The Index Is Dead. Long Live The Index.*

How to build a better benchmark

By Lubos Pastor, John Heaton and Aaron Foss
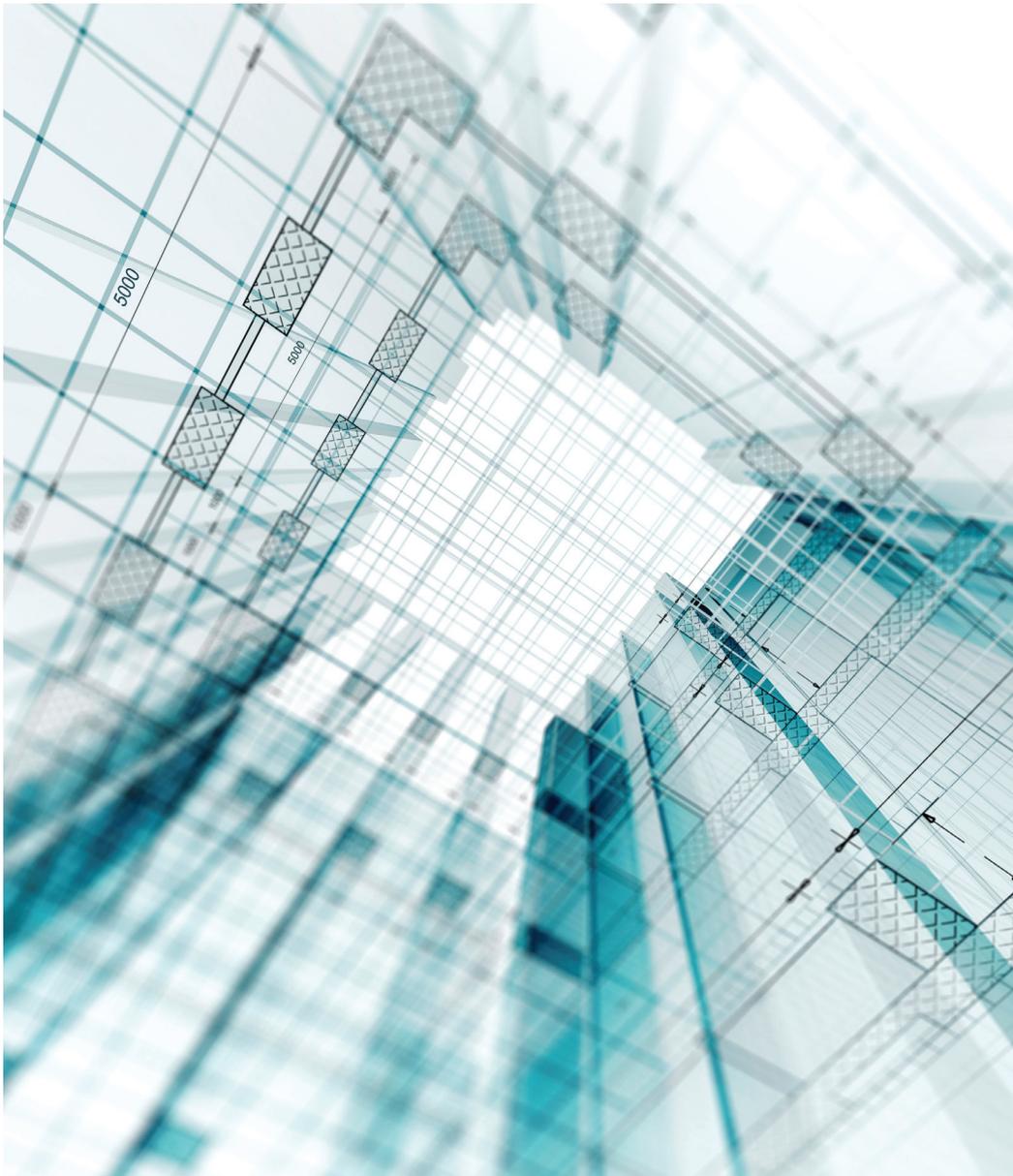
Gus Sauter, Vanguard's recently retired CIO, argues that theory provides little guidance as to how investing, beyond holding the total market, is actually done (Sauter, 2002). His point, that managers define the opportunity set, is best articulated in his cardinal rule of indexing: *An index must reflect the way that money managers actually invest.*

Building the Ideal Index

The Center for Research in Security Prices (CRSP) at the University of Chicago Booth School of Business sought, with its new index products, to combine current academic thought and practice with Sauter's cardinal rule, while paying considerable attention to the material constraints faced by investors. The result is a family of indexes that are both theoretically justifiable and a practical representation of those securities in which a manager, subject to a related mandate, could invest.

