Fil rouge:

* Experimental designs – need for control
* In finance, BM or peer group
* Example of size
* Alphabet portfolios
* Performance analysis is inherently relative.
* A random portfolio forms an obvious choice of control
* Idea of random portfolio is not new

Inhaltsverzeichnis:

* The need for a control, i.e., the problem with classical performance analysis
* the cross-sectional distribution of investment styles.

%Ideas:

% Distribution of cross-section of alpha's. Even if active investment is a zero-sum game, what's relevant is the distribution and not the average.

% FF run bootstrap on mutual-fund alphas

% alpha is a relative perspective

% good thing about using alpha's: t-test. simpler than Sharpe ratio test (test whether two SR are equal)

% Dartboard papers:

%Bing Liang The "Dartboard" Column: The Pros, the Darts, and the Market %38 Pages Posted: 27 Nov 1996

%B. M. Barber and D. Loeffler 1993. The "Dartboard" Column: Second-Hand Information and Price Pressure. The Journal of Financial and Quantitative Analysis

Vol. 28, No. 2 (Jun., 1993), pp. 273-284

%10.2307/2331290

%J. Greene and S. Smart Liquidity Provision and Noise Trading: Evidence from the “Investment Dartboard” Column 1999 The Journal of Finance 54, 5, 1885-1899

%10.1111/0022-1082.00171

G. E. Metcalf and B. G. Malkiel. 1994. The Wall Street Journal contests: The experts, the darts and the efficient market hypothesis. Applied Financial Economics 4 371-374

Glenn N. Pettengill and John M. Clark Estimating Expected Returns in an Event Study Framework: Evidence from the Dartboard Column. 2001. Quarterly Journal of Business and Economics, Vol. 40, No. 3 3-21

M. R. Thomas and W. I. Ghani. 1996. The dartboard column: Analysts’ earnings forecasts and the information content of recommendations. Journal of Business and Economic Studies 3 33-42.

[What “Smart Beta” Means to Us | Research Affiliates](https://www.researchaffiliates.com/publications/articles/292_what_smart_beta_means_to_us)

thoughts about the nomenclature of Smart Beta

Towers Watson, a leading global investment consulting firm, coined the expression smart beta

Brinson: “is probably worth the cost of trying to outperform”

Etiology:

is the study of the causes, origins, or reasons behind the way that things are, or the way they function, or it can refer to the causes themselves. The word is commonly used in reference to the causes or origins of various [phenomena](https://en.wikipedia.org/wiki/Phenomena).

“extremely unlikely to be a result of sampling variability (luck).”

% Surz:

% Most important, statistical significance is determined much more quickly than with indexes because inferences are drawn in the cross-section rather than across time. In other words, the ranking of actual performance against all possible portfolios is a measure of statistical confidence.

%The ranking within the manager’s customized opportunity set answers the question, Does the manager do it well?

% Most important, statistical significance is determined much more quickly than with indexes because inferences are drawn in the cross-section rather than across time.

% waiting times: it takes many decades to determine the statistical significance of outperforming a benchmark, even a benchmark that is customized to an individual manager.

% Leippold:

%use multi-factor models to understand the potentially different risk and style exposures of active and index funds

%Style describes the exposures of an investment to different sources of risk. The style of a fund can be inferred from past returns or from portfolio holdings. Return-based style analysis uses a constrained regression to estimate the risk exposures to different asset classes. The quality of the analysis crucially hinges on the selection of benchmark indices. An alternative approach is to perform a fundamental analysis of portfolio holdings data. Style benchmarks can be applied to evaluate the performance and to monitor the style consistency of a fund.

%Returns-based style analysis rests on the assumption that a manager’s style determines the shape of her return series, whereas her skill results in a constant or near-constant addition to the returns.

