Journal of Business Finance & Accounting, 27(1) & (2), January/March 2000, 0306-686X

An Analytical Confidence Interval for the Treynor Index: Formula, Conditions and Properties

MATTHEW R. MOREY AND RICHARD C. MOREY*

1. INTRODUCTION

The Treynor index (1965) is well known as one of the most widely used measures of portfolio performance. Based on the assumptions of the Capital Asset Pricing Model, the Treynor index is simply the ratio of the mean excess rate of return of the portfolio to the portfolio's Beta value. Since the true value of beta is rarely known, the portfolio performance assessment must rely on a point estimate of this value to arrive at a point estimate of the Treynor index. Hence, the accuracy of point estimate of the Treynor index is not known. This paper addresses an analytical method that yields, under a very reasonable condition, rigorous, analytical confidence intervals on the index. The analytical method uses as inputs: the point estimates of the various parameters; uncertainties in these point estimates; size of the sample; and the level of confidence desired.

The motivation for a confidence interval on the Treynor index is twofold. First, a confidence interval provides additional information for comparing portfolios.[1] For example, consider an investor who typically uses the Treynor index to evaluate

* The authors are respectively from the Department of Economics, Smith College, Northampton, MA, USA; and the School of Tourism and Hotel Management, Griffith University, Queensland, Australia. They would like to thank Rajiv Banker, Chris Blake, John Hughes, Emily Rosenbaum, Yusif Simman, participants at the 1998 Eastern Finance Meetings in Williamsburg VA, and especially Stephen Satchell, a referee, for helpful comments on the paper. (Paper received September 1998, revised and accepted April 1999)

Address for correspondence: Matthew R. Morey, Department of Economics, Smith College, Northampton, MA 010163, USA.
e-mail: mmorey@sophia.smith.edu

portfolios. Such an investor will identify two different portfolios as having very similar performance because they have similar Treynor index point estimates. However, such an appraisal does not at all consider the degree of accuracy in these Treynor index point estimates. One portfolio may have a very narrow confidence interval on its Treynor index while the other has a wide confidence interval. Such knowledge, if known, may sway an investor to prefer the former portfolio over the latter. To illustrate the importance of the above point, consider an analogy dealing with an evaluation of portfolios based on Jensen's (1968) alpha. Rating portfolios by their Treynor indices without regard to the degree of uncertainty in the point estimate values is somewhat similar to rating portfolios with Jensen's alpha without regard to their respective significance levels.

Second, a confidence interval on the Treynor index can be used, of course, to perform hypothesis testing. For example, if the 95 percent confidence interval of the Treynor index covers the value zero, then one cannot reject, at the 5 percent level of significance, that the Treynor index is different from zero. Up to now, the Treynor index has not permitted conventional hypothesis testing.[2]

The organization of the rest of this paper is as follows. Section 2 provides a literature review of past efforts on understanding uncertainties in the Treynor index and also provides a brief overview of our approach. Section 3 briefly describes the Treynor index itself. Section 4 derives the confidence interval for the Treynor index, yielding a closed form formula for computing the confidence interval under a very reasonable condition. Section 5 lists some of the properties of the Treynor confidence interval. Section 6 provides a numerical example on a real mutual fund. Section 7 illustrates the use of the confidence interval in the strategic choice between mutual funds with essentially identical Treynor point estimates. Section 8 provides some insights, via simulations, on the types of situations where our confidence interval is likely to work. Section 9 consists of some sensitivity excursions, via simulation, dealing with the changes to the width of the confidence interval when one varies the sample size and the standard error of the basic Jensen regression (used to obtain the α and β for a portfolio). Finally the last section provides some caveats and possible extensions. An Appendix includes a proof of a (well-known) result needed in the development of the confidence interval.

2. REVIEW OF RELEVANT LITERATURE AND OVERVIEW OF OUR APPROACH

Previous work on developing a confidence interval on the Treynor index is almost non-existent. In a related study, Jobson and Korkie (1981) attempt to conduct hypothesis testing on certain transformations of the Treynor index. In one approach, they conduct hypothesis testing on the transformed difference of two portfolio's Treynor indices. Their main conclusion, however, was that the asymptotic distribution that they derived was not well-behaved for sample sizes less than 480 and, in general, did not recommend its use. In another approach, they attempt to conduct hypothesis testing of one portfolio's transformed Treynor index against many other portfolio's transformed Treynor indices. They find that the test properties of this multi-comparison method are not very satisfactory. Hence, while their paper attempts to conduct hypothesis testing, their method is only for the transformed difference of the Treynor index and says very little about constructing confidence intervals on the actual index. Moreover, they do not recommend the use of the Treynor index hypothesis tests.

In other related work, Cadsby (1986) comments on Jobson and Korkie (1981) and finds that their tests should be interpreted with caution because they possess no power to distinguish between the null and certain plausible alternative hypotheses. Finally, Kryzanowski and Sim (1990) develop an approach similar to that of Jobson and Korkie using non-synchronous trading.

Our approach, in contrast, employs a clever device, originated by Roy and Potthoff (1958), to construct confidence intervals for the ratio of means from a correlated bivariate normal distribution. Nothing is required for our approach beyond the standard output of the regression packages used to estimate the beta parameter. These inputs allow us to generate a quadratic equation, the roots to which (if real) result in the confidence interval limits for a portfolio's Treynor index. A similar approach is used by Morey and McCann (1983) to develop a confidence interval related to the optimal marketing mix, a la the Dorfman-Steiner Theorem and a log-linear regression specification, and by McCann, Morey and Raturi (1991) to develop a confidence interval for total advertising impact, using Koyck regression modeling.

3. THE TREYNOR INDEX

The Treynor index is defined as the risk premium earned per unit of systematic risk, where systematic risk is measured in terms of the Beta, β, of the portfolio. This index has an advantage over the Jensen index in that the Treynor index takes into account that an investor can borrow at the risk-free interest rate and leverage a position in a particular stock portfolio by selling the risk-free bond and using the funds to invest back in the particular stock portfolio. The motivation for the Treynor index comes from the well-known equilibrium relationship between non-diversifiable risk and expected return popularly referred to as the Capital Asset Pricing Model (CAPM):

$$E(R_p - R_f) = \beta_p[E(R_m - R_f) \quad (1)$$

where:

R_p = the return on the pth portfolio,
R_m = the return on the market portfolio,
R_f = the risk-free rate of return,

$$\beta_p = \text{Covariance}\,(R_p, R_m)/\text{Var}\,(R_m).$$

By dividing both sides of equation (1) by β_p (where it is assumed $\beta_p \neq 0$) we obtain:

$$[E(R_p - R_f)]/\beta_p = E(R_m - R_f). \quad (2)$$

Since the right-hand side of equation (2) does not depend upon p, it implies that in equilibrium, portfolios should be priced such that:

$$[E(R_p - R_f)]/\beta_p = [E(R_q - R_f)/\beta_q, \text{for all } p, q. \quad (3)$$

