

AN ARBITRAGE RATIONALE FOR TESTS OF MUTUAL FUND PERFORMANCE

K.V. PEASNELL, L.C.L. SKERRATT AND P.A. TAYLOR*

1. INTRODUCTION

This paper examines the methods currently employed to assess investment performance in the light of recent developments in the theory of capital asset pricing. There have been a considerable number of studies in the last decade on whether or not mutual funds¹ are able to achieve superior investment performance ("beat the market" in some sense). The verdict is virtually unanimous: mutual funds do not consistently outperform the market. This evidence is having a traumatic impact on the US securities industry, resulting in proposals for far-reaching changes in the management and organisation of mutual funds. The validity of these research findings is, of course, dependent on the appropriateness of the methodology underlying the empirical tests. Recent work by Roll (20), Ross (21) and Fama (7) suggests that the methodology is suspect and can yield highly misleading results. Our purpose is to assess this claim and to see whether anything can be rescued. We argue that the damage is not as serious as first appearances suggest.

The paper is organised as follows. The next section outlines (i) the state of the art of empirical research into fund performance, with particular reference to the methodology employed by Jensen (12) in this pathbreaking study and (ii) the main features of Roll's critique of Jensen's methodology. Section III provides a different interpretation of the Jensen study, based on the theory of arbitrage, which goes some way to meeting Roll's criticisms. An arbitrage-based test procedure has the additional advantage of being able to be extended to deal with industry effects on security returns. Section IV considers a hitherto little recognised problem concerning the mechanics of securing excess investment returns which has implications for the measurement of beta. The section also reports some new empirical evidence, obtained by reworking some of Jensen's findings in order to facilitate the interpretation of his study in terms of arbitrage theory. Section V summarises the paper's contribution and suggests promising avenues for future research.

**K. V. Peasnell is Wolfson Professor of Accounting and Finance in the Department of Accounting and Finance at the University of Lancaster. L.C.L. Skerratt is Spicer and Pegler Professor of Accounting in the Department of Economics at the University of Durham. P.A. Taylor is a lecturer in the Department of Accounting and Finance at the University of Lancaster. An earlier version of this paper was presented to the Northern Accounting Group at the University of Sheffield, 27 September 1978. The authors wish to thank Dr. W.H.K. Fung for penetrating and constructive criticisms of a previous draft. (Paper received December 1978, revised April 1979).*

II. THE STATE OF THE ART

From the investment community's viewpoint, the primary purpose of measuring mutual fund performance is to assess the quality of the service provided to investors : the funds charge investors with commissions to finance administration and research expenses, and the obvious question to ask is, "Do investors get value for money?" Additionally, from society's viewpoint there is the need to know whether or not mutual funds are furthering the efficient allocation of resources.

The aspect of mutual fund performance which has commanded most attention is whether or not fund managers have access to superior information (from either inside knowledge or superior analysis) to other investors – testing the so-called "strong-form" of the Efficient Markets Hypothesis, see Fama (6).

The approach used in the studies is as follows. Fund performance has to be judged in terms of a standard or benchmark. The standard developed by Jensen (12), outlined below, is based on the Sharpe-Lintner Capital Asset Pricing Model (CAPM) of market equilibrium. The CAPM provides a measure of the equilibrium expected returns on portfolio j in period t (where the period is very small), and this is the norm against which performance is judged. Important assumptions underlying this approach include :

- (a) For the CAPM to provide a meaningful benchmark it is necessary that
 - (i) the securities markets are dominated by investors who are averse to risk and that
 - (ii) the coefficient β is a valid measure of risk. In order to test (ii), (i) is assumed to be true; see Jensen (12, p.220).
- (b) The process generating security returns is stationary and all investors are assumed to agree upon it; hence "superior information" is limited here to private knowledge about particular realisations of this stochastic process (12, pp. 187, 212).

In order to employ in practice the Jensen benchmark it is necessary to substitute realisations for the market's beliefs. It is here that point (b) is of crucial importance : the generating process underlying returns is assumed to be stable over time, such that realisations can be treated as sample observations of the ex ante joint distribution of returns.

Jensen's study has served as the model for most subsequent work. It concludes that funds' average returns were below what would have been obtained from investment in naive portfolios with the same systematic risks. Studies by Friend, Blume, Crockett (10), Williamson (24) and McDonald (15) in the US more or less confirm Jensen's results; these results have been replicated in a recent UK study by Firth (9).²

The importance of the conclusion of the studies of mutual fund performance is now widely recognised in the US securities industry. One consequence of the research has been the proposal of the "index fund" which is designed "to do no more than match the performance of the widely based indices of stock market performance by pursuing a passive buy/hold investment policy and minimising expenses" (11, p.5); index fund managers are not expected to pick "winners". This proposal has been picked up in the UK by Investment Research of Cambridge who are now providing an index-matching service (11, p.17).

A recent paper by Roll (20) casts strong doubt on the generally accepted economic interpretation of the mutual fund studies by Jensen and others which form the basis of these far-reaching proposals. Roll (p.132) claims:

"If the 'market' proxy used in the calculation [of mutual fund performance] is exactly (not significantly different from) ex-post efficient, *all* of the individual Jensen performance measures gross of expenses will be identically (not significantly differently from) zero. They can be (significantly) non-zero only if the proxy market portfolio is (significantly) not efficient. But if the proxy market portfolio is not efficient, what is the justification for using it as a benchmark in performance evaluation?"

This claim by Roll follows directly from the main thesis of his three-part article (for the complete paper, see 19³) which is to demonstrate that the testing of the CAPM is difficult and currently infeasible: "The two-parameter asset pricing theory is testable *in principle*; but arguments are given here that : (a) No correct and unambiguous test of the theory has appeared in the literature, and (b) there is practically no possibility that such a test can be accomplished in the future" (20, pp. 129–130).

Prior to Roll, the CAPM was accepted to be a positive scientific theory of share pricing which fitted the facts extraordinarily well. To be sure, empirical research shows that the Sharpe Lintner model *in its most elementary form* does not provide an adequate description of the structure of security returns (13, p.391). However, theoretical work by Black (1) showed that the Sharpe-Lintner model is but a special case of the Zero Beta formulation of the CAPM which is thought by many to account for the empirical discrepancy. This Zero Beta CAPM seemed to provide a positive theory of equilibrium security pricing which would yield empirical estimates which could serve as benchmarks for assessing fund performance. The benchmark is interpreted as the minimum average return required by investors for a specified level of risk. Roll, though, claims to have destroyed the case for using the CAPM as such a benchmark.

Roll's principal result (Corollary 6 to Theorem 1) is that the mean return on *any* security i , \bar{r}_i , can be expressed as an *exact* linear function of the mean return on *any* portfolio p on the efficient frontier, \bar{r}_p , viz:

$$\bar{r}_i \equiv a_{0i} + a_{1i} \bar{r}_p \dots \dots \dots (1)$$

(N.B. : A portfolio on the efficient frontier must have the following property : no portfolio with the same expected one period return has a lower variance of return). Furthermore, he proves that for every efficient portfolio p there exists (in any non-trivial situation) *a unique* portfolio z ; it is unique in the sense that (a) its return is uncorrelated with the return on p and (b) of the set of such zero-correlation portfolios, portfolio z has the smallest variance. Using this relationship Roll proves that

$$a_{0i} \equiv \bar{r}_z (1 - \beta_i) \dots \dots \dots (2)$$

and that

$$a_{1i} \equiv \beta_i, \dots \dots \dots (3)$$

where $\beta_i \equiv \text{cov}(r_i, r_p) / \text{var}(r_p)$ is the regression slope coefficient of returns of security i on those of the efficient portfolio p , and \bar{r}_z is the mean return on portfolio z .

It is crucial to note that equations (1), (2) and (3), taken together, imply that:

- (i) the linear relationship posited by Black's Zero Beta CAPM holds *exactly* for *every* efficient portfolio p ;
- (ii) this linear relationship holds *if and only if* portfolio p is efficient or, to put it differently, the relationship is a *consequence* of p 's efficiency.

From this it must necessarily follow that the Zero Beta CAPM can hold *if and only if* the market portfolio is efficient. In operational terms, a Zero Beta CAPM will be found to hold exactly⁴ (approximately) if and only if the proxy for the return on the market portfolio is exactly (nearly) efficient. This is so even if the value-weighted market portfolio is not even remotely efficient.