% [1] William F. Sharpe, “Determining a Fund’s Effective Asset Mix,” Investment Management Review, (December 1988), pp. 59 – 69

% [2] William F. Sharpe, “Asset Allocation: Management Style and Performance Measurement,” The Journal Of Portfolio Management, 18 (1992), pp. 7 – 19

% return-based style analysis and customized multi-factor model

%Harvey Liu 2021:\\

%&\cite{bib:KosowskiEtAl2006} and \cite{bib:FamaFrench2010} ask the question do outperforming funds exist? They try to answer it by analyzing the cross-section of alphas. The corresponding null hypothesis is that all funds generate a zero alpha.

%As opposed to “how many funds outperform?” (see, e.g., Barras, Scaillet, and Wermers, 2010, 2020; Ferson and Chen, 2020; and Harvey and Liu, 2018).

%We subtract the estimated alpha from each fund to obtain a pseudo panel of funds that have an in-sample alpha of exactly zero. We then treat this as the return population and resample to generate the cross-section of test statistics (i.e., t-statistics) under the null hypothesis

%The bootstrap allows us to obtain the null (empirical) distribution of a percentile statistic. If this percentile statistic for the actual data is too large to be explained by this null distribution, we reject the null and declare that some fund managers must possess skill. Skill in our context is measured by after fee excess returns.

%Throughout our paper, we also follow KTWW’s and FF’s main specifications and use the Carhart (1997) four-factor model as the benchmark model to risk adjust fund returns.

%Surprisingly, the two studies arrive at differing conclusions even though they employ similar statistical inference methods (bootstrap) applied to similar data. \cite{bib:HarveyLiu2022} explain the difference in detail.

%\cite{bib:KosowskiEtAl2006} suffer from high Type I error rates (falsely saying a fund outperforms) whereas the test setup in \cite{bib:FamaFrench2010} has little to no power.

Garleanu and Pedersen 2019:

The rise of delegated management raises several questions: What is the optimal portfolio of active (i.e., informed) and passive (i.e., uninformed) managers, respectively?

%Me: are you doing the right thing vs are you doing things right?

%Me: in the whole discussion we have to remind ourselves that the benchmark is not the risk-free. E.g., 70% of MSCI World in US. Huge bet.

%Me: Issue with bm-relative analysis is that it doesn’t question the bm itself. If you start with say 10 active strategies which you compare to the bm on a regular basis, most likely all of them will at some point, i.e., over some period, underperform the bm. Removing the losing strategies from the sample, one eventually ends up with just the bm.

%Me: If a relative approach is chosen to discern skill, the bm should be skill-free and a strategy should display a consistent edge over the zero skill bm, i.e., over multiple periods.

%// Compare the distribution of cross-sectional portfolio returns to the distribution of cross-sectional portfolio alpha’s. Depending on the geometry of the investable set, certain investment styles are more represented than others. 🡪 **the cross-sectional distribution of investment styles.**

%--------

%How to generate a set of random investment strategies obeying certain constraints such that one could statistically compare some elaborated strategy with this random control group. It seemed like an obvious thing to do; after all, the idea of using randomised controls is standard in various scientific fields. Nevertheless, the financial literature does not offer any solution to create such an experimental design\footnote{The tenor in finance is to analyse investment strategies by benchmark or peer group comparisons. If random portfolios are used, they very rarely consider constraints.}.

%Experiments are at the heart of any scientific insight.

Any research conducted under scientifically acceptable conditions uses experimental methods. The crux of experimental designs is the inclusion of a control, i.e., a base case where a certain idea is not applied, where a variable of interest is not manipulated, in order to assess the cause-and-effect relationships and to determine the value or validity of the proposed idea.

Such procedure is not part of the standard repertoire in the investment industry. Investment professionals implement all sorts of ideas, but it is rarely clear whether those ideas are indeed the source of their performances. Skill and luck are hard to disentangle.

Rather than by experimental designs, the way performances are evaluated in practice is by comparing the realizations of past profits and losses to those of other managers or products (peer group comparison) or to general market evolution, usually proxied by indices like the S\&P 500 (benchmark comparison). Such comparisons absolutely have their purpose, but this cannot be to declare one strategy better than another because they always come with severe drawbacks and are never without bias. \cite{bib:Surz2006} provide a thorough documentation of the problems with traditional performance evaluation methods.

