Some aspects of active portfolio management

David J Buckle

January 2003

Director of Research

Lee Overlay Partners

38 Wellington Road

Dublin

Ireland

+353 1 660 2852

None of the views expressed by the author in this article are necessarily shared by Lee Overlay Partners.

Some aspects of active portfolio management

Abstract

We propose a framework for considering active portfolio management, consistent with existing literature in the single period environment, but extended to the multi-period environment that most practicing active managers operate in. We derive several active portfolio management results including the *Generalised Law of Active Portfolio Management* (a generalisation that accounts for correlations and varying information coefficients and volatilities), and the *Theory of Active Portfolio Management* that suggests that mean/second-moment optimisation is the best active portfolio management approach in a multi-period setting, not mean/variance. In developing these results we uncover some inappropriate practices, and several simplistic practices that are theoretically supported by our framework.

Some aspects of active portfolio management

1. Introduction

Active portfolio management has not had the same level of mathematic rigour applied to it as have more mainstream areas of finance. The notable contributions being the security selection model of Treynor and Black (1973) and the work of Grinold and Kahn (1999). In this article we propose a structure for active portfolio management equivalent in the single time period to the above, but extended into the multi-period setting. Within this structure we uncover some interesting and controversial conclusions regarding active portfolio management. We list these conclusions in order of increasing importance:

1. Variances of our forecasts and their errors are related to the asset variance by the information coefficient
2. Distinguishing between residual returns that are uncorrelated with the market and returns in excess of a benchmark can be important.
3. The practice of using total asset return covariance matrices in active portfolio position optimisations is theoretically unsound, and can distort the positions.
4. The practice of using a linear allocation rule where an active position in an asset is proportional to its forecast is theoretically reasonable.
5. We derive the *Generalised Fundamental Law of Active Management* as an extension of Grinold and Kahn's Fundamental Law of Active Management. Using

this law we find that the use of independent active manager specialists is better than one general active manager who can account for asset correlations.

6. The naïve practice of taking active positions of constant size regardless of the size of forecast can produce higher information ratios than using modern portfolio theory.
7. The practice of active portfolio managers of using by mean/variance optimisation to generate active positions do *not* maximise information ratio, it is the positions generated by mean/second-moment optimisation that do. We derive this in a theorem that we have termed the *Theorem of Active Portfolio Management*.

We consider active portfolio management in the context of an active manager being given an investment wealth to manage on behalf of the trustee of that wealth. This context is vastly more prevalent than active managers using their own wealth. It is then helpful to be clear from the outset why a trustee of an investment wealth would consider active portfolio management, and what then is the role of the active portfolio manager. Therefore in Section 2 we provide the theoretical underpinnings of the objectives of both the trustee and the active manager. That being done we develop in Section 3 our mathematical structure for considering active portfolio management. This structure is equivalent to the existing work of Treynor and Black and Grinold and Kahn, but is extended to the multi-period setting where the active portfolio manager typically operates. Conclusions 1 – 3 are drawn in this Section. In Section 4 we explores three alternative approaches to active portfolio management and identify two important results, namely the Generalised Fundamental Law of Active Management and the Theorem of

Active Portfolio Management. Conclusions 4 – 7 are drawn in this section and stem from these two results. In Section 5 we make a critical review of the assumption we have adopted. In Section 6 we relate our results to existing literature. Section 7 contains our conclusions. An Appendix contains additional mathematical detail.

2. Objectives of the trustee and active managers

Consider the two participants of active portfolio management. The first is the trustee of an investment wealth who potentially allocates the wealth to an active manager. The second is the active manager who manages the investment wealth on behalf of the trustee. Given that we are attempting to provide rigour to active portfolio management, prior to commencing our study it behoves us to formally identify their objectives.

2.1. The trustee

Let us begin by considering the decision facing the trustee. It is how to allocate an investment wealth amongst a universe of assets. This decision is considered a single holding period decision, normally over a long-term horizon. As such it has been addressed by a series of extraordinarily important academic articles.

Modern portfolio theory (Markowitz 1952) says that in order to make the single period allocation decision the trustee should select an efficient portfolio. The market separation property (Tobin 1958) goes further to say that the trustee should hold some combination of the risk free asset and the *single* efficient portfolio of maximum Sharpe ratio (the so-called consumption decision).

If the efficient market hypothesis holds, then the CAPM (Sharpe 1964) says that this portfolio is the market portfolio. Alternatively, if asset prices are inefficient, an active manager may be able to beat the market portfolio. The ability to do this is measured by information ratio (Treynor and Black 1973). So, additional to the consumption decision, the trustee should choose the active managers with highest information ratios in order to maximise the Sharpe ratio (see for example Bodie, Kane, and Marcus 1996 or Grinold and Kahn 1999).

There is nothing particularly controversial about the above framework. It is considered a foundation of finance, and the objective of selecting the active manager with maximum information ratio is industry standard (Grinold and Kahn 1999).

What is controversial is our proposal that the typical active manager does not attempt to maximise the information ratio.

2.2 The active manager

Let us now take the perspective of the active manager, who has been given an investment wealth by the trustee. The active manager's objective is clear from the previous section: maximise the information ratio.

The typical active manager also uses modern portfolio theory in order to construct the managed portfolio (see for example Dybvig and Ross 1985, and Grinold and Kahn 1999). However in order to outperform the market, the active manager must take advantage of perceived mispricings by making regular changes to the holdings in the trustee's portfolio. So the active manager has a multi-period problem.

But modern portfolio theory is for the single period decision problem, and applying it to each period of the multi-period problem does not maximise information ratio. As we shall see there is an alternative investment approach that maximises the information ratio.

3. Mathematical framework for active management

In order to make any inference on active portfolio management we need a rigorous mathematical framework to work within. Ours consists of several parts. First we provide a framework for modelling asset returns where we essentially follow the well-known models proposed by Treynor and Black (1973) and Grinold and Kahn (1999). After developing this structure we use the information coefficient to calibrate our input parameters. Next we provide a framework for active portfolio positions, and illustrate how the asset return structure can be applied to portfolio return thereby hinting at the validity of using information ratios. Finally we provide an expression for information ratio, the measure of active portfolio management performance. In particular we identify an expression for information ratio as a function of the active portfolio positions and asset return forecasts.

3.1. A framework for modelling asset returns

Let the column vector $\boldsymbol{r}_t$ denote the returns from our universe of assets over the single time period $(t, t + 1)$. We decompose these returns into a market component and an uncorrelated residual component, so that

$$\boldsymbol{r}_t = \mu_t + \theta_t$$

where μ_t is the market component, and θ_t is the vector of single period residual returns that are uncorrelated with the market return. We shall denote $V[\theta]$ by Σ.

ASSUMPTION 1: We shall assume that in the long run the efficient market hypothesis holds and therefore we apply the assumption from the equilibrium asset pricing models such as the CAPM (Sharpe 1963) or APT (Ross 1976) that $E[\boldsymbol{\theta}] = \mathbf{0}$.

The CAPM and other equilibrium models are myopic in that they only consider one holding period. For our application that holding period is long. It *may* be the case that the market is inefficient over much shorter time horizons in which case θ_t will contain some single period abnormal return in addition to idiosyncratic noise. Adopting the notation of Treynor and Black we have that

$$\theta_t = \alpha_t + \phi_t$$

where vector ϕ_t denotes idiosyncratic noise and vector α_t denotes abnormal returns over the period (t to t+1). Observe that the abnormal returns are indexed by time. In this regard we have extended Treynor and Black's model to a multi-period setting.

ASSUMPTION 2: We assume that abnormal returns are only available in the short term and not indefinitely. Hence $E[\boldsymbol{\alpha}] = \mathbf{0}$.

3.2. Active manager forecasts

It is important, conceptually, to note that our model thus far incorporates no active management forecasts. For active management to exploit the opportunity to beat the market we must forecast α_t. So let us consider such a forecast, which we shall denote by the vector $\mathbf{m}_t$. We relate α_t and $\mathbf{m}_t$ by the equation

$$\alpha_t = \mathbf{m}_t + \mathbf{e}_t$$

where $\mathbf{e}_t$ is the error in our forecast.

*ASSUMPTION 3: We assume that we are unbiased forecasters in the long run which implies that E[**e**] = **0**.*

We can combine equations (2) and (3) to give us that

$$\theta_t = \mathbf{m}_t + \varepsilon_t$$

where ε_t denotes the error in our forecast plus the idiosyncratic noise in the market place. We shall denote V[**m**] by **Ω** and V[ε] by **V**.

