PERFORMANCE MEASUREMENT AND EVALUATION

Auke Plantinga

University of Groningen

26 September 2007

CHAPTER 10. PERFORMANCE MEASUREMENT AND EVALUATION[1]

Auke Plantinga
Faculty of Business and Economics
Department of Finance
University of Groningen
The Netherlands

10.1 Introduction

This chapter discusses methods and techniques for measuring and evaluating performance for the purpose of controlling the investment process. However, many of the methods discussed in this chapter are also used in communicating investment performance between the investment management company and it's (potential) customers. Therefore, performance measurements also play an important role in the competition between investments management companies. Substantial evidence from the net sales of mutual funds shows that investors buy mutual funds with good past performance records although they fail to sell funds with bad past performance[2].

This chapter provides an overview of commonly used methods for measuring and evaluating the performance of investment management processes. In section 10.2 we discuss methods of calculating the returns of investment portfolios. We focus on the precision of such measures and the practical considerations in choosing a measure. In section 10.3 we discuss methods for evaluating performance, which are base on standard measures of risk such as standard deviation and Beta. In section 10.4 we present several alternative performance measures, which are based on measures of downside risk.

If the performance of a portfolio deviates from the benchmark, it is desirable to figure out what the causes of this deviation are. Performance attribution is the process of allocating the deviations from the benchmark performance to specific causes. In section 10.5 and 10.6, we discus performance attribution methods, which aim to link the outcome of the evaluation to specific parts of the investment management process. Section 10.5 deals with methods that use information on the portfolio composition in the analysis. Section 10.6 deals with methods that rely solely on the time series of returns of the portfolio. Finally, in section 10.7, we relate the performance of a portfolio to well-defined investment objectives that are related to the liabilities. These liability-driven attribution models are based on customized benchmarks.

---

[2] See for example, Chevalier and Ellison (1997), Sirri and Tufano (1999), and Del Guercio and Tkac (2002).

10.2 Measuring returns

10.2.1 Introduction

The total rate of returns of a portfolio is the relative change in the value of the portfolio including cash returns such as dividends and coupon payments. Total rate of return includes both realized and unrealized capital gains. The total return of a portfolio is defined as:

$$R = \frac{P_1 - P_0 - C}{P_0}, \qquad (10.1)$$

where P_0 and P_1 represent the market value of the portfolio at the start and end of the evaluation period and C represents the net cash inflow (funding payments minus withdrawals) in the portfolio. In order to calculate the return of a portfolio, it is important to have a clear definition of the portfolio with respect to the inclusion or exclusion of cash. Some investors prefer to calculate the return of the portfolio inclusive of the cash position associated with the management of the portfolio.

Consider the following example of a portfolio that includes a cash position of 1000 at the beginning. Suppose that this portfolio is invested in 100 shares of Unilever stocks with a market price of 100 at the beginning of the period and 110 at the end. In addition, suppose that Unilever paid a cash dividend at the end of the period of 2 per share and the investor added another 500 to the portfolio. Therefore, the market value at the beginning is 11,000 and the market value at the end is 12,700. The total return on this portfolio is now equal to (12,700 – 11,000 – 500)/11,000 = 10.91%.

Alternatively, the return of the portfolio exclusive of the cash position is based on a beginning value of 10,000 and an ending value of 11,000 and a cash outflow of 200. This results in a return of (11,000-10,000+200)/10,000 = 12%.

For reporting purposes, it is convenient to summarize returns realized in subsequent periods. Returns are usually expressed as the annualized average return, using either arithmetic or a geometric return. An important issue it the way to handle intermediate cash flows to the portfolio, using either a time weighted or a money weighted return measure.

10.2.2 Arithmetic and geometric return

The difference between geometric and arithmetic return is important in measuring the performance of investment portfolios. Geometric return is defined as:

$$MGR = \left[\prod_{t=1}^{T}(1+R_t)\right]^{1/T} - 1 = \left[\frac{P_T}{P_{T\text{-}1}} . \frac{P_{T\text{-}1}}{P_{T\text{-}2}} . . . . \frac{P_2}{P_1}\right]^{1/T} -1 \;=\; \left[\frac{P_T}{P_1}\right]^{1/T} -1 \qquad (10.2)$$

From equation 10.2 it becomes clear that the geometric return is not affected by the price path leading to the final value P_T: the outcome of the geometric return is determined by the start and end value of the portfolio. For this reason, the geometric return is a suitable measure for an investor who wants to maximize his wealth at time T.

The (unweighted) arithmetic mean return is defined as:

$$RGR = \frac{1}{T} \sum_{t=1}^{T} R_t \tag{10.3}$$

The arithmetic mean is a 1st order approximation based on a Taylor series expansion. A 2nd order approximation is:

$$MGR \approx \frac{1}{T} \sum_{t=1}^{T} R_t - \frac{1}{2}\sigma^2, \tag{10.4}$$

where σ is the standard deviation of returns in individual sub periods.

Equation 10.4 implies that the geometric average is always smaller than the arithmetic average. Assuming that it is safe to ignore higher order terms from the Taylor series expansion, the difference between the arithmetic average return and the geometric average return is determined by the standard deviation. Consequently, an investor selecting risky alternatives based on arithmetic averages will choose the alternative with the highest standard deviation if both alternatives have the same terminal value.

The arithmetic average can be seen as an upward biased approximation of the geometric average. The following example is a clear illustration of this. Suppose that a stock returns +50% in year 1 and –50% in year 2. The arithmetic average return over both periods is equal to 0%. However, this is a rather optimistic representation of the facts. Starting with an investment of € 1000, the investment grows to € 1500 at the end of period 1 and ends with only € 750 at the end of period 2. Effectively, the investor lost money, a fact that will be reflected in the geometric average mean return of –13.4%.

Although the geometric average mean appears to be the superior measure, it is also associated with some deficiencies. For example, section 10.4 discusses the problems with the use of geometric average return in attribution models. Another problem is that the normality assumption no longer holds for the distribution of geometric average. A solution for both problems is the use of continuously compounded return, which we denote with the lowercase symbol r:

$$r = \ln\left(\frac{P_1}{P_0}\right) \tag{10.5}$$

The average continuously compounded return over several sub periods is equal to:

$$CGR_{1,T} = \frac{1}{T} \sum_{t=1}^{T} r_t \tag{10.6}$$

where R_t is the return over period t calculated according to equation 10.5. The value of an invested of € 1 at time T is equal to:

$$V_T = e^{T\ CGR_{1,T}} \qquad (10.7)$$

Similar to the geometric average return, the average continuously compounded return is determined by the terminal value of the investment.

10.2.3 Time weighted and money weighted return

The difference between time weighted return and money weighted return is determined by the way both measures handle intermediate cash flows. A time weighted return measure accounts for the precise amount and timing of the intermediate cash flows, whereas the money weighted return measure is based on assumptions regarding amount and timing. Although time weighted return usually yields the most accurate return estimates, practical considerations may be in favor of a money weighted return measure.

Time weighted return

Time weighted return requires the evaluator to determine the market value of the portfolio each time a cash flow occurs. Next, the return on each time frame between two consecutive cash flows can be calculated and finally the evaluator can calculate the geometric average return over the different time frames.

Suppose that a cash flow occurs at time t, P_t is the market value of the portfolio an instant before time t, and P^*_{t-1} is the market value just after the occurrence of the previous cash flow at $t-1$. Then, the return of the portfolio over the time frame starting at t -1 and ending at t is:

$$R_t = \frac{P_t - P^*_{t-1}}{P^*_{t-1}} \qquad (10.8)$$

The annualized time weighted return over the period from t = 0 to T is equal to the geometric average of the returns of the individual time frames[3].

$$TWR_{0,T} = \left[(1+R_1)\,(1+R_2)\ \ldots\ (1+R_T)\right]^{1/T} - 1 \qquad (10.9)$$

The following example illustrates the use of time-weighted return. Consider a portfolio with value € 100 at t = 0 that receives a funding cash flow of € 10 at t = 0.5. An instant before t = 0.5, the portfolio has a market value of € 105. At t = 1 the portfolio has a market value of € 120.75. Based on these numbers, we can calculate the time weighted return over the individual time frames as:

t = 0 .. 0.5	$R_{0.5}$ =	(105 - 100) / 100	= 5%
t = 0.5 .. 1	R_1 =	(120.75 - [105 +10]) / (105 + 10)	= 5%

The annualized time weighted return over the period from t = 0 to t = 1 is:

t = 0 .. 1	$TWR_{0,1}$ = (1.05) * (1.05) -1	= 10.25 %

[3] In practice, time weighted return is usually based on discrete compounding. However, it is also possible to calculate time weighted return using continuos compounding.

Money weighted return

Money weighted return requires market values of the portfolio at the start and end of the evaluation period: there are no market values required at other moments. Alternatively, assumptions are made regarding the timing of the cash flows and the return path during the evaluation period. Money weighted return measures are convenient in a portfolio with many cash flows, since they avoid frequent valuations of the portfolio.

A well-known example of a money weighted return measure is the so-called Dietz algorithm[4]. The assumptions in this algorithm are that the cash flows occur in the middle of the evaluation period at $t = 0.5$ and that the cash flow is reinvested at a return equal to the return of the portfolio, which is constant during the evaluation period. Given these assumptions, the terminal value of the portfolio is:

$$P_T = P_0\left(1+ MWR_{0,T}\right) + C\left(1+ 0{,}5\ MWR_{0,T}\right) \tag{10.10}$$

This can be rewritten as:

$$MWR_{0,T} = \frac{P_T - P_0 - C}{P_0 + 0.5C}, \tag{10.11}$$

where P_0 and P_T are the market value of the portfolio at the begin and end of the evaluation period, and C is the net cash flow during the evaluation period.

We can also calculate the money weighted return for the example introduced for explaining the time weighted return:

$MWR_{0,1}$ = (120.75 - [100 + 10])/(100 + 5) = 10.24 %

This example seems to suggest that the difference between time weighted return and money weighted return is very small. The small difference in this example is due to the fact that the assumptions made in the money weighted return are consistent with reality: cash flows indeed occur at the middle of the evaluation period, and the return in the first half of the period is exactly equal to the return in the second half. The remaining difference is due to fact that time weighted return is based on geometric average and the Dietz algorithm is based on arithmetic average.

The Banker's Administrative Institute proposed a variant of the Dietz algorithm that uses compounded interest calculations. This so-called BAI algorithm calculates the return by solving MWR_b from the equation below:

$$P_T = P_0\left(1+MWR_b\right) + C\left(1+MWR_b\right)^{0{,}5} \tag{10.12}$$

Recalculation of our example using the BAI-algorithm yields exactly the same outcome as the time weighted return based on geometric average.

Table 10.1 provides some intuition on the impact of deviations from the assumptions for the outcomes of money weighted return measures. The table contains the outcomes of 4 examples of return calculations over an evaluation period consisting of two sub periods. The first column shows the example identifier, the second and third column provide the returns for the sub periods. The start value of the portfolio is 100 in all 4 examples, and size of the cash flow is 10. The fourth column gives the terminal market value of the portfolio, and the fifth column gives the time *w* of the cash flow. The last three columns present the outcomes for the time weighted return, the Dietz algorithm and the BAI algorithm.

Table 10.1: Examples with time and money weighted return calculations

Example	$R(0,0.5)$	$R(0.5,1)$	Terminal value	*w*	*TWR*	(Dietz)	(BAI)
1	10%	2%	122.44	0.5	12.20%	11.82%	11.83%
2	2%	10%	123.20	0.5	12.20%	12.57%	12.59%
3	5%	5%	121.01	0.25	10.25%	10.49%	10.50%
4	5%	5%	120.50	0.75	10.25%	10.00%	10.01%

Examples 1 and 2 differ with respect to the order of the return path. In example 1, the return is high in the first sub period and low in the second sub period, whereas example 2 shows a reversed scenario. Comparing the outcomes of examples 1 and 2 shows that example 2 gives the highest terminal market value, which is due to the fact that the cash flow can be invested at a higher return. The time weighted return measure is the most appropriate measure for a portfolio manager who cannot control the timing (and size) of the cash flows. The table also shows that the Dietz measure underestimates the time weighted return if the return in the first sub period is higher than in the second sub period, and is an overestimation in the reverse case.

