BNY MELLON℠
ASSET MANAGEMENT

Tactical Equity Size Rotation: A Systematic Approach For Enhancing Alpha

By

Kenneth A. Barker
Senior Vice President and
Director of Quantitative Analysis
and Research

Mellon Equity

Summary

Over history the "size premium" between large and small capitalization stocks has alternated back and forth, even though it is not clear exactly what phenomena underlie such variation. But the intervals when either small or large capitalization stocks outperform tend to persist. So significant value-added is attainable for an investor with the capability to accurately shift between small and large capitalization stocks on a timely basis.

Using the Russell 1000 Index as a proxy for large capitalization stocks and the Russell 2000 Index for small stocks, an investor with perfect information could have produced 600 basis points of annualized value-added since 1983 by shifting only ±10% of her portfolio on a weekly basis to the out-performing asset versus a 90/10 large/small blended benchmark, before transaction costs. This perfect information strategy would require a shift about every two weeks on average.

This paper presents a methodology for building a model to provide a "size timing" signal. The results are quite robust, with the resulting model being correct more than 75% of the time over the past 21 years. A simulated track record shows annualized outperformance in excess of 100 basis points versus a static 90/10 blend of large vs. small cap stocks, whether measured using the S&P 500 and S&P 600 benchmarks or the Russell 1000 and 2000 (assuming a ±10% shift). Both produce information ratios greater than 1.0 after transaction costs by performing an asset shift about every 30 weeks. Detailed results for the Russell case are presented.

The Rationale

Any underlying rationale for such a model must be grounded in economics. Mellon Equity has a long history of econometric modeling with its Tactical Asset Allocation process, which shifts between large capitalization stocks and intermediate term bonds in order to provide value added relative to a static mix. In order to facilitate rapid development of a "Size TAA" process, much of the data, software and other infrastructure from Mellon Equity's traditional TAA modeling process were utilized.

Based on a client request, we investigated the use of Mellon Equity's stock valuation model (called the EVR, short for Equity Valuation Report) as a tool for selecting between small and large stocks. The EVR produces an alpha (expected excess return) for every stock in Mellon Equity's 3,500 stock universe. Initially, we tested to see if the aggregate alpha of a basket of small cap stocks relative to a basket of large cap stocks could be used as a size-timing signal. Preliminary results showed that such a signal held promise.

Initially published June 2006

As of July 1, 2007, Mellon Financial Corporation and The Bank of New York Company, Inc. merged into a newly created entity, The Bank of New York Mellon Corporation. Accordingly, the information in this publication relates to the respective predecessor company.

So we decided to incorporate this new alpha-spread variable as an additional input to our stock/bond TAA model. And, as described in more detail on the following pages, we revised the dependent, or response, variable of the model to be the three-month forward-looking spread in return (the size premium) between the same large and small capitalization baskets of stocks.

The Universe

The definition of large and small cap has evolved over the past 20 years. Also, many of today's small cap benchmarks did not exist prior to the 1990's. So we decided to construct our own large and small cap universes for estimating this model using Mellon Equity's historic equity database. The primary advantage of this database, since it was constructed real-time, is that it suffers no survivorship or look-ahead biases. The estimation subset of this universe contains on average 1,650 stocks, and was refreshed quarterly over most of the past 20 years.

We concluded that the largest 400 companies from our database would be the best proxy for large cap. And for small cap we chose those stocks that are below 750 in rank order on market capitalization from the estimation subset, which works out to be the smallest 900 stocks. By choosing large/small cap universes that are not adjacent, we hoped to obtain a clearer size-timing signal.

After constructing both universes weekly for the past 21 years, we then computed the corresponding weekly returns for each. We then examined the correlation of these returns vs. the S&P 500 and the S&P 600 Small-Cap benchmarks. The weekly returns of the Mellon Equity "Large Cap 400" are 99% correlated with that of the S&P 500 over the time period from 1989-present. For the Mellon Equity "Small Cap 900," the correlation is 95% vs. the S&P 600 Small Cap Index for the same time period. So we feel comfortable that we have developed large and small cap universes that are representative of these asset classes. As a result, we decided to use the Mellon Equity Large Cap 400 and Small Cap 900 as the baskets for computing both the weekly size premium and the alpha spread over the entire back test period from early 1982 to present.