To calculate an estimate of the Treynor index for a given portfolio, we use Jensen's (1968) approach. By assuming that the CAPM holds period by period, and that the returns on securities are generated by the market model, Jensen utilized a linear relationship between the realized returns on any portfolio at time

t, $R_{p,t}$, the realized returns on the market portfolio at time t, $R_{m,t}$, and the realized risk-free returns at time t, $R_{f,t}$. The specification is:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \mu_t \qquad t = 1, 2 \ldots, T \tag{4}$$

where α_p, β_p are the *true values* of the parameters, and μ_t is the error term. Then $\hat{\alpha}_p$, $\hat{\beta}_p$, are, of course, the point estimators of α_p, β_p, where:

$$\hat{\beta}_p = \frac{\sum(X_t - \overline{X})(Y_t - \overline{Y})}{\sum(X_t - \overline{X})^2} \text{ and } \hat{\alpha}_p = \overline{Y} - \hat{\beta}_p\overline{X} \text{ where } X_t = R_{m,t} - R_{f,t} \text{ and } Y_t = R_{p,t} - R_{f,t}. \tag{5}$$

Hence, the *estimated* Treynor Index for portfolio p, is given by:

$$\hat{T}_p = E(R_p - R_f)/\hat{\beta}_p = \frac{\hat{\alpha}_p + \hat{\beta}_p E(R_m - R_f)}{\hat{\beta}_p} \qquad \hat{\beta}_p \neq 0 \tag{6}$$

where $\hat{\alpha}_p, \hat{\beta}_p$ are the point estimates of α_p and β_p, the true but unknown parameters for portfolio p.

Then, the true Treynor index for this portfolio, evaluated at the point $\overline{R}_m - \overline{R}_f$, is of course:

$$T_p = \frac{\alpha_p + \beta_p(\overline{R}_m - \overline{R}_f)}{\beta_p} \qquad \beta_p \neq 0 \tag{7}$$

and it is this parameter for which we wish to develop a $1 - \gamma$ confidence interval, i.e., a random interval that covers the true parameter (7), $1 - \gamma$ percent of the time.

4. THE CONFIDENCE INTERVAL APPROACH FOR THE TREYNOR INDEX

Our confidence interval generation approach utilizes an approach developed by Roy and Potthoff (1958) to generate a confidence interval on the ratio of means for random variables from a correlated bivariate normal population.

The approach, customized to the Treynor index, is as follows. First consider the numerator of the right hand side of equation (6), i.e., the random variable:

$$\hat{\alpha}_p + \hat{\beta}_p(\overline{R}_m - \overline{R}_f). \tag{8}$$

It is well known that $(\overline{R}_p - \overline{R}_f)$, the mean excess return for the *portfolio*, equals (8), since the regression line always goes exactly through the means. For ease of notation, let $\overline{Y}$ = sample mean of $\{R_{p,t} - R_{f,t}\} = \hat{\alpha}_p + \hat{\beta}_p(\overline{R}_m - \overline{R}_f)$. Next, following Roy and Potthoff, we define a *new* random variable:

$$W = \overline{Y} - q\hat{\beta}_p \tag{9}$$

where q is the unknown *constant* (7).

Then we observe that $E(W) = 0$, since:

$$E(\hat{\alpha}_p) = \alpha_p$$

$$E(\hat{\beta}_p) = \beta_p$$

and that W is normally distributed since $\hat{\alpha}_p$ and $\hat{\beta}_p$ are normally distributed. Next consider the variance of the random variable W:

$$\mathrm{Var}(W) = \mathrm{Var}\overline{Y} + q^2\mathrm{Var}(\hat{\beta}_p) - 2\mathrm{Cov}(\overline{Y}, q\hat{\beta}_p)$$

$$= \frac{\alpha_Y^2}{n} + q^2\,\mathrm{Var}(\hat{\beta}_p) - 2q\,\mathrm{Cov}(\overline{Y}, \hat{\beta}_p), \text{ where } \sigma_Y^2 = \mathrm{Var}(Y). \tag{10}$$

But it is well known (see Appendix for proof) that if Y is the dependent variable in the regression, $Y = \alpha_p + \beta_p X + \mu$ (where μ is the error term), the covariance of $\overline{Y}$ and $\hat{\beta}_p$ is zero, i.e., they are orthogonal.

Hence:

$$\sigma_Y^2 = \mathrm{Var}(W) = \frac{\sigma_Y^2}{n} + q^2\mathrm{Var}(\hat{\beta}_p). \tag{11}$$

Also, $\frac{W}{\sigma_W}$ is normally distributed, with mean 0, standard deviation of 1.

Moreover, letting $s_w^2 = \frac{s_Y^2}{n} + q^2 s_{\hat{\beta}_p}^2$ where s_W^2, s_Y^2 and $s_{\hat{\beta}_p}^2$ are the *estimated* variances of W, Y and $\hat{\beta}_p$ respectively, then $\frac{W}{S_W}$ is distributed according to the t random variable with $n - 1$ degrees of freedom (where of course n is the size of the sample in the above estimations). Hence we have:

$$\text{probability}\left\{\frac{\overline{Y} - q\hat{\beta}_p}{\sqrt{\frac{s_Y^2}{n} \le q^2 s_{\hat{\beta}_p}^2}} \le t_{\frac{\gamma}{2}, n-1}\right\} = 1 - \frac{\gamma}{2}, \tag{12}$$

where $t_{\frac{\gamma}{2},n-1}$ is the deviate corresponding to the $1-\frac{\gamma^{th}}{2}$ percentile from a t distribution with $n-1$ degrees of freedom. To simplify the exposition, let $t = t_{\frac{\gamma}{2},n-1}$.

Then upon squaring both sides of the inequality in (12), we have:

$$\text{probability}\left(\frac{(\overline{Y} - q\hat{\beta}_p)}{\sqrt{\frac{s_Y^2}{n} + q^2 s_{\hat{\beta}_p}^2}} \leq t^2\right) = 1 - \frac{\gamma}{2}, \tag{13}$$

$$\text{or probability}\left((\overline{Y} - q\hat{\beta}_q)^2 - t^2\left(\frac{s_Y^2}{n} + q^2 s_{\hat{\beta}_p}^2\right) \leq 0\right) = 1 - \frac{\gamma}{2}, \tag{14}$$

where it is recognized that $\overline{Y}, \hat{\beta}_p, s_Y$ and $s_{\hat{\beta}_p}$ are all truly random variables. Then the left side of the inequality in (14) is a quadratic in q, namely:

$$\left(q^2(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2) - (2\hat{\beta}_p\overline{Y})q + \left(\overline{Y}^2 - t^2\frac{s_Y^2}{n}\right)\right) \tag{15}$$

where the coefficients of q squared, q, as well as the constant term are all random variables.