Other authors have done a good job of describing important features of benchmarks: completeness, objectivity, investability, etc. The CRSP Indexes have all of these features; however, this article delves deeper. CRSP seeks to explain our index design process and present the indexes' mechanics in the context in which the solutions arose. We believe our process can be most easily digested by understanding our theoretically guided and empirically validated approach and the balances struck that make CRSP's indexes valuable.

The Approach

CRSP's approach to index construction directly combines theory and empiricism:

Theoretically And Logically Guided

The CRSP Indexes aim to be "current." Company performance and valuation fluctuate with economic conditions, firm decisions and investor expectations. The result: Companies that looked cheap/expensive/small/big/U.S.-domiciled/liquid at one point likely will not remain so indefinitely. We could simply state that an index should be as close to current as possible, but that would have clear drawbacks in terms of turnover. A common industry compromise between being current and limiting turnover is to reconstitute indexes semi-annually or annually. The CRSP Indexes' quarterly reconstitution places a relative premium on being current; we married reconstitution to a novel migration strategy that limits turnover. The resulting indexes reflect changes to the investment opportunity set quickly while keeping turnover low.

Free float is another example of a constraint derived theoretically/logically rather than empirically. It has become widely adopted by index providers because it makes sense. Shares that are not available for trading cannot possibly lie in the investment opportunity set. While a departure from pure cap-weighted indexes, free-float-adjusted indexes are a more appropriate representation of those assets that investment managers should consider in their decisions.

A robust way to set breakpoints is another theoretical

problem. Breakpoints reflect the ability to discriminate between the opportunity set contained in one index and the next. Ideally, breakpoints would arise naturally along some dimension of investor concern. However, the dimension of investor concern itself may not be obvious.

We begin with a simple example: market cap. Academics and practitioners have long noted that equities with different market capitalizations display significant differences in average returns. In addition, small stocks and large stocks, as groups, have tended to move together. Domestic index providers have historically made the decision on cap-index membership in an ad hoc fashion by using counts of securities as proxies for market capitalization. For example, the S&P 500 Large-Cap Index contains 500 stocks, while the Russell 1000 Large Cap Index contains 1,000 stocks when reconstituted. However, most of us scratch our heads when asked why an investor should care whether a security is ranked 999 or 1,001 by size; it would seem hard to argue that such arbitrary cutoffs reflect genuine investor concerns. Furthermore, count-based indexes remain anchored at a point in time and always reflect the relationship that count had with

Figure 1a

Counts Of U.S. Stocks

The number of listed stocks changes significantly over time, falling by almost half over the past 12 years. Concurrently, the market risk represented by a portfolio with a fixed number of stocks changes.

Source: CRSP

Figure 1b

Market Capitalization Of Common Stock Counts

A large-cap index of the top 500 stocks jumped from around 75% of cumulative market capitalization in the mid-1990s to almost 85% five years later.

Source: CRSP

Figure 2

Presumed Opposites

Shuffling the Gordon model so that there is a scaled price ratio (Price / Cash Flow) on the left shows a clear relationship with the discount rate and expected growth.

Value investor's concern

$$\frac{P}{CF} = \frac{1}{r - g}$$

Growth investor's concern

Source: CRSP

market cap when the index was conceived. Since the number of listed securities fluctuates dramatically over time, as shown in Figure 1a, the economic significance of a given count-based index changes, too. Figure 1b shows that today's Russell 1000 Index represents a substantially larger proportion of the total market than it did 15 years ago. In contrast, CRSP's capitalization-based indexes use cumulative capitalization breakpoints, a solution common among other providers in the international space. Cumulative cap maps closely to more reasonable investor concerns and has the advantage of keeping the indexes current from an economic perspective.

Value and growth style index breakpoints and categorization present different theoretical issues. Numerous academic articles exploring the differences in returns between stocks have shown that value managers are on to something—stocks that are "cheaper" than their peers based on scaled price ratios tend to have higher future returns. These "value" stocks tend to move together, as do

stocks at the opposite end of the spectrum, the "growth" stocks. The co-movement seen in these value and growth stock portfolios provides evidence of a "value" factor in stock returns. Still, "value" remains ill defined.