The Modeling Methodology

Mellon Equity's preferred method for macro-economic modeling is to use the non-linear framework called neural networks. This approach is often misunderstood, as much of the terminology harkens back to its initial development, which occurred in the health and social sciences. So terms such as "training," "machine learning," etc. give this method of model estimation an almost anthropomorphic feel. But mathematically, the technique is just one way of performing a non-linear regression. It has certain advantages over some non-linear methods, not the least being the availability of high quality, off-the-shelf software for performing the analysis. Moreover, it does not force the analyst to presuppose a specific form (such as quadratic) for the solution. The downside is that it requires many more observations than most other non-linear modeling approaches.

The Data

In any modeling attempt such as this we must first ensure that the data is meaningful and of high quality. For the "response" or dependent variable (the one which we are trying to explain), we will look at the "size premium" between large and small cap stocks discussed above. Specifically, we define this to be:

Size Premium: The difference in cumulative return over the next 13 weeks between the Large Cap 400 and the Small Cap 900 (as described above under *The Universe*).

Using a 13-week forward-looking return implies the intrinsic investment horizon of the model is quarterly. We believe that this definition is appropriate for an institutional investor who traditionally monitors performance on a quarterly basis.

As explanatory, or "independent" variables, we selected those used in the Mellon Equity Tactical Asset Allocation Model in addition to the new EVR Alpha Spread variable described above. These variables, while macro-economic in nature, are not traditional government economic statistics like GDP, unemployment or inflation rates. All of these types of statistics are subject to estimation error, reporting lags, revisions and other sources of noise. Mellon Equity's approach is to instead use only market-priced economic variables like interest rates, commodity prices, exchange rates, etc. In addition to incorporating timely market information, these types of variables implicitly capture consensus expectations regarding inflation, economic outlook and market sentiment. We chose these variables as inputs to our "size rotation" model because they have shown their worth in a stock/bond valuation framework, and as such are fundamentally important for explaining relative valuations of domestic assets.

The frequency of all data is weekly, and runs from 1982 to present, yielding over 1,000 observations for model estimation.

Measures of Relative Value of Stocks

- Interest rate adjusted P/E ratio relative to its moving average
- Overbought/oversold indicator based on a proprietary "fair-value" model
- Proprietary expected return (EVR alpha) differential between large cap & small cap stocks

Measures of Relative Performance and Economic Strength

- Current versus historical equity premium over bonds
- Commodity futures prices relative to its moving average
- U.S. Dollar exchange rates relative to its moving average
- Mortgage yield spread relative to its moving average
- Corporate bond credit quality spread relative to its moving average

Interest Rate Environment

- Steepness of the term structure of interest rates relative to its moving average
- 3-Month Treasury bill rate

The Model Setup

One of the hallmarks of using neural networks for non-linear modeling is that it is not always a deterministic process. The set of data used for estimating the model is broken down into three components: training, test and validation. The training component is the data that the modeling software uses for estimating the relationships and interactions between the inputs and the output of the model. This is referred to as the "in sample" data. The test component is used during training as a pseudo out-of-sample set to check for over fitting. If the model fit degrades significantly between the training and test set, then over fitting of the training data has most likely occurred. So the modeling software alternates between the training and test data until an appropriate, generalizable fit is attained. The third component, validation, is data the modeling process never sees. It is set aside for true out-of-sample checking of the final model.

The method the analyst uses for creating these three subsets of the data can impact the resulting model. In order to mitigate this effect, Mellon Equity uses an approach referred to as "bagging" (bootstrap aggregation.) What this means is that we train multiple models, where each one has access to slightly different subsets of the data for training and testing. Then the output signals of the models are averaged, in the belief that this average is more representative of the "true" underlying relationships than any of the individual models. We segregate the training/test/validation data in a round-robin framework, where every week of history is out-of-sample (part of the validation set) for at least three of the models. Then the last year of data is incorporated into the validation set for all the models. In this way, we have an average out-of-sample signal for every point in history.

Another advantage of bagging is that we can also look at the agreement between the various models. If all the models have similar structure and produce very similar results, then we have higher confidence that we have detected stable relationships in the data. But if the various models are all over the map, then there's a higher probability that we have instead picked up on unstable or spurious relationships.