Hence we can write:

$$\text{Prob.}\left(q^2(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2) - (2\hat{\beta}_p\overline{Y})q + \left(\overline{Y}^2 - t^2\frac{s_Y^2}{n}\right) \leq 0\right) = 1 - \gamma/2. \tag{16}$$

Following Roy and Potoff's logic, expression (16) provides an acceptance region for the hypothesis that the Treynor index has a specific value of q. Subject to a very reasonable condition to be discussed shortly and that q has real values, the statement in (16) yields confidence bounds on:

$$q = T_p = \frac{\alpha_p + \beta_p(\overline{R}_m - \overline{R}_f)}{\beta_p}.$$

There is also the further restriction that (16) is a probabilistic statement on $\hat{\beta}_p, s_{\hat{\beta}_p}, \overline{Y}, s_Y$ for all *real* values of:

$$q = \frac{\alpha_p + \beta_p(\overline{R}_m - \overline{R}_f)}{\beta_p},$$

except for $\beta_p = 0$. Then, under the side condition to be stated momentarily, the quadratic in (15) will have *real* roots of:

$$2\overline{Y}\hat{\beta}_p \pm \frac{\sqrt{4\overline{Y}^2\hat{\beta}_p^2 - 4(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\left(\overline{Y}^2 - t^2 \frac{s_Y^2}{n}\right)}}{2(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)}. \tag{17}$$

But the amount under the square root in (17), (since the $4\overline{Y}\hat{\beta}_p^2$ terms cancel out), simplifies to:

$$4t^2\left(\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right), \tag{18}$$

thereby yielding roots to (15) of:

$$2\overline{Y}\hat{\beta}_p \pm \frac{\sqrt{4t^2\left(\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right)}}{2(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)} \tag{19}$$

or

$$\overline{Y}\hat{\beta}_p \pm t\frac{\sqrt{\left(\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right)}}{(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)}. \tag{20}$$

Now, we shall discuss our 'reasonable side condition':

it is that the 'β_p be statistically different from zero (at a level of significance of γ)' i.e.:

$$\left|\frac{\hat{\beta}_p}{s_{\hat{\beta}_p}}\right| > t_{\frac{\gamma}{2}, n-1}. \tag{21}$$

Then note first that if (21) holds, the coefficient of q^2 (in (15)), i.e., $\hat{\beta}_p - t^2 s_{\hat{\beta}_p}^2$ is non-negative.

Under the same condition (21), the amount under the radical in (20) is non-negative, and hence the roots to (15) are real. It follows then, under assumption (21), that in the interval along the q axis *between* the roots of the quadratic (15), the probabilistic expression of (14) holds.[3] Therefore (20) constitutes the $1 - \gamma$ confidence limits and interval for (7), the true Treynor index. Hence we have *if β_p is statistically different from zero* (at the level of significance of γ):

$$\text{Prob.}\left(\frac{\overline{Y}\hat{\beta}_p - t\sqrt{\overline{Y}^2 s^2_{\hat{\beta}_p} + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s^2_{\hat{\beta}_p})}}{\hat{\beta}_p^2 - t^2 s^2_{\hat{\beta}_p}} \leq T_p \leq \frac{\overline{Y}\hat{\beta}_p + t\sqrt{\overline{Y}^2 s^2_{\hat{\beta}_p} + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s^2_{\hat{\beta}_p})}}{\hat{\beta}_p^2 - t^2 s^2_{\hat{\beta}_p}}\right) = 1 - \gamma \qquad (22)$$

where $t = t_{n-1, \frac{\gamma}{2}}$.

The condition for β_p to be statistically different from zero (in order for our procedure to work) is very intuitive in retrospect since β_p is of course the denominator for the true Treynor index. If β_p is not statistically different from zero, the Treynor index has severe definitional problems.

In summary, if β_p is *not* statistically different from zero (at a level of significance of γ), we should not use (22). Other possibilities for the generation of a confidence interval, if the above is the case, will be discussed in the conclusions. We also note that even if a 95% confidence interval is not available, a confidence interval at a lower level of confidence (e.g. a 90% confidence level) might be available if $\hat{\beta}_p$ is significantly different from zero at the less stringent levels of significance. Section 9 provides some insights, via simulation, on situations where (21) is likely to hold.

5. PROPERTIES OF CONFIDENCE INTERVAL (22)

(It will be assumed that β_p is statistically different from zero at a level of significance of γ in what follows.)

(i) Note, if there are no uncertainties whatsoever, i.e., $s_{\hat{\beta}_p} = 0$, $s_Y = 0$, then:

$$\hat{\beta}_p = \beta_p, \overline{Y} = \alpha_p + \beta_p(\overline{R}_m - \overline{R}_f),$$

and the interval (22) collapses correctly to the constant (7).

(ii) Next, consider the *center* of the confidence interval, (22) namely:

$$\frac{\overline{Y}\hat{\beta}_p}{(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)} = \frac{\hat{T}_p \hat{\beta}_p^2}{(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)}. \tag{23}$$

Note that quite interestingly this is to the *right* of the point estimate $\hat{T}_p$ and not at $\hat{T}_p$ itself, as one might have guessed. However, in most cases, since $s_{\hat{\beta}_p}$ is quite small relative to $\hat{\beta}_p$, the shift is small.

(iii) Next, consider the *width* of the confidence interval (22). It is from (22):

$$\text{width of confidence interval} = 2t\frac{\sqrt{\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)}}{(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)}. \tag{24}$$

Note, for example, that as s_Y increases, the interval gets wider as expected. Also, it has a width of zero if the variances are zero. Also, as expected, (24) increases as $1 - \gamma$ gets larger (and hence as t increases), e.g. the 99% confidence interval is wider than a 95% confidence interval. A numerical example, for a real Fund, shows how much wider. The last section of the paper provides some sensitivity results, via simulation, on the average width of the 95% confidence interval as n and σ (the standard error of the Jensen regression) are varied.

(iv) Next consider the important necessary and sufficient conditions for the Treynor index to be *statistically* different from 0, at a level of significance of γ. This will be true only if either the left hand endpoint of the interval (22) is *strictly* greater than zero, or the right hand endpoint is strictly less than zero so that the confidence interval on T_p does not cover zero.

From (22), we see the first condition requires:

$$\overline{Y}\hat{\beta}_p - t\sqrt{\left(\overline{Y}^2 s_{\hat{\beta}_p}^2) + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right)} > 0 \tag{25}$$

$$\text{or } \overline{Y}^2 \hat{\beta}_p^2 > t^2\left(\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right) \tag{26}$$

$$\text{or } \overline{Y}^2 > t^2\left(\overline{Y}^2\left(\frac{s_{\hat{\beta}_p}}{\hat{\beta}_p}\right)^2 + \frac{s_Y^2}{n}\left(1 - t^2\left(\frac{s_{\hat{\beta}_p}}{\hat{\beta}_p}\right)^2\right)\right) \tag{27}$$

$$\text{or } \overline{Y}^2 > t^2\left[\left(\frac{s_{\hat{\beta}_p}}{\hat{\beta}_p}\right)^2\left(\overline{Y}^2 - t^2\frac{s_Y^2}{n}\right) + \frac{s_Y^2}{n}\right] \tag{28}$$

$$\text{or } \left(\overline{Y}^2 - t^2\frac{s_Y^2}{n}\right) > t^2\left[\left(\frac{s_{\hat{\beta}_p}}{\hat{\beta}_p}\right)^2\left(\overline{Y}^2 - t^2\frac{s_Y^2}{n}\right)\right]. \tag{29}$$

But if $\overline{Y}$ is also statistically different from zero (at a level of significance of γ), then:

$$\left|\frac{\overline{Y}}{\frac{s_Y}{\sqrt{n}}}\right| > t, \text{so that } \overline{Y}^2 - t^2\frac{s_Y^2}{n} > 0. \tag{30}$$

Hence upon dividing both sides of (29) by $\overline{Y}^2 - t^2\frac{s_Y^2}{n}$, (without reversing the direction of the inequality of (29)), (29) becomes:

$$1 > t^2\left(\frac{s_{\hat{\beta}_p}}{\hat{\beta}_p}\right)^2 \text{ or } \left(\frac{\hat{\beta}_p}{S_{\hat{\beta}_p}}\right) > t^2. \tag{31}$$

But this is the assumed condition (21) for use of the approach.

