & Bauman [1986] consider the earnings/price ratio and a 'neglect factor' as stock attributes. The neglect factor (the log of the number of institutions owning the corresponding stock, indicating investor popularity) appeared to be a proxy for size. According to their premia, α and β proved to be poor attributes, but the inclusion of earnings/price ratio and size improved the results. However, the relevance of the latter attributes emerged clearly from longer term holding period returns on a portfolio, composed of small size and low price/earnings ratio stocks.

---

[48] Confusingly, Sharpe [1982, p.6] labels the cross-sectional regression coefficients of the attributes as 'factors'. We prefer to reserve the term 'factor' for an explanatory variable in a time series context. When imposing a cross-sectional restriction on the relationship between security returns and security attributes (as is done in a cross-sectional regression like Sharpe's), the fitted coefficients are 'premia' for exposures to the corresponding attributes. Compare the two-stage procedures to test asset pricing models (as employed by Fama & MacBeth [1973] for testing the CAPM, e.g.).

Bower & Bower [1991] investigate the attribute model that Salomon Brothers developed and used for the electric utility industry over the period 1977:3 through 1990:12. The incorporated attributes are the ratio of dividend to book value, an equity return estimate, a rating and quality ranking and a revenue estimate. By relating these attributes in cross-sectional regressions to price/book ratios, over- and undervalued stocks are identified by comparing predicted price with actual price. Bower & Bower confirm that the model can detect undervalued stocks and are able to link this in part to a high dividend yield.

A final example is BARRA, who develops risk models which are widely used in the investment community. In the categorization as outlined above, the approach as followed by BARRA can be termed a mixed approach. In fact, the BARRA risk models are based on attributes, and the way in which these attributes are further used in the models makes the distinction between direct return related and indirect return related attributes fade.

The general procedure is as follows.[49] The first step is to construct the relevant attributes. In the BARRA model, these constructed attributes are termed 'risk indices' and they measure the stocks' exposures to common factor influences. Each risk index in turn is composed as a combination of various explicit attributes or 'descriptors'. These descriptors are underlying fundamental data items in the sense that they reflect characteristics of the corresponding company or its stock return. The set of risk indices is suppleted with a number of industry designators (dummy variables indicating industry membership).

The second step is to estimate monthly cross-sectional regressions between the excess returns on the stocks at the one hand and the risk indices and industry dummies at the other. The estimated coefficients from each monthly cross-sectional regression represent estimates of the corresponding attribute premia, which are termed 'factor returns'. The regression residuals represent specific returns. Given the stocks' risk indices (which can be considered as factor sensitivities), the variance-covariance matrix of factor returns and the specific variances, the variance-covariance matrix of the stock returns can be estimated. The latter information can finally be used for investment analysis.

The BARRA 'E2' model is developed for the US and is estimated for over 1,000 large capitalization stocks using twelve risk indices and fifty-five industry groups. To fit the model for another 4,500 stocks, a thirteenth risk index is added. The incorporated risk indices are shown in Table 1.1.[50] BARRA has adapted its model for various other

[49] For more details we refer to BARRA [1993].

[50] Cf. Arnott, Kelso, Kiscadden & Macedo [1989] and Fogler [1990].

Table 1.1: Attributes in the BARRA risk models.

Panel A: the US:[a]

	risk index	description
1.	market variability	return volatility
2.	success	relative price performance
3.	size	assets & market capitalization
4.	trading activity	share turnover
5.	growth orientation	subsequent earnings growth estimate
6.	earnings/price	earnings per share divided by price
7.	book/price	book value per share divided by price
8.	earnings variability	earnings & cash flow variability
9.	financial leverage	balance sheet & operating leverage
10.	foreign income	proportion earnings from foreign sources
11.	labor intensity	labor cost versus capital cost
12.	yield	expected dividend yield
13.	low capitalization	additional index for low cap stocks

Panel B: the Netherlands:[b]

	risk index	descriptors
1.	size	- log market capitalization - log total assets
2.	volatility	- historical β - historical β times historical sigma - historical sigma - one-year cumulative range
3.	success	- relative strength - historical α - five-years earnings growth
4.	yield	- current dividend yield - dividend payout - five-year average dividend yield
5.	value	- earnings/price ratio - book/price ratio - sales/price ratio - cash flow/price ratio
6.	growth	- five-year asset growth - change in assets - five-years earnings growth - recent earnings change - sales growth
7.	dollar sensitivity	- return sensitivity to the US$/Dfl exchange rate

[a] From Arnott, Kelso, Kiscadden & Macedo [1989] and Fogler [1990].
[b] Source: BARRA [1993].

countries.[51] For the Netherlands, the 'Dutch Equity Model' was introduced in 1993. The model includes 21 descriptors, organized in seven risk indices, plus eight industry classifications. The incorporated risk indices and their descriptors are also shown in Table 1.1. As a seventh risk index, we see an element from a multi-factor model (and hence a direct return related attribute): the stock return's sensitivity to changes in the US dollar/guilder exchange rate.

By selecting a set of security attributes according to their power of adequately explaining cross-sectional differences in returns, the BARRA approach makes the distinction between direct return related and indirect return related attributes obsolete. Only attributes that can prove their relevance in a cross-sectional context are considered. Although the transversal approach is the starting point, the attributes and the (co-) variances of the derived 'factors' (premia) are used to construct the variance-covariance matrix of returns. Hence, the BARRA approach is also a mixed cross-sectional time series approach.

