DOI: 10.1111/eufm.12210

WILEY EUROPEAN FINANCIAL MANAGEMENT

ORIGINAL ARTICLE

Manager skill and portfolio size with respect to a benchmark

Andrei Bolshakov[1] | **Ludwig B. Chincarini**[2,3]

[1]Wedge Capital Management, 301 South College Street, Suite 3800, Charlotte, NC 28202
Email: ABolshakov@wedgecapital.com

[2]University of San Francisco School of Management, 101 Howard Street, Suite 500, San Francisco, CA 94105
Email: chincarinil@hotmail.com; lbchincarini@usfca.edu.

[3]United States Commodity Funds, 1999 Harrison Street #1530, Oakland, CA 94612

Abstract

Investment managers often manage a portfolio with respect to a benchmark. Typically, they use a mean-variance optimization framework to maximize the information ratio of their portfolio. We develop an unconventional approach to this question. Given a set of assumptions, we ask what optimal percentage of the benchmark stocks the portfolio manager should select. This optimal portfolio depends on Fisher's and Wallenius's noncentral hypergeometric distributions. We find that the optimal selectivity of a benchmark universe varies from 50% to 80%. These results are provocative, given that many enhanced index portfolio managers select a low percentage of the benchmark universe.

KEYWORDS

enhanced indexing, information ratio, portfolio management, active management

JEL CLASSIFICATIONS

G0, G13

1 | INTRODUCTION

The investment world is broadly divided into active management and passive management. In recent years, there has been a push toward passive management. One can think of portfolio management as being very active or completely passive. For example, a portfolio manager who purchases every stock in the benchmark with the same weights as the benchmark is not providing any security picking ability. On the other hand, a portfolio manager who selects very few stocks with a variety of weighting schemes is creating a portfolio that is very different than the benchmark, perhaps with the ultimate objective of outperforming the benchmark.

Traditionally, we measure a portfolio's distance from pure passive management by the tracking error of the portfolio; that is, by the standard deviation of the difference between the portfolio's and the

We would like to thank Bruce Levin and Richard Gonzalez for useful discussions. We also thank an anonymous referee for valuable suggestions.

benchmark's expected or historical returns.[1] In this paper, we examine the question of what percentage of the benchmark portfolio is optimal to own for an enhanced index portfolio manager, given his skill. We depart from the typical portfolio optimization formulation and consider a simple world in which some stocks will do well and others will not and ask how much of the benchmark universe should be selected. We find that the skilled portfolio manager should select between 50% and 80% of the benchmark universe regardless of his skill level.

Our analysis of the optimal enhanced index selectivity is based on a series of assumptions that in subsequent work might be relaxed. We also bring some interesting new tools to the enhanced asset management framework, including the noncentral hypergeometric distributions of Fisher (FNCH) and Wallenius (WNCH). To our knowledge, this is the first time that these statistical tools have been applied to portfolio selection.

The paper is organized as follows. Section 2 discusses our framework for enhanced index portfolio selection. Section 3 discusses the optimal selectivity of the enhanced index portfolio. Section 4 provides a more general discussion about the assumptions in the model and suggests ideas for future work in this area. Section 5 concludes.

2 | A FRAMEWORK FOR ENHANCED INDEXING

The fundamental law of active management (FLAM; Grinold, 1989) establishes the following relationship between a manager's skill (i.e. his information coefficient, IC), the number of bets he is making (i.e. his breadth, BR) and the resulting information ratio (IR) of a certain strategy:

$$IR \approx IC\sqrt{BR}. \tag{1}$$

There has been disagreement as to the meaning of ''breadth.'' Grinold (1989) did not precisely state what it represented, except that it would be the number of independent bets of the portfolio manager. Different interpretations have been given to the breadth, including the number of relevant factors (Chincarini & Kim, 2007) and the number of stocks in one's universe (Clarke, Silva, & Thorley, 2002).[2] One can imagine breadth as having two components: a cross-sectional dimension (the number of potential bets made at a given time) and a time-series dimension (how many times bets are evaluated during each year or period). Hallerbach (2014) considers breadth along a time dimension in terms of the number of equal size periods in which a manager makes bets per year. He uses the variable N for his time intervals per year (not to be confused with the number of stocks necessarily). Thus, his expression for the fundamental law is $IR \approx IC\sqrt{N}$, where N is the number of betting intervals per year.[3]

The information coefficient is also at times misunderstood and at times requires unrealistic estimations of the correlation between the manager's forecasts and subsequent returns.[4] Some academic studies have been devoted to establishing a different metric for a manager's skill. For example, Constable and Armitage (2006) offered the interpretation of skill as a binomial variable where a manager has

[1] There are problems with this type of analysis that have been documented (Roll, 1992; Jorion, 2003; Bertrand, 2010).

