OCCAM Financial Technology

Updated 24th February 2003

Optimisation and Risk Analysis
Nineteen Notes for the Nervous

In these notes, we try to answer the legitimate concerns of those contemplating risk management for the first time, particularly private and smaller investors. Given OCCAM's philosophy of keeping things simple, we believe we are uniquely placed to cater for their needs, and have not been shy of saying so, especially in relation to our **POW!** *software. But in case you'd like to switch off the commercials, we have placed them in Italics, to allow you to read the remainder in peace. The only exception is the Appendix II, whose diagram you may find helpful in understanding how the risk modelling process works in general, even though it does specifically relate to our software.*

The main thing to remember is that risk management is not as difficult as it is cracked up to be, which is just as well, as it is something you are going to have to get grips with. It's no good installing a black box and thinking your worries are over - you've got to understand what it is doing, its strengths and limitations.

We hope you find these notes helpful, but if you do have any further concerns or queries we'd be happy to answer them - just give us a ring on 01892 783862, or drop an email to POW@occamsrazor.com.

Here are the questions we try to answer:

Risk in general
1. Do I need to worry about risk ?
2. There seem to be lots of different measures of risk. Which are the ones that matter ?
3. Is risk difficult to calculate ?
4. If monitoring risk is a good idea, shouldn't it be factored into portfolio construction too ?
5. Do I really need a benchmark ?

Risk Models
6. Don't I need very expensive software ?
7. Can I have something simple that meets my needs ?
8. Do stocks and bonds require different kinds of model ? And what about unit trusts and so on?
9. How do I work out which model is best for me ?
10. How can I make the model work for my clients ?
11. How can I ensure that the model produces reliable results ?

Data
12. Do I have to keep data on every single index constituent ?
13. Won't I have to pay a lot for data ?
14. What happens if there isn't enough data on some stocks ?

Maintenance
15. If I build my own model, won't it take ages to set up and maintain ?
16. Do I need a different model for each client ?
17. Isn't it dangerous for me to build my own model ?
18. Won't training be very expensive ?
19. What happens if the software goes wrong ?

Questions about Risk in General

1. Q Do I need to worry about risk ?

A Those looking after others' investments have always had a responsibility not to subject them to undue risk. In the old days, the focus was on **qualitative** measures, such as creditworthiness (hence "blue chips", "gilts" and so on), but following Markowitz's Nobel-prize winning work in the 1950s, professionals have increasingly realised the importance of **quantitative** measures. These help identify which investments or combinations of investments are volatile, that is, likely to fluctuate in value.

Even if you don't think about this kind of risk before choosing shares, you still need to understand it and monitor it carefully after the event - the Merrill/Unilever case is a warning of what can happen if you don't. Although this case concerned a very large institutional pension fund, it has focused the attention of the regulators on the area, and risk measurement is likely to be a formal requirement for all managers in the future.

For convenience these notes will refer mostly to just "shares" or "stocks", but the need to worry about risk applies to all sorts of investment - even gilts and property can go down in value as well as up.

2. Q There seem to be lots of different measures of risk. Which are the ones that matter ?

A By far the most popular measure - and the one that investment managers generally mean when they use the word risk without qualification - is **volatility**, or more formally, the **standard deviation of returns**. At first sight, this looks rather an odd choice, as it includes the chance of a share going up, as well as the chance of its going down. But for many investments these chances are approximately equal, and more importantly, volatility is easy to calculate and (subject to one or two assumptions) allows one to calculate more familiar types of risk, such as the chance of losing money.

It can also be used to calculate the proportion of one's wealth one might lose on a worst case basis over a given period, aka **Value at Risk**. Unlike volatility which is usually expressed as a percentage of the value of the investment, Value at Risk is often expressed in money terms - which helps bring home its significance.

If one is interested in hedge funds, options and similar investments, the assumption that returns are symmetrically distributed may break down - they may be skewed or have "fat tails". If so, it may make sense to look at measures that allow for this, such as **Downside Risk**, which calculates volatility taking into account only those returns below some prescribed minimum rate. But be aware that even though these risks are measurable by **analytic** software, portfolio **construction** tools which reflect them properly are still in their infancy. Until they are fully developed (or proven, as some would claim, to be unnecessary), this must raise a further question as to the suitability of such investments for institutions, let alone small investors.

Managers who are asked to outperform benchmarks may worry not about the chance of losing money, but the chance of underperforming. For them, relative versions of these measures are available, of which the most common is **tracking error**, that is, the volatility of relative returns.

Two other popular measures we owe to Bill Sharpe, who shared the Nobel prize with Markowitz. The first is the **beta**, which measures the sensitivity of a share or portfolio to the market. If a share has a beta above 1, it will tend to outperform the market when the market is rising, but fall faster when it is falling; shares with betas below 1 are defensive. The only factor in this model apart from the market is each stock's idiosyncratic or **specific risk**. Because of this, it is called the **single index model**. More complicated **multi-factor models** consider the beta of a share to more than 1 factor.

The second is the **Sharpe Ratio**. This recognises that risk is not necessarily bad, as long as it is properly rewarded. The Sharpe Ratio is the ratio of a share's return (over and above the risk-free rate, eg on deposits) to its volatility. Its counterpart in relative space and in multi-factor models is often called an **information ratio**.

All of these statistics, and many more, are available in ***POW! Risk*** *- a flexible tool which allows you to choose the numbers most relevant to you, and set them out in reports and charts in whatever way suits you best.*

3. Q Is risk difficult to calculate ?

A Whichever of these measures is chosen, the risk a portfolio **has** been subject to ("**ex post** risk" or "historic risk") is easy to calculate. Even if you don't have a GCSE in statistics, you can do the basic calculations in Excel or on a pocket calculator, although in practice it is easier to use special-purpose software.

The risk a portfolio **is** or **will be** subject to ("**ex ante** risk" or "predictive risk") is more difficult. The problem is that although the historic risk of a share is a reasonable indicator of how risky it will be in future, it is not perfect. For this reason, "quants" like to build factor and other models to work out what part of the historic risk is likely to continue, and what is mere noise.

Talking to this new professional breed, or even just reading some newspaper columns, it easy to conclude that building such models is very complicated and best left to the expert. But in fact it is straightforward, especially with the aid of software designed to make it easy.

Producing such software, and helping our clients use it, is our business. Six hundred years ago, a Franciscan monk from Ockham in Surrey proposed the rule known to philosophers as Occam's Razor - **don't make things complicated if you don't have to**. *Nowhere is this truer than in risk management, and that is why, to set our goals apart from those who would like to make it seem difficult, we named our firm* ***OCCAM Financial Technology***.