Modeling Results

We trained multiple models using the setup described on the previous pages. The results are in Table I, at right.

As indicated in Table I, we trained 30 models using the round-robin approach previously described. All the results displayed are for the validation (out-of-sample) set only. The column labeled Correl. is the correlation of the output signal of each model vs. the actual realized size premium. The next column is the R-Square of each model. The column labeled Records is the number of out-of-sample observations for each model. Next, the column labeled Input shows how many of the 10 input variables were utilized by each model (the neural network approach has the ability to disregard inputs deemed non-explanatory.) Some of the models show more than 10 inputs. This is because a preprocessing stage produces multiple transformations of each input variable, and the neural net then has the ability to select more than one transformation of the same input. The last column labeled Hidden shows how many "nodes" occur in the "hidden" layer of the neural net. This number is a good indicator of how many interactions were detected between the inputs. The general agreement between all 30 models, as evidenced by small standard deviation of the correlations relative to their average, is quite encouraging.

Table I: Tactical Equity Size Rotation Model

Model Id	Correl.	R-Square	Records	Input	Hidden
V01	0.33	0.11	163	12	7
V02	0.33	0.11	159	14	13
V03	0.18	0.03	156	8	0
V04	0.25	0.07	156	9	1
V05	0.22	0.05	156	7	0
V06	0.39	0.15	156	9	0
V07	0.37	0.14	156	11	0
V08	0.48	0.23	156	9	15
V09	0.50	0.25	156	9	0
V10	0.22	0.05	156	10	0
V11	0.36	0.13	156	9	9
V12	0.37	0.14	156	8	1
V13	0.28	0.08	156	10	4
V14	0.45	0.21	156	7	1
V15	0.40	0.16	156	12	0
V16	0.35	0.12	156	10	0
V17	0.33	0.11	156	9	3
V18	0.31	0.10	156	7	2
V19	0.33	0.11	156	12	0
V20	0.36	0.13	156	8	1
V21	0.42	0.17	156	12	4
V22	0.41	0.17	156	12	0
V23	0.45	0.20	156	8	0
V24	0.34	0.11	156	9	4
V25	0.16	0.03	160	13	2
V26	0.07	0.01	165	6	0
V27	0.36	0.13	169	11	3
V28	0.54	0.29	169	8	4
V29	0.60	0.35	169	12	9
V30	0.53	0.28	167	16	3
Average	0.36	0.14	158	10	3
Median	0.36	0.13	156	9	1
Stdev	0.12	0.08			
u/sd	3.07	1.72			
Minimum	0.07		156	6	0
Maximum	0.60		169	16	15

