Portfolio Risk Measurement and Reporting: An Overview For Pension Funds

Prepared by:
Malcolm Kemp (Chairman), Martin Cumberworth, Daniel Gardner, Julie Griffiths and Christopher Sandford

January 2000

Contents

Foreword by the Chairman of the Portfolio Risk Measurement and Reporting Working Party of the Faculty and Institute of Actuaries

This Overview seeks to explain portfolio risk measurement to pension fund trustees and their advisors. It forms a companion guide to the Faculty and Institute of Actuaries' Proposed Standard for Portfolio Risk Measurement and Reporting. A copy of this Proposed Standard is enclosed in Appendix A.

The Overview has been drafted by a Working Party with considerable fund management and investment consulting experience. We have also been very fortunate to receive many helpful comments from other professionals within the pension fund industry, which we hope has made the Overview easier to read and more likely to succeed.

As should be clear from the Overview, monitoring the level of risk in a portfolio is an important activity and one that trustees would be well advised to give a high priority. Time needs to be spent at the outset of the relationship defining suitable benchmarks and objectives, to ensure that the interests of trustees and fund managers are aligned. Risk relative to the benchmark should be measured regularly thereafter.

There are many different ways of measuring risk. No one method is ideal in all circumstances. The best risk measure, and indeed the underlying market risks themselves, may change over time. Usually portfolio risk is quantified in the form of tracking errors or variants such as Value-at-Risk. Trustees should note that there are quite a number of variations or assumptions that may go into the calculation of any specific figure being quoted. Careful choice of these may be needed to ensure that the resulting figures are as robust and as useful as possible. It is therefore the view of the Working Party that professional input into the risk measurement process is highly valuable.

Malcolm Kemp

1. Portfolio Risk

Investment managers generally seek to "add value", i.e. to produce above average performance, without taking "undue risk". Quantitative monitoring of fund managers has often concentrated merely on the first, without necessarily paying much attention to the second, i.e. the degree of risk being undertaken. This document seeks to help trustees redress this balance.

There are several reasons why more emphasis has traditionally been placed on performance measurement than on portfolio risk measurement:

(a) *UK pension fund performance measurement is now a very well established industry with most funds participating in such services. In contrast, risk measurement of UK pension funds is less well developed, although it is becoming more common.*

Commercial providers of UK pension fund investment performance measurement first appeared in the early 1970's. Industry-wide standards have existed for some time on how to measure and report investment performance. For example, time-weighted rates of return are now the norm and there is the AIMR Standard used in the USA and its world-wide equivalent called GIPS (Global Investment Performance Standard). In the UK, the NAPF have published standards on the presentation of investment performance, most recently UKIPS (modelled closely on GIPS). These standards and the academic theory behind them have gained wide acceptance, not least from the two dominant UK investment performance measurers, CAPS and WM.

In contrast, risk analysis of pension fund portfolios is a newer industry, with less standardisation of approach. Standards are only not being established in this area (see Appendix A).

However, both of the main UK investment performance measurers are actively promoting portfolio risk analysis alongside their performance measurement services, as are some of the leading investment consultants.

(b) *Risk measurement and reporting is more complicated than performance measurement and reporting.*

The measurement of investment performance is, strictly speaking, the measurement of *past* performance. It provides, at best, an uncertain guide as to likely future performance. Measurement of risk has the added complexity of involving two complementary, but different, perspectives. These are:

- Measurement of *past* portfolio risk. This attempts to answer questions such as "what level of risk did the manager adopt, was it appropriate (relative to the trustees' guidelines) and was the reward worth the manager taking these risks", and

- Estimation of likely *future* portfolio risk. This attempts to answer the question "what level of risk might the portfolio experience, looking forwards, were it to remain as currently structured".

(c) *Risk measurement is an inherently imprecise science.*

Given sufficiently accurate data we can calculate historic portfolio returns arbitrarily accurately. The same is not true with many sorts of risk measures. Any reasonable definition of risk will take into account the likelihood or otherwise of various (adverse) outcomes. Even after the event we will only know with certainty what actually happened. We still won't know what might have happened, or at least the likelihood that it might have happened. Although we talk about risk measures, it should be remembered that these are just formulas for estimating risk.

"Risk" is also a more complicated concept than investment performance. There are several possible definitions of risk (see Appendix B). There are also usually several possible ways of measuring each type of risk. Different people may have different views on which sorts of risk (and which sorts of risk measures) are most important.

However, just because risk measurement is more complicated than performance measurement does not make it any less important for pension fund trustees and their advisors. Trustees have fiduciary responsibilities to look after the assets in their pension fund prudently on behalf of the beneficiaries. The Pensions Act 1995 requires trustees to monitor the organisations they employ to carry out the day-to-day management of their investments. The trustees' policy on risk must be covered in their Statement of Investment Principles. This naturally leads to questions such as how should risk be monitored, measured and reported on. In time, we would expect risk monitoring and analysis to become just as prevalent as performance measurement, indeed the two should be well integrated.

But the sheer variety of meanings people attach to the word "risk" does mean that in this Overview we need to be focused on what we will cover. **In this Overview we have concentrated on a specific form of risk, namely the sort of *portfolio risk* that is of particular importance to investment managers**.

We do not cover, except in passing, other sorts of "investment" risk that the trustees themselves may be interested in, for example, the risk relative to the liabilities (if the liabilities do not behave in the same way as the benchmark given to the fund manager). Nor do we cover operational risk. Both are very important topics on which the trustees should have a policy. Indeed they can have a greater impact on the overall finances of the pension fund than the sort of portfolio risk we concentrate on in this document. However, they are both areas that are specifically outside the scope of the Proposed Standard to which this Overview is a companion. Some of the further reading mentioned Appendix D does address these types of risk.

We also specifically do not cover "absolute" risk, e.g. the risk of a large fall in the capital values of the investments (although if a fall in asset values coincides with a fall in liability values then the financial impact on the fund may not be as undesirable as appears at first sight). **Trustees need to**

be aware that investment managers will generally not concentrate on such risks. Instead they will concentrate on portfolio risk relative to the benchmarks and objectives that their clients have given them. Ensuring that the manager concentrates on the "right" sort of risk is a very important first step in risk control.

2. Benchmarks and Objectives

Trustees (and sponsoring employers) are interested in many sorts of risks. These include the risk of the liabilities growing faster than the assets, the risk of the fund failing to satisfy the Minimum Funding Requirement and the trustees' own personal risk of breaching fiduciary responsibilities.

They often carry out asset/liability exercises to identify an appropriate bespoke investment policy or set of objectives for their fund (for example because the degree of maturity of their scheme differs from the average). Or they may decide that comparison with a suitable peer group still provides the best guide to the manager on how to structure the portfolio. This latter approach is implicit for funds that are managed under what is often termed a "discretionary balanced" mandate. Such funds include many larger and medium sized segregated portfolios and nearly all pooled pension funds.

