



# User Guide

18.3. Using PowerPoint Scripts to Create and Update Presentations .....	298
18.3.1. Exporting One StyleADVISOR Workbook to One Office Document.....	298
18.3.2. Exporting Multiple StyleADVISOR Workbooks to One Office Document .....	299
18.3.3. Exporting StyleADVISOR Workbooks to Multiple Office Documents .....	303
18.3.4. Running a StyleADVISOR Export Script File.....	304
<b>19 DEFINITIONS OF ZEPHYR DEFINED UNIVERSES .....</b>	<b>305</b>
<b>20 DYNAMIC TEXT ADVANCED MODE.....</b>	<b>307</b>
20.1. Chart Dependent Functions.....	307
20.2. Non-Chart Dependent Functions .....	314
20.3. Math.....	317
20.4. Queries.....	319
20.5. Operators.....	321
20.6. Objects.....	322
<b>21 STATISTICS AND CALCULATIONS .....</b>	<b>324</b>
21.1. Introduction.....	324
21.2. Return Distribution .....	324
21.2.1. Annualized Return .....	324
21.2.2. Cumulative Return .....	324
21.2.3. Mean .....	324
21.2.4. Variance .....	324
21.2.5. Standard Deviation.....	325
21.2.6. Annualized Standard Deviation .....	325
21.2.7. Skewness.....	325
21.2.8. Kurtosis.....	325
21.2.9. Semi Standard Deviation and Upside Deviation.....	326
21.2.10. Market Line.....	326
21.2.11. Histogram Bin Size .....	326
21.3. Style Analysis .....	327
21.3.1. William F. Sharpe's Returns-Based Style Analysis.....	327
21.3.2. Single Window vs. Rolling Window Style Analysis .....	327
21.3.3. In-Sample vs. Out-Of-Sample.....	328
21.3.4. Zephyr's Style Drift Statistic .....	329
21.3.5. Style Points .....	331
21.3.6. Exponential Weighting .....	332
21.3.7. R-Squared .....	332
21.3.8. Correlation Squared .....	333
21.3.9. Variance Explained.....	333
21.3.10. Adjusted R-Squared.....	334
21.3.11. The R2 Definitions in StyleADVISOR.....	334
21.4. Alpha and Beta.....	335
21.4.1. Beta.....	335
21.4.2. Alpha.....	335
21.5. Excess Return .....	336
21.5.1. Variance and Standard Deviation of Excess Return .....	336
21.5.2. Annualized and Cumulative Excess Return .....	336
21.5.3. Relative Cumulative Excess Return.....	336
21.6. Information Ratio and Related Statistics .....	337
21.6.1. Information Ratio.....	337
21.6.2. Sharpe Ratio.....	337
21.6.3. Treynor Ratio.....	337
21.6.4. T-Statistic.....	337
21.6.5. Significance Level .....	337
21.7. Asset Turnover.....	338

21.8. Best and Worst Statistics .....	339
21.9. Up-Down Statistics .....	340
21.9.1. Number of Up and Down Periods .....	340
21.9.2. Average Up and Down Returns .....	340
21.9.3. Up and Down Capture .....	340
21.10. Batting Average .....	341
21.11. Downside Statistics .....	342
21.11.1. Upside and Downside Deviation (MAR = const) .....	342
21.11.2. Upside and Downside Deviation (MAR = cash) .....	342
21.11.3. Sortino Ratio .....	342
21.12. Drawdown and Runup Statistics .....	343
21.12.1. Pain Index .....	343
21.12.2. Pain Ratio .....	344
21.12.3. Relative Pain Ratio .....	344
21.12.4. Calmar Ratio .....	344
21.12.5. Relative Calmar Ratio .....	344
21.12.6. Omega (MAR = 0%) .....	345
21.12.7. Sharpe Omega (MAR = 0%) .....	346
21.12.8. Gain to Loss Ratio .....	346
21.12.9. Risk-Adjusted Performance .....	346
21.12.10. Arithmetic Risk-Adjusted Performance .....	347
21.12.11. High Water Mark .....	347
21.12.12. Maximum Drawdown .....	348
21.12.13. Longest Drawdown .....	348
21.12.14. Maximum Runup .....	349
21.12.15. Longest Runup .....	349
21.12.16. Maximum Uninterrupted Loss .....	349
21.12.17. Longest Uninterrupted Loss .....	350
21.12.18. Maximum Uninterrupted Gain .....	350
21.12.19. Longest Uninterrupted Gain .....	350
21.13. Ranks and Percentiles .....	351
21.13.1. Percentile Series .....	351
21.13.2. Rank Series .....	352
21.13.3. Volatility of Rank .....	353
21.14. Advanced Parameters .....	354
21.14.1. Market Benchmark: Adjust For Beta .....	354
21.14.2. Market Beta: No Limit .....	354
21.14.3. Annualize Results: Less than one Year .....	354
21.15. Model Selection .....	355
21.15.1. Maximize Adjusted R2 .....	355
21.15.2. Minimize $C_p$ .....	356
21.15.3. Minimize SC .....	356
21.15.4. Minimize AIC .....	356
<b>22 INDEX .....</b>	<b>357</b>

## 22 Statistics and Calculations

---

### 22.1. Introduction

This documentation explains the statistics and calculations used in StyleADVISOR.

### 22.2. Return Distribution

In the following,  $r_1, \dots, r_n$  is a return series, i.e., a sequence of returns for  $n$  time periods.

#### 22.2.1. Annualized Return

The annualized return is the geometric mean of the returns with respect to one year. If we denote by NumYears the number of years covered by the returns, the formula becomes:

$$\text{AnnRtn}(r_1, \dots, r_n) = \left( \prod_{i=1}^n (1 + r_i) \right)^{1/\text{NumYears}} - 1$$

#### 22.2.2. Cumulative Return

The cumulative return is simply the compound return of the series.

$$\text{CumRtn}(r_1, \dots, r_n) = \left( \prod_{i=1}^n (1 + r_i) \right) - 1$$

#### 22.2.3. Mean

The mean, or average, of the return series is the sum of the returns divided by their number.

$$\text{Av}(r_1, \dots, r_n) = \frac{1}{n} \sum_{i=1}^n r_i$$

#### 22.2.4. Variance

Variance can be calculated in two ways:

**Variance** assumes that the returns series is a sample of the population. This is the calculation most commonly used.

If we denote by  $\bar{r}$  the mean of the returns, the variance is:

$$\text{Var}(r_1, \dots, r_n) = \frac{1}{n-1} \sum_{i=1}^n (r_i - \bar{r})^2$$

**Population Variance** assumes that the return series is the population.

$$\text{PVar}(r_1, \dots, r_n) = \frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^2$$

### 22.2.5. Standard Deviation

Standard deviation of return measures the average deviations of a return series from its mean, and is often used as a measure of risk. A large standard deviation implies that there have been large swings in the return series of the manager.

Standard deviation can be calculated in two ways:

**Standard Deviation** assumes that the returns series is a sample of the population. This is the calculation most commonly used. The standard deviation of the return series is the square root of the variance:

$$\text{StdDev}(r_1, \dots, r_n) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (r_i - \bar{r})^2}$$

**Population Standard Deviation** assumes that the return series is the population. Population Standard Deviation is the square root of the population variance:

$$\text{PStdDev}(r_1, \dots, r_n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^2}$$

### 22.2.6. Annualized Standard Deviation

The annualized standard deviation is the standard deviation multiplied by the square root of the number of periods in one year.

$$\text{AnnStdDev}(r_1, \dots, r_n) = \text{StdDev}(r_1, \dots, r_n) * \sqrt{\text{NumPeriodsPerYear}}$$

### 22.2.7. Skewness

Skewness characterizes the degree of asymmetry of a distribution around its mean. Positive skewness indicates a distribution with an asymmetric tail extending toward more positive values. Negative skewness indicates a distribution with an asymmetric tail extending toward more negative values.

$$\text{Skewness}(r_1, \dots, r_n) = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left( \frac{r_i - \bar{r}}{\sigma} \right)^3$$

### 22.2.8. Kurtosis

Kurtosis characterizes the relative peakedness or flatness of a distribution compared with the normal distribution. Positive kurtosis indicates a relatively peaked distribution. Negative kurtosis indicates a relatively flat distribution.

$$\text{Kurtosis}(r_1, \dots, r_n) = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left( \frac{r_i - \bar{r}}{\sigma} \right)^4 - \frac{3(n-1)^2}{(n-1)(n-3)}$$

### 22.2.9. Semi Standard Deviation and Upside Deviation

The semi standard deviation and upside standard deviation differ from the ordinary standard deviation insofar as the sum is restricted to those returns that are less than the mean (semi standard deviation) or more than the mean (upside deviation):

$$\text{SemiStdDev}(r_1, \dots, r_n) = \sqrt{\frac{1}{n-1} \sum_{r < \bar{r}} (r - \bar{r})^2}$$

where  $\bar{r}$  is the mean of  $r_1, \dots, r_n$ , which are the returns BELOW the mean.

$$\text{Upside Deviation}(r_1, \dots, r_n) = \sqrt{\frac{1}{n-1} \sum_{r > \bar{r}} (r - \bar{r})^2}$$

where  $\bar{r}$  is the mean of  $r_1, \dots, r_n$ , which are the returns ABOVE the mean.

To annualize these deviations, one multiplies by the square root of the number of periods in a year, just as in the case of the ordinary standard deviation.