*ASSUMPTION 4: We assume that **m** and **ε** are independent so that*

$$\boldsymbol{\Sigma} = \boldsymbol{\Omega} + \boldsymbol{V}.$$

We can see from this assumption that our ability to beat the market will be reflected in the magnitude of **Ω** relative to **V**. It should be strongly noted that we have made no assumption regarding market efficiency. If the market is efficient then **Ω** = 0 and all of the residual return is attributed to idiosyncratic noise and therefore our active management framework collapses down to standard equilibrium pricing equations.

*ASSUMPTION 5: For mathematical convenience and to be consistent with existing mainstream literature we shall assume that **r**, **θ**, and **m** all follow a multivariate Normal distribution.*

3.3. Relating returns to forecasts: the information coefficient

The information coefficient, denoted IC, is an indicator of the quality of the active portfolio management (see for example Grinold and Kahn 1999). We can use it to relate residual returns to forecasts. It is defined as the correlation between the residual return of an asset and the abnormal return forecast. By definition it is applicable only in a univariate sense. Adopting our notation we define IC as

$$IC = E[m\theta] / (V[m]V[\theta])^{1/2}$$

where m and θ are an element of **m** and $\boldsymbol{\theta}$ respectively.

Applying equation (6) to our framework it is easy to show that

$$IC = \omega/\sigma$$

where ω^2 and σ^2 are the appropriate elements of $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}$ respectively. As the variance of residual return is attributed more to the abnormal return forecasts and less to the error term so IC will increase from a lower limit of zero when the active manager has no insight at all, to an upper limit of 1 when the active manager has perfect insight and all the residual return is abnormal return.

3.4. Identifying values for Σ, Ω and V

The matrix **V** is the covariance matrix of the asset residual returns conditional on our forecast of abnormal return. Together with $\mathbf{m}_t$ it defines our forecast distribution of the residual returns at any point in time. We therefore need to identify **V** to fully specify the forecast distribution. This can be done using IC, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$. We first identify the variances in **V**, then the correlations.

Variances

A variance embedded in $\boldsymbol{\Sigma}$, denoted by σ^2, represents the variance of residual return. It is observable and can therefore be estimated by empirical analysis of historical asset return data. Provided we know the information coefficient, then we can use equations (5) and (7), to identify a variance embedded in $\boldsymbol{\Omega}$, denoted by ω^2, thus

$$\omega = IC\sigma.$$

Being a volatility of our forecast, ω may not satisfy equation (8). In this event theoretical consistency can be achieved by scaling the forecasts such that equation (8) is satisfied (see Section 5.3.).

Having identified ω along with σ, we can again use IC and equation (5) to identify the variance of the forecast errors. Such a variance is embedded in **V**, denoted by v^2, and is trivially shown to be given by

$$v = (1\text{-IC}^2)^{½}\,\sigma.$$

<u>*Correlations*</u>

Having identified the variances we now consider correlations. The correlations in Σ are the correlations of the asset residual returns. It should be noted that they are <u>not</u> the asset return correlations. There is often significant difference between the two. In fact many financial models including Treynor and Black (1973) and Grinold and Kahn (1999) plus the single index model (Sharpe 1963), the multi-factor model (Cohen and Pogue 1967) assume the residual returns to be uncorrelated. If an assumed structure is not desired the correlation in Σ can be estimated from observed data.

The correlations in $\boldsymbol{\Omega}$ are the correlations between the forecasts. These too can be estimated from a time series of forecasts.

Now, from equations (5), (8) and (9) we have that the correlations between any two assets can be represented as

$$\rho_\Sigma = \rho_\Omega\, \text{IC}_1\, \text{IC}_2 + \rho_V\, (1\text{-IC}_1^{\,2})^{½}(1\text{-IC}_2^{\,2})^{½}$$

where ρ_Σ, ρ_Ω, and ρ_V denote the correlations between two residual returns, two forecasts, and two forecast errors respectively, and IC_1 and IC_2 denote the information coefficients of the two assets. *Therefore an observed asset residual return correlation can be attributed to both the correlation of the forecast and the correlation of the error*.

In viewing asset returns in this fashion we can clarify a deeply misunderstood issue regarding the role of correlations in practical investment management. In practice, the temptation is to substitute total asset return covariance, $V[\mathbf{r}]$, for $\mathbf{V}$. Such an action can erroneously embed high correlations into $\mathbf{V}$, since the correlations in Σ are likely to be much closer to zero (or even assumed to be zero). Moreover, provided the ICs are moderately high, any non-zero correlation may be attributed to the forecasts. The correlation of $\mathbf{V}$ should then be set to zero *even though the observed correlation of* Σ *is non-zero*.

3.5. A framework for modelling portfolio weights and resultant returns

In order to beat the market the active manager must construct a portfolio that differs from the market portfolio. Conceptually the active manager increases exposure to those assets that are forecast to produce abnormal return and decreases exposure to those that are not. Given our assumption that abnormal returns do not exist over the long run, the active manager must regularly adjust the exposures to account for periodic opportunities as they come and go.

The total portfolio positions, denoted by $\mathbf{w}_t$ and satisfying $\mathbf{w}_t^T\mathbf{1} = 1$, can be decomposed into market portfolio positions, denoted by $\mathbf{x}_t$ and satisfying both $\mathbf{x}_t^T\theta_t = 0$ and $\mathbf{x}_t^T\mathbf{1} = 1$, and active portfolio positions denoted by $\mathbf{y}_t$ respectively, thus

$$\mathbf{w}_t = \mathbf{x}_t + \mathbf{y}_t.$$

Then the total portfolio return can be decomposed as follows

$$\mathbf{w}_t^T\mathbf{r}_t = \mathbf{w}_t^T\mu_t + \mathbf{y}_t^T\theta_t .$$

ASSUMPTION 6: The active positions depend on our forecasts of residual return distribution but are independent of the market portfolio. So the deviations, $\mathbf{y}_t$, are a function of the forecasts, $\mathbf{m}_t$, but are independent of μ_t and $\mathbf{x}_t$.

Given that $\mathbf{y}$ is independent of $\mathbf{x}$ and μ we observe that $E[(\mathbf{w}^T\mu)(\mathbf{y}^T\theta)] = 0$ and therefore the decomposition of the total portfolio return has a market component ($\mathbf{w}_t^T\mu_t$) and an independent residual return ($\mathbf{y}_t^T\theta_t$). This essentially is the theory that Treynor and Black (well outlined in Bodie, Kane, and Marcus 1996) used in their original security selection model, and well aligned with the work of Grinold and Kahn. From here the proof that maximising the information ratio will maximise total portfolio Sharpe ratio is trivial.

In practice the market component is replaced with a specified benchmark portfolio, denoted in our notation by the position vector $\mathbf{x}_t$ rather than $\mathbf{w}_t$. The remainder term from the total portfolio return, $\mathbf{y}_t^T\mu_t$, is then attributed to the active manager. This action is sound provided the $\mathbf{y}_t^T\mu_t = 0$, i.e. active positions are "beta neutral" but not otherwise. For example the active manager with no insight could create an active portfolio which is simply the market portfolio but with higher beta. This active portfolio would exhibit positive information ratio but the total Sharpe ratio will remain unchanged (bar an active management fee). We believe this to be an issue not appropriately measured in practice.

Almost directly from Assumption 6 and equation (12) we can see that the task facing active managers is identifying the functional form of $\mathbf{y}_t$ as a function of $\mathbf{m}_t$ that maximises information ratio. This will be the meandering objective of the remainder of our article.

As a simplifying aside: by definition the active positions must sum to zero, i.e. $\mathbf{y}_t^T\mathbf{1} = 0$. We can drop this restriction with one assumption.

ASSUMPTION 7: The active manager can take an active position of any size in the market portfolio.

We can decompose the active positions into the position in the market portfolio and the positions in the other assets. Our vector $\mathbf{y}_t$ denotes the latter. From Assumption 7 we can set our active position in the market portfolio equal to $-\mathbf{y}_t^T\mathbf{1}$, so that the active positions always sum to zero, yet $\mathbf{y}_t$ is unrestricted.

