

Orchestrating Alpha Melodies

Ensemble Approach for Stock Selection Models

The ensemble approach, in a broad sense, refers to combining multiple units (which could be models, methods, or components) to work together to achieve more accurate and reliable outcomes than any individual unit could achieve alone. This is akin to making a decision based on the collective wisdom of a diverse group, rather than relying on a single opinion. Imagine you are trying to compose a song but are stuck on the melody. You decide to get opinions from your musician friends who play different instruments. The pianist friend suggests a soft, flowing melody with rich chords; the guitarist friend prefers a catchy, rhythmic tune while the drummer friend thinks a strong, consistent beat is crucial. Instead of strictly following one friend's suggestion, you decide to blend their ideas to compose your song, symbolizing the ensemble approach. You might integrate parts of the soft, flowing melody from the pianist with the catchy tune from the guitarist. Perhaps you'll give more emphasis to the rhythm in the chorus (guitarist and drummer) and focus on melody in the verses (pianist). By combining these various perspectives, you can compose a song that is well-rounded and harmonious.



The ensemble method offers several intuitive advantages. First and foremost, it leads to improved accuracy, as multiple models capture various patterns in the data that a single model might miss. When combined, these patterns provide a more comprehensive view that avoids overfitting and false positives more effectively, leading to better predictions. Secondly, by combining these models, the ensemble method effectively reduces individual biases, ensuring a more balanced and objective outcome. An ensemble is less susceptible to the pitfalls of any single model. For instance, if one model fails or makes an erroneous prediction due to certain anomalies in the data, other models in the ensemble can compensate for that error, leading to a more rigorous overall prediction. Lastly, this method provides enhanced robustness. Every model has its strengths and weaknesses. An ensemble approach enables capitalizing on the strengths of several models, thereby mitigating their individual limitations.

October 2023

Authors



Bin Cheng, CFA, CAIA
Vice President, Team Lead,
Quantitative Research Group



Serena Li, CFA
Quantitative Analyst,
Systematic Global Equities



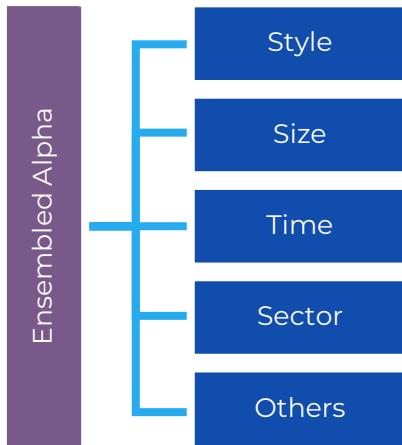
Sumali Sanyal, CFA
Managing Director, Senior Portfolio
Manager, Systematic Global Equities



Cameron McLennan, CFA
Director, Senior Portfolio Manager,
Systematic Global Equities



Chart 1 | Dimensions for Building Stock Selection Models



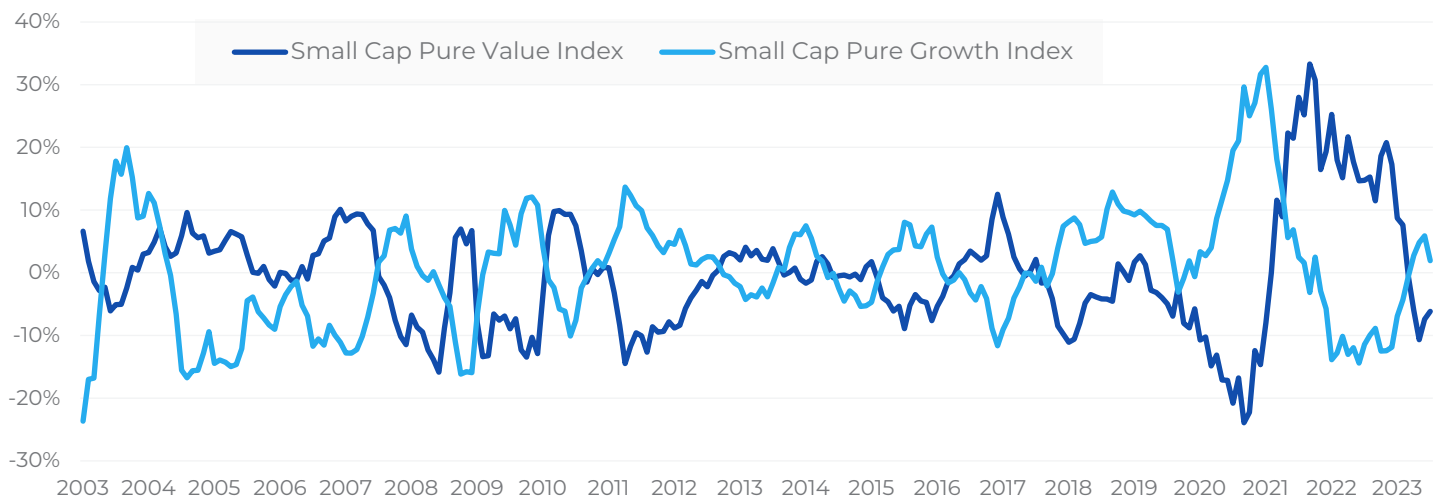
Diving into the world of investment management, we have identified an important role that an ensemble approach can play in building stock selection models. We leverage the advantage that an ensemble approach offers to increase the alpha prediction and mitigate overfitting and false positives that can be associated with stock selection models. The avenue to implement ensemble approaches in building stock selection models is to first identify the set of significant dimensions such as style, size, time horizon, and sector etc., which profoundly influence stock performance across universes (as illustrated in **Chart 1**). Subsequently, tailored stock selection models can be developed to pinpoint the top performers within each dimension. We advocate for constructing bespoke stock selection models for each dimension and then integrating them using an ensemble approach. This ensures that no single investment style overshadows the others, fostering a balanced and diversified portfolio.