### 22.2.10. Market Line

The Market Line can be plotted on the Manager Risk/Return graph. The capital market line is a straight line that can be drawn in any coordinate system where risk is the x-axis and return is the y-axis. The capital market line is the straight line that connects the point for the cash equivalent and the point for the Market Benchmark in this coordinate system. Any point on this line corresponds to a portfolio consisting of part cash equivalent (possibly shorted) and part Market Benchmark. In other words, the capital market line indicates what one could achieve by mixing cash (long or short) into the Market Benchmark.

If a manager lies to the right of this line, it means that the same return could have been achieved with less risk using a blend of cash and Market Benchmark. If a manager is to the left of this line, it means that any blend of cash and Market Benchmark would have had lower returns than the manager, higher risk than the manager, or both.

### 22.2.11. Histogram Bin Size

The default histogram bin size is estimated using optimal histogram theory as described by Isenman<sup>x</sup>. The formula used is:

$$W = 2 (\text{iqr}) n^{1/3}$$

Where:

W = histogram bin width

iqr = Interquartile range of data in the histogram

n = number of data points

The idea of this formula is to give a good bias versus variance tradeoff in selecting the bin size. The bins should not be too wide as they will represent biased estimates of their data. The bins should also not be too narrow as the variance of the bins will go up. "W" is rounded to one significant figure in order to obtain bin sizes that are human readable. For example, 1.6% would be rounded to 2% and 0.14% would be rounded to 0.1%.

## 22.3. Style Analysis

### 22.3.1. William F. Sharpe's Returns-Based Style Analysis

StyleADVISOR implements returns-based style analysis as set forth by Stanford professor and Nobel Prize winner William F. Sharpe. Returns-based style analysis calculates a Style Benchmark for a manager from the manager's return series and the return series of the indices that are to be used in the Style Benchmark. Thus, we are given:

- a sequence  $M = m_1, m_2, \dots, m_t$  of manager returns over  $t$  periods, and
- $N$  sequences of index returns over  $t$  periods:

$$C_1 = c_{11}, c_{12}, \dots, c_{1t}$$

$$C_2 = c_{21}, c_{22}, \dots, c_{2t}$$

...

$$C_N = c_{N1}, c_{N2}, \dots, c_{Nt}$$

The goal of style analysis is to determine weights  $x_1, x_2, \dots, x_N$  such that the composite series

$$S = x_1 C_1 + x_2 C_2 + \dots + x_N C_N$$

becomes a best fit, or *Style Benchmark*, for the manager series  $M$ . The crucial question is which of the many possible mathematical definitions of "best fit" is to be chosen. William F. Sharpe proposed as the criterion for best fit to *minimize the variance of excess return*, i.e., one should determine the weights  $x_1, x_2, \dots, x_N$  so as to minimize the expression:

$$\text{Var}(M - (x_1 C_1 + x_2 C_2 + \dots + x_N C_N))$$

From the definition of variance, it is clear that the above expression is quadratic in the unknown weights  $x_1, x_2, \dots, x_N$ .

Hence, determining the weights to minimize the expression requires a quadratic optimization. This is a very complex algorithm whose details are beyond this documentation. If you wish to verify StyleADVISOR's calculation independently, you may want to use the "Solver" add-in that is provided with Microsoft Excel.

As a plausibility condition on the weights, the StyleADVISOR program always enforces the constraint that the weights add up to 1, i.e., the quadratic optimization is performed subject to the constraint:

$$x_1 + x_2 + \dots + x_N = 1$$

By default, StyleADVISOR also constrains the weights to be between 0 and 1. However, this default can be changed in the Constraints group on the Parameters tab in the Analysis Parameters window to allow the optimizer to go short on some or all indices.

For a more in-depth discussion of returns-based style analysis, see Hardy<sup>v</sup>, Sharpe (1988)<sup>vi</sup> and Sharpe (1992)<sup>vii</sup>.

### 22.3.2. Single Window vs. Rolling Window Style Analysis

For single window style analysis, the program calculates one set of style weights per manager, using the entire selected date range. In this case, the style map graph shows one point per manager, and the asset allocation graph shows one set of weights per manager.

For rolling window style analysis, the program performs a whole sequence of style analyses for each manager. To this end, a small date range, or *window*, whose width is specified by the user, is placed at the beginning of the selected date range. This window is then moved forward by an increment that is again specified by the user, and a style analysis is performed for each resulting window. This results in a sequence of sets of weights for each manager. When such a sequence of weights is displayed in the style map graph or the asset allocation graph, the date that is associated with each set of weights is the last date inside the respective window.

### 22.3.3. In-Sample vs. Out-Of-Sample

The difference between an in-sample analysis and an out-of-sample analysis lies in the way that the rolling window Style Benchmark is calculated. In the one window case, the Style Benchmark is simply the composite series

$$S = x_1C_1 + x_2C_2 + \dots + x_NC_N$$

where  $C_1, \dots, C_N$  are the index series and  $x_1, \dots, x_N$  is set of corresponding style weights.

In the rolling window case, we have a whole sequence of sets of style weights rather than one set of style weights. When calculating the Style Benchmark, the question is, which set of style weights must be used for which date?

In the in-sample case, the Style Benchmark is calculated as follows: the first window's worth of data is filled using the style weights that correspond to that first window. After that, each set of weights is used for the last date inside the corresponding window. If the rolling window is advanced by  $n$  time periods, each set of weights is repeated  $n$  times to fill the gaps between window end dates.

In the out-of-sample case, each set of weights is used for the date that immediately follows the last date inside the corresponding window. Again, if the rolling window is advanced by  $n$  time periods, each set of weights is repeated  $n$  times to fill the gaps. Note that the in-sample Style Benchmark covers the entire selected date range, while the out-of-sample Style Benchmark does not start until the earliest date that immediately follows the last date inside the first window.

### 22.3.4. Zephyr's Style Drift Statistic

Since the inception of returns-based style analysis and the ensuing development of sophisticated software that brought returns-based style analysis to the masses, investors have used rolling asset allocation graphs and style maps to gain a visual feeling for the style consistency of managers and mutual funds. The opposite of style consistency is style inconsistency or *style drift*. Rolling asset allocation graphs and style maps are good tools for developing an intuitive understanding of a fund's style consistency; however, they do not replace the need for a quantitative measure of style drift.

The Zephyr Style Drift Score measures the variability of style through time. A Zephyr Style Drift Score of 0 indicates perfect style consistency and is equivalent to buying and rebalancing the indices that constitute the style benchmark each period.

The Zephyr Style Drift Score is calculated by measuring the aggregate variability of the individual asset class coefficients.

$$ZSDS = \sqrt{\text{Var} \begin{bmatrix} c_{1,1} \\ \vdots \\ c_{1,T} \end{bmatrix} + \text{Var} \begin{bmatrix} c_{2,1} \\ \vdots \\ c_{2,T} \end{bmatrix} + \dots + \text{Var} \begin{bmatrix} c_{K,1} \\ \vdots \\ c_{K,T} \end{bmatrix}}$$

where

$c_{k,t}$  is the  $k$ th asset class coefficient at time  $t$ .

Unfortunately this formula masks the underlying intuitiveness of the Zephyr Style Drift Score. An alternative calculation of the Zephyr Style Drift Score yields the same answer but reveals more of the intuition.

Start by calculating the average value of the asset class coefficients for time 1 through  $T$ . Let  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_K$  be the average values of the  $K$  asset class coefficients.  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_K$  can be thought of as the *center of gravity* in  $K$ -dimensional space.

$$\begin{aligned} \bar{c}_1 &= \text{Average}\{c_{1,1}, c_{1,2}, \dots, c_{1,T}\} \\ \bar{c}_2 &= \text{Average}\{c_{2,1}, c_{2,2}, \dots, c_{2,T}\} \\ &\vdots \\ \bar{c}_k &= \text{Average}\{c_{k,1}, c_{k,2}, \dots, c_{k,T}\} \end{aligned}$$



Next, subtract the average values of the  $K$  asset class coefficients from the respective style coefficients at time 1 through  $T$ .

$$\begin{pmatrix} (c_{1,1} - \bar{c}_1) & (c_{2,1} - \bar{c}_2) & \cdots & (c_{K,1} - \bar{c}_K) \\ (c_{1,2} - \bar{c}_1) & (c_{2,2} - \bar{c}_2) & \cdots & (c_{K,2} - \bar{c}_K) \\ \vdots & \vdots & \vdots & \vdots \\ (c_{1,T} - \bar{c}_1) & (c_{2,T} - \bar{c}_2) & \cdots & (c_{K,T} - \bar{c}_K) \end{pmatrix}$$

This results in a  $K \times T$  matrix of calculations in which each row represents the component wise distance from the center of gravity for that time period. For each time period, the individual  $K$  distances are squared and the sum is calculated resulting in a  $T \times 1$  vector. Each element of the  $T \times 1$  vector represents the squared Euclidian distance between the asset class coefficients at time  $t$  and the center of gravity.

$$\begin{aligned} (c_{1,1} - \bar{c}_1)^2 + (c_{2,1} - \bar{c}_2)^2 + \cdots + (c_{K,1} - \bar{c}_K)^2 &= Distance_1^2 \\ (c_{1,2} - \bar{c}_1)^2 + (c_{2,2} - \bar{c}_2)^2 + \cdots + (c_{K,2} - \bar{c}_K)^2 &= Distance_2^2 \\ \vdots & \vdots \\ (c_{1,T} - \bar{c}_1)^2 + (c_{2,T} - \bar{c}_2)^2 + \cdots + (c_{K,T} - \bar{c}_K)^2 &= Distance_T^2 \end{aligned}$$

Next, calculate the sample average of the  $T \times 1$  vector of squared Euclidian distances from the center of gravity.

$$Average\ Squared\ Euclidian\ Distance = \frac{\sum_{t=1}^T Distance_t^2}{T - 1}$$

Finally, calculate the square root of the average squared Euclidian distances.

