The Problem of Alpha and Performance Measurement

by:
Wayne E. Ferson[*]

first draft: May 27, 2008
this draft: September 22, 2009

Abstract

Studies of investment performance routinely use alpha, yet the literature has only partially resolved two fundamental questions about the use of alpha. First, if an investor faces a fund with a positive (negative) alpha, will the investor want to buy (sell) some of that fund? Second, if a manager has superior information, will he or she generate a positive alpha? This note revisits these fundamental questions. It shows that when alpha is defined based on the client's marginal utility function, then a client faced with a positive alpha would want to buy under general conditions. It also discusses operationalizing alpha and concludes with some suggestions for future research and practice in performance measurement.

** Ivadelle and Theodore Johnson Chair in Banking and Finance, and Research Associate, NBER, Marshall School of Business, University of Southern California, 3670 Trousdale Parkway Suite 308, Los Angeles, CA. 90089-0804. ferson@marshall.usc.edu www-rcf.usc.edu/~ferson (213) 740-6515.*

This was initially prepared as a keynote address for the 16-th Annual Pacific Basin Economics, Accounting and Management Conference and revised for the 2009 Northern Finance Association Meetings. I am grateful to George Aragon, Stephen Brown and Derek Horstmeyer for suggestions and discussion.

Introduction

Finance researchers have an easy familiarity with alpha, which measures the expected abnormal return of an investment. Alpha is so ubiquitous that it has become a generic, like Zerox™ or Google. Studies refer to CAPM alpha, three-factor alpha or four-factor alpha, assuming the reader hardly requires a definition. Investment practitioners routinely discuss their strategies in terms of their quest for alpha. Alpha can be active, conditional or portable. The number of investment firms with alpha in their names is truly staggering.

Despite the apparent familiarity with alpha, I think that the current literature too often fails to think rigorously about how alphas can and should be interpreted. The goal of this note is to review and synthesize the main issues with alphas in performance measurement, to offer modest extensions of existing results and to make some suggestions for future research and practice in the use of alphas.

There are two Fundamental Questions about the use of alpha. The first question is: When faced with a fund that has a positive (negative) alpha, should the investor want to buy (sell) that fund? The second question is: If a manager has superior information, will he or she generate a positive alpha? While the concept of alpha may be traced in some form back to Coles (1933), a substantial literature grappled with these two questions after alpha was developed within the CAPM (Sharpe, 1964) by Jensen (1968, 1972) and others. But this work, with a few exceptions, essentially died out in the late 1980s, leaving these two fundamental questions only partially resolved. Without a clean answer to these questions, it would seem that a large part of the literature on investment performance lacks a rigorous foundation.

As to the first question, whether an investor would wish to buy a positive-alpha fund, the literature offers some hopeful examples, but also many counterexamples. The simplest intuition for the attractiveness of a positive alpha is taught with the CAPM, where a combination of a positive-alpha fund, the market portfolio and cash can "beat the market" in a mean variance sense (higher mean return given the variance). However, as Roll (1978) emphasizes, a positive alpha implies that the market index is not mean variance efficient, so a mean variance investor would not want to hold it anyway, and if we change the index we can almost arbitrarily change the alphas. If the index is mean variance efficient the alphas are all zero. Considering an arbitrary (inefficient) benchmark, Dybvig and Ross (1985b) show (their Theorem 5), that a positive alpha measured relative to the benchmark implies that buying (shorting) some of the fund *at the margin*, will result in a higher Sharpe ratio than the benchmark, if the benchmark excess return is positive (negative). However, this marginal result does not extend to a discrete investment change.

Jobson and Korkie (1982) showed that given an inefficient index, a portfolio with weights proportional to the vector of assets' alphas, premultiplied by an inverse covariance matrix (the optimal orthogonal portfolio), can be combined with the index to generate a mean variance efficient portfolio. However, the weight in the optimal orthogonal portfolio for a positive alpha asset can be negative, and Gibbons Ross and Shanken (1989) provide examples where it is. So, even if a positive alpha is attractive at the margin to a mean-variance investor, it might not imply buying a positive alpha fund given a realistic discrete response.

Of course, the real world of investing is not the static one in which the above results apply. Returns are not normally distributed and mean-variance preferences, while a convenient analytical tool, are not very realistic. We would like a justification for alpha that works more generally. The results in this note are derived in a multiperiod setting, for general investors, and consider the optimal discrete response of the investor to an investment fund with alpha.

It seems natural to think that a portfolio manager may have better information about returns than the client investor, and I make that assumption here. When the literature addresses differential information, the problem of alpha becomes richer. The portfolio of a better-informed manager expands the opportunity set of the less-informed client, so the client would generally like to use the managed portfolio return. The problem is, the client might wish to short the fund even if it has a positive alpha (Chen and Knez, 1996).

The literature is replete with examples that suggest that alpha can't reliably indicate investments a client would want to buy. In many examples performance within the model is neutral but alpha is not zero. Jagannathan and Korajczyk (1986) and Leland (1999) show how you can get nonzero CAPM alphas by trading fairly priced options with no special skill. Ferson and Schadt (1996) show you can record negative alphas when performance is neutral if you don't account for public information. Roll (1978), Dybvig and Ross (1985b) and Green (1986) give examples of nearly arbitrary alphas when there is no ability. Goetzmann et al. (2007) show how most performance measures can be gamed to produce positive alphas through informationless trading. Thus, the existing literature suggests that the general answer to the first Fundamental Question is negative. It is typically not true that investors should buy a positive-alpha fund. But, if alpha doesn't indicate attractive investments, why do we routinely use it in the current literature as if it did?