4. Q If monitoring risk is a good idea, shouldn't it be factored into portfolio construction too?

A Yes ! Markowitz's key contribution was to show how to do this, using a technique called portfolio optimisation. The idea is not just to **choose**, but also to **weight**, the set of assets which minimise the risk of the portfolio for any desired level of return; or, putting it the other way, to maximise the return on the portfolio for the level of risk your client is comfortable with - his level of "risk aversion". This risk/return trade-off is a marginal version of the Sharpe Ratio (how much **extra** return would you like for each **extra** unit of risk?)

A good optimiser will produce not just one optimal or **efficient** portfolio, but the complete set of such portfolios - the so-called **Efficient Frontier**. You can then read off the portfolio which suits your client, according to his particular level of risk aversion, but also put it into context - maybe somewhere a little lower or higher on the frontier would be theoretically inferior, but preferable for some less quantifiable reason.

To optimise a portfolio, you need for each asset its

- expected return
- ex ante risk, and
- ex ante correlation with each other asset (how the two tend to go up and down in tandem)

plus some optional extras such as:

- maximum or minimum holding constraints
- turnover constraints
- transactions costs

and so on.

Apart from those in charge of index funds, most managers are comfortable with the idea of forecasting asset returns in some form, but are not used to calculating ex ante risk, so you will need either to have risk estimates packaged with your optimiser software and updated by the vendor regularly, or, if you go for a flexible optimiser, some additional software to help you.

Either way it is quite easy to assemble the inputs needed. Couple this with the fact that maximising return while minimising risk sounds like a lot like the holy grail, and you might be pardoned for supposing that all managers must use optimisation, especially when their colleagues elsewhere in the office use exactly the same information to monitor risk after the event. But many do not, and there are two main reasons why:

- what C.P.Snow called **the two cultures**. Many investment managers have a background in subjects like history and non-financial economics, and are disposed to believe that investment is an art rather than a science. It is certainly true that some investment managers seem to have that intangible quality called "flair" but, even when they do, it must be right to impose some control on quantifiable risks, and to do so before rather than after the event, in order to ensure that if risks are taken, they are also rewarded.

- distrust of the **black box**. This is not irrational: optimisers do sometimes produce non-intuitive answers. But the solution is not to abandon a useful tool (which is really just a solver of simultaneous equations), but to learn about and allow for its limitations. In particular, you must remember that if you feed garbage forecasts in, you will get garbage portfolios out.

To protect yourself from this danger, there are a number of solutions. Starting with the easiest, these include:

- make sure that the optimiser you use is not a black box, but **transparent**, so that you can see what is going on, or at the very least comes from a vendor who is happy to explain it in terms that you can understand and ideally replicate. This seems obvious, but sadly it is a feature missing from much risk management software.

- eyeball the results the optimiser produces, and if they look odd, go back and **check your inputs**. An optimiser is a literal-minded animal, and if you say you want to maximise portfolio return, but forget to mention that your client is risk-averse, it may well assume you are happy to have 100% in the highest performing market, even if it is Hong Kong.

- in addition to any constraints which the regulators and your client impose, it is often wise to add some **common-sense limits** on maximum holdings, turnover and the like. The ability to do so is a standard feature of most optimisers. Purists regard such limits as second-guessing; practitioners know that belts are useful as well as braces.

- use a **robust** procedure which recognises that risk and return forecasts, and therefore any Efficient Frontier based on them, are fuzzy; for more about this see Q11

- use a **factor mode**l to "add structure" to your forecasts of return as well as risk, in effect imposing a discipline which requires you (for example) to ask not just how ICI will do, but how the UK market will do, how the chemical industry will do relative to the market, and finally how ICI is likely to do relative to other chemical stocks. For more about such models, see Q8-10.

- use a Bayesian technique such as Black-Litterman which makes it possible to weight one's forecasts by the degree of confidence one feels, and not to make forecasts at all where one has no views.

One other important detail: you need to tell the optimiser what sort of benchmark you have chosen: if you go for an absolute (cash) benchmark, it will maximise the total return on your portfolio for any given degree of (absolute) risk; if you go for something like the FTSE index, it will maximise the return on your portfolio relative to that index for any given degree of tracking error. The results will look quite different: for an investor with a cash benchmark, any holding of shares is risky; for an investor with an equity benchmark, the reverse is true: any holding of cash will be risky. We discuss benchmarks in answer to the next question.

POW! Frontier *is our oldest and most popular product. It will allow the user to optimise using almost* any *return-risk ("mean-variance") model based on* any *data in* any *input/output format desired, although most users take advantage of the Excel version, where a series of dialog box "helpers" guide you through the process. It does however assume that you have already estimated your risks - if you have not, you will need to use* ***POW! Toolbox*** *to calculate them, with or without factors, as a first step.*

POW! Bayes *provides a generalised version of the Bayesian techniques mentioned above, but it is relatively advanced, and we would recommend that you get started on some of our other software first.*

5. Q Do I really need a benchmark ?

A All investors need to set themselves some sort of goal for their investment activities, and it is particularly important for a private client fund manager to establish just what his client's requirements are. These will generally be quite different from those of the institutional investor, and to make this clear we will review institutional benchmarks first in (a), and only then turn to benchmarks for the private client in (b).

(a) For an institutional investor, three kinds of benchmark are common:

(1) **A single stock market index**. This is the automatic choice for a unit trust or other commingled fund investing in one market, say Japan, where the manager's job is either to come as close to the benchmark as possible (if it is a tracker) or to outperform it.

It will also be the choice where a pension fund or the like has a number of managers, each of whom is assigned to a particular market, and any one manager's job is simply to come as close to his benchmark as he can, or to outperform it. In this case the benchmark of the fund as a whole will be a weighted combination of the individual benchmarks - see (2) below.

(2) **A weighted composite of two or more such indices**. Often this will have been established as a result of an Asset/Liability study by a firm of actuaries, who will have looked at the funds' likely inflows and outflows over time. In addition they will help the trustees decide between implementing a "no-brainer" tracker or an active fund. In either case they may also recommend that the fund be farmed out to a number of specialists (as in (1) above), or handled by a single manager (as is likely if the fund is relatively small)

(3) **A peer-group benchmark**, often of balanced funds. The best known examples are the WM and CAPS Medians. This is essentially an old-fashioned approach, which has two major disadvantages:
- it assumes that "one size fits all", and
- it leads to herding - one manager copying another, and in the end everyone running "closet" index funds for fear of being out of step.

Industry standard software, and industry standard thinking, tends to be oriented to meeting the demands of these three approaches, yet…

(b) To the private investor none of these institutional benchmarks may be relevant.