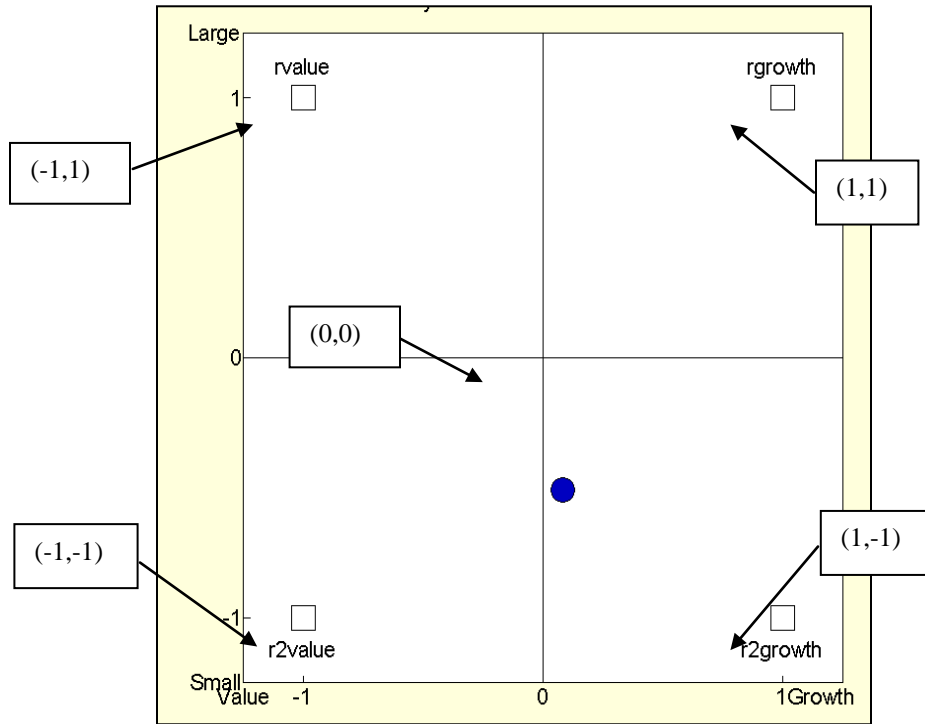
$$ZSDS = \sqrt{Average\ Squared\ Euclidian\ Distance}$$

To summarize, the Zephyr Style Drift Score is the square root of the average squared Euclidian distance of the  $T$  rolling window asset class coefficients from the center of gravity in  $K$ -dimensional space.

Additional mathematical details of the Zephyr Style Drift Score are available in a Zephyr research white paper, “The Style Drift Score: A Quantitative Measure,” which is available upon request (support@styleadvisor.com).

### 22.3.5. Style Points

To understand how StyleADVISOR calculates style points as shown in the style map, it is necessary to understand the concept of a Style Basis. A style basis consists of a set of indices, each of which is assigned to a point in a coordinate system, the so-called Manager Style Graph, or Style Map. The example below shows StyleADVISOR's default style basis. Here, the Style Basis is the Russell Generic Corners, so each of the four Russell style indices are assigned to points in the style map as indicated.



To arrive at the general formula for calculating a style point, let us assume that the style basis contains  $m$  indices. Assume further that  $p_1, \dots, p_m$  are the points that have been assigned to the indices. Note that each  $p_i$  is in fact a two-dimensional point such as  $(-1,1)$ . Finally, let  $w_1, \dots, w_m$  be the weights that correspond to the  $m$  indices. The style point  $p$  for this set of weights is then given by the linear combination

$$p = w_1 p_1 + w_2 p_2 + \dots + w_N p_N$$

### 22.3.6. Exponential Weighting

When a style analysis is performed with exponential weighting, all input series to the analysis (i.e., manager and index series) are multiplied by an exponential function. Thus, if the original series was  $r_1, \dots, r_n$ , the analysis now uses the series

$$r_1 e^{-(N-1)\ln 2/h}, r_2 e^{-(N-2)\ln 2/h}, \dots, r_{N-1} e^{-\ln 2/h}, r_N$$

where  $h$  is the chosen half-life. Although this can be done with both a positive and negative half-life, the most useful choice for style analysis is a positive half-life. In this case, the style analysis places more emphasis on more recent returns. The smaller the half-life, the stronger this effect. For example, a half-life of 12 months means that the returns that lie one year in the past count half as much as the most recent ones, the returns that lie two years in the past count one fourth as much as the most recent ones, and so on. A shorter half-life allows the optimizer to pick up shifts in style more quickly, but also adds more noise to the optimization.

#### Dynamic Exponential Weighting

When a style analysis is performed with exponential weighting, the rolling window analysis employs exponential weighting with dynamically changing half-life. The half-lives for the individual windows are chosen according to a rather complex heuristic that aims at maximizing the overall R2. Dynamic exponential weighting helps strike a balance between capturing changes in style and creating a stable Style Benchmark.

### 22.3.7. R-Squared

Generally speaking, the R-Squared (R2) of a manager versus a benchmark is a measure of how closely related the variance of the manager returns and the variance of the benchmark returns are. If the benchmark is a Style Benchmark, this can be rephrased by saying: the R2 is a measure of how well the variance of the Style Benchmark explains the variance of the manager.

The program uses three different definitions of R2. The next three sections explain the three definitions, followed by an explanation of where the program uses what definition.

### 22.3.8. Correlation Squared

This is the classical statistical method for measuring how closely related the variances of two series are. The correlation is defined as:

$$\frac{\text{(covariance of manager and benchmark)}}{\text{(standard deviation of manager) * (standard deviation of benchmark)}}$$

More explicitly, this is:

$$\frac{\sum_{i=1}^n (m_i - \bar{m}) * (b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (m_i - \bar{m})^2} * \sqrt{\sum_{i=1}^n (b_i - \bar{b})^2}}$$

Covariance of manager and benchmark with the denominator cancelled out.

Standard deviation of manager with the denominator cancelled out

Standard deviation of the benchmark with the denominator cancelled out

where

$n$  = number of returns

$b_i$  = i-th benchmark return

$m_i$  = i-th manager return

$\bar{b}$  = average benchmark return

$\bar{m}$  = average manager return

The correlation squared is of course the *square* of the above expression.

### 22.3.9. Variance Explained

The variance explained is also referred to as Standard R2 in StyleADVISOR. This is usually very close to the correlation squared. To understand what variance explained means, think of a manager and a Style Benchmark. Any variance in the difference between manager and Style Benchmark, i.e., any variance in the excess return of manager over benchmark, represents a failure of the Style Benchmark variance to explain the manager variance. Hence, the quotient of variance of excess return over variance of manager represents the unexplained variance. The variance explained is 1 minus the unexplained variance:

$$\text{Variance Explained} = 1 - \text{Var}(e) / \text{Var}(M)$$

where

$\text{var}(M)$  = variance of manager returns

$\text{var}(e)$  = variance of excess return of manager over benchmark

### 22.3.10. Adjusted R-Squared

This applies only to a Style Benchmark; for Market Benchmarks, it is the same as the standard R2. The Adjusted R2 is based on the Standard R2, but it imposes a penalty for each additional index that is used to build the Style Benchmark.

$$\text{Adjusted R2} = 1 - \left( (m - 1) / (m - n) \right) * \left( \text{var}(e) / \text{var}(M) \right)$$

where

$\text{var}(M)$  = variance of manager returns

$\text{var}(e)$  = variance of excess return of manager over benchmark

$n$  = number of indices used in building the Style Benchmark

$m$  = number of returns

### 22.3.11. The R2 Definitions in StyleADVISOR

The Advanced Parameters tab in the Edit Analysis Parameters window allows the user to select one of the three R2 definitions. The Performance Attribution graph displays the type of R2 that the user has chosen. Similarly, the column “Explained Variance”<sup>1</sup> in the Performance table displays the chosen type of R2. By contrast, the columns labeled “R-Squared” in the Risk Return and Up-Down tables always display correlation squared.

---

<sup>1</sup> The choice of “Explained Variance” as the default column header is perhaps somewhat unfortunate, because it seems to imply that we always display the standard R2 here.

## 22.4. Alpha and Beta

The alpha and beta of a manager vs. a benchmark are obtained by fitting a straight line to the points in a scatter plot of the market returns vs. the manager's returns. Alpha is the intercept of this straight line, while beta is the slope. Hence, if the market returns change by some amount  $x$ , then the manager returns can be expected to change by  $\beta \cdot x$ .

StyleADVISOR includes two Alpha and Beta statistics: Alpha and Beta and Cash-Adjusted Alpha and Beta. The Cash-Adjusted Alpha and Beta, which are the most commonly used, subtract the risk-free rate from the returns of both the manager and benchmark. Otherwise, the formulas are the same.

### 22.4.1. Beta

Beta is defined as 
$$\frac{(\text{covariance of manager and benchmark})}{(\text{variance of benchmark})}$$

More explicitly, this is:

The diagram shows the explicit formula for Beta: 
$$\frac{\sum_{i=1}^n (m_i - \bar{m}) * (b_i - \bar{b})}{\sum_{i=1}^n (b_i - \bar{b})^2}$$
 A callout bubble points to the numerator  $\sum_{i=1}^n (m_i - \bar{m}) * (b_i - \bar{b})$  with the text "Covariance of manager and benchmark with the denominator cancelled out". Another callout bubble points to the denominator  $\sum_{i=1}^n (b_i - \bar{b})^2$  with the text "Variance of benchmark with the denominator cancelled out".

where

$n$  = number of returns

$m_i$  = i-th manager return

$\bar{m}$  = average manager return

$b_i$  = i-th benchmark return

$\bar{b}$  = average benchmark return

Beta is a measure of systematic risk, or the sensitivity of a manager to movements in the benchmark. A beta of 1 implies that you can expect the movement of a manager's return series to match that of the benchmark used to measure beta.