Classical performance analysis takes a relative perspective. It's tools try to identify sources of excess returns, i.e., relative to a benchmark, and to attribute them to active bets undertaken by the portfolio manager. Any benchmark replicating strategy, since it involves no active deviation from the capitalization-weighted allocation structure in a market, is therefore called passive. Given the relative perspective, skill, if it exists at all, is reserved to active decision makers and the average investor has, by definition, no skill. Further, skill is commonly corrected by factors which are widely accepted to explain security returns and are associated with a positive premium.

Whenever assets are not allocated in the proportion they are held in an index.

Reasons as to why assets in a portfolio should not be held in the proportion they are allocated in an index can be numerous. Regulation may prohibit holding of overly concentrated portfolios (e.g., 5/40/40 rule), the investor’s risk profile asks for a risk-reducing allocation (e.g., minimum variance portfolios), market frictions like tradability issues or sustainability requirements.

return differences coming from discretionary decisions to not hold assets in proportion to as they are held in an index, are adjusted for factors other than the market.

Dividing a manager’s performance into the part resulting from her ability to select the best stocks and stock characteristics (e.g., sector or country belongingness, currency, etc.) and factor exposures (market, value, size, etc.)

This relative perspective holds for both camps of performance measurement schools, i.e., holdings- or transactions based and return-based factor models. Active therefore means deviation from passive, obviously. And since passive is just the average of all investors, active and passive must have the same return, on average and before costs. Whether the trivial arithmetic holds has been the subject of many sophisticated studies.

Skill, under this perspective, is measured in terms of outperformance vis-à-vis the average investor. Did the manager "beat the market"?

match the performance of certain market indexes rather than trying to outperform them

Passive managers simply seek to own all the stocks in a given market index, in the proportion they are held in that index.

Classical performance analysis tools deal with relative return attribution and try to identify sources of excess returns and to relate them to active decisions by the portfolio manager. A more general attribution would explain the sources of return of a portfolio in isolation rather than in comparison with a benchmark.

\textit{add Text on performance analysis methods used in the industry. Holding- or Transaction-based Performance Attribution vs Returns-based (factor) Attribution, typically concducted by means of linear regression (Fama-French-Carhart).}

Random portfolios form an obvious choice of control as by design, they incorporate no penchant to any strategy; any one portfolio structure is just as likely to occur as any other. A random portfolio should represent the set of counterfactual results, i.e., the entirety of performances that could have been achieved over a certain period given a universe of investable assets subject to investors constraints.

In principle, this allows for a description of the probability distribution of cross-sectional performance measures under the null hypothesis of no skill which provides a means to statistically test the benefit of having a certain strategic portfolio feature.

Any approach to portfolio selection deemed skillful should then systematically outperform the bulk of random allocations, at least if outperformance is measured in terms of the objective seeked for by the strategy. Instead of comparing a particular investment strategy to the average investor (i.e., a capitalization weighted market index) or to competitor products, performance analysis is vis-à-vis an exhaustive control. Of course, this is in a first instance a theoretical construct. The practical challenge is to find a good representation.

%In the most general formulation, random portfolios should be the outcome of a uniform sampling on the domain of feasible allocations. The feasibility and complexity of such a task depends on the dimensionality of the asset space and on the shape of the region induced by the constraints. In the following, we will present several solutions and solution procedures that are computationally tractable, each tailored to a specific problem type. Exact uniformity will not always be attainable but this is not per se a problem because our interest lies in a pragmatic application where near-uniformity may be gauged sufficient for certain statistical inference.

Clearly, the idea of creating random portfolios for performance evaluation is not new. Ever since the claim of the economist Burton Malkiel \cite{bib:Malkiel1973} that “a blindfolded monkey throwing darts at a newspaper’s financial pages could select a portfolio that would do just as well as one carefully selected by experts” random portfolios have been promoted to probe investment skill. Most prominently, The Wall Street Journal's Dartboard Contest, a column published by the business newspaper between 1988 and 2002, put Malkiel's claim to the test by letting their staffers (acting as the allegoric monkeys) throw darts at a stock table, while investment experts picked their own stocks. After a holding period of six months, performances were compared. If nothing else, the game added another animal symbolism to the jargon at Wall Street, emblematic in the ongoing debate on active versus passive management and the underlying hypothesis on market efficiency.