Common academic valuation frameworks dictate that value and growth are negatively correlated. CRSP, too, believes that value and growth are two distinct categories. We go slightly further, however, and specify that value and growth should be determined relative only to other securities in the universe, rather than against some cap-weighted aggregate metric for a universe. Our specification means a security's placement is a statement of cap-specific relative value and/or growth versus other stocks.

Investors should be able to immediately recognize how CRSP's style placement decisions make the indexes meaningful measures of manager style performance. To clarify the intuition: CRSP understands that a large-cap value manager may only choose securities that look like value stocks within his or her universe; the use of a relative value score specific to the large-cap universe provides a good representation of his or her opportunity set. It also means that the same security may have a different style assignment in our midcap or large-cap portfolios, which use their own relative scores.

Empirically Validated

Specifying what can reside within the opportunity set is a problem different from determining the best way to map to what investors see. The investor's vantage point must be inferred empirically.

Even a question as simple as "What is a domestic company?" has a largely empirical answer. Companies that all of

Figure 3

Respecting Convention

CRSP's cumulative cap breakpoints (shaded background) are aligned with industry convention to allow for easy translation from count-based indexes. Russell, MSCI and S&P columns show cumulative capitalization against the most complete index in each respective column (e.g., Russell Top 50 shows cumulative capitalization vs. the Russell 2000).

Sources: CRSP; Russell and MSCI data as of 3/22/2013 from Bloomberg; S&P data as of 3/22/2013 from S&P fact sheets.

us would consider American are incorporated overseas and recognize revenue abroad for tax purposes. Others, clearly foreign, go public through reverse mergers with U.S.-listed shell companies. We turned to data in an attempt to develop a transparent methodology. The domicile scheme we rely upon is a product of testing thousands of models that examine hundreds of company-level variables.

Other areas are less clear still—cap breakpoints, for example. As mentioned before, large and small stocks behave differently, but even after significant research into potential breakpoints, no clear statistical discontinuities emerge. CRSP's conclusion: Cap segments are a matter of convention. In an effort to stay practical, we adopted cumulative cap breakpoints at levels that should look very familiar to practitioners. A comparison with other major indexes is provided in Figure 3.

Importantly, though, CRSP also included "bands" around these breakpoints and a migration plan called "packeting." We studied banding and migration in depth and let the data provide guidance as to scale and mechanics. We require that a security pass a threshold beyond the breakpoint before a 50 percent "packet" of the security's weight is moved to the adjacent index. It follows that a small-cap stock must move beyond the mid/small breakpoint and the mid/small band threshold before triggering the first 50 percent move to the midcap index; the migration of the remaining 50 percent depends on remaining beyond the threshold next quarter. Among all index providers that rely on banding, turnover reduction is touted as a benefit. We agree, but we believe there is also a deeper purpose. As the indexes age, banding and packeting capture something important: There is no unanimity of opinion among managers as to which marginal companies qualify as large or small. Beyond the reduction in turnover, our migration strategy improves the fit of the index to manager behavior.

If there is a lack of unanimity on cap breakpoints, value and growth styles are downright fuzzy. CRSP's work suggests that, in contrast to the model in Figure 2, value and growth managers do something other than hold portfolios of single-dimension "value" or "growth" stocks. Value managers typically describe their process as one that involves trying to buy assets or cash flows at inexpensive prices. Growth managers, on the other hand, look for fast-growing firms. These managers follow separate, though related, processes. Accordingly, we treat value and growth separately. This two-dimensional method allows us to generate a richer description of the style-specific investment opportunity set.

Recent empirical research provides useful insights into the factors managers consider. CRSP was the first to introduce investment rate ("INV") and return on assets ("ROA") as growth factors. Academic studies show that firms that invest more tend to grow faster, as do firms that are more profitable. Additionally, economic theory links both INV and ROA to expected stock returns.

CRSP studied its factors in typical empirical fashion: portfolio sorts along factor dimensions, cross-sectional Fama-MacBeth regressions, cross-sectional and predic-

Figure 4a

The Latest Styles

CRSP's multifactor model is designed to be easily intelligible to investors. The model acknowledges that all factors are not equally important.

VALUE	GROWTH
EP = (2/3) FEP + (1/3) HEP	FG = (1/3) FLGE + (1/3) FSGE + (1/6) INV + (1/6) ROA
V1 = (2/3) EP + (1/3) BP	HG = (2/3) HGS + (1/3) HGE
V2 = (2/3) SP + (1/3) DP	G = (2/3) FG + (1/3) HG
V = (2/3) V1 + (1/3) V2	

Source: CRSP

Figure 4b

Decomposition Of CRSP Style Model

While the actual model is hierarchical, the flow-through weights expressed in our factor model are visible in the chart.