It is this bespoke or peer group based policy (and any control limits incorporated within it) that the fund manager will concentrate on when taking investment decisions. Poor performance relative to it can jeopardise the client's objectives. It can also result in the investment manager being sacked. This means that the policy given to the fund manager needs to be:

- *Carefully specified at outset*. It is a key driver in terms of ensuring that the attitudes to risk being taken by fund managers (and the sorts of "risk" they concentrate on) are in line with the attitude the trustees have themselves. Trustees should also seek to ensure that formal investment objectives given to fund managers are consistent with the style of management that the fund managers are actually incentivised to provide, particularly if performance-related fees are in operation.

- *Carefully documented*. This minimises the possibility of subsequent disagreement between the client and the fund manager over precisely what the fund manager is supposed to be trying to achieve.

- *Internally consistent, practical and achievable*. The client should not expect the manager to be trying to do two mutually incompatible things simultaneously. Nor should the client expect the impossible.

The term "benchmark" is sometimes used loosely to refer to the entire package of objectives that trustees give to their fund managers. It usually aids clarity of thinking to limit the use of the term "benchmark" to the first two of seven components that will generally be present in any objective, i.e.:

(a) A benchmark asset mix/portfolio (and a rebalancing regime that identifies how the mix changes in response to market movements)
(b) Indices/peer group averages defining the structure of the benchmark within a particular asset category
(c) A target outperformance level
(d) A time-scale over which the client hopes this target outperformance will be achieved
(e) Control ranges or other restrictions on how far the portfolio asset allocation is allowed to deviate from the benchmark asset mix

(f) Restrictions on stock or sector positions within individual components of the portfolio (or guidelines on target tracking errors in individual markets, see section 5), and
(g) Prohibitions on particular stocks or asset types.

For example, one set of trustees might set its investment manager the objective to outperform a scheme-specific asset mix by 1% p.a. over a rolling 3-year period. Suppose that in each market the manager is to outperform a suitable market index. In this instance, the benchmark asset mix is the scheme-specific asset mix. Within individual markets the benchmark is the stock portfolio underlying the relevant index for that market.

Another set of trustees might set its investment manager the objective to be upper quartile in a chosen peer group over rolling 3-year periods. In this instance, the benchmark is effectively the *average* portfolio of the peer group. The upper quartile component of the objective is the target, i.e. component (c) of the above list. Such a benchmark does in principle extend down to the level of individual stock holdings. However, getting hold of stock level information in individual markets on other managers' portfolios can be problematic. Fund managers given such a benchmark may need to use proxies such as weights in suitable market indices for risk management purposes.

3. Deviating from the Benchmark and "Active Risk"

The key reason why the benchmark is so important to fund managers is that it forms their neutral position. To outperform fund managers need to deviate from the benchmark. But the further they deviate, the more they might underperform.

The extent to which a fund manager will be prepared to deviate from the benchmark is likely to depend mainly on:

(a) The target level of outperformance
(b) The time-scale over which this outperformance is being sought
(c) The control ranges/parameters around the benchmark set by the trustees.

Trustees need to ensure that these are consistent with each other. For example, if trustees only allow their investment manager minimal freedom to deviate away from the benchmark but at the same time expect the manager to achieve a very demanding outperformance target then they are almost certain to be disappointed.

The sizes of deviations from the benchmark will also be linked to how strongly a manager holds a particular investment view. If managers hold views very strongly then their stances might be near the limit imposed by the control ranges. If they only holds a view quite weakly then they are unlikely to deviate far from the benchmark/neutral position.

The deviation from the benchmark portfolio for a specific position is usually termed the "active money" of that position in isolation. If all the active money positions are zero then the fund is being run in a perfectly passive, index-like fashion. Another somewhat older term for the same concept that is used by some fund managers is "load difference". This is because the active money is also the difference between the weight (or "load") a stock, industry or entire market has within the manger's portfolio and its weight (load) in the benchmark.

The active money of a particular position is directly relevant to the impact such a position might have on future investment performance. For example, if 5% of a portfolio is invested in a particular stock, but the benchmark only has 3% in that stock, the "active money" for that stock is 5% - 3%, i.e. 2%. If that stock then outperforms the benchmark as a whole by 10% then this will generate outperformance at the total portfolio level of 2% $\times$ 10% = 0.2%. However, if the stock underperforms by 20% then the total portfolio will, all other things being equal, underperform by 2% $\times$ 20% = 0.4%.

However, the active money of a particular stock does not provide a complete picture of the "risk" of such a position versus the benchmark. Some stocks may be much more likely to perform very differently to the benchmark than others. A 1% active money position in the former would, all other things being equal, be riskier than a 1% active money position in the latter. Actively managed portfolios are also likely to have active money positions in lots of stocks or markets simultaneously. Some of these positions may tend to exacerbate each other, whilst others may tend to counteract each other. The concept of a prospective tracking error, which we discuss further in Section 5, is an attempt to meld together into one number all of these different contributory factors.

Such concepts apply to any type of portfolio, whether incorporating equities or bonds (or a combination). In some cases, lots of different positions may contribute to risk. In other instances, a few key portfolio characteristics may dominate. How a bond portfolio's *duration* differs from that of the benchmark is often a particularly helpful measure in this respect, as often it is the single most important determinant of portfolio risk for such a portfolio.

4. Backward Looking Risk Measures

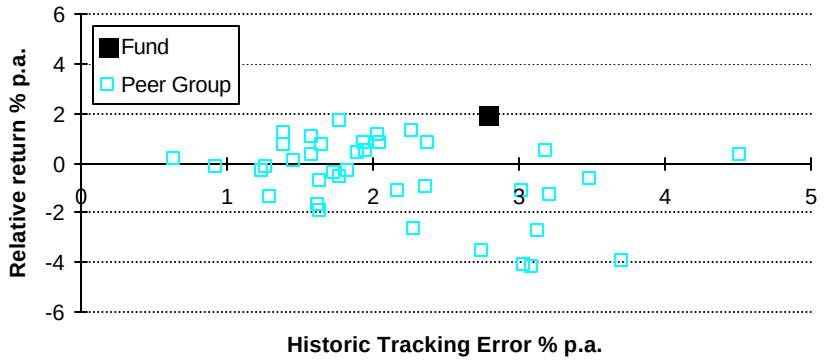
One way of analysing the riskiness of a portfolio is to analyse how it has behaved in the past. We could for example, calculate the maximum underperformance in any given month during, say, the last three or five years, or the average size of any underperformance. Alternatively, we could calculate the worst cumulative amount of underperformance during the period under analysis. Some sectors of the fund management industry do just this (often referring to such concepts by the term "drawdown").