Rigorous versions

Alternatively to forming random portfolios bottom-up by sampling weights, a common practice is to synthesize the price or return paths of a hypothetical random portfolio directly. However, such Monte Carlo simulation necessitate an assumption about the underlying stochastic process and thus disqualify as a model-free control.

% Hence, choosing constituent weights randomly

%It is done all the time in the form of Monte Carlo Simulations or, as in case of The Wall Street Journal’s dartboard game\footnote{In 1988, the Wall Street Journal decided to see if the claim of the economist Burton Malkiel \cite{bib:Malkiel1973} that “a blindfolded monkey throwing darts at a newspaper’s financial pages could select a portfolio that would do just as well as one carefully selected by experts” would hold up, and created the Dartboard Contest.

However, both approaches fail to capture the essence and inherit value of random portfolios as a model-free control. Monte Carlo simulations typically synthesize the price or return paths of a hypothetical random portfolio, thus necessitating an assumption about the underlying stochastic process. Although this is overcome by the random stock selection approach chosen by the dartboard game, thus allowing for an agnostic position with respect to (w.r.t.) the data generating process, random stock picking forms no acceptable control for it fails to account for real-world constraints that strategies built from theory typically need to satisfy. This is where previous work on random portfolios fell short and may explain why the great break trough has never materialized.

In what follows we will build upon the deficiencies of the dartboard game in order to devise a proper randomized control.

%However, despite the ubiquitous academic literature on alpha potential, many of which, nota bene, are empirical in nature, the aspect of a proper randomized control is missing across the board.

%Rather than as a random stock picking excercise (where hypothetical amounts are invested equally among the randomly selected firms), random portfolios should be understood as a random sampling from the feasible set of portfolio allocations. And importantly, the stochastically generated portfolios should obey the same constraints as the strategy built from theory.

%Even more should it surprise that the financial literature does not offer a ready solution for anyone wanting to conduct a comparison of an investment approach not only to a benchmark or peer group, but to a set of stochastically generated portfolios that are unbiased by design and contain no investment style, yet respect the same constraints as the strategy in question.

%the game has not met scientific standards, thorough studies have analyzed the contest in detail \cite{bib:BarberLoeffler1993}, \cite{bib:GreeneSmart1999} and

%\textit{cite papers on active management}. \textit{Maybe mention rise of ETF, Vanguard}

%https://econs.online/en/articles/economics/heads-or-tails/

%although their experiments have not always met scientific standards

%the professionals won: their returns were higher in 87 of 142 six-month contests, and in 76 rounds their returns were higher than the Dow Jones Industrial Average. The returns of the professional team's portfolios averaged 10.2\% over the rounds, outperforming the Dow Jones’ rise of 5.6% and the dart throwers’ gain of 3.5%.

%http://www.investorhome.com/darts.htm

%effectively disappears if you (1) account for the fact that the pros pick relatively riskier stocks and (2) measure returns from the day after the column appears (thereby eliminating the announcement effect).

First, close to no financial literature exists treating the mathematical challenges of generating random portfolios, which are in effect considerable. Everyone intuitively understands that random portfolios should obey certain properties like being equally likely to occur and being composed of components that are limited in some way to make economic sense, but it is not a priori clear how one should actually build them. Technically, what we search to do boils down to uniform sampling over some constrained domain, i.e. finding a set of uniformly distributed points in an n-dimensional space bounded by k linear or nonlinear constraints without knowing the boundaries of the solution space. This is a problem in its own right for which no general solution exists. Interestingly, although originally a well established problem in mathematical programming, most recent advances come from the field of bioinformatics and the modelling of metabolic networks. Overlooking the semantic issues….