Results (i) and (ii) have been proved independently by Fama (7). The essential point is that it is now known the CAPM can be deduced from a *single* assumption, viz. that the market portfolio or proxy thereof is efficient. Ross (21) also should be credited with the independent discovery of this important result. Roll's contribution is in drawing out the full implications of this result.

The implication of the above analysis for Jensen's methodology is potentially devastating. If the value-weighted market portfolio is efficient *ex post*, and if $\bar{r}_z = r_f$,⁵ then Jensen's fund performance measure α_j can only be non-zero when the market proxy is not efficient. Roll (20, p. 132) queries: "But if the proxy market portfolio is not efficient, what is the justification for using it as a benchmark in performance evaluation?" However it is possible to obtain a non-zero α_j *even* when the market proxy is efficient. Roll (19, p. 87) shows that when $\bar{r}_z \neq r_f$, then the effect on α_j of Jensen's assuming the contrary will be :

$$\alpha_j = (\bar{r}_z - r_f) (1 - \beta_j) \dots \dots \dots (4)$$

Accepting this line of reasoning, and if, following Black (1), it is assumed that $\bar{r}_z > r_f$, then

- $\alpha_j > 0$ if and only if $\beta_j < 1$ (i.e. fund *j* is of below-average riskiness),
- $= 0$ if and only if $\beta_j = 1$ (*j* is invested in shares of average riskiness),
- < 0 if and only if $\beta_j > 1$ (*j* is of above-average riskiness).

Thus observed fund performance is dependent on the systematic risk of the fund's portfolio – which is completely contrary to Jensen's intentions. α_j is an exact negative linear function of β_j . Conversely, if $\bar{r}_z < r_f$ then the above inequalities would be reversed; α_j would be an exact positive linear function of β_j , which is equally undesirable. It is interesting, in this context, to note that Jensen found that fund performance is negatively correlated with beta – which he puts down to much of the variability in estimates of performance being "due to random factors or sampling error in the estimates of β_j " (12, p.225), whereas the above analysis suggests a quite different and much more worrying possibility.

Whether or not this criticism is a fatal one is, in the final analysis, an empirical issue. As already noted, what evidence there is suggests that the difference $\bar{r}_z - r_f$ is not insignificant.⁶ Note though that when β_j is close to unity, then the consequential error is trivial; see equation (4). In the UK, fund betas rarely substantially depart from unity; see Firth (9).⁷ In the US, on the other hand, there are more outliers (12, Table 5; 16, Table 3). Consequently, US findings might be quite suspect in this respect.

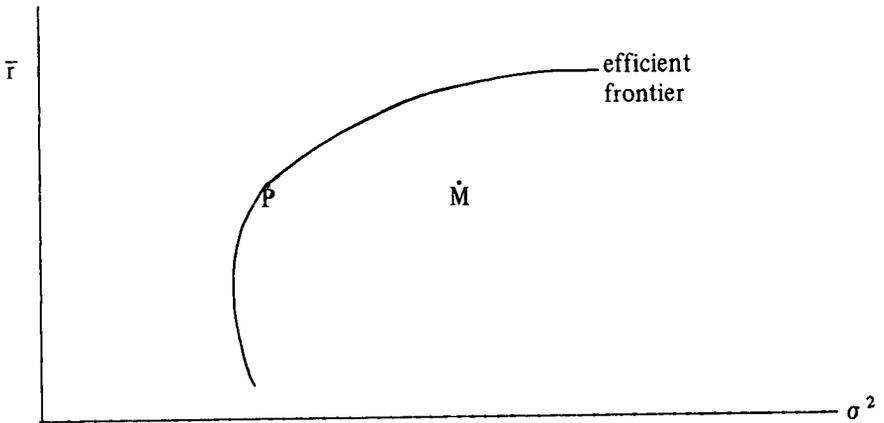
Although this problem is clearly of some importance, the difficulties encountered by relaxing the assumption that the proxy for the market portfolio is efficient are more fundamental. After all, relaxing the assumption about the location of the *z* portfolio merely allows α_j to be non-zero (admittedly in a misleading fashion); relaxing the assumption of the market's efficiency kicks away the very foundations of the Jensen methodology, as can be seen as follows.

The CAPM is an equilibrium theory and only makes sense if it is assumed that the market is efficient. The situation can be summarised in terms of the following paradox. If the market is efficient then the only way in which a fund can appear to earn excess (risk adjusted) returns is by comparing its performance with a standard based on an inefficient proxy for the market. What is the justification though for such a benchmark? On the other hand, if the market is not efficient then the CAPM does not hold and the test procedure is apparently meaningless.

In order to see this more clearly consider Diagram I. For simplicity, assume $\bar{r}_z = r_f$ and that r_f is less than the mean return on the global minimum variance portfolio. Let portfolio M be the inefficient market portfolio (or its inefficient proxy) used for evaluating fund performance, and let P be an efficient portfolio with $\bar{r}_z = r_f$ such that $\bar{r}_P = \bar{r}_M$. It follows that $\sigma_M^2 > \sigma_P^2$ by the

DIAGRAM I

PORTFOLIOS IN MEAN VARIANCE SPACE



definition of efficiency. As P is efficient \bar{r}_j can be expressed as an exact linear function of β_{jP} (the estimated beta of fund j with P),

$$\bar{r}_j \equiv r_f + (\bar{r}_P - r_f) \beta_{jP} \dots \dots \dots (5)$$

Identity (5) can be rewritten as⁸

$$\bar{r}_j \equiv r_f + (\bar{r}_M - r_f) \beta_{jM} + (\bar{r}_M - r_f) (\beta_{jP} - \beta_{jM}) \dots \dots \dots (6)$$

Note that if $\sigma_P^2 = \sigma_M^2$, then M must be the efficient portfolio P; in which case j must have the same beta with both and the last term in (6) vanishes, (6) reducing to (5). On the other hand, when M is inefficient (as assumed) then the third term is highly unlikely to vanish.⁹

Using (6) yields excess performance of

$$\alpha_j \equiv (\bar{r}_M - r_f)(\beta_{jP} - \beta_{jM}). \quad \dots \dots \dots (7)$$

Notice that (7) is the general case of Roll's special result that when M is the efficient portfolio P then α_j is exactly zero.¹⁰ Unfortunately, nothing really useful can be inferred from the converse situation when $\alpha_j \neq 0$; the sign and magnitude of $(\beta_{jP} - \beta_{jM})$ does not tell us anything about the extent of M's inefficiency: betas of j with different "market" portfolios (one of which is efficient) can be of considerably different magnitudes (cf. footnote 9). Put differently, it is impossible to predict the sign or magnitude of $(\beta_{jP} - \beta_{jM})$ without detailed knowledge of the composition of M and P, even if we are told that M is highly or minimally inefficient.

It would seem that if the market proxy used to calculate betas is inefficient then the fund performance measure α_j will be nonzero. Unfortunately nothing can be inferred about the fund's performance from the sign or magnitude of the fund performance measure. What Roll has established is that the theoretical form of the CAPM must hold exactly if the market portfolio is efficient, thereby ruling out the possibility of abnormal gains and losses. Conversely, identity (6) demonstrates that a regression of the returns on a security (or fund) on the returns on a market portfolio will depart from the theoretical model if the market portfolio is inefficient, resulting in possibly nonzero performance measures of the form of (7) which are functions of beta. Roll shows that when the market portfolio (or proxy thereof) is inefficient then such regression lines will not be linear functions of beta.

Does this mean that reliance on Jensen's CAPM-based benchmark when evaluating fund performance is misguided, as Roll implies? The answer would seem to be that it is. But it has to be acknowledged that it may be just as misleading to attempt to evaluate fund performance without first specifying a model of some kind of the security pricing process, as some studies do. For example, Cranshaw (4) makes a straight comparison of the performance of UK unit trusts with that of the market as a whole for the period 1963–1969, without adjusting for risk differences;¹¹ he concludes that the trusts did no better than the market index. He remarks (4, p. 467):

“It is a conclusion difficult to reconcile with a belief in differences in management, because it would imply that for each good manager there was one actively bad manager, that is, one who consistently managed his portfolio worse than an unmanaged portfolio. This is a curious skill to be possessed by a large number of our investment managers”.