[2] Many papers have extended the discussion of the original FLAM. For example, Qian and Hua (2004) and Ye (2008) study the effect of variation in the information coefficient. Buckle (2004) covers correlated forecasts, Zhou (2008a) focuses on estimation risk, and Lam and Li (2004) and Zhou (2008b) analyze unconditional optimality in the context of the FLAM.

[3] With certain assumptions, it can be shown that N represents the number of securities in one's universe (see Clarke et al., 2002, for a derivation).

[4] Chincarini and Kim (2007) propose one way to interpret the parameters. The most commonly used formulation assumes that IC is the same for all securities, despite being highly unrealistic.

a probability p of picking a winning stock or strategy. The authors construct a binomial tree showing the possibilities as the manager takes new bets at different intervals through time. They then show that the information ratio of the manager depends on the probability or skill in picking a good stock. In the case of a single period with only two possible outcomes (either the manager picks well or the manager picks badly), the information ratio is given by

$$IR = \frac{2p-1}{2\sqrt{p(1-p)}}. \tag{2}$$

Van Loon (2018) extends this relationship further by proposing another variable x which is defined as the expected absolute ratio between the return of a winning stock or strategy and a losing one (to reflect the skewness in the return generating process). With his specific extension, the information ratio can be approximated as[5]

$$IR \approx 2\left(p - \frac{1}{1+x}\right). \tag{3}$$

For modeling purposes, given a value of x, we could choose to normalize it to 1 and adjust p upward so that the overall information ratio according to Eq. (3) remains the same. For example, from an information ratio standpoint the strategy with $p = 0.52$ and $x = 1.2$ would be equivalent to the strategy with $p = 0.56546$ and no skewness ($x = 1$).

With these new definitions of skill, the academic literature has mostly focused on the practical application of the fundamental law by extending one-bet information ratios to multiple bets over time, resulting in the return generating process based on the Cox–Ross–Rubinstein binomial tree (Cox, Ross, & Rubenstein, 1979). For example, Hallerbach (2014) established that if a manager makes N consecutive bets per year, the information ratio of the strategy would be[6]

$$IR \approx \sqrt{\frac{2}{\pi}} IC\sqrt{N} = \sqrt{\frac{2}{\pi}} IR_1\sqrt{N}, \tag{4}$$

where IR_1 is the one-bet information ratio (with one bet, $IR \approx IC$).

Another critical corollary of these binomial-like models is that ultimately the IR is completely independent of the actual fixed return assumption associated with a win or a loss (Constable and Armitage, 2006). Stated differently, p and x are the most important variables that define the ultimate information ratio of an investment strategy. This has been found to be accurate empirically across many asset classes and types of strategies (Van Loon, 2018).

In this paper we attempt to study the existing framework not across many time periods but across many securities in a given universe over just one time period. In other words, we focus more on cross-sectional breadth as opposed to time-series breadth using the fundamental blocks of FLAM, while adopting a very unique approach. We attempt to answer questions of what happens to the overall probability of beating the benchmark and the resulting information ratio (IR) if an investor makes many bets with a certain skill (p) across a universe of stocks and over a single period of time. The question of how many stocks to select into a portfolio from a fixed given universe has an added relevance in today's investment world because of the proliferation of so-called index funds.

[5] To derive this expression, he uses the approximation that $\sqrt{p(1-p)} \approx 0.5$, when p is close to 0.5. When $x = 1$ (i.e. no skewness in returns), then Eq. (3) simplifies to $IR \approx 2(p - 0.5)$.

[6] Hallerbach (2014) also considers cross-sectional variation in terms of the number of markets that the portfolio manager is taking positions in.

