

LUCK VERSUS SKILL

. . .

John H. Cochrane and Tobias J. Moskowitz

We include one paper on performance evaluation, “Luck versus Skill in the Cross-Section of Mutual Fund Returns,” written with Ken French.

Gene’s 1970 essay introduced fund performance evaluation as a test of “semi-strong form” efficiency. If prices already reflect publicly available information, then managers using such information won’t be able to generate special or “abnormal” returns.

One must control for risk (that’s the “normal” part). Funds can earn apparently better returns by leveraging, by investing in higher beta stocks, and so forth. But such strategies don’t require any skill or special information—the investor can borrow and invest in a market index, without paying any fees.

The CAPM inaugurated the alpha-beta performance evaluation to control for risk. Betas measure a portfolio’s exposure to risk and the return the investor can obtain by indexing; alphas measure the average return after correcting for risk and the return that requires skill or special information of the manager. Returns don’t reveal skill or justify fees, alphas do. But measuring alphas and betas, especially in portfolios that are frequently changing, is not as easy as it seems, and this difficulty is especially strong as performance evaluation is extended to hedge funds and other less-stable investments, and to indices beyond the market portfolio.

The most fundamental issue is highlighted in the title of Gene and Ken’s paper: Even if one controls for risk or benchmarks by looking at alphas, many funds will have large alphas just by chance. How do we tell skill from luck? The previous academic answer is straightforward: find a group of funds and then follow them—all of them, even the ones that go out of business—through time.

Overcoming the selection and survivor bias—we are much more likely to have data on ex post successful funds—is also not easy. Nonetheless, correcting for risk and selection bias, as Gene reviewed in 1970 and Gene and Ken remind us here, funds don’t outperform indices on average.

We have gotten so used to this finding that we forget how remarkable it is. It should have come out the other way! In any other field of human endeavor, seasoned professionals systematically outperform amateurs. Tiger Woods will beat you at golf. You should hire a good plumber and a better tax lawyer. Why not hire a stock picker? But other fields are not as ruthlessly competitive, and free to enter, as financial markets.

Here the conversation between academics and practitioners has always run into a rough patch. The practitioner (or want-to-be practitioner) will ask, "Well, what about Warren Buffet? If markets are so efficient, how did he get so rich?" The academic answer "well, maybe he just got lucky" is not satisfying. How do we tell if a single investor, ex post, was skilled or lucky? But pointing out shared ignorance is not an effective rebuttal to the seemingly solid evidence of Mr. Buffet's vast wealth.

Furthermore, the practitioner may admit that the *average* fund doesn't outperform indexing. But surely the "good" funds have skill. But, as reviewed by Gene in 1970, there is a way to study this question: name an ex ante criterion for finding "good" funds, and look at the performance of that whole group, adjusting for risk, over time. There is a bit of fishing danger here too, as one out of twenty selection criteria will have a significant *t*-statistic. Selection criteria such as "invest with guys named Buffett" or "follow all funds in Nebraska" will lead to spurious results. But that's the same problem as all alpha-producing characteristics that we examine for stocks using historical data, as we must, rather than announcing skill indicators and waiting 20 years for the returns to roll in.

This procedure is an example of the clever inversion that Ray Ball describes, and shows quite how much genius had to go between the idea of informational efficiency and teasing out its predictions for markets. Rather than look for the information that made Buffet rich, look at the performance of all investors using a similar strategy, averaging out the luck.

The most natural indicator of skill or information advantage is past performance: look at groups of funds that did well in the past and see if they collectively do well in the future. Mark Carhart's (1997) study of one- and five-year performance persistence stands as a modern classic of this approach. The answer is, past performance tells you essentially nothing about performance going forward.

Still, the practitioner might note that the uncertainty in measuring average returns makes return histories very noisy signals of skill. The one-year return histories that Carhart found most powerful for predicting returns before risk

adjustment are and should be nearly meaningless measures of long-term average performance. So the practitioner might retort, “I know the good funds, you’re just using very noisy signals of who is good.” So far, a large number of other indicators of skill have been tried in the academic literature, such as Morningstar Ratings, education of the manager, and so on, and few have found substantial skill or information advantage. (Inside information does work, but poses legal risks.) But the practitioner, or salesman, trying to convince you that his fund, with an eye-popping return history, really will violate the adage that past performance is no guide to the future, can still claim that the academic literature just isn’t good enough at shopping for the “good” funds to notice them.