To see that this is a somewhat premature judgement, let us for argument's sake assume that we know that investors in general behave according to the dictates of the CAPM, i.e. that most investors' decisions are based on the two parameters of mean and variance of portfolio returns such that their portfolios are chosen in order to promise the combination of mean returns and beta optimal to the individual. Let us further assume that realisations are generally in accordance with expectations, i.e. that ex post and ex ante return distributions are identical. Now, Firth (9, pp. 170–176) reports in his appendix estimated betas for over 300 UK trusts, *not one* of which was greater than unity.¹² Given our above assumptions about the applicability of the CAPM,¹³ it would seem clear that the equilibrium mean returns for *all* these trusts should be less than the mean return on the market. In short, Cranshaw's test does not allow for risk differences; according to theory, the funds *should* report lower average returns than the market.

The dilemma is obvious. If it is desired to measure fund performance such that all factors influencing performance other than skill and availability of information are held constant, then a model of the investment process is needed. Without such a model, we cannot say how much of Cranshaw's results are due to attempts to minimise or maximise risk. (CAPM's apparent capacity to handle this was its basic attraction to empirical researchers in the first place.) We thus seem to be back at square one. The CAPM – or something like it – is needed to control for confounding factors. But, as Roll points out, if the CAPM fits ex post then superior or inferior fund performance is impossible, and if it does not fit then the tests are without support from economic theory.

III. AN ARBITRAGE APPROACH TO MEASURING INVESTMENT PERFORMANCE

The only way of escaping this dilemma is to abandon the CAPM along with its baggage of mean-variance theory and its attendant efficient frontiers. Whatever alternative approach is adopted, a benchmark must be found which (a) is computable from observable variables and (b) can be interpreted in the context of a recognisable investment strategy. CAPM certainly satisfies both of these requirements; see Ohlson (17). One attraction of using CAPM as the benchmark is that it offers a method of computing risk adjusted required returns such that the (positive or negative) excess returns can be interpreted as pure economic quasi-rentals, i.e. arbitrage profits (something for nothing). If at all possible, it is

desirable to retain this arbitrage feature of the CAPM benchmark; the strength of our approach set out below is that it is *entirely* based on arbitrage.

The economic logic of arbitrage can be illustrated by the following example. Assume that an investor knows that the return on fund *j* will exceed the return on investments in the same risk class; in which case he can “earn” excess returns without committing any of his own funds. He will achieve this by selling short other securities in *j*’s risk class and using the proceeds to purchase shares in *j*. Subsequently he will sell the shares in *j* to cover the short position. If his private information about *j* is correct, he will obtain an abnormal return. Conversely, if the investor had known that *j* would earn inferior returns, he would have adopted the reverse strategy, i.e. gone short in *j* and long in the other securities.

As noted above, the CAPM methodology can be viewed in arbitrage terms. Interestingly, we can interpret the CAPM parameters as investment weights. The expected return of fund *j* can be rearranged as

$$\bar{r}_j \equiv (1 - \beta_{jM}) r_f + \beta_{jM} \bar{r}_M \dots \dots \dots (9)$$

A return of the same risk class as *j* can be obtained by investing in *f* and *M* in the proportions $(1 - \beta_{jM})$ and β_{jM} . In arbitrage terms, this would mean going short in *f* and *M* in these proportions and using the proceeds to go long in *j*. As the CAPM holds only when *M* is efficient this must imply zero gains to arbitrage.¹⁴

The arbitrage approach, of course, does not depend on the existence of the CAPM. If the investor has “certain” (private) knowledge that the fund will do better (or worse) than average then he should be able to develop a “hedge” whereby he goes long (short) in the fund and short (long) in other assets in such a manner that his own wealth commitment is zero. When an arbitrager “knows” of a superior (inferior) profit opportunity and invests without committing any of his own wealth, then there is no risk. In practice, of course, arbitrage operations are risky because : (a) arbitragers cannot be sure that the information about r_j upon which they are acting is both private and certain; (b) it is impossible to be sure that one’s knowledge of the multivariate generating process is reliable. To the extent that futures markets are incomplete, thereby inhibiting short selling, the possibility and extent of arbitrage is reduced. However the risks of and restrictions on arbitrage operations, although very real, are not relevant for the purpose of evaluating fund performance. It is not to be supposed that investors will necessarily use (or get) private information about superior/inferior fund performance to make arbitrage gains; the above difficulties may be of sufficient magnitude to discourage them. The purpose of evaluating fund performance is to isolate the information aspect from risk and other considerations.

Briefly consider, in formal terms, the investment manouvres underlying the proposed notional arbitrage logic. Let x_j denote the amount of resources (available to the arbitrageur) invested in j such that it can be "covered" by short or long positions in $N-1$ other assets :

$$x_j = \sum_{i \neq j}^N x_i \dots \dots \dots (10)$$

N must be sufficiently large to average out random fluctuations in j 's return; Ross (22, p. 197) points out that it is necessary to hold a well-diversified portfolio in order to eliminate "noise". Also, the hedge must be such as to eliminate any systematic elements in security returns.

The arbitrage profit or loss from going long in j is

$$D_j = x_j \bar{r}_j - \sum_{i \neq j}^N x_i \bar{r}_i \dots \dots \dots (11)$$

A difficulty immediately presents itself: it is not possible to ascertain how much of the profit $D_j > 0$ is due to the superior performance of j and how much is due to the inferior performances of some of the shorted investments. There is difficulty in allocating D_j over the long position and the short portfolio. What is needed is an arbitrage strategy which involves shorting an investment called the "market" in order to be able to interpret the results as the extent to which fund j outperforms the market in general. As it is necessary to hold a well-diversified portfolio it might be possible to treat the second term on the RHS of (11) as the short position in assets including the market portfolio.

It is now possible to provide an arbitrage rationale for Jensen's α_j measure. The arbitrage strategy can be viewed from more or less the same position as the CAPM but without being burdened by the CAPM's mean-variance equilibrium restrictions. The RHS of (10) can be reduced to a portfolio of two assets, the riskless asset f and the market portfolio M , such that

$$x_j = x_f + x_M \dots \dots \dots (10a)$$

(f is included in order to complete the hedge. Any portfolio uncorrelated with M would serve the same purpose.) In which case, by (11), the arbitrage profit is

$$D_j = x_j \bar{r}_j - x_f r_f - x_M \bar{r}_M \dots \dots \dots (11a)$$

Rescaling by x_j gives

$$D_j/x_j = \bar{r}_j - (x_f/x_j)r_f - (x_M/x_j)\bar{r}_M \dots \dots \dots (11b)$$

which can be written more simply as

$$D_j/x_j = \alpha_j = \bar{r}_j - (1 - w_M)r_f - w_M \bar{r}_M \dots \dots \dots (11c)$$

where

$$w_M = x_M/x_j.$$

The problem reduces to the choice of w_M .

It is known that a significant proportion of the variation in security¹⁵ returns is accounted for by economy-wide factors. Hence we can write j 's return in terms of the following model¹⁶

$$r_j = a_j + b_j r_M + e_j \dots \dots \dots (12)$$

or, in terms of means,

$$\bar{r}_j = a_j + b_j \bar{r}_M \dots \dots \dots (13)$$

In order to eliminate the systematic element $b_j \bar{r}_M$, we must set $w_M = b_j$. Substituting (13) into (11c) gives

$$\begin{aligned} \alpha_j &= a_j + b_j \bar{r}_M - (1-b_j)r_f - b_j \bar{r}_M \\ &= a_j - (1-b_j)r_f \dots \dots \dots (14) \end{aligned}$$

This result depends on equation (12) being an adequate representation of the pattern of realised security returns. For arbitrage purposes it is necessary that e_j can be diversified away; this requires that the e_j 's be sufficiently independent across securities to allow the law of large numbers to hold, see Ross (23). (The Sharpe diagonal model more than meets this requirement in that the off-diagonal elements in the variance-covariance matrix are zero; this is a stronger assumption than is required.¹⁷) It is interesting to see what happens if the Sharpe-Lintner CAPM holds, i.e. when $r_z = r_f$ and M is efficient. In this case b_j is the beta of j with M and $a_j = (1-b_j)r_f$, thereby ensuring that $\alpha_j = 0$ (as efficient set mathematics demands). However, it is important to recognise that the arbitrage approach is *not* dependent on the CAPM holding; hence excess returns are possible.