3.6. Deriving the information ratio

The variable of interest to the trustee in measuring the performance of the active manager over any single time period is the active portfolio residual return, defined as $\mathbf{y}_t^T\theta_t$. Measuring active portfolio returns over a large multiple of time intervals eliminates the element of luck, where the residual return may happen to come from the tail of its probability distribution. Treynor and Black showed that the appropriate measure is information ratio, denoted by IR, which is defined as the expected residual return divided by the risk of residual return, written as

$$\text{IR} = \text{E}[\mathbf{y}^T\theta] / \text{V}[\mathbf{y}^T\theta]^{1/2}.$$

In our analysis we require that the information ratio be expressed as a function of the abnormal return forecasts. We do this by noting that IR is related to a function *h* defined as

$$\text{IR} = 1/(h - 1)^{½},$$

and in Appendix A.1 we show that we can express h as

$$h = \text{E}[\mathbf{y}^{\text{T}} (\mathbf{mm}^{\text{T}} + \mathbf{V})\, \mathbf{y}] \,/\, \text{E}[\mathbf{y}^{\text{T}}\mathbf{m}]^2.$$

This will be a useful expression for us because different active management approaches are defined by the function relating **y** to **m**. We can then substitute that function into equation (15) and measure the information ratio of that active management approach.

4. Considering various approaches to active portfolio position taking

Having derived an expression for information ratio as a function of the active positions, let us now apply it to a variety of approaches to active portfolio management. We shall consider three approaches. The first is what might be considered the "standard" approach, the second is a naïve alternative, and the third is what we identify to be the optimal approach.

4.1. A modern portfolio theoretic (MPT) approach to active portfolio management

4.1.1. Deriving the active positions

Modern portfolio theory says that the optimal active position vector over a single time period is the result of a mean/variance optimisation. We show in Appendix A.2 that the functional form of $\mathbf{y}_t$ that does this is

$$\mathbf{y}_t = c\, \mathbf{W}^{-1}\, \mathbf{m}_t.$$

where c^{-1} denotes the investor's coefficient of (active) risk aversion and **W** denotes the investors covariance matrix for the forecast errors.

It is critical to note that we have <u>*not*</u> used the matrix **V** as our covariance matrix of forecast errors in equation (16). This is because although **V** is the theoretically correct variable there are two reasons why it might not be used in practice. First, the covariance matrix of forecast errors is the choice of the active manager and there is nothing to inhibit the active manager from using any values for **W**. Second, different classes of assets are typically awarded to different active managers. Therefore the active manager must assume a zero correlation with any asset that he/she does not manage so that the active positions are not a function of another active manager's forecast of some other assets. The cost to performance of this separation is discussed in the next section.

4.1.2. Identifying the information ratio

With the functional form of $\mathbf{y}_t$ identified we can substitute equation (16) into our equation for information ratio, equation (15), and therefore identify the information ratio of the mean/variance optimal approach to active management.

We show in the Appendix A.3 that if $\mathbf{y}_t = c\, \mathbf{W}^{-1}\mathbf{m}_t$ then

$$\mathrm{IR} = \mathrm{tr}(\boldsymbol{\Omega}\mathbf{W}^{-1}) \,/\, (\mathrm{tr}(\boldsymbol{\Omega}\mathbf{W}^{-1}\boldsymbol{\Sigma}\mathbf{W}^{-1}) + \mathrm{tr}((\boldsymbol{\Omega}\mathbf{W}^{-1})^{2}))^{½}$$

where tr( ) denotes the trace of a matrix.

This is a very useful result. As we shall show in the next section it can be considered a generalisation of Grinold and Kahn's fundamental theorem of active management. For this reason we term equation (17) the *Generalised Fundamental Law of Active Management*.

Our expression is general in that it allows for different ICs across the assets and also allows for non-zero correlations between asset returns, forecasts, and forecast errors. This is particularly useful in identifying whether or not it is optimal to allow two correlated assets to be managed by one manager who accounts for this correlation, or to separate out the two assets to two specialist active managers. The trade off is that whilst the two specialists cannot account for the correlation and therefore the overall IR will be reduced, the introduction of two specialists will increase the individual IRs relative to the single generalist by enough to offset this correlation.

For example, suppose our residual returns are correlated but our forecasts are uncorrelated so that all of the residual return correlation appears in the correlation of the forecast errors. Then equation (17) can provide us with the total information ratio in this correlated situation (for simplicity we have assumed that IC, σ, ω, and v are equal for both assets). Chart 1 illustrates how the total IR will increase as either the individual asset IR increases or the residual return correlation increases. Now suppose we ignore the forecast error correlation. It is easy to show that the total IR is independent of the residual return correlation, and is therefore a function solely of individual asset IR. These cases are represented in Chart 1 by the points mapping to correlation equal to zero. We can then observe from Chart 1 that a 0.2 increase in information ratio (annualised assuming monthly time steps) will be sufficient to offset a residual return correlation of as much as 0.65.

Buckle(2000a) also shows the increase in information ratio through using specialist managers is sufficient to offset even quite large correlations. This example is just one application of the generalised fundamental law of active management. Buckle (2000a) goes on to discuss many others.

4.1.3. Assuming zero correlations

As we have discussed, assuming zero correlations in $\mathbf{\Sigma}$, $\mathbf{\Omega}$, and $\mathbf{V}$ may be reasonable. In fact Grinold and Kahn (1999) make this assumption in their seminal textbook on active portfolio management. Doing so provides two interesting results, which we shall now review.

Reconciling with Grinold and Kahn

In addition to the zero correlation assumption, let us assume that the residual return variances and the ICs are equal. Substituting these assumptions into equations (8), (9), and (17) gives us that

$$\text{IR} = [\text{BR}\times\text{IC}] / [\text{BR}(1 + \text{IC}^2)]^{½}$$

where BR denotes the dimension of **Σ**. Noting that the IC in practice is rarely higher than 0.3 we adopt Grinold and Kahn's assumption that IC^2 is negligible. This the gives us that

$$\text{IR} \approx \text{BR}^{½}\ \text{IC}$$

and we have Grinold and Kahn's fundamental law of active management.

Validating the linear rule

Given the zero correlation assumption, **W** is a diagonal matrix in which case the active position in the *i*th asset is given by

$$y_{it} \propto m_{it}.$$

This is an interesting result as it validates the use in practice of the simple but robust rule of taking an active position in an asset in proportional to the forecast of that asset return. See Buckle (1999) and Buckle (2002) for further details.

4.1.4. The univariate case

Some of the conceptual simplicity of our argument is lost in the multi-dimensional setting. For this reason we consider the special case of one dimension. In this case we can write that $y_t = c\ m_t / w^2$, and

$$\text{IR} = \text{IC} / (1+\text{IC}^2)^{½}.$$

As we might expect IR is an increasing function of IC with a lower limit of IR = 0 when the forecaster has no ability as measured by IC = 0. It is perhaps a little more surprising that there is an upper bound to IR when the forecaster has perfect foresight. It is easy to see that IR will increase to an upper bound of $1/\sqrt{2} = 0.707$ when IC=1.

4.2. A naïve approach to active portfolio management

4.2.1. The active position

As an alternative to the MPT approach to position taking, let us consider a more naïve strategy. Let us assume that we completely disregard any information in our forecast distribution with the exception of the direction of the forecast. We represent this by taking positions proportional to the sign of the forecast, thus

$$\mathbf{y}_t = K \,\mathrm{sign}(\mathbf{m}_t)$$

where K denotes some arbitrary constant of aggressiveness.

4.2.2. Deriving the information ratio

Again considering the univariate scenario, we show in Appendix A.4 that the information ratio for this naive strategy is given as

$$\mathrm{IR} = \mathrm{IC} / (\pi/2 - \mathrm{IC}^2)^{1/2}.$$

As in the MPT approach IR is an increasing function of IC also with a lower limit of IR = 0 when IC = 0. Like the MPT approach, but perhaps less surprising, the information ratio of the naïve approach also has an upper bound when the forecaster has perfect foresight. What is surprising is that in this case the upper bound is $1/(\pi/2-1)^{1/2}$ = 1.324.

4.2.3. The paradox?

The upper limit of the information ratio is _greater_ using the naive approach (1.324) than the MPT approach (0.707). Given that the latter strategy is considered the optimal approach, there appears to be a paradox here. The superior performance of the naïve approach does not occur just at the limit. It is trivial to show that the point at which the naive approach yields a higher information ratio is when the IC is greater than $(\pi/4 - ½)^{½} = 0.53$. Even below this level the naïve approach does not under-perform the MPT approach by very much.

This result has an appeal in that it is consistent with, and adds credence to, the adage that it is more important to forecast direction correctly than it is to forecast the magnitude of the returns.