THE PAPER

Enter Fama and French. This paper is especially innovative for asking the question in a completely different way, but with the usual “now why didn’t I think of that?” simplicity. Fama and French ask, if funds really have no alpha, what is the chance of seeing funds with good performance? How likely is it in a universe of 6,000 funds, that we see a fund with Buffet’s performance?

They tabulate the distribution of fund alpha t -statistics—alpha / standard error of alpha—to put funds on the same basis despite varying lifespans and tracking errors. Then they ask how many funds should by luck alone produce large alpha t -statistics if the true alphas are all zero. The heart of the paper is a careful simulation approach to this question, which handles cross-correlation and non-normality. However, if you read the tables, the numbers are not far from standard t distributions. About 2.5% of funds should, if there is no skill at all, have alpha t -statistics above about 1.96.

The result: when adjusted with the Fama-French three- or four-factor model, there is a tiny amount of skill in gross returns, before fees. The upper tail of alphas is just a little bit bigger than it should be if there were no skill at all. After fees, there is essentially no positive skill in net returns to investors. And there is a puzzle, in our view: the average and median fund has negative alpha, both gross and net. And the *left* tail is much too big. There are far too many funds with *negative* alpha than pure chance would allow. In an efficient market, investors can’t systematically lose money either, except by lack of diversification or by blowing money on trading costs and fees!

The immense novelty of this paper is that it does not rely on the researcher to come up with a signal of skill. So it answers the practitioner’s challenge directly—no matter how clever you are at finding skill *ex ante*, the number of

funds that have it is quite small. The weakness, from an investor's point of view, is that it does not tell you how to find the skilled funds, if there are any.

THE FUTURE

A half century after it started, performance evaluation remains a hot topic. A whole new class of funds has sprung up to study, including hedge funds, sovereign wealth funds, university endowments, private equity, and institutional proprietary trading. These are more active and spread across more markets. More importantly, the fundamental ideas behind performance evaluation are changing quickly.

Fama and French's paper reports a controversy with Berk and Green (2004), which will, we think, expand. For 50 years, academics have been deploring active management. But how does it survive? If efficient marketers would not allow "irrationality" as an "explanation" for a 50-year pricing puzzle, why do we allow it for a quantity puzzle? And though individuals are finally moving to indexing, supposedly sophisticated investors, including university endowments, are moving to more active and higher-fee management.

Berk and Green's article is, in our eyes, a watershed, because it writes down a simple, coherent supply-and-demand model in which the basic facts of performance evaluation and the larger academic study of active management make some sense. In their model, some managers do have skill, but only at limited scale. Investors move to the skilled managers until returns to investors are no more than investors can achieve by indexing. Managers capture all the skill in fees.

However, the Berk and Green model has not been quantitatively matched to the data. You can see this process beginning in Fama and French's comments. In Berk and Green's model, investors receive zero alpha after fees. In Fama and French's paper, the average fund delivers negative alpha after fees, which they cite as evidence against Berk and Green's model. Berk and van Binsbergen (2013) retort that Fama and French's factors don't include transactions costs, investors didn't know about value and momentum factors in the 1970s, investors could not trade those factors if they did know about them, and we should measure skill by alphas times assets under management not by raw alphas. Ten basis points on a billion dollars is a lot of skill. Using available tradeable factors, they climb back up to zero alpha. The argument has a long way to go.

Negative alphas and harmful fees are a deep puzzle. The fund literature (e.g., Carhart 1997) finds that fees and turnover cut returns to investors one-for-one. While we have repeated this mantra so often it seems natural, reflect a moment

that this result is really puzzlingly bad. The fund manager says, “We charge fees and turnover to achieve our alpha. We then pass on the good returns to our investors.” The most cynical Chicago economist would predict zero correlation between fees, turnover, and returns to investors: investors should get paid their outside opportunity, namely the market index, just as we each pay the same price to the electric company and the cleaners. Funds would chew up all the extra performance in fees, but no more. The fact that fees and turnover lower returns to investors is a huge puzzle even to that most cynical Chicago view. Why don’t the investors just leave? The janitors don’t take a pay cut at high-fee firms. Berk and Green assert the puzzle isn’t there. Clearly, this discussion will continue.

