Approximating the Confidence Intervals for Sharpe Style Weights

Angelo Lobosco and Dan DiBartolomeo

Style analysis is a form of constrained regression that uses a weighted combination of market indexes to replicate, as closely as possible, the historical return pattern of an investment portfolio. The resulting coefficients, called Sharpe style weights, are used to form inferences about a portfolio’s behavior and composition. This technique has been widely adopted in the investment industry, despite the fact that no explicit confidence interval measures have been available to describe the results. We derive an approximation for the confidence intervals of these weights and, using Monte Carlo simulation, verify its efficacy. The estimation of these confidence intervals can help practitioners assess the statistical significance of their results and aids in determining which indexes to include in the analysis. It may also encourage the use of daily return data to meaningfully reduce the size of the confidence intervals.

Style analysis, pioneered by William Sharpe (1992), has gained widespread acceptance among investment practitioners. First described by Sharpe (1988) as “effective asset mix analysis,” the method is a constrained form of regression analysis used to form inferences about the influences on and composition of investment portfolios, based solely on the historical returns to those investment portfolios. The approach is widely used by investment professionals seeking to understand the composition and investment approach of portfolios for which they have no data on the securities held therein. One of the problems with using the technique is that no explicit measures of the confidence interval on the resultant coefficients, called style weights, have been available.

A PRIMER ON STYLE ANALYSIS

Style analysis is a way to estimate an asset class factor model. Such a model would take the form

\[ R = b_1F_1 + b_2F_2 + b_3F_3 + b_4F_4 \ldots + b_nF_n + e, \]  

where

- \( R \) = time series of returns for a given portfolio
- \( F_k \) = time series of returns to factor (asset class) \( k \)
- \( b_k \) = sensitivity of the portfolio returns to factor \( k \)
- \( e \) = time series of portfolio residual returns

So far, we have a standard multiple regression equation in which the factors are the independent variables and the return to the portfolio is the dependent variable. In style analysis, we first define the factors as market indexes representing various asset classes or approaches (“styles”) of investing. For example, we could select indexes to represent investment in stocks, bonds, and money-market instruments. With our factors so defined, the value of a factor during any time period is merely the return on the market index for the time period in question.

Traditional multiple regression calculations would provide a set of coefficients \((b_k, k = 1,n)\) that would estimate how sensitive the return on the subject portfolio is to the returns of the set of market indexes selected as factors. If we hypothesize a causal relationship between the returns on the market indexes and the return on the subject portfolio, we can say that these sensitivity coefficients express the extent to which the returns on the subject portfolio are influenced by returns on the market indexes. Coefficients arising out of the traditional regression process may take on either positive or negative values. Investment practitioners often find it intuitively unappealing that a market index exerted a negative influence on the return on a portfolio. In addition to a frequent lack of intuitive appeal, traditional regression also may exhibit multicollinearity. If the independent variables are highly correlated, as two indexes representing different approaches to investing with the same asset class (say, corporate bonds and Treasury bonds) are likely to be, the reliability of the estimated coefficients in meaningfully describing the underlying
true relationship is very much in doubt.

In style analysis, we place two constraints on the range of values that the coefficients \( (b_k) \) may take. The first is that the coefficients have only values between 0 and 1. The second constraint is that the sum of the coefficients must equal 1; that is,

\[
0 < b_k < 1,
\]

and

\[
\sum_{k=1}^{n} b_k = 1.
\]

In this way, the coefficients have an intuitive interpretation as the weights of the assets within a portfolio and can be conveniently displayed as “slices” in a pie chart. So, style analysis can be thought of as the process of forming a portfolio of indexes that mimics, as closely as possible, the historical performance of a given portfolio. This “mimicking portfolio” is commonly referred to as the Sharpe style index. From the weights ascribed to the various market indexes (known also as Sharpe style weights), we form inferences regarding the behavioral influences on and likely composition of the given portfolio.

By placing these constraints on the values of weights, we also somewhat mitigate the multicolinearity problem of traditional regression. Placing bounds on the weights reduces the likelihood that high correlations between the independent variables will cause the coefficients to “blow up” to unrealistic values. The price we pay for this advantage is that weights that are calculated to be 0 or 1 are not unbiased estimators. A coefficient (weight) on a market index that is calculated to be 0 as derived by traditional regression might have come out negative but is prevented from doing so. Similarly, a value of 1 could have been greater than 1 in a traditional regression.