For such an investor, the most natural benchmark is cash, or rather building society deposits, treasury bills or the like: there is no point in investing in risky securities if one can get a better return on assets like these that are riskless. Since they are riskless, the calculation of the risk of an investment or portfolio relative to them is no different from calculating its risk in the ordinary way - absolute risk, as originally envisaged by Markowitz.

But of course, cash is only riskless in nominal terms, so another alternative is an **inflation index** - in this case the riskless asset is the basket of goods underlying the index, which is riskless in real terms. However while such an index may be more appropriate as a benchmark of **return**, in practice inflation itself usually has such low volatility that **risk** relative to inflation is almost identical to the much simpler absolute risk, so in a low, or rather steady, inflation environment, it may make more sense to use the latter.

For the longer term investor, the best benchmark may be **index-linked bonds**. This is not to say that they are the best investment, simply that they are the lowest risk, a point that is exhaustively confirmed by a recent mathematical study by Professor John Campbell of Harvard. Campbell also confirms that the longer an investor's horizon, the more he should consider switching into equities to enhance his return.

For this reason, he may prefer to set his fund manager a simple **composite benchmark** such as 50% index linked, 50% equities - his own no-brainer solution, which his manager must beat or emulate. He may also adopt a life-style approach, gradually reducing his equity exposure as he gets closer to retirement or the time he will actually need the capital.

Finally, if he is wealthy enough, he may contemplate a significant investment in overseas markets, and at this point he will be tempted to adopt a full-scale **multi-index benchmark** like the institutional investor in (2) above; but before he does so, he should imagine the following scenario:

- building society deposits have been paying 5%
- the benchmark falls by 20%
- his own portfolio falls by 15%

and ask himself whether he will

(1) wish he had held cash and not lost 10% relative to it; or

(2) be very happy that his manager has outperformed the benchmark by 5%.

If (1) predominates over (2), he would probably be better off with a cash benchmark.

This does not mean that a private client investment manager should put all his clients exclusively into cash, but simply that he should realise that when he moves away from cash he is indeed risking his clients' money. To ensure that he allows for this, he should always measure risk and return in absolute as well as relative terms. Alternatively it may be wise for the client to specify cash as his **benchmark**, against which risk and return are to be measured, and to specify the index or index composite as a **target**, the sort of portfolio he would like to hold other things being equal.

One way the fund manager can harness this idea (and incidentally provide a multitude of clients with bespoke treatment that is still cost-effective) is to implement a **model portfolio** approach. The idea here is to run a single optimisation (usually absolute) say once a month and select say three portfolios along the Efficient Frontier - one for investors with a conservative profile (eg those nearing retirement), one for the aggressive (typically youthful or wealthy), and one in between for those with a middling risk aversion profile.

Once the clients have been sorted into one of the three profiles, each personal portfolio will then be managed to keep it close to the relevant model portfolio, but with some freedom to allow for different tax positions, favourite shares, ethical preferences and so on.

The one type of benchmark[1] that a private client manager should avoid if at all possible is the **peer-group**. Such benchmarks create an "agency" problem, where the client's interests are not aligned with those of the manager. As already discussed, the client's primary concern will probably be not to lose money, but the manager's concern will be not to underperform the peer-group.

---

[1] If comparative performance tables continue to be published, then even if peer groups are not specified as benchmarks, there will remain a danger (indeed a certainty) that fund managers will watch them, with the result that the peer median (to the extent that is known) will become an unofficial benchmark; but at least (if this advice is followed) it will be unofficial, and the longer the client himself can resist the temptation to make comparisons, the longer it will remain so. The agency problem however is inherent in the structure of the fund management industry, and although it might be further alleviated if the league tables were of Sharpe or information ratios rather than returns, even these can be gamed, as a recent study by Welch et al has shown; similar remarks apply to league tables of relative (ie de-benchmarked) returns.

Questions about Risk Models

6. Q Don't I need very expensive software ?

A Institutional investors have access to an enormous range of investments across the globe and often need to control tracking error very tightly, for example when running an index fund. To gather and maintain precise data on all these investments, and to develop models to cover the most abstruse of them was once a time-consuming and expensive task.

But the demands of private and smaller investors are much simpler. The range of investments held is much smaller, and their portfolios are too concentrated for tight control of risk or tracking error to be feasible - unless of course they have opted to make use of unit trusts and the like whose managers will do the job for them.

If they do, they will need software which can handle such investments as well as shares, a capability missing from most institutional-strength software - so the message is that they should buy software which is designed specifically for the sort of problems faced by private and smaller investors, and as such will bear the depth of their pockets in mind.

Even if they do opt for more complex models, good quality data and models have now become commodities, and no longer **need** be as expensive as once they were, even for institutional investors.

Typical industry standard software starts at over £20,000 for a single model, but can cost much more; and you'll need more than one such model if you are investing in different markets and instruments.

By contrast, ***POW!*** *prices start at £4,536, and although you may need one or two additional* **modules** *(at half the price of the first), you will still need only one piece of software, however many* **models** *you wish to create.*

7. Q Can I have something simple that meets my needs ?

A For purposes such as asset allocation, a model without factors or **full covariance model** is often appropriate, and for this purpose is the standard choice of small and institutional investors alike.

Those wishing to control currency risks may wish to use a special form of full covariance called a **currency factor model**, which breaks each investment into two components; for example, an investment in Germany might be characterised by the DAX index in Euro terms (the market factor) plus the Euro/£ exchange rate (the currency factor).

For stock allocation, an institutional investor may turn to a more complex **multi-factor** model (see Q9 below). For example, a well-known model of the European market has no less than 78 factors. Even though each individual stock can only be exposed to 10 of these plus the stock's specific risk, this is still overkill for a small investor.

For such an investor, something like Bill Sharpe's **single index model** may well be as detailed as he needs to get. As noted earlier, this breaks down the risk of each stock, and therefore the portfolio as a whole, into just three components:

- the market risk
- the stock or portfolio's beta or sensitivity to that risk
- the stock or portfolio's specific (or residual) risk

In calculating the portfolio's specific risk, it is normal nowadays to take each stock's residual correlation with each other stock into account as well, at least in part: oil shares, for example, even after discounting their sensitivity to the market as a whole, tend to behave more like other oil shares than they do like financials, and they may be negatively correlated with airlines, whose costs suffer when the oil price is high. But some models, like that used in London Business School's Risk Measurement Service, follow Sharpe's original model and ignore these correlations.

A popular half-way house, particularly with funds of funds, is the **style model**. It was Bill Sharpe once again who showed how it could be done, but two other US academics, Fama and French, who suggested the key factors - size and value (versus growth). These two factors, they demonstrated in a famous article published in 1992, capture the major part of the cross-sectional (ie single-period) variation in stocks' returns, to the extent that not much is left for market beta. This is not necessarily true of variation through time, and one useful type of style model does make use of the market as a third factor.