Examples 3 and 4 show the impact of deviations from the timing assumption The Dietz measure overestimates the return if the cash flow occurs earlier than assumed. If the cash flow occurs later than assumed, the Dietz measure is an underestimation of return. The modified Dietz measure allows for a more realistic assumption on the time of the cash flows, by choosing a value τ that relates to the actual average timing of the cash flows:

$$R_{0,T} = \frac{P_T - P_0 - C}{P_0 + \tau C} \tag{10.13}$$

where τ represents the average time a cash flow is present (or absent) in the portfolio. This average time is measured as a fraction of the length of the evaluation period. For example, the return according to the modified Dietz algorithm in example 3 can be calculated as follows. Since the cash

[4] See Dietz (1968). Another well-known example of a money weighted return measure is the internal rate of return, also known as the ‘yield to maturity’.

flow occurred at t = 0,25, the cash flow is 75% of the evaluation period present in the portfolio. The modified Dietz return is no equal to:

$R_{0,1}$ = (121.01 - 100 - 10)/(100+7.5) = 10.24%,

which is very close to the geometric average return.

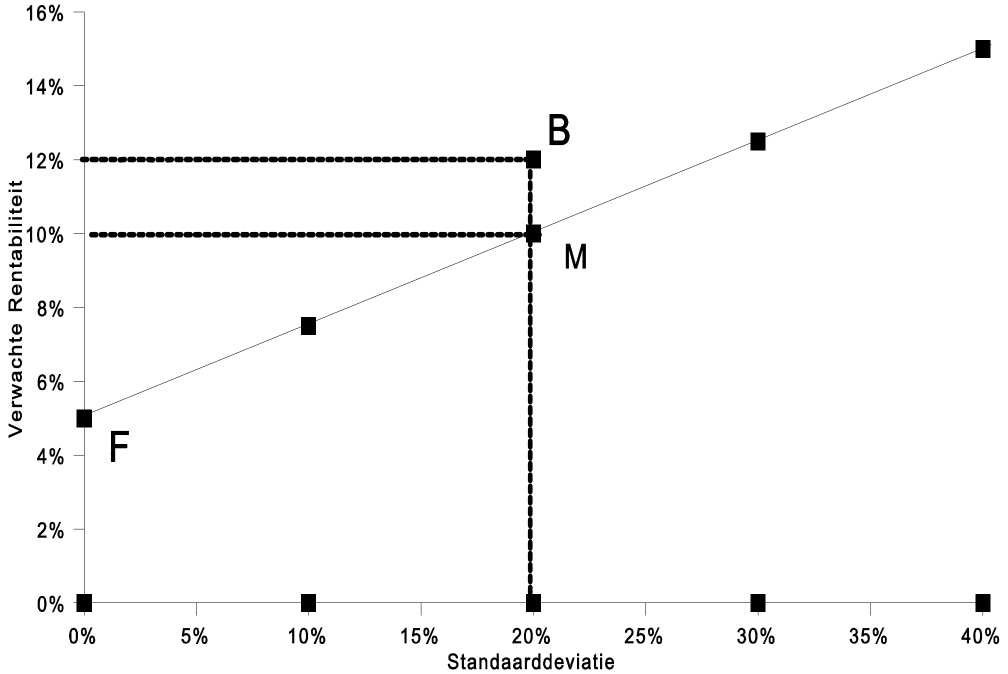
10.3 Evaluating portfolio returns

Having calculated the total rate of return on a portfolio, the next step is to evaluate the return and identify whether or not the performance is satisfactory. Usually, this is accomplished by translating the investment objective into a benchmark portfolio. In the academic literature there has been a lot of emphasis on the difference between the risk of the portfolio and the benchmark. First, we discuss the methods developed by Jensen (1968), Treynor (1966), and Sharpe (1966), which focus on the differences in the levels of absolute risk of the portfolio and the benchmark. Next, we discuss the so-called relative measures of performance, which measure risk relative to the benchmark.

10.3.1 Jensen, Treynor and Sharpe performance measures

Modern portfolio theory is an important source of inspiration for constructing risk-adjusted performance measures. Several measures have been derived from the Capital Asset Pricing Model (CAPM). The simple version of the CAPM assumes that investors have homogeneous expectations, which implies that they share equal expectations regarding risk and return. One of the main predictions of the CAPM is that investors choose portfolios from the so-called capital market line (CML), which plots expected return against risk. Portfolios on the CML maximize the expected return for a given level of risk, where risk is expressed as the standard deviation of returns. In the presence of a risk free asset, the CML is linear. The CML intersects the vertical axis at the risk free rate and is tangent at the efficient set of risky assets. The tangency point of risky assets is also called the market portfolio. Figure 10.1 is an example of a CML with a risk free rate of 5% and a market portfolio with an expected return of 10%. **F** represents the risk free fund, and **M** represents the market portfolio.

Figure 10.1: The capital market line



The Sharpe ratio

In the simple version of the CAPM, investors can lend or borrow risk free at the same interest rate. Therefore, an investor views all investment opportunities positioned on the CML as equally attractive, since risk free lending and borrowing enables him to replicate any position on the CML from another position on the CML. The only way to improve performance is to increase the angle of the CML.

The angle of the CML is also known as the Sharpe ratio which is defined as the ratio of expected return E(R_p) in excess of the risk free rate R_f and the standard deviation σ_p of the portfolio:

$$S = \frac{E\lfloor R_p \rfloor - R_f}{\sigma_p} \qquad (10.14)$$

With homogeneous expectations, it is not possible to find portfolios with a Sharpe ratio exceeding that of the market portfolio. However, active management implies heterogeneous expectations, which means that some portfolio managers may be able to construct portfolio with a Sharpe ratio that exceeds the market portfolio. For example, portfolio **B** in figure 10.1 has a Sharpe ratio equal to (12% - 5%) / 20% = 0.35 that exceeds the Sharpe ratio of the market portfolio (10% - 5%) / 20% = 0.25.

It is important to notice that the Sharpe ratio is suitable for evaluating the performance of the portfolio of all assets of an investor. In empirical applications, this is not always true. Many portfolios are not designed to represent the total asset holdings of an investor. Portfolio managers can specialize in specific markets, and investors should consider such portfolios only as part of a larger portfolio. A Sharpe ratio of 0.25 for a portfolio of South-American stocks in comparison with a Sharpe ratio of 0.30 for European stocks does not justify the conclusion that the South American portfolio is not a suitable investment. A more relevant criterion is to consider the Sharpe ratio of a portfolio of both European and South American stocks.

Despite this limitation, the Sharpe ratio is frequently used in performance reporting. In order to judge the Sharpe ratio of a particular portfolio, it is useful to have the distribution of Sharpe ratios of alternative portfolios. Table 10.2 presents a distribution of Sharpe ratios for 145 Dutch mutual funds calculated using monthly total returns over the period 1994-1998. The risk free rate is equal to the average money market return in this period. Most mutual funds have a Sharpe ratio between 0 and 0.34 and there are no funds with a Sharpe ratio over 0.68.

Table 10.2: Frequency distribution of Sharpe ratios of Dutch mutual funds over 1994-1998.

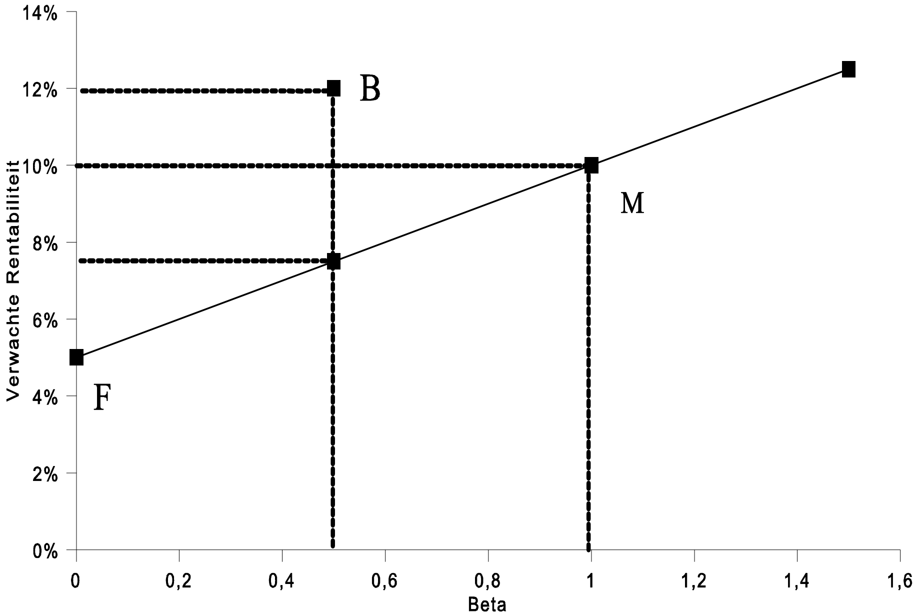
Interval	Relative frequency
$x \leq -2.39$	2.8%
$-2.39 < x \leq -2.00$	0.7%
$-2.00 < x \leq -1.67$	0.0%
$-1.67 < x \leq -1.30$	1.4%
$-1.30 < x \leq -0.68$	2.1%
$-0.68 < x \leq 0.00$	9.0%
$0.00 < x \leq 0.34$	57.2%
$0.34 < x \leq 0.68$	26.8%
$0.68 < x \leq$	0.0%

Alternatively, the Sharpe ratio can be interpreted as a t-test for the hypothesis that the return on the portfolio is equal to the risk free rate. Since we do not observe Sharpe ratios over 1.96, we cannot reject this hypothesis for any of the mutual funds.

Jensen's alpha

The Sharpe ratio uses the capital market line as a reference point for evaluating investment portfolios. Portfolios on the CML also minimize the risk for a given level of return, which implies that all diversifiable risk has been eliminated and that they face only systematic risk. Since it is very likely that portfolio managers specialize into specific markets and industries, it is very likely that their portfolio contain a considerable amount of diversifiable risk. Since ultimate risk reduction is not an objective of those portfolio managers, it is reasonable to use a performance measure that gives meaningful evaluations for inefficient portfolios. In the context of the CAPM, a suitable measure can be derived from the security market line (SML). The security market line plots the expected return against the beta of the portfolio, the index of systematic risk. In figure 10.2 we present an example of the SML consistent with the example used for figure 10.1. Assuming equilibrium, the security market line must hold for all portfolios, both efficient and inefficient.

Figure 10.2: The security market line



With heterogeneous expectations, some investors may have an informational advantage resulting in return expectations that differ from the market's expectations. Eventually, the informational advantage allows them to realize average returns above the security market line. For example, portfolio B has an expected return of 12% that exceeds the market expectation, which is 7.5% for a portfolio with a Beta of 0.5. The difference of 4.5% is Jensen's alpha, which is defined as the difference between the realized return and the expected return given the beta of the portfolio.

Usually, Jensen's alpha is measured with the following regression equation:

$$R_p - R_f = \alpha_p + \beta_p (R_m - R_f) + \varepsilon, \qquad (10.15)$$

where R_m is the return on the market, β_p is the slope of the regression equation (the index of systematic risk), and α_p is Jensen's alpha. Using this equation we can calculate Jensen's alpha for portfolio B as

$$12\% - 5\% = \alpha_p + 0{,}5\ (10\% - 5\%)$$
$$\alpha_p = 4.5\%$$

A convenient property of Jensen's alpha is that is presented as a return number. It can also be interpreted as the difference between the actual portfolio and a customized benchmark portfolio. For example, since portfolio **B** presented in figure 10.1 and 10.2 has a Beta of 0.5, its corresponding benchmark is an investment of 50% in the market portfolio and 50% in the risk free asset. The return on such a portfolio is equal to 0.5 * 10% + 0.5 * 5% =7.5%. The difference between the return on the actual and the benchmark portfolio is Jensen's alpha.

