

## The price all investors pay for benchmarking

---

Michael Edesess | Fair Advisors | 30 January 2015

Could the practice of measuring and evaluating manager performance by comparing it to a market index be distorting prices across the whole market? That is the conclusion reached in a recent paper entitled "[Asset Management Contracts and Equilibrium Prices](#)" by three academic researchers, Andrea M. Buffa of Boston University and Dimitri Vayanos and Paul Woolley of the London School of Economics.

A warning – the paper is heavily mathematical. I have spent enough time poring over it to be convinced that the math is sound and the paper is innovative and important – possibly very important.

But unfortunately, neither the recent non-technical article about it in the *Financial Times* nor the authors' own non-technical explanation sent to me by one of the authors does a good job of explaining it for the layman. I believe it is onto something, so I will explain the core idea and why serious attention should be paid to the methodology and the conclusions.

### HOW BENCHMARKING COULD DISTORT THE MARKET

Start with a very simplified example. Suppose the market consists of only two securities or assets. Call these assets A and B. Further, suppose that the total market capitalization is 50% security A and 50% security B.

Now suppose that all investors start investing in these securities – in the market – by hiring fund managers. The managers are evaluated by comparing their performance to a benchmark or index. The benchmark happens, for whatever reason, to be an index consisting of 60% security A and 40% security B.

What will happen? Let's add some suppositions about the managers, their personal objectives and their incentives. Suppose that the managers are awarded a basic fee, but are also in effect rewarded (or penalized) for their performance by giving them (or taking away from them in the case of underperformance) a small percentage of the difference between their performance and that of the index.

Now, suppose these managers are loss-averse, meaning that they hate a loss in their personal incomes more than they love a gain of equivalent size. (1. The mathematical form of this is that their utility function  $U$  satisfies the inequality  $-U(-x) > U(x)$  for any  $x$ , that is, their utility is more negative for a loss than it is positive for an equivalent gain.)

Finally, suppose also that the managers don't really believe they can control whether they outperform the index or underperform it. Then their utility function dictates that they will simply hug the index as closely as possible. That way, if underperformance by  $x$  is as likely as outperformance by  $x$ , they will maximise their expected utility by neither outperforming nor underperforming.

#### **NOW THAT WE'VE MADE ALL THESE SUPPOSITIONS...**

What happens is rather odd – though not really, when you think about it. All managers want to be 60% A and 40% B. But the market is only 50%/50%. So the managers must compete to buy the undersupplied security A, thereby driving up its price. Ultimately, the price of A will be driven up enough so that its capitalisation becomes 60% of the market cap. Once security A forms 60% of the market cap and security B is 40% of the market cap, all managers can hug the benchmark and maximise their expected utility.

This is what Buffa, Vayanos, and Woolley call the equilibrium result. It's what they mean when they say their model determines price endogenously. That is, the model is all about the desire of the managers to hug the benchmark, in this process the price of a security is not a given – not determined "exogenously" – but is the result of the benchmarking process itself (that is, endogenous).

This may seem an unrealistic example. If the market is 50% security A and 50% security B, investors are not going to set a 60%/40% mix as a benchmark.

Because of this, Buffa et al introduce another element to their basic assumptions. They assume there is another class of investors – buy-and-hold investors. These investors remove a portion of the market securities from availability for trading. They might be specialised mutual funds or something else. The portfolio of securities they remove does not have the same market cap proportions as the total market. Therefore, the securities that are available to the money managers do not mirror the total market – but the managers are nonetheless measured by their performance relative to the total market.

This does not change the situation from my simplified example. Just as in the simple example, when the managers try to hug the benchmark, they will discover that there is not a sufficient supply of some securities for them all to do it, while other securities are in oversupply. The ones that are in undersupply relative to the benchmark will have their prices driven up (and their expected returns down), while the ones in oversupply will have their prices driven down (and their expected returns up).

#### **MESSING UP THE RISK–RETURN RELATIONSHIP**

The above is the most bare-bones possible explanation of the Buffa et al model. In fact, their model includes much more. In particular, it claims to be able to explain why higher-volatility and higher-beta stocks have been observed to get lower returns.

The argument goes like this. The manager takes more risk of departing from index weights by underweighting high-volatility stocks than by underweighting low-volatility stocks because the high-vol stocks drive more of the volatility of the whole index than the low-vol stocks do. Managers will take special care not to be underweighted in high-volatility stocks, especially those that comprise a large share of the index. This will drive up their price if they are in short supply (that is, if the buy-and-hold investors have overweighted them relative to the index).

The result, conclude Buffa et al is that large-cap, high-volatility stocks will have their prices driven up and their expected returns down. Voila! High volatility will correspond to low return.

As the *Financial Times* article states, "Benchmarking therefore gives managers a huge incentive to stay fully weighted in large, risky securities. This eliminates any risk premium these stocks may have had, and leads to the inversion of risk and return."