We are however left with the question as to what is the approach to active portfolio management that maximises the information ratio?

4.3. The maximum information ratio approach to active portfolio management

We have seen that the MPT approach to active portfolio management does not maximise the information ratio because the naïve approach produces a greater one. We now derive the approach that maximises the information ratio and analyse its characteristics.

4.3.1. Deriving the active positions

The key result of this article is that we have found the function relating active positions to forecasts that maximises the information ratio. We shall express this result in the form of a theorem that we have termed the *Theorem of Active Portfolio Management*. We have deferred the proof of this theorem to Appendix A.5.

Although we have made an assumption on the distribution of returns (Assumption 5) we would like to point out that this theorem requires no distributional assumptions.

THEOREM (Theorem of Active Portfolio Management)

The single period investment position that maximises the multi-period information ratio is the result of a mean/second-moment optimisation.

Mathematical Representation of the Theorem

Given a vector of abnormal return forecasts, denoted by $\mathbf{m}_t$*, and a vector of asset residual returns, denoted by* θ_t*, then the vector of active positions, denoted by* $\mathbf{y}_t$*, that maximises the information ratio, defined as* $\mathrm{IR} = \mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta] / \mathrm{V}[\mathbf{y}^{\mathrm{T}}\theta]^{1/2}$*, is given as*

$$\mathbf{y}_t = c\,(\mathbf{m}_t\mathbf{m}_t^{\mathrm{T}} + \mathbf{V})^{-1}\mathbf{m}_t$$

where $\mathbf{V}$ *is the covariance matrix of residual returns conditional on the forecast, and* c^{-1} *is the scalar coefficient of (active) risk appetite.*

In the univariate case we can write this as

$$y_t = c\, m_t / (m_t^2 + v^2)$$

where the notation is taken from previous sections.

It is interesting to study the characteristics of this optimal position function. To aid that study Chart 2 illustrates the active positions in the univariate case for each of the three approaches as a function of the forecast. As can be observed, like the MPT and naïve approaches, the approach of using the theory of active portfolio management (TAPM) generates zero positions when the forecasts are zero. As the forecasts increase so the positions increase in a similar fashion to the MPT approach. Then, contrary to modern portfolio theory, as the forecasts increase further, the positions begin to level off exhibiting a similar profile to the naïve approach. Then as the forecasts increase yet further the positions from the TAPM approach decrease toward zero again.

Probably the most significant characteristic of positions generated from a mean/second-moment optimisation is that there is a maximum position. It can be shown that in the univariate case the maximum occurs when the forecast is one standard deviation in size. In the multivariate case the maximum position occurs when $\mathbf{m}_t\mathbf{m}_t^{\mathrm{T}} = \mathbf{V}$, although this maximum can only be achieved if **V** has a particular form. From a trustee's perspective the fact that the active manager has a limit to the position size provides an extra level of implicit risk control.

The most significant negative feature of mean/second-moment optimisation is that we have lost the attractive feature of mean/variance that when the covariance matrix is diagonal each asset position will only be a function of the forecast of that asset.

4.3.2. Deriving the information ratio

Unfortunately we have not been able to derive an expression for the information ratio of the mean/second-moment approach in the multivariate case, but have however been able to derive it in the univariate case. In Appendix A.6 we show that the information ratio of the univariate mean/second-moment approach is given by

$$\mathrm{IR} = [\phi(x) / (x\Phi(-x)) - 1]^{1/2}$$

where $x = (1\text{-}IC^2)^{½} / IC$ and $\phi()$ and $\Phi()$ are the standard normal pdf and cdf respectively.

This expression is not easy to interpret, and is clearer in graphical form. Chart 3 plots the IR of the TAPM approach together with the IRs of the MPT and naïve approaches as the information coefficient increases. Just as for the MPT and naive approaches, the value of IR for the TAPM approach equals 0 when IC equals 0. However, unlike the first two approaches, the IR for the TAPM approach has no upper bound. As the information coefficient tends to 1, IR tends to infinity. Implying that risk free non-zero portfolio residual returns are available.

Graphically, the optimal property of the TAPM approach is clear, although when the information coefficient is no more than moderate, all three approaches have similar information ratios.

4.3.3. An explanation of our findings

The key to explaining the surprising results that we have outlined above lies with the time varying residual return expectations. In the case of the MPT approach, the varying expectations imply that the amount invested must also vary through time. However, modern portfolio theory considers the investment amount a fixed quantity. Therefore, what the MPT approach does not take into account is the additional portfolio return volatility that arises from varying the investment amount (a point noted by Bodie, Kane and Marcus 1993).

In the environment of low volatility of abnormal return forecast (reflected by IC close to 0), the MPT approach must be near to the optimal IR (because the investment amount will hardly vary in which case MPT is known to be optimal). However, as this volatility increases, the volatility of the active position increases, so the active portfolio return volatility also increases. Such an increase is not present in volatility of the active portfolio returns from the naïve approach return because the active position is constant. Eventually then, it can be shown that the volatility of the returns from MPT approach will offset that approach's return expectation and, in terms of maximising the information ratio, the naïve strategy will begin to dominate.

The TAPM approach goes even further than the naïve approach as it dampens down the positions that are generated from large expected return forecasts. This dampening reduces the likelihood of very large returns, positive or negative, and therefore reduces the volatility of the resulting active portfolio returns. Recall that we earlier observed that the MPT approach was negatively impacted by not accounting for the volatility in the expected returns. The TAPM approach accounts for this.

5. A critical review of assumptions

We recognise that we make a strong statement in suggesting that active managers should not be using the MPT approach to generate investment positions, and it is incumbent of us to critically assess our work.

5.1. Assumptions on the trustee's objective

Our key assumption is that the trustee maximises Sharpe ratio. This assumption requires only that the trustee assumes a single period investment, uses modern portfolio theory, and uses the market separation property. These last two assumptions are essentially beyond question, although the market separation property does require unlimited borrowing and lending of the risk free asset, and both require trading to be free of taxes, transactions costs, and other frictions.

For modern portfolio theory to be consistent with the decision theoretic approach of maximising expected utility we do require that utility of wealth is quadratic or that returns are normally distributed. For many practical purposes asset returns are considered normal (although evidence of heavier tails has been found, see for example Fama 1965 and Blattberg and Gonedes 1974). However, even if asset returns are normally distributed, our active portfolio returns (made up of the product of a normal distribution and the distribution of our positions) may not be normally distributed (a point noted by Dybvig and Ross 1985). It is then theoretically arguable that the trustee should be using the utility theoretic approach directly, rather than mean/variance, in which case the Sharpe ratio may not be the appropriate measure. In practice this issue is ignored, although the academic community have recently indicated the risk of that stance. In particular skew and kurtosis can render the expected return and variance poor measures of return location and risk respectively. Ergo the Sharpe ratio becomes an inaccurate measure of risk-adjusted return. For example Goetzmann et al (2002) illustrate how better than market Sharpe ratios are obtainable without any forecasting ability simply by using sophisticated derivative strategies to make the portfolio return distribution heavily skewed and fat tailed. To some extent this does not impact us since we have not explicitly attempted to alter the distribution of portfolio returns so as to improve information ratio, but rather have optimally used actual forecasting ability. We do however need to be careful that our active portfolio returns do not exhibit excessive skew and kurtosis. It is difficult to specify the exact distribution of the portfolio returns, but in the MPT approach we can show that the coefficient of skew is approximately 6×IC and the coefficient of kurtosis is approximately 9. We take comfort in the fact that the kurtosis of the other two approaches should be closer to normality, and skew is only significant once IC becomes large, which is not borne out in practice.

We are left to justify our single period investment assumption. Certainly over an infinite time period we should expect the trustee to make rebalances to the allocation of investment wealth, but the spirit of the task is intended as a single period decision and over a shorter period of say three to five years our assumption is accurate.

5.2. Assumptions on the active manager's objective

We have assumed that the active manager should maximise the information ratio in a multi period environment.

Much has been written about the correct objective for the multi period problem (early articles include Samuelson 1969, Hakannson 1970, and Fama 1970, followed by the textbooks of Merton 1990 and Duffie 1996), and the Sharpe ratio has been identified as sub-optimal in a multi-period context (see for example Bodie, Kane and Marcus 1996). But the crucial point is that these theories are not relevant here, for the portfolio management objective is not something that the active manager can choose, but instead remains the responsibility of the trustee.

Given that the trustee faces a single period decision we take maximising the information ratio as the appropriate criterion from the original work of Treynor and Black (1973), and the modern textbooks of Bodie, Kane, and Marcus (1996) and Grinold and Kahn (1999) for example.