The Treynor ratio

The Treynor ratio is also derived from the SML. The Treynor ratio is the slope of the line connecting the actual portfolio with the risk free rate. If this slope exceeds that of the security market line, the portfolio has added value. The Treynor ratio is a measure of the return per unit systematic risk, and is defined as:

$$T = \frac{R_p - R_f}{\beta_p} \tag{10.16}$$

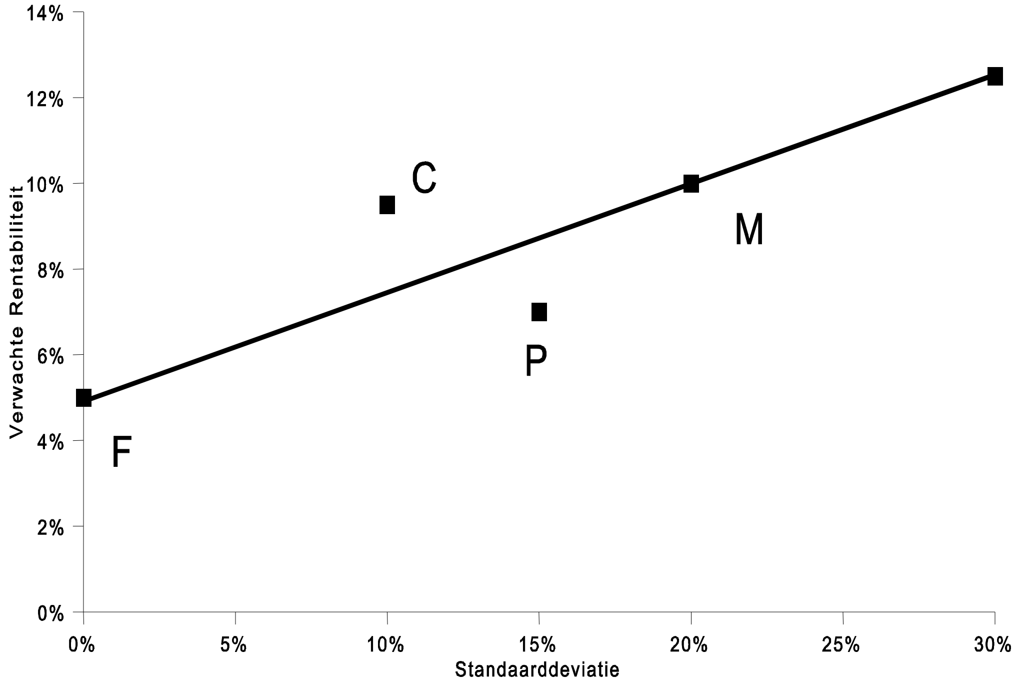
The outcome of the Treynor ratio is directly comparable with the equity premium, which is usually measured as the average of the difference between the market return and the risk free rate. If the Treynor ratio exceeds this equity premium, the portfolio manager has added value relative to a passive manager.

Problems with the use of JTS-measures

Normality and stationarity of the return distribution are important assumptions associated with Jensen's alpha, the Treynor and the Sharpe ratio. In particular the assumption of normality may cause problems in case of portfolios managed with derivative instruments[5]. Bookstaber en Clarke (1985) used a simulation study to show that an investor without any forecasting skills can generate a Sharpe ratio in excess of the market's Sharpe ratio by buying call options on the market portfolio. They consider three strategies, which are plotted in figure 10.3.

[5] Jensen's alpha, the Treynor and Sharpe ratio are also less suitable for fixed income portfolios. An appropriate method of analysis for fixed income portfolio is presented in section 10.4.3.

Figure 10.3: The capital market line with options



The first strategy is to invest 100% in the market portfolio. The second strategy is a portfolio **C** with 50% invested in the stock market portfolio and 50% in a long call option on the markt portfolio. The outcomes of the simulation study shows that this strategy results in a position above the capital market line, which implies that the Sharpe ratio of the portfolio exceeds that of the market portfolio. The third strategy is a portfolio **P** that invests 50% in the market portfolio and 50% in a long put option on the market portfolio. Figure 10.3 shows that this strategy actually plots below the capital market line.

10.3.2 Relative performance measurement

The Jensen, Treynor, and Sharpe measures of performance have been utilized extensively in both theory and practice. However, practitioners seem to favor relative performance measures. In particular, they focus on the performance relative to a benchmark without adjustments for systematic risk. Alternatively, they utilize the tracking error, which is the risk relative to the benchmark. They combine the relative return and relative risk measure into the so-called information ratio, which is the equivalent of the Sharpe ratio in relative risk and return space.

The information ratio

The information ratio is a widely used measure for portfolio managers specialized in a particular sector or market. Those portfolio managers are usually evaluated against a previously determined representative market index. Since the investor made the choice for the particular market segment and its benchmark, the total return of the portfolio is not an issue in evaluating the performance. Alternatively, the focus is on performance relative to the index. The information ratio is defined as:

$$IR = \frac{\mathrm{E}[R_p] - \mathrm{E}[R_m]}{\sigma_{te}} \tag{10.17}$$

where R_p is the return on the portfolio, R_m is the return on a representative market index and σ_{te} is the tracking error. Tracking error can be calculated using a time series of returns:

$$\sigma_{te} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(R_{p,t} - R_{m,t}\right)^2} \tag{10.18}$$

where $R_{p,t}$ is the portfolio return in period t, $R_{m,t}$ is the market index return in period t and T is the total number of observed periods.

The information ratio can be interpreted as the t-test associated with the hypothesis that the returns on the portfolio do not significantly deviate from the market index. An information ratio larger than 1.96 implies that a portfolio manager has a 95% probability of beating the market index in any period.

As an illustration, table 10.3 shows information ratios for 4 mutual funds specialized in Dutch stocks. The bottom row presents data on the MSCI Netherlands index, which will be used as a benchmark for these funds. The second column presents the average monthly return. The ABN Amro Netherlands Fund and the ING Bank Dutch Fund offer the highest average monthly returns, with AXA Aandelen Nederland taking a middle position and the Delta Lloyd Top-20 Nederland Fonds closing ranks.

Table 10.3: Example of mutual fund evaluation using the Sharpe and information ratio

	E[R]	σ	σ_{te}	Sharpe	IR
ABN Amro Netherlands Fund	2.23%	5.02%	1.73%	0.37	0.24
AXA Aandelen Nederland	1.95%	4.56%	2.47%	0.35	0.05
Delta Lloyd Top-20 Nederland	1.52%	4.08%	3.14%	0.28	-0.10
ING Bank Dutch Fund	2.24%	4.73%	1.41%	0.40	0.30
MSCI Netherlands.	1.82%	4.93%	0.00%	0.30	0.00

Calculations are based on data from Standard & Poor's Micropal

The third and fourth column presents respectively the standard deviation and the tracking error of the monthly returns. The numbers show that ABN Amro Netherlands Fund is the most risky fund in terms of total risk, although it's tracking error ranks lower as compared to the other funds. There appears to be an association between the total level of risk and average return, since funds with a high total risk also have a high average return. It is also clear that the tracking error is not related to the level of total risk. Despite the fact that the risk measures result in different outcomes, the ranking of performance based on the Sharpe ratio and the information ratio is the same. However, the information ratio results in larger differences between the funds in terms of the score on the performance measure.

10.4 Alternative performance measures

This section discusses several risk-adjusted performance measures that allow for deviations from the traditional mean-variance framework. A major criticism of the mean-variance framework is its reliance on either the assumption of normally distributed returns or on the assumption of quadratic utility.

Several alternative performance measures are based on the concept of downside deviation. Downside deviation is a risk measure that deviates from standard deviation in two ways. First, it defines risk relative to an exogenous reference point. This reference point, which is also called the minimal acceptable rate of return, is used to distinguish 'risk' from 'volatility'. According to Sortino and Van der Meer (1991), realizations above the reference point imply that goals are accomplished and, therefore, are considered 'good volatility'. Realizations below the reference point imply failure to accomplish the goals and should be considered 'bad volatility' or risk. Second, downside risk only accounts for deviations below the reference point, and ignores deviations above the reference point. Downside deviation is defined as:

$$\delta = \sqrt{\tfrac{1}{T}\sum_{t=1}^{T}\left(R_t - R_{mar}\right)^2} \qquad \forall \; R_t < R_{mar}, \tag{10.19}$$

where δ is downside deviation and R_{mar} is the minimal acceptable rate of return.

The Sortino ratio is probably the most well-known measure utilizing downside risk, and it is calculated as follows:

$$Sort = \frac{E[R] - R_{mar}}{\delta} . \quad (10.20)$$

Another measure based donwside risk is the Fouse index, which is described in Sortino and Price (1994):

$$Fouse \quad = E[R] - B\delta^2 , \quad (10.21)$$

where B is a parameter representing the degree of risk aversion of the investor.

The Sortino ratio and the Fouse index rely on the use of expected return and downside risk. Expected return is used as a measure of the potential reward of an investment opportunity. An alternative for using the expected return is the so-called upside potential ratio, which is the probability weighted average of returns above the reference rate. The upside potential ratio was developed by Sortino, Van der Meer, and Plantinga (1999) and is defined as:

$$UPR = \frac{\sum_{t=1}^{T} \iota^{+} \frac{1}{T}\left(R_t - R_{mar}\right)}{\sum_{t=1}^{T} \iota^{-} \frac{1}{T}\left(R_t - R_{mar}\right)^2} \quad (10.22)$$

where T is the number of periods in the sample, R_t is the return of an investment in period t, $\iota^+ = 1$ if $R_t > R_{mar}$, $\iota^+ = 0$ if $R_t \leq R_{mar}$, $\iota^- = 1$ if $R_t \leq R_{mar}$ and $\iota^- = 0$ if $R_t > R_{mar}$. An important advantage of using the upside potential ratio rather than the Sortino ratio or the Fouse index is the consistency in the use of the reference rate for evaluating both profits and losses.

Finally, an important difference between downside risk and standard deviation is the use of an exogenous reference rate versus the mean return. The investor's objective function motivates the choice of the reference rate. As a result, a part of the investor's preference function is introduced into the risk calculation. Therefore, the resulting calculation is only valid for individuals sharing the same reference rate. Investors with different minimal acceptable rates of return will have different rankings.

10.5 Performance attribution

10.5.1 Introduction

Relative performance measures such as the information ratio are used to evaluate performance relative to a benchmark. Performance attribution aims to identify the causes of deviations between the return of the benchmark and the actual portfolio. In this section we discuss performance attribution methods that analyze both the portfolio composition and returns in identifying causes of deviations from the benchmark.

Most methods used in practice are derived from the framework used by Brinson and Fachler (1985). Many institutional investors use this method. The Brinson and Fachler framework is based on a top-down investment management process. Although the original version of this framework is only useful to support a very general analysis, it can be extended quite easily to support more realistic assumptions regarding the special properties of specific asset classes such as fixed income instruments and currency exposure.

In section 10.4.2 we discuss the Brinson & Fachler framework. In the next sections, we discuss the extensions to the simple framework. In section 10.4.3 we discuss how to analyze fixed income securities using the model of Fong, Pearson and Vasicek (1983) and in section 10.4.4 we deal with currency exposure using the method of Singer and Karnosky (1995).