As in traditional regression, the weights are calculated so as to minimize the sum of the squared differences between the actual dependent variable (the subject portfolio) and the combination of the independent variables (the Sharpe style index). Because of the constraints on the weights, the problem of actually calculating the weights is solved using the technique of quadratic programming. Once the Sharpe style index has been determined, the historical returns on the subject portfolio can be compared with the returns of the Sharpe style index. The portion of the return variability of the subject portfolio that is explained by the Sharpe style index is the \( R^2 \) statistic.

Unless we have an index to represent every possible asset behavior, it is unrealistic to hope for a perfect fit (\( R^2 = 1 \) and standard error = 0). This statement is true even under the ideal circumstances of knowing exactly the style weights for a given portfolio and set of market indexes.

Empirical style analysis cases calculated in this fashion typically show \( R^2 \) values that are only marginally lower than \( R^2 \) values derived through traditional unconstrained multiple regression. What style analysis has lacked is a measure of the level of confidence that the individual coefficients (which express the weights in the Sharpe style index) are accurate estimates of the true underlying relationships. Other concerns that surround the use of style analysis are discussed in more detail by Christopherson (1995) and Trzcinka (1995). This article focuses on the issue of establishing confidence intervals for the Sharpe style weights.

The Problem of Linear Combinations

The solution to a style analysis problem is indeterminate if any of the market indexes used as factors can be formed as a linear combination of other indexes. Imagine doing a style analysis of some portfolio using three indexes: A, B, and C. Assume that Index B is formed by combining Indexes A and C in equal proportions. A style analysis that showed the subject portfolio was best mimicked by a portfolio of 30 percent Index A, 40 percent Index B, and 30 percent Index C would be exactly equivalent to a result of 50 percent Index A, 0 percent Index B, and 50 percent Index C, because the 40 percent weight given to Index B can itself be replicated by an equal combination of A and C. In each case, the time series error terms and the \( R^2 \) statistic would be identical. Style analysis can only reliably attribute portfolio returns to the portions of the market index returns that are themselves not attributable to the returns of the other indexes.

In 1992, Sharpe stated “asset class returns should have either low correlations with one another or, in cases in which correlations are high, different standard deviations.” This statement is a qualitative description of selecting the market indexes so as to avoid the likelihood that any one index would be a linear combination of others.

Unfortunately, many practitioners fail to heed this warning and routinely select market indexes that are highly correlated and often are prima facie likely to be linear combinations of other indexes. For example, intuition might tell us that an index of mid-capitalization stocks is likely to be closely reproduced by a weighted combination of a large-cap stock index and a small-cap stock index. To describe more fully the portfolios under study, practitioners often increase the number and similarity level of the market indexes used in the analysis. Without explicit confidence intervals on the style weights, the damage done to the reliability of
the analysis can go undetermined and undetected.

**Calculating the Confidence Intervals**

In doing a style analysis, we are seeking to identify the “true” underlying relationship between the returns on the subject portfolio and the returns on the market indexes that we choose as factors. Unfortunately, as in any statistical procedure, the best we can hope to do is to estimate the underlying relationship. The style weights that result from performing a style analysis can be thought of as estimates of the true style-weight combination of market indexes. The standard deviation of these estimates is approximated by the following formula:

\[
\frac{\sigma_{\text{a}}}{\sigma_{B_i} \sqrt{n - k - 1}},
\]

where

\(i\) = index corresponding to the style weight being estimated
\(\sigma_{\text{a}}\) = standard error of the style analysis
\(\sigma_{B_i}\) = “unexplained Sharpe style index volatility” for index \(i\)
\(n\) = number of returns used in the style analysis
\(k\) = number of market indexes with nonzero style weights

The derivation of this formula and the definition of unexplained Sharpe style index volatility are covered in detail in the appendix. As we might intuitively expect, however, the confidence interval for a style weight of a particular market index

- increases with the standard error of the style analysis,
- decreases with the number of returns used in the style analysis, and
- decreases with the “independence” of that market index from the other market indexes used in the analysis (as represented by the unexplained Sharpe style index volatility).

**THE SIMULATION PROCEDURE**

To test the efficacy of the measure derived above, we performed Monte Carlo simulations. We first formulated a “portfolio” of return time series that is a known combination of our market indexes. As such, its “true” style weights are known. If we were to perform a style analysis of this portfolio against our market indexes, we would get estimates equal to the true weights. If we were to compare the portfolio returns with the combination of market indexes, we would have a perfect fit, with \(R^2\) equal to 1 and standard error equal to 0.