A similar analysis applies for the second condition, i.e.:

$$\overline{Y}\hat{\beta}_p + t\sqrt{\left(\overline{Y}^2 s_{\hat{\beta}_p}^2 + \frac{s_Y^2}{n}(\hat{\beta}_p^2 - t^2 s_{\hat{\beta}_p}^2)\right)} < 0 \tag{32}$$

and also yields condition (26).

Hence working backwards, we can see that a *sufficient* condition for T_p to be *statistically* different from zero (at a level of significance of γ) is that both $\overline{Y}$ and β_p be statistically different from zero. It is also easily seen that if $\overline{Y}$ is *not* statistically different from zero, the confidence interval will indeed cover zero so (30) is also a necessary condition for T_p to be statistically different from zero. This makes perfect sense in retrospect, as $\overline{Y}$ is of course the numerator of T_p.[4]

Note also that one can perform now a hypothesis test where the null hypothesis is of the form:

$$T_p \leq K \text{ or } T_p = K \qquad (K \text{ unrestricted}). \tag{33}$$

In order to reject (33), we require the left hand endpoint of (22) to be to the right of K (in the first case) or the interval not cover K (in the second case).

6. A NUMERICAL EXAMPLE

In order to illustrate all aspects of our approach, consider the Treynor index for the Janus Fund over the period July 1, 1990–June 30, 1995, where the regression is performed using monthly return data.[5] First of all, the mean *excess* monthly percentage return over the 60 periods for the Janus Fund was 0.6417%; additionally the monthly excess returns for the Janus Fund had a standard deviation of 3.2998. This compares favorably with the respective measures of 0.3707% and 3.3094 for the Standard and Poor's excess return percentages.[6]

(i) Regression Results

When one performs the regression (5), (where the independent variable is of course the monthly excess percent returns for the Standard and Poor index, and the dependent variable is the excess monthly percent returns for the Janus fund), the $\hat{\alpha}_{Janus}$ and $\hat{\beta}_{Janus}$ have point estimates of 0.2976[7] and 0.928 respectively, with standard deviations of 0.158 and 0.0480 respectively. The respective p values are 0.0648 and 0.0000 so that the α_{Janus} is not statistically different from 0 (at the 5% level of significance), but β_{Janus} is significant at levels of significance as low as 0.0001.

Hence our confidence interval approach applies and will yield confidence interval at levels of confidence up to 99.99%. Completing the regression results, the R^2 of the regression is 0.8663, the standard error of the regression is 1.217, and the correlation between $\hat{\alpha}_{Janus}$ and $\hat{\beta}_{Janus}$ is -0.1122.

(ii) The Point Estimate of the Janus Treynor Index Over July 1, 1990–30 June, 1995

First, consider the point estimate of the Treynor index for the Janus Fund.

One obtains:

$$\hat{T}_{Janus} = \overline{Y}/\hat{\beta}_{Janus} = \frac{\hat{\alpha}_{Janus} + \hat{\beta}_{Janus}(\overline{R}_m - \overline{R}_F)}{\hat{\beta}_{Janus}}$$

$$= \frac{0.298 + 0.928(0.371)}{0.928} = \frac{0.642}{0.928} = 0.6915. \quad (34)$$

However, one's intuition says very little about how accurate the above estimate, is, or even if its true Treynor index is statistically different from zero.

(iii) Confidence Intervals for the Janus Fund Over July 1, 1990–June 30, 1995

Consider first the 95% confidence interval on the parameter *estimated* in (34). Then upon summarizing the pertinent information, we have:

$$
\begin{aligned}
n &= 60 \\
t_{\frac{\gamma}{2},n-1} &= t_{0.025,59} = 2.001 \\
\hat{\alpha}_{Janus} &= 0.2976 \\
s_{\hat{\alpha}_{Janus}} &= 0.1581 \\
\hat{\beta}_{Janus} &= 0.9280 \\
s_{\hat{\beta}_{Janus}} &= 0.0479 \\
\overline{Y} &= 0.6417 \\
s_Y &= 3.2998 \\
\frac{s_Y}{\sqrt{n}} &= 0.4260.
\end{aligned}
$$

Hence $\overline{Y}$ is not statistically different from zero (at the 0.05 level of significance) since:

$$\frac{\overline{Y}}{\frac{S_Y}{\sqrt{n}}} = \frac{0.6417}{0.426} = 1.506 < 2.001.$$

Then, upon calculating (22), one obtains as the 95% confidence interval over the period July 1, 1990–June 30, 1995:

$$0.6989 \pm 0.9264 = (-0.2275, 1.6253). \tag{35}$$

Note first the interval (35) is real (as we knew it must be since $\hat{\beta}_p$ was statistically different from zero at the 5% level of significance). Note also the center of the interval (35), is 0.6989 and indeed is slightly to the right of the point estimate of 0.6915, as we knew it must be (from (23)). Also the width of the interval is 1.8528.

Other confidence intervals for the Janus Fund are included in Table 1 (for 3 other levels of confidence, ranging from 80% to

Table 1

Confidence Intervals for the Janus Fund Over the Period July 1, 1990–June 30, 1995

Confidence Level	*t factor*	*Confidence Interval*	*Approximate Width of Confidence Interval*
80%	1.297	(0.0965, 1.293)	1.197 (35% less than benchmark)
90%	1.672	(−0.0702, 1.4633)	1.5032 (19% less than benchmark)
95% (Benchmark)	2.001	(−0.2275, 1.6253)	1.8528 (n.a)
99%	2.661	(−0.5118, 1.921)	2.4328 (31.3% more than benchmark)

99%). Note, e.g., that a change from 95% to 90% in the confidence level (a reduction of about 5.3%) reduces the width of the interval for this case by about 19%. Further, an increase of 4.2% in the confidence factor (from 95% confidence to 99%) increases the width by over 31%. It should be stressed these particular magnitudes are only strictly applicable to this one mutual fund, over the one time period investigated. But it does clearly demonstrate the strong non-linearity of the relationship between the width of the confidence interval and the level of confidence desired. In a subsequent section, the sensitivity of the width of the interval to changes in the sample size and the standard error of the regression will be explored via simulation using the Janus Fund's set of parameters.