Source: Mellon Equity Associates, LLP

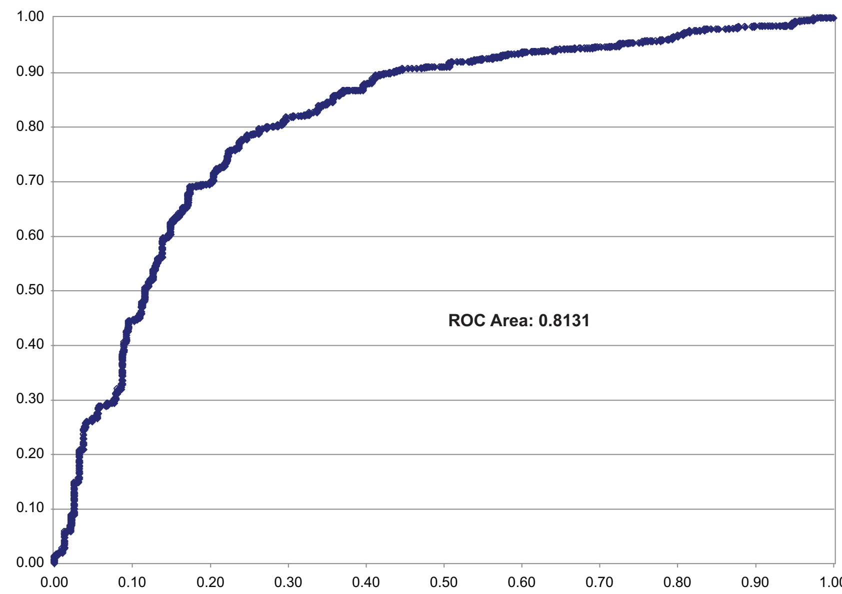
Quality of the Model

Aside from correlation and R-Square, there are other methods for quantifying the quality of the resulting model. One is called the ROC chart (short for Receiver Operating Characteristics.) ROC charts were originally developed in the early days of Radar to gauge the quality of radar operators. It is a useful non-parametric way of comparing the quality of two signals, regardless of whether they are produced by a computer model or a person. The approach is to first sort the signals from the most positive to the most negative. In our case, the more positive the signal, the more we expect large cap stocks to outperform small caps. Then, starting at the origin, we plot the corresponding premia (outcomes.) If the premium is positive (large cap stocks did well), we plot the point by moving upwards. If the premium is negative, we plot the point by moving to the right. Both axes are then rescaled to 1.0 by dividing each by its corresponding total premium. So all ROC charts begin at the origin (lower left corner), and terminate at the (1,1) point (upper right corner.) See Figure I on the following page.

In the case of a perfect model, we would expect the line of the graph to first rise vertically (all the positive premia occur while the signal is positive) to 1.00. Next, as all the negative premia occur, the line would move horizontally across to the upper right corner. The area under this curve would be 1.0, the size of the entire box. In the case of a random, or uninformed model, the chance of a correct outcome is 50-50, so positive or negative premia occur with equal probability. This will generate a line sloping at a 45˚ angle from the lower left corner to the upper right corner, yielding an area under the curve of 0.50. So, the closer the area of a model in question is to 1.0, the better the model. And the closer its area is to 0.50, the worse. In the case of our Tactical Equity Size Rotation (TESR) model described above, the area is 0.81, which is quite strong.

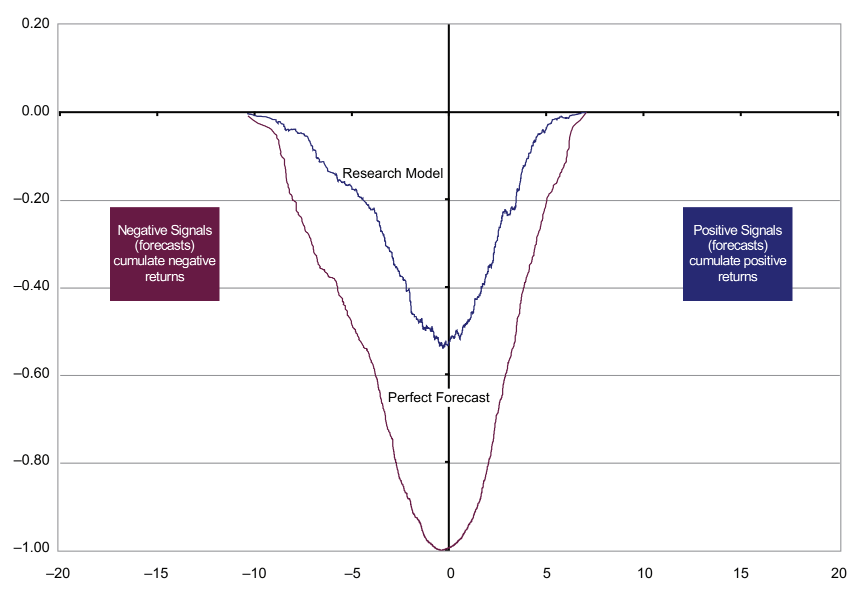
Another approach is to construct what we whimsically refer to as a "bowl" chart. This is a similar approach to the ROC chart, but in this case we first sort the signals from lowest to highest. Then we plot these signals on the X-axis, and cumulate the corresponding premia on the Y-axis. So as we move from left to right, we first have negative signals, and expect to have negative premia (small cap stocks doing better than large cap.) So our cumulative plot should be heading downward. As the X-axis passes through the origin our signals turn positive, and so the "bowl" should turn up as we experience positive premia. In this case, a random model would just be a straight line that lies along the X-axis, as positive and negative premia would occur with equal probability. We also plot the perfect model, by using the actual realized premia as the ranking variable. Then we rescale both plots so the perfect case drops to –1.0. The area of the "bowl" of the model in question vs. the area of the perfect model is a goodness-of-fit measure signifying how much of the available performance is being captured. The shape of the bowl is also instructive, for it provides insight as to regions where the model's signal is more or less predictive. See Figure II, at right.