The arbitrage approach does not avoid the problems involved in the (limiting)

case where the CAPM holds in its zero-beta form. As given in equation (4), fund performance is equal to the excess of \bar{r}_z over r_f times the proportion of shorts held in the riskless asset, whereas it should be identically zero. Following Roll's critique, there is no practicable way to ascertain empirically which of the two versions of CAPM is to be preferred. However, we have already pointed out in section II that in the case of UK fund performance the impact is likely to be small, particularly in comparison to the effects of relaxing the assumption of the efficiency of the market portfolio. In addition, there are other reasons for believing that the biasing effect is not likely to be large. In terms of the CAPM theory, \bar{r}_z should be substituted for r_f when there are (absolute) restrictions on going short in the riskless asset, i.e. on borrowing. Jensen (13, p.377) points out that whilst in times of inflation it is strictly true that a riskless asset is not available, the error made by assuming that there is a riskless asset is likely to be very small. In the empirical studies, average returns are calculated as the mean of logarithms of proportionate changes (i.e. average return is the logarithm of the geometric mean return over the sample observations) thus implying that holding of an investment is over an infinitesimally small interval – in which case, uncertainty introduced by inflation will be very minor. Thus we can usefully treat the riskless asset as synonymous with the zero-beta portfolio.

More important, even if it were strictly true that the riskless asset could not in practice be shorted, it does not really matter when assessing fund performance. All that our arbitrage approach shows is the returns to be made *if* shorting were possible, whereas a CAPM rationale for α_j places much emphasis on the identification of a model of security pricing. All that the arbitrage approach depends on is the identification of the systematic element in security returns. It should be recalled that if the zero-beta CAPM holds, resulting in excess returns of $\alpha_j = (\bar{r}_z - r_f)(1 - \beta_j)$, then all that our measure states is that abnormal gains/losses are possible *if* the the riskless asset f can be shorted. However such gains/losses can be made on all quoted securities, not just fund j . An appropriate check would be to compare the α of j with that of a "control" security i with the same beta. If $\alpha_j > \alpha_i$ then we can say that fund j 's performance is superior.

Indeed, the arbitrage approach can be used to incorporate any number of (linearly independent, i.e. orthogonal) systematic elements in security returns. There is evidence to suggest that there are a number of systematic elements; King (14), for example, has shown historically that whilst market-wide movements accounted for the largest proportion of variation in New York Stock Exchange security returns, industry and other factors together were nearly as important. This result is not consistent with the diagonal model, but is consistent with an expanded model of the form

$$r_j = a_j + b_{j1} r_M + b_{j2} r_I + e_j \dots \dots \dots (15)$$

where r_M is the market-wide factor and r_I is the industry-specific factor in which fund j is heavily invested. (If fund j invests heavily in more than one industry, then more than one factor would be appropriate.) Taking means, equation (15) becomes

$$r_j = a_j + b_{j1} \bar{r}_M + b_{j2} \bar{r}_I \dots \dots \dots (16)$$

As before, the investment in j is financed by short and long positions – in this case, in f , M and I :

$$x_j = x_f + x_M + x_I \dots \dots \dots (17)$$

Rescaling, we get

$$\begin{aligned} x_M/x_j &= w_M, \\ x_I/x_j &= w_I, \\ x_f/x_j &= 1 - w_M - w_I. \end{aligned}$$

In which case x_j is given as :

$$\alpha_j = \bar{r}_j - (1 - w_M - w_I) r_f - w_M \bar{r}_M - w_I \bar{r}_I \dots \dots \dots (18)$$

Setting

$$w_M = b_{j1}$$

and

$$w_I = b_{j2}$$

then (18) can be rearranged as

$$\alpha_j = a_j - (1 - b_{j1} - b_{j2}) r_f \dots \dots \dots (19)$$

As with equation (14), excess performance is defined as the firm specific factor minus a proportion of the return on the riskless asset.¹⁸ It is important to appreciate the nature of equations (12) and (15). In both cases, the firm-specific element in r_j is $a_j + e_j$; $b_{j2} r_I$ in (15) is that part of the firm-specific element in (12) which it has in common with other firms invested in the same industry. In other words, equation (12) decomposes r_j into two elements, whereas equation (15) further decomposes it into three elements.

The key issue is the identification of the market-wide and industry-wide factors. One approach would be to regress security returns on proxies for the market and industry factors; the difficulty with this though is that results are sensitive to the choice of proxies. On the other hand, the approach used by King employs factor-analytic techniques to identify the hidden systematic elements. The regression approach, in effect, imposes a model on the data, i.e. the factors are *predetermined*. On the other hand, factor analysis is a data-search technique which *identifies* the factors; it is a sample-specific method and is not concerned with the underlying generating process.

The role of the CAPM in the arbitrage approach to evaluating fund performance is clearly coincidental – more accurately, a CAPM-based approach gives the same results in the single-factor case. Where there is more than one systematic element (factor) in security returns the CAPM is misleading: notional arbitrage profits will not be riskless since non-market systematic elements will not have been hedged away. It is worth noting in this context that the existence of more than one important factor is itself evidence against the CAPM. The CAPM predicts, in effect, that there is only one systematic element in the returns, the general market factor. To the extent that there is known to be industry and other systematic effects on security returns, the CAPM does not fit the facts. And to the extent that the CAPM does not fit the facts, the CAPM is an imperfect arbitrage model for adjusting for risk. Roll's view on the untestability of the CAPM notwithstanding, findings such as King's *are* evidence against the CAPM in that a proxy for the market is not required (a market portfolio is not required). Conversely, Ross (21) shows that the efficiency of the market requires one-factor linearity. A word of caution: the factor-analytic studies are based on subsets of securities, not on the complete set, and therefore one cannot be certain that the results would hold for all securities. Intuitively though, it would be difficult to envisage the finding of strong industry effects in randomly selected subsamples when there is only one factor in the complete data set.

There are also difficulties concerning the plausibility of the CAPM. The foundations of the CAPM are to be found in portfolio theory, where it is assumed that the mean and the variance of portfolio returns are sufficient statistics. For mean-variance analysis to make economic sense either of two conditions must hold: (i) the distribution of returns can be completely specified by two or less independent parameters and utility functions are concave (risk aversion over the entire range of outcomes); or (ii) utility functions can be approximated by a quadratic which is monotonic non-decreasing over the entire range of outcomes. Whether these strong assumptions do grave injustice to the facts is far from clear – what is clear is that they are strong enough to cause us to think again about the appropriateness of assessing fund performance in terms of mean-variance portfolio theory. Given Roll's demonstration of the untestability of the CAPM, and also given the restrictiveness of its assumptions (which include i and ii above)

it is unwise to assess fund performance in terms of the CAPM without further inquiry into the appropriateness of the CAPM as a vehicle for adjusting for risk.

The arbitrage approach outlined in this section avoids all these difficulties. Furthermore, CAPM-based tests can be viewed as special cases of the arbitrage approach, appropriate when industry and other systematic effects (excluding the general market effects) are deemed to be of trivial magnitude. The arbitrage approach therefore offers a means of rescuing an extensive body of empirical research from the dustbin, the place to which Roil's results seemed to have consigned it.

IV. MEASUREMENT OF SYSTEMATIC ELEMENT(S)

Another difficulty with the CAPM-based tests concerns the way in which it is supposed that funds create excess returns. In Section II we pointed out an assumption underlying Jensen's tests, viz. that the generating process is stationary over time and that the CAPM describes this process exactly. Hence, as Jensen notes, superior and inferior fund performances are random disturbances around the expected value and thus "superior information" can only be with respect to particular realisations of the process. Therefore if a fund follows a buy-and-hold policy then superior returns must eventually be averaged out with inferior returns. The only way to avoid this is for the fund to make *at least one switch*; in which case the laws of large numbers need not apply. This might lead, though, to a fund's beta changing through time, contrary to Jensen's assumption. If so, the fund's risk will not have been held constant.

In order to see this more clearly, let us assume (a) that the Sharpe-Lintner CAPM fits "in the long run" and (b) that a "sufficient" number of observations have been taken to obtain reliable estimates of the parameter values of the process. This means that market portfolio M is efficient. Assume, first, that fund j adopts a buy-and-hold investment strategy over the investigation interval. The fund's mean return over the period can be written as:

$$\bar{r}_j = (1 - \beta_j) r_f + \beta_j \bar{r}_M \dots \dots \dots (20)$$

$$= (1 - w_M) r_f + w_M \bar{r}_M \dots \dots \dots (21)$$

where w_M can be interpreted as fund j 's proportionate "equivalent investment" in M .¹⁹ w_M is a *fixed weight* over the holding period. In these circumstances α_j is constrained to be zero. In order to obtain a positive excess return, the fund must vary the weight w_M . For example, assume that the fund "knows" at the beginning of the period that the market return will be temporarily depressed for part of the period, such that the average return during say the first half of the period will be below the (fixed) return on the riskless asset. Logically, the fund should invest heavily²⁰ during the first half of the period in the riskless asset; at the

end of this interval the fund should sell part or all of these fixed-return securities and reinvest the proceeds in the market portfolio, in order to take advantage of the rise which is expected to occur in the market. Notice though that as the market portfolio M is efficient over the holding period, equation (21) must hold. However we can no longer interpret the observed beta, β_j in equation (20), as the proportion of j's resources invested in M during the period; the portfolio weights change. Similarly, j's beta changes during the period; β_j is just the average beta. Regardless of these problems of interpretation, efficient set mathematics ensures that equation (20) must hold; hence $\alpha_j = 0$.