If we select some arbitrary value, \(x\), for the standard error, we can create a series of simulated error terms by simply taking random draws from a unit normal distribution and multiplying by \(x\). We can then take this series of simulated error terms and add it to the portfolio return time series for which the true style weights are known. Because of the “noise” we will have added to our portfolio return time series, a repeat of the style analysis will yield new weights that are, of course, estimates of the true weights.

By repeating this procedure many hundreds or thousands of times, we obtained the probability distribution for each of the style weights. The mean values of each style weight are approximately equal to the true style weights. We also observed that the standard deviations of the style weights that we used as our confidence-interval measure are very nearly equal to the values predicted by our derived equations. As would be expected, the predicted and simulated style weights’ means and confidence interval values converged as the number of iterations increased through the Monte Carlo procedure. In a typical example, a simulation of 10,000 trials produces convergence in the fourth or fifth decimal place.

The predictive formula for the confidence interval on style weights overestimates the standard error if the true style weight is very close to either 0 or 1. In those cases, the constraints truncate the estimated values, making the estimated standard deviation upward biased and biasing the style-weight estimate itself. Quantifying the extent of the bias remains problematic. Except for this truncation effect, we can expect the style weight estimates to be approximately normally distributed, just as in a standard multiple regression, so that a traditional \(t\)-statistic can be used.

**Sample Case**

To illustrate how these measures can be calculated in practice, consider the example of the Reich & Tang Equity Fund for the 60-month period from December 12, 1989, through December 31, 1994. We have no special knowledge about the investment style of this fund; we know only that it is a domestic equity mutual fund. For illustrative purposes, assume that the fund has the flexibility to invest in international equities and domestic fixed-income securities. We performed a broad style analysis using the following four indexes: 90-day Treasury bills, the Lehman Aggregate Index, the Russell 3000 Index, and the Morgan Stanley Capital International (MSCI) Europe/Australia/Far East (EAFE) Index.

For the 90-day T-bill index, we determined the Sharpe style index that is composed of the other three indexes. We next calculated the standard
deviation of the active return of 90-day T-bills relative to this Sharpe style index (which we called the unexplained Sharpe style index volatility). This procedure was repeated for the remaining indexes, and the results are shown in Table 1.

Table 1. Unexplained Sharpe Style Index Volatility (percent per month)

<table>
<thead>
<tr>
<th>Index</th>
<th>Unexplained Sharpe Style Index Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>90-day Treasury bills</td>
<td>1.01%</td>
</tr>
<tr>
<td>Lehman Aggregate</td>
<td>0.85</td>
</tr>
<tr>
<td>Russell 3000</td>
<td>2.60</td>
</tr>
<tr>
<td>MSCI EAFE</td>
<td>4.59</td>
</tr>
</tbody>
</table>

We then performed a Sharpe style analysis on the Reich & Tang Fund, using these four market indexes to obtain an active standard deviation of 1.28 percent a month for the fund relative to its Sharpe style index. For each index, the standard deviation of its Sharpe style weight can then be expressed as

\[ \sigma_{w_i} = 1.28 \times \left( \frac{1}{60} \times \sqrt{1 - 3 \times \left( \text{Unexplained Sharpe style index volatility for index } i \right)} \right) \]

because three indexes had nonzero weights. The results of this calculation are shown in Table 2 in the row titled Index Group 1.

Now, suppose that we were told in advance that the fund did not invest in fixed-income securities (other than cash). We would then perform the same analysis but excluding the Lehman Aggregate Index. This calculation would produce the results in Table 2 shown in the row titled Index Group 2.

We can see the dramatic decrease in the T-bill weight’s standard deviation that comes from removing the fixed-income index. This result makes intuitive sense: It is much less likely for T-bill returns to be replicated only by equities than by a combination of equities and fixed-income securities. This example illustrates the importance of limiting the number of indexes used for Sharpe style analysis; advance knowledge of a portfolio’s investment constraints can be quite valuable in this regard.

In a style analysis using only the equity and T-bill indexes, we first separated domestic equities into large and small indexes; the corresponding style weights and standard deviations appear in Table 2 in the Index Group 3 row. We could also introduce a “growth/value” breakdown (Index Group 4) or further extend the size dimension to include mid-cap equities (Index Group 5).