%Current text treats the mathematical derivation, programmatic implementation and empirical application of randomized algorithms for near-uniform sampling from the feasible set.

%The feasible set is defned as the solution space to a portfolio optimization problem consisting of all possible values of the choice variables that satisfy the problem’s constraints. The complexity of such a sampling procedure depends on the number of variables, the combined size of domains of variables, and the structure formed by the constraints. I will restrict analysis to convex sets that can be expressed as a bounded intersection of a fnite number of closed halfspaces and/or ellipsoids (i.e. linear and quadratic constraints). The fundamental geometric bodies of study are thus polytopes and ellipsoids with the characterising property of being defned in high dimensional Euclidean space.

%After an optimal portfolio is located, remaining stochastic portfolios uniformly spreading the entire feasible set can be used for performance evaluation as by design, they do not display any penchant to any one portfolio strategy. This idea that random portfolios form an optimal control for performance analytics has long been recognized. Yet, to the best of my knowledge, none of the designs found in the literature were setup to uniformly cover the full range of candidate portfolios. Being able to do so would correct the inevitable biases encountered in traditional peer group comparisons and allow for proper hypothesis testing.

**Literature:**

**Performance analysis:**

Holdings- and transaction based:

\cite{bib:Dietz1966} - Dietz’s (1966) Pension Funds: Measuring Investment Performance.

\cite{bib:BAI1968} - The Bank Administration Institute (BAI 1968) report Measuring the Investment Performance of Pension Funds for the Purpose of Inter-Fund Comparison.

\cite{bib:Brinson1985}

\cite{bib:BrinsonEtAl1986}

Many articles have since been written critiquing and defending existing methods and offering refinements and new methods: Multiperiod (cite) and Multicurrency \cite{bib:Allen1991} attributions.

Ankrim and Hensel (1992, 1994)

Karnosky and Singer (1994). Arguably the single most important paper published on attribution in the 1990s

%Menchero, Jose G. 2000. “An Optimized Approach to Linking Attribution Effects over Time.” Journal of Performance Measurement 5 (1): 36–42.

Seminal paper of \cite{bib:Fama1972} suggested breaking down observed return into the part resulting from ability to pick the best securities at a given level of risk and the part that is the result of predictions of general market price movements.

%\cite{bib:FamaFrench1993}, size value and momentum \cite{bib:Carhart1997} and size, value, and two quality factors \cite{bib:FamaFrench2015}.

% \cite{bib:Sharpe1998}

% Eugene F.Fama and Kenneth R.French 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 1, 3-56

% William F. Sharpe 1998. The Journal of Portfolio Management, 18 (2) 7-19; DOI: https://doi.org/10.3905/jpm.1992.409394

% Fama (1972) “Components of Investment Performance”.

%Brinson, Hood, and Beebower’s (1986) “Determinants of Portfolio Performance”.

%Brinson and Fachler (1985).“Measuring Non-US Equity Portfolio Performance”.

% Allen (1991). “Performance Attribution for Global Equity Portfolios”

**Active management:**

By Sharpe's arithmetic of active managment \cite{bib:Sharpe1991}

% W. F. Sharpe 1991. The Arithmetic of Active Management. Financial Analysts Journal. 47, 1, 7-9 10.2469/faj.v47.n1.7

**Dartboard game:**

the game has not met scientific standards, thorough studies have analyzed the contest in detail \cite{bib:BarberLoeffler1993}, \cite{bib:GreeneSmart1999} and

%Bing Liang

%The "Dartboard" Column: The Pros, the Darts, and the Market %38 Pages Posted: 27 Nov 1996

%B. M. Barber and D. Loeffler

%The "Dartboard" Column: Second-Hand Information and %Price Pressure

%The Journal of Financial and Quantitative Analysis

%Vol. 28, No. 2 (Jun., 1993), pp. 273-284

%10.2307/2331290

J. Greene and S. Smart

Liquidity Provision and Noise Trading: Evidence from the “Investment Dartboard” Column 1999. The Journal of Finance 54, 5, 1885-1899. 0.1111/0022-1082.00171