10.5.2 The attribution model of Brinson and Fachler

The framework of Brinson and Fachler (1985) is based on well-known techniques from management accounting aimed at analyzing the differences between budgets and realizations. The framework corresponds to a top-down approach in the investment management process, starting from a general investment plan that describes planned portfolio weights for asset classes and assigns benchmarks to asset classes. The top level of management authorizes this general investment plan. The plan provides room for the lower levels in the organization to deviate from the plan in order to capture changing conditions in financial markets. Many professional investors have monthly or weekly meetings of portfolio managers and researchers to decide on tactical issues regarding deviation from the asset allocation. This decision is called the tactical asset allocation decision. Portfolio managers within asset classes also have different levels of discretion to deviate from their benchmarks. Actual deviations from the benchmark with an asset class are called stock selection decisions.

The analysis is based on four different portfolios:

I. The benchmark portfolio;
II. The stock-selected portfolio;
III. The timing portfolio;
IV. The actual portfolio.

Portfolio **I** is the overall benchmark portfolio, which is derived directly from the general investment plan. Portfolio **I** defines the desired asset allocation also known as the strategic asset allocation and the benchmarks for individual asset classes. Portfolios **II** and **III** are the outcomes of a 'what if' analysis that aim to measure the impact of decisions in isolation from the other decisions.

The return on portfolio **I** is the result of investing the portfolio exactly according to the strategic asset allocation and the benchmarks for the asset classes. The return of this portfolio is defined as:

$$R(I) = \sum_i w_i^p \; R_i^p \, , \qquad (10.23)$$

where w_i^p is the strategic weight of asset class *i* and R_i^p is the return of the benchmark for asset class *i*.

Portfolio **II** measures the return of implementing the stock selection decision, whilst ignoring the tactical asset allocation decision. Let the actual return of an investor in asset class *i* be R_i^a, then the return on portfolio **II** is:

$$R(II) = \sum_i w_i^p \; R_i^a \qquad (10.24)$$

Portfolio **III** measures the return of implementing the tactical asset allocation decision without the stock selection decision. The asset class weights for this portfolio are equal to the actual weights w_i^a, and the securities within an asset class are exactly equal to the benchmark for the asset class. The return on this portfolio is:

$$R(III) = \sum_i w_i^a \; R_i^p \, . \qquad (10.25)$$

Portfolio **IV** is the actual portfolio subject of the analysis. The return on the actual portfolio is:

$$R(IV) = \sum_i w_i^a \; R_i^a \qquad (10.26)$$

The difference between the return of portfolio **I** and portfolio **III** is the so-called 'timing effect', and represents the additional return due to the tactical asset allocation. The return difference between portfolio **I** and portfolio **II** is the so-called 'selection effect' and shows the additional return due to stock selection.

Since the timing effect and selection effect are calculated independently, we also need to capture the joint impact of both decisions, which is called the 'interaction effect'. This effect is calculated as the sum of the return of portfolio **I** and **IV** *minus* the sum of the return of portfolio **II** and **III**. Since the interaction effect cannot be attributed to a single person or department, some practitioners allocate this effect in equal parts to both the timing and selection effect. The effects and their calculations are summarized in table 10.4.

Table 10.4: Components of the attribution model

Source	Calculation	
Timing	R(III) -R(I)	$\Sigma\ (w_i^p - w_i^a)\ R_i^p$
Selection	R(II) -R(I)	$\Sigma\ (R_i^a - R_i^p)\ w_i^p$
Interaction	R(I)-R(II)-R(III) + R(IV)	$\Sigma\ w_i^p\ (R_i^p - R_i^a) + w_i^a\ (R_i^a - R_i^p)$
Total contribution	R(IV) -R(I)	$R_a - R_p$

Suppose for example that an investor utilizes a strategic asset allocation of 50% bonds, 20% domestic stocks and 30% foreign stocks. Furthermore, assume that the actual allocation, the actual returns, and the benchmark returns are presented in table 10.5.

Table 10.5: Actual portfolio weights, benchmark returns and benchmark portfolio weights

	actual allocation	Return benchmark	Return actual portfolio
Bond	30%	8%	7%
Domestic stocks	20%	12%	15%
Foreign stocks	50%	24%*	22%*

* Returns are measured in terms of the reporting currency.

The outcomes of the calculations for this example are given in table 10.6.

Table 10.6: Outcomes of the attribution model

Portfolio	Return		
I	13.6%	Timing	3.2%
II	13.1%	Selection	-0.5%
III	16.8%	Interaction	-0.2%
IV	16.1%	Total contribution	2.5%

This analysis shows that the investor in this example outperformed the overall benchmark by 2.5%, which is mainly due to considerable timing skills and not to selection skills.

Multi-period attribution

It is often desirable to summarize the performance attribution report over multiple periods. For example, many institutional investors generate quarterly attribution reports that they like to summarize on an annual basis. Since returns over multiple periods are usually calculated using a geometric average, problems may arise in summarizing the performance attribution. Table 10.7 provides an illustration of these problems.

Table 10.7: Multi-periodattribution

Period	1st half of year	2nd half of year	overall
I	20.00%	10.00%	32.00%
II	17.00%	12.00%	31.04%
III	24.40%	12.00%	39.33%
IV	21.00%	14.25%	38.24%
Timing	4.40%	2.00%	(a)
Selection	-3.00%	2.00%	(b)
Interaction	-0.40%	0.25%	(c)
Return difference	1.00%	4.25%	(d)

Table 10.7 presents performance attribution data on two periods. Performance presentation standards[6] suggest that the return has to be summarized using a time weighted return measure, which involves the geometric average. In the fourth column of table 10.7 we present the geometric average return for the four portfolios from the attribution models.

The challenge to be faced is how to calculate the summarized outcomes (a), (b), (c), and (d) for the attribution components. An obvious solution is to calculate (a) as the difference betweenportfolio III and I over two periods: 39.33% - 32.00% = 7.33%. Unfortunately, it is not possible to link this number to the values for the timing component for each individual period: summing the individual answers results in a total of 6.4% and a geometric average on an annual basis results in 6.48%. Similar results can be found for the other components of the attribution model.

A solution to this problem is to calculate the return for the four portfolios based on continuously compounded average return, which is a time weighted return measure. In section 10.2.2 we found that the continuously compounded average return over multiple periods is an arithmetic average of the continuously compounded returns over the sub periods.

10.5.3 Portfolios with currency exposure

In this section we discuss the proposal by Singer en Karnosky (1995) to extend the model of Brinson and Fachler in order to deal with currency exposures. A distinct feature of the model is that local returns are analyzed in terms of risk premia. Singer and Karnosky argue that international interest parity conditions imply that currency returns depend on interest rate differences between countries. In particular, the pricing of currency forward contracts is determined to a large extent by interest rate differences. For this reason, they consider the effects of interest rate differences jointly with currency returns.

The return of an internationally diversified portfolio in terms of a base currency equals:

$$R_{bc} = \sum w_i \left(R_{l,i} + \varepsilon_{bc,i}\right), \qquad (10.27)$$

where w_i is the weight of country i as a fraction of the total portfolio value measured in terms of base currency, $R_{l,i}$ is the local currency return of an investment in country i and $\varepsilon_{bc,i}$ the relative value change of the local currency i against the base currency.

Using currency forward contracts to hedge the currency risk completely, results in the following return in terms of base currency:

$$HR_{bc} = \sum w_i \left(R_{l,i} + f_{bc,i}\right) \qquad (10.28)$$

where $f_{bc,i}$ is the base currency return of the currency forward contract.

---

[6] Time weighted return measures are required in order to be consistent with performance presentation standards such as GIPS.

In hindsight, hedging provides the highest return if the realized currency return is lower than the return on the forward contract:

$$f_{bc,i} > \varepsilon_{bc,i} \tag{10.29}$$

According to international interest parity, the return of the forward contract is equal to the difference in interest rates between the foreign country and the base currency:

$$f_{bc,i} = c_{bc} - c_i \tag{10.30}$$

where c_{bc} is the interest rate for the base currency and c_i is the interest rate for the foreign currency i. Therefore, equation 10.26 can be written as:

$$\begin{aligned} c_{bc} - c_i &> \varepsilon_{bc,i} \\ c_{bc} &> c_i + \varepsilon_{bc,i} \end{aligned} \tag{10.31}$$

These equations show that the interest rate differences and the currency returns are related. The return of a hedged foreign investment in terms of the base currency can be rewritten as:

$$HR_{bc,i} = R_{l,i} + f_{bc,\,i} = (r_i - c_i) + c_{bc} \tag{10.32}$$

The return of a non-hedged foreign investment in terms of the base currency is

$$R_{bc,i} = R_{l,i} + \varepsilon_{bc,i} = (r_i - c_i) + (c_i + \varepsilon_{bc,i}) \tag{10.33}$$

In other words, the return of any foreign investment in terms of the base currency can be written as a combination of (1) the local risk premium and (2) the relative change of an investment in the foreign currency. The second component is determined by the foreign interest rate and – if the foreign currency position is unhedged - the relative change in the value of the foreign currency. This decomposition is essential in the attribution model of Singer en Karnosky. Both components can be analyzed in terms of timing and selection attributes resulting in the market attribution for component (1) and the currency attribution for component (2). In table 10.8 we present the attribution framework for the market and the currency attribution.

Table 10.8: Attribution with currency exposure

	Market attribution	Currency attribution
Timing	M(II)-M(I)	C(II)-C(I)
Selection	M(III)-M(I)	C(III)-C(I)
Interaction	M(I)+M(IV)-M(II)-M(III)	C(I)+C(IV)-C(II)-C(III)

Following the method of Brinson and Fachler, each attribution requires 4 portfolios. In table 10.9 we show how to calculate the return on each portfolio.

Table 10.9: Return of the portfolios needed for calculating the attribution model

M(I)	$\sum w_i^p (r_i^p - c_i^p)$

M(II)	$\sum w_i^a (r_i^p - c_i^p)$
M(III)	$\sum w_i^p (r_i^a - c_i^p)$
M(IV)	$\sum w_i^a (r_i^a - c_i^p)$
C(I)	$\sum(w_i^p + h_i^p)(c_i^p + \varepsilon_i^p)$
C(II)	$\sum(w_i^a + h_i^a)(c_i^p + \varepsilon_i^p)$
C(III)	$\sum(w_i^p + h_i^p)(c_i^a + \varepsilon_i^p)$
C(IV)	$\sum(w_i^a + h_i^a)(c_i^a + \varepsilon_i^p)$

10.5.4 Fixed income securities

The nature of a portfolio of fixed income securities justifies a tailor-made evaluation method. For most fixed income instruments, the most important factor driving returns is interest rate risks. Unfortunately, the exposure to this factor is not constant since portfolio managers can change this exposure with a few transactions. Furthermore, the exposure to interest rate risk is declining with the passing of time. Therefore, we need a procedure specific for fixed income instruments for dealing with systematic risk. Fortunately, pricing of fixed income instruments is facilitated by the fact that prices can be approximated adequately with analytical price expressions. This allows us to deal with the interest rate risk exposure.

The price of a plain vanilla government bond is:

$$P = \sum_{t=1}^{T} \frac{C_t}{\left(1+r[0,t]\right)^t} + \frac{H_T}{\left(1+r[0,T]\right)^T} . \qquad (10.34)$$

where C_t is the coupon payment at t, H_T is the face value at maturity T, and $r[0,t]$ is the interest rate for a zero coupon bond originating at t=0 and ending at t.

Fong, Pearson en Vasicek (1983) proposed a method for analyzing the returns of a portfolio of fixed income instruments. The following procedure is based on their model.

Table 10.10: Attribution model for fixed income instruments

<table>
<tr><td>V</td><td>Bond selection</td><td rowspan="2">Credit risk and other factors</td><td rowspan="3">Management</td></tr>
<tr><td>IV</td><td>Sector management</td></tr>
<tr><td>III</td><td>Maturity management</td><td rowspan="2">Interest rate</td></tr>
<tr><td>II
I</td><td>Unexpected return</td><td>External</td></tr>
</table>

Expected return		

In table 10.10 we present the schedule for analyzing the returns of fixed income portfolios. The schedule is based on five additional benchmark portfolios that facilitate the decomposition of the return into external factors and skills of the portfolio manager. The external factors refer to the level and changes in the level of the interest rates. The skills of the portfolio manager reflect the effect of management on the portfolio return. For example, the effect of a rise in interest rates is an external factor since it is not a skill of the manager. However, the anticipation of the manager on this increase in interest rates by decreasing the portfolio duration is a skill.