Finally, we note that in the cross-sectional regression context, various biases can arise when assessing the relevance of attributes. Keenan [1970, pp.256-260] and Jacobs & Levy [1988a, p.24] point at several biases. An important source of bias they do not mention (and that seems neither mentioned elsewhere in the literature to our knowledge and surprise), is the possibility of spurious correlation. This bias arises from the fact that in the cross-sectional regressions both the independent variable (actual or expected return) and many explanatory variables (the attributes) are price ratios. On the LHS of the equation, next period's (expected) price is related to current price, and on the RHS variables like dividends, earnings, cash flow and book value per share are related to the same current price. This will induce what Karl Pearson termed 'spurious correlation' between the return and the attributes (cf. Yule & Kendall [1953, pp.330-331]). So even when there exists no relationship at all between the variables in the numerator of the price ratios and price, one can expect to find considerable correlations between the price ratios and, hence, significant regression results.[52] This does not imply that estimated correlations are fully spurious, but it makes it difficult to determine what part of the correlation comes from a real underlying relationship between returns and attributes and what part is induced by a spurious source.

---

51) For Germany, for example, BARRA developed in cooperation with Schröder Münchmeyer Hengst & Co the BARRA/SMH model with 10 risk indices. In the Netherlands, ABN-AMRO Bank adviced BARRA in developing the Dutch Equity Model.

52) As high as approximately .5 for equi-variance variables in the numerators and the denominators! This effect is not only relevant for cross-section studies but also for time series studies (like Bauman & McLaren's [1982]).

Anomalies can help...

Many attributes are considered important, not only from a practical point of view, but also from an academic point of view because they represent 'anomalies'. An attribute is an anomaly with respect to an asset pricing theory when that attribute possesses power to explain cross-sectional variation in expected returns in addition to the risk measures as specified by the pricing model at hand.[53] An attribute is an anomaly with respect to the efficient market hypothesis when it can be used to forecast future returns. Below, we briefly review some studies on these matters.[54]

Dividend yield can be a relevant attribute because of the investor's tax position, or because the investor does not consider dividends and capital gains as perfect substitutes (cf. Shefrin & Statman [1984]). In addition, however, there seems to be a relationship between expected returns and dividend yields.[55] Furthermore, the importance of dividend yields in a forecasting context is shown by Rozeff [1984] and Fama & French [1988], among others. Fama & French [1989] remove the issue of the dividend yield's forecasting ability from a market efficiency context by arguing that this effect may reflect time-varying expected returns. This opens the way for studying whether this phenomenon conforms to predictions of rational asset pricing theories. So far, however, the evidence on this matter is by no means clear. Goetzmann & Jorion [1993] recently uncovered several problems in the methodologies employed and conclude that there is no strong statistical evidence for the forecasting ability of dividend yields.

Size (log of market capitalization) is a long-time notorious variable. Since Banz [1981], this easily measured variable is studied extensively. Chen [1988] Jacobs & Levy [1989] review the overwhelming evidence on the size effect and its relation to other pricing anomalies; the latter authors also investigate the predictability of the payoffs connected to the size effect. A long history of empirical research also has documented the ability of the **earnings/price ratio** to explain expected returns in cross-section (Basu [1983]) and to identify superior future performance (Basu [1977]).

Several studies indicate the importance of the **book/price ratio** (book value of common equity per share divided by market price per share). Rosenberg, Reid & Lanstein [1985] test a 'book/price' strategy by buying (selling) stocks with a high (low) book/price ratio, and the abnormal performance of the strategy leads them to reject market efficiency. For the Japanese stock market, Chan, Hamao & Lakonishok

[53] For example CAPM-β, or factor sensitivities in an APT-context.

[54] We concentrate on seminal studies and most recent studies. For more detailed overviews and many additional references we refer to Jacobs & Levy [1988a], Keim [1988], Fama [1991] and Hawawini & Keim [1994].

[55] See for example Ang & Peterson [1985] and the references cited therein.

[1991, 1993] also find the book/price ratio (and to a lesser extent the **cash flow/price ratio**) relevant for explaining expected returns. The role of the earnings/price ratio and size appears to be mixed. In the latter study, they also report abnormal returns from strategies based on book/price ratio and cash flow/price ratio. Capaul, Rowley & Sharpe [1993] also report abnormal returns from book/price strategies in six countries.

Bhandari [1988] presents evidence that, even when controlling for beta and firm size, the **debt/equity ratio** is relevant for explaining expected returns and suggests that the (positive) premium for this leverage attribute is not likely to be a risk premium.

Fama & French [1992] evaluate the relevance of size, earnings/price ratio, leverage and book/price ratio in a joint context. Although the attributes in separation do a very good job in explaining cross-sectional differences in expected returns (unlike CAPM-β), the combination of size and (above all) the book/price ratio seems to absorb the influences of the earnings/price ratio and leverage. These results are confirmed by Fama & French [1993], using a time series regression context.[56] They first sort stocks according to size and book/price ratios and divide them into two size groups and three book/price ratio groups. Next, they construct six portfolios from the intersections of these groups. Finally, they regress stock returns on the returns of these mimicking portfolios. The estimated sensitivities are now risk measures and take the place of the stocks' scores on size and book/price ratio. The results point at the existence of common risk factors for which size and book/price act as proxy. Fama & French [1993] argue that both size and book/price ratio are related to the firm's profitability (smaller firms tend to have lower earnings, just like high book/price ratio firms). In addition, most of the price/earnings effect and the dividend yield effect seems to be absorbed by size and book/price.

Finally, the extensive study of Jacobs & Levy [1988b] examines the relationship between actual stock returns, expected returns from a dividend discount model (DDM) and a set of as much as 25 equity attributes. From their results (which are comparable to those cited above), it is most notable that they found that the "DDM expected return is nothing more than an additional equity attribute" (Jacobs & Levy [1988b, p.56]). This strengthens the notion (as stated before) that indirect return related attributes can act as proxies for expected return components and that it hence makes sense for an investor to consider these attributes in addition to expected returns estimates from dividend discount models.