Even if the answer to the first Fundamental Question is negative, alpha could still be useful if it reliably indicates managers with superior information. This leads to the second Fundamental Question, whether a manager with superior information will produce a positive alpha. Mayers and Rice (1979) argued for an affirmative answer. They assumed complete markets, quadratic utility and the CAPM. They also assumed that the manager either has no information about the market return (i.e., no timing information) and that either the expected conditional beta is the unconditional beta (generally not true) or that the agent's optimal consumption is unaffected by the information. Dybvig and Ingersoll (1982) showed that you can't marry complete markets with quadratic utility because it leads to negative state prices, and Verrechia (1980) gave a counterexample with quadratic utility to the more general proposition that the informed earn higher returns than the uninformed expect, based on the Mayers and Rice set up. Dybvig and Ross (1985a) generalized the Mayers and Rice result to avoid the complete markets assumption (their Theorem 2) but assumed that the manager has no information about the mean or variance of the uninformed client's portfolio. This early debate, which involved several authors including Admati and Ross (1985a,b), Cornell (1979) and Roll (1979), leaves much ambiguity about the second fundamental question.

Grinblatt and Titman (1989) provide perhaps the best positive answer in the literature to the question of whether an informed manager will generate a positive alpha. They use a single period model under the assumption of normality, and thus, mean variance preferences.[1] They consider alpha measured relative to a benchmark that is mean variance efficient given the uninformed client's information, and assume that the manager has nonincreasing Rubinstein (1976) risk aversion as a function of the portfolio return. Even under these assumptions, sadly, they find that alpha can be negative for an informed manager. Grinblatt and Titman then introduce a positive period weighting measure[2] and show that alphas relative to this measure will be positive if the better-informed manager has constant Rubinstein risk aversion (their Proposition A1) or has no timing information, or has selectivity information that is independent of both the benchmark and the weighting measure and optimally increases beta when receiving a positive timing signal about the efficient portfolio (their Proposition 2).

It seems that the conditions under which a manager with superior information will generate a positive alpha are fairly special, and there are examples where it won't be true. Dybvig

---

[1] They do allow for nontraded human capital, and thus a hedging demand related to human capital. The informed agent's optimal portfolio in their set up is conditional multifactor minimum variance efficient (Ferson, Siegel and Xu, 2006) given the client's information. With the additional assumption of normality, that will also be the case in the model developed below.

[2] The positive period weighting measure is a set of scalars $\{w_t\}$ that are strictly positive, sum to 1.0 and are bounded in the sample size T: $|\text{plim}(Tw_t)|<\infty$. The alpha for a portfolio with excess return r_t is defined as $\text{Plim}(\Sigma_t\ w_t r_t)$. A positive period weighting measure produces a zero alpha for a portfolio that is mean variance efficient conditioned on the uninformed client's information set, r_{Et}: $\text{Plim}(\Sigma_t\ w_t r_{Et}) = 0$. The alpha advocated in this note is an example of a positive period weighting measure, and thus has the properties described by Grinblatt and Titman when a single period model is assumed.

and Ross (1985a) and Grinblatt and Titman (1989) show that a manager that is a positive market timer can generate a negative alpha. Dybvig and Ross (1985a) and Hansen and Richard (1987) show that a portfolio can be mean variance efficient given the informed manager's knowledge, but appear mean variance inefficient to the uninformed client. In general, the answer from the literature to the second Fundamental Question is also negative.

This note revisits the two Fundamental Questions about alpha using a natural definition based on the stochastic discount factor (SDF) approach. This definition of alpha was proposed as early as Beja (1971), but the SDF approach has only become common in asset pricing in the years following the earlier literature that tried to address the Fundamental Questions. The SDF approach turns out to offer new insights on the Fundamental Questions about alpha. The insights are general, in that they are based on a multiperiod model and do not require normality. A mean variance efficient benchmark is not required. There is no need to rule out timing ability, nor is selectivity information required to be independent of timing information or otherwise restricted. The agent is allowed to have a general consumption response to the introduction of the managed portfolio. The results are not limited to marginal changes, but consider the optimal discrete responses.

The main result is a pretty clean answer to the first Fundamental Question. Under fairly general conditions, if alpha is positive (negative) from the client's perspective, he should buy (sell) some discrete amount of the fund. Given this result, there is no need to fully answer the second Fundamental Question for many practical purposes, although I do address it below for some special cases. The key to obtaining such a simple result is using the "right" definition of alpha, based on the client's marginal utility function. This rules out most of the pathological examples that the earlier literature struggles with. I do require some plausible restrictions on the client's portfolio response, as described below. And since the SDF alpha is client specific in general, the results raise issues for implementation as discussed below.

The closest work that I know of to this is Glosten and Jagannathan (1994) and Chen and Knez (1996). Glosten and Jagannathan (1994) start with the definition of alpha studied here. They then assume that client's SDFs are functions only of some traded benchmark portfolios and a small set of options strategies, and they focus on the resulting consensus or representative agent valuation of informed-manager strategies that may have option-like characteristics. A key focus of their analysis is to approximate the functional form of the expected payoff of the fund, given the benchmark returns. Among other things their approach provides insights about models for measuring market-timing ability, as discussed below.

Chen and Knez (1996) characterize general classes of "admissible" and positive admissible performance measures in the presence of information, concluding that performance measurement is essentially arbitrary. Following Grinblatt and Titman (1989) they focus on the case where the SDF is represented by a mean variance efficient portfolio conditional on the client's information. However, they do not provide explicit answers to the two Fundamental Questions posed above.[3]

Of course, since the definition of alpha here is so natural, a number of studies have used special cases. Cumby and Glen (1990) use an exponential utility example, Ferson, Henry and Kisgen (2006) examine exponential functions from term structure theory and Goetzmann et al. (2007) advocate a power utility example. Farnsworth, et al. (2002) examine several specifications.