2.1 | A Simple Model

The spectrum of active and passive equity investment strategies goes from a highly concentrated selection of individual securities (the most active) to indexing (the most passive). Suppose that an investor is leaning toward the latter but, instead of becoming a pure indexer, this investor is contemplating an enhanced indexing strategy. He believes that he has found a manager who can marginally improve on the performance of a completely passive index by applying a certain skill. The skill in this sense can be thought of as a quantitative model with one or multiple factors, or some intrinsic ability of the manager to select winners across all stocks in the index. The manager controls what proportion of the index is selected for ownership based on the skill mentioned above. In this case, the wise investor might pose the following question: what proportion of stocks selected out of the index results in the best utilization of the manager's skill to outperform the index over some future period?

Traditionally, one would answer this question by optimizing a portfolio with respect to the benchmark and curtailing the tracking error *vis-à-vis* the benchmark to be within a specified range. Thus, a typical enhanced index manager might ask an optimizer to choose a group of securities within the universe so as to maximize the expected excess return of the portfolio over the benchmark subject to a tracking error of, say, 5% per annum. In this paper, we take a different approach in addressing this question.

Suppose we take a universe of N stocks and make four simplifying assumptions about the investment process.

Assumption 1. *Half of the stocks will outperform the index over some period in the future and half will underperform.*

Assumption 2. *The performance of each stock over that period is expressed as a binary event. Either it beats the index (winner) or not (loser). The magnitude of out- or underperformance does not matter.*

Assumption 3. *The skill of the manager lies in having a constant higher probability of picking a winner versus a loser across the universe of stocks.*

Assumption 4. *The benchmark and portfolio are weighted equally.*

These assumptions follow a similar framework of previous academic research discussed above. At this point, we apply a unique set of tools (the FNCH and WNCH distributions) to create the model(s) that will help us answer the important question of portfolio size given manager skill.

Given our assumptions, we can cast all N stocks in the index as stocks of two types: good and bad. According to our first assumption, half of the stocks are good stocks and half are bad stocks. Mathematically, there are $n_g = N/2$ good stocks (winners) and $n_b = N - n_g = N/2$ bad stocks (losers). The skill of the manager in this setting can be expressed as an imparted weight differential. Each good stock has a weight $\omega_1 = \omega$ and each bad stock has a weight $\omega_2 = 1$. The relative picking ability between the two stocks can be expressed as ω. Thus, if the manager has skill at picking good stocks, then $\omega > 1$. The probability of selecting a good stock versus a bad one is proportional to ω. We use the term ω instead of the something easier to understand, like the probability of picking a good stock, because this is consistent with the probability distribution theory that we will use in our model. However, as just stated, there is a natural mapping from ω to the probability that the portfolio manager picks a good stock. But this probability changes as stocks are chosen. Thus, we choose to use ω. In the case of a portfolio manager choosing one stock out of a basket of 50% good and bad stocks, the effective

probability of picking a good stock would be

$$P(X=1)=\frac{n_g}{n_g+\frac{1}{\omega}n_b}=\frac{1}{1+\frac{1}{\omega}\frac{n_b}{n_g}}=\frac{\omega}{1+\omega}.$$

Thus, for ω=1.1, 1.2, and 1.3, the probability of picking a good stock is 0.5238, 0.5455, and 0.5652, respectively. Of course, this will change when the second stock is selected in the portfolio.

2.2 | A bulk selection perspective

We assume that a portfolio manager managing a portfolio of stocks with respect to a benchmark draws all of the stocks in his portfolio, n, at the same time. In this case, the numbers of good and bad stocks in the portfolio manager's portfolio are distributed as a univariate FNCH distribution (Wallenius, 1963; Fog, 2015). The key difference between this distribution and the more commonly known hypergeometric distribution is that the manager has a weight bias in picking the good stocks and the sample is taken simultaneously, rather than each stock being picked sequentially without replacement. The distribution in this particular case is given by

$$\Pr(X=x)=\frac{\binom{n_g}{x}\binom{N-n_g}{n-x}\omega^x}{\sum_{u=\max(0,n-n_b)}^{\min(n_g,n)}\binom{n_g}{u}\binom{N-n_g}{n-u}\omega^u}, \tag{5}$$

where n_g is the number of good stocks (winner stocks) in the universe, n_b is the number of bad stocks (loser stocks) in the universe, n is the number of stocks randomly selected (stocks picked), N is the total number of stocks in the universe, x is the number of good stocks in the selected sample, and ω represents the ratio of the probability of picking a ''good'' stock to that of picking a ''bad'' one. Thus, if $\omega=1$, then the stock picker is equally likely to pick a good stock or bad stock. If $\omega>1$, the stock picker is more likely to pick a good stock.[7] One will also notice that because we assume 50% of the stocks are ''good'' stocks and 50% are ''bad'' stocks, then $N=n_g+n_b=2n_g$ and $n_b=N-n_g$.