### 22.4.2. Alpha

Alpha is the mean of the excess return of the manager over beta times benchmark:

$$\text{Alpha} = \text{mean}_{(i=1, \dots, m)}(m_i - \beta \cdot b_i)$$

Alpha is a measure of risk (beta)-adjusted return.

### 22.4.3. Alpha\*

Alpha\* is the arithmetic mean of the excess return of the Manager over a risk-adjusted Market Benchmark. For Alpha\*, the market benchmark is replaced with a composite made up from the original Market Benchmark and cash, with the quotient of the manager's and the benchmark's standard deviation as the Market Benchmark's weight and  $(1 - \text{the quotient of the manager's and the benchmark's standard deviation})$  as the weight for cash.

## 22.5. Excess Return

If  $r_1, \dots, r_n$  and  $s_1, \dots, s_n$  are two return series (typically a manager and a benchmark), then the excess return series  $e_1, \dots, e_n$  is defined as

$$e_1, \dots, e_n = r_1 - s_1, \dots, r_n - s_n$$

A positive excess return implies that the manager outperformed the benchmark.

### 22.5.1. Variance and Standard Deviation of Excess Return

The variance and standard deviation of excess return are simply the variance and standard deviation as defined earlier, applied to the excess return series.

$$\text{Var}(e_1, \dots, e_n) = \frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2$$

$$\text{StdDev}(e_1, \dots, e_n) = \sqrt{\text{Var}(e_1, \dots, e_n)}$$

$$\text{AnnStdDev}(e_1, \dots, e_n) = \text{StdDev}(e_1, \dots, e_n) * \sqrt{\text{NumPeriodsPerYear}}$$

Standard deviation of excess returns measures the deviations of an excess returns series from its mean.

### 22.5.2. Annualized and Cumulative Excess Return

It is very important to realize that annualized and cumulative excess return are *not* calculated in the naive way, by taking the annualized or cumulative return of the excess return series. Instead, one must take the annualized and cumulative return of the two original series and *then* form the difference between the two:

$$\text{AnnExRtn} = \text{AnnRtn}(r_1, \dots, r_n) - \text{AnnRtn}(s_1, \dots, s_n)$$

$$\text{CumExRtn} = \text{CumRtn}(r_1, \dots, r_n) - \text{CumRtn}(s_1, \dots, s_n)$$

### 22.5.3. Relative Cumulative Excess Return

The relative cumulative excess return is the ordinary cumulative excess return as a percentage of the benchmark's return.

$$\text{RelCumExRtn} = \frac{\text{CumRtn}(r_1, \dots, r_n) - \text{CumRtn}(s_1, \dots, s_n)}{\text{CumRtn}(s_1, \dots, s_n)}$$

## 22.6. Information Ratio and Related Statistics

### 22.6.1. Information Ratio

The Information Ratio of a manager series vs. a benchmark series is the quotient of the annualized excess return and the annualized standard deviation of excess return.

$$\text{Information Ratio} = (\text{AnnRtn}(r_1, \dots, r_n) - \text{AnnRtn}(s_1, \dots, s_n)) / \text{AnnStdDev}(e_1, \dots, e_n)$$

where

$r_1, \dots, r_n$  = manager return series

$s_1, \dots, s_n$  = benchmark return series

$e_1, \dots, e_n = r_1 - s_1, \dots, r_n - s_n$

The Information Ratio measures the consistency with which a manager beats a benchmark.

### 22.6.2. Sharpe Ratio

The Sharpe Ratio of a manager series is the quotient of the annualized excess return of the manager over the cash equivalent and the annualized standard deviation of the manager return.

$$\text{Sharpe Ratio} = (\text{AnnRtn}(r_1, \dots, r_n) - \text{AnnRtn}(c_1, \dots, c_n)) / \text{AnnStdDev}(r_1, \dots, r_n)$$

where

$r_1, \dots, r_n$  = manager return series

$c_1, \dots, c_n$  = cash equivalent return series

The Sharpe Ratio is a risk-adjusted measure of return which uses standard deviation to represent risk.

### 22.6.3. Zephyr K-Ratio

Zephyr K-Ratio is a ratio that is used in the performance evaluation of a manager relative to its risk. The ratio examines the consistency of return over time. The numerator is the slope of a regression line of the cumulative return on a log scale (“trend return”). The denominator is the standard error of the slope.

$$\text{Zephyr K - Ratio} = \frac{\text{slope of regression line to cumulative return series}}{\text{standard error of slope of regression line to cumulative return series}}$$

$$= \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \frac{[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})]^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}{(n-2) \cdot \sum_{i=1}^n (x_i - \bar{x})^2}}}}$$

where

- $x_i = i$
- $y_i = i$ -th element of the cumulative return series
- $\bar{x}$  = mean of the  $x$ -values
- $\bar{y}$  = mean of the  $y$ -values

### 22.6.4. Treynor Ratio

The Treynor Ratio differs from the Sharpe Ratio insofar as the beta to the Market Benchmark is used as a measure of risk rather than the standard deviation of the manager series.

$$\text{Treynor Ratio} = (\text{AnnRtn}(r_1, \dots, r_n) - \text{AnnRtn}(c_1, \dots, c_n)) / (\text{beta of manager to market})$$

where

$r_1, \dots, r_n$  = manager return series

$c_1, \dots, c_n$  = cash equivalent return series

The Treynor Ratio is a risk-adjusted measure of return which uses beta to represent risk.

### 22.6.5. T-Statistic

The T-Statistic of a manager series vs. a benchmark series is the information ratio, multiplied by the square root of the number of periods in a year.

$$\text{T-Statistic} = (\text{Information Ratio}) * \sqrt{\text{Number of years}}$$

The T-Statistic is used to calculate the significance level.

### 22.6.6. Significance Level

The significance level of a manager series vs. a benchmark series indicates the level of confidence with which the statement “the manager’s annualized excess return over the benchmark is positive” or “the manager’s annualized excess return over the benchmark is negative,” as the case may be, holds true. The significance level is calculated from the T-Statistic using a numerical approximation known as the *incomplete beta function*. For further details, we refer the reader to Press et al.<sup>viii</sup>, Chapter 14.2.

### 22.7. Asset Turnover

Asset turnover is a statistic that applies to any fund (i.e., a combination of assets) over a period of time. Asset turnover is a measurement of how much rebalancing was necessary over time to achieve the actual weightings of the assets in the fund.

To calculate the turnover from one period to the next, we first calculate the weight that each asset would have in the next period if no rebalancing had taken place. From that, we subtract the weight that the respective asset actually had in the next period. The total asset turnover is the sum of the absolute values of these differences divided by two. The division by two is necessary because by adding the turnovers for all assets, we have actually counted every shifted dollar twice: once when it was taken from one asset and then again when it was added to another.

The annualized cumulative asset turnover is the sum of the turnovers for all periods as described above, divided by the number of years.

When calculated for a Style Benchmark, a high Turnover denotes a shift in style.

Asset turnover from period j to period j+1 =

$$\frac{\sum_{i=1}^m \frac{w_{ij}(1+r_{ij})}{\sum_{k=1}^m w_{kj}(1+r_{kj})} - w_{i(j+1)}}{2}$$

weight that asset i would have in period j+1 had there been no rebalancing

where

$m$  = number of assets that make up the fund

$w_{ij}$  = weight of i-th asset for j-th period

$r_{ij}$  = return on i-th asset in j-th period



## 22.8. Best and Worst Statistics

### Best and Worst Period Returns

The best period return for a time window is simply the maximum of the returns inside this window. Similarly, the worst period return for a time window is the minimum of the returns inside this window. Thus, if the return series is  $r_1, \dots, r_n$ , we have:

$$\text{Best Period Return} = \max(r_1, \dots, r_n)$$

$$\text{Worst Period Return} = \min(r_1, \dots, r_n)$$

### Best and Worst One-Year Returns

To calculate the best one-year return for a return series, the program moves a one-year time window along the series and calculates the compound return for each of these windows. The best one-year return is the maximum of the returns thus found. Similarly, the worst one-year return is the minimum of the returns thus found.

Note that best and worst one-year returns do *not* refer to calendar years. Rather, they refer to arbitrarily placed one-year periods.

## 22.9. Up-Down Statistics

### 22.9.1. Number of Up and Down Periods

The number of up periods for a given return series  $r_1, \dots, r_n$ , is the number of positive returns in the series. Similarly, the number of down periods is the number of zero and negative returns in the series.

### 22.9.2. Average Up and Down Returns

To calculate this statistic for a given return series  $r_1, \dots, r_n$ , the program partitions the series in two parts, one made up of the positive returns, the other of the zero and negative returns. The average up and down returns are the respective averages of these two series.

$$\text{Average Up Return} = \frac{1}{l} \sum_{i=1}^l u_i$$

$$\text{Average Down Return} = \frac{1}{m} \sum_{i=1}^m d_i$$

where

$l = \text{NumUpPeriods}$

$u_1, \dots, u_l = \text{the positive returns in } r_1, \dots, r_n$

$m = \text{NumDownPeriods}$

$d_1, \dots, d_m = \text{the non-positive returns in } r_1, \dots, r_n$

### 22.9.3. Up and Down Capture

The up and down capture is a measure of how well a manager was able to replicate or improve on phases of positive benchmark returns, and how badly the manager was affected by phases of negative benchmark returns.