What can we learn from these pricing results? That book/price ratio and size (and perhaps other attributes) just happen to describe the cross-section of expected returns by chance cannot be ruled out, but is not so likely. The question for academics then is whether the results can be

[56] See also Fama, French, Booth & Sinquefield [1993].

explained either by rational asset pricing (the attributes proxy for some risks) or by irrational asset pricing (as a result of overreaction).[57] For an investor, these results justify considering the inclusion of book/price ratio and size (among others) as relevant attributes. This, however, is not to say that no other (non-priced) attributes may be relevant. The pricing argument to selecting attributes is further elaborated on in section 1.5.

1.5 OF MARKET PRICING AND PROCRUSTES

In section 1.4, the transversal approach for deriving relevant security attributes was described. According to this method, first a set of potential attribute candidates (factor sensitivities and/or fundamental firm characteristics, e.g.) is selected. Then, the cross-sectional relationship between returns and characteristics is analyzed. Finally, according to their power of adequately explaining the cross-sectional differences in returns, a set of relevant attributes is obtained. These are the only attributes considered important in the investment decision. Only attributes that can prove their relevance in a cross-sectional context are considered. Assuming a rational pricing framework, it can be hypothesized that these attributes proxy risk measures that through their premia contribute to a security's expected return.

In this section we investigate whether pricing models provide a standard for judging the relevance of attributes. Stated otherwise: what is the role of market pricing in the distinction between 'relevant' and 'non-relevant' attributes?

Pricing as a constraint

We can draw a sharp distinction between (multi-) factor models and attribute models on one side and pricing models on the other. Factor models or return generating processes in general involve a description of the structure of uncertain (i.e. unexpected) security returns. Pricing models (comprising equilibrium-based models like CAPMs[58] or Equilibrium-APTs[59], and pure arbitrage-based models like the APT[60]) by their very nature specify a structure in expected returns. So, whereas a factor model is purely a statistical model,

---

57) Cf. Fama & French [1992, p.451] and Jacobs & Levy [1988b, p.50].
58) Sharpe [1964, 1977], Lintner [1965a,b] and Mossin [1966].
59) See for example Ross [1982], Grinblatt & Tittman [1983], Dybvig [1983], Chen & Ingersoll [1983], Connor [1984,1989] and Shanken [1985]. Ehrhardt [1987], however, does **not** fit this category.
60) Ross [1976, 1977b].

"the economic content of any equilibrium asset pricing model can be expressed in the form of a potentially refutable constraint on the set of parameters of the process generating security returns. This constraint specifies that, given a set of security specific characteristics, there exists a linear relationship common across securities relating expected returns to these characteristics. The coefficients of this linear relationship represent the market prices [premia] of the different characteristics."

(Brown & Weinstein [1983, pp. 711-712]). In expectation forms of factor models (as will be described in chapter 3), or in multi-attribute models in general, this cross-sectional constraint is clearly absent. On this basis, factor models or attribute models could be criticized because of their arbitrariness and termed *'ad hoc'*. In this pessimistic view, the specification of factor models (or attribute models) then reduces to data mining, data fishing, factor dredging or, euphemistically, to an attempt to 'let the data speak for themselves'.

One could attempt to defend the specification and estimation of factor models by raising the argument that these models are often used as underlying return generating processes, in order to make the necessary transition from ex-ante defined pricing models to ex-post return data (cf. chapter 2). However, as pricing models are by definition of a descriptive or positive nature, not their assumptions are to be subjected to empirical tests, but their implications. Furthermore, the cross-sectional restriction that a pricing relationship places on the parameters of the return generating process (or more general, on the relationship between security attributes) as specified in the first stage, would separate 'relevant' from 'non-relevant' attributes in the second stage. Portfolio selection models, in contrast, are of a (conditional) normative nature, implying that the validity of the underlying assumptions (in the form of a set of relevant attributes) is crucial to their adequate applicability.

Equilibrium and arbitrage pricing

The question whether market pricing provides a standard for judging the relevance of attributes brings us to the complexity of the pricing process and its underlying mechanisms. A pricing model specifies premia for security attributes, i.e. rewards that offer no arbitrage opportunities and that clear the market. In this sense, 'non-relevant' attributes are 'non-systematic' because they can be diversified away, either in large, well diversified arbitrage portfolios (in case of arbitrage pricing) or in the optimal aggregate or market portfolio (in case of equilibrium pricing).

Pure arbitrage-based pricing models stipulate that a change in a portfolio's composition without changing invested wealth should not lead to a change in its expected return without altering at the same time the risk characteristics that fully describe its stochastic return. Otherwise the revised portfolio stochastically dominates (in first order) the initial portfolio, making the zero-cost risk free revision portfolio an arbitrage portfolio. In this sense, an arbitrage pricing theory can be considered as a 'normative portfolio revision theory' (Hallerbach [1987, p.3]). It restricts the way in which investors can revise their portfolios, but it does not make any assumptions on the nature or composition of the portfolios investors hold. So, investors may be holding portfolios that are very diverse in composition, without violating the no-arbitrage condition. Diversity in portfolio composition in turn implies that investors exhibit diverse tastes and preferences, also with respect to the relevance of the characteristics describing its stochastic return component.[61)]

Turning to equilibrium pricing theories, one can doubt or dispute the validity of the underlying equilibrium assumption. Static models assume that the market is in equilibrium; comparative-static models assume that if we are in disequilibrium, we will move back; and in a dynamic context it is assumed that the market moves from one equilibrium to another. Findlay & Williams [1980, p.11; 1985, p.9] contrast this equilibrium view with the possibility of a continuous disequilibrium, where the market moves from one disequilibrium to another. This indeed would render most analytical tools developed sofar useless and throw us back to the no-arbitrage condition (which is necessary but not sufficient for the existence of equilibrium). One could argue, in defense, that not the assumption of equilibrium is relevant, but the empirical validity of the implied pricing relationships. But then again, empirical tests are frustrated by the intrinsic tautological nature of equilibrium pricing relationships, induced by the within-sample optimality of the aggregate portfolio's proxy.[62)]

This brings us to the question whether pricing relationships critically depend on the optimality of the aggregate portfolio, as the sum of all investors' optimal portfolios. One can argue that, if the prevailing security prices are to have an equilibrium character, there must exist some degree of consensus between the market agents with respect to the value of security attributes (because this value is reflected in the market prices). Indeed, the assumptions of "homogeneity of investor expectations" (a term introduced by Sharpe [1964, p.433]),

---

61) In addition, pure arbitrage opportunities may be hard to find on a portfolio level since it is very difficult to assess its stochastic return characteristics exactly. This would replace 'arbitrage' with a 'switch', involving "an individual selling what he perceived as an inferior holding in order to purchase a superior one" (Findlay & Williams [1985, p.7]).