The next section sets up the general problem and Section 3 proves the main results. Since the stochastic discount factor used in the model may not be easy to empirically measure, Section 4 discusses operationalizing alpha. I advocate alphas based on investor clienteles. This section also describes some common mistakes in the use of alphas. Section 5 present some further proposals

---

[3] The version of alpha studied here is a "positive admissible" measure. Chen and Knez show that there can be funds that have positive alphas under some positive admissible measures and negative alphas under others. However, if an alpha is positive, there exists some agent with a monotone, concave utility function that would want to buy the fund *at the margin*. Optimal discrete responses are not addressed.

for future empirical work on performance measurement.

2. A General Model for Alpha

Agents make consumption and portfolio choices at each date t, to maximize a lifetime utility function, represented as the indirect value function:

$$J(W_t,\text{info}) \equiv \text{Max}_{\{c,x\}}\ u(C) + E\{\ \beta\ J(W_{t+1},s_{t+1})\ |\text{info}\}, \tag{1}$$

$$\text{s.t. } W_{t+1} = (W_t\text{-}C)\ x'R_{t+1},\ \ x'\underline{1}=1,$$

where W_t is the wealth at time t, C is the consumption expenditure at time t, R_{t+1} is the N-vector of gross (i.e., one plus the rate of) returns for the N assets, one of which can be risk-free, and $\underline{1}$ is an N-vector of ones. The K-vector of state variables in the model is s_t at time t and the conditioning information at time t, "info," takes one of two forms. The info is Z_t, representing public information that includes the current values of state variables s_t and the current risk-free rate if any, when referring to the uninformed client. The info is Ω_t when referring to the better-informed manager, assuming that Z_t is contained in Ω_t. The time subscripts are dropped except when needed to avoid ambiguity.[4]

Assuming that the uninformed agent is at an interior optimum in the N-asset economy, the first order conditions to the problem imply:

$$E(mR|Z) = \underline{1},\ \ \text{with } m = \beta\ J_w(W_{t+1})/u_c(C_t), \tag{2}$$

---

[4] There can be another component of wealth; for example nontraded human capital, and that component can imply hedging demands as in Grinblatt and Titman (1989) without affecting any of the results.

where m is the stochastic discount factor and subscripts denote derivatives. The notation $J_w(W)$ suppresses but allows for the dependence of the value function on the state variables and Z_t.

Consider now presenting the client with a new investment opportunity, the managed portfolio with return $R_p = x(\Omega)'R$ where $x(\Omega)$ is the vector of the informed manager's portfolio weights. We define alpha for any portfolio R_p as:

$$\alpha_p = E(mR_p|Z) - 1. \tag{3}$$

Clearly, if the manager has no superior information in the sense that Z includes Ω, then α_p is zero. Let the managed portfolio return be $R_p = v_{t+1}/P_t$, where P_t is the price that the manager offers the client at time t and v_{t+1} is the random payoff one period later. From the definition of alpha we see that $(1+\alpha_p)P_t = E\{v_{t+1}\ \beta\ J_w(W_{t+1})/u_c(C_t)\}$, so that if alpha is zero the client would find the offer price "fair," relative to the previous equilibrium. A positive alpha intuitively suggests a "low" price, or an attractive investment *at the margin.* It is shown below that this intuition holds when the client's consumption and investments in all assets can change by discrete amounts in response to the introduction of the managed portfolio.[5]

3. Addressing the Fundamental Questions

3.1 Resolving the First Question

How will the client behave when confronted with a new investment opportunity? When faced with a new investment opportunity R_p with a nonzero alpha, the client will generally adjust to new optimal consumption and portfolio choices, until the alpha is zero at the new optimum.[6]

---

[5] It is assumed, as is common in the literature starting with Mayers and Rice (1979), that the manager's trading based on superior information does not affect the market prices of the underlying assets.

[6] This is a normative, not an equilibrium analysis. Berk and Green (2004) make an argument

Consider a situation where we allow the client to adjust current consumption and to buy or sell some amount of the manager's fund. The client feels the effects of these decisions in his future wealth, and thus the marginal utilities of current consumption and future wealth must change. We assume that the client is a price-taker, so there is no effect on the market prices of assets or consumption goods. Let Δ be a reduction in current consumption used to buy the fund, leading to a random wealth increment at time t+1, $W(\Delta) = W_t + \Delta R_p + (W_t\text{-}C)[x(\Delta)\text{-}x]'R$, where $x(\Delta)$ is the new optimal portfolio weight vector for the N base assets, normalized to sum to 1.0, and x is the old optimal weight vector. Let $J(\Delta)$ be the lifetime utility on the left hand side of Equation (1) with the perturbation in place and let J refer to the previously optimal utility before the new investment opportunity was introduced. The first order condition for an optimal response that maximizes the lifetime utility increment implies:

$$- u_c(C_t\text{-}\Delta) + E\{\beta J_w(W(\Delta))\,[R_{pt+1} + (W_t\text{-}C_t)\,(\partial x(\Delta)/\partial\Delta)'R]\,|Z_t\} = 0. \tag{4}$$

Assuming regular utility functions, we can use the mean value theorem to represent

$$\begin{aligned} &u_c(C_t\text{-}\Delta) = u_c(C_t) - u_{cc}^{*}\,\Delta \text{ and} \\ &J_w(W(\Delta)) = J_w(W_t) + J_{ww}^{*}\,[W(\Delta)\text{-}W_t], \end{aligned} \tag{5}$$

where * indicates that the functions are evaluated at points in the intervals $(C_t\text{-}\Delta, C_t)$ and $(W(\Delta), W)$ respectively. The appendix considers this general case, but to see the logic more simply we restrict here to the special case where $[W(\Delta)\text{-}W] = \Delta R_p$, abstracting from the change in the weights on the original N assets.

---

where equilibrium adjustment comes from flows of new cash across funds and diseconomies of scale in fund management, which drive informed manager's alphas to zero in equilibrium.