To calculate the up capture, we first form new series from the manager and benchmark series by dropping all time periods where the benchmark return is zero or negative. The up capture is then the quotient of the annualized return of the resulting manager series, divided by the annualized return of the resulting benchmark series. The down capture is calculated analogously.

$$\text{UpCapture} = \frac{\left( \prod_{i=1}^{n_p} 1 + r_i \right)^{1/y} - 1}{\left( \prod_{k=1}^{n_p} 1 + s_k \right)^{1/y} - 1}$$

where

$n_p$  = number of positive benchmark returns

$s_k$  = k-th positive benchmark return

$r_i$  = manager return for the same period as the i-th positive benchmark return

$y$  = number of years, counting periods of positive benchmark returns only

For the *down capture*, just use the non-positive returns instead of the positive ones.

### 22.10. Batting Average

The batting average of the manager is the ratio between the number of periods where the manager outperforms a benchmark and the total number of periods. In other words,

$$\text{Batting Average} = \text{numOutperform} / \text{numTotal}$$

where

numOutperform = number of periods where the return of the manager is greater than the return of the benchmark

numTotal = total number of periods selected

### 22.11. Downside Statistics

#### 22.11.1. Upside and Downside Deviation (MAR = const)

Here, MAR stands for “minimum acceptable return.” To calculate this, we first determine the sum of the squared distances between the returns and the MAR constant, where the sum is restricted to those returns that are more than MAR (upside deviation) or less than the MAR (downside deviation). This sum is then divided by  $n$ , and the square root of the result is taken:

$$\sqrt{\frac{1}{n} \sum_{r < c} (r - c)^2}$$

where

$c$  = MAR

$n$  = total number of returns

This result is annualized in the same way as standard deviation.

You can select the constant used as the MAR on the Advanced Parameters tab of the Analysis Parameters window. See Section 7.8, Setting the Minimum Acceptable Return (MAR).

### 22.11.2. Upside and Downside Deviation (MAR = cash)

This is very similar to the upside and downside deviation with constant MAR. The only difference is that the constant in the formula is replaced by the cash return series.

### 22.11.3. Sortino Ratio

The Sortino Ratio is an analog to the Sharpe Ratio, with the standard deviation replaced by the downside deviation. Accordingly, there are two versions: one uses the downside deviation with constant MAR, the other uses the downside deviation with cash as the MAR.

SortinoRatioConstantMAR =

$$\left( \text{AnnRtn}(r_1, \dots, r_n) - c \right) / \text{DownsideDeviationConstantMAR}(r_1, \dots, r_n)$$

where

$r_1, \dots, r_n$  = manager return series

$c$  = annualized constant MAR

SortinoRatioCashMAR =

$$\left( \text{AnnRtn}(r_1, \dots, r_n) - \text{AnnRtn}(c_1, \dots, c_n) \right) / \text{DownsideDeviationCashMAR}(r_1, \dots, r_n)$$

where

$r_1, \dots, r_n$  = manager return series

$c_1, \dots, c_n$  = cash equivalent return series

You can select the constant used as the MAR on the Advanced Parameters tab of the Analysis Parameters window. See Section 7.8, Setting the Minimum Acceptable Return (MAR).

## 22.12. Drawdown and Runup Statistics

### 22.12.1. Pain Index

The Pain Index and Pain Ratio are statistics that have been developed by Zephyr and are available exclusively in StyleADVISOR. The Pain Index captures much of the information available in Zephyr's drawdown graph. It represents the frequency, the depth, and the width of the manager's drawdowns.

The Pain Index is the area enclosed by the drawdown graph and the zero drawdown line, divided by the length of the time interval.

$$\text{Pain Index} = \frac{\int_{t_1}^{t_2} D(x) dx}{(t_2 - t_1)}$$

where:

$t_1$  = beginning of period analyzed

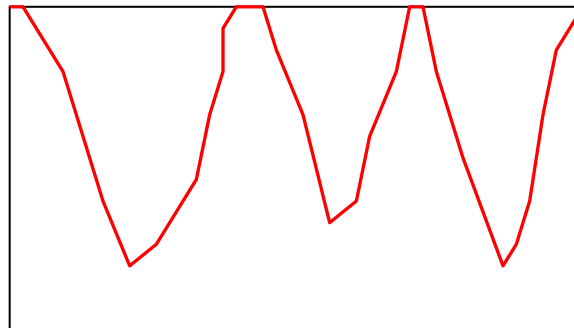
$t_2$  = end of period analyzed

$D(x)$  = drawdown function as shown in StyleADVISOR's drawdown graph

To better understand the Pain Index, consider the following two drawdown graphs:



Manager 1



Manager 2

If one compares these two managers by maximum drawdown, then Manager 1 would be worse than Manager 2. While this is important and useful information, it is clear from the graphs that Manager 2 has in some sense “more” (or “worse”) drawdown than Manager 1. That is of course why we have the drawdown graph in the first place: it contains more information than just the maximum drawdown number. Our pain index is an attempt to capture in one single number as much of the information that is contained in the drawdown graph as possible. Our reasoning is as follows: what makes a manager's drawdown worse is the frequency, the depth, and the width of the downward spikes in the drawdown graph. The one number that simultaneously measures these three things is the area enclosed by the curve in the downward graph: this number increases as the spikes grow more frequent, deeper, or wider. Of course, the area also tends to increase as one considers wider time intervals, hence the division of the area by the length of the time interval. In the example above, Manager 2 would have a much higher pain index than Manager 1, reflecting the fact that Manager 2 has “more” drawdown than Manager 1.

**22.12.2. Pain Ratio**

The Pain Ratio is a risk-return ratio which uses the Pain Index as the measure of risk.

$$\text{Pain Ratio} = (r_{\text{ann}} - c_{\text{ann}}) / p$$

where:

$r_{\text{ann}}$  = annualized manager rate of return

$c_{\text{ann}}$  = annualized risk-free rate of return

$p$  = pain index

**22.12.3. Relative Pain Ratio**

The Relative Pain Ratio is similar to the Pain Ratio, but it uses the geometric difference.

$$\text{Relative Pain Ratio} = (1 + (r_{\text{ann}} - c_{\text{ann}})) / (1 + \text{abs}(p)) - 1$$

where:

$r_{\text{ann}}$  = annualized manager rate of return

$c_{\text{ann}}$  = annualized risk-free rate of return

$p$  = pain index

**22.12.4. Calmar Ratio**

The Calmar Ratio is a risk-return ratio which uses the maximum drawdown as the measure of risk.

$$\text{Calmar Ratio} = r_{\text{ann}} / \text{abs}(m)$$

where:

$r_{\text{ann}}$  = annualized manager rate of return

$m$  = maximum drawdown

**22.12.5. Relative Calmar Ratio**

The Relative Calmar Ratio is similar to the Calmar Ratio, but it uses the geometric difference.

$$\text{Relative Calmar Ratio} = (1 + r_{\text{ann}}) / (1 + \text{abs}(m)) - 1$$

or, equivalently,

$$\text{Relative Calmar Ratio} = (r_{\text{ann}} - \text{abs}(m)) / (1 + \text{abs}(m))$$

where:

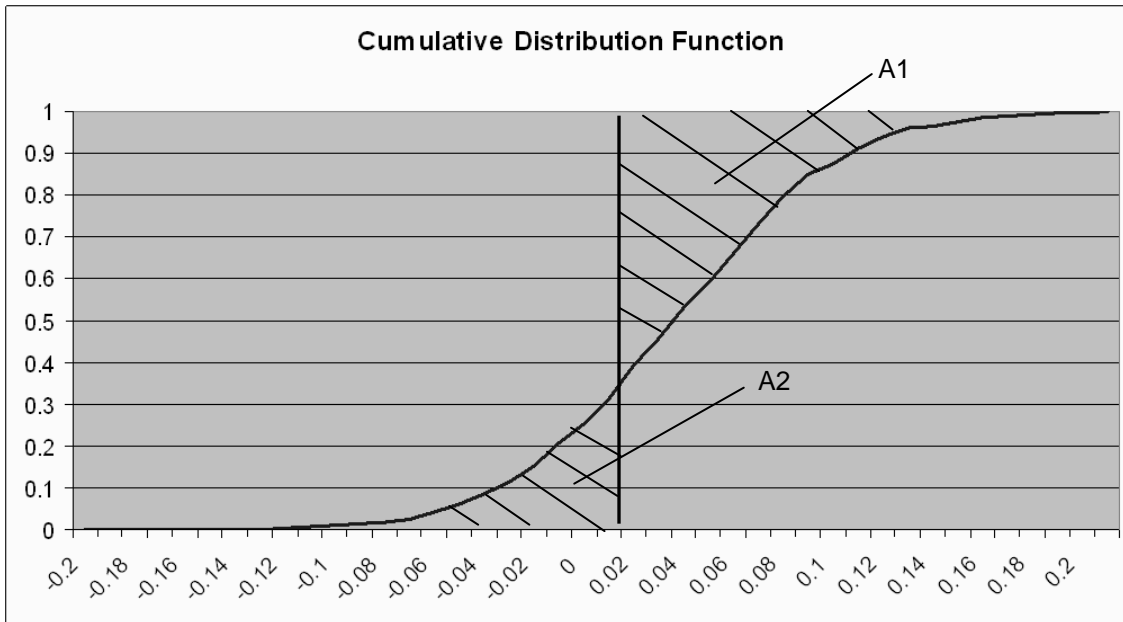
$r_{\text{ann}}$  = annualized manager rate of return

$m$  = maximum drawdown

**22.12.6. Omega (MAR = 0%)**

Omega relative to a given minimal acceptable return (MAR) is the ratio between the price of a European call option written against the investment and the price of a European put option written against the investment, with the strike price being equal to the MAR in both cases.