62) Cf. Roll [1977] and chapter three, section 3.4.

and equal portfolio opportunities play a major role in the development of asset pricing theories. The function of these assumptions, however, is merely to simplify the analysis. Invoking homogeneous anticipations and portfolio opportunities allows to focus on 'the representative investor'. Solving the portfolio selection problem for this investor solves the portfolio problem for the aggregate of investors. Combined with the market clearing condition, the equilibrium relationships between the (relevant) security attributes follow from the specification of the optimal aggregate portfolio.

Although the homogeneous anticipations assumption plays a major role in simplifying the analysis, it "plays no **essential** role in either the determination of a market equilibrium or the analysis of the details of [investor] equilibrium. (...) [M]arket equilibrium simply requires a set of asset prices such that supply equals demand for every asset" (Fama & Miller [1972, pp.292-293], our emphasis). Indeed, several studies have analyzed market equilibrium and asset pricing under heterogeneous anticipations. In the (E,σ^2)-framework we have the seminal work of Lintner [1969] who, for the sake of analytical tractability, balances heterogeneous anticipations with restrictions on investor preferences.[63] Although market equilibrium prevails on the basis of weighted averages of heterogeneous expectations, the investors in the market hold risky portfolios that differ in composition.[64] William's [1977] continuous trading model assumes that for some investors the opportunity set is partly unknown so that the errors in estimating mean returns give rise to heterogeneous anticipations. This model also implies a consensus pricing relationship, although investors' risky portfolios differ, beta no longer remains a complete measure of risk and residual risk enters the pricing relationship.

The more recent model of Merton [1987] is more general in its specification: imperfect information is replaced by incomplete information. This indeed is one aspect of investors' bounded rationality as we discussed it in section 1.3. According to Merton's assumptions, each investor only knows about a subset of the available securities, implying that investors select portfolios that are (E,σ^2)-efficient relative to the securities contained in their information sets. Like Williams [1977] (who is not cited by Merton), Merton [1987] finds that the equilibrium pricing relationship includes other characteristics (depending on residual return variance) in addition to the market risk that prevails under homogeneous anticipations. Hence, beta no longer remains a complete measure of risk and the market portfolio is not mean-variance efficient.

These examples clearly show that investors solve their portfolio

---

[63] He assumes either negative exponential utility and normality, or quadratic utility.

[64] Instead of restricting preferences, Gonedes [1976] imposes (quite severe) restrictions on the heterogenity of perceived return distributions and finds that the minimum-variance frontier is the same for all investors.

problem in the context of their own anticipations, beliefs, preferences and restrictions, confirming Fama & Miller's [1972, p.293] early remark that "[i]n essence, the risk of an individual asset to an investor is always measured relative to the portfolio that he holds, regardless of whether his expectations or his portfolio opportunities are the same as everyone else's." This is also stressed by Sharpe [1970, p.109]. Then, however, Sharpe [1970, p.113] continues that "a model based on disagreement (...) has little value in a positive role. But it provides the motivation for normative applications of more complicated portfolio-analysis techniques. Unless an investor can make superior predictions about the future, there is very little to be gained from the use of complicated mathematical procedures." Sharpe equates 'superior predictions' with the ability to detect seriously mispriced securities. When an investor does not posses this ability, he then should act according to the 'complete agreement' model (Sharpe [1970, p.113]). This implies the advice that the investor become a passive investor. But this can be contrasted with the active investors in Merton's [1987, p.507] model who certainly do not posses superior information but instead select their portfolios to be (E,σ^2)-efficient relative to the securities contained in their information sets.

The argument can be extended from the (E,σ^2)-framework to a general multi-attribute framework. The evaluation of an investor's portfolio should take place considering his personal context; hence, the evaluation of a portfolio's attributes crucially depends on the expectations, tastes and preferences of the particular investor holding it. Invoking no-arbitrage or market clearing conditions can result in pricing relationships on the market level, but investors' portfolios can differ from the aggregate portfolio, implying that the relevance of attributes (and their number) on a micro and a macro level may differ. Indeed, referring to pricing models in order to delineate the set of relevant attributes is reversing the argument.

From normative theory to positive theory

The very statement that non-systematic risks (with systematic defined with respect to the aggregate portfolio) are not priced and hence should be avoided, rests on circular reasoning. It leads to the obvious question whether non-systematic risks are not priced because they can be avoided, or whether non-systematic risks are avoided because they are not priced. By definition, non-systematic risks are risks that actually **are** avoided on the macro level of the aggregate portfolio. But this does not imply that an individual investor **will** avoid 'non-systematic' risks. In general, the question is: what behavioral recommendations can be obtained from pricing models? In general, the answer is: none.

The natural line of reasoning goes from normative models to positive models. This is summarized in Figure 1.4. Starting point is the portfolio selection problem of individual investors. Aggregating microequilibria into a macro-equilibrium yields a market pricing relationship. The most simple example is the standard CAPM. The micro-equilibrium is studied by Markowitz [1952,1959]. Sharpe [1964] then extends this to a macro-equilibrium by studying a representative investor and invoking market clearing. Using the macro-equilibrium in order to advice individual investors to hold combinations of a risk free portfolio and the market portfolio then rests on false reasoning.