However, these sorts of measures can be particularly sensitive to one or two extreme movements within the period being analysed. Two funds may have been adopting equally risky sorts of positions in the past. The first may have been particularly "unlucky" in that its positions might have been particularly hard hit by the market circumstances that it encountered. The second may have been more "fortunate", without necessarily running any less risk in some fundamental sense of the word. Using a limited number of valuations at discrete points in time can also flatter one fund over another without there being any underlying difference in the amount of risk either has run.

All practical risk measures suffer from these sorts of difficulties. They are only imprecise measures of the "intrinsic" risk that the portfolio has been running. Statisticians faced with this problem tend to prefer risk measures that are not overly sensitive to a small number of extreme movements, and have other intuitively appealing mathematical characteristics, whilst still being appropriate for the task in question.

The most usual sort of backwards risk measure adopted in the fund management industry is the *retrospective* or *backwards looking tracking error*. Tracking errors are based on the statistical concept of *standard deviations* (see Appendix C). If the returns relative to the benchmark are what are known as Normally distributed then in roughly two periods out of every three we would expect the return to be within plus or minus one standard deviation of the average. The retrospective tracking error is merely another way of describing this standard deviation, usually annualised, so that it refers to the *actual spread of returns experienced in the past*.

Often investors are interested in how a given portfolio compares with other similar ones. One common tool is to show both historic risk and historic return (relative to the benchmark) in a scatter-plot such as the following:



A fund that has performed well relative to its benchmark and with low risk relative to its benchmark will appear towards the top left-hand corner of this chart. A statistic that is sometimes quoted in this context is the *information ratio*. It is the ratio between the relative return and the historic tracking error. If the fund manager concerned could have doubled the sizes of all the positions (relative to the benchmark) then both the risk and the return of the portfolio (relative to the benchmark) would be doubled, leaving this ratio unchanged. Appendix D contains references to other more complicated risk statistics that trustees may sometimes come across.

It is worth noting that different organisations may calculate historic tracking errors or standard deviations of relative returns in different ways. For example:

(a) The time-scales used can vary (both the total period length and the frequency). Using data merely over a short period prior to the date of analysis may produce a poor estimate of risk, since the period could exclude any particularly extreme market conditions. Conversely, if you use too long an overall period then the current portfolio management characteristics may be quite different to those ruling at the start of the period.

(b) The weighting attached to different sub-periods can vary. Some people try to overcome the weaknesses inherent in (a) by giving greater weight to, say, the most recent months and least weight to periods a long time ago.

Usually tracking errors will be calculated by reference to the actual average return that has been achieved. Thus a fund which achieves a very steady 0.3% p.a. outperformance of its benchmark each month would usually be said to have a tracking error of 0.0% p.a.. Most investors would consider this fund to be less "risky" than a fund that performs within a range ±0.2% p.a. of its benchmark. It is less obvious that investors would also consider a fund which achieves a very steady *underperformance* of 0.3% p.a. to be lower risk than the more variable fund, as it will consistently underperform, albeit in a very steady manner. Sometimes tracking errors will incorporate an explicit assumption that the average expected performance is the benchmark performance (or some other figure), to overcome this sort of problem.

Trustees should therefore be aware that there can be some subjective input in the calculation of retrospective tracking errors, even though they may appear at first sight to be "hard facts".

5. Forward Looking Risk Measures

Backwards looking risk measures are a powerful way of analysing the risks that have been run within a portfolio. However, they are of less help in identifying what are the likely risks that the portfolio will face in the future, especially if the portfolio management style has changed a lot in the recent past.

The further you diverge from the benchmark, the greater is the risk that you are running. So, to measure the current riskiness of a portfolio you in some way need to measure how far its return might vary from that of the benchmark in the future. This in turn means that you must in some sense identify how far away the portfolio currently is from its benchmark.

How should we measure "distance" in this context? Are some measures of distance better than others?

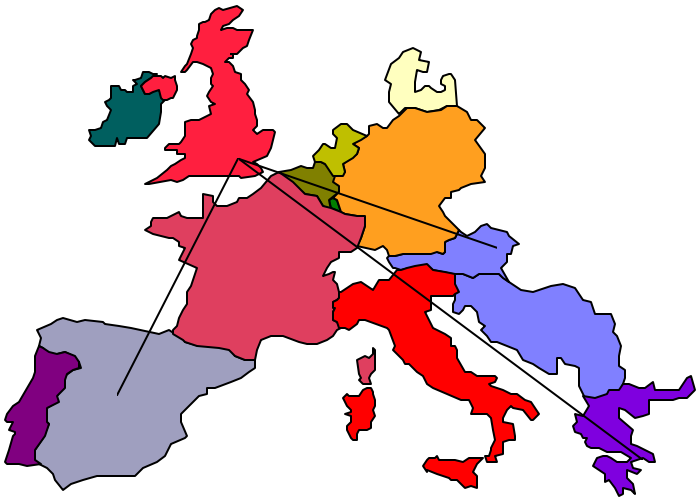
We can split the situations we might face into two main sorts:

(a) Cases where we are focusing solely on how a specific portfolio might behave. This might be particularly important to trustees who have given their fund managers a bespoke, scheme-specific benchmark.

(b) Cases where we are principally interested in ranking how the risk (relative to its benchmark) of our portfolio compares with the risk that other competing portfolios are running. This might be of key interest to trustees who have given their fund managers objectives linked to peer group performance.

The precise choice of risk measure is likely to be very important when we are concentrating on bespoke benchmarks, but may be less important if we are principally interested in the *relative* ranking of funds. An analogy, based on measuring distance in the real world helps to explain why (see box titled "Real World Distances").

Real world distances



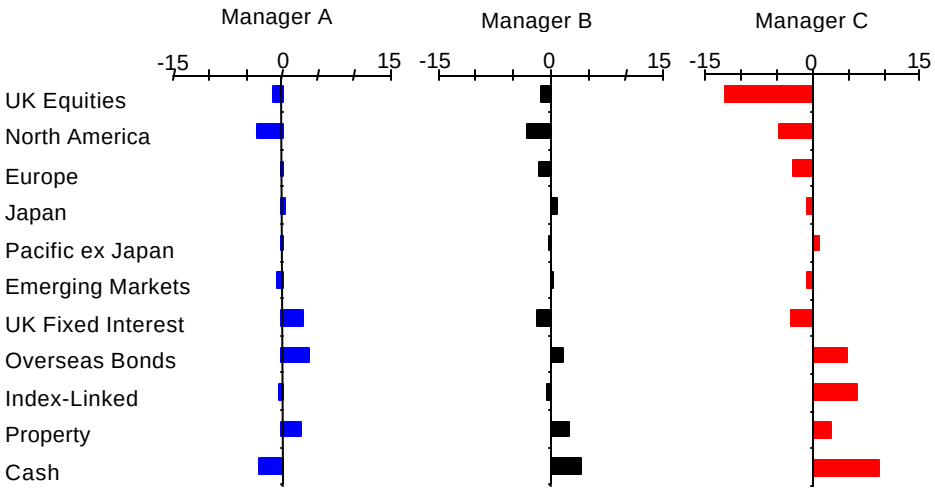
Suppose we wanted to plan a trip from London to different European locations. Looking at a map will give us a good feel for which place would take the longest to get to *by flying*. It would also give us a good feel for which would be the longest to get to *by train*, even though trains do not necessarily travel by direct routes, and the speeds at which they travel can vary quite considerably.