Substituting from (5) into (4) we have for this special case:

$$u_c(C_t) - u_{cc}^* \Delta = E\{\beta [J_w(W_t) + J_{ww}^* \Delta R_p] R_p |Z_t\}. \tag{6}$$

Now substituting in the definition of α_p,

$$u_c(C_t) - u_{cc}^* \Delta = u_c(C_t)(1+\alpha_p) + E\{\beta J_{ww}^* \Delta R_p^2 |Z_t\}, \tag{7}$$

and solving for the optimal investment in the new fund we have:

$$\Delta = \alpha_p [u_c(C_t)/(-A)], \text{ with } A = u_{cc}^* + E\{\beta J_{ww}^* R_p^2 |Z_t\} < 0. \tag{8}$$

Thus, the sign of the optimal investment in the new fund is the same as the sign of alpha, and the optimal investment is zero only when the alpha is zero. The optimal investment is the alpha, scaled by a term related to "risk tolerance." In the general case we have:

Proposition 1:

Under the assumptions leading to equations (4) and (5), and assuming further that the response of the optimal portfolio weights on the original N assets is small enough to guarantee that $[R_{pt+1} + (W_t - C_t)(\partial x(\Delta)/\partial\Delta)'R][R_{pt+1} + (W_t - C_t)(x(\Delta)-x)/\Delta)'R] > 0$, Then the agent when confronted with a new investment with an alpha equal to α_p, will optimally purchase (sell) some of the investment only when α_p is positive (negative).

Proof: See the Appendix.

The restriction that $[R_{pt+1}+(W_t-C_t)(\partial x(\Delta)/\partial\Delta)'R][R_{pt+1}+(W_t-C_t)(x(\Delta)-x)/\Delta)'R]>0$ says that the derivatives of the optimal portfolio weights on the N base assets are closely approximated by the discrete changes divided by the optimal Δ. There are special cases where the restriction is guaranteed to hold, such as when the relative allocation to the original assets does not change, or when $x(\Delta)$ is well-approximated by a linear function of Δ.[7]

It is important to note that while Proposition 1 relies on the definition of alpha, it does not assume that the alpha is optimally generated from the portfolio weights of an informed manager. This is important in view of the examples cited in the introduction, where using other definitions of alpha it is shown that positive or negative alphas can be obtained when there is no true performance, and negative alphas can be obtained when there is actually positive performance. The definition of alpha used here should preclude most these pathologies. This does not, of course, rule out statistical biases in measuring alpha. For example, return smoothing or nonsynchronous trading can make it difficult to accurately estimate alpha because the returns of the fund are not accurately measured.

If funds can trade within the return measurement period, interim trading can create biases in observable alphas. This issue is raised by Goetzmann, Ingersoll and Ivkovic (2000) and Ferson and Khang (2002), and examined in detail by Goetzmann et al. (2007). However, Ferson, Henry and Kisgen (2006) show that if the right time-aggregated SDF for the return measurement period is used, this problem is avoided. The definition of alpha here involves the right, time-aggregated SDF, and so should not be subject to interim trading bias. The use of a time-additive utility function is important for this property, and it will not generally hold for non time separable utility functions, such as have been used to represent habit persistence.

---

[7] The analysis can accommodate the case where the investor does not change the current consumption, but only the portfolio weights in response to the new investment. In this case the weights $x(\Delta)$ do not sum to 1.0 and $\Delta = W(1-x(\Delta)'\underline{1})$.

3.2 The Second Question

The fact that Proposition 1 does not rely on alpha being generated optimally from an informed manager would seem to obviate the need to resolve the second Fundamental Question. All the client needs to know is that if he sees a positive alpha, he should buy. Nevertheless, the model speaks to the second question, at least in special cases. For example, it follows from the definition of α_p in (3) and the client's first order condition that if R_u is any other portfolio that is feasible to the client, then
$\alpha_p = E[m(R_p - R_u)|Z]$. Since $m>0$ it follows that $\alpha_p>0$ if the manager's return first order stochastically dominates R_u (Chen and Knez, 1996).

Since the uninformed portfolio is feasible for the informed manager, we must have for the same initial wealth, W_0 and consumption C_0, that $E[J(W^I)|\Omega] > E[J(W^u)|\Omega]$, where W^I is the future wealth of the informed manager and W^u is the future wealth of the uninformed client. These are related as $W^I = W^u + [W_0 - C_0][x(\Omega) - x(Z)]'R$. By the mean value theorem, $J(W^I) = J(W^u) + J_w^{\#}\,[W_0 - C_0][x(\Omega) - x(Z)]'R$, where $J_w^{\#} = J_w(aW^I + (1-a)W^u)$ for some $a \varepsilon [0,1]$. Substituting implies $E\{\beta J_w^{\#}\,[R_p - R_u]|\Omega\} = E\{\beta J_w(aW^I + (1-a)W^u)\,[R_p - R_u]|\Omega\} > 0$. If $a=0$ so there is no wealth effect associated with having the superior information Ω, then by the client's first order condition we have $\alpha_p > 0$. But this is not a very realistic case, as informed portfolio managers are often highly compensated for their work. A more interesting special case is:

Proposition II:

Under the assumptions of Proposition I, if the indirect value function J(.) is quadratic in wealth, and informed manager produces a positive alpha.

Proof:

When J(.) is quadratic in wealth, then J_w is a linear function and using the first order conditions again we have $E\{\beta\,[aJ_w(W^I) + (1-a)J_w(W^u)]\,[R_p - R_u]|\Omega\} =$

$E\{\beta(1-a)J_w(W^u)\,[R_p - R_u]|\Omega\} = (1-a)u_c\,\alpha_p > 0$, implying that alpha is positive. QED.

Proposition II generalizes results of Mayers and Rice (1979) and Grinblatt and Titman (1989) to a multiperiod model. A quadratic J(.) function would occur in a continuous-time diffusion setting or under conditional normality of returns given Ω. This does not require that returns appear normal from the client's perspective.