As Bodie, Kane, and Marcus (1996) point out; to be strictly theoretically correct the trustee should continually allocate differing amounts to the active manager as a function of the size of active manager forecasts. Then both the trustee and the active manager have multi period problems, but it is operational impractical, and prohibitively expensive, for the trustee to do this and the single period problem remains the pertinent one.

5.3. Assumptions on returns and forecasts

Our first three assumptions: that the efficient market hypothesis holds in the long run; that any abnormal return is only expected to exists in the short run (if at all); and that our forecasting of any abnormal return is unbiased, all appear to be reasonable. It is Assumption 4 where we assume that our forecasts and the errors in our forecasts are independent that is debatable. We are unsure of the implications of a violation of this assumption.

Assumption 5, the normal distribution assumption, is not a concern to us as we only use it in parts of our article and it is entirely in keeping with mainstream theory although we recognise that it might not hold up to empirical study.

Not only is Assumption 6 (active positions are independent of the market) reasonable, but a violation should be of concern to a trustee, because the active manager should be attempting to maximise the *independent* residual return (risk adjusted). In practice the activities of the active manager can become connected to the market if the trustee imposes active position constraints that are dependent on the market. In this event it is particularly important to ensure that the *independent* residual returns are being measured.

Our final assumption that the active manager can take any size position in the market portfolio is commonplace in practice, although the results in the article could be derived subject to the constraint that $\mathbf{y}_t^T\mathbf{1} = 0$ if required.

Our biggest concern with the framework we have identified here is the fact that we have assumed the covariance matrices be known and fixed in time. That they are fixed through time does not permit us to be more confident about our forecasts at one point in time relative to another. However, the alternative, assuming time varying covariance matrices leads to non-normal distributions and all the complications that that entails.

More of a concern in practice is the fact that the covariance matrices are assumed known. This can lead to significant problems. In particular it is not possible to identify ex-ante market return sensitivities (i.e. betas) and therefore the two components of the return decomposition may not be independent ex-post. This leads to all the problems discussed in Section 3.5. Additionally, implicit in taking forecast variance as known is the assumption that our forecast variance is reflective of our ability to forecast (i.e. that $\omega = IC\sigma$). If this is not so, and in practice it almost certainly will not be, then we need to scale our forecasts in our relationship between forecast and return so that $\theta_t = \mathbf{\Lambda m}_t + \boldsymbol{\varepsilon}_t$, where Λ denotes a diagonal scaling matrix that might be termed the humility multiplier. It is then possible to recalculate all of the above equations incorporating this new scaling factor. Of course ex-post we can simply multiply the forecasts by Λ and therefore proceed without it.

Also embedded in our framework is the assumption that the ICs are known. Once again in an ex-ante context this assumption is violated. In terms of our variances we could adjust forecast variance with our humility multiplier, Λ, but in this situation this too is unknown. In terms of correlations, in an ex-ante context it may be that we cannot use ICs at all and must resort to $\mathbf{\Sigma} = \mathbf{\Omega} + \mathbf{V}$ directly in order to identify $\mathbf{V}$. Even this relies on the independence assumption between forecasts and errors.

6. Existing literature

Following Treynor and Black's original security selection model there is little written on a structure for active portfolio management, with a key exception being the textbook of Grinold and Kahn (1999). We have followed their structure quite closely, and have generalised and expanded many of their results. More broadly, much of the work in active management can be cast in a total portfolio context (such as the link between Sharpe ratios and information ratios) and as such we can refer to literature in that area.

The most well known article in the financial literature that focuses on the issue of changing forecasts is by Dybvig and Ross (1985). Although the article focuses on the role of the security market line, there are some key results with relation to our own work.

Dybvig and Ross discuss the impact on portfolio performance from changing positions that inevitably arise from changing return forecasts. They also identify the difference between the trustee's single period decision and the active manager's multi period decision, and find that varying positions in a portfolio add to the portfolio risk from a trustee's perspective.

Dybvig and Ross demonstrate that mean/variance efficient positions generated from return forecasts where the active manager has information could yield an inefficient portfolio in the mean/variance sense (i.e. a portfolio with lower Sharpe ratio than the market), but share the conclusions of Mayers and Rice (1979) in that, for the special case of security selection, the managed portfolio will have a higher Sharpe ratio than the market, assuming the active manager has information and uses mean/variance efficient positions.

Hansen and Richard (1987) provide results relating to portfolio efficiency (in a mean/variance sense) through the use of conditional information in an asset-pricing context. In the context of our work they find that the unconditional Sharpe ratio is not necessarily maximised by a sequence of single period mean/variance optimisations if the expected returns change through time.

Our interest is in furthering these results by finding the portfolio of *maximum* information ratio. Ferson and Siegel (2001) have considered a similar problem but in the total portfolio context where positions must then sum to one. They also find that single period investment positions must exhibit dampening to maximise long run Sharpe ratio, and by making certain assumptions such as the risk free rate is zero (which is entirely appropriate for residual returns) we can reconcile our results with their equations for optimal positions.

We would like to make it clear that we have endeavoured to find the active portfolio management approach that converts insight in asset returns into additional portfolio return in a way that maximises information ratio, and therefore Sharpe ratio. To this end we consider our work largely detached from the recent articles such as the one by

Goetzmann et al (2002) where active portfolios positions can be generated using sophisticated derivative strategies that generate high information ratios even though no insight into asset returns exists.

7. Conclusions

There are several main conclusions to be drawn from our work. We have detailed a formal mathematical framework for considering active portfolio management. It is consistent with the existing literature, and in particular deals with the multi-period setting. Using this framework we have derived a formula for identifying the covariance matrix of forecast errors to be used in our active portfolio construction. We noted that this matrix may need to be adjusted if the active manager has only been allocated a subset of the asset universe. We have argued that the practice of using the total asset return correlations is poor, and in fact suggested that zero correlations may be reasonable thereby validating linear allocation rules for active management. We also showed that there may be a difference between residual portfolio returns that are uncorrelated with the market, and returns in excess of the benchmark. As such the trustee should ensure that performance is appropriately attributed.

Assuming an MPT approach to active portfolio management we have derived a formula for the information ratio. We showed that this formula is a more general version of the existing fundamental law of active management and have therefore termed it the generalised fundamental law of active management. Using our generalised law we showed that it is better to use specialist active managers at the expense of ignoring correlations than to use a lower quality manager who accounts for asset correlations.

Also by using our generalised law we showed that the MPT is not the active portfolio management approach that maximises information ratio as it can be beaten by a naïve approach, thereby validating the folklore that it is more important to forecast direction of return correctly than it is to forecast magnitude. Further, we put forward our theory of active portfolio management, which states that the active positions from a mean/second-moment optimisation are the ones that maximise information ratio.

In terms of further work, our primary task is to identify the impact of the violations of our assumptions through ex-ante uncertainty, or by measurement over a short time period where sample averages do not match distributional expectation. This is particularly important for performance measurement of active managers, and to this end we have already found that we can add a term to the fundamental law of active management to account for either the positions or the residual returns not averaging zero. It seems that the new term dominates the information ratio with only a small impact coming from IC.

Appendix

A.1. Expression for information ratio

By definition

$$\mathrm{IR} = \mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta] \,/\, \mathrm{V}[\mathbf{y}^{\mathrm{T}}\theta]^{½} = \mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta] \,/\, (\mathrm{E}[(\mathbf{y}^{\mathrm{T}}\theta)^2] - \mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta]^2)^{½}.$$

So, by defining h as

$$\mathrm{IR} = 1 \,/\, (h - 1)^{½}$$

we can express h as

$$\mathrm{h} = \mathrm{E}[(\mathbf{y}^{\mathrm{T}}\theta)^2] \,/\, \mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta]^2.$$

Now, using conditional expectations

$$h = E_m[E[\mathbf{y}^T\theta\theta^T\mathbf{y}|\mathbf{m}_t]] / E_m[E[\mathbf{y}^T\theta|\mathbf{m}_t]]^2.$$

Given that $\mathbf{y}_t$ is a function of $\mathbf{m}_t$, the conditional expectations can be written as

$$E[\mathbf{y}^T\theta|\mathbf{m}_t] = \mathbf{y}_t^T E[\theta|\mathbf{m}_t] = \mathbf{y}_t^T\mathbf{m}_t, \text{ and}$$

$$E[\mathbf{y}^T\theta\theta^T\mathbf{y}|\mathbf{m}_t] = \mathbf{y}_t^T E[\theta\theta^T|\mathbf{m}_t]\mathbf{y}_t = \mathbf{y}_t^T (\mathbf{m}_t\mathbf{m}_t^T+\mathbf{V})\, \mathbf{y}_t.$$

We can therefore write h as

$$h = E[\mathbf{y}^T (\mathbf{m}\mathbf{m}^T+\mathbf{V})\, \mathbf{y}] / E[\mathbf{y}^T\mathbf{m}]^2.$$

A.2. The function form of positions using an MPT approach

The MPT approach involves optimising the expectation of forecast return distribtuion offset by a penalised variance of forecast return distribution. Thus we optimise an objective function, denoted by f_1, given as

$$f_1(\mathbf{y}_t) = \mathbf{y}_t^T\mathbf{m}_t - ½\, c^{-1}\, \mathbf{y}_t^T\mathbf{W}\mathbf{y}_t$$

where c^{-1} denotes the investor's coefficient of risk aversion, and $\mathbf{W}$ denotes the covariance of forecast errors as specified by the investor.