The standard deviations of the Sharpe style weights for the domestic equity indexes increase with the inclusion of more similar indexes (the weights and standard deviations for the other indexes are relatively unaffected). As expected, the

Table 2. Style Weights and Standard Deviation of Style Weights (standard deviations in parentheses)

<table>
<thead>
<tr>
<th>Index Group 1</th>
<th>Lehman Aggregate</th>
<th>90-Day Treasury Bills</th>
<th>MSCI EAFE</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R² = 0.868)</td>
<td>0.0%</td>
<td>10.0%</td>
<td>2.2%</td>
<td>87.8%</td>
<td>(6.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(20.1)</td>
<td>(16.9)</td>
<td>(3.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index Group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R² = 0.868)</td>
<td>10.0%</td>
<td>2.2%</td>
<td></td>
<td>87.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.2)</td>
<td>(3.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index Group 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R² = 0.872)</td>
<td>11.6%</td>
<td>2.9%</td>
<td></td>
<td>70.0</td>
<td>15.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.3)</td>
<td>(3.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index Group 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R² = 0.873)</td>
<td>9.9%</td>
<td>2.7%</td>
<td></td>
<td>41.5</td>
<td>30.5</td>
<td>10.0%</td>
<td>5.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.0)</td>
<td>(3.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index Group 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R² = 0.872)</td>
<td>10.1%</td>
<td>2.4%</td>
<td></td>
<td>29.2</td>
<td>16.3</td>
<td>24.2</td>
<td>17.8</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>(6.0)</td>
<td>(3.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Domestic Equity Index Number Assignments

Index Group 1,2 1 = Russell 3000
Index Group 3 1 = Russell 1000
2 = Russell 2000
Index Group 4 1 = Russell 1000 Value
2 = Russell 1000 Growth
3 = Russell 2000 Value
4 = Russell 2000 Growth
Index Group 5 1 = S&P/BARRA Value
2 = S&P/BARRA Growth
3 = Sharpe/BARRA Medium Value
4 = Sharpe/BARRA Medium Growth
5 = Sharpe/BARRA Small Value
6 = Sharpe/BARRA Small Growth
$R^2$ of the Sharpe style index increases (albeit ever so slightly) with the number of indexes. This result illustrates the need for practitioners to strike a balance between the number of style weights and the likely reliability of each of those weights.

CONCLUSION

Although users of style analysis have used the $R^2$ statistic as a measure of goodness of fit, no measures for the quality of fit for individual style weights have been available. We derived an expression for the approximate confidence interval on style weights. Through extensive Monte Carlo simulation, we verified that the predicted and simulated values converge.

How can this information—that confidence intervals on style weights can, in fact, be approximated—be of use to practitioners? First, we can make statements about the statistical significance of the results. For example, we can now determine if the differences in the style weights of two different portfolios (analyzed against the same market indexes) are different to a statistically significant extent. In recent years, the potential misclassification of mutual funds into objective categories has received considerable attention in academia, regulatory agencies, and the financial press. Clustering of mutual funds into categories based on their style analysis weights is one technique that has already been used in such studies. The availability of explicit confidence intervals on the style weights should improve the robustness of the results of those efforts.

Another use would be to determine whether differences in style weights for the same portfolio over different time periods are statistically significant. Practitioners often use changes in style weights over time as an indication that a portfolio manager has meaningfully shifted his strategy. Without explicit confidence intervals, apparent shifts in strategy may simply be the result of the level of imprecision in the data and the technique.

The confidence interval measures may also be useful in disallowing certain combinations of market indexes. If the indexes to be used are too similar, the confidence intervals will be unacceptably large. Practitioners can either find other market indexes used in the analysis. To isolate the extent to which our estimates for the style weights do not match the true style weights:

Another use would be to determine whether differences in style weights for the same portfolio over different time periods are statistically significant. Practitioners often use changes in style weights over time as an indication that a portfolio manager has meaningfully shifted his strategy. Without explicit confidence intervals, apparent shifts in strategy may simply be the result of the level of imprecision in the data and the technique.

The confidence interval measures may also be useful in disallowing certain combinations of market indexes. If the indexes to be used are too similar, the confidence intervals will be unacceptably large. Practitioners can either find other market indexes to represent the range of possible assets and strategies or disallow the use of style analysis for those situations for which this technique is genuinely unsuited.

These measures also may encourage us to use daily return data in our style analysis in order to obtain a meaningful reduction in the confidence intervals of the resulting style weights.

In some coarse sense, this technique is most appropriate for determining a fund’s effective asset mix when the market indexes will be used to span a wide range of asset behaviors, but the appropriateness of the technique becomes less certain as we work within a single asset class, using style analysis to analyze finer distinctions in investment-management style.