4. Operationalizing Alpha

The main idea in most of the applied measures of investment performance is quite simple. The measures essentially compare the average return of a managed portfolio over some evaluation period to the return of a benchmark portfolio. Formally, alpha in practice is defined as $E\{R_p - R_B|Z\}$, where R_B is the benchmark portfolio for fund p. Alpha should be conditioned on the client's information, Z, sometimes assumed to be no information at all. The benchmark portfolio represents a feasible investment alternative to the managed portfolio, and should be equivalent in all of the return-relevant aspects to the managed portfolio, except that it should not reflect the investment ability of the firm or manager. Aragon and Ferson (2006) call such a portfolio an "*Otherwise Equivalent*" (OE) portfolio. This definition of alpha is equivalent to Equation (3) if the "right" benchmark portfolio is chosen. The right benchmark portfolio requires that $\mathrm{Cov}(m,R_p|Z) = \mathrm{Cov}(m,R_B|Z)$. This follows because a feasible benchmark R_B has a zero alpha. Writing:

$$\alpha_p = E(mR_p|Z)-1 = E(m[R_p-R_B]|Z) = E(m|Z)E(R_p-R_B|Z) + \mathrm{Cov}(m,R_p-R_B|Z), \qquad (9)$$

then α_p is proportional to $E(R_p-R_B|Z)$ when the final covariance term is zero. The problem in practice is to operationalize this idea.

In order to operationalize alpha it is necessary to have a model for m, in order to determine what aspects of a portfolio should lead to higher or lower client expected returns. In general this is client-specific and may involve more than covariances with m, especially when taxes and trading costs are considered, but the literature has sought models in which clients could agree on the relevant covariances. For example, the Capital Asset Pricing Model of Sharpe (CAPM, 1964) implies that all investors should hold a broadly diversified "market portfolio," combined with safe assets or "cash" according to the investor's tastes for risk. It follows that an OE portfolio is one with the same market risk exposure, or "beta" as the fund. The result is Jensen's (1968) alpha. Merton (1973) and Long (1974) develop models where investors should not simply hold a broad market index and cash, but should also invest in "hedge portfolios" for other economically relevant risks, like interest rate changes and commodity price inflation. It follows from such a model that the benchmark portfolio should have the same exposures as the managed portfolio, not just with respect to the overall market, but also with respect to the relevant risk factors or their hedge portfolios. Glosten and Jagannathan (1994) as mentioned above, take this one step further by adding option payoff factors. This can also motivate the use of "style" or characteristics-based benchmarks, on the assumption that controlling for the styles or the characteristics implies controlling for the relevant covariances.

4.2 A Proposal for Improving Alpha

Fundamentally, performance measures always amount to some specification of the benchmark. It is appealing to seek benchmarks about which there can be unanimity, as the performance measure can then be relevant for all investors. However, this is more of a modelling fiction than a practical reality. While some investment managers try hard to divine their clients preferences, the right client-specific stochastic discount factor is likely to be difficult to measure in practice. It may be practical however, to closely approximate the "right" alpha by identifying cohorts of investors based on investor characteristics, who would be expected to have similar m's, and then to construct cohort-specific performance measures. It would be interesting to see how different the measures are for different cohorts. The literature using firm-specific characteristics and country-specific characteristics has developed dramatically over the past two decades, so maybe the time is right for investor-specific characteristics.

4.3 Common Mistakes with Alphas

When you think about performance measures in terms of the right benchmarks some new insights emerge and a few common mistakes become evident. Two examples of common mistakes are made using the most popular classical models of market timing ability. Formal models of market timing ability were first developed in the 1980s, following the intuitive regression model of Treynor and Mazuy (1966). In the simplest example, a market timer has the ability to change the market exposure of the portfolio in anticipation of moves in the stock market. When the market is going up, the timer takes on more market exposure and generates exaggerated returns. When the market is going down, the timer moves into safe assets and minimizes losses.

Merton and Henriksson (1981) model market-timing behavior as like put option on the market. A successful market timer can be seen as producing "cheap" put options. The Merton-Henriksson market timing regression is:

$$r_{pt+1} = a_p + b_p\, r_{mt+1} + \Lambda_p\, \mathrm{Max}(r_{mt+1},0) + u_{t+1}. \tag{10}$$

The coefficient Λ_p measures the market timing ability. The intercept has been naively interpreted in may studies as a measure of "timing-adjusted" selectivity performance. This only makes sense if the manager has with "perfect" market timing, defined as the ability to obtain the option-like payoff at zero cost. But in reality no one has perfect timing ability, and the interpretation of a_p as timing adjusted selectivity breaks down. For example, a manager with some timing ability who picks bad stocks may be hard to distinguish from a manager with no ability who buys options at the market price. Indeed, without an estimate of the market price of a put option on the market index, the intercept a_p has no clean interpretation.

In the model of Merton and Henriksson, the right benchmark portfolio is a combination of the market index, the risk-free asset and options on the market index. The benchmark portfolio has a weight equal to b_p in the market index returning R_m, a weight of $\Lambda_p P_0$ in an option with beginning-of-period price P_0 and return $\{Max(R_m - R_f, 0)/P_0 - 1\}$ at the end of the period, and a weight of $(1-b_p-\Lambda_p P_0)$ in the safe asset returning R_f. The option is a one-period European call written on the relative value of the market index, $V_m/V_0 = 1+R_m$, with strike price equal to the end of period value of the safe asset, $1+R_f$.

Given a measure of the option price P_0 it is possible to estimate returns in excess of the benchmark portfolio. In practice, the price of the option must be estimated from an option pricing model. For equity options the Black Scholes (1972) option pricing model is a simple choice. Let r_0 be the return on the option measured in excess of the safe asset. The difference between the excess return of the fund and that of the benchmark portfolio may be computed as:

$$\alpha_p = E(r_p) - b_p\, E(r_m) - \Lambda_p\, P_0\, E(r_0). \tag{11}$$

The measure α_p captures "total" performance in the following sense. If an investor holds the benchmark portfolio he obtains the same market beta and nonlinear payoff with respect to the market as the fund. The difference between the fund's expected return and that of the benchmark portfolio reflects the manager's ability to deliver the same beta and nonlinearity at a below-market cost, and thus with a higher return. The essence of successful market timing is the ability to produce the convexity at below-market cost.