Alternatively, you can look at Omega graphically with a cumulative distribution function:



The dark curve shows the cumulative probability distribution of the returns. The returns are depicted on the x-axis. For each x on the x-axis, the corresponding y-value is the probability to achieve a return  $\leq x$ . The vertical line is drawn at an arbitrarily selected MAR (minimal acceptable return). Omega is the quotient of the A1 hatched area and the A2 hatched area, that is,

$$\Omega = A1 / A2 = \frac{\int_{M}^{\infty} (1 - F(x)) dx}{\int_{-\infty}^{M} F(x) dx}$$

where:

$M$  = minimal acceptable return

$F(x)$  = cumulative distribution function

For a more detailed discussion of Omega, see “Omega as a Performance Measure” by Kazemi, Schneeweis, and Gupta, online at <http://cisdm.som.umass.edu/research/pdffiles/omega.pdf>.

### 22.12.7. Upside Omega

Upside Omega is the area on a cumulative distribution between the vertical Minimum Acceptable Return (MAR) line and the distribution above the MAR. This is shown as A1 in Subsection 22.12.6, Omega. It is the numerator in the Omega calculation.

$$A1 = \int_M^{\infty} (1 - F(x)) dx$$

where:

$M$  = minimal acceptable return

$F(x)$  = cumulative distribution function

### 22.12.8. Downside Omega

Downside Omega is the area on a cumulative distribution between the vertical Minimum Acceptable Return line (MAR) and the distribution above the MAR. This is shown as A2 in Subsection 22.12.6, Omega. It is the numerator in the Omega calculation.

$$A2 = \int_{-\infty}^M F(x) dx$$

where:

$M$  = minimal acceptable return

$F(x)$  = cumulative distribution function

### 22.12.9. Sharpe Omega (MAR = 0%)

The Sharpe Omega is a variant of the Omega that bears more of a resemblance to the classical Sharpe Ratio. The plain Omega and the Sharpe Omega are rank-equivalent.

$$\text{Sharpe Omega} = (E(r) - M) / P(M)$$

where:

$E(r)$  = expected value of the manager's returns

$M$  = minimal acceptable return

$P(M)$  = price of a European put option with strike price  $M$

The Kazemi/Schneeweis/Gupta paper mentioned above has more details.

### 22.12.10. Gain to Loss Ratio

The Gain to Loss Ratio is the ratio of the average gain in an up period to the average loss in a down period.

$$\text{Gain to Loss Ratio} = \text{abs}(u / d)$$

where:

$u$  = average up return (as seen e.g. in StyleADVISOR's Up-Down Table)

$d$  = average down return (as seen e.g. in StyleADVISOR's Up-Down Table)

### 22.12.11. Risk-Adjusted Performance

Risk-adjusted performance is the annualized rate of return of a blend consisting of the manager's portfolio and cash. The blend is built in such a way that the resulting portfolio has the same standard deviation as the market benchmark. In other words, risk-adjusted performance "normalizes" a manager's standard deviation to equal that of the market benchmark.

$$\text{Risk-Adjusted Performance} = \text{AnnRtn}(\sigma_b/\sigma_m \cdot M + (1 - \sigma_b/\sigma_m) \cdot C)$$

where:

AnnRtn() = annualized rate of return of a return series

$\sigma_b$  = standard deviation of market benchmark

$\sigma_m$  = standard deviation of manager

$M$  = manager return series

$C$  = cash equivalent return series

### 22.12.12. Arithmetic Risk-Adjusted Performance

Arithmetic risk-adjusted performance differs from risk-adjusted performance insofar as the annualization is performed arithmetically, as opposed to the (more common) annualized rate of return, which is calculated geometrically. The arithmetic risk-adjusted performance most closely resembles the original M2, on which this type of risk adjustment is based: the original M2 took the arithmetic mean of the manager-cash blend, but it did not annualize. In other words, you get the original M2 from the formula below if you replace  $a$  with 1.

$$\text{Arithmetic Risk-Adjusted Performance} = (a / n) \cdot \sum (\sigma_b/\sigma_m \cdot m_t + (1 - \sigma_b/\sigma_m) \cdot c_t)$$

where:

$a$  = number of periods per year

$n$  = number of periods analyzed

$\sigma_b$  = standard deviation of market benchmark

$\sigma_m$  = standard deviation of manager

$m_t$  = manager return for period  $t$

$c_t$  = cash equivalent return for period  $t$

### 22.12.13. High Water Mark

The high water mark is often used in the structure of hedge fund incentive fees. With a high water mark, a hedge fund must recover any losses (return to the last high-water mark) incurred before incentive fees can be charged.

- **High Water Mark Date** is the date at which the high water mark was reached, that is, it is the date of the peak of the cumulative return series as seen in StyleADVISOR's Cumulative Return Graph.
- **Under Water Loss** is the loss incurred between the high water mark date and the end date of the period analyzed.
- **Under Water Length** is the length of the time interval that begins at the high water mark and ends at the end of the period analyzed.
- **To High Water Mark** is the percentage gain that the manager needs to regain the peak level (that is, the high water mark) of the cumulative return series. In other words, it is the percentage gain that is needed to offset the under water loss:

$$\text{To High Water Mark} = 1 / (1 + u) - 1$$

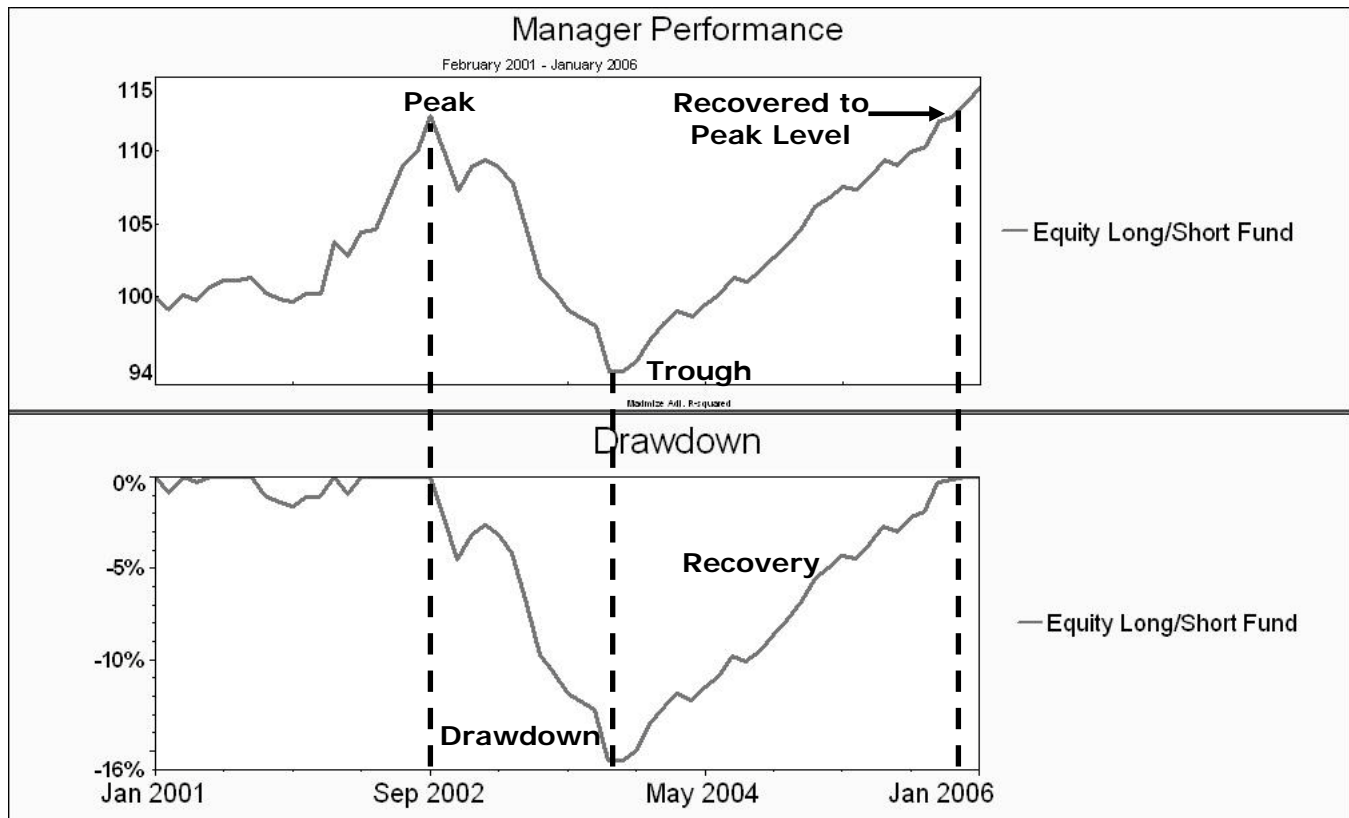
where:

$u$  = under water loss



### 22.12.14. Maximum Drawdown

This is the maximum loss (compounded, not annualized) that the manager incurred during any sub-period of the entire time period. Conceptually, this is the biggest “peak to trough” loss. The calculation looks at all subperiods of the time period in question and calculates the compound return of the manager over that period. The maximum drawdown is the minimum of zero and all these compound returns.



- **Maximum Drawdown Begin Date** is the first date of the sub-period used to calculate the maximum drawdown.
- **Maximum Drawdown End Date** is the last date of the sub-period used to calculate the maximum drawdown.
- **Maximum Drawdown Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the maximum drawdown.
- **Maximum Drawdown Recovery Date** is the date at which the compounded returns regain the peak level that was reached before the maximum drawdown began.
- **Maximum Drawdown Recovery Length** is the number of periods it takes to reach the recovery level from maximum drawdown end date.