Consider, for instance, a large-cap growth stock within the consumer discretionary sector, it would be an oversimplification to have one generic model, trying to predict its forward-looking excess return, relative to the entire universe which includes its value, smaller cap, and financial counterparts. The complexity arising from the myriad

of influencing factors or dimensions could cause a generic model to make unintended style bets or encounter challenges in identifying reliable and consistent patterns. We advocate for a thoughtful ensemble approach in this case to specially address some of the concerns presented. The approach involves isolating significant dimensions and constructing respective targeted models to rank the subject within its refined peer groups—namely, growth, larger cap, and consumer discretionary dimensions. Consequently, the alpha score of this stock will be assembled to reflect a balanced perspective, emanating from three distinct stock selection models.

The rationale behind this approach is intuitive and straightforward. Stocks with pronounced value characteristics can exhibit markedly different behaviors than those aligned with growth characteristics, a divergence strikingly depicted in **Chart 2**, which contrasts the rolling 12-month excess returns between the pure value index and the pure growth index in the small cap space. Market capitalization demonstrates another significant variation. Stocks with a larger market cap often perform differently than their smaller market cap counterparts. Additionally, sector affiliation can play a pivotal role; for example, stocks in the technology sector yield different return streams compared to those in the financial sector, as **Chart 3** aptly illustrates.

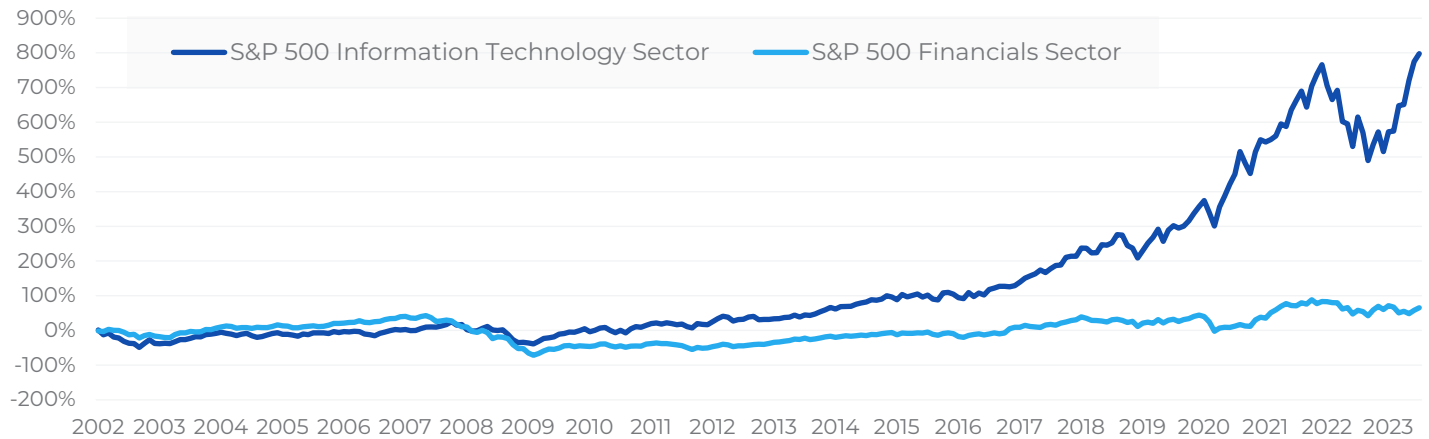
Chart 2 | Rolling 12-month Excess Return