APPENDIX: Approximating the Confidence Interval for Sharpe Style Weights

The return series for the “true” style weight combination of market indexes is defined as

$$ S = \sum w_i r_i , $$

where

- $w_i = \text{true style weight of index } i$
- $r_i = \text{time series of returns on index } i$

The values of $S$ and $w_i$ are unknown; we are using Sharpe style analysis to estimate these values. Now, define

$$ A = R - S, $$

where

$$ R = \text{time series of returns to the subject portfolio} $$

Series $A$ can be thought of as the residual term of a suboptimal constrained regression. Note that the values that compose $A$ are fixed a priori, even though they are unknown to us (because we do not know what the true style weights are). Now, consider the extent to which our estimates for the style weights do not match the true style weights:

$$ \omega_i = w_i + \Delta w_i , $$

where

- $\omega_i = \text{estimate of the true style weight for index } i$
- $\Delta w_i = \text{amount of error in the estimate of style weight for index } i$

As in the three-market-index example in the main text, shifting the style weights will not affect the goodness of fit if the market indexes are linear combinations of one another. Therefore, our fitting process must somehow isolate the portion of the market indexes’ returns that are independent of the other market indexes used in the analysis. To isolate this independent portion of each index, we define

$$ T_i = \sum v_m r_m \text{ (for } m \neq i), $$

and

$$ \sum v_m = 1 \text{ (for } m \neq i), $$

where

- $T_i = \text{returns on the Sharpe style index for market index } i \text{ analyzed against all market indexes exclusive of } i$
\[ \nu_m = \text{style weight on index } m \]
\[ r_m = \text{returns on market index } m \]

In defining the extent to which one market index is a linear combination of the others, the intuitiveness of the explanation is no longer necessary. We thus remove the constraint that the style weights must be in the range of 0 to 1. We can now define
\[ B_i = r_i - T_i, \] (A6)
where
\[ B_i = \text{portion of the returns on index } i \text{ not attributable to the other market indexes, subject to the constraint in Equation A5.} \]

In our later example, we refer to the standard deviation of this return series as "unexplained Sharpe style index volatility."

We now have expressions for both the errors in the style weights \( \Delta w_i \) and the independent portions of the market index behaviors \( B_i \).

Only through the interaction of these two sets of values can the goodness of fit be varied. It can be shown that the operative process in style analysis is to try to minimize the variance of \( R - S - (\Delta w_i B_i) \) or \( A - \Delta w_i B_i \). We can set an objective function, \( Z \), to this expression:
\[ Z = \text{Var} (A - \Delta w_i B_i) \] (A7)
\[ = \sigma^2_A + \Delta w_i \sigma^2_B - 2 \Delta w_i \sigma_A \sigma_B \rho_{AB}, \] (A8)
where
\[ \sigma = \text{sample standard deviation} \]
\[ \rho = \text{sample correlation coefficient} \]

To solve for the minimum of the variance, we set the derivatives of the variance with respect to the style weights equal to zero:
\[ dZ/d\Delta w_i = 2 \Delta w_i \sigma^2_B - 2 \rho_{AB} \sigma_A \sigma_B, \] (A9)
\[ dZ/d\Delta w_i = 0 \text{ (if and only if } \Delta w_i = \rho_{AB} \sigma_A / \sigma_B). \] (A10)

Because the standard deviation for \( \rho \) is approximately 1/\( \sqrt{n-2} \), the standard deviation of \( \Delta w_i \) is approximated by
\[ \sigma_{\Delta w_i} \equiv \sigma_A / (\sigma_B \times \sqrt{n-2}), \] (A11)
where
\[ n = \text{number of data points in the return time series} \]

Because we do not know the true style weights, we do not know \( \sigma_A \). We do know \( \sigma_A \), however—the standard error of our style analysis computation, where
\[ a = R - \sum (w_i + \Delta w_i) r_i. \] (A12)

Because \( a \) has \((n - k)\) degrees of freedom and \( A \) has \((n - 1)\) degrees of freedom, we use the relation
\[ \sigma^2_a = \frac{\sigma^2_A (n - k)}{(n - 1)}. \] (A13)

where
\[ k = \text{number of market indexes with nonzero style weights} \]

Rearranging Equation A13 and substituting back into Equation A11, we arrive at
\[ \sigma_{\Delta w_i} \equiv \frac{\sigma_a}{(\sigma_B \times \sqrt{n - k - 1})}. \] (A14)

REFERENCES