(iv) Hypothesis Testing for the Janus Fund

Note from the results of Table 1, we cannot reject the hypothesis that T_{Janus} is different from zero at any reasonable levels of significance. From the results displayed in Table 1, only at the 20% level of significance can one reject $T_{Janus} = 0$, as its 80% confidence interval does not cover zero.

7. USE OF THE CONFIDENCE INTERVAL IN STRATEGIC CHOICES

To illustrate a strategic use of the Treynor index confidence interval, consider an individual attempting to choose between the following two mutual funds with very similar Treynor index point, estimates. Table 2 shows: (1) the $\hat{\alpha}$ and $\hat{\beta}$ values, their

Table 2

A Comparison of Two Mutual Funds: Twentieth Century and Fidelity Puritan

Fund	$\hat{\alpha}_p$	*t-value for* $\hat{\alpha}_p$	$\hat{\beta}_p$	*t-value for* $\hat{\beta}_p$	*Standard Error of Regression*	R^2	$\overline{Y}$	S_Y	$\frac{s_Y}{\sqrt{n}}$	*t value for* $\overline{Y}$
Twentieth Century	0.887 (0.536)	1.653	1.461* (0.162)	8.993	4.129	0.582	1.428	6.335	0.818	1.746
Fidelity: Puritan Fund	0.425* (0.145)	2.919	0.698* (0.044)	15.850	1.120	0.866	0.683*	2.563	0.331	2.060

Fund	*Treynor Index Point Estimate*	*95% Confidence Interval*	*Width of the 95% Confidence Interval*
Twentieth Century	0.9775 (1.428/1.461)	(−0.143, 2.200)	2.343
Fidelity: Puritan	0.9780* (0.683/0.698)	(0.029, 1.958)	1.929

Notes:
Data period for regressions is July 1990 to June 1995. The spread between the market and the risk-free rate, $\overline{R}_m - \overline{R}_f$, being used to calculate the Treynor indices is the average monthly excess return over the 60 periods. Standard errors are in parenthesis. A* indicates the variable is significantly different from zero at the 5% level of significance.

standard errors, and associated t-values for each of the two funds; (2) the R^2 values for the regressions; (3) point estimates of the Treynor index values for each of the funds; and (4) the 95 percent confidence intervals for the Treynor index point estimates for the two funds. The numbers calculated in Table 2 are calculated from regression results using excess return monthly data from July 1, 1990 to June 30, 1995.

Table 2 shows that the mutual funds have essentially the same Treynor index point estimates. By relying only on point estimates of their Treynor indices, one would be indifferent between the two mutual funds. However, the lower R^2 and lower t-values for the Twentieth Century Fund all indicate that this fund has more uncertainty. But how much more, and what is its impact on the respective confidence intervals? Upon examination, the first mutual fund's 95% confidence interval is about 21.5% wider than that for the second mutual fund.

Also the first Fund has a confidence interval that covers zero so that its Treynor index is not statistically different from zero. Note also that the big difference is in the left hand end points of the two 95% confidence intervals[8]. Hence one could make a strong case to recommend the Puritan fund over Twentieth Century fund over this time period, relying on this added information previously unavailable to the investor.

8. SITUATIONS WHERE THE CONFIDENCE INTERVAL APPROACH IS LIKELY TO APPLY

We have seen that we require $\hat{\beta}_P$ to be statistically different from zero (at the level of significance of γ) in order to be able to generate a $1-\gamma$ confidence interval on the pth portfolio's Treynor index. Hence if one had already done the Jensen type regression, and one desired say a 95% confidence interval, one could simply ascertain whether or not the estimated beta was statistically significant at the 5% level of significance. For example, consider the case where $\hat{\beta}_p = 1, s_{\hat{\beta}_p} = 0.555$, from a regression where $n = 60$. Then

$$\frac{\hat{\beta}_p}{s_{\hat{\beta}_p}} = 1.8$$

which is not more than 2.001, i.e. $t_{59,\,0.025,}$ the reject threshold at the 5% level of significance. Hence a 95% confidence interval on the above stock's Treynor index is not available from this approach. But note the test statistic value of 1.8 *does* indeed pass the reject test at the 10% level of significance where the appropriate t threshold, $t_{n-1,\frac{\gamma}{2}}$, i.e. $t_{59,\,0.05}$, is about 1.672. Hence this means that indeed a 90% confidence interval would be available for this particular stock. Hence one use of our method is to be able to ascertain, *after already having performed the Jensen regression,* just how low the confidence factor must be dropped in order to yield confidence limits.

A possible second use of our basic approach, more tactical in nature, is illustrated in Tables 3–7. Here the issue being addressed is: 'Given a particular type of stock (e.g., airlines, gold mining, etc.), are there any guidelines for the stock analyst on how many time *periods* he should include in his Jensen regression in order to be able to have a good chance of being able to compute a $(1-\gamma)$th confidence interval on the stock's Treynor index?' In other words, the question being raised is:

(i) Given likely α and β values for the type of stock of interest;
(ii) Given a plausible value σ for the standard error of the Jensen regression that will result;
(iii) Given the means and standard deviations of the excess market returns *for various time horizons,* then what number of periods (n) should the analyst include in the Jensen-type regression in order for the resulting beta estimate to be statistically different from zero, at the level of significance of γ?

In order to provide some preliminary (but by no means exhaustive) insights on this issue, we did the following simulations:

For given values of α, β, n, and σ, we created the following draws:

(i) We first created n monthly observations of the excess market returns $\{R_{m,t}-R_{f,t}, t=1,2,\ldots,n\}$ assuming a normal distribution. To obtain reasonable parameter values, we used the *mean* and *standard deviation* of the excess monthly return for the Standard and Poor's for the period July 90–June 95. These values were once again 0.37% for the mean and 3.309 for the standard deviation.

(ii) We next needed to create n monthly observations for the excess returns of the stock of interest: $\{R_{m,t} - R_{f,t}, t = 1, 2, \ldots, n\}$. In order to do this, we first drew n disturbance (or error) terms from a normal distribution with mean 0, and standard deviation of σ, to yield the monthly error terms, $\mu_t (t = 1, 2, \ldots, n)$. Then we *calculated* n values of the stock's observed excess returns from the formula:

$$R_{p,t} - R_{f,t} = \alpha + \beta(R_{m,t} - R_{f,t}) + \mu_t \quad (t = 1, 2, \ldots, n).$$

(iii) We next performed the Jensen-type regression to obtain its Beta and its standard error, i.e. $\hat{\beta}_p, s_{\hat{\beta}_p}$ respectively.

(iv) We next computed $\left|\frac{\hat{\beta}_p}{s_{\hat{\beta}_p}}\right|$ and compared it to $t_{n-1,\frac{\gamma}{2}}$ for $\gamma = 0.05$.

(v) We repeated steps (i)–(iv) 100 times, each time using the same draws for step (i)[9]. We recorded the percent of times that β_p was statistically different from zero (at level of significance of 0.05).