Setting the differential of f_1 with respect to $\mathbf{y}_t$ equal to zero yields the mean/variance optimal position vector as

$$\mathbf{y}_t = c\, \mathbf{W}^{-1}\, \mathbf{m}_t.$$

A.3. The Generalised Fundamental Law of Active Management

We take the active positions vector from the MPT approach and therefore use equation (35) and for notational convenience take c=1 observing that a constant multiplier of positions does not affect information ratio. Hence we adopt

$$\mathbf{y}_t = \mathbf{W}^{-1}\mathbf{m}_t.$$

The residual return on our portfolio is the product of the positions and the asset residual returns and therefore given by

$$\mathbf{y}_t^{T}\theta_t = \mathbf{m}_t^{T}\mathbf{W}^{-1}\theta_t.$$

Using our normal distribution assumption, the portfolio return distribution, conditional on our forecast of expected return, is given as

$$\mathbf{y}^{T}\theta|\mathbf{m}_t \sim \mathrm{N}(\mathbf{m}_t^{T}\mathbf{W}^{-1}\mathbf{m}_t, \mathbf{m}_t^{T}\mathbf{W}^{-1}\mathbf{V}\mathbf{W}^{-1}\mathbf{m}_t).$$

Using the reasoning as in Appendix A.1 we express the information ratio in the form

$$\mathrm{IR} = (\mathrm{E}[(\mathbf{y}^{T}\theta)^2] / \mathrm{E}[\mathbf{y}^{T}\theta]^2 - 1)^{-½}$$

and are left to derive $\mathrm{E}[\mathbf{y}^{T}\theta]$ and $\mathrm{E}[(\mathbf{y}^{T}\theta)^2]$. In order to do this we continue the reasoning laid out in Appendix A.1 noting that the marginal expectations of $\mathbf{y}^{T}\theta$ and $(\mathbf{y}^{T}\theta)^2$ are given as

$$\mathrm{E}[\mathbf{y}^{T}\theta] = \mathrm{E}_{\mathbf{m}}[\mathrm{E}[\mathbf{y}^{T}\theta|\mathbf{m}_t]] \text{ and}$$

$$\mathrm{E}[(\mathbf{y}^{T}\theta)^2] = \mathrm{E}_{\mathbf{m}}[\mathrm{E}[(\mathbf{y}^{T}\theta)^2|\mathbf{m}_t]].$$

The first two zero centred moments of $\mathbf{y}^{T}\theta$ which are extracted directly from the distribution of $\mathbf{y}^{T}\theta$ thus

$$\mathrm{E}[\mathbf{y}^{T}\theta|\mathbf{m}_t] = \mathbf{m}_t^{T}\mathbf{W}^{-1}\mathbf{m}_t \text{ and}$$

$$\mathrm{E}[\mathbf{y}^{T}\theta^{\,2}|\mathbf{m}_t] = \mathbf{m}_t^{T}\mathbf{W}^{-1}\mathbf{V}\mathbf{W}^{-1}\mathbf{m}_t + (\mathbf{m}_t^{T}\mathbf{W}^{-1}\mathbf{m}_t)^2.$$

All that remains is to derive the marginal expectations of $\mathbf{y}^{T}\theta$ and $(\mathbf{y}^{T}\theta)^2$ from our conditional distributions. To do this we need to find $\mathrm{E}[\mathbf{m}^{T}\mathbf{V}^{-1}\mathbf{m}]$, $\mathrm{E}[\mathbf{m}^{T}\mathbf{W}^{-1}\mathbf{V}\mathbf{W}^{-1}\mathbf{m}]$ and $\mathrm{E}[(\mathbf{m}^{T}\mathbf{V}^{-1}\mathbf{m})^2]$. The first two of these expressions are easily shown to be

$$\mathrm{E}[\mathbf{m}^{T}\mathbf{W}^{-1}\mathbf{m}] = \mathrm{E}[\mathrm{tr}(\mathbf{m}\mathbf{m}^{T}\mathbf{W}^{-1})] = \mathrm{tr}(\mathrm{E}[\mathbf{m}\mathbf{m}^{T}]\mathbf{W}^{-1}) = \mathrm{tr}(\mathbf{\Omega}\mathbf{W}^{-1}), \text{ and}$$

$$\mathrm{E}[\mathbf{m}^{T}\mathbf{W}^{-1}\mathbf{V}\mathbf{W}^{-1}\mathbf{m}] = \mathrm{tr}(\mathbf{\Omega}\mathbf{W}^{-1}\mathbf{V}\mathbf{W}^{-1}).$$

The third expression is considerably harder. In Buckle (2000b) it is shown to be given by

$$E[(\mathbf{m}_t^T\mathbf{V}^{-1}\mathbf{m}_t)^2] = 2\mathrm{tr}((\mathbf{\Omega V}^{-1})^2) + \mathrm{tr}^2(\mathbf{\Omega V}^{-1})$$

These expressions can be plugged into equations (42) and (43) and then into equation (27) so the information ratio can then be expressed as

$$IR = \mathrm{tr}(\mathbf{\Omega W}^{-1}) \,/\, (\,\mathrm{tr}(\mathbf{\Omega W}^{-1}\mathbf{VW}^{-1}) + 2\mathrm{tr}((\mathbf{\Omega W}^{-1})^2))^{½}.$$

This expression can be tidied up by noting that the $\mathbf{\Sigma} = \mathbf{\Omega} + \mathbf{V}$ thus

$$IR = \mathrm{tr}(\mathbf{\Omega W}^{-1}) \,/\, (\,\mathrm{tr}(\mathbf{\Omega W}^{-1}\mathbf{\Sigma W}^{-1}) + \mathrm{tr}((\mathbf{\Omega W}^{-1})^2)\,)^{½}.$$

A.4. Deriving the information ratio of the naïve approach

We take the active positions vector from the naive approach and therefore use equation (22) and for notational convenience take K = 1 observing that a constant multiplier of positions does not affect information ratio. Hence we adopt

$$\mathbf{y}_t = \mathrm{sign}(\mathbf{m}_t).$$

Once again using the reasoning as in Appendix A.1 we express the information ratio in the form

$$IR = (E[\mathbf{y}^T\theta^2] \,/\, E[\mathbf{y}^T\theta]^2 - 1)^{-½}$$

noting that

$$E[\mathbf{y}^T\theta] = E_{\mathbf{m}}[E[\mathbf{y}^T\theta|\mathbf{m}_t]] \text{ and}$$

$$E[\mathbf{y}^T\theta^{\,2}] = E_{\mathbf{m}}[E[\mathbf{y}^T\theta^{\,2}|\mathbf{m}_t]].$$

The first two zero centred moments of $\mathbf{y}^T\theta$ which are extracted directly from the distribution of $\mathbf{y}^T\theta$ thus

$$E[\mathbf{y}^T\theta|\mathbf{m}_t] = \mathrm{sign}(\mathbf{m}_t)^T\mathbf{m}_t = |\mathbf{m}_t|^T\mathbf{1} \text{ and}$$

$$E[\mathbf{y}^T\theta^{\,2}|\mathbf{m}_t] = |\mathbf{m}_t|^T|\mathbf{m}_t| + \mathrm{sign}(\mathbf{m}_t)^T\mathbf{V}\,\mathrm{sign}(\mathbf{m}_t).$$

Specifically in one dimension, equations (53) and (54) can be simplified to

$$\mathrm{E}[y\theta|m_t] = |m_t| \text{ and}$$

$$\mathrm{E}[(y\theta)^2|m_t] = m_t^2 + v^2.$$

Trivially then we have that

$$\mathrm{E}[|m_t|] = \omega(2/\pi)^{½} = \mathrm{IC}\sigma(2/\pi)^{½} \text{ and}$$

$$\mathrm{E}[m_t^2 + v^2] = \sigma^2$$

Substituting these results into equation (27), the information ratio can then be expressed as

$$\mathrm{IR} = \mathrm{IC}\sigma(2/\pi)^{½} / (\sigma^2 - \mathrm{IC}^2\sigma^2(2/\pi))^{½}$$

This expression can be tidied up thus

$$\mathrm{IR} = \mathrm{IC} / (\pi/2 - \mathrm{IC}^2)^{½}.$$

A.5. Theorem of Active Portfolio Management

THEOREM (Theorem of Active Portfolio Management)

The single period investment position that maximises the multi-period information ratio is the result of a mean/second-moment optimisation.