### 22.12.15. Longest Drawdown

This is the manager’s compounded return for the longest sub-period of the entire time period in which the manager incurred a loss.

- **Longest Drawdown Begin Date** is the first date of the sub-period used to calculate the longest drawdown.
- **Longest Drawdown End Date** is the last date of the sub-period used to calculate the longest drawdown.
- **Longest Drawdown Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the longest drawdown.
- **Longest Drawdown Recovery Date** is the date at which the compounded returns regain the peak level that was reached before the longest drawdown began.
- **Longest Drawdown Recovery Length** is the number of periods it takes to reach the recovery level from the longest drawdown end date.

### 22.12.16. Maximum Runup

The Maximum Runup is the maximum gain (compounded, not annualized) that the manager incurred during any sub-period of the entire time period. The calculation looks at all sub-periods of the time period in question and calculates the compound return of the manager over that period. The maximum runup is the maximum of these compound returns.

- **Maximum Runup Begin Date** is the first date of the sub-period used to calculate the maximum runup.
- **Maximum Runup End Date** is the last date of the sub-period used to calculate the maximum runup.
- **Maximum Runup Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the maximum runup.
- **Maximum Runup Lost Date** is the date at which the compounded returns fall to the level where the maximum runup began.
- **Maximum Runup Lost Length** is the number of periods between the maximum runup end date and the maximum runup lost date.

### 22.12.17. Longest Runup

This is the manager's compounded return for the longest sub-period of the entire time period in which the manager incurred a gain.

- **Longest Runup Begin Date** is the first date of the sub-period used to calculate the longest runup.
- **Longest Runup End Date** is the last date of the sub-period used to calculate the longest runup.
- **Longest Runup Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the longest runup.
- **Longest Runup Lost Date** is the date at which the compounded returns fall to the level where the longest runup began.
- **Longest Runup Lost Length** is the number of periods between the longest runup end date and the longest runup lost date.

### 22.12.18. Maximum Uninterrupted Loss

The calculation of the Maximum Uninterrupted Loss looks at all sub-periods in which there was a loss for all periods and calculates the compound return of the manager over that period. The maximum uninterrupted loss is the minimum of zero and all these compound returns.

- **Maximum Uninterrupted Loss Begin Date** is the first date of the sub-period used to calculate the maximum uninterrupted loss.
- **Maximum Uninterrupted Loss End Date** is the last date of the sub-period used to calculate the maximum uninterrupted loss.
- **Maximum Uninterrupted Loss Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the maximum uninterrupted loss.
- **Maximum Uninterrupted Loss Recovery Date** is the date at which the compounded returns regain the peak level that was reached before the maximum uninterrupted loss began.
- **Maximum Uninterrupted Loss Recovery Length** is the number of periods it takes to reach the recovery level from maximum uninterrupted loss end date.

### 22.12.19. Longest Uninterrupted Loss

This is the manager's compounded return for longest sub-period in which there was a loss for all periods.

- **Longest Uninterrupted Loss Begin Date** is the first date of the sub-period used to calculate the longest uninterrupted loss.
- **Longest Uninterrupted Loss End Date** is the last date of the sub-period used to calculate the longest uninterrupted loss.
- **Longest Uninterrupted Loss Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the longest uninterrupted loss.
- **Longest Uninterrupted Loss Recovery Date** is the date at which the compounded returns regain the peak level that was reached before the longest uninterrupted loss began.
- **Longest Uninterrupted Loss Recovery Length** is the number of periods it takes to reach the recovery level from the longest uninterrupted loss end date.

### 22.12.20. Maximum Uninterrupted Gain

The calculation of the Maximum Uninterrupted Gain looks at all sub-periods in which there was a gain for all periods and calculates the compound return of the manager over that period. The maximum uninterrupted gain is the maximum of zero and all these compound returns.

- **Maximum Uninterrupted Gain Begin Date** is the first date of the sub-period used to calculate the maximum uninterrupted gain.
- **Maximum Uninterrupted Gain End Date** is the last date of the sub-period used to calculate the maximum uninterrupted gain.
- **Maximum Uninterrupted Gain Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the maximum uninterrupted gain.
- **Maximum Uninterrupted Gain Lost Date** is the date at which the compounded returns fall to the level where the maximum uninterrupted gain began.
- **Maximum Uninterrupted Gain Lost Length** is the number of periods between the maximum uninterrupted gain end date and the maximum uninterrupted gain lost date.

### 22.12.21. Longest Uninterrupted Gain

This is the manager's compounded return for longest sub-period in which there was a gain for all periods.

- **Longest Uninterrupted Gain Begin Date** is the first date of the sub-period used to calculate the longest uninterrupted gain.
- **Longest Uninterrupted Gain End Date** is the last date of the sub-period used to calculate the longest uninterrupted gain.
- **Longest Uninterrupted Gain Length** is the number of periods (days, months or quarters depending on the periodicity of the data) the sub-period used to calculate the longest uninterrupted gain.
- **Longest Uninterrupted Gain Lost Date** is the date at which the compounded returns fall to the level where the longest uninterrupted gain began.
- **Longest Uninterrupted Gain Lost Length** is the number of periods between the longest uninterrupted gain end date and the longest uninterrupted gain lost date.

## 22.13. Ranks and Percentiles

The Manager vs. Universe graph and table allow you to present the data as a percentile or ranked series. In both cases, every manager is ranked according to the statistic being shown on the graph or table. The universe is then separated into percentiles according to those ranks. The number of percentiles used is selected in the Advanced Parameters. By default, the universe is separated into quartiles.

### 22.13.1. Percentile Series

The percentile series plots the actual statistic (return, i.e.) on the vertical axis of the Manager vs. Universe Graph. The universe is represented by colored bands that represent the span of returns for each quartile of the universe.

Here, we are given:

- a universe of  $N$  manager series, each running over  $t$  periods:

$$M_1 = m_{11}, m_{12}, \dots, m_{1t}$$

$$M_2 = m_{21}, m_{22}, \dots, m_{2t}$$

...

$$M_N = m_{N1}, m_{N2}, \dots, m_{Nt}$$

- a percentage  $p$

We are trying to construct a return series

$$S = s_1, s_2, \dots, s_t$$

such that for each  $i$  with  $1 \leq i \leq t$ , the return  $s_i$  is the  $p$ -th percentile of the returns  $m_{1i}, \dots, m_{Ni}$ . To calculate  $s_i$  for any  $i$  with  $1 \leq i \leq t$ , follow these steps:

1. Order the manager returns  $m_{1i}, \dots, m_{Ni}$  in ascending order, say  $r_0, \dots, r_{N-1}$ .
2. Find the position between 0 and  $N - 1$  that the percentage  $p$  corresponds to:

$$\text{pos} = (N - 1) * (100 - p) / 100^r$$

3. Find the two indices that pos lies between, say  $k$  and  $k+1$ :

$$k \leq \text{pos} \leq k + 1$$

4. Interpolate the return between  $r_k$  and  $r_{k+1}$  according to pos:

$$s_i = r_k + (\text{pos} - k) * (r_{k+1} - r_k)^r$$

### 22.13.2. Rank Series

The rank series plots the rank (0% to 100%) on the vertical axis of the Manager vs. Universe Graph. Each of the universe percentiles are represented by equally sized bands.

Here, we are given:

- a universe of  $N$  manager series, each running over  $t$  periods:

$$M_1 = m_{11}, m_{12}, \dots, m_{1t}$$

$$M_2 = m_{21}, m_{22}, \dots, m_{2t}$$

...

$$M_N = m_{N1}, m_{N2}, \dots, m_{Nt}$$

- a single manager series  $s_1, \dots, s_t$  running over the same  $t$  periods.

We are trying to construct a series

$$P = p_1, p_2, \dots, p_t$$

of percentage ranks such that for each  $i$  with  $1 \leq i \leq t$ ,  $p_i$  is the percentage rank of the manager return  $s_i$  within the set of  $m_{1i}, \dots, m_{Ni}$  of corresponding universe manager returns. To calculate  $p_i$  for any  $i$  with  $1 \leq i \leq t$ , follow these steps:

1. Order the manager returns  $m_{1i}, \dots, m_{Ni}$  in ascending order, say  $r_0, \dots, r_{N-1}$ .
2. Find the two universe managers that the manager falls between, say:

$$r_k \leq s_i \leq r_{k+1}$$

3. Interpolate the manager's rank between  $k$  and  $k+1$  according to the returns:

$$\text{rank} = k + (s_i - r_k) / (r_{k+1} - r_k)^T$$

4. Convert rank to a percentage rank with 0 % being the top rank and 100 % being the bottom rank:

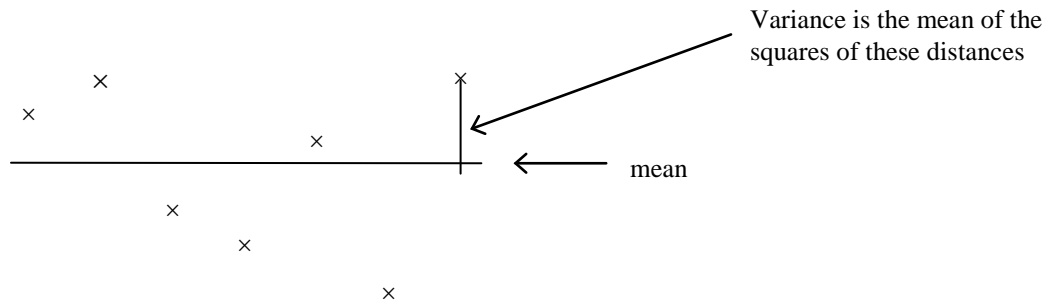
$$p_i = 100 - (100 * \text{rank} / (N - 1))^T$$