Figure I: Out of Sample Rolling Test ROC Chart 19820104 – 20030526



Source: Mellon Equity Associates, LLP

Figure II: Value Added Bowl Chart



Source: Mellon Equity Associates, LLP

Implementation

The final step of the process is implementation – how to take the output of the model and translate that into investment decisions. This requires a set of decision rules for implementing the signals. To the right, are some considerations that are important:

- Because making a shift entails transaction costs, we want to ensure we don't make unnecessary moves.
- Depending on the volatility of the signal, we may want to limit overall turnover so we are not shifting excessively.
- The possibility of a false signal is always present, so it might be prudent to wait for confirmation of a signal before making a move.
- The shift from the maximum to the minimum large cap position can be made in steps so as to reduce transaction cost while reducing the chance of responding to a false or transient signal.

The methodology Mellon Equity has used successfully for our TAA process is a combination of several of these approaches. So as our first choice we adopted the TAA decision rules for TESR:

- We employ a "Schmidt Trigger" to filter small movements of the signal near the decision point. Rather than have a single decision threshold (say zero) for indicating when to overweight large cap vs. small cap, we use double thresholds of –1.0 and +1.0. The signal must drop below –1.0 to constitute a small cap signal, and must return above +1.0 to signal large cap. This is similar to the way the thermostat on your furnace works at home. If you set the temperature at 70°, the furnace won't kick on until the temperature drops to about 68°. And then the furnace will run until the temperature rises to near 72° before shutting off. This prevents the furnace from shutting off and on for very brief periods of time as the temperature crosses 70°.
- We use a voting rule: a signal is not considered valid until three out of the past five weekly signals indicate a shift from the current position. So assume for a minute that we have been at our maximum large cap position for some time, and the model has been producing positive signals. When the first negative signal (below –1.0, indicating a shift towards small cap) is received, we do nothing. If two confirming negative signals are repeated within the next four weeks, we shift.
- Shifts from the large to small cap position are implemented in two steps. Following the above example, after the first valid signal is confirmed, we first shift from large cap to the neutral position. Then we enforce a wait in that state (or any state for that matter) for at least three weeks. This rule ensures that we never shift more than once per month. If after three weeks three out of the last five signals still signal a shift to small cap, we complete the second step of the shift to the maximum small cap position. If not, we remain in neutral until a confirmed valid signal is received for either direction.

Although many possible decision rules could be chosen, those described above have served us well in our TAA framework, and so were chosen as the initial approach to use for TESR.