Athens is obviously much further from London than both Madrid and Vienna and would almost certainly take the longest to get to, whatever method of travel we are using. Maybe Madrid would take a little more (or less) time to get to than Vienna, and maybe this would depend on whether we were flying or going by train, but intuition would suggest that the times would probably not be too dissimilar. This is indeed the case:

	London to Madrid	**London to Vienna**	**London to Athens**
Air distance (miles)	**774**	**790**	**1,500**
Train times (hours)	**c.15**	**c.15**	**c.44**

So it is with investment risk. Different risk measures are like different ways of assessing travel times. Visually inspecting the strategy, or indeed using any sensible sort of risk measure, should usually tell us if one fund is much riskier than another. What a particular risk measure won't necessarily do is tell us the *absolute* level of risk involved, i.e. just how much a portfolio might underperform, just as you can't necessarily tell travel times purely by looking at a map.

For example, suppose there were three fund managers, each aiming to outperform their peer group, with positions in the major markets as follows:



Asset Allocation Position (%) Relative to Peer Group Average

Even just glancing at these positions generally gives an intuitively reasonable ordering of risk (although care might be needed if the positions in very volatile markets, like Emerging Markets, are very different). Manager C seems to be materially "riskier" relative to the benchmark than either of the other managers. However, it is difficult to get a feel for the absolute magnitude of underperformance that might be experienced by any of the managers merely by looking at such charts.

The method most usually adopted within the fund management industry to quantify forward looking risk is to calculate an estimated *forward looking tracking error*. Again the concept is linked to standard deviations of returns, usually annualised. A forward looking (i.e. *prospective*) tracking error is an estimate of the standard deviation of returns (relative to the benchmark) that the portfolio might experience *in the future were its current structure to remain unaltered*.

Forward looking tracking errors calculated in this way have some nice mathematical properties that tie in well with a desire to measure "distance" away from a benchmark (see Appendix C). They are analogous to backward looking track errors in the sense that they are estimates of what we might calculate for the backward looking risk measure at some stage in the future, assuming that the current portfolio structure remains unaltered in the meantime.

Forward looking risk measures are dependent on assumptions about the how individual stocks or markets might behave in the future. These will include:

(a) Assumptions about the likely future volatility of individual stocks or markets relative to the benchmark, and

(b) Assumptions about correlations between different stocks/markets.

These assumptions will usually be based on statistical analyses of past data. **However, the past may not be a good guide to how stocks or markets might behave in the future. There may be no past data available at all for an entirely new stock. Predictions about the future are often very imprecise. Appropriate assumptions about, for example, likely future volatility will change over time. Systems for calculating portfolio risk will often change the assumptions they make quite slowly over time, and may therefore not react rapidly enough to changing market conditions**.

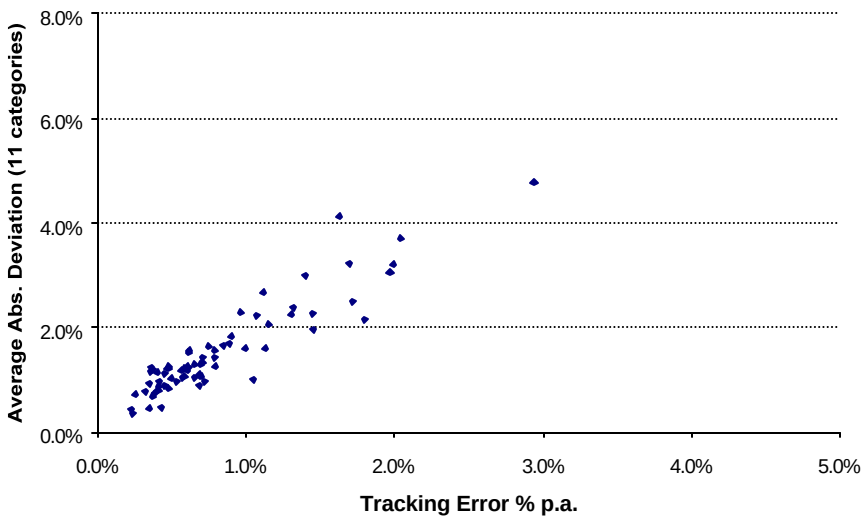
Fortunately, trustees will often mainly be interested in how risky one portfolio is relative to another (or in changes in the level of risk of a single portfolio between different points in time). As we mentioned above, relative rankings are usually fairly insensitive to the precise methodology used. An example of this in practice is given in the box titled "How sensitive are risk rankings to particular ways of defining forward looking risk?"

How sensitive are risk rankings to particular ways of defining forward looking risk?

If we try to rank the risks being run by a variety of funds, we find that relative rankings are often similar whatever the risk measure being adopted. For example, suppose we calculate a tracking error for each fund in the CAPS Mixed With Property pooled pension fund universe, concentrating merely on the fund's asset allocation positions versus the peer group average. Using the default assumptions on likely future volatilities and correlations as supplied by, say, BARRA (a leading supplier to investment managers of software to analyse portfolio risk) will give us one measure of asset allocation risk.

A second, less sophisticated, measure of asset allocation risk can be obtained by calculating merely the average sizes of the asset allocation positions for each fund (ignoring whether the positions are positive or negative in the averaging process). For the three funds shown in the box "Real World Distances", these average sizes are: 1.7% (Manager A), 1.6% (Manager B) and 4.3% (Manager C).

The chart below plots the two measures against each other for all the funds in this universe as at 30 September 1998. The correlation between the two risk measures is high despite the significant differences involved in the two calculations.



Source: CAPS, BARRA, Threadneedle (as at 30 September 1999).

6. Value at Risk

One weakness that is often cited for tracking errors is that they give a false sense of comfort. As mentioned earlier there is about a two-thirds chance of performing no worse than one tracking error away from the benchmark (if relative returns are Normally distributed with a mean equal to the benchmark return). The corollary is that there is a roughly one in six chance of underperforming by more than the tracking error.

A concept that is closely related to tracking error is *Value-at-Risk* (VaR). VaR generalises the likelihood of underperforming to any desired confidence interval, e.g. a 95% or a 99% risk of not underperforming by more than a given amount.

Estimating VaR, like estimating forward-looking tracking error, involves subjective inputs that may not be immediately apparent. It is helpful to realise that VaR is frequently calculated using the same Normal distribution assumption as the forward looking tracking error. In this case it is just an alternative representation of the same information, in a format which is more intuitively appealing, albeit perhaps with a different (often shorter) time horizon.