We conducted the simulations using four different settings of β. For each setting of beta, we investigated the impact of varying three levels of α, four levels of n, and three levels of σ, on the likelihood that the resulting estimate of beta was statistically different from zero. In order to choose the settings of β to investigate, we used as a guide Mullins (1982) study. He found that most portfolios can be divided into three valuations of β: low risk (i.e., β values between 0.35 for Gold Mining stocks to 0.85 for banks); medium risk (liquor at 0.90 to 1.05 for steel); and high risk (shipping at 1.20 to airlines at 1.80). To provide insights on various types of portfolios, we investigated β levels of: 0.20 (low-beta); 0.60 (low-medium beta); 1.00 (medium beta); 1.50 (high beta). For the choice of alpha, Sharpe (1985, p. 587) displays the range of alphas (from regressions being done on a monthly basis). The great bulk of the alphas range from -0.25 to 0.25, with peaks at or near zero. Hence we varied α from -0.25, 0 and $+0.25$. Finally, moving to the selection of the non-systematic risk,

σ, we chose a wide range of σ's, ranging from 1, 3, 5 and 10. The justification for these values was twofold: first, in a random sample of growth mutual funds, using monthly data, we calculated their σ values; the σ's ranged from a low of 0.1 (and a beta of 0.999 and a R^2 of 0.999 for the Vanguard 500 Fund), to a high of 4.13 (with a beta of 1.46 and a R^2 of 0.58 for the Twentieth Century Fund). Moreover, of these Funds, 70 percent had a σ-value between 1.0 and 2.0. Secondly, Sharpe (1985, p.370) comments that a large number of securities have σ values (from monthly data) of between 5 and 10. Finally for the values of the n to investigate, we simply chose 36, 60, and 120; these values of course correspond to 3, 5 and 10 years of monthly data.

Tables 3–7 present the results. Note that for small beta values, the size of n and σ has the most impact on the likelihood that the resulting beta is significant. For example, in the lower part of

Table 3

Simulation Study of when Confidence Interval Procedure is Applicable: Low Beta Case ($\beta_p = 0.20$)

Examines the % of 100 Regressions when β_p is Significantly Different from 0 at the 0.05 Significance level

	$n = 36$	$n = 60$	$n = 120$
For $\alpha = -0.25$			
$\sigma = 1$	99%	100%	100%
$\sigma = 3$	23%	41%	68%
$\sigma = 5$	9%	18%	32%
$\sigma = 10$	7%	7%	11%
For $\alpha = 0.00$	$n = 36$	$n = 60$	$n = 120$
$\sigma = 1$	97%	100%	100%
$\sigma = 3$	31%	44%	64%
$\sigma = 5$	13%	20%	36%
$\sigma = 10$	4%	7%	9%
For $\alpha = 0.25$	$n = 36$	$n = 60$	$n = 120$
$\sigma = 1$	99%	100%	100%
$\sigma = 3$	20%	51%	73%
$\sigma = 5$	16%	22%	28%
$\sigma = 10$	4%	8%	15%

Table 4

Simulation Study of when Confidence Interval Procedure is Applicable: Low-Medium Beta Case ($\beta_p = 0.60$)

Examines the % of 100 Regressions when β_p is Significantly Different from 0 at the 0.05 Significance Level

	$n = 36$	$n = 60$	$n = 120$
For $\alpha = -0.25$			
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	96%	100%	100%
$\sigma = 5$	69%	93%	99%
$\sigma = 10$	26%	46%	56%
For $\alpha = 0.00$			
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	98%	100%	100%
$\sigma = 5$	60%	93%	99%
$\sigma = 10$	19%	35%	59%
For $\alpha = 0.25$			
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	68%	94%	100%
$\sigma = 10$	20%	31%	68%

Table 3, (where $\beta = 0.2, \alpha = 0.25$), small σ values ($\sigma = 1$) nearly always yield significant betas regardless of how small n is. However when $\sigma = 3$, 10 years of monthly data is needed (i.e., $n = 120$) to have a good chance (73% chance) that the β will be significant.

Note also in general the rather small impact of the size of α. Also for small betas, if $\sigma = 5$ or larger, the chances are quite low (see Table 3) that the resulting beta will be significant, even if 10 years of data are used. In summary, our approach will work most of the time, having difficulty only when the beta is small, the n is small, and the σ is large. In this rare case, the resulting beta will be insignificant and hence our procedure will not apply. Similar types of results were obtained for the 90% confidence interval case and are available from the authors on request. Some recourses when β is insignificant are mentioned in the Conclusion.

Table 5

Simulation Study of when Confidence Interval Procedure is Applicable: Medium Beta Case ($\beta_p = 1.00$)

Examines the % of 100 Regressions when β_p is Significantly Different from 0 at the 0.05 Significance Level

	$n=36$	$n=60$	$n=120$
For $\alpha = -0.25$			
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	96%	100%	100%
$\sigma = 10$	49%	87%	94%
For $\alpha = 0.00$	$n=36$	$n=60$	$n=120$
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	97%	100%	100%
$\sigma = 10$	52%	81%	95%
For $\alpha = 0.25$	$n=36$	$n=60$	$n=120$
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	100%	100%	100%
$\sigma = 10$	55%	74%	95%

9. SENSITIVITY OF THE CONFIDENCE INTERVAL WIDTH TO CHANGES IN n AND σ

The width of the Treynor confidence interval varies, of course, based on changes in either n and to changes in the standard error of the regression, σ. However, the degree and magnitude of the sensitivity is difficult to determine analytically. Consequently, we conduct a simulation study to assess the sensitivity of the confidence interval to changes in n and σ. For the simulations we use a procedure similar to that used in the last section, namely:

To access the sensitivity of the confidence interval width to changes in n:

(1) We first generate n observations of monthly excess market returns, $R_{m,t} - R_{f,t}\,(t=1\ldots n)$, from a normal distribution.

Table 6

Simulation Study of when Confidence Interval Procedure is Applicable: High Beta Case ($\beta_p = 1.50$)

Examines the % of 100 Regressions when β_p is Significantly Different from 0 at the 0.05 Significance Level

For $\alpha = -0.25$	$n = 36$	$n = 60$	$n = 120$
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	100%	100%	100%
$\sigma = 10$	87%	99%	100%
For $\alpha = 0.00$	$n = 36$	$n = 60$	$n = 120$
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	100%	100%	100%
$\sigma = 10$	85%	98%	100%
For $\alpha = 0.25$	$n = 36$	$n = 60$	$n = 120$
$\sigma = 1$	100%	100%	100%
$\sigma = 3$	100%	100%	100%
$\sigma = 5$	100%	100%	100%
$\sigma = 10$	87%	100%	100%

Again, to obtain reasonable parameter values, we used the mean and standard deviation of the excess monthly returns of the Standard and Poor's 500 for the period July 1990 to June 1995 ($n = 60$).