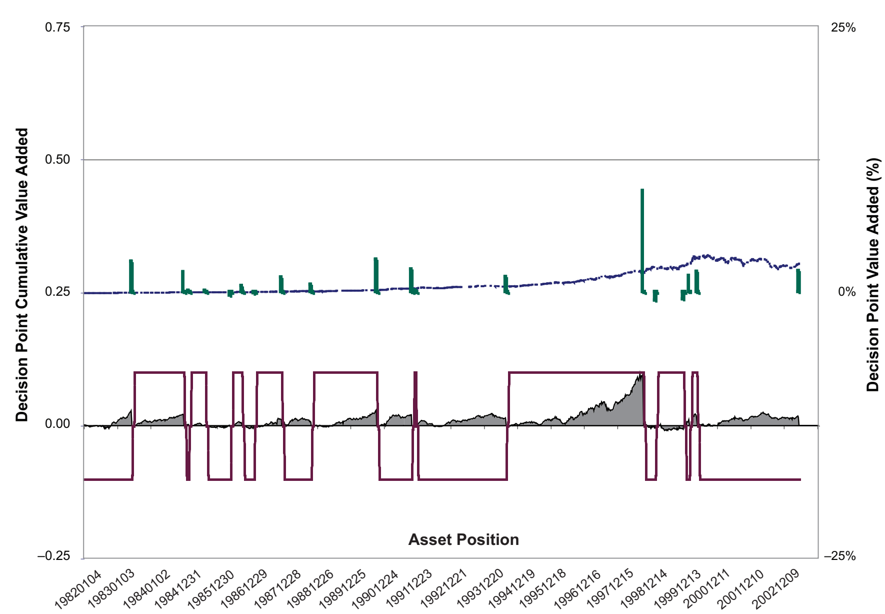
Portfolio Results

Coupled with a set of decision rules, we can now use the model to simulate an historic track record. Although the model has been "trained" using the size premium of the custom size baskets discussed earlier, real-world implementation (using a futures overlay strategy, for example) would require the use of traditional indices like the S&P 500 and 600 Small Cap, or the Russell 1000 and 2000. Using the historic weekly returns of these Russell benchmarks we have constructed a simulated track record.

Using this approach, the results are summarized in Figure III at right. The oscillating red line indicates the decisions. When the red line is high, we are at our maximum large cap allocation, and when it is low, we are at the maximum small cap position. For the purpose of this simulation, we have assumed our neutral position is 90/10 large vs. small cap, with a +/– 10% shift. So our large cap allocation would move from as high as 100% to as low as 80% of the total portfolio.

The gray areas show the cumulative valued added after each shift relative to the neutral blend. Once the next shift is made, this total value added is saved as the green bar. Looking at the green bars, we can see how many decisions were made, how often, and how many of them paid off. The dark blue line is the cumulative value added of all the decisions over the entire 21-year period.

Figure III: Tactical Equity Size Rotation Decision Return Analysis



Source: Mellon Equity Associates, LLP

There are two points of note on this chart:

- The infrequency and small size of the downward green bars. This indicates that the model is not often wrong, and when it is, the cost is not high.
- The model adds value in both the large cap and small cap states. So there is no bias in the model towards either of the two asset classes.

Final Summary of Results

As a final summary of the model and back test results, see Table II below. The box on the upper left indicates the distribution of the raw signals from the model. They are evenly distributed between large and small cap over the 21-year period. The next box on the left shows what fraction of time was spent in each position using the decision rules. Again, they are fairly evenly distributed, spending about half of the time in either the maximum large cap or maximum small cap position. The third box down on the left summarizes how many decisions were made, their average duration, and the annual turnover.

On the right side, the top box shows the breakdown of the 37 decisions that were made over the 21-year period. The next box below shows the decision results as percents of the total number of decisions. Importantly, the distribution of correct/incorrect is comparable for both large cap and small cap decisions, again demonstrating lack of bias. The third box down on right shows the percent of the value-added depending on the type of decision. The model accrues more of the value-added while at the maximum large cap position. This is influenced by the long period during the 1990's when large cap stocks significantly outperformed small caps.

Table II: Tactical Equity Size Rotation: Summary Model Results

August 15, 2003
RESEARCH MODEL

Signal	Weeks	%
Small Cap Signal	427	38%
Neutral Signal	249	22%
Large Cap Signal	441	39%

Decision Position	Weeks	%
Large Cap	519	46%
Neutral	54	5%
Small Cap	544	49%
	1117	100%

Turnover	Count
Total Decisions	37
Weeks/Decision	30
Decisions/Year	1.72
Turnover (Round Trip)	0.43

Decision Rules

3-out-of-5 / Confirming Signal

Steps Per Move	**2**
Trigger +/-	**1.00**
Minimum Holdtime Weeks	**3**
Round Trip Transaction Cost	**0.40**
Raw Signal	

Decision Results Count

Position	Correct	Incorrect	
Large Cap	8	1	9
Neutral			18
Small Cap	7	3	10
Total	15	4	37

Decision Results Percent

Position	Correct	Incorrect	
Large Cap	42%	5%	47%
Neutral			
Small Cap	37%	16%	53%
Total	79%	21%	100%

Decision Results Payoff (Additive)

Position	Correct	Incorrect	
Large Cap	18%	-1%	18%
Small Cap	11%	-1%	9%
Total	29%	-2%	27%

	Large Cap	Small Cap	Model	Benchmark	Excess
			90 +/-10	**90 / 10**	(less transcost)
Annualizd Return	13.1%	10.7%	13.9%	12.9%	1.01%
Annualizd Standard Deviation of Weekly Returns	15.9%	17.1%	15.8%	15.8%	0.96%
Sharp Ratios	0.82	0.62	0.88	0.82	
Information Ratio					1.05
Largest Upside Run					9.5%
Largest Downside Run					-0.8%