POW! Toolbox *is designed to allow the user to create a wide variety of models, which can then be applied in* ***POW! Frontier****. He is not tied to one type of asset, or one market area.*

8. Q Do stocks and bonds require different kinds of model ? And what about unit trusts etc?

A Yes. Although the risk of bonds is in many ways much better understood than that of equities, "the industry standard" method of modelling them is much more obscure, typically involving a technique called Principal Components (see Q9 below) or an approach based on options theory. This is probably due to the fact that the main impetus behind bond models has come from two sources:

- the United States where a wide variety of exotic instruments is available, with features that can be handled only by quite intricate models
- banks, which are preoccupied with managing diverse portfolios of specialist paper

They are not designed for the investor holding at most a few instruments, mostly governments plus a few corporate bonds, largely concentrated in his home market, with perhaps one or two holdings in the major overseas markets.

For such investors, a model based on the volatility of the relevant indices, adjusted for any difference in duration between the bonds held and their respective indices, will be quite sufficient. OCCAM has produced a paper on this subject, available on request[2].

Unit trusts and other commingled vehicles can be handled in a number of ways. In the rare case where their full constituents are known, they can be handled as a portfolio of the shares or other instruments of which they are composed. But almost equally good results can be obtained by treating them as if they were shares themselves (which of course investment trusts are); this will be convenient where they form part of a portfolio which also holds shares directly. Where the only other holdings are other unit trusts, a fund of funds technique such as that described in Q7 will be appropriate.

There are two key points to note here:

- there is no one right answer
- given a suitable model for one type of asset (say shares) it is easy to combine it with one or more similar models of other assets (say bonds) into a single unified model, since the underlying Markowitzian structure is exactly the same.

However until recently, the pre-packaged models used by industry standard software precluded the easy development of models to handle "new" types of asset such as unit trusts, and the integration of different types of instrument in the same model. Portfolios holding different types of instrument - that is, the majority of real world portfolios - were left at least partially in the cold. Integration is now available in such software - but at a price.

POW! *software is flexible and under the user's control; it therefore has not, and never has had, any problem accommodating new instruments and integrating the heterogeneous.*

---

[2] There is one other type of bond model often used for attribution , based on the "bucket" system. This sorts bonds into a set of buckets, say one for 1-5 year bonds, one for 5-10 year, and one for over 10 year. This approach springs from an unthinking adaptation of equity models: a chemical company **is** arguably completely different from an electrical company; but not only is a 10.1 year bond not very different from a 9.9 year bond: in two months' time it will actually be one.

9. Q How do I work out which model is best for me ?

The answer to this section is quite detailed and has therefore been relegated to Appendix I. It may be skipped on a first reading

10. Q How can I make the model work for my clients ?

A ***POW!*** *software, as well as being technically flexible, also makes it easy to customise reports and charts, so that they can be presented in a way suited to every type of client. For example, if the manager wishes to report just two numbers (say return and risk) in a mail-merged letter, he can do so.*

If you do need help in explaining risk numbers to clients, or preparing explanatory literature for the purpose, we will be happy to help, although some consultancy fee may be appropriate.

11. Q How can I ensure that my model produces reliable results ?

A There are four main reasons why models may produce unreliable results, which we will consider in turn:

(a) Fat tails and similar phenomena

The distribution of returns on investments or portfolios as a whole may turn out to be:

1. negatively skewed (more low returns than high)
2. positively skewed (more high returns than low)
3. fat-tailed (more extreme returns than normal)
4. thin-tailed (less extreme returns than normal)

Obviously models can be built to handle these phenomena when they are expected, and in fact the standard investment model does expect positively skewed returns, since it is based on a **lognormal**, rather than the more familiar **normal** distribution.

A normal distribution assumes it is as likely for a share to fall in price by 200% as it is to rise by 200%: this is implausible, as it is difficult for a share to have a negative price. A lognormal distribution assumes it is as likely for a share to **halve** in price as it is to **double**, which is much more sensible, and is by and large borne out empirically.

The problems described above arise then when shares or portfolios do not conform to a lognormal distribution. Once again, where this is expected, as in the case of options, there should be no difficulty - the risks can be calculated by Black-Scholes or some similar formula.

But they may arise for other reasons. It is easy to see how this can happen in the case of hedge funds, where the manager follows some specialised rebalancing strategy (which may indeed include holding or replicating options); it may also happen to more conventional portfolios when their managers bet very heavily on particular factors; and it may happen to individual shares or the market as a whole as a result of feedback from investor behaviour - examples are the October 1987 crash and the LTCM débâcle.

To understand one way of coping with this sort of problem, consider Value at Risk (VaR).The easiest way to calculate it is to assume a (log)normal distribution, and then, if one is interested for example in the return which one can be 95% confident of achieving, simply feed in the mean return and volatility, and read off the answer from the distribution formula. But it is also wise to check what actually happened to a share, or what would actually have happened if one had held a particular portfolio. To do this, supposing one has 200 return observations to go on, simply sort them by size from highest to lowest, and read off the actual 190th return - this is then the **Empirical VaR** at 95%. The great advantage here is that one does not have to make any assumptions about the shape of the distribution, and with the right software it is just as easy.

Another approach, already discussed in Q2, is to use measures such as Downside Risk, which do not require an assumption of symmetry. However such approaches will not in themselves cure problems associated with fat and thin tails.

(b) Autocorrelation

Positive autocorrelation occurs when shares trend - once they start going up (or down) they tend to continue in the same direction. This phenomenon is particularly likely when returns on the investment in question are calculated by reference not to market prices, but to valuations. Asset classes prone to this include property and private equity. In such cases, the eventual risk of the investment may turn out significantly higher than conventional calculations suggest. The solution in relatively predictable cases like property is to make a suitable adjustment based on historic experience.

Negative autocorrelation is when shares "mean-revert", that is when a rise one day tends to be followed by a fall the next, and vice versa. This is a common feature of markets within short periods such as a day or week, due to fluctuations in demand and supply, and is one reason why it is unwise to use too frequent observations to measure risk - they will tend to exaggerate it.

(c) Risky risk

Another problem is that the volatility of shares does not remain constant - there may be violent crashes followed by periods of relative tranquillity. A whole series of complicated models under the generic name **GARCH** has been designed to handle this, but they suffer from two main problems:

- they require an impracticably large amount of data to estimate
- over typical investment manager rebalancing periods such as one month, there is little predictable effect.

Such models are best left to banks and others for whom short term fluctuations are mission-critical.