Note that alpha in this example is not the same as the intercept in regression (10). The problem is that the term $Max(r_{mt+1}, 0)$ in the regression is not an excess return. Taking the expected value of (10) and comparing it with the expression for alpha in (11), the intercept in (10) is related to the "right" alpha in this model as:

$$\alpha_p = a_p + \Lambda_p P_0 R_f. \tag{12}$$

Only if the fund has perfect timing ability does the intercept in the regression (10) measure timing-adjusted selectivity. If a manager had perfect timing ability she would deliver the same payoff as the benchmark portfolio, while "saving" the cost of the option, $\Lambda_p P_0$. Increasing the position in the safe asset by this amount leaves the beta unchanged and produces the additional return, $\Lambda_p P_0 R_f$. The additional return is the difference between the intercept, a_p, and the alpha in (11). If a manager had perfect timing ability and could generate a higher return in excess of the benchmark portfolio than $\Lambda_p P_0 R_f$, the extra return could then be presumed to be attributed to selectivity. Under this interpretation, when $\alpha_p > \Lambda_p P_0 R_f$, then $a_p > 0$ measures the selectivity-related excess return, on the assumption of perfect market timing ability. Since in general the return to timing activity will be less than $\Lambda_p P_0 R_f$, then for a given total performance the intercept in (10) is less than the selectivity performance. The literature typically finds that funds with positive timing coefficients have negative intercepts, consistent with understated selectivity performance.

Treynor-Mazuy model

The Treynor-Mazuy (1966) market-timing model is a quadratic regression:

$$r_{pt+1} = a_p + b_p\, r_{mt+1} + \Lambda_p\, r_{mt+1}^2 + v_{t+1}. \tag{13}$$

Treynor and Mazuy (1966) argue that $\Lambda_p > 0$ indicates market-timing ability. Like the intercept of

Equation (10), the intercept in the Treynor-Mazuy model has been naively interpreted as a "timing adjusted" selectivity measure. However, as in the Merton-Henriksson model, the intercept does not capture the return in excess of an benchmark portfolio because r_m^2, in this case, is not a portfolio return.[8] However, the model can be modified to capture the difference between the return of the fund and that of an appropriate benchmark portfolio.

Let r_h be the excess return of the maximum correlation portfolio to the random variable r_m^2 and let Λ_h be the portfolio's regression coefficient on r_m^2. The benchmark portfolio that replicates the beta and convexity of r_p has a weight of Λ_p/Λ_h in r_h and b_p in r_m, with $1 - b_p - \Lambda_p/\Lambda_h$ in the safe asset, assuming $\Lambda_h \neq 0$.[9] The fact that the benchmark portfolio has the same beta and convexity coefficient as r_p can be seen by substituting the regression for r_h on r_m^2 into the combination of r_m and r_h that defines the benchmark excess return. The means of r_p and the benchmark portfolio excess returns differ by $\alpha_p = a_p - a_h \Lambda_p/\Lambda_h$, where a_p is the intercept of (13). Thus, α_p measures the total return performance, in the presence of timing ability, on the assumption that timing ability may be captured by a quadratic function. A modified version of the model is the system:

$$r_p = [\alpha_p + a_h \Lambda_p/\Lambda_h] + b_p r_m + \Lambda_p r_m^2 + \varepsilon_p, \qquad (14)$$

$$r_h = a_h + \Lambda_h r_m^2 + \varepsilon_h.$$

---

[8] Note that the intercept in (13) can be interpreted as the difference between the fund's average return and that of a trading strategy that holds the market index and the safe asset, with a time-varying weight or beta in the market index equal to $b_p + \Lambda_p r_{mt+1}$. However, this weight is not feasible at time t without foreknowledge of the future market return, so this strategy is not a feasible OE portfolio.

[9] As Λ_h approaches zero the weight of the OE portfolio in r_h becomes infinite. If $\Lambda_h = 0$, no portfolio can be formed with a nonzero correlation with r_m^2. In this unlikely event the model of Equation (14) is undefined.

This system is easily estimated using the Generalized Method of Moments (Hansen, 1982). No study has yet examined these performance measures empirically, so it would be interesting in future research to implement the measures in (11) and (14) using data on managed portfolios where market timing is likely to be present.

5. More Proposals for Future Research

5.1 Weight-based Performance Measures

Substituting $x(\Omega)'R$ for R_p and expanding the expectation of the product in Equation (3) into the product of the expectations plus the covariance, the alpha may be written as:

$$\alpha_p = \mathrm{Cov}(mR'; x(\Omega)|Z), \tag{15}$$

where the notation indicates the sum of covariances across the assets. Alpha is the sum of the covariances of the manager's weights with the future "abnormal" returns of the assets, mR. This is not what the literature on weight-based performance measures looks at. Instead, the weight-based performance literature looks at things like $\mathrm{Cov}(R - R_B\underline{1}; x(\Omega))$ or $\mathrm{Cov}(R; x(\Omega)-x_B)$, where R_B is a benchmark return and x_B is a benchmark weight, such as a lagged portfolio weight (e.g. Grinblatt and Titman, 1993). The average of the conditional version of this covariance has also been used (e.g. Ferson and Khang, 2002). Often the benchmark is somewhat ad hoc. It is likely to be "anomaly adjusted," matching stock characteristics like firm size, book/market, momentum etc. (e.g. Daniel, Grinblatt, Titman and Wermers, 1997). The idea is to ensure that a fund that merely exploits well-known anomalies will not record abnormal performance. Nevertheless, it would be interesting to see results from using Equation (4) directly. The results are likely to differ even if the SDF is assumed to be linear in factors -- which leads to beta pricing models (e.g., Ferson, 1995) -- because then the covariances of portfolio weights with the products of returns and factors appear in Equation (4).