Source: Mellon Equity Associates, LLP

Further Implementation Considerations

Actual implementation in a client portfolio could be approached in several ways. Simulation results show that comparable results are obtained whether using either S&P or Russell benchmarked assets for large and small cap. Some of the implementation options available include the following:

- Futures overlay strategy – the underlying physical assets could be either passively or actively managed equity portfolios to facilitate the 90/10 blend. Futures would be used to implement the shifts. Due to liquidity, Russell futures would be best for small cap, and S&P 500 for large.
- ETFs – exchange-traded funds could also be used to implement the size shifts. Again, the underlying physical assets could be either actively or passively managed portfolios, with ETFs being used for implementing the shifts. This approach would be more expensive to implement than using futures due to higher trading costs and the underlying ETF management fees.
- If the underlying assets are actively managed, an 80/20-large/small cap blend might be more appropriate for the physical assets. Value-added through stock selection is typically better in small cap. So to maximize this potential, the physical assets could be implemented using the maximum small cap allocation. Then buying or selling large cap futures would implement the shifts. S&P 500 futures, which are highly liquid, would be best for this approach.

Next Steps

Additional research is planned on these following fronts to further the power of this Tactical Equity Size Rotation model:

- Gather further back history to the mid 1970's to capture more of the volatile interest rate environment during that period.
- Refine the EVR alpha spread variable by looking at the current spread relative to a typical spread. In general, small cap stocks have higher alphas than large cap. So there may be more information content when the spread becomes atypical, rather than in the absolute level of the spread.
- Explore additional explanatory variables that could prove useful for forecasting the size premium.
- Given its apparent success, investigate the applicability of this approach to Style Rotation (Value vs. Growth).

Disclosure

Model results have certain inherent limitations. Unlike an actual performance record, model results do not represent actual trading and may not reflect the impact that material economic and market factors might have had on Mellon Equity's decision making if actual client funds were being managed. The model performance represented is simulated and is not an AIMR composite.

The author wishes to thank Warren Carlson, Charles Hendrix, Harry Grosse, and Pavani Reddy for their valuable research assistance.

Kenneth A. Barker

Ken is Senior Vice President, Director of Quantitative Analysis and Research within Mellon Equity. Ken leads our quantitative research team, drawing on his investment background and 35 years of experience. He serves on the Investment Management Committee and works with the chief investment officer and other senior managers to shape our investment strategy. Ken uses his background in systems, statistics and finance to assess the suitability, feasibility and likely significance of ongoing enhancements to our investment process. He and his team keep abreast of practical developments in finance, statistical forecasting, decision sciences, artificial intelligence and non-linear modeling.

Before joining Mellon Equity in 1989, Ken was director of research at Triangle Portfolio Associates, a Mellon subsidiary that made pioneering use of quantitative techniques in aggressive, concentrated portfolios. Before that, he held related positions at Deloitte Haskins & Sells, Tektronix, McDonnell Douglas Astronautics and The Rand Corporation. He frequently speaks at industry conferences, is a member of the Chicago Quantitative Alliance and serves on the board of directors of the Institute for Quantitative Research in Finance. Ken earned a BS in engineering from Case Western Reserve University and an MS in operations research at the University of Southern California.

The Mellon Equity Small Cap 900 and the Mellon Equity Large Cap 400 are proprietary indexes developed by Mellon Equity Associates for this study. You cannot invest directly in them on in any other index. The S&P 500, S&P 600 Small Cap, Russell 1000 and Russell 2000 Indexes are trademarks of Standard & Poor's Corp. and Frank Russell Co. The foregoing index licensers do not sponsor, endorse, sell, or promote the investment strategies or products mentioned in this article and make no representation regarding the advisability of investing in the products or strategies described herein.

The statements and opinions expressed in this article are those of the author as of the date of the article, and do not necessarily represent the view of Mellon Equity Associates or any of its affiliates. This article does not constitute investment advice, and should not be construed as an offer to sell or a solicitation to buy any security. For more information about any product or services of Mellon Equity Associates, please contact Robert Brinker at 412 234-7276.