A better solution here is a procedure known as **bootstrapping**. This can be applied to static risk analysis, but is particularly useful in optimisation. The idea is not just to estimate risk based on the actual set of historic returns, but to draw random samples from history, so one can see what risk would have arisen if some returns had not been observed at all, while others had happened twice over. After taking each sample, one recalculates the risks and correlations, and then does a fresh optimisation. Then, after conducting a sufficient number of bootstraps (say 1000), one takes the 1000 resulting frontiers and averages them. The result is a set of portfolios which, in cases where the risk is risky, are much more stable and less prone to small differences in levels of expected returns.

Like the empirical calculation of VaR discussed above, an advantage of this approach is that it does not require any assumption about the shape of the return distribution, in contrast to another attempt to address this problem called **resampling**.

(d) Risky holdings

A problem that was ignored until quite recently concerns the difference between the way ex ante and ex post risk are calculated. Ex post risk looks at the way the returns on a portfolio have fluctuated over a period of time; during this period, not only will the returns on the shares held have gone up and down, but so also will the size of the holdings. Ex ante risk, on the other hand, normally calculates the risk of a portfolio allowing for variations in the returns on the shares but assuming the size of the holdings remain constant. Recent research by Lawton-Browne suggests that this may understate portfolio risk on moderately actively managed portfolios by around 30%.

This effect has not yet been fully explored, but its strength is likely to vary with the degree of portfolio turnover, and the rebalancing strategy followed. In either case, the solution is to investigate the historical behaviour of the manager in question, adjust it for any known changes, and use the resulting information to calibrate future ex ante risk estimates. The phenomenon will not affect optimisation as such, since Markowitzian optimisation handles only one period at a time, but it will be important to ensure that the level of risk aversion specified by the client is understood to be on an ex ante basis, or is recalibrated to be so.

OCCAM offers two modules which address these issues: ***POW! Robust****, which allows the user to bootstrap returns and produce more stable frontiers; and* ***POW! Empirical*** *which allows the calculation of*

- *risk by direct reference to actual historic returns*
- *over differing periods to check for stability*
- *autocorrelation, skewness etc*
- *downside, upside and related risk measures*
- *and so on*

Both are intended not as replacements for their model-based counterparts, ***POW! Frontier*** *and* ***POW! Risk****, but as supplements and cross-checks to them.*

Questions about Data

12. Q Do I have to keep data on every single index constituent ?

A It is only necessary to have data on every index constituent if the client:
1. is investing in individual shares rather than unit trusts or commingled funds

and 2. has a benchmark which contains an index

and 3. the client wishes to track the index, or manage his risk relative to it, very closely

and 4. he has a portfolio large and diversified enough for this to be a realistic proposition

As will be evident from the remainder of these notes, the number of private clients who fulfill all 4 conditions will, or at least should, be very small. For all other clients, it will be quite sufficient to use the published index as a whole, or if greater precision really is desired, its individual broad sectors.

There are two reasons why it is theoretically desirable to have data on every index constituent:
- it allows one to see if one is over or underweight a particular stock. But this is not really relevant if the portfolio only holds 30 or 40 stocks in a 100-share, let alone a 750-share, index.
- it allows one to calculate tracking error and betas more accurately for some purposes (although for other purposes such "up to date" betas can be misleading).

13. Q Won't I have to pay a lot for data ?

A Almost certainly no, for two reasons.
1. You probably do not need even 10% of the data required by the institutional manager, as explained above; and
2. What you do need - historic prices of the securities you actually hold, and any relevant index levels, you almost certainly have already. The only thing you may need to do is to set up a database - a simple spreadsheet will be quite adequate - to hold the historic data, and then refresh it weekly or monthly, as you choose. This might account for a man-day to set up, and about an hour a week or month (according to how often you choose to refresh it) to maintain.

14. Q What happens if there isn't enough data on some stocks ?

A In ex ante analysis and portfolio construction, if it is to be statistically significant, there should be many more data observations than there are stocks in the portfolio. For example, to analyse a portfolio which holds, or is considering holding, 30 stocks, it is desirable to have at least 60 months worth of monthly data. Failing, this 60 weeks of weekly data will do, but the shorter the overall period, the greater the danger that it will not cover all the market situations likely to occur.

Monthly data is good, because it is the best approximation to most portfolio rebalancing intervals; daily data is less good, as it will tend to pick up too much noise, while to get enough observations with quarterly and annual data, one will have to go back so far in time that the data may no longer be representative of the market as it is now. To cope with this problem, even with monthly data it may be desirable to ask the software to weight more recent data more heavily.

Despite one's best endeavours, there will always be one or two holdings with short or missing data. In these cases, there are a number of options which are easy to set up in ***POW!'s*** *Excel environment. These include:*

- *linear interpolation (the easiest, but not recommended*
- *mapping the short asset to another similar asset or index*
- *using* ***POW! Toolbox's*** *new maximum likelihood gap-filling functions, which insert estimates of the missing observations based on the structure of the data that is available. (These use a technique similar to that employed by RiskMetrics, but with some enhancements to make them more rigorous in certain cases)*

Questions about Maintenance

15. Q If I build my own model, won't it take ages to set up and maintain ?

A *The amount of time required to maintain data is very small, as discussed in Q13 above. Apart from this there are only two other requirements:*

- *running* **POW! Toolbox** *to update the risk matrix. This should not be done more than once a week, and we would in fact recommend once a month or once a quarter. Even allowing for some time to eyeball the model once it has been created, the process should not take more than an hour at the outside; and the actual model creation will take only a minute or two, maybe less.*

- *running the optimisation (***POW! Frontier** *and/or* **POW! Robust***) or the risk analysis (***POW! Risk** *and/or* **POW! Empirical***) to process each portfolio. Again allowing for eyeballing, this is bound to take a little time whatever software is used; but the advantage of* **POW!** *is that the whole process can be automated, including input and output - see Q16 below.*

16. Q Do I need a different model for each client ?

A It is important to distinguish **risk models** from the sort of **model portfolios** described in Q5. As explained in Q9, it is the type of asset, and (to some extent) the investment process applied by the manager to that asset, which determines the type of risk model required. Once it has been developed, the same risk model can be applied to any number of portfolios, model or otherwise, whatever their benchmarks.

POW! Frontier *and* **POW! Risk** *allow the creation of a single* **Universe Workbook** *to handle a model of the available assets, and then to set up as many* **Portfolio Workbooks** *as are required to contain details of the particular subset of those assets held by each portfolio.*

To make the manager's task even easier where a large number of individual portfolios are involved, **POW! Mandate Machine** *allows the user to combine the details of up to 120 portfolios in a single* **Mandate Workbook***, so that they can be optimised and analysed simultaneously. The manager can then choose between:*

- *reports and charts on each portfolio separately (suitable for client or individual manager consumption), and*
- *summary reports covering all or chosen groups of portfolios (suitable for use by the CIO and others involved in supervision), or*
- *both.*

To obtain a fuller understanding of how these different components work together, please refer to Appendix II.