Source: CRSP

Figure 5

Thinking Inside The Box

The CRSP Style Box is an intuitive representation of our style assignment routine. Our two-dimensional style model allows us to plot any firm in a value/growth plane.

The 0.5 AR breakpoint separates value from growth, but it is not a hard line, as the investment styles themselves dictate.

After initial placement, a security's style characteristics may change. Given the inherent ambiguity in style, migration requires a significant change in characteristics; one that takes a firm from its current location through the band region (white) to the opposite band threshold. If the security remains beyond the threshold at rank, a 50% packet is moved.

This model can be easily extended to form indexes containing characteristics that are value-only, growth-only, a combination of value and growth, or neither.

Source: CRSP

tive rank correlations, etc. We conducted individual and multifactor regression tests. Validation depended on achieving the behavior expected for the process: Value factors should explain variation in future returns, while growth factors should predict future growth.

However, CRSP also recognizes that investment managers possess information beyond that contained in scaled price ratios and growth statistics. The decisions these investors make are shaped by this unobservable information. We aim to capture that information by choosing an appropriate model—one that proxies for information managers know, but that we do not know. Naturally, the best proxies should be those that best emulate active managers.

CRSP's exercise was to find a set of weights for our factors that tracks the most widely used active manager indexes, Lipper and Morningstar, with limited error and low turnover. We ultimately evaluated more than 2,500 differently weighted candidate factor models before coming to our current design. Here there was a risk of over-fitting—selecting a model that was just representative of the world that did occur as opposed to one that is a better representation of all worlds that may have occurred. Simple rank tests helped stratify our models in sample. However, we also conducted a novel cluster analysis to understand broad functional classes of possible models. We were able to map the performance of these functional model classes back to the underlying factors to understand what factors managers used in their style appraisal and in what proportions these factors were likely considered. The cluster and ranking analysis agreed in a large number of cases, which gave us increased confidence in our ultimate index design.

CRSP arrived at a model that uses five value and six growth factors (Figures 4a and 4b). For value, our model groups forward and historical earnings to price ("FEP" and "HEP," respectively) into an EP factor and combines that with book to price ("BP"), creating a primary value superfactor ("V1"); sales to price ("SP") and dividend yield ("DP") create a secondary value superfactor ("V2"). The two value superfactors merge into a composite value score ("V"). Growth builds a future growth superfactor

("FG") from analyst-estimate future long-term growth in earnings ("FLGE"), analyst-estimate future short-term growth in earnings ("FSGE"), INV and ROA. A historical growth superfactor ("HG") comprises three-year historical growth in sales ("HGS") and three-year historical growth in earnings ("HGE"). The two growth superfactors combine to make a composite growth score ("G").

CRSP assigns each composite score a rank value ("RV" or "RG") as a percent of the cumulative market cap with lower scores. The growth score is inverted and the scores averaged to arrive at an average rank ("AR"). High ARs (those above 0.5) are value securities, low ARs are growth. As mentioned earlier, there is little agreement on what, specifically, value and growth are, or when, as often happens, a growth security becomes a value security (and vice versa). Therefore, we employ our threshold packeting mechanism here as well. As in the cap indexes, a security must pass a threshold beyond 0.5 AR before a 50 percent packet is moved to the adjacent style index (Figure 5). Again, this improves fit with manager behavior and dramatically decreases turnover.

The Balancing Act

Steven Schoenfeld, in "Perfection Impossible," highlights that index development is a process with inherent

Figure 6a

What Trade-Off?: Morningstar
Annualized Turnover Vs. Annualized Monthly Tracking Error

As claimed earlier, CRSP's migration strategy serves two purposes: 1) reduce turnover; and 2) improve fit to manager behavior. Figures 6a and 6b use simple averages of CRSP's style indexes' performance to illustrate the success of our methodology against popular measures of manager behavior.

The point labeled "Ideal" is our multifactor model with no banding or packeting. "Packeting" has no banding, but introduces the 50% packet. Each subsequent point represents band width. Our selection, point "33" (red), is an intuitive, obvious improvement on a no-migration-strategy model.

Source: CRSP, periods 6/30/01-3/31/11
Note: Tracking error, as presented above, is the simple average of each CRSP style index's tracking error with respect to the appropriate Morningstar active manager benchmark for a given banding/packeting conformation.