Figure 1.4: Positive versus (conditional-normative theory.

positive theory	<———	(conditional-) normative theory
macro equilibrium social level		micro equilibrium individual level
aggregate ('market') behavior		(individual) investor behavior
market equilibrium absence of arbitrage opportunities		portfolio optimality

Of course, because the investors are agents on the capital market, the way they form their portfolios -and thus the way in which they appraise its attributes in the light of their personal context- has implications for the valuation and pricing of these portfolios (and hence the incorporated securities). But this implies that pricing is a result of complicated multi-dimensional decision processes of the market agents, i.e. the multifarious investors with varying opportunities, constraints and preferences. The existence of short sales and of options, e.g., necessarily reflects heterogeneous anticipations and differences in preferences. Under heterogeneous anticipations, the perception of attributes is subjective and can differ among investors. Because of taxation, flexibility in portfolio revision (depending on liquidity and marketability) and because of information costs, for example, the settings (i.e. personal context) in which investors evaluate the desirability of attributes, differ. Under heterogeneous anticipations, there is no guarantee that the aggregate portfolio is the optimal (risky) portfolio for all investors in the market. By definition, the aggregate of all investors' portfolios is the market portfolio but, as stressed by Ross [1978, p.887] in the context of the CAPM, nothing guarantees that the market portfolio is efficient. The investors' subjective appraisals of securities amalgamate to market prices, which in an efficient market reveal 'objective valuation'. It is then quite restrictive to accept this melting pot of evaluations as a delimiting criterion on the individual investor level. In their attribute appraisals, investors attach subjective premia to attributes and there is no reason that subjective premia equal market premia. When the market premium of an attribute is zero, its subjective premium may be non-zero for some investor. In trying to match an investor to the market level in this respect, by either stretching or reducing the set of relevant attributes, one would act like the madman Procrustes.

Resume

An attribute's ability to contribute to the explanation of cross-sectional return differences appears to be a convincing criterion for the selection of relevant attributes. However, an attribute will only carry a significant premium when it is 'priced' in the market. An investor can face an opportunity set of investment opportunities that is different from the market. Hence this investor is only interested in the relevance of this attribute in his opportunity set. Futhermore, partly connected to the former argument, the reward that an investor attaches to the exposure to an attribute (a 'subjective premium') may well be different from the premium that the market as a whole attaches to that attribute (the 'objective premium'). This leads us back to the starting point that the selection of attributes depends on the personal circumstances of the investor, as summarized in his profile. In a (E,σ^2)-context, Markowitz [1983, p.294] remarked that "[t]he use of portfolio analysis by an individual investor can be desirable even if no one else in the world uses it. Thus the success of any particular [no short-sales mean-variance] CAPM as a positive theory is in no way a prerequisite to the desirability of [no short-sales mean-variance] analysis for normative purposes." We hope that this section offers convincing arguments for supporting that claim and for extending its validity beyond (E,σ^2)-analysis. It also means a further challenge in developing positive pricing theories.

1.6 RISK MANAGEMENT IN AN EFFICIENT MARKET

The efficient market hypothesis and the informational efficiency of capital markets have great relevance for shaping investment decisions. According to the efficient market hypothesis, all available information is incorporated in the security prices (where the delineation of the information set determines the degree of efficiency). Whereas security pricing theories specify some structure in security prices or security returns, the market's informational efficiency determines the degree in which market prices reflect the underlying pricing structure.

In designing a portfolio strategy, an investor must confront his information concerning the future with the information set that is expected to be already incorporated in the market prices of the securities (cf. section 1.3). The efficient market entails a hypothesis, not a dogma, so we can not preclude that an investor possesses information that is not (yet) reflected in the market prices of securities. Note, however, that the prices in a (fairly) efficient market do not reflect the consensus expectations (or information of the **average** investor, whoever that may be), but the expectations of the **marginal** investors who trade in the specific securities. In addition, investors' judgements concerning the return on an investment for example, are (at least indirectly) judgements about the reactions of other investors to the investment. When expectations are actively used in the decision process, it is important to evaluate the investor's expectations with respect to the efficient market hypothesis: in what degree do his expectations differ from the information that is already reflected in the market prices by the trading activities of marginal investors?

The efficiency of the market also plays a role in the stage of security analysis, where the securities' exposures to attributes are determined and evaluated. An important point in this issue is the difference between data and information. Data are objective in the sense that they can be observed in an objective way as 'facts'. Information contributes to the knowledge of an investor and is the result of processing and interpreting the data. As argued in section 1.3, the gathering of data and the processing of these data into information is subjective and not only depends on the limitations of the investor's mind but also on his decision context.

Since long, then, we wonder why the essential difference between data and information, so familiar in information science, is not made in the context of market efficiency. One exception in this respect is Bernstein [1975], who on this ground attacks Lorie & Hamilton's [1973, p.100] statement that "the most general implication of the efficient market hypothesis is that most security analysis is logically incomplete and valueless." Bernstein's [1975, p.61] argument is based on the notion that "whereas the market on average mirrors the best in human skill and inginuity, it also mirrors human frailties and imperfections, including tendencies to under- or over-react to new developments." To this argument we then add the importance of an investor's decision context in providing specific relevance to specific sets of information.

In an inefficient market, investors can earn 'abnormal returns' by profiting from superior information, but in an efficient market, an investor gets value for money. This is often translated to the recommendation that investors should replace 'stock picking' by throwing darts at the financial pages. This 'stock pricking', however, can degenerate into a dangerous sort of financial roulette. Surely, an investor will end up with a combination of securities that is priced correctly, but the resulting portfolio is not likely to match his profile. Especially with respect to the portfolio's risk profile, an investor will display specific preferences. Accepting the multi-dimensional nature of the risk concept, there are many dimensions in which an investor can shape his portfolio. Even if all available securities are fairly priced (i.e., their prices accurately reflect the underlying pricing structure and fully embody 'market' information), risk management is important. The reason is that the subjective risk premia (i.e. as perceived by the investor) may differ from the objective (market) risk premia because of the specific circumstances (decision environment) of the investor, as set out in section 1.3. The same argument applies to indirectly related attributes.