17. Q Isn't it dangerous for me to build my own model ?

A There certainly are pitfalls in building one's own model from scratch, especially if a general purpose statistical package is used; but there are also pitfalls in using someone else's model which may be ill adapted to one's particular purpose and which one may not fully understand.

POW! *software provides a convenient halfway house. Its generic but investment-specific framework allows managers to build models suited to their particular needs, while still keeping them within the framework proposed by Markowitz and his successors.*

Even so, it is possible to build and attempt to apply models unsuited to the application in question, so OCCAM is regularly called on by clients to advise on the best type of model for their purposes, or to review one already chosen, services which we are happy to supply, often free of charge.

18. Q Won't training be very expensive ?

A *OCCAM will help you get up and running with our software, and provide basic training, free of charge. Thereafter telephone help is also free, and we will do our best to drop in if you need advice or would like an update on the software. If you do need more extensive consultancy or training, we will again be happy to oblige, but may need to make a charge.*

19. Q What happens if the software goes wrong ?

A *No software is infallible, and each client's computer environment and application of the software is slightly different, so we cannot promise that the software will never suffer from bugs, but if you do have a problem, we promise to investigate it and cure it as soon as possible - typically within 10 working hours.*

Often, as you might guess, it is a matter of user error, or features in the software which you have misunderstood, but we realise that from your point of view, this may be just as bad as a genuine bug, and we will do our best to put you back on the road as soon as possible.

Appendix I
Answer to Q9: How do I work out which model is best for me ?

The answer to this question is quite lengthy, in order to give the reader a proper feeling for the difference between the various types of model, but even so descriptions of the most complex models have had to be cut short for reasons of space. If you would like further details on any of the models, please do let us know.

The members of the ***POW! Frontier*** *and* ***POW! Risk*** *families can handle all the models described, and* ***POW! Toolbox*** *can be used to estimate all the models described in part (i) and (ii) except for (ii)(g) .* ***POW! Toolbox*** *is in process of modification to facilitate the handling of the models described in (iii).*

There are three main questions you should ask yourself in choosing a model:

(1) Which model for any given share or portfolio gives the best ex ante estimate of the **amount** of risk and correlation it is subject to (ex post calculations should be exact) ?

(2) Which model best explains or **attributes** the risk in terms which are relevant to my client and my process ? This is an issue ex post as well as ex ante.

(3) Which model is most **cost-effective** in terms of money if I buy it ready made, or in terms of time if I build it myself ?

We will consider the three main families of model discussed in Q7 and Q8, together with some additional variants, by reference to criteria (1) to (3), in progressive order of complexity and cost.

The models considered are:

i) Regression-free models
(a) Full covariance models
(b) Currency factor models
(c) Duration-based Bond models

ii) Time-series Single and Multi-Factor Models
(a) Single index models
(b) Manager models
(c) Returns-based Style models
(d) Cross-country models
(e) Macro-economic models
(f) Statistical Equity Factor models
(g) Statistical Bond Factor Models

iii) Cross-sectional Multi-Factor models
(a) Characteristic-based Style models
(b) Fundamental Factor models

i) Regression-free models

(a) Full covariance models based on stock market and similar indices should produce an accurate[3] estimate of the quantum of risk (criterion 1), and are well-suited to asset allocation over the corresponding asset classes (2), although it should be noted that this implies a top-down process. If the sub-portfolios (say for each country) are actively managed, it is theoretically preferable to use the actual manager's returns in place of the index - equivalent to bottom-up; but in practice this nicety is opften ignored. It becomes more important when one is constructing a fund of funds and there is a choice of managers to represent each country, although in this case one of the variations on the style and single index themes may be more suitable - see below.

Full covariance models may also be applied to industry and similar allocation exercises; but they are less well suited to allocating individual stocks, where they include too much unreliable noise; some sort of factor model, such as one of those described in (ii) and (iii) below, will usually be more suitable.

Full covariance models require

either	a single risk matrix of the various asset classes in the investor's currency
or	a time series, typically of 60 or more monthly returns, for each asset class from which such a matrix can be constructed

In either case this data will be easy, and therefore cheap, to assemble (3).

(b) Currency factor models, as discussed in Q7, despite their name are really a form of full covariance model. They will therefore produce an accurate estimate of risk (1), and will be useful when currency risk is to be actively managed or monitored (2)[4].

Currency factor models require

either	a single risk matrix containing the various asset classes in local currency plus the exchange returns on each currency
or	a time series for each asset class and currency from which such a matrix can be constructed

They are therefore slightly more onerous and complex than full covariance models and need not be used unless currency is an issue (3).

---

[3] Here and throughout Q9, this means "accurate within the limitations discussed elsewhere in these notes".

[4] Proponents of cross-country models (see (ii)(d) below) point out that currencies can be a factor driving markets and securities even in local currency terms: for example, the Yen/$ exchange rate may explain the behaviour of the export-oriented Japanese stock market even in Yen terms; in this case a currency-factor model may under- (or over-) state the currency exposure of a particular market. However this is something of a special case, and the total quantum of risk will remain correct - the issue is only about how it is broken down.

(c) Duration-based Bond models also have much in common with full covariance models, since like currency factor models, they require no regression, but simply the estimation of a full covariance matrix of cash and bond indices, which serves as a factor matrix, plus a set of prescribed betas of each stock to the relevant factors, which in these case are not 1s (as in the currency model) but based on the ratio of the duration of each bond (or bond fund) to its index. Optionally, currency may be broken out, again without regression, this time with betas of 1. Residual risk may also be handled, by subtraction rather than regression.

Such simple models can produce surprisingly accurate estimates of risk (1), and moreover ones that are much closer to the investment management process (2), and are dramatically more frugal of data than the statistical models discussed under (ii)(g) below (3)

Duration-based Bond Models were first developed by OCCAM.

Requirements are:

either — a single risk matrix containing the various indices (and currencies if required)
plus durations (or data from which durations can be calculated) of each bond and index
plus residual risks (with correlations if required) of each bond

or — a time series for each index (and currency if required) from which such a factor matrix can be constructed
plus durations (or data from which durations can be calculated) of each bond and index
plus a time series of each bond if residual risk are included

They are therefore slightly more onerous and complex than full covariance and currency factor models, but are basically easy to assemble, especially if no residuals are needed (3).

ii) Time-series Single and Multi-Factor Models

(a) Single index models are useful for investors concentrating on their local market and not requiring a great level of detail (perhaps the majority !), and may be useful **within** markets, in conjunction with full covariance or currency factor models **across** markets for managers who use a two-stage investment process (2). They may also be useful as a simple means of building a fund of funds within a single market.

.