(2) We next needed to generate n observations of the excess portfolio returns, $R_{p,t} - R_{f,t}$ $(t = 1 \ldots n)$. In order to do this, and to assure reasonable parameter values were used, we first locked in the estimated alpha, beta, and σ for the Janus Fund over the 60 month period July 1, 1990–June 30, 1995. These values were $\hat{\alpha} = 0.2976$, $\hat{\beta} = 0.9280$, and $\sigma = 1.217$. We then drew monthly n disturbances, $\mu_t(t = 1, 2, \ldots, n)$ from a normal with mean zero and $\sigma = 1.217$. Then using these disturbances and the assumed *known* parameter values, we *calculated* the monthly excess portfolio returns, $R_{p,t.} - R_{f,t} = 0.2976 + 0.928(R_{m,t} -$

Table 7

Sensitivity of the Width of the 95% Confidence Interval to Changes in n

n	*Mean of 100 Simulated Confidence Interval Widths* ($\sigma = 1.217$)	*% Change from* $n = 60$
24	3.113	+53.30
36	2.520	+24.20
48	2.220	+9.40
60	2.029*	n.a.
120	1.319	−34.99
240	0.937	−53.80
360	0.762	−62.40

Notes:
For each value of n, a simulation was conducted in which the original parameters of the Janus Fund were used to construct 100 confidence intervals. The simulation is described in the text. For each value of n, the mean of the 100 confidence intervals is displayed. Also displayed is the percentage change in the mean width from the base case of $n = 60$ (5 years of monthly data).

Table 8

Sensitivity of the Width of the Confidence Interval to Changes in the Standard Error of the Regression, σ

Sigma	*Mean of 100 Simulated Confidence Interval Widths*	*% Change from* $\sigma = 1.217$
0.100	1.885	−6.30
0.500	1.910	−5.90
1.000	1.981	−2.40
1.217	2.029	n.a.
2.000	2.248	+10.80
3.000	2.672	+31.70
5.000	3.910	+92.80

Notes:
For each value of σ, a simulation was conducted in which the original parameters of the Janus Fund (for $n = 60$) were used to construct 100 confidence intervals. The simulation procedure is described in the text. For each value of σ, the mean of the 100 confidence intervals is displayed. Also displayed is the percentage change in the mean width from the base case of $\sigma = 1.217$.
* The reason for the slight difference in the mean width from that of the Janus fund of Table 1, that is 2.029 versus 1.8528, is that the excess market returns used in the calculation of Table 1 use the actual excess monthly market returns for the period July 1, 1990–June 30, 1995 whereas those in Tables 7 and 8 rely on draws from a normal distribution with the same mean and standard deviation as that of the Janus Fund over the specified period.

$R_{f,t}) + \mu(t = 1, 2, \ldots, n)$ which serve as our n monthly draws for the excess portfolio returns.

(3) Next, using the above draws values, we performed an OLS regression of the type (11) to estimate $\hat{\beta}$ and $s_{\hat{\beta}}$. We also calculate the mean and the standard deviation of the excess portfolio returns, i.e. $\overline{Y}$ and s_Y. From this information, assuming the $\hat{\beta}$ was statistically different from zero (at the level of significance of γ), we construct the width of the Treynor confidence interval (22).

(4) We repeat the above step 100 times, using new realizations just for μ_t. (To have drawn new draws from (1) above would have introduced another source of variation not desired). Each time we record the width of the confidence interval.

(5) The mean of the 100 confidence interval widths is calculated.

(6) Repeat steps (1) through (5) for different values of n.

To access the sensitivity of the confidence interval width to changes in σ:

(1–5) Same as Steps (1–5) as above. The only change is that we used a fixed value of $n = 60$.
(6) Repeat steps (1–5) for different values of σ.

Table 7 presents the confidence interval width sensitivity results for changes in n. In every simulation, the β was significantly different from zero, so none of the results had to be dropped. The table shows, as expected, that there is a clear reduction in the width of the confidence interval as n increases. Using $n = 60$ as a benchmark case, an increase to an n of 360 (a 600% increase) decreases the mean confidence interval width by over 62%. Conversely, a sample size of 24 (a 60% drop) results in a 53% increase in the mean confidence interval width. Table 8 shows the width sensitivity results for changes in σ. The table illustrates that as σ increases, there is also, as expected, an increase in the mean width of the confidence interval. However, the effects on the width seem to be smaller than in the case of changing n. Using $\sigma = 1.217$ as the benchmark case, an increase in σ to 5.0 (a 311% increase) produces a 92.8% increase in the width of the interval. A decrease in σ to a level of 0.1, well over a

92% reduction from the benchmark case, produces just a 6.3% decline in the mean confidence interval width. Only when σ is raised to the very high level of 5 is there a large increase in the width of the confidence interval.

10. CONCLUSION

One of the shortcomings of the well-known Treynor portfolio performance index is that, due to the index's distribution, it was not possible to easily construct confidence intervals on the index. This paper has presented a methodology to easily construct confidence intervals on this index. The methodology is based upon the work of Roy and Potthoff (1953) and uses information from a simple regression equation. In the paper, we have shown the very reasonable condition when our methodology is applicable, described the analytical properties of the confidence interval, presented hypothesis testing applications, determined necessary and sufficient conditions for the Treynor index to be statistically different from zero, and illustrated the use of the approach with several actual mutual fund returns. Moreover, we have conducted some simulations to identify, in a preliminary way, both when the approach is likely to be applied (i.e. β_p significantly different from zero), as well as the sensitivity of the confidence interval width to changes in certain parameter values.[10]

It is possible (in the less interesting cases where beta is not statistically different from zero) that our method will be unable to construct a confidence interval. However, recent work by Vinod and Morey (1998) on bootstrapping the Treynor confidence interval may handle these rare situations. Also use of Jobson and Korkic's asymptotic approaches may be of help in cases of hypothesis testing. Such cases notwithstanding, our method provides a relatively easy, 'back of the envelope' method for constructing confidence intervals on one of the most well-known portfolio performance measures. Clearly more effort is needed on the very important aspect of assessing the accuracy of point estimates of popular financial indices. The clever device of Roy and Pothoff may well be able to be utilized for other such measures.

APPENDIX

Proof that Cov $(\overline{Y}, \hat{\beta}) = 0$, from regression (5)

$$\mathrm{Cov}(\overline{Y}, \hat{\beta}) = E[(\hat{\alpha} + \hat{\beta}\overline{X} - (\alpha - \beta\overline{X}))x(\hat{\beta} - \beta)] = E[(\beta - \beta)(\alpha - \hat{\alpha}) + \overline{X}E(\hat{\beta} - \beta)^2] = \mathrm{Cov}(\hat{\beta}, \hat{\alpha}) + \overline{X}\mathrm{Var}\hat{\beta}$$

but:

$$\mathrm{Var}\hat{\beta} = \frac{\sigma_\mu^2}{\sum(X_i - \overline{X})^2}$$

where σ_μ^2 is the variance of error term in (5). See Johnston (1972, p. 20).

Similarly Johnston (1972, p. 21):

$$\mathrm{Cov}(\hat{\beta}, \hat{\alpha}) = -\frac{\overline{X}\sigma_\mu^2}{\sum(X_i - \overline{X})^2}.$$

Hence we have the desired result.