5.2 Costs

A manager may be able to generate higher returns than an otherwise equivalent benchmark before costs and fees, yet after costs investors' returns may be below the benchmark. If a fund can beat the otherwise equivalent benchmark on an after cost basis, Aragon and Ferson (2006) say that the fund *adds value* for investors, to distinguish this situation from one where the manager has investment ability, but either extracts the rents to this ability in the form of fees and expenses, or dissipates it through trading costs. It would be useful to have clean measures of investment performance, both on a before-cost and an after-cost basis.

Mutual funds charge expenses that get deducted from the net assets of the funds and sometimes transactions fees paid into the assets of the fund by new investors to compensate existing shareholders for the costs of buying and selling the underlying assets. Funds sometimes also charge additional "load fees," such as those paid to selling brokers -- measured fund returns do not account for these additional charges. Funds' trading costs represent a drain from the net assets of the fund, and are reflected in the measured returns. These trading costs can be substantial, both for managed funds and even for common "passive" benchmarks. Table 1 shows that passive benchmark indexes typically experience annual turnovers larger than 25% and often more than 40%, as they are reconstituted annually to reflect the addition of new firms and the deletion of old ones.

It is important to recognize the role of costs in any comparison of a managed portfolio return with a performance benchmark, but most performance measures are crude in their treatment of investment costs and fees. In most academic studies the benchmark strategy does not reflect its costs. For example, the CRSP indexes pay no costs when their composition changes. Mutual fund returns, in contrast, are measured net of all the expenses summarized in the funds' expense ratio and also the trading costs incurred by the fund. The typical performance measure is therefore somewhere between a before-cost measure of investment ability and an after-cost measure of value added.[10] I think it makes sense to modify current measures to reflect the costs of trading the benchmarks. Then, we would have a better sense of the performance after costs.

Weight based performance measures are cleaner before cost measures than the current returns-based measures, as they amount to the average difference between a hypothetical portfolio and a benchmark that pays no costs. The hypothetical portfolio is constructed from a snapshot of the fund's weights and abstracts from trading costs. It

It seems that a logical next step for research on fund performance measures is to more carefully take into account the full range of costs associated with investing in funds, and to develop better measures of after-cost performance.

Measuring the managed portfolio's returns and the performance benchmark returns on a cost-equivalent basis can get complicated. But, there are asset pricing models that make predictions about the return-relevant measures of transactions costs and taxes, going back at least to Brennan (1970). The problem is that the incidence of many costs is likely to be different for different investors. For example, a pension plan pays no income taxes on the dividends or capital gains generated by a portfolio, so the manager and the plan client may care little about the form in which the gains are earned. An individual investor may be taxed more favorably on capital gains than on dividends, and the relative tax cost may depend on the investor's income profile. This implies that for a given fund, the benchmark portfolio for one investor may not be the same as the benchmark portfolio for another investor, and different investors may view the performance

---

[10] This common, "in between" measure is perhaps best interpreted as a partial value added for a "no load" fund investor whose tax payments for returns on the fund and the benchmark would be the same, when the costs of trading the benchmark portfolio are trivial. Roughly speaking, a manager whose performance just matches the benchmark then has enough ability to cover his or her trading costs and fees.

of the same fund in different ways. This suggests that future performance measures that reflect costs can, and I think should, be constructed for a range of hypothetical investors or clienteles. This goes back to and reinforces the earlier call for clientele-based alphas.

6. Conclusions

The main theoretical result and the practical suggestions in this address both point in the same direction for future research on performance measurement. The theoretical ambiguities in the interpretation of alpha found in the early literature are largely resolved when alpha is defined relative to the client's preferences. Properly accounting for costs and taxes in performance comparisons is also client specific. In evaluating managed portfolios, one size does not fit all. A challenge for future research, therefore, is to identify and characterize meaningful investor clienteles and to develop performance measures specific to the clienteles. Recently, behavioral finance studies have found access to refined data about individual investors, and I would like to see these data used to address this challenge.

Appendix

Proof of Proposition 1:

In the general case $W(\Delta)-W = \Delta R_p + (W_t-C_t)[x(\Delta)-x]'R$, where $x(\Delta)$ is the new optimal portfolio weight vector for the N base assets, normalized to sum to 1.0, and x is the old optimal weight vector in Equations (5) and (6). Substituting (6) into (5) yields:

$$u_c(C_t) - u_{cc}^{*}\Delta = E\{\beta\, [J_w(W_t) + J_{ww}^{*}\, (W(\Delta)-W)]\, [R_p + (W_t-C_t)(\partial x(\Delta)/\partial\Delta)'R]\}. \qquad \text{(A.1)}$$

Substituting in for $(W(\Delta)-W)$, using the first order condition $E\{\beta\, J_w(W_t)\, R\}=\underline{1}$ and the fact that $x(\Delta)'\underline{1}=1$ implies $(\partial x(\Delta)/\partial\Delta)'\underline{1}=0$, and using the definition of α_p, (A.1) reduces to:

$$u_c(C_t) - u_{cc}^{*}\Delta = (1+\alpha_p)u_c(C_t) + \Delta\, Q, \qquad \text{(A.2)}$$

$$Q = E\{\beta\, J_{ww}^{*}\, [R_{pt+1} + (W_t-C_t)\, (\partial x(\Delta)/\partial\Delta)'R]\, [R_{pt+1} + (W_t-C_t)\, (x(\Delta)-x)/\Delta)'R].$$

Solving for the optimal Δ we have:

$$\Delta = \alpha_p\, \{u_c /(-u_{cc}^{*} - Q)\}, \qquad \text{(A.3)}$$

and the conditions of the theorem guarantee that $Q<0$, which establishes the result. QED.