Single index models in their original form, without residual correlation, ignore a considerable amount of risk information, although they are easier to compute as a result. In their full form with residual correlation, they can be reassembled into a full covariance model. Whereas full covariance models are good for asset classes, as already noted they (and unadjusted equivalents) generally do not ignore enough noise when applied to stocks. To work out how much noise to eliminate, a multi-factor model may be used - see below; but a much simpler, and still very effective, approach was proposed by Ledoit: using a statistically calculated adjustment factor, the entire (off-diagonal) correlation matrix can be damped down (1).

Single index models require
either — a single risk number for the market
plus a beta and residual risk for each stock
plus (full form only) a residual correlation matrix
or — a time-series for each stock (as in the full covariance model) plus one extra time-series for the market.
They are therefore cheap to assemble (3).

(b) Manager models are a simple form of model developed by OCCAM, extending the single index model to cover more than one market. They are useful for those building multi-market funds of funds, or requiring a simple model to cover a number of stock markets. They can be estimated in one of 3 ways:

A. by a doing a multiple regression of each fund against all the markets being considered (this will result in each fund having a beta not only to its own market, but also to each other market), and then keeping the residual risk, but usually not the residual correlations;
B. by regressing each stock against its own market only, and then assigning all remaining risk and correlations to a full residual matrix, which may be damped à la Ledoit. This allows the fund of funds manager to break down each fund's risk into its beta to the market index plus a residual. Correlations with other funds can then be restricted to those mediated by the index correlations (if the residual correlations are damped to zero), or additionally allowed for in the residual matrix. This approach is likely to be closer to the way the fund manager thinks.
C a halfway house between A and B, which like B regresses each stock against its own market but then also regresses it against one or more of the other markets thought or found to be statistically relevant, before assigning all remaining risk and correlations to a full residual matrix, which again may be damped à la Ledoit.

Which of these methods is chosen will depend to some extent on the manager's process. Because it uses more factors, A is likely to estimate the total quantum of risk most accurately, with B still quite satisfactory, and C somewhere between (1); on the other hand B or C is likely to be a better fit to the investment process (2).

Manager models require
either risk numbers for each market
plus a market factor correlation matrix (eg 6 x 6)
plus one or more betas and a residual risk for each stock
plus (full form only) a residual correlation matrix
or a time-series for each stock (as in the full covariance model) plus one extra time-series for each market.

They are therefore cheap to assemble (3), but slightly more complicated to estimate than single index models.

(c) Returns-based Style models when applied to a single market are like single index models, but with each stock or fund regressed against (typically)
either a market neutral size index (small minus big)
and a market neutral value index (high minus low)
plus a market index
or a small growth index
a large growth index
a small value index
and a large value index
but other additional styles and permutations are possible.

When applied to more than one market, there should in principle be additional style indices for each market, but some sacrifice of within-market detail, especially overseas, may be pragmatically desirable[5].

Although this structure is not strictly a Fama and French model (for which see (iii)(a) below), like that model it has high explanatory power, and it is usually possible to dispense with residual correlations (1).

These models also fit in well with many investment processes (2), and have quite simple requirements, typically:
either risk numbers for each index factor
plus a factor correlation matrix (eg 4x4)
plus betas and a residual risk for each stock
or a time-series for each stock
plus one extra time-series for each index factor.

All quite straightforward (3). Indeed, since the index factors are often just "factor-mimicking portfolios", it may be possible for the user to construct the factors himself, if he has the constituent weights.

---

[5] Sharpe's original illustrative formulation when applied to US mutual funds contained the following style indices: Bills, Intermediate Bonds, Long-term Bonds, Corporate Bonds, Mortgages, Value Stocks, Growth Stocks, Medium Stocks, Small Stocks (all in the US) plus Foreign Bonds, European Stocks and Japanese Stocks.

(d) Cross-country models are similar to the manager multi-factor models described above under (b) method C, but with each stock being regressed against a wider selection of currencies, markets and industries. The regression may be block-wise, with some practitioners preferring to "take out" countries first, thereby maximising the explanatory power of country factors, but others arguing that with the globalisation of the markets, it makes better sense to take out industries first.

If well-constructed, they can have high explanatory power for many stocks, both in terms of the quantum of risk and, at first sight, its source. For example, a cross-country model might reveal that a company like Shell is exposed to the £, the US$ (in which oil is priced) and the Dutch Guilder among currencies, Oil, Chemicals and (negatively) Transport among industries, and the UK, US and Netherlands among stock markets. (1 and 2).

However with regard to sources of risk, these models are less transparent than may at first appear: each successive block of factors in the matrix is **orthogonalised** against the previous blocks; for example, any industry effect will be a "pure" industry effect with all currency effects washed out of it (and all country effects too if countries are taken first). This means that the factors are not as "real-world" as they look.

To ensure that such models work well for all stocks requires considerable skill, and quite substantial amounts of data in their construction - because more factors are involved, more observations are needed for statistical significance (3); and even then some explanations may owe more to statistical quirks than intuition or economic rationale (1 and 2 again). The chief commercial proponent of this type of model is QUANTEC, now part of the Thomson group.

Requirements are:

either — risk numbers for each index factor
plus a truly multi-factor correlation matrix (eg 48x48, although this may be sparse, ie contain blocks of zeroes)
plus multiple betas and a residual risk for each stock

or — a long time-series for each stock
plus one extra time-series (also long) for each index factor.

(e) Macro-economic models in their original form (due to Chen, Roll and Ross) seek to explain the behaviour of stocks in a single market in terms of macro-economic phenomena, such as interest rates, inflation, industrial production and so on. They are intuitively appealing to economists (2), but unfortunately as Connor showed in a useful 1995 article, their explanatory power in terms of quantum of risk is poor (1). However Salomon Brothers has developed a more useful version of the model in several individual markets (such as UK, US and Japan) by incorporating industry and market factors. (1 and 2)

Requirements are not too heavy (3):

either — risk numbers for each index factor
plus a multi-factor correlation matrix (eg 12x12)
plus multiple betas and a residual risk for each stock

or — a time-series for each stock
plus one extra time-series for each index factor.

(f) Statistical Equity Factor models are models in which a statistical technique called Factor Analysis, or a similar technique called Principal Component Analysis[6], is used to estimate from the stocks themselves which set of factors can best explain them. These factors are not real-world factors, but statistical constructs, and therefore score very low on criterion (2), even though they often score highly in their estimates of the quantum of risk (1).

However by using a further statistical technique called Factor Rotation, they can be mapped to real-world factors. This greatly improves their intuitive value, although the mapping may not be perfect, a deficiency which may be because that is the reality of the situation - real world factors are not good at explaining every nuance, or at least not those that we are familiar with - but it may also be because statistical models are prone to "high-jacking" by spurious or other one-off effects. Like the models in (e) they require skill to construct; the main ready-made vendors are APT and EM Applications.