VaR becomes a more useful concept if the expected distribution of returns is not Normal, in which case the tracking error is less useful and may even become misleading. There are two main ways in which relative returns differ from a Normal distribution; they may be more skewed or they may have fatter tails.

For example, the performance of a poorly diversified corporate bond fund is unlikely to differ hugely from its benchmark as long as all the issuers of the bonds held remain solvent. However, if one or more of the bonds in the portfolio should default then the portfolio as a whole could substantially underperform. The spread of relative returns that the portfolio might experience is "skewed" to the downside by the existence of credit risk. Similar skews can also exist with portfolios that contain significant exposure to options (or other financial instruments with option-like characteristics), which are often used to skew the distribution to the upside, by limiting downside risk.

Fatter tailed distributions generally occur when a systematic bias is introduced into the portfolio, relative to the benchmark, such as a greater exposure to certain sectors or factors such as value stocks. Such systematic biases can lead to sustained periods of outperformance or underperformance over many months, creating more extreme outcomes than would occur if the active positions were selected at random.

The usefulness of VaR in these situations depends on our ability to make assumptions about the skewedness or "fat tailed-ness" of the distribution of returns. These assumptions usually involve the use of historical data distributions or Monte Carlo simulations from a range of possible outcomes. If these assumptions are invalid (or become invalid because of changing market conditions) then Value at Risk statistics could also misstate the risks involved. Of course, the extremes are precisely the events that are most likely to stick in the minds of trustees and it is very difficult to develop risk measurement methods that are really robust for the extreme but very rare outcomes. One possible approach is to use scenario testing (or "stress testing") to identify the impact of some particular set

of events occurring simultaneously. However it is extremely difficult to identify objective criteria to be used when choosing suitable scenarios to test.

7. Downside Risk Measures

One of the reasons VaR is intuitively appealing is that it explicitly focuses on the probability of underperformance. Another criticism often levelled at tracking error as a risk measure is that it does not distinguish between upside "risk" (the chance of outperforming) and downside risk (the chance of underperforming).

Much recent work on risk has focused specifically on downside risk measures. Simple backward looking downside risk measures might be to calculate the worst monthly or quarterly underperformance over the preceding five years. The downside equivalent of tracking error would be the standard deviation of returns below the benchmark (also known as the *downside semi-standard deviation*).

However this measure seems to imply that a fund which has consistently outperformed each month has taken no risk, as the downside semi-standard deviation would be zero. This is at odds with the concept of risk relating to active money (see section 3). In order to achieve consistent substantial outperformance the fund must have taken positions away from the benchmark, thereby surely being exposed to some risk of underperformance, unless of course the manager has perfect foresight. As we noted in section 4, it can be difficult to distinguish between funds that were "fortunate" that their high risk stances did not come home to roost and funds that actually adopted a low risk stance. Hence although downside risk measures have some basic intuitive appeal, we would caution against the use of any risk statistic which assigns zero risk to periods of outperformance; indeed extremely high levels of outperformance are also likely to be indicative of high risk, and not just superior skill.

The use of downside risk measures is often mooted in circumstances where the assumption of a symmetrical distribution, inherent in the use of tracking error, is not justified, i.e. where the return is highly skewed. In these circumstances the use of VaR, with appropriate assumptions about the skewedness of the distribution, is probably more reliable.

8. Reporting and Presenting Risk Statistics

The Proposed Standard on Risk Measurement and Reporting set out in Appendix A contains some specific comments on how risk statistics should be presented. The presentation of the results should bear in mind the likely recipients of the analysis. For example, reporting to trustees should probably be more summarised in nature than the sorts of detailed internal reporting that might be appropriate for Chief Investment Officers or other senior fund managers. The presentation should also contain:

(a) Commentary on the models/methodology used. This should include commentary on when the methodology is likely to break down, on any exercise of professional judgement within the model construction and on the main tasks to which the analysis is or is not relevant.

(b) A statement on the sources of data, on whether there are any assets (or liabilities) omitted, and the degree to which this might influence results.

(c) Levels of fund turnover and/or other commentary indicating the extent to which the analysis may no longer be relevant because of changes in fund disposition.

The Proposed Standard also indicates that the monitoring of risk should, where possible, involve both backward and forward looking risk measures. Ideally there should be a reconciliation between the two to assess model risk. Significant changes in risk parameters since the previous analysis should be explained.

Many pension funds employ several different fund managers. Trustees are likely to find it helpful to calculate the overall risk of their entire portfolio. This cannot normally be calculated merely by adding together the risks of the individual portfolios, as they may interact with each other.

It is particularly important to appreciate the strengths and weaknesses of the risk measures being quoted. As we have explained above, even apparently simple risk measures often involve subjective input. For example:

(a) Calculation of backward-looking risk measures will be sensitive to the length of the overall period analysed. It will also be sensitive to whether greater weight is given to more recent data (and if so, how much more weight).

(b) Calculation of forward-looking risk measures will be sensitive to the method used to estimate likely magnitudes of relative returns for individual stocks, sectors or markets. It will also be sensitive to the method used to estimate how different stocks or markets might interrelate with each other.

Merely calculating a risk figure does not guarantee that it is particularly appropriate to the situation in hand. We have already highlighted the difficulties of telling whether the historic risk characteristics of a fund appear high or low because of “luck” rather than just the disposition of the portfolio concerned. Portfolio construction techniques can also accidentally introduce bias. An example is described in the box titled “The tracking errors of early index funds”.

The tracking errors of early index funds

In the early days of indexation, index fund managers often seemed to have difficulty keeping within their target tracking errors, particularly if the method used to construct the index fund involved "optimised sampling". In this sort of approach, the fund manager calculates the optimal portfolio to hold using a model such as BARRA's, and then estimates the likely tracking error using the *same* model.

It is impossible to identify an exact model of how stocks might behave in the future. Instead, any real-life model can be expected to underestimate the true tracking errors of some portfolios (and to overestimate the tracking errors of others).

When we use such a model to create an optimised portfolio, we are using it to select out of a whole range of portfolios ones that have as low tracking errors as possible. This selection process is more likely to choose portfolios for which the model will underestimate tracking errors than ones for which it will overestimate tracking errors. So, if we use the *same* model to prepare estimates of future tracking errors, we are almost certain to introduce a systematic understatement into our results. Such a bias does seem to have contributed significantly to the understatement of actual tracking errors experienced by early index fund pioneers. Nowadays, index fund managers usually adjust upwards any calculated tracking error to avoid quoting overoptimistic forecasts.

Monitoring risk is an important activity and one that trustees would be well advised to give high priority. However, it is clear from this Overview that it is often not particularly easy to do. Perhaps the most important conclusion we make in this Overview is the high value of professional input into this process. This conclusion underpins many of the points contained in the Proposed Standard on Portfolio Risk Measurement and Reporting.