Source: FactSet



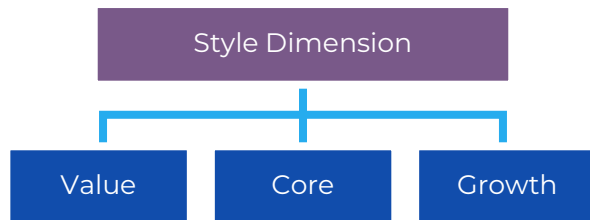
Chart 3 Cumulative Raw Return



Source: FactSet

Factors pivotal in discerning champions among value stocks might deviate considerably from those influential in pinpointing standouts among growth stocks. And even if the factors that matter for these two groups appear similar sometime, their relative importance or weight distribution can display substantial variations. Designing a singular generic model for the entire universe could lead to inefficiencies, especially when such a model grapples with both value and growth stocks to establish

Chart 4 Model for Identifying Top-performing Stocks Using Style Dimension



reliable and consistent patterns. The model can be easily misguided to eventually just make an unintended style bet in identifying winners, i.e., prioritizing growth over value stocks. But to construct a well-rounded portfolio with proper diversification, investors need top picks from both categories. We advocate for the formulation of distinct, specialized models for both value and growth segments, ensuring that each model's primary objective remains clear: identifying top-performing stocks within its designated refined subset (as shown in **Chart 4**). This degree of specialization ensures that the insights derived are more precise, relevant, and actionable.

Another dimension that is both critical and overt is sector affiliation. Factors that matter for identifying winners in technology markedly diverge from those relevant to top performers in the financial sector. Attempting a singular cross-sectional stock selection to pinpoint top performers simultaneously across a universe encompassing constituents from disparate sectors can lead to ambiguity. The factor distribution and pattern suitable for the technology sector may possess minimal efficacy for the financial sectors, leading the model to yield inconsistent and weak patterns that barely conform to either category. Consequently, we advocate for the separation of sectors into distinct segments and the construction of tailored models within each to discern top-performing technology stocks, for instance, in relation to their respective technology peer group. We have advanced further in this domain by developing a proprietary clustering methodology. This innovation provides insights on optimal segregation for dissecting sector dimensions (as shown in **Chart 5**), which are more pertinent beyond GICS classification, ensuring the preservation of a larger sample size conducive to a more stable model pattern.

Chart 5 Insights on Optimal Segregation for Dissecting Sector Dimensions

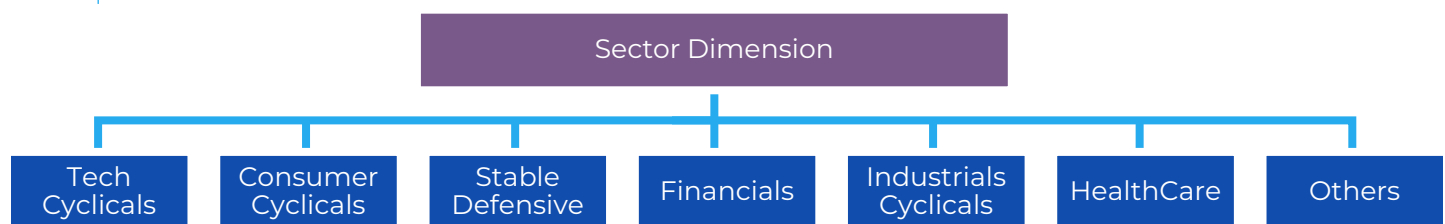
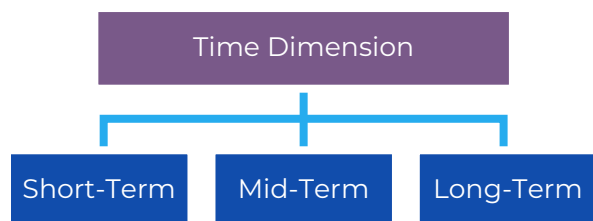




Chart 6 Capturing Short-, Mid-, and Long-term Market Dynamics with Time Dimension