At first sight, this result seems contrary to commonsense, even paradoxical. If the fund manager "knows" when the market is going to fall and rise, and he acts on this knowledge, then surely he must make an abnormal gain? The answer, of course, is that (20) *assumes away the very possibility* of abnormal returns : if the market portfolio M is efficient ex post, then there is no way of "beating the market". The trouble is that fund j is itself part of this universe of securities whose average returns are defined by equation (20) and as such offers only an equilibrium return. For the fund j to beat the market, j must be *outside* the universe of securities assumed efficient. To assume that j is part of the set of securities deemed to be in equilibrium (such that M is on the frontier) is to *eliminate the possibility* of j being out of equilibrium.

Assume therefore that all the mutual funds whose investment skills are being investigated are *excluded* from the universe of securities comprising the "market". (By market we mean the quoted securities of industrial and commercial companies.) This is not an unrealistic assumption : open-ended mutual funds in the US and unit trusts in the UK are not quoted in the stock market : the returns on funds used in the empirical studies are calculated from data concerning their portfolios. This is not so with closed-end mutual funds and investment companies; these funds have quoted market prices for their own shares. If, in accordance with CAPM, the value-weighted portfolio of quoted securities is on the efficient frontier (of the set of quoted securities) it is then just conceivable that, by switching, the funds could earn excess returns. But how are we to interpret these excess returns (assuming they arise)?

Let us continue to concentrate on fund j which, it will be recalled, invested all its resources in the riskless asset f and then switched all these resources to M for the remainder of the holding period. A number, a_j , can be found such that

$$\bar{r}_j = (1 - a_j) r_f + a_j \bar{r}_M \dots \dots \dots (22)$$

as long as $r_f \neq \bar{r}_M$ (which is so, given that M is on the positive-sloping part of the efficient frontier). It needs to be emphasized that a_j cannot be interpreted as an investment weight, for the reasons given above; there are difficulties interpreting it as β_j , given that now j is not part of the universe of securities comprising the market.

It is of course true that a slope coefficient could be obtained from a time series regression, but it is not possible to interpret this as the beta of CAPM theory. By switching investments the fund might be changing its risk (certainly so in the example given above) and it is difficult to see how the investigator can aggregate betas over different time periods; conversely, there are (apparently insoluble) problems in disaggregating an estimate of "average" risk. See Roll (19, p.88).

The arbitrage approach avoids these difficulties for the following reasons :

- (a) It is not assumed that the market is in long-run equilibrium; hence it is possible for funds with information of the kind in the above example to earn excess returns.
- (b) There are no special difficulties with changing risk consequent upon switching. The notional arbitrage operation is riskless, because the arbitrage investment weights are determined with the benefit of hindsight.

Point (b) needs amplification. In the CAPM, each time the fund changes its investments, there may be a change in beta. Consequently it is necessary to estimate j 's beta over these different intervals and somehow average them, whereas with the arbitrage approach, how the fund's returns came about is of no consequence; all the researcher has to do is to estimate the systematic components and use these to create a notional hedge.

A further virtue of the arbitrage approach is that it avoids the need to determine from sample data the nature of the underlying generating process. All that is required is to calculate ex post, from a given set of data, the systematic elements in all security prices. The most appropriate technique for this is factor analysis.

In the remainder of this section, we shall consider a potential source of error in interpreting Jensen's study from an arbitrage viewpoint, a source stemming from the desire understandable from a CAPM perspective to get accurate estimates of population betas.

For expositional purposes, assume that the Sharpe-Lintner CAPM holds "in the long run". Roll's critique of Jensen's methodology, discussed in Section II above, assumes implicitly that the sample data used for estimating betas are the same as those used for calculating excess performances. Empirical studies, though, have often used different data sets to estimate betas. For where a reliable estimate is sought of the underlying population betas then it would seem natural to use all the data available to estimate betas. Thus in Jensen's study "the risk coefficient is estimated from all data available on each fund [i.e. not only the data for the period 1955-64 used to evaluate fund performance but as far back as the data go] in order to minimise the sampling error in the estimates" (12, p.217). We shall show, though, that from an arbitrage viewpoint, using additional data could result in misleading estimates of fund performance, in theory at least. Whether the error is a substantial one in practice is

an empirical issue to which we shall also address ourselves.

Jensen's use of prior data is legitimised, from a CAPM standpoint, by the evidence he provides supporting the hypothesis that betas are stable through time. However, whether or not betas are stable is a matter of some controversy at the present time of writing. On the one hand, Brenner and Smidt (3) and Fabozzi and Francis (5) have found that betas in different periods could have been drawn from the same samples; Blume (2), on the other hand, argues that estimated betas tend to regress toward the grand mean, suggesting nonstationarity. Cranshaw's (4) findings suggest that the evidence concerning UK unit trust betas is consistent with the hypothesis that they are all drawn from the same population! More research on this issue is clearly needed. Later in this section we demonstrate that, from an arbitrage perspective, it is important that *sample* estimates be (more or less) constant over time. It is not clear whether all studies of mutual fund performance utilising prior information meet this requirement. For example, Jensen's Table 5 (12, p.218) shows that individual fund betas for the period 1955-64 often differ from those for 1945-64, even though they appear to be drawn from the same set of distributions.

First let us consider the problem from a theoretical perspective. Consider the following partitioned matrix of hypothetical security returns data. (Note: we are continuing to deal entirely with realisations; the matrix of returns below is to be viewed as observed ex post returns, not expectations.)

$$\underline{R}^E = \begin{bmatrix} \underline{R}^B \\ \vdots \\ \underline{R}^T \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1B} & r_{1,B+1} & r_{1,B+2} & \dots & r_{1,B+T} \\ r_{21} & r_{22} & \dots & r_{2B} & r_{2,B+1} & r_{2,B+2} & \dots & r_{2,B+T} \\ \vdots & \vdots \\ r_{N1} & r_{N2} & \dots & r_{NB} & r_{N,B+1} & r_{N,B+2} & \dots & r_{N,B+T} \end{bmatrix} \quad (23)$$

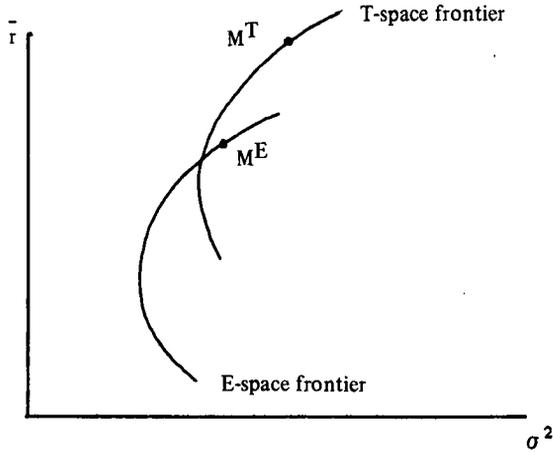
Each row of the matrix gives the observations available on a given security (there being N securities) and each column shows the security returns for a given period (there being $B+T$ periods). In the empirical studies, the β_j s could be computed in a variety of ways including (a) over an initial estimation interval of B periods or (b) over the entire observation interval $E = B+T$ or (c) over the test interval T . The β_j s thus computed are then used to construct the hedge for the test interval T .

To make this clearer, consider the procedure (b) where β_j s are computed from returns of the entire interval E . We shall write this β_j as β_j^E , the superscript E indicating that the beta is obtained from the pre-test and test data.