APPENDIX A
A Proposed Standard for Portfolio Risk Measurement and Reporting

1. Introduction & Summary

1.1 This Appendix contains a Proposed Industry Standard for portfolio risk measurement and reporting. The Proposed Standard is designed to:

(a) Improve the standards of portfolio risk measurement reporting and control within the fund management industry, and

(b) Provide practical guidelines for the implementation of such activities within a fund management house.

1.2 To adhere to the Proposed Standard a fund management company needs to satisfy the following:

(1) The risk and performance requirements of each fund should be documented and should be framed so that they are not inconsistent with each other.

(2) A formal monitoring of investment risk should be undertaken at least quarterly or whenever there is a major change to fund structure. More frequent monitoring should be undertaken if trading activity is high.

(3) The monitoring of risk should, where possible, involve both ex-post and ex-ante measurement. Ideally this would involve a reconciliation of the ex-post and ex-ante measurements to assess model risk. Significant changes in risk parameters since the previous analysis should be explained.

(4) The analysis should contain:

(a) Commentary on the models/methodology used, including when they are likely to break down (and hence suitable caveats/health warnings if appropriate), on any exercise of professional judgement and on the main tasks to which the analysis is or is not relevant.

(b) A statement on the sources of data and on whether there are any assets (or liabilities) omitted, and the degree to which this might influence the results.

(c) Levels of fund turnover and/or other commentary indicating the extent to which the analysis may no longer be relevant because of changes in fund disposition.

(5) The presentation of the results of the analysis should bear in mind the likely recipients of the analysis. In particular, it would generally be appropriate to supply

Chief Investment Officers with more detail than most third-party clients would wish to receive.

(6) The analysis could include commentary on the typical risk stances of several portfolios, all run in a similar fashion. If so, the methodology used to construct these composites should be consistent with standard rules on composite construction for performance measurement purposes. If, for example, a house median is to be calculated then the risk characteristics of each individual portfolio in the composite should be separately calculated, and then the median determined from the figures for each portfolio in isolation.

1.3 Brief rationales for each of these requirements are set out below.

1.4 This Proposed Standard was prepared by a Working Party of the Institute and Faculty of Actuaries. The members of the Working Party were:

Malcolm Kemp (Chairman), Martin Cumberworth, Daniel Gardner, Julie Griffiths, Christopher Sandford.

All have a wealth of experience in the area of portfolio risk measurement within the fund management industry.

2. The Justification for a Portfolio Risk Measurement and Reporting Standard

2.1 Active investment management is about adding value without taking undue risks. Several industry-wide standards exist on how to measure and report "added value", i.e. investment performance. For example, there is the AIMR Standard used in the USA and its world-wide equivalent called GIPS (Global Investment Performance Standard). The use of time-weighted rates of return is now very widely accepted and much of the AIMR and GIPS standards relate to the ways in which composite performances across a range of similar accounts should be constructed. In the UK, the NAPF have published standards on the presentation of investment performance, most recently UKIPS (modelled closely on GIPS). De facto industry standards also exist in the UK created by CAPS and WM.

2.2 The measurement and reporting of investment risk is less well developed. This is partly because investment risk is itself less clearly defined than investment performance.

There is also a clearly identifiable distinction between the measurement ex-post of past investment risks taken and the assessment ex-ante of risk stances currently being adopted. Past investment risk might perhaps be a matter of fact, like past investment performance, if we could agree on how past "risk" should be measured. Measuring current investment risk stances has no real analogue in terms of investment performance, except perhaps the investment target in the client's objectives. It involves subjective input.

2.3 However, this does not make measurement and reporting of investment risk any less important - it merely makes it more difficult to prepare suitable standards in this area. Two

funds which have the same past performance might have achieved this performance with vastly different levels of risk. Clients are becoming more aware of this fact, and are demanding better disclosure of information on risk. There has been significant growth in the activities of companies specialising in investment risk measurement. The leading UK institutional performance measurers are also developing services in this area.

2.4 One of the problems in preparing a Standard is the very concept of investment "risk". Widely understood, the concept can have many facets, as explained in Appendix B.

The Working party discussed which of these components might be feasible to cover in this Standard. The Working Party decided that it would not be practical to cover within the Standard aspects principally linked to operational risk (e.g. fraud, mis-pricing of unitised funds etc.).

Instead, the Working Party decided it should concentrate its endeavours on the risk of underperforming the investment benchmark or objective set by the client.

2.5 The Proposed Standard concentrates on mainstream institutional or retail fund management as this is the area that the Working Party thought would benefit most from such a Standard.

3. Brief Commentary on the Proposed Standard

3.1 Agreeing and Documenting Client Requirements

It is self evident that client requirements should be well documented and sensibly framed (from the perspective of both the client and the fund manager). Failure to do this leaves the fund management house open to the risk that the client interprets the requirements one way when the fund management house thinks that they mean something else. This is a recipe for client dissatisfaction and potentially litigation.

The main documentary repository for these requirements will normally be the client agreement between the client and the investment manager, although the manager might also prepare "norms" which it applies internally to clarify what may only be very general requirements expressed in client agreements/fund particulars.

3.2 Monitoring Frequency

Typically, third-party investment managers report to their institutional clients on a quarterly basis (although some report more frequently, e.g. monthly). Reporting is in the form of valuations, performance statements and investment commentaries. It would therefore be natural to report on investment risk somewhere within these statements.

How frequently a fund manager might want to carry out such analyses internally will depend on the speed at which portfolio risk measures change. If there is a major change in fund structure then the portfolio risk stance could, of course, change quite dramatically over a

very short time frame. This would also be true for actively traded portfolios. However, for many mainstream portfolios the risk stances do not seem to change much even over a quarterly time period, and there may be no need to carry out regular analyses more frequently than quarterly.

3.3 Type of Risk Measurement To Be Carried Out

Clients (and fund management houses) should be interested in both ex-post and ex-ante measurements of risk. This is because an ex-post analysis is important in understanding whether the risks that have been run have been rewarded in terms of outperformance. An ex-ante analysis answers a different question, i.e. what might happen in the future. Of course, for many mainstream portfolios we would expect the two to show relatively similar results.

The process of identifying why or how risk stances have changed since they were previously reported (particularly ex-ante risk analyses) is useful for indicating changes in how the portfolio is being managed.

3.4 Contents of Analysis

There are different sorts of models and methods of calculating investment risk, some more quantitative in nature than others. There are also different ways of expressing investment risk, e.g. tracking errors or "value-at-risk". The Working Party does not feel that it is appropriate to specify one method as being "the best", feeling that this should be left to professional judgement. However, it is important that the recipients of the analysis know in general terms the methodologies being used, their strengths and weaknesses and to what tasks they are relevant.

Any analysis is dependent on the data used. This is not always as complete as the recipient might expect or the person carrying out the analysis might desire. If a small proportion of the underlying portfolio is missed out this may have little impact on the accuracy of the analysis, but if a large proportion is missed out then the accuracy may be seriously compromised (and this fact may itself be suggestive of a breakdown in operational controls that might give rise to other sorts of concerns).