Alternatively, it is possible to distinguish between the factors that drive the prices of fixed income instruments, such as interest rates, liquidity and credit risk. In the schedule, we limit our attention to interest rates and credit spread, although it is possible to extend the schedule. Consistent with the usual approach in performance attribution models, Fong, Pearson and Vasicek use several auxiliary portfolios to facilitate their analysis.

Portfolio **II** is used to measure the external interest rate effect, which cannot be attributed to specific actions of the portfolio manager. This effect can be decomposed into an expected part (measured by portfolio **I**) and an unexpected part (the difference between portfolio **II** and **I**). The difference between portfolio **III** and portfolio **II** is the return due to maturity management, which measures the added value of anticipating interest rate movements.

The difference between portfolio **III** and **IV** measures the impact of a portfolio manager's choice for specific sectors in the bond market, such as the choice for bonds originated in the financial sector or specific industries. Finally, the difference between portfolio **IV** and **V** refers to the added value of bond picking.

The impact of interest rates on portfolio return

The level of the interest rates and changes in the level drives the impact of interest rates on the portfolio return. In addition, the portfolio manager can anticipate on changes in interest rates by increasing or decreasing the duration of his portfolio.

In order to separate the management effect from the external effect, it is necessary to choose a benchmark that corresponds with the investment objective. This implies that the benchmark has an adequate risk profile, by choosing appropriate levels of duration, convexity or other measures of interest rate risk. The external component is measured by calculating the present value of both the benchmark and the actual portfolio based on the term structure of risk free interest rates. This procedure eliminates the impact of other factors, such as credit risk and liquidity risk.

We attain this goal by deriving two hypothetical portfolios from the benchmark, which we label portfolio **II** (for the benchmark portfolio) and portfolio **III** (for the actual portfolio). We assume that these portfolios invest in government bonds with cash flows exactly equal to those of the original portfolio. The cash flows from both portfolios are discounted with the term structure of risk free interest rates.

Portfolio **I** measures the impact of the level of the interest rates. This portfolio is assumed to be invested at t = 0 in zero coupon bonds with maturity equal to the evaluation horizon. The return on this portfolio is equal to $R(\mathrm{I})=r[0, t]$. Portfolio **I** can also be seen as an investment in the money market.

The value of portfolio **II** is equal to the present value of the cash flows from the benchmark portfolio discounted with the term structure of risk free interest rates:

$$P_t(\mathrm{II}) = \sum_{i=t}^{T} \frac{C_{p,i}}{\left(1 + r[t,i]\right)^{i-t}}. \tag{10.35}$$

where $C_{p,i}$ the cash flow for the benchmark portfolio at time i and $r[t,i]$ the interest rate for a zero coupon originated at time t that matures at time i.

The return of portfolio **II** measures the external effect of the interest environment on the benchmark return. The return of portfolio **II** is also affected by the investor's choice for a benchmark with a particular duration. Consequently, the portfolio manager does not have control over this return. The external effect of the interest environment can be decomposed into an expected component $R(\mathrm{I})$ and an unexpected component $R(\mathrm{II})$ - $R(\mathrm{I})$. The expected component is equal to the rate for a bond with a maturity equal to that of the investment horizon, the unexpected component is equal to the impact of a change in interest rate levels.

The value of portfolio **III** is equal to the present value of the cash flows from the actual portfolio discounted with the term structure of risk free interest rates:

$$P_t(III) = \sum_{i=t}^{T} \frac{C_{a,i}}{\left(1 + r[t,i]\right)^{i-t}}, \tag{10.36}$$

where $C_{a,i}$ denotes the cash flow in the actual portfolio at time i.

The return of portfolio **III** measures the impact of the interest rate environment on the return of the actual portfolio, including its impact on the portfolio manager's decision to choose a maturity structure different from that of the benchmark. Therefore, the difference between the return of portfolio **III** and portfolio **II** is entirely due to the difference in the timing of the cash flows. Consequently, the difference between portfolio **III** and **II** measures the impact of the decision of the portfolio manager to choose a duration different from the benchmark. This difference is labeled as the 'effect of maturity management' or the 'duration effect'.

Example (1)
Suppose that investor **ZZZ** manages a portfolio of fixed income securities. During the past year, the actual return on this portfolio was equal to 15.38 %. The future cash flows of this portfolio and its

corresponding benchmark are presented in table 10.11. In addition, this table presents the term structure of risk free interest rates at *t=0* and *t=1*.

Table 10.11: Cash flows and termstructure for evaluating portfolio ZZZ

	Projected cash flows		Term structure	
Time	Portfolio	Benchmark	*t* = 0	*t* = 1
1		1 000	5.00 %	-
2		1 000	6.00 %	6.50 %
3		1 000	6.50 %	6.00 %
4	1 000	1 000	7.00 %	5.75 %
5	1 000	1 000	7.25 %	5.00%

	Portfolio **III**		Portfolio **II**	
	t=0	t=1	t=0	t=1
Present Value	1 468	1 668	4 138	4 497
Return		13.67 %		8.69 %

Based on this information, the present value of portfolio **II** and **III** can be calculated. The outcomes are given at the bottom rows of table 10.11. In this example, the effect of maturity management is $R(III)$ - $R(II)$ = 13.67% - 8,69% = 4.99%. The external effect of the interest environment is 8.69%, which can be decomposed into an expected component of 5.00% and an unexpected component of 3.69%.

Credit risk

Credit risk is the second important factor in determining the return of a portfolio of fixed income instruments. Credit risk refers to the possibility that the debtor is not able to service the scheduled coupon and redemption payments on a bond. Credit risk is measured by rating agencies such as Fitch, Standard & Poor's en Moody's. Investors require a premium for credit risk that is increasing with the amount of risk. Therefore, a portfolio of corporate bonds is expected to yield a higher return than a portfolio of government bonds with similar maturities. In addition, a portfolio manager can add additional return by anticipating changes in the credit risk of individual corporations or entire sectors of the bond market. The anticipation of changes in the credit risk of individual corporations is called 'bond picking' and the anticipation of sectors is called 'sector management'.

The contribution of sector sector management is calculated by valuing each individual bond at the begin and end date of the evaluation period based on the risk free discount rate and a premium for credit risk equal to that of the sector average. The return of the portfolio based on these valuations is the return on portfolio **IV**. The contribution of sector management is equal to the difference between the return on portfolio **IV** and **III**.

Example 1 (Continued)

Table 10.12: Cash flows and term structures for different credit risk classes

Time	Cash flow	Term structure *t* = 0	Term structure *t* = 1

	A	B	A	B	A	B
1			6.00 %	7.00 %	-	-
2			7.00 %	8.00 %	8.50 %	8.50 %
3			7.50 %	8.50 %	8.00 %	8.00 %
4	300	700	8.00 %	9.00 %	7.75 %	7.75 %
5	600	400	8.25 %	9.25 %	7.00 %	7.00 %

Suppose that the cash flows of portfolio ZZZ can be decomposed in two credit risk categories according to table 10.12. At $t = 0$ the credit risk premium is 1% for A-rated bonds and 2% for B-rated bonds. At t=1, the credit risk premium for both categories is equal to 2%. For reasons of simplicity, we assume that the credit risk premium is unrelated to the maturity of the loan. Based on this information, we can construct a term structure for both credit classes by adding the credit risk premium to the term structure of risk free discount rates.

The value of portfolio **IV** – decomposed into different sectors - is calculated using the term structures presented in table 10.12. The outcomes of these calculations as well as the return on portfolio **IV** are presented in table 10.13.

Table 10.13: Value of portfolio IV decomposed into sectors

	A	B	Total
Present value t=0	624	753	1 377
Present value t=1	698	865	1 563
Return R(IV)			13.45 %

The contribution of sector management is now equal to R(IV) - R(III) = -0.22 %. This contribution can be compared against the contribution of sector management in the benchmark portfolio.

Bond selection

Fong, Pearson, and Vasicek (1983) identify the remaining difference between the actual return (portfolio **V**) and portfolio **IV** as a measure of bond selection ability. The accuracy of this measure depends on the validity of the implicit assumptions underlying this method, such as the following assumptions:

the portfolio is invested in fixed income securities, without option features, such as call and put features.

the portfolio is invested in one currency.

the bond markets are very liquid and do not require liquidity premia.

If the actual portfolio deviates from these assumptions, other effects contaminate the measurement of the bond selection ability. The model can be refined further in order to correct for these mismeasurements. For example, the effect of option features can be separated from the 'bond selection' by creating another portfolio of similar bonds without option features. The difference between this portfolio and the actual portfolio represents the effect of 'option-adjusted spreads'[7]. In general, the actual decision to refine the measurements depends on the magnitude of the deviations from the assumptions and needs to be weighted against the required resources to implement the refinements.

Table 10.14: Attribution of a fixed income portfolio.

Component			**Portfolio**
Expected return	5.00 %		I
Unexpected return	3.69 %		II - I
External effect		**8.69%**	**II**
Maturity management	4.99 %		III - II
Sector management	- 0.23 %		IV - III
Selection	1.94 %		V - IV
Total management		**6.70%**	**V - II**
Total		**15.38%**	**V**

10.6 Attribution of performance using time series of returns

10.6.1 Introduction

In the previous section, we used information on the portfolio weights in order to get a performance attribution. However, the portfolio composition is not always available for the performance analyst. In this section we discuss performance attribution methods based on the analysis of time series of returns.

10.6.2 Treynor and Mazuy

Treynor and Mazuy (1966) proposed to measure timing and selectivity from portfolio returns in a different way. They argued that if the portfolio manager has timing ability, then he is able to avoid negative returns. For example, if the portfolio manager forecasts correctly a negative market return, then he will sell short stocks, and thus make a positive return on the managed portfolio. If the portfolio manager forecasts correctly a positive market return, then the portfolio manager will increase his exposure to the stock market by increasing his leverage. In such a case, the portfolio return will be an increasing function of the market return. Treynor and Mazuy model the return of the portfolio of a manager with forecasting skills as a quadratic function of the return of the market index.

7 This chapter does not explicitly deal with the calculation of option-adjusted spreads. We refer the interested reader to for example, Hayre (1990) or Babbel and Zenios (1992).

Based on this assumption, they propose to measure selection and timing ability, by fitting following curve on the time series of return of the portfolio and the market index:

$$R_a - R_f = TM_0 + TM_1\left(R_p - R_f\right) + RM_2\left(R_p - R_f\right)^2 \qquad (10.37)$$

where TM_0 is a measure of the portfolio manager's selection ability and TM_2 is a measure of the fund's timing ability. If the manager has selection ability, then $TM_0 > 0$, and if the manager has timing ability, then $TM_2 > 0$.

10.6.3 The parametric test of Henriksson and Merton

Henriksson and Merton (1981) developed two methods of performance evaluation based on the option-like payoff structure of a managed portfolio's return. One of these methods is called the parametric test, and can be used to measure selection and timing ability.

The parametric test of Henriksson and Merton (1981) is derived from a model of the value of forecasting abilities developed by Merton (1981). In this model, there is a perfect market timer who is able to predict whether or not the return on the stock market will exceed the risk free return. In effect, the market timer will invest in the risk free fund if a down-market is expected, and in the stock market if an up-market is expected. The return from this strategy corresponds to investing in a stock portfolio and put options or investing in a cash portfolio with call options. The perfect market timer is likely to be fictional, so the important issue is to assess to what degree market timers possess timing ability.