In Diagram II "E-space frontier" is a short-hand reference to the efficient frontier calculated from the whole $N \times E$ observations shown in the data matrix; likewise,

DIAGRAM II

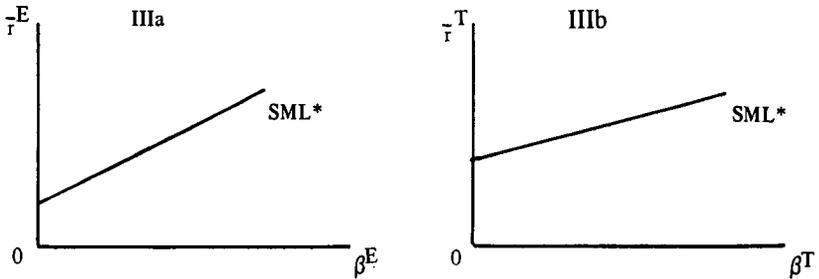
PORTFOLIOS IN MEAN-VARIANCE SPACES



“T-space frontier” refers to the efficient frontier for the last $N \times T$ observations. Assume initially that the proxy for the market portfolio M is efficient in both E-space and T-space, as indicated by the Diagram (the space involved is indicated by the superscript). It would seem that α_j must be zero given that M is efficient in both periods. This need not be so, though, as the beta obtained from E-space, β_j^E , is not necessarily that which would be obtained from T-space, β_j^T . This can be seen from Diagram III. Diagram IIIa gives the relationship in E-space between mean returns on any of the N securities (or combinations thereof), say j , and its beta β_j^E such that the intercept and slope have the required values of r_f^E and

DIAGRAM III

RETURN BETA RELATIONSHIPS IN E- AND T-SPACE



*SML = security market line

$(\bar{r}_M^E - r_f^E)$ respectively; all average returns on the N securities are given as points on this SML line. The same is equally true for T-space, as can be seen from Diagram IIIb with intercept and slope of r_f^T and $(\bar{r}_M^T - r_f^T)$; all the mean returns on the N securities are exactly "explained" by the SML line. But the slope and intercepts of the two SML lines can be significantly different in the two spaces, despite the market portfolio M being efficient in both sample spaces; this can be seen from the two diagrams. (For consistency with Jensen's analysis, we are assuming that in both spaces the riskless security is the appropriate orthogonal portfolio).

We can now see the impact on Jensen's performance index α_j of measuring beta in E-space and applying it to T-space. The performance measure is

$$\alpha_j = r_j^T - r_f^T - (r_M^T - r_f^T) \beta_j^E \dots \dots \dots (24)$$

In effect, the long-run expected return for j , r_j^E , has been updated to incorporate the realised returns on f and M ; the revised return is then deducted from the realised return on j in T-space, r_j^T . Substituting in (24) for r_j^E yields

$$\alpha_j = (\bar{r}_M^T - r_f^T) (\beta_j^T - \beta_j^E) \dots \dots \dots (25)$$

For large T , we would expect economic forces to ensure that $r_f^T < \bar{r}_M^T$. Hence the sign of α_j depends on the relative magnitudes of the betas in E-space and T-space. If beta is exactly (approximately) constant over the two intervals, then by (25) α_j will be exactly (approximately) zero. On the other hand, funds with decreasing sample systematic risk, i.e. with $\beta_j^T < \beta_j^E$, will appear as inferior performers, with $\alpha_j < 0$; conversely, funds with increasing risk will appear to outperform the market.

Jensen calculated betas from observations in what we have here called E-space, the period 1945-64, in those instances where the funds had been in existence that long; otherwise he used observations in T-space, 1955-64, and as far back in B-space as possible. However CAPM tests of fund performance may be interpreted from an arbitrage viewpoint only if the betas are calculated from observations in T-space. In order to estimate the error caused by using prior observations, a part of Jensen's study has been reworked, from data given in his paper. The results are given below.

Columns (i) and (ii) of Table 1 report, for those 56 funds (out of 115) which were in existence over the whole period 1945-64, betas computed by Jensen from both E-space and T-space; these were extracted from Jensen's Tables 2 and 5. Our Table 1 also gives in column (iii) Jensen's estimates of performance for these funds. Column (iv) reports our estimates of the "performance bias" which could have occurred *assuming that the market was efficient in both E-space and T-space*. They are defined as α_j in equation (25) and provide an indication of the degree of error in Jensen's methodology, error stemming from using extra observations to

compute "more reliable" beta estimates. Column (v) shows estimates of fund performances in accordance with our arbitrage rationale, in which betas are calculated from the 1955-64 T-space observations alone. The calculations in columns (iv) and (v) are based on Jensen's estimates (15, p.222) of r_f^T and r_M^T , viz. the ten-year risk free return is given as $r_f^T = \log(1 + .03)^{10} = .296$, and the estimated return on the market portfolio as $r_M^T = \log(1 + .126)^{10} = 1.187$. Fund returns, defined as logarithms of wealth relatives, are to be found in Jensen's Table 2.

The degree of error from using the wrong beta, the one using data prior to the test period, is considerable, as can be seen by comparing the elements in columns (iii) and (iv) of Table 1. In Jensen's sample, his performance measure (which he labels δ^*) is biased downwards.²¹ The mean of the excess returns calculated by Jensen for these 56 funds is -0.076 , i.e. a negative yield of approximately $\frac{3}{4}\%$ per annum, whereas the mean of the excess returns calculated by us (reported in column (v) is -0.051 , i.e. a negative yield of $\frac{1}{2}\%$ per annum.

It is interesting to note that Jensen found that 34 out of the 56 funds had $\delta_j^* < 0$ (negative excess returns); on our corrected basis, a further 5 funds had negative excess returns, despite the fact that on average our α_j is larger than Jensen's δ_j^* .

This reworking of Jensen's study in order to assess the funds' performances from an arbitrage standpoint, has produced substantially the same results: fund returns net of management expenses and brokerage commissions (but gross of loading charges) do not "beat the market" on average. The shortfall, though, is not as great as Jensen's study indicates. In general, it seems that although Jensen's methodology is theoretically unsatisfactory from an arbitrage viewpoint, in that using betas derived from E-space would not secure for an investor the elimination of all the systematic risk in T-space, his results are still generally reliable. Whether or not other studies of fund performance which have similarly used observations outside of T-space to calculate betas can also be relied upon remains a subject for further research.

It is interesting to consider a number of additional relationships. As previously noted, Jensen found that there was a relationship between his fund measure and beta which he attributed to "noise". We have recomputed for the reduced sample the corrected R^2 between δ^* and β_{45-64} , and done the same for our α_j and β_{55-64} , yielding R^2 's of 0.435 and 0.368 respectively (both highly significant at the .01 level). From our arbitrage standpoint, there is no reason to require that a fund performance measure be independent of beta (as there is with the CAPM). It is interesting to note that the extent of the bias is not significantly correlated with beta: the R^2 between the bias estimates in column (iv) and the beta estimates in column (ii) is -0.098 which is not significant at the .05 level.

TABLE I
JENSEN'S ESTIMATES OF FUND PERFORMANCES 1955-64 CORRECTED
FOR BETA BIAS

ID NO.	FUND NAME	JENSEN'S ESTIMATES			OUR ESTIMATES	
		β_{45-64} (i)	β_{55-64} (ii)	δ^* (iii)	BIAS (iv)	CORRECTED α_j (v)
140	Aberdeen Fund	.907	.935	-.024	.024	-.049
141	Affiliated Fund	.892	.729	.063	-.145	.208
142	American Business Shares	.508	.394	.074	-.101	.174
145	Associated Fund Trust	.834	.869	-.158	.031	-.189
147	Axe-Houghton Fund B	.834	.789	-.187	-.040	-.146
1148	Axe-Houghton Fund A	.777	.651	-.180	-.112	-.067
2148	Axe-Houghton Stock Fund	.934	.722	-.194	-.188	-.005
151	Boston Fund	.593	.512	.122	-.072	.193
152	Broad St. Investing	.828	.759	.052	-.061	.113
153	Bullock Fund	.962	.888	-.094	-.065	-.028
157	Century Shares Trust	.774	.819	.028	.040	-.012
162	Chemical Fund	.819	.890	.199	.063	.136
163	The Colonial Fund	.890	.881	.002	-.008	.010
166	Commonwealth Inv.	.687	.648	-.116	-.034	-.080
169	Corporate Leaders---	.712	.751	.207	.034	.172
171	Delaware Fund	1.011	1.037	-.245	.023	-.268
174	Diversified Inv. Fund	.821	.726	-.218	-.084	-.133
175	Dividend Shares	.768	.776	.087	.007	.079
177	Eaton & Howard Balanced	.562	.548	.034	-.012	.046
178	Eaton & Howard Stock	.838	.865	-.026	.024	-.049
180	Equity Fund	.912	.909	-.052	-.002	-.049
182	Fidelity Fund	1.034	1.049	-.219	.013	-.232
184	Financial Ind. Fund	1.118	1.104	-.392	-.012	-.379
185	Founders Mutual	.995	1.050	.016	.049	-.033
187	Fundamental Investors	1.051	1.024	-.171	-.024	-.147
188	General Investors Trust	.662	.652	-.078	-.008	-.068
190	Group Securities-Common	.962	.900	-.094	-.055	-.038
2191	Group Securities-Aero-space	.685	.689	-.018	.003	-.021
195	Income Foundation Fund	.664	.657	-.167	-.006	-.160
198	Incorporated Investors	1.262	1.179	-.712	-.073	-.638
200	Investment Co. of America	.952	.985	.002	.029	-.027
201	Investors Mutual	.586	.606	.036	.017	.018
205	Investment Trust of Boston	1.402	.998	-.569	-.359	-.209
215	Loomis-Sayles Mutual	.548	.524	.064	-.021	.085
216	Mass. Investors Growth	1.058	1.027	-.058	-.027	-.030

