NORTHFIELD modeling financial markets worldwide

Three Approaches to Finding Active Manager Skill

Dan diBartolomeo

PMAR Conference 2022

Introduction

- In recent years, there has been a large switch from active management to passive index funds by both institutional and retail investors.
- One plausible justification of this trend is that even if some active managers can persistently add value to investment outcomes, investors are unable to distinguish between active managers with such skills and those without.
- This presentation will provide three different methods of identifying active manager skill and therefore the likelihood of persistent superior risk-adjusted performance in the future.
- All three methods will be available from Northfield and some of our partners. All three can be implemented by institutional investors or consultants as they do not require "inside the manager only" data. Unlike conventional analyses, long track records are not required.

The Persistence Literature

- If markets are very efficient, there should be no persistence patterns in active management returns. While there are there are innumerable studies showing markets are relatively efficient, many fund studies show that some persistence does exist

- Hendricks, Patel and Zeckahuaser (1993)
  - Find positive persistence only over time horizons less than a year
  - Stronger persistence among worst managers who stay worst
- Elton, Gruber, and Blake (1996)
  - Persistence of “risk adjusted” returns over one to three year time horizons
  - Appears to be correlated with investor capital flows
- Goetzmann and Ibbotson (1994)
  - Persistence over one to two-year horizons
  - Effect is stronger for more volatile funds

The Persistence Literature is Persistent

- Carhart (1997)
  - Some persistence over a one year horizon but not longer
  - Investment style and expense ratios explain most persistence effects. No evidence of “stockpicking” skill
- Stewart (1998)
  - Funds that have consistently outperformed the S&P 500 over a screening period also outperform during subsequent periods
  - Consistent performers hold more diversified portfolios
- Brown & Goetzmann (1995)
  - Find performance persistence in mutual fund
  - Superior performance is correlated across managers (style herding)
- Detzel and Weigand (1998)
  - Some persistence in mutual fund returns, but after adjusting for manager investment style, all persistence in returns is explained.

PWER: Analysis of Return Data Only

- Our first method, called “Precision Weighted Excess Returns” (see Bolster, diBartolomeo and Warrick, FMA, 2006) involves only historical return data for the fund being evaluated and for a large set of similar funds needed for purposes of comparison.

- The PWER process involves statistical enhancements to increase the precision and accuracy of traditional evaluation of investment performance.

- We produce PWER scores for every mutual fund in the United States. Analysis of the history indicates statistically significant predictive power.

Problems in the Evaluation of Managers

- Much manager evaluation occurs relative to benchmarks that are often not suitable for the manager's investment approach
- Evaluation of past performance is based on standardized periods (i.e. 5 years) rather than periods that are relevant to the manager in question
- Many evaluation measures such as Sharpe ratio or information ratio correspond to meaningful investor utility for only a small fraction of investors
- The statistical significance of ex-post performance is measured in a simple time series fashion.
  - It does not include the context of whether the manager exists among a tightly bunched set of peers or a widely dispersed set
  - This is critical in examining the "luck versus skill" issue

A Simple Prescription for Success

- Classification
  - Make sure each fund is being measured against the right benchmark and the right peers. We use an augmented method of returns-based style analysis
- Process Control
  - Evaluate each manager over the evaluation period that is the best for that particular manager
  - We use a “Sequential Probability Ratio Test” called CUSUM to find the optimal evaluation period.
- Evaluate Past Performance
  - Use a return measure such as alpha, not the Sharpe Ratio or Information Ratio
  - Use a Bayesian framework adjustment to ex-post alpha to reflect contemporaneous dispersion across managers.
  - Does the CUSUM analysis show improving or declining efficiency?