NOTES

1 Note that several analytical approaches have been developed to construct, from an opportunity set of securities, the mix of securities that maximizes the given portfolio's average performance index. See for example Elton, Gruber and Padberg (1976), and Faaland and Jacob (1981).

2 Jobson and Korkie (1981) attempt to deal with this issue. See the literature review (Section 2) for more details.

3 It should be mentioned that condition (21) is needed to guarantee that the coefficient of q^2 is non-negative so that the parabola of (15) 'opens upwards'. It is this condition which guarantees that, for the interval between the roots, the value of the quadratic will be *below* zero, and hence the condition in (16) is maintained. Therefore, even if the amount under the radical in (18) is non-negative (as it sometimes will be, *even* when (21) does not hold), condition (21) is still needed for the confidence interval approach to work. This point was missed in Roy and Pothoff's 1958 analysis where their thrust was on generating a confidence interval on $\frac{\mu_1}{\mu_2}$, the ratio of means from a bivariate normal population. In their analysis, it also turns out that the coefficient of the squared term in their quadratic is:

$$\overline{X}_2^2 - \frac{t^2 s_2^2}{n}$$

which needs also to be non-negative for their approach to work. But this requires μ_2 to be statistically different from zero, i.e.:

$$\left|\frac{\overline{X}_2}{\frac{s_2}{\sqrt{n}}}\right| > t.$$

This property, in retrospect, namely that the denominator must be statistically different from zero (at the level of significance of γ) in order to yield a $(1 - \gamma)$ confidence interval on the ratio of means, is very intuitive.

4 We also point out here that empirically it is often the case for actual mutual funds that, while the point estimate of the Treynor index is positive, the 95% confidence interval covers zero. In a random sample of the growth mutual funds for the period July 1990–June 1995, 100 percent of funds had a significant Beta coefficient and hence a confidence interval could be calculated on the Treynor index. However, 80 percent of the funds examined had Treynor indices *not* significantly different from zero.

5 The Janus fund is used since it is a well-known, large mutual fund. The data is from the Morningstar Principia program.

6 The risk free-rate was calculated from the three-month US T-Bill monthly returns.

7 The high alpha estimate is due to the Fund's superior performance *vis-a-vis* that of the S & P index. Section 9 discusses the types of monthly alphas and betas observed for different types of stocks.

8 This difference would be even larger if we compared say the left hand endpoints of the two 99% confidence intervals from the two Funds.

9 We used the same draws in (i) as the key thrust was on assessing the impact of varying σ, n, α, β on the likelihood that the confidence interval approach would work. Redrawing the market excess returns would introduce another source of variation not desired.

10 The usefulness of the confidence interval approach shown here has been presented in two ways. One, that it allows an investor more information on the performance of the portfolio. Two, it allows an investor to conduct hypothesis tests on whether the Treynor index is significantly different from a certain specified value. Our methodology does not however, allow for hypothesis testing of the *difference* in Treynor index values of two portfolios. This is an important aspect for one of the problems with both the Treynor and Sharpe (1965) indexes is that they may provide very strange results in the case of perverse sampling. Consider for example, a situation where due to the sampling period used, two portfolios have equal but negative mean returns. However, one of the portfolio has a beta ½ of the size of the other. In such a case, the fund with the higher systematic risk will indeed be ranked *higher* according to the Treynor index. Such cases as these document the need for hypothesis testing on the difference between two funds. Jobson and Korkie (1981) present a methodology for testing the transformed difference of the Sharpe and Treynor indexes. However their results for the Treynor index show that their test statistic is not well-behaved in small samples. Vinod and Morey (1998), using a bootstrap methodology, are able to conduct such a test on the actual difference of the Treynor indexes.

REFERENCES

Cadsby, C.B. (1986), 'Performance Hypothesis Testing with the Sharpe and Treynor Measures: A Comment' *Journal of Finance*, Vol. 41, No. 5 (December), pp. 1175–76.

Elton, E., M. Gruber and W. Padberg (1976), 'Simple Criteria for Optimal

Portfolio Selection', *Journal of Finance*, Vol. 31, No. 5 (December), pp. 1341–57.

Faaland, B.H. and N.L. Jacob (1981), 'The Linear Fractional Portfolio Selection Problem', *Management Science*, Vol. 27, No. 12 (December), pp. 383–89.

Jensen, M.C. (1968), 'The Performance of Mutual Funds in the Period 1945–1964', *Journal of Finance*, Vol. 23, No. 2 (May), pp. 389–416.

Jobson, J.D. and B.M. Korkie (1981), 'Performance Hypothesis Testing with the Sharpe and Treynor Measures', *Journal of Finance*, Vol. 36, No. 4 (September), pp. 889–908.

Johnson, J. (1972) 'Econometric Methods', 2nd ed. (McGraw-Hill, New York).

Kryzanowski, L. and A.B. Sim (1990), 'Hypothesis Testing with the Sharpe and Treynor Portfolio Performance Measures Given Non-Synchronous Trading', *Economics Letters*, Vol. 32, No. 4 (April), pp 345–52.

McCann, J.M, R.C. Morey and A.S. Raturi (1991), 'Confidence Interval for the Total Advertising Impact and its Mean Duration under Koyck Models', *Journal of the Academy of Marketing Science*, Vol. 19, No. 4, pp 333–40.

Morey, R.C. and J.M. McCann (1983), 'Estimating the Confidence Interval for the Optimal Marketing Mix: An Application to Lead Generation', *Marketing Science*, Vol. 2, No. 2, pp. 193–202.

Mullins, D.W. Jr (1982) 'Does the Capital Asset Pricing Model Work?' *Harvard Business Review*, Vol. 60 (Jan Feb), pp. 105–14.

Roy S.N. and R.F. Potthoff (1958), 'Confidence Bounds on Vector Analogues of the 'Ratios of Means' and the 'Ratio of Variances' for Two Correlated Normal Variates and Some Associated Tests', *Annals of Mathematical Statistics*, Vol. 29, pp 829–41.

Sharpe, W.F. (1985), *Investments*, 3rd ed. (Prentice Hall International).

Treynor, J.L. (1965), 'How to Rate Management of Investment Funds', *Harvard Business Review*, Vol. 43 (January–February), pp. 63–75.

Vinod, H.D. (1993), 'Bootstrap Methods: Applications in Econometrics', in G. Maddala, C. Rao and H. Vinod (eds.), *Handbook of Statistics*, Vol. 11, chp. 23, (Elsevier Science Publishers), B.V., pp. 629–61

_____ and M.R. Morey (1998), 'Confidence Intervals and Hypothesis Testing for the Sharpe and Treynor Portfolio Performance Measures: A Bootstrap Approach', Working Paper (Fordham University, New York, NY, USA).

_____ _____ (1999), 'Confidence Intervals and Hypothesis Testing for the Sharpe and Treynor Portfolio Performance Measures: A Bootstrap Approach', in Y. Abu-Mostafa, B. LeBaron, A. Lo and A. Weigend (eds.) in *Computational Finance 1999* (MIT Press).