TABLE I /continued
JENSEN'S ESTIMATES OF FUND PERFORMANCES 1955-64 CORRECTED
FOR BETA BIAS

ID NO.	FUND NAME	JENSEN'S ESTIMATES			OUR ESTIMATES	
		β_{45-64} (i)	β_{55-64} (ii)	δ^* (iii)	BIAS (iv)	CORRECTED α_j (v)
217	Mass. Investors Trust	.962	.970	-.065	.007	-.072
219	Mutual Investing Foundation	.815	.842	.009	.024	-.014
220	Mutual Investment Fund	.821	.805	-.361	-.014	-.347
221	National Investors Corp.	.970	.864	.080	-.094	.174
222	National Securities Stock	1.185	1.109	-.407	-.067	-.339
2223	National Securities Income	.890	.826	-.297	-.057	-.239
225	Nation-Wide Securities	.490	.482	.204	-.007	.211
226	New England Fund	.523	.377	.017	-.130	.147
233	Pioneer Fund	.758	.816	.131	.051	.078
236	George Putnam Fund	.704	.735	.000	.027	-.027
240	Scudden, Stevens-Clark Bal.	.603	.665	-.034	.055	-.089
241	Scudden, Stevens & Clark Common	.946	.997	-.046	.045	-.091
243	Selected American Shares	.940	.944	-.172	.003	-.176
245	State St. Investment	.827	.879	-.034	.046	-.080
251	United Accum. Fund	1.004	.961	-.079	-.038	-.040
252	United Income Fund	.949	.921	-.065	-.024	-.039
257	Wellington Fund	.585	.584	.033	.000	.033
259	Winconsin Fund	.796	.813	-.011	.015	-.026
260	Composite Bond & Stock	.435	.422	.089	-.011	.100
2261	Dodge & Cox Balanced	.635	.664	-.062	.025	-.087
263	The Knickerbocker Fund	.806	.746	-.212	-.053	-.158

Note:

- (1) The numbers in columns (iii), (iv) & (v) are truncated; hence (iii) minus (iv) does not equal (v), due to rounding error.
- (2) Funds shown above existed for the whole of the period 1945-64.
- (3) The corrected R^2 between Jensen's δ^* and our α_j estimates in columns (iii) and (v) is 0.838 which is highly significant at the .01 level (t -statistic = 16.909 with 54 degrees of freedom).

It should not be supposed that our correction of Jensen's finding is dependent upon the assumption that the market portfolio M is efficient in both E-space and T-space. This assumption was made for expositional purposes in order to demonstrate that

Jensen's approach would result in the performance measure α_j being a linear function of β_j^T even in the limiting case where the market is efficient and hence excess returns cannot, by definition, be made. In the more realistic case where the market is not automatically assumed to be efficient – hence it is conceivable that funds could beat the market – then the correction is still necessary in order to ensure that the correct notional arbitrage weights are employed. The arbitrage approach assumes that the systematic elements in fund returns of a *particular test interval* can be eliminated. If not, the constructed notional hedge is not riskless. Future research in this area should analyse more closely the systematic elements in security returns in the chosen test interval (probably by factor analytic methods) than has hitherto been done. The arbitrage interpretation of Jensen depends upon (a) there being only a single factor across security returns, and that (b) this single factor is adequately proxied by the chosen market index. Our reworking of Jensen's study assumes that the Standard and Poor index adequately represents the assumed single factor, and this is very unlikely to be the case.

V. CONCLUSIONS

This paper examines the validity of a widely used method of assessing mutual fund performance. The method owes its popularity to having the virtue of seemingly being able to control for the confounding effects of differences in risk (and also for gains/losses due to general market movements, i.e. for good/bad luck). Recent theoretical research by Roll and others suggests that the CAPM is an inappropriate vehicle for assessing investment performance. Our study, whilst corroborating and extending this finding, offers a different rationale for the CAPM-based tests, one which owes nothing to the CAPM and hence sidesteps Roll's objections. This new rationale is based on a notional arbitrage strategy. Above all else, the proposed arbitrage approach avoids having to assume that the market portfolio is efficient. Unfortunately when the market portfolio is efficient then there is no possibility of abnormal returns; the absence of such a condition precludes the use of the CAPM. An arbitrage approach avoids this logical difficulty. We have thus turned a full circle. The paper starts by outlining Roll's critique of Jensen's approach to measuring fund performance. We go on to suggest that the tests are not lacking in economic logic, if re-interpreted in terms of a notional arbitrage strategy. Hopefully, this avoids the need to scrap an entire body of research.

Nevertheless, these tests are not without deficiencies. There may be more than one common factor in security returns; for example, in addition to the general factor affecting all securities, industry class may "explain" part of the variation in returns. These other factors need to be purged; future research might well employ factor analytic methods to identify these common elements in returns rather than using the ubiquitous linear regressions. Research is also needed into the extent to which the stock market index approximates the general factor obtained by principal components analysis.

The CAPM-based tests are only approximations of the kinds of studies which the arbitrage approach calls for. Hence their reliability is in doubt. One particular source of error in Jensen's mutual fund study, viz. the use of prior observations in computing betas, is investigated in Section IV above. A reworking of Jensen's findings indicates that, whilst he has underestimated excess returns by about 30 per cent on average, his conclusion that most funds do not beat the market is correct. Although incomplete, this is an example of the kind of corroborative research which is currently needed. Accordingly, it seems somewhat premature to reject out of hand all the past tests of fund performance just because they fall short of the ideal. That is to counsel perfection. One must consider the plausibility of alternative views. There are good economic reasons for doubting that most mutual funds can outperform the market. Our reworking of Jensen shows that his results are robust with respect to quite serious measurement errors.

This paper has focussed on the measurement of mutual fund performance. However the CAPM-based performance measure has also been used in the assessment of semi-strong efficiency in answering such questions as : Is the stock market capable of making proper allowance for the effects on published company profits of changes in reporting methods? Does the market react to earnings announcements? The arbitrage rationale of the CAPM-based tests of fund performance may serve equally well as a partial justification for these other studies. Further research is needed in this area.

Notes

- 1 Mutual fund is the American term for what is usually referred to in the UK as a Unit Trust. As most of the evidence is of US origin, we shall employ the American term in this paper.
- 2 Cranshaw (4) provides corroboratory evidence about the performance of UK unit trusts. However his methodology is different to Jensen's and therefore will be the subject of further comment below.
- 3 At the time of writing, only Part I has been published (20). Thanks are due to R. Stapleton for providing us with a copy of the full manuscript (19).
- 4 Will hold exactly, that is, in the sense that the regression parameters will have the precise magnitudes of (2) and (3). It does not however follow that the fit of the regression line counterpart of (1) will account for all or even most of r_i 's variation. In an empirical test of the CAPM the calculated Ordinary Least Squares (OLS) regression line will be

$$r_{it} = a_{oi} + a_{li} r_{pt} + e_{it} \quad \dots \dots \dots (1a)$$

when \bar{r}_p is on the efficient frontier, then (as shown below in this section) a_{oi} and a_{li} have the precise values given in (2) and (3), regardless of the 'goodness of fit' (given by the R^2 , say) of (1a). Equation (1) is just the sample mean of (1a).