Manager Classification Issues

- One very large plan sponsor using PWER wanted to break out simplistic classifications
  - No more “large-small”, “growth-value”, “international”
  - Strong desire to allow unconventional strategies that might add value
- Allowing unconventional strategies creates the risk of “gamed” peer comparisons
  - “The best way to win a contest for the largest tomato is to paint a cantaloupe red and hope the judges don’t notice”
- Forming manager peer groups:
  - For conventional managers, using iterated returns based style analysis from diBartolomeo and Witkowski (FAJ, 1997)
  - For unconventional managers, we synthesize a peer group using a form of Monte Carlo simulation from Surz (JOI, 1994)

Time Horizons for Investment Track Records

- Practitioner tradition in the investment industry is to evaluate active manager track records over a long period
  - **At least** 3 to 5 years
  - Some will argue a full “market cycle” is needed

- As we’ve seen, all the academic studies refute this
  - No evidence that long-term past performance is predictive of future performance
  - If there is any meaning to past performance at all, it’s short-lived, perhaps the last year

The Key Question

- What time portion of a track record do we really need to evaluate as part of our monitoring of manager “quality control”
- What we need is a procedure to draw the line between getting enough meaningful data within a manager’s record and older, stale data that should be ignored
- Enter CUSUM
  - Backward looking sequential probability ratio test
  - Created by E.S. Page in 1954
  - Reliably detects small process shifts
  - Insensitive to probability distribution
  - Provably optimal: detects process shifts faster than any other method.
  - Robust, good under almost any definition of optimality
  - Much better than exponentially weighted moving average.

Robust Monitoring Manager Returns: CUSUM

- CUSUM analysis defines key turning points in the active return time series, and defines statistical significance of results subsequent to the turning point
- Use of CUSUM to monitor active managers started with the IBM pension fund
  - Philips, Stein and Yashchin (JPM, 2003)
  - The PSY CUSUM method classifies managers into three categories: Good, We Don't Know, and Bad
  - Managers are reviewed whenever a class boundary is crossed, but is not an automatic "hire/fire" signal
- Our use of CUSUM is different
  - Focus on whether active performance is improving or declining since the last regime change

Now That We Know the "When," Let's Deal with "What"

- Many performance measures are not congruent to adding value for investors
  - deGroot and Plantinga (Journal Perf Meas, 2001)
  - Consider a manager that adds exactly one basis point of return in every time period. The information ratio is infinite, but very little investor wealth is added
- We chose to measure excess return above a carefully constructed benchmark that should reflect both risk and investing style
  - This directly measures added value for investors
  - Our CUSUM analysis is already a variation on information ratios

Separating Luck from Skill

- To maximally exploit our information about manager performance we need to separate skillful managers from the merely lucky

- We need to adjust for the fact that if manager returns are widely dispersed within a peer group, its easier to have a high excess return.
  - If the dispersion of returns is low, its harder

- We adopt a method a Bayesian framework of a “precision weighted” estimate that incorporates information about the dispersion of peer fund returns during the evaluation period for each fund
  - Similar to Shanken and Jones (FMA, 2004)

The Precision Weighted Excess Return Estimate

- Let's assume Manager X has an excess return (A) of 4% per year with a standard deviation (S) of 4%

- Over the same time period, the average peer manager had an annual excess return of .25% (Mean), and the dispersion (CSD) of the excess returns across the peer group is 1.5%

  $E = (A/S^2 + Mean/CSD^2) / (1/S^2 + 1/CSD^2)$

  $A = 4, S = 4, Mean = .25, CSD = 1.5$

  $E\ (precision\ weighted) = (0.361) / (0.5069) = 0.712$

- We assume the manager has skill sufficient to add 71 basis points per year over the benchmark

Empirical Conclusions BDW (2006)

- Large scale tests on three data sets
  - Domestic mutual funds
  - International mutual funds
  - Hedge funds
- The hypothesis that past returns can be used to predict future returns is supported to a degree of virtual statistical certainty
  - Using raw excess returns, the expected excess returns are about 20% of the observed past returns
  - Using precision weighted excess returns, the expected values are over 40% of the past values
- Given the observed dispersion among manager returns, large and economically significant excess returns should be available to investors

PODS: What were managers allowed to do?