<u>*Mathematical Representation of the Theorem*</u>

Given a vector of abnormal return forecasts, denoted by $\mathbf{m}_t$*, and a vector of asset residual returns, denoted by* θ_t*, then the vector of active positions, denoted by* $\mathbf{y}_t$*, that maximises the information ratio, defined as* $\mathrm{IR} = \mathrm{E}[\mathbf{y}^T\theta] / \mathrm{V}[\mathbf{y}^T\theta]^{½}$*, is given as*

$$\mathbf{y}_t = c\,(\mathbf{m}_t\mathbf{m}_t^T + \mathbf{V})^{-1}\mathbf{m}_t$$

where $\mathbf{V}$ *is the covariance matrix of residual returns conditional on the forecast, and* c^{-1} *is the scalar coefficient of (active) risk appetite.*

PROOF

If a position vector maximises IR, then that position vector must also minimise h where

$$\mathrm{IR} = 1/\,(h - 1)^{1/2} \text{ and } h = \mathrm{E}[(\mathbf{y}^{\mathrm{T}}\theta)^2]\,/\,\mathrm{E}[\mathbf{y}^{\mathrm{T}}\theta]^2.$$

Surprisingly, a multivariate version of the Schwarz inequality not only gives us the lower bound for h, but does so without requiring us to solve the two expectations in our formula. The implication is that the optimal investment approach is known and is independent of distributional assumption on the return forecasts.

THEOREM (multivariate Schwarz inequality)

For any two vector distributions $\mathbf{U}$ and $\mathbf{S}$

$$\mathrm{E}[\mathbf{U}^{\mathrm{T}}\mathbf{U}]\,/\,\mathrm{E}[\mathbf{U}^{\mathrm{T}}\mathbf{S}]^2 \geq 1/\mathrm{E}[\mathbf{S}^{\mathrm{T}}\mathbf{S}]$$

with equality holding if and only if $\mathbf{U} = c\mathbf{S}$ for some c.

Proof is deferred to Appendix A.7.

We set $\mathbf{U} = \mathbf{L}^{\mathrm{T}}\mathbf{y}$ and $\mathbf{S} = \mathbf{L}^{-1}\mathbf{m}$ where $\mathbf{L}$ is the Cholesky decomposition of $\mathbf{mm}^{\mathrm{T}} + \mathbf{V}$ so that $\mathbf{LL}^{\mathrm{T}} = \mathbf{mm}^{\mathrm{T}} + \mathbf{V}$ (see for example Rencher 1998 for details of Cholesky decompositions). Using our version of Schwarz's inequality, we have that

$$h \geq 1\,/\,\mathrm{E}[\mathbf{m}^{\mathrm{T}}(\mathbf{mm}^{\mathrm{T}} + \mathbf{V})^{-1}\mathbf{m}].$$

We know that equality only holds if $\mathbf{U} = c\mathbf{S}$, equivalently if

$$\mathbf{y}_t = c\,(\mathbf{m}_t\mathbf{m}_t^{\mathrm{T}} + \mathbf{V})^{-1}\mathbf{m}_t.$$

This expression represents the strategy that maximises information ratio.

The mean/second-moment optimisation approach has an objective function given as

$$f_2(\mathbf{y}_t) = \mathbf{y}_t^{\mathrm{T}}\mathbf{m}_t - ½\, c^{-1}\, \mathbf{y}_t^{\mathrm{T}}\, (\mathbf{m}_t\mathbf{m}_t^{\mathrm{T}} + \mathbf{V})\, \mathbf{y}_t$$

provided **V** is assumed as the covariance matrix (recall we used matrix **W** in the MPT approach).

Setting the differential of f_2 with respect to $\mathbf{y}_t$ equal to zero yields the mean/second-moment optimal position vector as

$$\mathbf{y}_t = c\, (\mathbf{m}_t\mathbf{m}_t^{\mathrm{T}} + \mathbf{V})^{-1}\mathbf{m}_t.$$

We observe that this is identical to our expression of the position vector that maximises information ratio.

A.6. Deriving the information ratio of the TAPM approach

Directly from equation (64) in Appendix A.5 we can see that in one dimension the minimum value that h can take is

$$1 / \mathrm{E}[m^2 / (m^2 + v^2)]$$

Where we once again have taken c = 1 for notational convenience. Since m_t is normally distributed, and using Rieman integrals (see for example Lindgren 1976, p119), the expectation can be written

$$\mathrm{E}[m^2 / (m^2 + v^2)] = \int_{(-\infty,\infty)} x^2 / (x^2 + v^2) \exp\{-½\, x^2 / v^2\} / (2\pi v^2)^{½}\, \mathrm{d}x.$$

Using the closed form integral result (see for example Gradshteyn and Ryzhik 1994, p383)

$$\int_{(0,\infty)} z^2 / (z^2 + \beta^2) \exp\{-z^2\, \mu^2\}\, \mathrm{d}z = \pi^{½} / (2\mu) - (\pi\beta / 2) \exp\{\mu^2\beta^2\}[1 - \mathrm{erf}(\beta\mu)],$$

we can show with some algebraic manipulation that

$$\mathrm{E}[m^2 / (m^2 + v^2)] = 1 - x\, \Phi(-x) / \phi(x)$$

where $x = \sigma/\omega = (1 - \mathrm{IC}^2)^{½} / \mathrm{IC}$, and ϕ and Φ are the probability density function and cumulative density function respectively of the standard normal distribution. We can then substitute this expectation into equation (68), and then substitute h into equation (28) to derive the result that for a given information coefficient, the maximum achievable information ratio is

$$\mathrm{IR} = [\phi(x) / (x\Phi(-x)) - 1]^{1/2}.$$

Appendix A.7 The multivariate Schwarz inequality

THEOREM (multivariate Schwarz inequality)

*For any two vector distributions **U** and **S***

$$E[\mathbf{U}^T\mathbf{U}] / E[\mathbf{U}^T\mathbf{V}]^2 \geq 1/E[\mathbf{V}^T\mathbf{V}]$$

*with equality holding if and only if **U** = c**S** for some c.*

PROOF

Note that a proof of the multivariate Schwarz inequality is provided here because we have been unable to find reference to this result in the statistical literature. We do not claim ownership of this result.

Our proof follows the logic of the univariate version in Lindgren (1976). We consider the non-negative random variable $(\mathbf{U} - c\mathbf{S})^T(\mathbf{U} - c\mathbf{S})$ where **U** and **S** are vector random variables and c is a scalar constant. The expectation of this expression must be greater, or equal to zero, and by expanding the quadratic this gives us that

$$E[\mathbf{U}^T\mathbf{U}] + c^2E[\mathbf{S}^T\mathbf{S}] - 2cE[\mathbf{U}^T\mathbf{S}] \geq 0.$$

For this quadratic inequality to hold for all c, the discriminant must be non-positive, so

$$E[\mathbf{U}^T\mathbf{S}]^2 - E[\mathbf{U}^T\mathbf{U}]\ E[\mathbf{S}^T\mathbf{S}] \leq 0$$

which, when rearranged in a form that suits are analysis, gives us that

$$E[\mathbf{U}^T\mathbf{U}] / E[\mathbf{U}^T\mathbf{S}]^2 \geq 1/E[\mathbf{S}^T\mathbf{S}].$$

For equality to hold we must have that the discriminant equals zero, so that

$$E[\mathbf{U}^T\mathbf{S}]^2 - E[\mathbf{U}^T\mathbf{U}]\ E[\mathbf{S}^T\mathbf{S}] = 0.$$

For this to hold for any value of **U** and **S** we can only have that $\mathbf{U} = c\mathbf{S}$.