- 5 Note that the assumption of $\bar{r}_z = r_f$ is not independent of the assumption of the efficiency of the market portfolio. The very existence of portfolio z is dependent on the market portfolio being efficient. Moreover, the difference $\bar{r}_z - r_f$ depends on the position of the market portfolio on the efficient frontier. The sign of this difference also depends on the position of r_f in relation to the return on the global minimum variance portfolio.

- 6 Implicit in such tests of the CAPM is the assumption that the proxy for the market portfolio is efficient.
- 7 This may be due, of course, to the choice of market proxy.
- 8 Identity (6) can be shown to be consistent with (and a simplification of) Roll's (20, p.168) identity A.23, as follows. Taking the j -th row of Roll's equations A.23, with suitable changes in notation, yields

$$\bar{r}_j \equiv r_f + (\bar{r}_M - r_f) \beta_{jM} + [(r_M - r_f) / \sigma_M^2] \sum_i [x_{pi} \sigma_M^2 / \sigma_P^2 - x_{Mi}] \text{cov}(r_j, r_i)$$

where x_{P_i} and x_{M_i} are the weights of security i in portfolios P and M, respectively, such that $\sum_i x_{P_i} = \sum_i x_{M_i} = 1$. The expression after the summation sign Σ reduces to

$$\text{cov}(r_i, r_P) \sigma_M^2 / \sigma_P^2 - \text{cov}(r_j, r_M).$$

Substituting the definitions of beta, the above reduces to identity (6).

- 9 This can easily be seen as follows. For the third term of (6) to vanish we require $\beta_{jP} = \beta_{jM}$, hence that

$$\text{cov}(r_j, r_P) / \sigma_P^2 = \text{cov}(r_j, r_M) / \sigma_M^2 \dots \dots \dots (6a)$$

Recall that $\sigma_P^2 \neq \sigma_M^2$, by definition. Equation (6a) can be expanded as

$$\sum_{i=1}^N x_{P_i} \text{cov}(r_j, r_i) / \sigma_P^2 = \sum_{i=1}^N x_{M_i} \text{cov}(r_j, r_i) / \sigma_M^2 \dots \dots \dots (8)$$

For M and P to be different portfolios, the weights of at least two securities must differ. If we assume that out of the universe of N securities only two do so (securities 1 and 2, say), we still cannot cancel out the remaining $N-2$ terms in (8) because they are normalised by variances of possibly different magnitudes. Hence all N terms must be considered, making the required equality highly improbable.

- 10 It is important to recognise that zero α_j does not imply the efficiency of the benchmark market portfolio. Roll shows that efficiency can only be inferred if return-beta linearity is observed for each individual security comprising the market portfolio. Thus whilst it is possible to infer that $\alpha_j = 0$ if the market portfolio is efficient, it is not possible to infer the converse that $\alpha_j = 0$ implies that the market portfolio is efficient. (We are grateful to Bill Fung for pointing this out to us.) For the purpose of evaluating fund performance, it is the first direction of causality which is important.

- 11 This is the first of the two tests Cranshaw reports; in the second test he uses the CAPM to adjust for risk.

12 cf. note 7.

- 13 Recall that for the CAPM to hold exactly (approximately) the market portfolio must be exactly (approximately) on the efficient frontier. What we are assuming here is that both the market portfolio and the proxy for the market portfolio used in assessing fund performance are efficient, viewed both ex ante and ex post. This enables us to disregard the lower-order problems arising from the market proxy behaving differently to the value-weighted

market portfolio.

- ¹⁴ Ross (22) demonstrates that when security returns can be explained by a Sharpe-type diagonal model and arbitrage profits are zero, then the CAPM holds.
- ¹⁵ Even more so with fund returns. See Modigliani and Pogue (16).
- ¹⁶ For our purposes, (12) is an empirical model concerning the joint conditional probability distribution of security returns; it is not to be treated as a theoretical model deduced from a priori specifications of investors' utility functions.
- ¹⁷ We are grateful to Bill Fung for pointing this out to us.
- ¹⁸ As $w_M + w_I$ can exceed unity this proportion can have a negative sign.
- ¹⁹ It is not being asserted that fund j has in fact invested the proportion w_M of its resources in the market portfolio—it is a property of the CAPM that any security can be viewed as equivalent to an investment in the riskless asset and the market portfolio.
- ²⁰ The fund might invest *more* than all its resources in f , by shortselling M (if possible.)
- ²¹ The mean of the bias estimates given in column (iv) of Table 1 is -0.235 . Using a two-tailed t -test with 54 degrees of freedom, the t -statistic of -2.514 is significantly different from zero at the .02 level.

References

- (1) Black, F., "Capital Market Equilibrium with Restricted Borrowing", *Journal of Business* (July 1972), pp.444-454.
- (2) Blume, M., "Betas and their Regression Tendencies", *Journal of Finance* (June 1975), pp.785-796.
- (3) Brenner, M., and S. Smidt, "A Simple Model of Non Stationarity of Systematic Risk", *Journal of Finance* (September 1977), pp.1081-1092.
- (4) Cranshaw, T.E., "The Evaluation of Investment Performance", *Journal of Business* (October 1977), pp.462-485.
- (5) Fabozzi, F.J., and J.C. Francis, "Stability Tests for Alphas and Betas over Bull and Bear Market Conditions", *Journal of Finance* (September 1977), pp.1093-1099.
- (6) Fama, E.F., "Efficient Capital Markets : A Review of Theory and Empirical Work", *Journal of Finance* (May 1970), pp.383-417.
- (7) Fama, E.F., *Foundations of Finance* (Basic Books, 1976).
- (8) Financial Accounting Standards Board, *Tentative Conclusions on the Objectives of Financial Statements*, 1976.
- (9) Firth, M., "Unit Trusts : Performance and Prospects", *Management Decision* (No. 3, 1978), whole issue.
- (10) Friend, J., M. Blume and J. Crockett, *Mutual Funds and Other Institutional Investors* (McGraw-Hill 1970).

- (11) Henfrey, A.W., B. Albrecht, and P. Richards, "The U.K. Stockmarket and the Efficient Market Model : A Review", *Investment Analyst* (September 1977), pp.5-24.
- (12) Jensen, M.C., "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios", *Journal of Business* (April 1969), pp.167-247.
- (13) Jensen, M.C., "Capital Markets : theory and evidence", *Bell Journal of Economics and Management Science* (Autumn 1972), pp.357-398.
- (14) King, B.F., "Market and Industry Factors in Stock Price Behaviour", *Journal of Business* (January 1966), pp.139-190.
- (15) McDonald, J.G., "Objectives and Performance of Mutual Funds, 1960-1969", *Journal of Financial and Quantitative Analysis* (June 1974), pp.311-333.
- (16) Modigliani, F., and G.A. Pogue, "An Introduction to Risk and Return : Concepts and Evidence", *Financial Analysts Journal* (March-April and May-June 1974), pp.68-78, 69-85.
- (17) Ohlson J.A., "On the Theory of Residual Analyses and Abnormal Performance Metrics", unpublished manuscript, School of Business, University of California at Berkeley, March 1977.
- (18) Paine, N.R., "A Case Study in Mathematical Programming of Portfolio Selections", *Applied Statistics*, Vol. 1, 1966, pp. 24-36.
- (19) Roll, R., "A Critique of the Asset Pricing Theory's Tests", unpublished manuscript, Graduate School of Management, UCLA, 1976.
- (20) Roll, R., "A Critique of the Asset Pricing Theory's Tests, Part I : On Past and Potential Testability of the Theory", *Journal of Financial Economics* (March 1977), pp.129-177.
- (21) Ross, S.A., "The Capital Asset Pricing Model (CAPM), Short-Sale Restrictions and Related Issues", *Journal of Finance* (March 1977), pp. 177-183.
- (22) Ross, S.A., "Return, Risk and Arbitrage", in J. Friend and J.L. Bicksler, eds., *Risk and Return in Finance : Volume I* (Ballinger, 1977), pp. 189-218.
- (23) Ross, S.A. "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory* (December 1976), pp. 341-360.
- (24) Williamson, J.P., "Measurement and Forecasting of Mutual Fund Performance : Choosing an Investment Strategy", *Financial Analysts Journal* (November-December 1972), pp.78-84.

Copyright of Journal of Business Finance & Accounting is the property of Blackwell Publishing Limited. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.