- The second method is "Portfolio Opportunity Distributions" (PODs) as first proposed by Surz (JPM, 1994).

- In PODs, we need the historical record of a manager's performance and a description of the limitations that were imposed on the active portfolio (security universe, benchmark, position size limits, liquidity constraints).

- Using a form of Monte Carlos simulation, we can create a broad range of alternative portfolio returns that could have arisen had the manager made different choices but under the same constraints.

- By comparing the realized performance to the distribution of "what if the manager had done things differently" returns, we can quicky make statistically significant evaluations, **including expectations of alpha.**

PODs: Bells and Whistles

- Our implementation of the PODs technique includes a number of mechanical improvements to the simulations.
  - Simulated portfolios can be weighted equally, by capitalization or by square root of capitalization.
  - The number of securities to be included can have an upper bound.
  - Security position sizes can be subjected to constraints either in absolute terms or active weight relative to benchmark.
- Our PODs implementation will eventually be combined with our "Optimized Scenario Analysis" tool to form a statistically significant form of "back-testing".
  - Produce statistically significant results (which conventional back tests do not) because we will have expanded the data exploration in randomizing both in time series (bootstrapping) and in cross-section (comparison portfolios).
  - See Bailey, Borwein, dePrado and Zhu: rnoti-p458.pdf (ams.org)

EIC: Were individual positions optimal?

- We will describe the method known as the "Effective Information Coefficient" (diBartolomeo, JPerfMeas, 2008) wherein we use risk models to estimate the *implied security return values* an active manager must have believed at each moment in time.
- Positions in an active portfolio must reflect manager skill in forecasting returns and in constructing portfolios that can benefit good forecasts.
- By correlating the implied returns with the subsequent realized returns we can evaluate skill inclusive of all constraints. Multiplying through by the cross-sectional volatility (variety) of security returns yields the expected alpha.
- Given that the EIC analysis occurs at the security level (i.e. very large sample size) it achieves statistically significant results quickly but does require full transparency of past portfolio holdings.

EIC = IC * TC

- A manager knows their input alphas and their constraints so they can separately calculate “Information Coefficient” and Transfer Coefficient
  - Grinold (JPM, 1989)
  - Clarke, DeSilva and Thorley (FAJ, 2002)
- If you assume a risk model (e.g. Northfield), then an outside observer (investor, consultant) can compute the product of IC and TC without knowing either item separately.
- You also have to assume how averse the investor is to active risk which you can infer from the portfolio tracking error
  - Derived from Rubinstein’s utility function (JoF, 1976)
  - Practitioner version in our webinar, Estimating an Investor’s Volatility/Return Tradeoff: The Answer is Always Six (northinfo.com).

Capacity Analysis

- As some investors are very large, we want to incorporate capacity limits into capital allocations
  - For background, see Vangelisti (JPM, 2006)

- We adjust manager alpha estimates as a function of capital allocated
  - Market impact of trading will rise decreasing net alpha (k values from Northfield trading cost model)
- $a = S_{i = 1 \text{ to } N} \, a_i w_i (1 - k_i w_i)$

- We frame the capacity upper bound in the form of a liquidation policy:
  - We want limit the amount of capital to the manager such that liquidation of P percent of the manager's portfolio could be accomplished in N days at cost of no more than C %, to begin within T days of the decision

Conclusions

- In today's presentation we have provided three methods available to investors (not just managers) to carefully evaluate the amount of skill associated with a particular active manager.
- We spent most of our discussion on the **Precision Weighted Excess Return** method because it has the least informational requirements. The only data needed are the past returns of the subject manager and a peer group of similar funds.
- The **Portfolio Opportunity Distributions method** allows us to reach statistically significant evaluations of skill but requires knowledge of the constraints under which the subject manager was required to operate.
- The **Effective Information Coefficient** method also allows rapid convergence on statistically significant conclusions by evaluating manager skill at the individual position level rather than portfolio returns.