The parametric test of Henriksson and Merton (1981) allows for market forecasters that have less than perfect forecasting skills. This test is based on the following assumptions:

The portfolio manager makes forecasts of a specific type: the forecast is that either the return on the stock market exceeds the return on the risk free asset (an up-market), or the return on the risk free asset exceeds the return on the stock market (a down-market).
The probability of a correct down-market forecast is identified by p_1 and the probability of a correct up-market forecast is identified by p_2. This means that the model allows for the possibility that the reliability of the forecasts can be different in up- and down-market forecasts.
The portfolio manager reacts on this forecast by choosing a portfolio with β_1 if a down-market is predicted, and β_2 if an up-market is predicted.
The return process of the underlying assets is consistent with CAPM[8].

Merton shows that if $p_1+p_2>1$, then the portfolio manager has forecasting skills and $p_1+p_2=2$ means that the portfolio manager has perfect forecasting skills.

The method is implemented using the following regression equation:

[8] This assumption is not crucial, in the sense that the methodology could be easily adapted to for example a multi-factor APT model.

$$R_a(t) = HM_) + HM_1(R_m(t) - R_f(t)) + HM_2(\max(R_n(t) - R_f(t),0)) + \varepsilon(t), \quad (10.38)$$

where $R_m(t)$ is the return on the market portfolio, $R_f(t)$ is the return on the risk less asset, and $\varepsilon(t)$ is a residual term satisfying the usual conditions. The term max($R_m(t)$-$R_f(t)$,0) can be seen as the payoff of a long put option.

With this regression equation, the evaluator is able to test whether or not the changes in β are on average adequate, by testing H_0: $HM_2 = 0$. If HM_2 is significantly positive, then the portfolio manager exhibits superior timing skills. If HM_0 is significantly positive, then the portfolio exhibits superior selection skills.

Henriksson and Merton showed that the probability limits of the least squares estimates of HM_1 and HM_2 can be written as:

$$p\lim \quad HM_1 = b + E[\omega_2] = p_2\beta_2 + (1-p_2)\beta_1$$

$$p\lim \quad HM_2 = E[\omega] - E[\omega] = (p_1 + p_2 - 1)(\beta_2 - \beta_1)$$

where $\omega = \beta$ -E[β]. These estimates of HM_1 and HM_2 are consistent with Merton's equilibrium value of market forecasts. If the manager has forecasting skills ($p_1 + p_2 > 1$) and the manager behaves rationally ($\beta_2 > \beta_1$) then $HM_2 > 0$.

10.7 Liability-driven performance attribution

10.7.1 Introduction

The design of a system of performance measurement and evaluation should be closely linked to the investment philosophy of the investor. The most obvious way of obtaining this link is through the choice of the benchmark portfolio. The objective of the investor is the main determinant of the benchmark choice. However, the investment philosophy not only affects the investment objective, but also the design of the investment process. The components of the performance attribution should be a reflection of the decisions in the investment process.

Institutional investors, such as life insurance companies and pension funds, face at least two different objectives. The first objective is to maximize the value to shareholders of the insurance company. The second objective is to preserve the value for policyholders and to guarantee the future payout of policy obligations. From a shareholder perspective, the objective of preserving the value of the policyholders may be of secondary importance. Consequently, there is a risk that shareholders may act against the interests of policyholders by choosing high risk asset portfolios. These portfolios offer a high return to shareholders, while a large part of the potential losses will be passed to policyholders.

Babbel and Hogan (1992) showed that the strategy of excessive risk-taking is not necessarily beneficial to shareholders. In particular, if policyholders are not convinced that their interests will be protected, they are likely to demand higher returns for their policies in order to compensate for the higher risks. Higher liability returns reduces the benefits of shareholders. Eventually, this could lead to a situation where shareholders' value is lower compared to the situation where the asset portfolio has low risk.

Babbel, Stricker and Vanderhoof (1999) proposed a model for evaluating and controlling the performance of insurance companies that focuses both on the objective of shareholder value maximization and liability preservation. Their model is based on an analysis of the difference between asset and liability returns, and starts from the construction of so-called liability benchmarks. A liability benchmark is based on a portfolio of actively traded securities, and the return of the benchmark mimics the return of the market value of the liabilities over time. Next, an overall asset benchmark is created based on a separate optimization algortihm. The resulting overall asset benchmark is a maximization of the total rate of return of the assets subject to the condition that asset returns should outperform the liability benchmark returns. An alternative way of constructing a benchmark can be found in Leibowitz, Kogelman, and Bader (1992), who construct a portfolio that maximizes the asset return subject to conditions on the surplus returns. By comparing the return of the actual asset portfolio with the return on the overall asset benchmark, the value added by the asset management process can be measured and attributed to different sources.

The model we propose is different from the model proposed by Babbel, Stricker, and Vanderhoof (1999), although it shares many similarities as well. The main similarity is the central role of the liability benchmarks. However, in contrast with the Babbel, Stricker and Vanderhoof model, we compare the realized asset returns directly with the returns of the liability benchmarks, instead of creating an additional overall asset benchmark. Our motivation for skipping the step of the additional overall asset benchmark is that the liability benchmark is the most efficient portfolio from the perspective of the policyowner. By using the overall asset benchmark instead of the liability benchmark for evaluating the asset portfolio, a potential agency problem is created which could result in excessive risk taking at the expense of the policyowner. In other words, our model places more weight on the interests of the policyowner[9].

Another difference is that our model is expressed in terms of surplus returns, which means that the degree of financial leverage plays an important role in the analysis. Sharpe and Tint (1992) showed that in the context of a Markowitz portfolio optimization, the degree of financial leverage has an impact on the optimal asset allocation for an investor with liabilities. This conclusion is also shared by Leibowitz, Kogelman and Bader (1994), who use a related measure, the funding ratio return, in the context of pension fund asset allocation. From a more practical point of view, financial leverage should play a role in a performance attribution model as it emphasizes the fact that leverage increases

[9] The degree to which policy holders' interests are at stake in the Babbel, Stricker, and Vanderhoof (1999) model depends largely on the optimization algorithm used to construct the overall asset benchmark. For example, the use of the dual-shortfall approach proposed by Leibowitz, Kogelman and Bader can result in an overall benchmark portfolio that will also protect the interests of the policy holders.

the effect of any decision within the insurance company on shareholder value. Our model also makes an explicit distinction between the part of the asset portfolio that is used to cover the liability claims and the part of the asset portfolio that can be used to pay dividends to the shareholder. This distinction is caused by the extra emphasis on the objective of the policyholders. As a consequence the investment manager in our model can decide to deviate from the benchmark with respect to the allocation of funds over these two portfolios.

10.7.2 Liability-driven performance attribution

In this section, we develop a return attribution model that accounts for both the objective of liability preservation and the objective of surplus maximization. Taking a shareholder perspective, Elton and Gruber (1992) showed that part of the optimal asset portfolio of an investor with liabilities consists of cash flow matching assets. The market value of the cash flow matching assets in the optimal portfolio will be equal to the market value of the liabilities. We define the asset portfolio with cash flow matching assets as the liability-driven asset portfolio. The liability-driven asset portfolio combined with the liabilities forms a riskless portfolio.

Elton and Gruber showed that the remaining part of the optimal asset portfolio, which we will define as the surplus-driven portfolio, consists of risky assets. Depending on the institutional investor's risk aversion, he can choose to replace part of the risky assets in the surplus-driven portfolio with an investment in zero-coupon treasury bonds. Based on the distinction between liability-driven and surplus-driven assets, we develop the following performance evaluation model[10]. The variables included in this model are presented in the following balance sheet:

Surplus-driven assets	A_s	Surplus	S
Liability-driven assets	A_l	Liabilities	L
Total	A	*Total*	*S+L*

The benchmark of the institutional investor in this model has L/S cash matched investments and S/S assets with limited liability. All assets and liabilities are valued at market value. The liability-driven investments are the assets that are intended for replicating the return and risk characteristics of the liabilities. This is an important assumption in the analysis: it should be possible to replicate the return distribution of the liabilities with some asset portfolio. Liabilities expressed in nominal terms, such as bank deposits and life insurance, can be replicated reasonably well with default-free fixed income securities.

[10] The liability-driven attribution model was developed by Plantinga and Van der Meer (1995). See for a more thorough derivation also Plantinga (1999).

The actual surplus return generated by the portfolio of assets and liabilities can be calculated as follows:

$$r_s^a = r_{a_s}^a + \lambda\left(r_{a_l}^a - r_l^a\right) + \tau\left(r_{a_l}^a - r_{a_s}^a\right) \tag{10.39}$$

where r_s^a is the actual return on surplus, $r_{a_s}^a$ the actual return on surplus-driven assets, $r_{a_l}^a$ the actual return on liability-driven assets, r_l^a the actual return on liabilities, λ is the degree of financial leverage *L/S*, and τ is the so-called funding mismatch *(A*$_l$*-L)/S*. The actual return on liabilities is calculated as $r_l^a = L_{t+1}^a / L_t^a - 1$, where L_t^a is the value of the liabilities at time *t*, based on the actuarial assumptions at *t* and the term structure at *t+1*, and L_{t+1}^a is the value of the liabilities at time *t+1*, based on actuarial assumptions and the term structure at *t+1*.

This equation gives us a starting point for an analysis of the surplus return of the institutional investor. In particular, components $\lambda\left(r_{a_l}^a - r_l^a\right)$ and $\tau\left(r_{a_l}^a - r_{a_s}^a\right)$ deserve some attention. Note that λ and r_l^a are given from the perspective of the portfolio manager. The actions of the portfolio manager only affect $r_{a_s}^a$, $r_{a_l}^a$ and τ. This means that the performance of the portfolio manager can be measured in terms of these variables.

In order to calculate the performance of the portfolio manager, it is necessary to construct an appropriate benchmark. The surplus return of the benchmark can be calculated as:

$$r_s^p = r_{a_s}^p + \lambda\left(r_{a_l}^p - r_l^p\right) + \tau\left(r_{a_l}^p - r_{a_s}^p\right) \tag{10.40}$$

If the benchmark includes a strategy of perfect cash matching, then $r_{a_l}^p$ will be equal to r_l^p and τ will be zero. As a result, equation (10.41) reduces to:

$$r_s^p = r_{a_s}^p \tag{10.41}$$

If the portfolio manager does not pursue a cash matching strategy, it is interesting to determine the nature and the extent of the mismatches and the impact on surplus return. Mismatches may occur for several reasons. For example, the portfolio manager can choose an asset portfolio with a different duration than that of the liability portfolio, or the portfolio manager can choose to incur credit risk. This would result in a deviation between $r_{a_l}^a$ and r_l^a.

By subtracting equation (10.41) from equation (10.39) the following return attribution results:

$$r_s^a - r_s^p = \left(r_{a_s}^a - r_{a_s}^p\right) + \lambda\left(r_{a_l}^a - r_l^a\right) + \tau\left(r_{a_l}^a - r_{a_s}^a\right). \tag{10.42}$$

If it is not possible to hedge the liability return perfectly, then it is not fair to evaluate the portfolio manager based on the realized return on liabilities. Mortality risk[11] is a reason why the asset manager cannot hedge the liability return. For this reason, we substitute r_l^p for r_l^a, where $r_l^p = L_{t+1}^p / L_t^p - 1$. L_t^p is the present value of the liability cash flows at time *t* and L_{t+1}^p is the present value of the liability cash flows at time *t+1* based on the actuarial expectation regarding the survival frequencies at time *t*. As a result, the effect of unexpected mortality in the liabilities is ignored, and equation (10.42) becomes:

$$r_s^a - r_s^p = \left(r_{a_s}^a - r_{a_s}^p\right) + \lambda\left(r_{a_l}^a - r_{a_l}^p\right) + \tau\left(r_{a_l}^a - r_{a_s}^a\right). \tag{10.43}$$

This equation has three terms, which identify 'causes' for differences between the surplus return on the actual and the benchmark portfolio. The first term refers to the decision to choose a surplus-driven asset portfolio that is different from the benchmark. The second term refers to the decision to invest the liability-driven asset portfolio in assets that are not perfectly matched to the liabilities. The third term involves the funding mismatch, or the decision to allocate more (or less) asset to the surplus-driven asset portfolio.