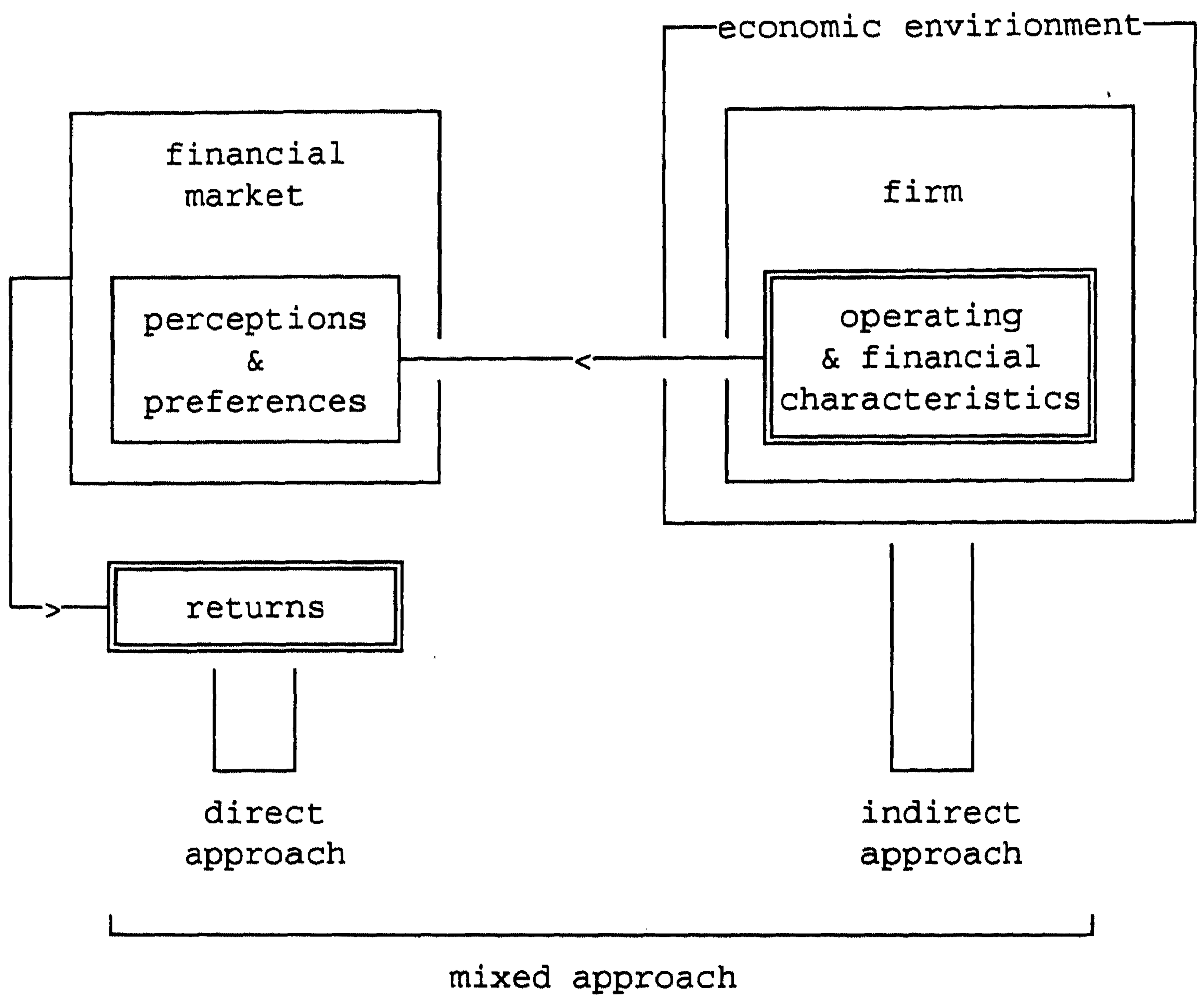
We can even make a stronger claim: a certain degree of informational efficiency is required in order to adequately determine the risk characteristics of securities and portfolios. We therefore stipulate another 'efficient market paradox' in addition to Lorie & Hamilton's [1973] paradox:[65] the possibilities for adequate and 'active' investment risk management increase with the degree of informational market efficiency. The argument for this statement lies in the importance of informational market efficiency in the process of determining the securities' exposures to attributes.

The direct return related security attributes take the form of expected return and sensitivities to changes in common factors. Combined, these attributes constitute the joint return distribution, conditional on the joint distribution of the factors. Specializing to common stocks, returns are determined by operational and financial firm characteristics (both related to external influences or factors) **and** the market evaluation of these characteristics. The market evaluation of these fundamental characteristics, which may be partly rational and partly 'irrational', materializes in the observed stock returns. This is depicted in Figure 1.5.

The direct return related attributes may be assessed in a direct way. This implies that stock returns are related to changes in the firm's environmental factors in order to estimate its sensitivities to these factor changes. So, although hunches about the relevance of factors may be derived from analyzing the firm and its position in the economy, the firm characteristics are assessed by analyzing its stock return behavior. The market evaluation of attributes is accepted in the sense that stock returns are assumed to reflect adequately the (changing) position of a firm in the changing economic environment in which it operates its activities and generates its income. The estimated

[65] Lorie & Hamilton's [1973, p.98] efficient market paradox stipulates that a market can only be efficient when sufficient investors realize that the market is not efficient. Otherwise, investors would have no drive to collect 'information'. See also Grossman & Stiglitz [1980] on this point.