Figure 6b

What Trade-Off?: Lipper
Annualized Turnover Vs. Annualized Monthly Tracking Error

Source: CRSP, periods 6/30/01-3/31/11
Note: Tracking error, as presented above, is the simple average of each CRSP style index's tracking error with respect to the appropriate Lipper active manager benchmark for a given banding/packeting conformation.

predictable process, and several leading indexes see front-running by active managers looking to trade securities transitioning into or out of the index ahead of passive managers who must make these trades. We randomize the pricing date during reconstitution, which alleviates front-running by making the inclusion/exclusion of marginal securities in a specific index harder to predict. The randomization itself follows a transparent algorithm—it should preclude manipulation of index membership without introducing any methodological opacity.

Tracking error, too, comes at a cost, albeit in a slightly less salient "risk" dimension. This can be thought of as the potential for variance in returns versus the instantaneous opportunity set. Active, benchmarked investors call this "active" risk, as it represents a decision to deviate from a naive position in the asset. Measuring the quality of investing decisions then becomes a question of the return-to-risk ratio.

In some sense, index providers make similar decisions to those made by active managers. The index obviously deviates from the instantaneous "true" opportunity set; the goal of the index provider is to strike a balance between the cost of turnover and risk from tracking error. Developing transparent, mechanical rules for banding and migration precludes a simple functional approach, but the idea pays homage to standard mean-variance optimization. To determine the appropriate mechanics, CRSP ran a large number of experiments that studied problems ranging from the symmetry of bands at breakpoints to the ideal band width to the use of thresholds versus continuous transitions to transitional packet size. We evaluated more than 40 different approaches before settling on our breakpoints, threshold bands and 50 percent packeting. We compared all approaches on several metrics: tracking errors versus appropriate Lipper and Morningstar manager indexes, tracking error versus "pure" (i.e., no migration strategy) indexes, aggregate turnover and bad turnover (Figures 6a and 6b).

contradictions and tensions. One tries to provide a good measure of the opportunity set for an asset class or style at a given point in time, but must balance the abstract desire for a perfect measure against practical considerations of fund operations. All major index providers tackle operational issues with features such as free-float adjustments and banding/migration schemes. CRSP is no different in this respect, though the attention we pay to costs borne by the investor separates the CRSP Indexes from others.

The two most important costs of indexing are turnover and tracking error. Turnover is obvious; irrespective of cause, turnover has a direct dollar cost and eats into returns. The bulk of turnover occurs at reconstitution, when constituents move in or out of an index. To track the index, the indexer must access capital markets. Capital-market access comes with price impact, especially when the volume of a transaction is relatively large.

As one could imagine, turnover and price impact are different. An added dimension of our analysis looked at, for lack of a better name, "bad turnover." Bad turnover is that which demands transacting in volumes that are difficult for the market to meet and are thus expected to have greater price impact cost.

The CRSP Indexes also address turnover costs associated with front-running. Reconstitution is a fairly

Tying It All Together

We believe the CRSP Indexes represent a meaningful improvement in index usability. True to the cardinal rule, the indexes should prove to be a cost-effective approximation of the actual investing process.

For purely passive investors, the cost of implementing the CRSP Indexes is low. From an active investor perspective, CRSP's efforts to keep the indexes current and remove ad hoc constraints mean managers will no longer expend their risk budgets by allocating to securities that have recently changed category or by simply following their investment style.

Other improvements are easy to assess. When considering risk, CRSP's cumulative market-cap breakpoints give a more meaningful stock market "size" exposure. Similarly, CRSP's style indexes use many factors and a carefully validated weighting scheme to better approximate "value" and "growth" as styles.

As for market impact and trading costs, lower turnover is a big step in the right direction. CRSP conducted a study looking at the indexes we replaced, and calculated the turnover characteristics of the funds as if they had been historically managed against the CRSP Indexes. CRSP's style indexes show impressive reductions in turnover. Keep in mind that the turnover reductions were achieved even with CRSP's more rapid quarterly reconstitution, which provides a better measure of the current investment opportunity set.

We have attempted to provide new insight into the motivation and practical elements that underlie our index construction. While there has been significant convergence in methodologies among the major index providers, that is solely an end—one that provides little insight into means. We believe that CRSP's academic heritage and our desire to both study and represent the "real" world are distinguishing characteristics of *our* means.

A response to "Index Rex," by Gus Sauter, published in the 2Q 2002 issue of the Journal of Indexes.