The actual liability-driven investment portfolio may deviate from this benchmark in several ways. For example, the portfolio manager may change the average duration of the portfolio in order to anticipate future interest rate movements, or he may choose for corporate bonds with default risk. In order to separate the effects from the anticipation of future interest rate movements from those from default risk, we use a method suggested by Dietz, Fogler, and Hardy (1980). They construct an additional benchmark portfolio based on default-free fixed income assets, which is cash flow matched with the actual portfolio. The return on this additional benchmark portfolio is denoted by r_{a*}^p. This results in the following specification of the liability-driven return attribution:

$$r_s^a - r_s^p = \underbrace{\left(r_{a_s}^a - r_{a_s}^p\right)}_{(I)} + \underbrace{\lambda\left(r_{a_l}^a - r_{a*}^p\right)}_{(II)} + \underbrace{\lambda\left(r_{a*}^p - r_l^p\right)}_{(III)} + \underbrace{\tau\left(r_{a_l}^a - r_{a_s}^a\right)}_{(IV)} \tag{10.44}$$

The four components reflect:

I Surplus selection: the ability to forecast the returns of individual securities in the portfolio of surplus-driven assets;
II Credit Risk: the size of the premium for credit risk in the portfolio of liability-driven assets;
III Maturity mismatch: the return of the ability to forecast future interest rate

[11] Mortality risk is the risk that the survival frequencies of the policyholders will be different than projected at the beginning of the evaluation period.

movements;

IV Funding mismatch: the ability to forecast the relative performance of liability-driven assets versus surplus-driven assets.

The forecasting abilities result in a portfolio that deviates from the benchmark portfolio. Deviations in the composition of the surplus portfolio result, as reflected in component I, are called the surplus selection component. A deviation in the liability-driven portfolio that results in credit risk will be revealed in component II, which is called the credit risk component. Deviations in the liability-driven portfolio that result in a duration different from the benchmark portfolio, as reflected in component III, are called maturity mismatches. Deviations that result in an allocation of assets over surplus-driven and liability-driven assets different from the ratio of liabilities to surplus, as reflected in component IV, are called funding mismatches.

In this model we do not specifically address the issue of options on interest rates. These options can have big effects on the performance of the insurance companies. If options are present, the calculation of component III has to be changed in order to facilitate value changes in the options. This requires an appropriate model of the term structures. In order to keep the presentation of our model simple and to avoid the complex valuation issues that arise from choosing and estimating an appropriate term structure model, we presented the model for an insurer without option feature in the liabilities.

10.7.3 An application of the model

In this section we will illustrate our model of liability-driven performance attribution with data from Last Benefit. The liabilities of this funeral insurance company are typical for many life insurance companies. Life insurance companies sell policies that pay the policyowners a nominal amount conditional on the life or death of the policyowner. Funeral insurance companies sell policies that pay the nominal amount if the death of the beneficiary occurs.

As an illustration of the practical feasibility of the liability-driven performance attribution model, the performance of this funeral insurance company, Last Benefit, is evaluated over the year 1997. The market conditions prevailing in 1997 can be represented by the developments on the bond and stock markets. The return on the stock market index and several bond market indices are presented in table 10.15. This table shows that the market conditions in 1997 were rather favorable for stocks compared to the average performance during the period 1980-1987. Except for bonds with a maturity longer than 10 years, the return on bonds during 1997 was slightly lower than the historical mean return.

Table 10.15: Returns on stocks and bonds 1980-1997

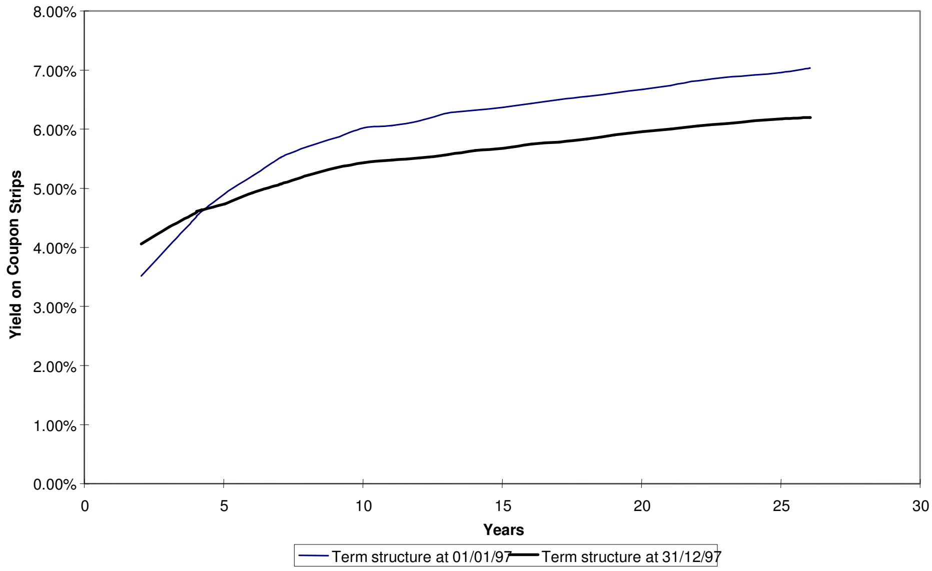
	Return 1997	**Mean Return** 1980-1997	**Standard Deviation** 1980-1997
Government Bonds			
1-3 yr	3.07%	7.11%	2.20%
3-5 yr	4.13%	8.77%	4.10%

5-7 yr	5.98%	9.39%	5.19%
7-10 yr	8.16%	9.52%	6.16%
10 + yr	12.00%	11.85%	6.22%
CBS Total Return Index	43.14%	19.18%	16.4%
Corp. Bonds (financials)		7.04%	5.42%

The mean and standard deviation of returns for bonds with a maturity over 10 years is based on return data for 1991-1995.

The term structure of default free interest rates has an important impact on both the market value of assets and liabilities. In figure 10.4, we present the term structure at the start, and end of 1997. The term structure is based on the yields of strips of coupons and principals of Dutch government bonds.

Figure 10.4: Term structure of interest rates at start and end of 1997



Liabilities

The liabilities of Last Benefit are the result of the sale of funeral insurance policies. The policies are intended to cover the costs of the funeral of the beneficiary. If the costs of funeral services rise, the policyowner is required to pay an additional premium. Effectively, this means that the insurance company does not incur any inflation risk. Premiums for this product are collected through a level premium. At the beginning of 1997, the estimated future cash flows from the current policies in the liability portfolio are calculated. The projections and the present value of these cash flows are presented in table 10.16. The present value of the liabilities is calculated based on the term structure of interest rates at the beginning of 1997. Based on the term structure of interest rates, the present value of the liabilities is 33.3 million guilders. The present value of the liabilities based on the actuarial discount rate of 4% is 80.1 million guilders.

Table 10.16: Nominal cash flows and present values of liability cash flows

Years	Nominal Cash Flows	Present Value
1997-2001	-/- 9.15	-/- 8.31
2002-2006	2.90	1.67
2007-2011	14.28	6.31
2012-2016	24.78	7.72
2017-2021	33.44	7.32
2022-2026	40.52	6.27
2027-2031	44.86	4.94
2032-2036	46.60	3.65
2037-2041	46.32	2.59
2042- …	26.89	1.14
Total	271.44	33.32

Amounts in million NLG.

As can be seen in table 10.16, the expected cash outflows rise throughout the first 45 years. However, due to the effect of discounting, the present value of the cash flows after 2036 accounts for less than 5% of the total present value of the liabilities. The cash flows of Last Benefit are fixed by the terms of the insurance policy. Early redemption is not allowed. Variation in cash flow for the level premium product is due to the unexpected development of mortality. The duration of the liabilities of the funeral insurance company is 30.84 years, indicating that the present value of the liabilities declines about 31% in value if the interest rate rises by 1%.

Assets

The portfolio of assets of Last Benefit contained only fixed income instruments. A large part (82.3%) of the portfolio was invested in government bonds, while the remaining part (17.7%) was invested in low-risk bonds issued by semi-governmental organizations or banks. Most of the portfolio of bonds were traded on the public market (90%), and the rest were bonds traded in the OTC market. The cash flows and present value of the cash flows of the asset portfolio are presented in table 10.17.

Table 10.17: Nominal cash flows and present values of asset cash flows

Years	Nominal Cash Flows	Present Value
1997-2001	65.6	58.1
2002-2006	51.8	36.9
2007-2011	20.4	11.3
2012-..	0	0
Total	137.8	106.3

Values in million NLG

The composition of the asset portfolio is mainly the result of an ‘asset-only’ investment policy. The portfolio manager is evaluated based on the Salomon Brothers government bond market index for the

Netherlands. Therefore, the manager has choosen a duration of the asset portfolio equal to 4.67 years, which is quite close to his benchmark. Although the portfolio manager is aware of the long maturity of the liability portfolio, the 'asset-only' benchmark makes an asset portfolio with a higher duration than the general government bond market index very risky from the perspective of his benchmark.

Performance analysis

In table 10.18 we show the present value of assets and liabilities for Last Benefit. The present value of the asset and liability portfolio at the end of 1997 is calculated exclusive of sales during 1997. Therefore the market value of the surplus can be considered to be the embedded value of the in force liability portfolio. Based on this balance sheet, the return on the surplus is equal to (68.77-73.12)/73.12=-6.33%.

Table10.18: Assets and liabilities of Last Benefit

	Estimated Market Value 1/1/1997	31/12/1997	Duration 1/1/1997
Surplus-driven Assets	**0**	**0**	
Government Bonds	95.97	101.32	
Other Bonds	10.47	11.11	
Liability-driven assets	**106.44**	**112.43**	**4.7**
Total Assets	106.44	112.43	4.7
Liabilities	33.32	43.66	30.8
Surplus	**73.12**	**68.77**	**-16.7**

Market value in million NLG, duration in years

Based on the liability-driven performance attribution, the contribution of portfolio management to the investment process is measured in comparison with a benchmark portfolio derived from the liability structure. The desired benchmark contains a fraction *L/A* of a portfolio that is duration matched with the liabilities and a fraction *S/A* of a passively managed portfolio of risky assets. Usually, the maturity of the liabilities exceeds the maximum maturity of available government bonds. Currently, the longest maturity available in most markets is a government bond with a maturity of 30 years. As a result, it is not possible to pursue a cash-flow matching strategy. The problem of the lack of cash-flow matching assets can be solved in several ways. One solution is to choose a duration matching strategy. Duration matching strategies immunize portfolios against small parallel shifts in the term structure. We have chosen therefore a portfolio that is cash flow matched for the first 30 years, while investing the remaining assets in 30 year bonds. The resulting benchmark portfolio has a duration equal to 27.97 years, which is quite close to the duration of the liabilities.

The benchmark for the portfolio of surplus-driven assets is CBS Total Return Index for Dutch stocks. A motivation for this choice is that the institutional investor could also have returned a large part of the surplus to the shareholders. The shareholders then would have needed an alternative investment opportunity for these funds. For the shareholders of a Dutch company, an investment in Dutch stocks is a plausible alternative.

For the performance analysis, we will use the liability-driven performance attribution model proposed in equation (10.44). The first component of this attribution model measures the effect of choosing a portfolio of surplus-driven assets that deviates from the benchmark portfolio of surplus-driven assets. As Last Benefit does not have any surplus-driven assets, the first component equals zero.