Just the revelation that the practice of benchmarking may have as much or more undue influence on market prices as any other factor is enough to make Buffa et al's research give us pause. We should have known of course, but sometimes things that are as plain as the nose on your face do not become plain until you look straight at them.

The Buffa et al model not only determines securities prices endogenously. It also determines the optimal manager fee contract – optimal in the sense that it should be the one agreed upon by manager and client if they have symmetric information. Not only that but if they don't have symmetric information – if there are agency frictions, with the manager being the agent and the client the principal – then it determines an optimal strategy for the manager to shirk. Shirking means that the manager does something to put a little extra in his pocket at the expense of the client, like routing trades to high-commission brokers who give the manager a perk in exchange for the favor, even if it's only free lunches.

## THE VALUE OF THOUGHT EXPERIMENTS FORMULATED IN MATHEMATICS

I am often highly critical of mathematics as it is used in finance. So perhaps it is therefore incumbent on me to explain why this paper is different. The main reason is that it is sufficiently rigorous and well-formulated with clear definitions of all the terms. It can be checked thoroughly, even if with great effort and difficulty. Its assumptions seem as realistic as they can be and still enable the mathematics to be tractable. Of course, if tractability of the maths were the first objective and realism were completely subordinated to that objective

that would be a black mark against it. But this research doesn't exceed an acceptable ratio of unrealistic assumptions to mathematical tractability.

The important thing is that the research is valuable as a contribution to thought experimentation. Ideally, research that relies heavily on mathematics should be understandable in non-mathematical terms. The Buffa et al paper falls short on that measure, but that may be because not enough efforts have been made yet to explain it. The miraculous thing about mathematical formulation is that sometimes you discover real-world relationships by working through the mathematics that you would not have thought of otherwise. This happens more often in physical sciences than in economics, but it can happen in economics too. One should not jump to conclusions until it has all been checked out, the mathematics against the reality, and the reality against the mathematics.

#### WHY BENCHMARK AGAINST A MARKET INDEX ANYWAY?

The *Financial Times* article concludes that the Buffa et al paper makes "a strong case against benchmarking". It then asks, "But what can replace it? In most industries, benchmarking is good practice. If funds are not to benchmark against their peers, or against an index, then how should we judge them?"

A good way to look at this may be to borrow a leaf from the field of industrial engineering. W. Edwards Deming was a statistician and product quality expert who became one of the most important consultants to post-war Japan in its ultimately successful effort to transform a ravaged country producing cheap, low-quality goods into one of the most successful economies in the history of the earth. Deming used an instructional tool he called the Red Bead Game to illustrate the travesty of management trying to make workers produce more products of high quality when the production process is inherently a very noisy process and quality actually varies uncontrollably at random.

"Workers" in the Red Bead Game "produce" a product - white beads - by pushing a spatula underneath a pile of 80% white beads and 20% red beads in a mixing bowl, and drawing out a sample of 50 beads. The spatula has 50 holes in it. A bead falls into each hole when the spatula is pushed under the beads. White beads are regarded as intact products while red beads are defective ones. The Red Bead Game can only be truly appreciated by participating in it or [watching it in action on video](#). The video has six parts, so it would require a significant commitment of time to watch it all, but watching the first part will give you the flavor. It may remind you of client meetings with investment managers.

In the Red Bead Game, management does everything it can to minimise the percentage of the product that employees produce that are red beads and maximize the percentage that are white. They try carrot incentives to minimize red bead production, and they try sticks. Of course, nothing actually works because the production of red beads in this process is random. But because it is random, sometimes the incentives seem to work for a while.

Deming's message was that there's no point wasting effort trying to change the results of an industrial process that are due to common causes – causes that can't be controlled. Effort should only be devoted to controlling the effects of special causes, those which can be controlled. Controlling the special causes can both improve the product and reduce the variability in product quality introduced by the common causes.

In investment management, we know at least one special cause – fees. Statistical tests have conclusively shown that fees lower performance. We're not really sure we know any other special causes. So, all the efforts at performance evaluation are just playing the Red Bead Game. As Buffa et al have shown, playing this game may not only be futile, it could be messing with market prices.

Facing up to these facts squarely might be a good start at reform of the performance evaluation process. One inescapable conclusion is that the practice of evaluating managers by monitoring their performance against an index benchmark should be jettisoned – even if there's no immediately obvious alternative to replace it.



*Michael Edesess, a mathematician and economist, is a visiting fellow with the Centre for Systems Informatics Engineering at City University of Hong Kong, a partner and chief investment officer of Denver-based [Fair Advisors](#) and a research associate at EDHEC–Risk Institute.*

*This article is abridged and reproduced with permission from [Advisor Perspectives](#).*

---