Figure 1.5: The direct and indirect approach for deriving direct and indirect return related attributes.



characteristics will only be unbiased when the return behavior adequately reflects the changes in the firm's environmental factors. The latter condition depends on the degree of market efficiency.

Indirect return related attributes can be obtained by analyzing the firm and choosing firm characteristics of a fundamental nature. In this way, the role of the financial market in translating scores on attributes to return is bypassed. Instead, the investor relies on his own judgment in assessing the relevance of attributes. Here, the degree of informational market efficiency is not crucial in this assessment process. However, direct return related attributes like the sensitivity for changes in, say, the dollar/guilder exchange rate could also be assessed by performing a sensitivity analysis of the firm's cash flows to these exchange rate changes. Using a present value model (dividend discount model), the effect on the stock price or the expected return could be estimated. As the market's evaluation of this factor change is bypassed, it is difficult to assess in what degree the indirectly estimated change in price or return will materialize. This suggests a mixed approach of assessing a stock's exposure to attributes. The returns, resulting from changes in a firm's environment are assessed in both a direct and indirect way. By confronting the outcomes of both methods, the relevance of possible differences can be evaluated. It is then the investor's choice to disregard one of these estimates or to weigh both estimates in a way to obtain an adequate estimate.

1.7 GOAL AND OUTLINE OF THE STUDY

Goal of the study

The discussion in the last sections leads us towards a general approach to portfolio selection problems. In this study, we describe a coherent conditional-normative framework that can serve as a decision aid in the portfolio selection process. In connection with the central problems as set out in section 1.1.3, our approach must deal with the closely related issues concerning the preference structure of the investor (comprising his goals, preferences and restrictions), his beliefs in the form of the (subjective) joint distribution of security returns, and other security characteristics that the particular investor considers relevant.

In this spirit, we propose a basis for a general approach to the investment decision under uncertainty. The main feature of the framework is that we look for a close correspondence with decision making in practice. This translates into four essential characteristics of the approach.

First, we present an approach that is **unified** and **coherent** in the sense that it covers all of the stages of the investment decision process as outlined in Figure 1.2. Furthermore, the approach is **consistent**, not only with decision theory but also with financial theory. Theoretical insights and practical considerations allow for manoeuvring between a straightforward and restrictive implementation of mean-variance analysis on one side and 'practical' *ad hoc* approaches on the other. Below, we will touch again on this issue.

Second, we explicitly recognize the limitations to the investor's **perceptions**. Hence, we do assume neither perfect insight in the real world nor the availability of perfect information about the joint distribution of future investment returns. Instead, the approach allows the use of all partial information.[66] We do assume, however, that the investor perceives a set of attributes which (from his point of view) adequately characterize the securities in the opportunity set. In this respect it is important to recognize that an attribute set has no all-time validity; it is the result of an active and continuing search process, during which an investor learns and accordingly may adjust the set.

[66] Cf. section 1.3.

Third, we deliberately avoid making restrictive assumptions with respect to the **preference structure** of the investor. Given a set of perceived security attributes, we assume that the investor strives to attain various goals. These goals are related to (a number of) these perceived attributes. Depending on the investor's insights and preferences, the relative importance of each of these goals may vary. In this respect, we allow for multiple and possibly conflicting goals. In addition, we neither require an explicit representation or specification of the investor's preference functional, nor an explicit quantitative representation or specification of the trade-offs among conflicting goals.

Accepting these weak behavioral assumptions on the side of investor's perceptions and preferences removes much of the rigidity and limitations on the part of both preference analysis and security analysis. Indeed, the approach is tailored to multifarious preferences on the side of the investor and to available information with respect to the investment opportunities. At the same time, however, it necessarily implies an extensive burden on the process of portfolio selection in terms of increased complexity. Hence, the generality of the approach calls for **systematical search procedures** in order to efficiently confront feasible portfolios with the investor's desires. The application of flexible optimization techniques is the fourth characteristic of the approach.

In order to give practical content to a multi-attribute approach to portfolio selection, we therefore adress the following issues in our study:

(i) **a multi-attribute representation of securities:**
how can (portfolios of) financial securities adequately be represented by scores on a limited number of attributes? This implies a detailed security analysis.
In this respect, we make a distinction between two classes of security attributes: *(a)* direct return related attributes and *(b)* attributes that may only be indirectly related to future investment returns.

(ia) **direct return related attributes:**
the classical assumption is that risky investment alternatives can be fully characterized by the (joint) probability distribution of their returns. The locus of a portfolio's distribution is a measure for its 'expected return'. As a portfolio's 'riskiness' is directly related to the shape of the associated probability distribution, how can we **model** this investment risk? Our goal is to derive measures of risk on the basis of a limited number of features or characteristics of a joint probability distribution. The risk measures result from conditioning security returns on multiple state variables or factors. In that way, a linkage is provided

between the joint distributions of the factors and the security returns. By this mechanism of multivariate conditioning, risk is decomposed and a multi-dimensional risk concept is implied. This multi-dimensional conditional risk concept acts both as a simplifying distribution assumption and as a simplifying preference assumption. From the side of the preference analysis, a multi-dimensional risk concept is attractive since risk can be decomposed depending on the specific decision environment of the investor. Furthermore, by conditioning on appealing variables in the economic environment, the security analysis is enhanced, thus enlarging the economic transparancy of the investment decision. From the side of portfolio analysis and portfolio selection, finally, at least part of the perceived joint distribution of security returns is represented by the multi-dimensional risk measures. This reduces the combinatoric problem of confronting preferences with beliefs.

(ib) **other attributes:**
in addition to these risk attributes, we leave room for the incorporation of other attributes that in practice are considered as relevant for some prototypical investors. These attributes were covered in section 1.4.

The multi-attribute representation of securities according to *(ia)* and *(ib)* implies that securities with the same bundle of characteristics are perfect substitutes for the investor.