Because the factors are created out of the stocks themselves, time-series data requirements are the same as full covariance, except to the extent that real-world factors are introduced (3):

either — risk numbers for each index factor (eg 20)
but no factor correlation matrix, as each factor is orthogonal to (ie uncorrelated with) each other factor
plus multiple betas and a residual risk for each stock
plus further information on each real world factor

or — a time-series for each stock
plus an extra time-series for each real world factor

6 Factor analysis proceeds by trying to minimise the amount of unexplained **covariances**, while Principal Components tries to minimise unexplained **variances**.

(g) Statistical Bond Factor Models are similar to statistical equity factor models, except that they typically require only 3 factors per country (estimated as Principal Components) plus currencies and corporate spread factors. The description which follows is based on the BARRA Cosmos model.

As noted in Q8, although these models are accurate in estimating the amount of risk (1), they are low in real-world explanatory power: they are typically not mapped to real world factors although labels associated with real world factors are affixed to them (2), and some adjustment is made to make these as congruent as possible.

Before the PC factors can be created, a bond term structure model such as Cox-Ingersoll-Ross is applied to the universe of individual bond data to estimate the term structure at each moment in time, with 20 vertices, or points along the spot curve, being chosen to define it. The bond data is adjusted for any embedded optionality. It is the changes in these vertices over time that are used to create the PC factors, to which are added the spread factors, and the currency factors (which may themselves be separately modelled using GARCH), to give the raw material for the factor covariance matrix.

Finally each bond is decomposed into a series of zero-coupon bonds whose betas to the individual factors can be calculated; these betas are then summed to give the full exposure of that bond to each factor. The residual risk is then modelled separately.

Data requirements are:

either — risk numbers for each index factor (eg 4 per each of 21 markets plus 26 currencies = 110 in all)
a factor correlation matrix, to the extent that the factors are not orthogonal
plus full details of each bond in the universe (including those not held) to permit the calculation of each bond beta to be built up out of the betas of its individual cash flows
plus the residual risk for each stock

or — a time-series for each bond (including those not held) and currency
plus full details of each bond in the universe (including those not held)

The data and pre-processing demands of this type of model are enormous, let alone the construction of the factor model itself. Even supposing it was worthwhile in other respects, it would be out of bounds for all but the very largest banks to construct.(3)

iii) Cross-sectional Multi-Factor models

(a) Characteristic-based Style models

All models are ultimately time-series models, since the manager's concern is variations in return over time. However such models can be built up from a series of cross-sectional models, each of which explains variations in returns from one stock to another at a single moment in time.

The basic technique is to regress the returns on all the stocks at the end of period t on the characteristics of the same stocks at the beginning of that period. This answers the question, "How strongly associated was the presence of characteristics x and y in the stock with a good or bad return ?" A simple example, which formed the basis of the Fama and French article already mentioned, is a cross-sectional model using the two factors size (market capitalisation) and value (book to market). The importance of each factor in each period can be thought of as the return to that factor during that period, and the series of these returns can be "stacked up" into a factor time-series.

Once this is done, no further regression work is necessary, since each stock's sensitivity to each factor ("beta") is simply its current size, value or other characteristic (these will typically be normalised). The residual is simply what is left unexplained by the factors.

As Fama and French showed, the explanatory power of the size-value model is high (1), and these factors have the additional benefit of being transparent to the manager and often associated with his process (2). In cross-sectional form, it is more onerous than its pure time-series counterpart, the returns-based style model describe in (ii)(c), since each stock's characteristics are required in each period (3). On the other hand its "beta" will in some sense be more up to date, and stocks which have short histories can be incorporated without any problem, so long as their current characteristics are known.

Requirements are:

either	a factor covariance matrix (eg 2x2)
	each stock's current characteristics (including those not held) (eg n)
	its residual risk and (if desired) correlation (eg n, n x (n-1)/2)
or	a time-series of each stock's return (eg (n +dead) x 60)
	+ a time-series of each stock's characteristics (eg (n + dead) x 2 x 60)
	(in both cases including those not held; it is also desirable to include those now dead but then alive, for the purpose of estimating historic factor returns.(3)

(b) Fundamental Factor models were first developed by Barr Rosenberg et al in the 1970s, and continue to form the basis of the BARRA stock models, although these days the latter often incorporate additional bells and whistles.

Essentially they are elaborations of the cross-sectional style models described in the previous section. The characteristics ("Descriptors") used (at least by BARRA) are more numerous, so much so that they are aggregated together into a dozen or so "Risk Indices". These are then used as the basis of a cross-sectional regression, which incorporates a large number of dummy factors, principally to identify membership of different industries; in some versions stocks may be allowed to belong to more than one industry, as determined by fundamental characteristics. Cross-country versions also include country and currency factors.

In the BARRA equity models specific risk is modelled separately; residual correlations are discarded as insignificant given the number of factors.

BARRA models impound a very large amount of data, so it would be surprising if they were not in general successful in providing an accurate estimate of the quantum of risk, although they have come under fire (not always with justification) over the last few years (1).

In terms of real-world relevance they are more problematic.There are 4 main difficulties:

- the Risk Indices are too heavily processed to be readily recognisable as what their labels indicate they once were; they are no longer real-world factors
- the Risk Indices may relate to factors such as Trading Activity, Size Nonlinearity and Labour Intensity which even if recognisable have little relevance to the average fund manager's investment process
- the factor exposures or betas are not real-world betas. They are equivalent to those arising from a multi-factor regression, whereas the true betas are those which would arise from a series of single-factor regressions. For example, the size beta of Vodafone may be stated as 0.29, although its real beta, taking into account the correlation of size with other factors, is 1.19.
- for these reasons, together with the overall complexity of the models, it is often difficult even for a well-trained user to have a good idea of what is going on inside the box - much has to be taken on trust.

These deficiencies, which are not unique to BARRA or even to fundamental factor models, can be most effectively overcome by an intelligent user constructing his own model in a simpler and more transparent form (2).

The data requirements of fundamental factor models, as we have mentioned, are large but not impossible for those with funds under management sufficient to justify the development of such a model: they will typically have access to most of the data anyway, so that the main challenge is marshalling it in the form required.

Requirements are:

either	a factor covariance matrix (eg 50 x 50)
	each stock's current characteristics (eg n x 50)
	its residual risk (eg n)
or	a time-series of each stock's return (eg (n+dead) x 60)
	+a time-series of each stock's characteristics (eg (n + dead) x 40 x 60)

Once again the stocks should include those not held, together with those that are dead, for the purpose of estimating historic factor returns.(3)