References

Blattberg, R. C., and Gonedes, N. J. (1974) A comparison of the stable and student distributions as statistical models of stock prices. *Journal of Business*, **47**, pp224-280.

Bodie, Z., Kane, A., and Marcus, A. J. (1996) *Investments*. Third edition. Irwin.

Buckle (1999) A discussion on allocation rules, and a recommended approach for use in currency overlay. *Lee Overlay Partners Research Report*, buc-dj-0799b.

Buckle (2000a) Active management in more than one asset and the role of correlations. *Lee Overlay Partners Research Report*, buc-dj-0600a

Buckle (2000b) The fourth moment of a multivariate normal distribution. *Lee Overlay Partners Research Report*, buc-dj-0300a

Buckle (2002) The relationship between linear allocation rules and modern portfolio theory. *Lee Overlay Partners Research Report, in progress*.

Cohen, K. J. and Pogue, G (1967) An empirical evaluation of alternative portfolio models, *Journal of Business*, **40**, pp166-193

Duffie, D. (1996) *Dynamic asset pricing theory*. Second edition. Princeton University Press.

Dybvig, P. H., and Ross, S. A. (1985) Differential information and performance measurement using a security market line. *Journal of Finance*, **60**, pp383-399.

Fama, E. F. (1965) The behaviour of stock market prices. *Journal of Business*, **38**, pp35-105.

Fama, E. F. (1970) Multiperiod consumption-investment decisions. *American Economic Review*, **60**, pp163-174.

Ferson, W. E., and Siegel, A. F. (2001) The efficient use of conditioning information in portfolios. *Journal of Finance*, **56**, pp967-982.

Goetzmann, W. J., Ingersoll, M., Spiegel, M., and Welch, I. (2002) Sharpening Sharpe ratios, *Yale School of Management*, Working Paper.

Gradshteyn, I. S., and Ryzshik I. M. (1994) *Tables of Integrals, Series, and Products*. Fifth edition. Academic Press.

Grinold, R. C., and Kahn, R. N. (1999) *Active portfolio management: a quantitative approach for providing superior returns and controlling risk*. Second edition. McGraw-Hill.

Hakansson, N. (1970) Optimal investment and consumption strategies under risk for a class of utility functions. *Econometrica*, **38**, pp587-607.

Hansen, L. P., and Richard, S. F. (1987) The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, **55**, pp587-613.

Lindgren, B. W. (1976) *Statistical inference*. Third edition. Macmillan Publishing Co.

Markowitz, H. M. (1952) Portfolio selection. *Journal of Finance*, **7**, pp77-91.

Mayers, D., and Rice, E. (1979) Measuring portfolio performance and the empirical content of asset pricing models. *Journal of Financial Economics*, **7**, pp3-28.

Merton, R. C. (1990) *Continuous time finance*. Blackwell Publishers.

Samuelson, P. A. (1969) Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, **51**, pp239-246.

Sharpe, W. F. (1963) A simplified model for portfolio analysis. *Management Science*, **9**, pp277-293.

Sharpe, W. F. (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**, pp425-442.

Rencher, A. C. (1998) *Multivariate statistical inference and applications*. Wiley Series in Probability and Statistics.

Ross, S. A. (1976) The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, **3**, pp343-362.

Tobin, J. (1958) Liquidity preference as behaviour towards risk. *Review of Economic Studies*, **25**, pp65-68.

Treynor, J., and Black, F. (1973) How to use security analysis to improve portfolio selection. *Journal of Business*, **46**, pp66-86.

Chart 1 illustrates how an improvement in information ratio through the use of specialist active managers, rather than one generalist, will offset even large correlations between the assets, even though the specialists act independently and therefore cannot account for this correlation. For example two specialists with an information ratio of 0.8 will produce a superior total information ratio than one generalist with information ratio 0.6 who accurately accounts for a correlation as large as 0.65.

Chart 1. Improved information ratio offsets even large correlations

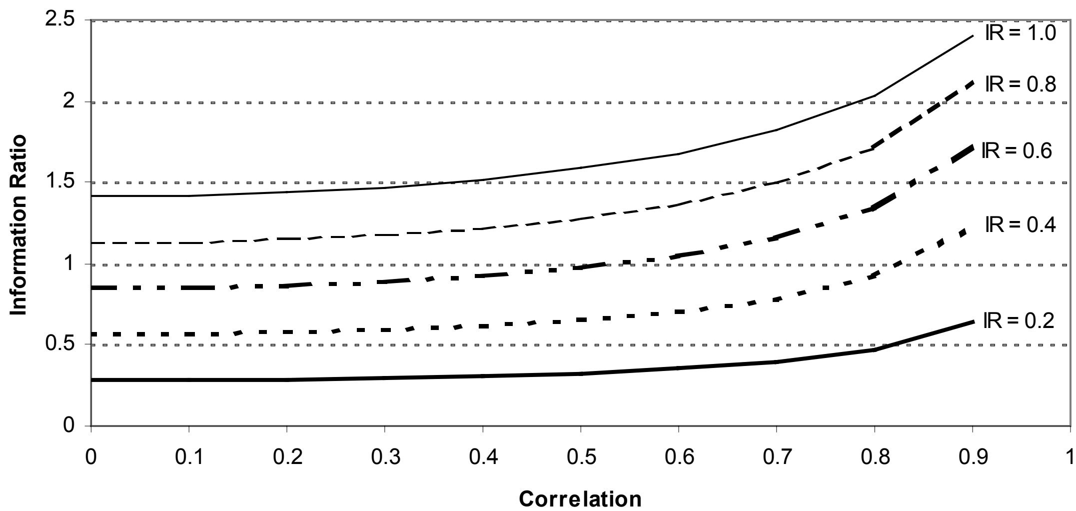

Chart 2 illustrates the positions taken under three approaches to active portfolio management. The modern portfolio theoretic approach is linear function of the forecast; the naïve approach is a constant function of the forecast; and the theory of active portfolio management approach is a dampened linear function of the forecast.

Chart 2: Positions from three apporaches to active portfolio management

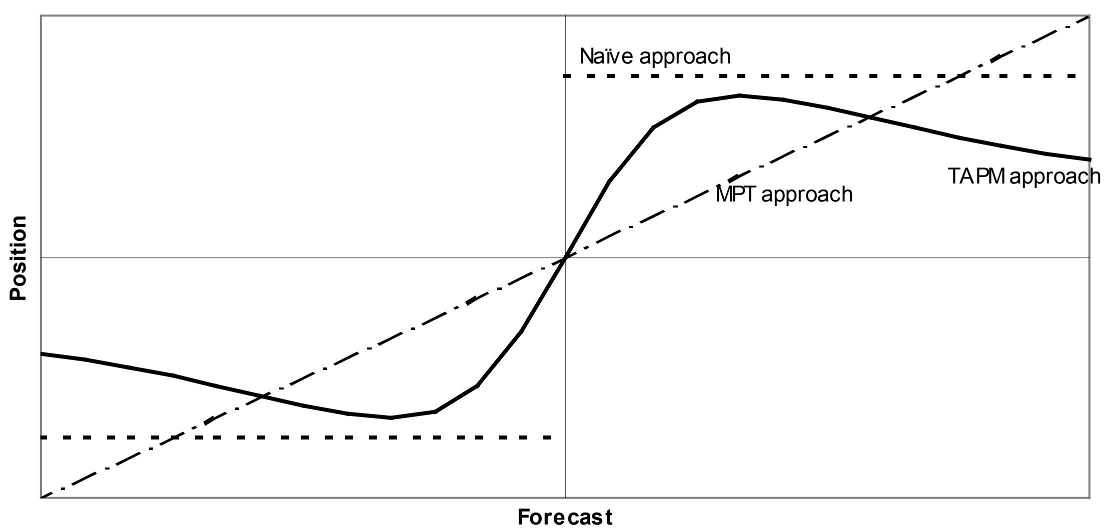

Chart 3 illustrates the information ratio from three approaches to active portfolio management. The modern portfolio theoretic approach outperforms the naïve approach until IC is greater than about 0.5. However, the theory of active portfolio management approach outperforms both of the other two. In practice IC is less than 0.3 in which case all the approaches perform similarly.

Chart 3: Information Ratios from three apporaches to active portfolio management