More questions remain. Does the strength of fund flows following performance quantitatively match Berk and Green’s Bayesian learning theory? We’ll see. To our casual eyes, money flows faster than real information about skill in the latest returns. But nobody has checked this question yet.

A deeper intellectual shift is hiding in the transition from CAPM to multifactor models for performance evaluation. When the CAPM was the only model, well, you used the CAPM to risk adjust managers like everything else. Now that multifactor models, following Fama and French’s size and book-to-market models, are the standard, which model should you use to adjust performance? Or should we be using deeper models, like the consumption-based model?

The answer in Fama and French is that the choice of factors for *performance evaluation* reflects much different criteria than, say, a model one would use to decide if booms and busts were “rational” or not. As the best example, Fama and French here use a momentum factor, as did Carhart (1997), despite their clear reservations whether the momentum premium is rational.

Why? Because the question here is, can a manager deliver something that I cannot get by mechanical trading strategies, for which I don’t need to pay high fees? Any “factor” that the investor can trade mechanically or get in a passive, low-fee vehicle goes on the right-hand side in this thinking. If a manager offers a better, more efficient, lower cost of providing value or momentum, that is still valuable. However, net of fees, and benchmarked to a tradeable portfolio, that manager will still show alpha.

But is this logic correct? Adjusting with the market return makes abundant sense, whether or not the CAPM holds, because most investors have access to cheap market indices and have thought through their allocation to the market portfolio. They don’t need to pay for beta through a manager.

But value is alpha to an investor who hasn't thought about value. How many investors have thought about momentum or short-volatility exposure? How many are in any position to trade those "factors," "mechanical" though they appear to us academics, without getting swamped by transactions costs?

As another example, should we benchmark with fixed or time-varying betas? Suppose a manager bought only the S&P 500 ETF, but changed his exposure based on some publicly observable signal, which he is not going to tell you about, and suppose his signal works. Suppose further we could measure his beta *ex post* with high-frequency data. Should we benchmark him to this time-varying beta and conclude he has no alpha? Should we at least worry about time-varying risk exposure? Or should we listen to his sputtering complaints that knowing when to increase beta *is* his alpha, and we should only benchmark to factors that we the investors already understand and trade?

The larger picture painted by the second generation of empirical work in the multifactor context is, in our evaluation, that "alpha," interpreted as pure information that is not incorporated in market prices, is hard to find. But the model of market equilibrium question, raised in Gene's 1970 essay, remains and has grown in importance. There are now many flavors of "beta." The function of most funds is simply to understand betas that investors don't understand or don't know how to trade efficiently. If that is our world, then just how one should benchmark, and the very definition of "skill," changes dramatically.

You can see hints of this fact in Fama and French's appendix tables, where they risk adjust only using the CAPM. Lo and behold, now there is a substantial tail of funds with positive "skill." Who are they? Well, they are value funds of course, perhaps even including Dimensional Fund Advisors (the fund specializing in small and value stocks based on Fama and French's research).

Finally, the "equilibrium accounting" that Fama and French emphasize is a deep puzzle of all active trading. Positive alphas are not really a puzzle. Negative alphas are a puzzle! It is natural to think that professional traders can, by paying attention, find information not incorporated into prices—somebody's got to do it—buy a little low, and sell a little high. In doing so, they take a little money from the "liquidity" trader, as all models of trading specify.

But the average investor must hold the market portfolio. The average alpha, relative to the market portfolio, must be zero. For every winner, there must be a loser. For every seller, there must be a buyer. Beating the market is a completely zero-sum game. (We repeat because practitioner analysis so routinely ignores this simple fact.)

There is a way for any of us to avoid being the liquidity trader with negative alpha: hold only the value-weighted market index, and refuse any offer to trade away from those weights. If we did, traders with information would simply bid prices up to efficient values, with no trading and no profit. Why does anyone agree to be the negative alpha? That remains the basic puzzle of information trading.

In sum, Fama and French give us a deeply innovative analysis and point the way to an ongoing, active area of research, not one that is just repeating 45-year-old arguments.

REFERENCES

- Berk, Jonathan B., and Richard C. Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112, 1269–1295.
- Berk, Jonathan, and Jules H. van Binsbergen. 2013. "Measuring Managerial Skill in the Mutual Fund Industry." Manuscript, Stanford University.
- Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52, 57–82.