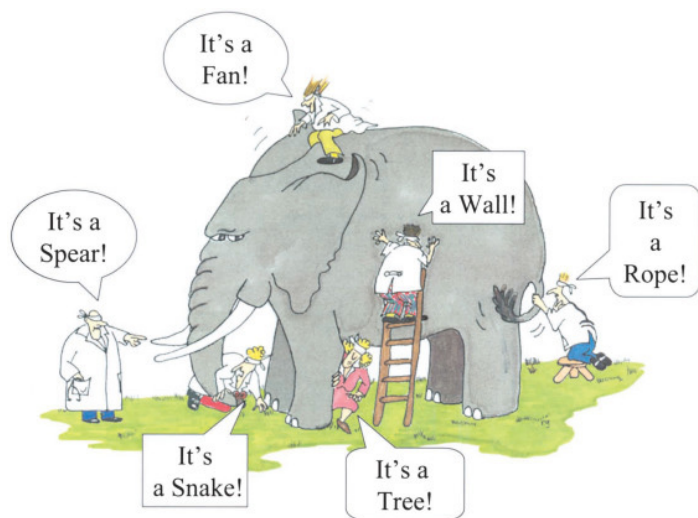


The final example to elaborate upon pertains to the dimension of time horizon. Systematic investors often find themselves debating and determining the time horizon upon which a model is constructed and trained. Some investors favor a shorter term to remain as responsive as possible to market changes, fully aware of the downside – a low signal-to-noise ratio. Conversely, some prefer a longer term to prioritize model stability, understanding the tradeoff involved - the lag in how the model responds to potential trend alterations. We recognize the merits in both scenarios and, as such, design a dedicated time dimension to construct tailored

models, each with respective rolling model training horizons to capture short-, mid-, and long-term market dynamics (as illustrated in **Chart 6**). The alpha score for each stock, derived from the time dimensions, will yield a more balanced ranking, accommodating varying market environments throughout a full market cycle.

When robust dimensions are operating efficiently to deliver targeted predictions, taking into account a spectrum of significant market drivers, the ensemble approach consolidates these varying perspectives to furnish an unbiased and diversified alpha score for each stock. This method permits each dimensional model to contribute their insights, ensuring no singular investment style predominates. The ensemble approach enables each dimension to validate or challenge their respective rankings, thus, the ensuing alpha score exhibits heightened accuracy and diminished biases, circumventing overfitting and mitigating false positives at the same time. To conclude on a lighter note, consider the allegory of the blind men and the elephant, illustrated in the cartoon below. The blind men each describe an elephant from their limited perspective. While their descriptions are relevant, they remain fragmented and incomplete. A more accurate and holistic understanding of the elephant would emerge if they consolidated their various descriptions. Similarly, we, as investors, confront another “elephant” — the market. We assert that the ensemble approach for building active stock selection models can augment our capability to outperform it more sustainably.

This paper serves as the first in a series of practical research insights that we intend to publish, articulating our thoughts and approaches for systematic active investment. There are many other critical aspects of a systematic investment process that requires extensive research, in addition to the ensemble approach and dimension construction we’ve expounded upon in this post. What factors should investors vet through to build an active stock selection model? How should one allocate weight distribution to the chosen set of factors within the stock selection model? What methodologies should investors adopt for alpha prediction calibration? How should portfolio optimization be crafted, factoring in risk budgets and constraints? These are all important questions to address for building out a robust systematic investment process. We will continue to share our insights, rooted in rigorous internal research, in future posts.



Source: *The blind men and the elephant*. Poem by John Godfrey Saxe (Cartoon originally copyrighted by the authors; G. Renee Guzlas, artist). <https://www.sciencedirect.com/science/article/pii/S0085253815487095>

This report is neither an offer to sell nor a solicitation to invest in any product offered by Xpounce® and should not be considered as investment advice. This report was prepared for clients and prospective clients of Xpounce® and is intended to be used solely by such clients and prospects for educational and illustrative purposes. The information contained herein is proprietary to Xpounce® and may not be duplicated or used for any purpose other than the educational purpose for which it has been provided. Any unauthorized use, duplication or disclosure of this report is strictly prohibited.

This report is based on information believed to be correct, but is subject to revision. Although the information provided herein has been obtained from sources which Xpounce® believes to be reliable, Xpounce® does not guarantee its accuracy, and such information may be incomplete or condensed. Additional information is available from Xpounce® upon request. All performance and other projections are historical and do not guarantee future performance. No assurance can be given that any particular investment objective or strategy will be achieved at a given time and actual investment results may vary over any given time.