### 22.13.3. Volatility of Rank

To understand this calculation, it is necessary to recall that there are two ways of measuring the “instability” of a sequence of numbers:

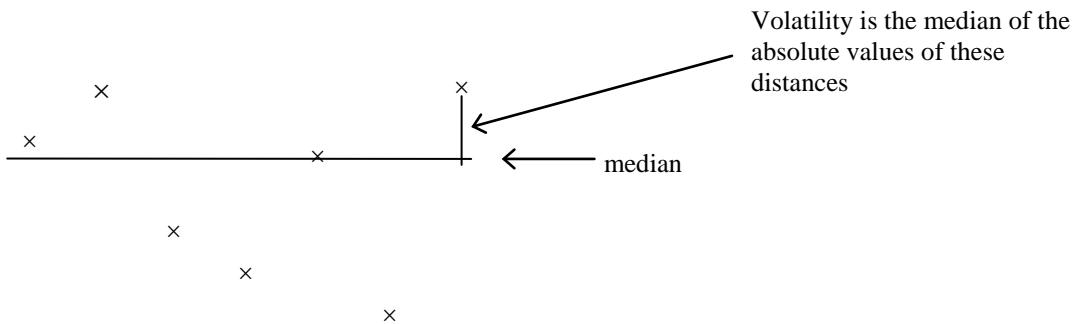
- *Variance and Standard Deviation*

Here, one takes the mean of the numbers, then takes the square of the distance of each individual number to that mean, then looks at the mean of those squared distances. This gives the variance. The standard deviation is simply the square root of the variance.



- *Volatility*

Here, one takes the median of the numbers, then takes the absolute value of the distance of each individual number to that median, then looks at the median of those distances. This is the volatility.



Thus, the two measures differ only insofar as variance uses the mean and the squared distances from the mean, whereas volatility uses the median and the absolute value of the distance from the median.

In the case of the manager ranks in our manager vs. universe calculation, we use the volatility rather than the variance. This is because we are already displaying the median of rank (because we think that this is the appropriate measure of “average” in this case). Obviously, the volatility is a better companion to the median than the standard deviation would be.

$$\text{Volatility of rank} = \text{median}_{(i=1, \dots, m)}(|R_i - \text{median}_{(i=1, \dots, m)}(R_i)|)$$

where

$m$  = number of dates

$R_i$  = manager’s rank number for the  $i$ -th date

## 22.14. Advanced Parameters

The Advanced Parameters tab in the Edit Analysis Parameters window allows the user to set a number of parameters that modify the way the calculations are done. Following is an explanation of these parameters.

### 22.14.1. Market Benchmark: Adjust For Beta

Some people feel that calculating any statistic of a manager vs. a Market Benchmark is inherently unfair because in all likelihood, the risk characteristics of the manager and the benchmark are different. Adjusting the Market Benchmark for beta is one possible way to eliminate this unfairness.

Adjusting for beta means replacing the Market Benchmark with a composite made up from the original Market Benchmark and cash, with beta as the Market Benchmark's weight and  $(1 - \text{beta})$  as the weight for cash.

In StyleADVISOR, the option to adjust the Market Benchmark for beta affects the Scan, Performance Attribution and Performance charts as well as the Style and Performance tables. Generally speaking, the effect is that whenever a statistic vs. Market Benchmark *other than alpha or beta* is calculated, the beta-adjusted Market Benchmark is used instead. The Performance graph is different insofar as the Market Benchmark is not replaced; rather, the beta-adjusted Market Benchmark is offered as an additional choice in the right-click menu.

### 22.14.2. Market Beta: No Limit

This is related to adjusting the Market Benchmark for beta as described above. When the option is selected, which is the default, the true beta of the respective manager vs. the Market Benchmark is used for the adjustment. If the option is unselected, the true beta of the manager vs. the benchmark is used only if it is between 0.0 and 1.0, including 0.0 and 1.0. If it is less than 0.0, 0.0 is used instead. If it is greater than 1.0, 1.0 is used instead.

### 22.14.3. Annualize Results: Less than one Year

If this option is unselected, which is the default, the program does not annualize any returns, standard deviations (including downside risk) or alphas if the underlying time period is less than one year. If the option is selected, all these statistics are annualized regardless of the underlying time period.

## 22.15. Model Selection

StyleADVISOR allows the user to modify the mathematical calculation used to analyze a manager's style. This is done on the Advanced Parameter tab in the Edit Analysis Parameters window under Model Selection. Changing the model affects all style weights and style bases, i.e., it affects the Manager Style and Asset Allocation graphs, as well as everything that involves a Style Benchmark.

When the Standard Model is chosen, StyleADVISOR performs the classical style analysis as set forth by William F. Sharpe. This means that StyleADVISOR determines the style weights (i.e. the asset allocation) so as to minimize the variance of excess return. Looking at the formula for the standard R2 as explained earlier, it is easy to see that minimizing the variance of excess return is equivalent to maximizing the standard R2.

The four other models all have one thing in common: they modify the calculation of the style weights so that there is a penalty for having a larger number of non-zero weights. In other words, there is a reward for dropping indices from the analysis.

To see an example, open the workbook Models.zsa. The page named "Standard Model" shows a standard style analysis of the Acorn fund. Here, StyleADVISOR has determined the weights for the four indices so as to maximize the standard R2. In the left hand style table, we have chosen "Standard" as our definition of R2. Hence, the value 88.87 that is shown in the left hand style table is in fact the value that StyleADVISOR has maximized.

Now contrast this with the results on the page named "Adjusted R2 Model." Here, we have chosen the second one of the five models. The weights for the four indices are different from the first page. As is to be expected, the standard R2 as shown in the left hand style table is lower than before: it has dropped slightly to 88.86. That is because only the weights of the standard style analysis give the maximal standard R2. All other weights result in a lower standard R2. However, since we have chosen a non-standard model, StyleADVISOR has given a reward for dropping one index, namely the Russell 1000 Value index. That reward has outweighed the small decrease of the R2.

The right hand style tables on the two pages of Models.zsa are explained below.

The overall structure of the calculation is the same for all four non-standard models. StyleADVISOR first determines all possible combinations of dropping one or more of the chosen indices. In other words, it looks at all possible subsets of the set of chosen indices. For each of these subsets, it performs a regular style analysis, and it calculates a "utility function" which rewards high standard R2 and penalizes for using a large number of indices. The final result is obtained by picking the subset for which the utility function has the optimal value. (This optimum can be the maximal or minimal value depending on the nature of the utility function). Finally, the indices that were dropped are assigned zero weights.

What distinguishes the four non-standard models is the way in which the reward for dropping indices is determined. In other words, different models use different ways to calculate the utility function as explained in the previous paragraph. It is important to understand that there are no compelling mathematical reasons to prefer one model over the other. The criteria employed by the different utility functions are contentious even among mathematicians. The choice of a particular model is therefore highly discretionary.

For a deeper mathematical discussion of the four models, we refer the reader to Judge et al.<sup>ix</sup>, Sections 20.4.1-20.4.3.

### 22.15.1. Maximize Adjusted R2

In this model, the utility function is the adjusted R2, i.e. StyleADVISOR chooses the subset of indices that gives the highest adjusted R2. Recall that the adjusted R2 is defined as follows:

$$\text{Adjusted R2} = 1 - ((m - 1) / (m - n)) * (\text{Var}(e) / \text{Var}(M))$$

where

$\text{Var}(M)$  = variance of manager returns

$\text{Var}(e)$  = variance of excess return of manager over benchmark, using current subset

$n$  = number of indices in current subset

$m$  = number of returns

Again, the workbook Models.zsa provides an example. On the page named "Standard Model," the adjusted R2 as seen in the right hand style table is 88.68. On the page named "Adjusted R2 Model," the adjusted R2 has increased to 88.74. That is why StyleADVISOR chose to drop the Russell 1000 Value index in this model.



**22.15.2. Minimize  $C_p$** 

This is also known as Mallows's  $C_p$ -criterion. In this model, the utility function is given by:

$$C_p - n = (m - n_i) * \text{Var}(e) + (n - m) * \text{Var}(e_i)$$

where

$n_i$  = total number of indices

$\text{Var}(e_i)$  = variance of excess return of manager over benchmark, using all indices

$n$  = number of indices in current subset

$\text{Var}(e)$  = variance of excess return of manager over benchmark, using current subset

$m$  = number of returns

Mallows's criterion calls for minimizing the utility function, i.e. StyleADVISOR chooses the subset of indices where  $C_p - n$  assumes its minimal value.

**22.15.3. Minimize SC**

This is the Schwarz Criterion. In this model, the utility function is given by:

$$SC = m * \log(\text{Var}(e) / m) + n * \log(m)$$

where

$n$  = number of indices in current subset

$\text{Var}(e)$  = variance of excess return of manager over benchmark, using current subset

$m$  = number of returns

In this model, StyleADVISOR chooses the subset of indices where SC assumes its minimal value.

**22.15.4. Minimize AIC**

This is the Akaike Information Criterion. In this model, the utility function is given by:

$$AIC = m * \log(\text{Var}(e) / m) + 2 * n$$

where

$n$  = number of indices in current subset

$\text{Var}(e)$  = variance of excess return of manager over benchmark, using current subset

$m$  = number of returns

Again, StyleADVISOR chooses the subset of indices where the utility function assumes its minimal value.