References

Admati, A. and S.A. Ross, 1985, Measuring performance in a rational expectations equilibrium model, *Journal of Business* 58(11), 11-26.

Aragon, George O., and Wayne E. Ferson, 2006, Portfolio Performance Evaluation, *Foundations and Trends in Finance* 2:2, 1-111.

Beja, Avraham, 1971, The structure of the cost of capital under uncertainty, *Review of Economic Studies* 4, 359-369.

Berk, Jonathan and R.C. Green, 2004, Mutual Fund flows in Rational Markets, *Journal of Political Economy* 112, 1269-1295.

Black, F. and M. Scholes, 1972, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-654.

Carino, J. Christopherson and W. Ferson, 2009, Performance Measurement and Benchmarking, McGraw Hill Publishers ISBN 978-0-070149665-0..

Chen, Z., and P.J.Knez. 1996. "Portfolio performance Measurement: theory and applications." *Review of Financial Studies*, 9: 511-556.

Coles, A., 1933, Can stock market forecasters forecast? *Econometrica* 1(3), 309-325.

Cornell, B., 1979. Asymmetric information and portfolio performance measurement. *Journal of Financial Economics* 7, 381-390.

Cumby, Robert, and Jack Glen. 1990. "Evaluating the performance of international mutual funds" *Journal of Finance*, 45: 497-521.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52, 1035-1058.

Dybvig, P.H. and J. Ingersoll, 1982, Mean variance theory in complete markets, *Journal of Business* 55(2), 233-251.

Dybvig, Philip H., and Ross, Stephen A. 1985a. Performance Measurement Using Differential Information and a Security Market Line. *Journal of Finance* 40: 383-99.

Dybvig, Philip H., and Ross, Stephen A. 1985b. The Analytics of Performance Measurement Using and a Security Market Line. *Journal of Finance* 40: 401-416.

Farnsworth, Heber, W. Ferson, D. Jackson and S. Todd, 2002, Performance Evaluation with Stochastic Discount Factors, *Journal of Business* 75, 473-504 (July).

Ferson, 1995, Theory and Empirical Testing of Asset Pricing Models, Chapter 5 in *Finance, Handbooks in Operations Research and Management Science*, by Jarrow, Maksimovic and Ziemba (editors), Elsevier, 145-200.

Ferson, Wayne and K. Khang, 2002, Conditional Performance Measurement Using Portfolio Weights: Evidence for Pension Funds, *Journal of Financial Economics* 65, 249-282 (August).

Ferson, W., Tyler Henry and Darren Kisgen, 2006, Evaluating Government Bond Fund Performance with Stochastic Discount Factors, *Review of Financial Studies* 19, 423-456.

Ferson, W., and Rudi Schadt. 1996. "Measuring Fund Strategy and Performance in Changing Economic Conditions." *Journal of Finance*, 51: 425-462.

Ferson, Wayne, A. Siegel and P. Xu, 2006, Mimicking Portfolios with Conditioning Information, *Journal of Financial and Quantitative Analysis* 41, 607-636.

Gibbons Ross and Shanken (1989)

Glosten, L. and R. Jagannathan, 1994, A contingent claim approach to portfolio evaluation, *Journal of Empirical Finance* 1, 133-160.

Goetzmann, W., Ingersoll J., Ivkovic Z., 2000, Monthly measurement of daily timers, *Journal of Financial and Quantitative Analysis* 35, 257-290.

Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I, 2007, Portfolio performance manipulation and manipulation-proof measures, *Review of Financial Studies* (forthcoming).

Green, R.C., 1986, Benchmark portfolio inefficiency and deviations from the security market line, *Journal of Finance* 41 (2), 295-312.

Grinblatt, Mark and Titman, Sheridan. 1989. Portfolio Performance evaluation: old issues and new insights. *Review of Financial Studies* 2: 393-422.

Grinblatt, M., Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. *Journal of Business* 60, 97-112.

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029-1084.

Hansen, Lars P. and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587-613.

Jagannathan, R., Korajczyk, R. 1986, "Assessing the market timing performance of managed portfolios." *Journal of Business*, 59: 217-236.

Jensen, M. 1968. "The performance of mutual funds in the period 1945 -1964." *Journal of Finance*, 23: 389-46.

Jensen, M., 1972, Risk, the pricing of capital assets and the evaluation of investment portfolios, *Journal of Business* 42(2), 167-247.

Leland, Hayne. 1999. Performance beyond mean-variance: Performance measurement in a nonsymmetric world. *Financial Analysts Journal* 55: 27-36.

Long, John B., 1974, Stock prices, inflation and the term structure of interest rates, *Journal of Financial Economics* 1, 131-170.

Mayers, D. and E. Rice, 1979, Measuring portfolio performance and the empirical content of asset pricing models, *Journal of Financial Economics* 7(1), 3-28.

Merton, Robert C., 1973, An intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867-87.

Roll, Richard, 1978, Ambiguity when performance is measured by the security market line. *Journal of Finance* 33: 1051-1069.

Roll, Richard, 1979, A reply to Mayers and Rice, *Journal of Financial Economics* 7, 391-400.

Rubinstein, Mark, 1976, The valuation of uncertain income streams and the pricing of options, *Bell Journal of Economics and Management Science* 7, 407-425.

Sharpe, William. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-42.

Verrechia, R., 1980, The Mayers-Rice conjecture: A counterexample, *Journal of Financial Economics* 8, 87-100.

Table 1

Percentage turnover rates at end-of-June index reconstitutions, averaged over the 1994-2004 period. LG refers to the large growth style, LV to large value, SG to small growth and SV to small value. Source: Carino, Christopherson and Ferson (2009).

Index	LG	LV	SG	SV
Russell 1000/2000 Style	15.6	18.1	41.2	33.5
MSCI	20.7	20.0	41.5	28.2
Russell 2500 style	--	--	34.4	28.1