Taming the Factor Zoo

Overview

Since the introduction and subsequent first test of the Capital Asset Pricing Model (CAPM) market factor over 40 years ago, the asset pricing literature has produced hundreds of potential risk factors. Along with the rise of this "factor zoo," distinguishing between useful, useless, and redundant factors from among them has become increasingly important. The authors propose a new methodology that allows researchers to systematically test and evaluate potential new factors for asset pricing. The model's usefulness stems from its explicitly accounting for possible mistakes (like omitted variables) in model selection as opposed to assuming the researcher's model is perfectly chosen. This is especially important as the number of potential factors is large (and growing), so researchers cannot be sure to have chosen the correct model. The authors then apply their approach to a large set of factors recently proposed in the literature and show that appropriately selecting the benchmark model against which to evaluate new factors represents a robust way to determine the usefulness of new factors.

Investigation

With the rise of factors in asset pricing, it has become critically important to discriminate between new, useful factors that should be employed versus those that are not useful, or redundant. That is, are there really 300 useful factors as the cumulative asset pricing literature suggests, or are they all some combination of a smaller set? A robust framework is needed to select from among the myriad of factors found in the literature and to discipline the proliferation of factors. The dominant asset pricing approach currently in place to evaluate any new factor is to compare it to standard models in the literature, like the Fama-French 3- and 5-factor models. But, how do we know that this model is the correct benchmark? The authors propose another approach that tries to compare the contribution of a new factor to all existing factors while simultaneously trying to avoid data snooping biases.

Such a comparison is infeasible using standard approaches. The authors apply and extend recent model selection econometric techniques in order to systematically (not arbitrarily or data mined) select the best possible control model out of the existing, known set of factors. The contribution of any new factor is then compared to the control model. The proposed methodology termed double-selection LASSO draws from machine learning to employ a two-pass regression approach without assuming prior knowledge of which factors to include as controls from among the possible hundreds of factors found in the literature.

The authors apply their methodology to a large set of more than 150 factors proposed over the past 30 years to evaluate the marginal contribution of new factors proposed in just the past five years. The new factors examined include profitability, betting against beta, quality minus junk, and investment, among others. One might expect that with over 30 years of research, newly "discovered" factors might contribute little, if anything, to explaining the cross section of expected returns; however, the results show otherwise. The authors find:

- Several new factors make a significant contribution to the explanatory power for expected returns including profitability, investment, and quality minus junk even after accounting for the large set up factors proposed in the years prior to five years ago.
- The evaluation of the usefulness of these significant factors were found to be robust, or stable across different samples.
- Applying the proposed test over time to some 150 factors proposed in the literature deems only a small number of them significant, thus significantly reducing the "zoo" of factors.
- The authors' model results differ meaningfully from the conclusions one would obtain by using the well-known Fama-French 3-factor model as the control benchmark — fewer new factors would be considered to contribute significantly on the margin.

Conclusion

Making meaningful progress in asset pricing research requires an improved framework for evaluating and disciplining the proliferation of asset pricing factors. This suggests a robust approach for studying the marginal contribution of new factors relative to the immense set of existing ones is needed to allow a way to screen new factors as they are proposed. The authors offer a two-step model-selection method to bring discipline to this challenge and to help researchers organize the current zoo of factors. Their approach, based on machine learning and model selection techniques, systematically evaluates the contribution of any new factor to asset pricing. Using their approach to evaluate a large set of factors proposed in the literature reveals that only a small number of factors capture truly new dimensions of risk.

This document is not intended to, and does not relate specifically to any investment strategy or product that AQR offers. It is being provided merely to provide a framework to assist in the implementation of an investor's own analysis and an investor's own view on the topic discussed herein.

This document has been provided to you solely for information purposes and does not constitute an offer or solicitation of an offer or any advice or recommendation to purchase any securities or other financial instruments and may not be construed as such. The factual information set forth herein has been obtained or derived from sources believed by the author and AQR Capital Management, LLC ("AQR") to be reliable but it is not necessarily all-inclusive and is not guaranteed as to its accuracy and is not to be regarded as a representation or warranty, express or implied, as to the information's accuracy or completeness, nor should the attached information serve as the basis of any investment decision. This document is not to be reproduced or redistributed to any other person. The information set forth herein has been provided to you as secondary information and should not be the primary source for any investment or allocation decision. Past performance is not a guarantee of future performance. Diversification does not eliminate the risk of experiencing investment losses.

This material is not research and should not be treated as research. This paper does not represent valuation judgments with respect to any financial instrument, issuer, security or sector that may be described or referenced herein and does not represent a formal or official view of AQR. The views expressed reflect the current views as of the date hereof and neither the author nor AQR undertakes to advise you of any changes in the views expressed herein.

The information contained herein is only as current as of the date indicated, and may be superseded by subsequent market events or for other reasons. Charts and graphs provided herein are for illustrative purposes only. The information in this presentation has been developed internally and/or obtained from sources believed to be reliable; however, neither AQR nor the author guarantees the accuracy, adequacy or completeness of such information. Nothing contained herein constitutes investment, legal, tax or other advice nor is it to be relied on in making an investment or other decision. There can be no assurance that an investment strategy will be successful. Historic market trends are not reliable indicators of actual future market behavior or future performance of any particular investment which may differ materially, and should not be relied upon as such. Diversification does not eliminate the risk of experiencing investment losses.

The information in this paper may contain projections or other forward-looking statements regarding future events, targets, forecasts or expectations regarding the strategies described herein, and is only current as of the date indicated. There is no assurance that such events or targets will be achieved, and may be significantly different from that shown here. The information in this document, including statements concerning financial market trends, is based on current market conditions, which will fluctuate and may be superseded by subsequent market events or for other reasons.