Forward looking analyses also require the portfolio to be reasonably stable over time to give meaningful information. The client needs to be aware if this is not likely to be the case.

3.5 Presentation of Results

Any presentation of results should reflect the likely recipients of the analysis. Otherwise the analysis could merely confuse rather than help. We would envisage fund managers supplementing raw numerical information with presentations and descriptions of the methodologies tailored to the client concerned.

3.6 Composite Construction

We can envisage fund managers who run large numbers of similar portfolios (e.g. segregated balanced pension fund portfolios) wanting to provide information on house average risk levels and variations in these risk levels. When doing so, it is essential that the composites used for risk measurement purposes match those used for performance measurement purposes. There are already standard rules on composite construction for performance purposes (e.g. AIMR/GIPS or the NAPF standards) which can therefore be carried over directly into these standards.

APPENDIX B
Different types of investment "risk"

Risk in an investment context can be split into two main sorts:

(a) **Operational risk**, e.g. fraud, mis-pricing of funds, breaching of regulatory constraints. Monitoring and controlling these sorts of risk properly fall within a compliance or internal audit role.

(b) **Portfolio risk**, i.e. the risk that the fund performs "poorly", not because it (or the fund manager) has broken any laws or regulations, but because of poor choice of investments to hold within the fund.

This Overview and the Proposed Standard set out in Appendix A concentrate on Investment Risk of the sort described in (b). The main sorts for a fund management house are:

(1) Risk of underperforming other similar funds

(2) Risk of underperforming relevant market indices

(3) Risk of loss of capital or failure to maintain an adequate level of income

All three can be thought of as variants of risk relative to some suitable *benchmark*. Thus measurement of risk nearly always involves some assessment of how far away from the benchmark the portfolio is, or has been, e.g. in terms of concentration to a single situation/company/sector.

Usually one of the three will be more important than the other two. This should be set out in the relevant client agreement.

For a pooled fund with several different owners, the fund manager may sometimes specify precisely which type of objective is most important. However, usually the fund's objectives are worded more generally. The fund manager should then document internally which sort of objective takes precedence.

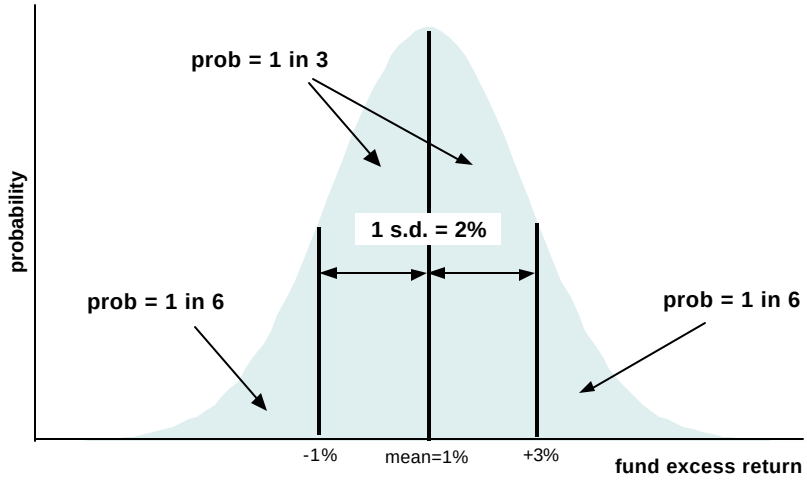
APPENDIX C
Normal Distributions, Standard Deviations and Tracking Errors

1. Although a variety of definitions and interpretations of risk are possible, the most commonly discussed measure of portfolio risk is tracking error. Tracking error is a statistical measure that indicates the range of possible outcomes surrounding the central expectation. It is equivalent to the standard deviation of returns relative to a specified benchmark. It can be calculated on either a forward-looking basis (using some form of predictive model) or a retrospective basis, using actual observed data.

2. It is helpful to illustrate this concept by a simple example. Consider a specialist UK equity fund where the fund manager is seeking to outperform some suitable peer group median with a tracking error of 2% per annum. How might fund performance vary year by year?

   Suppose that the manager has some stock-picking skill and on average will achieve an outperformance of 1% per annum. This will only be observed over relatively long periods of time. Over shorter periods there will be some variation around this level. The degree of variation is measured by the tracking error.

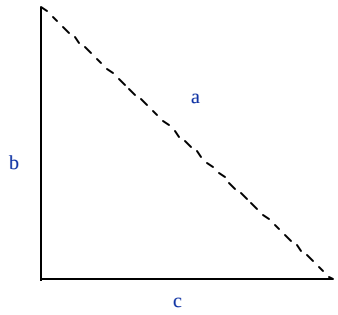
   If we assume that the returns are Normally distributed, the bell-shaped curve in the chart below will show the range of possible outcomes. There will then be an equal probability of returns being higher or lower than the average outperformance. There is also approximately a 1 in 3 chance of returns being greater than one standard deviation away from the mean. Thus the probability of returns above 3% (=1%+2%) is the same as the probability of returns below –1% (=1%–2%). Both are approximately 1 in 6.



3. Of course, it may be optimistic to assume that the mean outperformance level is 1% p.a.. Trustees may want to adopt a more cautious assumption that it will be 0% (e.g. because they think it prudent to assume that their selection of fund manager will prove with hindsight to be no better than average). The 3% and –1% would then be replaced by 2% and –2% respectively.

4. In Section 5 we explained that a good analogy for measuring risk is measuring real world distances. Tracking errors are particularly closely related to the physical concept of distance.

Suppose we have two independent sources of tracking error. We can calculate the cumulative tracking error from the two combined using exactly the same formula as Pythagoras' celebrated theorem for calculating lengths in a right-angled triangle. We would use the formula $a^2 = b^2 + c^2$ where b and c are the magnitudes of the tracking errors of the two independent sources and a is the magnitude of the tracking error from the two combined. If the two sources of tracking error are not independent then a further refinement is needed. This refinement is exactly akin to the equivalent refinement needed to Pythagoras' Theorem when the two sides of a triangle are no longer perpendicular to each other



Tracking errors "add" in the same sort of way as distances in the real world, i.e using Pythagoras' formula $a^2=b^2+c^2$ or variants when the sources are not independent of each other.

APPENDIX D
Further Reading

1. Other types of portfolio risk measures

1.1 In sections 4 to 6 we discussed several different measures of portfolio risk, including tracking error (both retrospective and prospective), Value-at-Risk and Drawdown. In section 4 we also discussed the "information ratio", i.e. the return in excess of the benchmark divided by the historic tracking error (versus the same benchmark). The information ratio is an example of a risk adjusted performance measure, because it incorporates both risk and return into a single number. Other risk adjusted performance measures that trustees may sometimes come across include the *Sharpe ratio*, *beta* and the *Treynor ratio*. Definitions are given in the glossary below.