The second and the third components of the attribution model refer to the effect of default risk and the effect of interest rate movements. In calculating these components, we follow the procedure proposed by Dietz, Fogler, and Hardy (1980). According to this procedure, the effect of interest rate risk can be measured by assuming that the cash flows from the asset portfolio are default-free. The first step is to calculate the present value of the expected cash flows from both the fixed income portfolio and the benchmark portfolio at the start and end of 1997. The second step is to calculate the return of assets and liabilities from their present values at the start and end of 1997. The difference between both returns is the effect of interest rate movements.

The effect of default risk can be measured by calculating the difference between the realized return on the asset portfolio and the return on the asset portfolio under the assumption that the cash flows are default-free. In order to quantify the effect of the absence of cash matching assets for the cash flows with maturities over 30 years, we have also included the original liability cash flows in the calculations. The results of these calculations are presented in table 10.19. The return from assets, benchmark and liabilities are then calculated based on these present values. It should be noticed that the returns presented in this table are exclusive of operational components such as costs and new sales.

Table 10.19: Present value and return of assets and liabilities.

	1/1/1997	31/12/1997	return	
Market value of liability-driven assets	106.44	112.43	$r_{a_l}^a$	5.63%
Present value of assets (assuming no credit risk)	112.02	117.96	$r_{a_*}^p$	5.31%
Present value of liability cash flows	33.32	43.66	r_l^a	31.03%
Present value of benchmark cash flows	33.32	43.00	r_l^a	29.06%
Present value of surplus cash flows	73.12	68.77	r_s^a	-6.33%

Amounts in million NLG.

The difference between the return on the actual asset portfolio and the return on the benchmark portfolio is related to the effect of interest rate management. The contribution of interest rate management to the difference between the surplus return of the actual portfolio and the benchmark portfolio is defined by component III of equation (10.44). This component involves λ, the degree of financial leverage, which is equal to 33.32/73.12. Component III is equal to

$$\lambda\left(r_{a*}^{p}-r_{l}^{p}\right)=0.46(5.63\%-29.06\%)=-10.82\% .$$

The effect of the non-availability of matching assets for maturities over 30 years can be quantified by deriving the return r_l^p from the present value of the liability cash flows. Based on this return, the contribution is $\lambda\left(r_{a*}^{p}-r_{l}^{p}\right)=0.46(5.63\%-31.03\%)=-11.73\%$. The difference between the return on the benchmark for liability-driven assets and the return on the liability portfolio can be attributed to the absence of a complete market.

In order to determine component II, the present value of the cash flows has to be calculated using a discount rate including a spread for credit risk. Based on the relevant credit spreads for non-government bonds, the present value of the asset portfolio at the start of 1997 is 106.44 and at the end of 1997 is 112.43. The return on the assets including credit spreads is 5.63%. The contribution to surplus return equals $\lambda\left(r_{a_l}^{a}-r_{a*}^{p}\right)$ = 0.46 (5.63% - 5.31%) = 0.15%.

The return on the funding mismatch is measured as the difference between the actual return on the liability-driven asset portfolio and the benchmark for the surplus-driven asset portfolio. As the return on the benchmark portfolio for the surplus-driven assets, the CBS Total Return Index, amounts to 43.14%, the contribution of the funding mismatch to the difference between the surplus return of the actual portfolio and the benchmark portfolio is 5.63%-43.14%=-37.51%[12].

Table 10.20: Liability-driven return attribution

Benchmark Return			43.14%
I	Surplus selection	0.00%	
II	Credit risk	0.15%	
III	Maturity mismatch	-10.82%	
IV	Funding mismatch	-37.51%	
	Lack of matching asset	-0.91%	
Actual Return			-6.23%

[12] The procedure for calculating the funding mismatch ratio is different from the specification in equation (6). Due to the absence of an actual surplus-driven portfolio, the liability-driven return attribution presented in equation (6) becomes:

$$\underset{}{r_s^a - r_s^p} = \underset{(I)}{-r_{a_s}^p} + \underset{(II)}{\lambda\left(r_{a_l}^a - r_{a*}^p\right)} + \underset{(III)}{\lambda\left(r_{a*}^p - r_l^p\right)} + \underset{(IV)}{\tau r_{a_l}^a}$$

As a result, the return components I and IV have become meaningless. Therefore, as τ =1 in the case of the absence of a surplus-driven asset portfolio, we can rewrite this equation in the following way:

Based on the analysis of returns, it is fair to conclude that the asset portfolio performed rather poorly as compared to the liability portfolio. The main reason for this under-performance is the large difference in interest rate sensitivity between assets and liabilities. The return on the maturity mismatch contributed -10.82% to the surplus return. This is a large amount, in view of the fact that the long term interest rates have declined only 80 basis points.

The second reason for the under-performance is the absence of surplus-driven assets. The contribution of the funding mismatch to the difference between actual surplus return and benchmark surplus return is considerable. The 1997 stock market return may be considered exceptionally high. Under normal market conditions, the expected funding mismatch of the current portfolio of Last Benefit will result in underperformance equal to the risk premium for stocks. The positive contribution of portfolio management is the result of credit risk. Due to the decline in credit risk premium during 1997, a small positive effect of 0.15% on surplus return could be measured.

10.7.4 Conclusion

The application of asset-liability models by institutional investors such as insurance companies has increased the awareness of these investors of the relationship between the composition of the asset portfolio and the liability portfolio. There has been limited attention given to closely linking performance evaluation and the liability portfolio. A notable exception is the model developed by Babbel, Stricker, and Vanderhoof (1999). We developed a framework for evaluating the performance of investors with liabilities, such as insurance companies. The model is different from the model of Babbel, Stricker, and Vanderhoof, as it puts more emphasis on the interests of policyholders. The model is useful in identifying the investment skills of the investment portfolio managers of insurance companies. This is accomplished by developing a dual benchmark for the investor. The dual benchmark is focussed on the two objectives of the investor, namely the maximization of shareholder value and the protection of the value for the policyholders. For each objective, we develop a different benchmark. As a result, our model achieves a balance between these two objectives, and reduces agency costs.

The benchmark implies a distinction between ‘surplus-driven’ assets and ‘liability-driven’ assets. Based on this distinction, there are four different decisions that affect the investment performance. The first decision is related to the out performance of the surplus-driven assets. The second and third decisions are related to the liability-driven assets, and relate to the decision to incur credit risk and interest rate risk. The fourth decision relates to the allocation over surplus-driven and liability-driven assets.

$$r_s^a - r_s^p = \underbrace{\lambda\left(r_{a_l}^a - r_{a*}^p\right)}_{(II)} + \underbrace{\lambda\left(r_{a*}^p - r_l^p\right)}_{(III)} + \underbrace{\left(r_{a_l}^a - r_{a_s}^p\right)}_{(IV)} .$$

10.8 Conclusion

In this chapter, we discussed a wide variety of methods for measuring, evaluating and analyzing the performance of investment portfolios. Peformance measurement and evaluation is a science in the sense that it involves may technical calculations that allow the analyst to draw inferences about investment portfolios and investment processes. It is also an art, since the choice for a particular method requires an intuition of the analysts for the amount of effort that can be spend on performance evaluation as well as the type of analysis needed for the particular investment context. Perhaps the most important contribution of this chapter is section 10.7, where the performance analysis is explicitly linked to the liabilities of the investor. The liabilities of many investors are directly connected to the future consumption possibilities of the individuals that have claims on this investor, such as pension claims and life-insurance claims.

ENDNOTES

--------------------------------ooo---------------------------

Literature

Association for Investment Management and Research, *Performance Presentation Standards*, 1993.

Association for Investment Management and Research, *AIMR Performance Presentations Standards Handbook*, 1997.

Babbel, David F. and Arthur B. Hogan, 'Incentive conflicts and portfolio choice in the insurance industry', *Journal of Risk and Insurance*, Vol. LIX, No. 4, 1992, pp. 645-654.

Babbel, David F., Robert Stricker, and Irwin T. Vanderhoof, 'A performance measurement system for insurers', In: *Investment management for insurers*, David F. Babbel and Frank J. Fabozzi (eds.), Frank J. Fabozzi Associates, 1999, pp. 61-76.

Bookstaber, R., en R. Clarke, 'Problems in evaluating the performance of portfolios with options', *Financial Analysts Journal*, januari-february, 1985.

Brinson, Gary P. en Nimrod Fachler, 'Measuring non-US equity portfolio performance', *Journal of Portfolio Management*, spring, 1985.

Chevalier, Judith and Glenn Ellison, 'Risk taking by mutual funds as a response to incentives', *Journal of Political Economy*, Vol. 105, No. 6, 1997, pp. 1167-1200.

Del Guercio, Diane, and Paula A. Tkac, 'The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds', *Journal of Financial and Quantitative Analysis*, Vol. 37, No. 4, 2002, pp. 523-557.

Dietz, Peter O., 'Components of a measurement model: rate of return, risk and timing', *Journal of Finance*, Vol. XXIII., No. 2, 1968.

Dietz, Fogler, and Hardy, 'The challenge of analyzing bond portfolio returns', *Journal of Portfolio Management,* 1980.

Elton, Edwin J. and Marin J. Gruber, 'Optimal investment strategies with investor liabilities', *Journal of Banking and Finance*, Vol. 16, 1992, pp. 869-890.

Fong, Gifford, Charles Pearson en Oldrich Vasicek, 'Bond performance: Analyzing sources of return', *Financial Analysts Journal*, spring, 1983.

Henriksson, R.D., and R.C. Merton, 'On market timing and mutual fund performance. II. Statistical procedures for evaluating forecasting skills', *Journal of Business*, Vol. 54, No. 4, 1981, pp. 513-533.

Jensen, Michael C., 'The performance of mutual funds in the period 1945-64', *Journal of Finance*, Vol. XXIII, No. 2, 1968, pp. 389-416,.

Leibowitz, Martin L., Stanley Kogelman, and Lawrence N. Bader, 'Asset performance and surplus control: A dual-shortfall approach', *Journal of Portfolio Management*, winter, 1992, pp. 28-37.

Leibowitz, Martin L., Stanley Kogelman, and Lawrence N. Bader, 'Funding ratio return', *Journal of Portfolio Management*, Fall, 1994, pp. 39-48.

Panning, William H., 'The paradox of performance measurement', In: *Financial Management of Life Insurance Compantics*, J. David Cummins and Joan Lamm-Tennant (ed.), Kluwer Academic Publishers, 1993, pp. 69-100.

Plantinga, A. en R.A.H. Van der Meer, 'Liability-driven performance attribution', *Geneva Papers on Risk and Insurance*, Vol. 20, No. 74, 1995, pp 10-23.

Plantinga, A., *Performance evaluation of investment portfolios: the measurement of forecasting abilities and the impact of liabilities*, Uitgeverij De Marne B.V., Leens, 1999.

Sharpe, William F., 'Mutual fund performance', *Journal of Business*, Vol. 39, No. 1, 1966, pp. 119-138.

Sharpe, William F. and Lawrence G. Tint, 'Liabilities - a new approach', *Journal of Portfolio Management*, winter, 1990, pp. 5-10.

Singer and Karnosky, 'The general framework for global investment management and performance attribution', *Journal of Portfolio Management*, Vol. 21, No.2, 1995, pp. pp. 84-92.

Sirri, Erik R. and Peter Tufano, 'Costly search and mutual fund flows', *Journal of Finance*, Vol. 53, No. 5, 1998, pp. 1589-1622.

Sortino, Frank A. and Robert van der Meer, 'Downside risk', *Journal of Portfolio Management*, summer, 1991.

Sortino, Frank, Robert van der Meer, and Auke Plantinga, 1999, 'The Dutch Triangle', *Journal of Portfolio Management*, Vol. 26, No. 1, pp. 50-58.

Treynor, Jack L., 'How to rate management of investment funds', *Harvard Business Review*, Vol. 43, No. 1, 1965, pp. 63-76.

Treynor, Jack L. and Kay K. Mazuy, 'Can mutual funds outguess the market', *Harvard Business Review*, Vol. 44, No. 4, 1968, pp. 131-136.