(ii) **multi-attribute portfolio selection:**
given the relevant attributes of the individual investment alternatives, how can the scores on these attributes be aggregated to the portfolio level and how can an investor **select** an optimal (satisficing) portfolio, i.e. a portfolio that suits his tastes best and fits best in his personal context?
On a less ambitious level, one could consider the elimination of dominated (i.e. 'bad') portfolios, implying a partial ordering of the choice alternatives.

A caveat

The main ideas in the direction of formulating a general approach to conditional-normative portfolio selection problems were developed quite some time ago.[67] Since the mid-eighties, the topic enjoyed widespread attention and increased popularity, especially were the link is made with the (positive) Arbitrage Pricing Theory. However, as many developments and practical applications take place in investment

[67] See Hallerbach & Spronk [1986]. The general ideas were further accentuated in Hallerbach & Spronk [1991,1992].

practice, the ins and outs are hidden from published research. Obviously, the (paramount) commercial interests form one reason for this. Another reason is that many empirical relationships and findings, judged by their degree of significance, do not meet accepted academic standards, but still can provide useful information for applications in investment practice.[68]

As a result, a **unifying** framework in which various approaches can be embedded and which provides (where possible) theoretical underpinnings to the various aspects, is dearly missed. So aside from giving practical content to a general (and in our opinion both realistic and tractable) approach to portfolio selection, we aim to provide some theoretical justification by embedding the approach in financial theory. The contribution of this study, then, is not developing and presenting the ultimate 'practical theory' as a panacea for portfolio selection but instead, more modest, combining several building blocks in the overall puzzle of the portfolio selection process.

In the relatively short period of writing this study (but especially in the long contemplative period before), we repeatedly discovered that many useful ideas and precious insights, presented in the recent literature, were already expressed in the 'early' literature. This early literature sometimes even extended before the 'official' inception of 'modern' portfolio theory. This not only enforces the well-known fact that progress is a result of the efforts of many; it also stresses the fact that some extent of historical regard, though time consuming, effectively stimulates **coherence** in the development and extension of the body of ideas.

As a telling example of the 'back to the future' phenomenon, we became aware that we are at this moment in fact celebrating the twentieth birthday of "a multiattribute approach" to portfolio selection (Smith [1974, p.53]). Smith's seminal contribution to a multi-attribute representation of securities in part mirrors our view on multifarious security analysis and futher strengthened our belief that multi-attribute portfolio selection calls for an **integrated** approach to multifarious security analysis, non-restrictive preference analysis and flexible portfolio analysis.

Outline of the study

The representation and management of investment risk is a central topic in this study. The multi-dimensional risk concept arises from assumptions regarding the behavior of security returns in relation to the behavior of other variables. By specifying a relationship between

[68] See also chapter four, section 4.1, on this point.
Relating to both the issues of commercial interests and academic standards, Stephen Ross remarked: "I keep my academic work carefully separated from my commercial and consultancy work" (personal communication).

security returns and these state variables or factors, a return generating process is assumed. In the literature, several of these return generating processes are proposed. These models differ in the specific types of underlying assumptions, for example with respect to the incorporated generating variables. Chapter two provides an overview and a typology of these models.

Chapter three starts from the observation that generating security returns from underlying factors and conditioning security returns on these factors are two sides of the same coin. By assuming price functions that relate security prices to a set of state variables, a multi-factor representation of security returns is derived and its characteristics are investigated. Given a multi-factor model of security returns, we then point out its role as a simplifying distribution assumption, both in the context of the security analysis as in the context of portfolio analysis. Finally, we discuss the twofold role of multi-factor models in the context of preference analysis. Firstly, we address the question how the sensitivity coefficients, linking the security returns to the factors, can be interpreted as risk measures. Secondly, we substitute a present value function for the hitherto abstract price function. A present value function explicitly identifies a financial security as a claim on a specific but uncertain pattern of future cash flows. By relating future cash flows from securities to economic variables, we show how economic content can be given to multi-factor models.

Chapter four is devoted to several empirical aspects of multi-factor models and illustrates an important part of the security analysis stage. We use a sample of 36 stocks from the Amsterdam stock exchange, whose returns span the period 1970:1 through 1994:5. In the light of our theoretical analysis in chapter three, special attention is paid to estimating interest rate sensitivities. Next, the model is extended to a multivariate context by considering additional (financial-) economic variables.

As the copestone of the decision process, the multi-attribute portfolio selection is treated in chapter five. The outputs of a security analysis, for example as described in section 1.4 and chapter three, serve as inputs for the portfolio analysis. In this view, when buying a security, an investor is actually buying an exposure to various attributes. Likewise, when composing a portfolio, an investor is actually composing an exposure to the various attributes. These exposures are measured in terms of scores on the attributes. The first issue in portfolio analysis, then, is aggregating the security attribute scores to the portfolio level. Given the non-linear nature of some attributes, this is not at all a trivial step. The selection of an optimal or a satisficing portfolio entails the combination of investor preferences and portfolio attribute scores. Because of the weak assumptions and the flexibility allowed on the field of investor preferences, standard portfolio selection approaches are inadequate. We opt for an interactive systematic search procedure that even in the most general context allows for an effective and efficient scan of feasible, non-dominated portfolios. Although an application of the multi-attribute approach depends on the investor's specific decision environment (including his goals, tastes and restrictions), we discuss some general applications. These applications range from mere portfolio analysis to the actual selection of an optimal portfolio, and from defensive strategies to aggressive strategies.

Chapter six evaluates the multi-attribute approach to portfolio selection and contains our conclusions and a summary. We realize that the selection of the topics from the overall framework, as discussed in this study, is somewhat biased towards personal interests and personal experience. Still, we tried to maintain some 'equilibrium', even in this conditional-normative setting. Furthermore, we only dented the surface of a comprehensive and truly multifarious topic. In an attempt to offer some compensation for this, we furnish suggestions for lines of fruitful future research. We hope that the material covered by this study provides a stimulus to fill in the missing parts.