1.2 Carrying out performance attribution can also help to identify some of the risks that have been run within a portfolio. Performance attribution involves identifying the main sources of out-performance or under-performance relative to a benchmark. Large contributions (both positive or negative) will often give indications of where the portfolio has differed most from the benchmark.

1.3 The following additional reading contains further details on these topics:

Risk-Adjusted Performance Measurement – A Practical Viewpoint (Daniel Gardner)

Value-at-Risk Handbook (J.P. Morgan)

Investment Performance Measurement (William G. Bain, in association with The WM Company, Woodhead Publishing Limited, 1996)

Downside Risk – David Damant and Dr Stephen Satchell (Professional Investor, May/June 1996)

[Details of latest available documents to be added shortly before printing]

2. Operational Risk

This Overview concentrates on portfolio risk and does not seek to cover operational risks (e.g. fraud).

In the USA a consultancy firm, Capital Markets Risk Advisors, Inc, has set up a Risk Standards Working Group which has prepared some industry-wide Risk Standards. These Risk Standards do cover portfolio risk measurement, but not in as much depth as the Proposed Standard set out in Appendix A (and not with the same emphasis on reporting issues). The CMRA Standards are more focused on operational risk.

Copies of their Standards are available on http://www.cmra.com, or from Capital Markets Risk Advisors, Inc, 565 Fifth Avenue, New York, NY 10017, USA.

Glossary

Active Management — Portfolio selection process designed to produce higher returns than are purely available from a passive strategy aiming to replicate benchmark performance.

Active Risk (or Relative Risk) — The risk of performance deviating from the benchmark return due to active management.

Absolute Risk — The risk of a large fall in the capital value (or absolute performance) of an investment portfolio.

AIMR — Association for Investment Management and Research

Average Shortfall — The average level of underperformance relative to the benchmark (or target).

Backward Looking Risk Measure — An approach to measuring portfolio risk based on historical analysis (aka retrospective, ex-post analysis).

Benchmark — A notional portfolio of assets normally comprising three components:

(i) Asset class mix
(ii) Market indices/peer group average returns within those assets classes
(iii) A rebalancing regime.

Benchmarks can be set by explicitly applying fixed percentage weights to various asset classes and index returns within those asset classes, or indirectly by reference to peer group performance, or by a combination of peer group performance within each asset class and fixed asset class weights.

The rebalancing regime defines the rules for re-setting the benchmark asset class weights which change due to relative market movements.

Beta — A measure of the extent to which a portfolio's return moves in line with the market return. A beta of 1 means that each upward or downward movement of 1% in the market should, all other things being equal, generate a 1% upward or downward movement in the portfolio. A beta less than 1 (greater than 1) implies that a 1% market movement should, all other things being equal, generate a less than (greater than) 1% movement in the portfolio.

Downside Risk	The risk of portfolio performance falling below the benchmark (or target) return.
Downside Target	A maximum underperformance level relative to the benchmark return, e.g. -3.0% in any single calendar year.
Duration	The weighted average tie to payment of cashflows arising from the securities held within a bond portfolio. Changes in economic conditions usually affect bonds of similar durations in similar ways, and often affect bonds of longer duration more than bonds of shorter duration. Hence, how the average duration of a bond portfolio differs from the duration of the benchmark is a very important component of the risk of such a portfolio.
Forward Looking Risk Measure	An approach to measuring portfolio risk based on forecast analysis (aka prospective, ex-ante analysis).
GIPS	Global Investment Performance Standards. Ethical standards for the presentation and calculation of performance. Designed to ensure fair representation and full disclosure of an investment organisation's performance history. A voluntary standard effective from 1 January 2000
Information Ratio	Portfolio return in excess of a benchmark rate divided by the standard deviation of return relative to the same benchmark. The benchmark is fund specific, e.g. cash, a market index return or a peer group average.
Jensen's Alpha	Portfolio return in excess of the beta-adjusted benchmark return
Market Risk	The exposure to loss due to market movements
Normal Distribution	A probability distribution commonly used within statistics to represent the likelihood of certain events occurring. It has the familiar bell-shape shown in Appendix C.
Performance Attribution	The process of allocating a portfolio's return (absolute or relative to the benchmark) to the various components in the investment decision-making process, e.g. to asset allocation, industry exposures and to individual stock bets.
Performance Measurement	The calculation of a portfolio's investment return over a given period to determine the growth in asset values and to facilitate comparison with funds that have similar investment objectives.

Portfolio Risk	The risk that a fund's investment objectives are not achieved due to poor performance, excessive volatility etc.
Relative Return	The performance of a fund compared to the return on the benchmark.
Risk Constraints	Formal control ranges on asset mixes, sector or individual stock limitations which are intended to act as a limit on the amount of portfolio risk that can be taken. Constraints can also involve explicit performance volatility controls (e.g. a maximum 0.2% difference in the performance of a fund and its benchmark)
Risk/Reward Trade-Off	The preferred balance between risk and return, specific to each investor. Determined with reference to the point where the incremental risk is justified by increased relative performance
Risk-Adjusted Performance Measurement	The calculation of investment return with explicit allowance for the level of risk incurred. Commonly used measures include: Sharpe Ratio, Information Ratio, Treynor, Jensen's Alpha and Sortino Ratio
Sharpe Ratio	Portfolio return in excess of a "risk-free" rate divided by the standard deviation of return relative to this same "risk-free" rate. Risk-free is usually taken to mean cash or short-dated treasury bills, even though these may not be "risk-free" in the context of a pension fund.
Shortfall Risk	The probability of failing to achieve the benchmark (or target) return.
Sortino Ratio	Portfolio return in excess of a benchmark rate divided by the downside standard deviation of returns relative to the same benchmark.
Standard Deviation	A statistical measure of dispersion or volatility of performance returns (the square root of variance).
Target Return	An outperformance level in excess of the benchmark performance, e.g. +1.0%pa over rolling three year periods.
Tracking Error	A measure of active risk. The annualised standard deviation of the difference between portfolio return and benchmark return. The measure can be used to substantiate past performance or to predict future experience. Forecast tracking errors rely on quantitative

	modelling techniques; historical tracking errors are based on the observed relative performance. For example, a forecast tracking error of 3.0% implies there is a 67% probability that the portfolio performance will be within 3.0% of the benchmark return (plus or minus) over the following 12 months (as long as relative returns are Normally distributed).
Treynor Ratio	Portfolio return in excess of the risk-free rate divided by the portfolio beta
Value at Risk (VaR)	A statistical measure of downside risk. The potential losses on a portfolio over a given future time period with a given degree of confidence. Measured in either absolute terms or relative to a benchmark.
Variance	A statistical measure of dispersion or volatility of performance returns (the square of the standard deviation).
Volatility	A measure of an asset's propensity to rise/fall in value over a specified period of time.