Don't Throw out Anomaly Studies Just Yet!

September 11, 2017

by C. Thomas Howard

Advisor Perspectives *welcomes guest contributions. The views presented here do not necessarily represent those of* Advisor Perspectives.

In his recent article, Michael Edesess argued that multiple empirical "anomaly" studies and the wide use of regression are ruining finance research. While some of his points are valid, his conclusion that the entire set of academic studies should be discarded goes too far.

I am a former academic researcher. But this is not an esoteric argument between researchers. It is important to understand the benefits and limitations of such research as it has strong bearing on the debate raging on whether or not the markets are informationally efficient, and, in turn, if it is possible to use anomalies for earning excess returns.

I am well aware of the limitations when conducting empirical studies in a field like finance with its extensive, historical primary and secondary databases, where highly sought after promotion, tenure, and academic prestige depend upon publishing referred articles.

On the one hand, these immense data sets provide an opportunity for testing theories both cross-sectionally and through time. In many respects, this is a researchers dream. But it also makes data-snooping possible and thus the publication of selective, mostly positive results. This is the problem Edesess highlighted in his article.

This problem clearly exists in the anomaly research studies on which he focuses. But depending upon the question being addressed, these studies still provide useful information.

Is the stock market informationally efficient?

Another way to pose this question is, "can public information, such as P/E ratios and momentum, be used to earn excess returns?" The anomalies empirical literature answers a resounding "yes!" The Campbell Harvey talk Edesess referenced is based on this paper in which over 300 statistically significant, academically reviewed anomalies studies are catalogued. On its face, this is conclusive evidence that markets are not informationally efficient, since numerous pieces of public information can be used to earn excess returns.

But Harvey argues that, as Edesess parrots in his article, in order to avoid a data-snooping bias, the minimum t-value should be raised from the traditional 2 to 3. But even when this is done, half of the anomalies remain statistically significant, still providing overwhelming evidence against market efficiency.

Can a specific reported anomaly be used for earning superior returns?

While this question sounds similar to the one above, it is fundamentally different. If only half of the reported anomalies actually work, then the odds are not good that a successful portfolio can be built around a single anomaly. However, managers rarely rely on just one anomaly, what I refer to as the ingredients used for designing an equity strategy, but include a combination of anomalies. For example, if four anomalies are used for managing a portfolio, the chance that the portfolio will earn an excess return increases from 50% to over 90% (the chance that at least one of the anomalies generates an excess return), assuming little correlation among anomalies, as confirmed by Harvey. Thus a mixture of reported anomalies provides an excellent starting point when designing an equity strategy.

Minimizing false positives at the expense of false negatives

Increasing the minimum t-value means fewer false positives (an anomaly does not work but turns out to be statistically significant, referred to as a type I error) but unfortunately also fewer false negatives (one that works but is statistically insignificant, referred to as a type II error) are published in the academic literature.

So as the minimum t-value is increased from 2 to 3, fewer anomalies will be published but some of those will be effective anomalies that by chance are not statistically significant. In the case of the 150 anomalies in Harvey's study that did not clear the t=3 hurdle, some number of them actually worked but were false negative victims.

Such is the nature of statistical testing in that there is always a tradeoff between false positive and false negative errors. Traditionally, researchers choose to minimize false positives at the expense of false negatives. That is, it is argued that it is better to reject an anomaly that actually works than to accept one that does not. Increasing the t-value means that many effective anomalies never see the light of day.

Imperfect but effective

Finally, Edesess' contention that we should stop relying on regressions and simply look at pictures when determining significance is silly. Even in light of the many problems with regression and other statistical tools highlighted by Harvey and others, they remain a powerful set of tools for identifying useful relationships in market data. This is particularly true of academic studies, which conduct a wide range of robustness tests prior to publication.

A final quote from Harvey sums it up:

> The key…is to design appropriate statistical methods to adjust for biases, not to eliminate

> research initiatives. The multiple testing framework detailed in our paper is true to this advice.

Thomas Howard, PhD is professor emeritus at Daniels College of Business, University of Denver, CEO of AthenaInvest Advisors and Chief Investment Officer at AthenaInvest.