The term ω, which is the relative ability of picking good stocks versus bad stocks, can vary. When $\omega=1.1$, this can be interpreted as a slight edge to picking good stocks. For example, suppose one were to pick only one stock from the universe of good and bad stocks. A portfolio manager with $\omega=1.1$ would have a 52.38% probability of picking a good stock.[8] A manager with ω equal to 1.2 or 1.3 would have a 54.54% or 56.52% probability of picking one good stock, respectively.

Let us illustrate this formula with a simple example. Suppose there are 10 stocks in the universe, with 5 good stocks and 5 bad stocks. Suppose that a stock picker is picking a portfolio of five stocks. If this portfolio manager has no ability (i.e. $\omega=1$), then the probability of picking three good stocks (out of five) is given by the hypergeometric distribution and is equal to 39.68%.[9] If the portfolio manager has specific skill, say, $\omega=1.1$, then the probability of getting 3 good stocks is equal to 41.49%, which is naturally a bit higher than in the case of ''no skill.''

Table 1 uses Eq. (5) to calculate the probabilities for different values of good stocks among a selection of five stocks. For example, the numerator for a case of $x=3$ can be calculated easily (i.e.

[7] The original approach to this was first described in Wallenius (1963, Section 5.3).

[8] In Wallenius's original derivation, this probability can be computed as $\Pr(X=1)=\frac{n_g}{n_g+(\frac{\omega^2}{\omega_1})n_b}$.

[9] The hypergeometric probability with no stock-picking skill is given by $\Pr(X=3)=\frac{\binom{n_g}{x}\binom{N-n_g}{n-x}}{\binom{N}{n}}=\frac{\binom{5}{3}\binom{5}{2}}{\binom{10}{5}}=0.3968$.

TABLE 1 A simple example of picking 5 stocks from the noncentral fisher distribution

This table shows a simple example of the noncentral Fisher distribution in the context of portfolio selection. The table shows the probability of selecting $0, 1, \ldots, 5$ good stocks in a portfolio of 5 stocks chosen from a 10-stock benchmark with an equal amount of good and bad stocks.

Number of Good Stocks	Numerator	Denominator	Probability of Event (%)
0	1.00	320.81	0.31
1	27.50	320.81	8.57
2	121.00	320.81	37.72
3	133.10	320.81	41.49
4	36.60	320.81	11.41
5	1.61	320.81	0.50

$\binom{5}{3}\binom{5}{2}\omega^3$).[10] The denominator is simply the sum of all such calculations from $x = 0$ to $x = 5$ (i.e. $\binom{5}{0}\binom{5}{5}\omega^0 + \binom{5}{1}\binom{5}{4}\omega^1 + \binom{5}{2}\binom{5}{3}\omega^2 + \binom{5}{3}\binom{5}{2}\omega^3 + \binom{5}{4}\binom{5}{1}\omega^4 + \binom{5}{5}\binom{5}{0}\omega^5$). Thus, the sum of all rows in the numerator column is the denominator and equal to 320.81. Dividing one by the other, we get the probability of the event of three good stocks as 41.49%.

The probability that the portfolio manager chooses more ''good stocks'' than ''bad stocks'' can be expressed as

$$\Pr(x > n - x) = \frac{\sum_{x=\frac{n}{2}+1}^{n} \binom{n_g}{x}\binom{N-n_g}{n-x}\omega^x}{\sum_{u=\max(0,n-n_b)}^{\min(n_g,n)} \binom{n_g}{u}\binom{N-n_g}{n-u}\omega^u}. \tag{6}$$

In the simple example above the probability would be the sum of three discrete probabilities in the last column of Table 1 for the three cases of a ''win'' (i.e. having more good than bad stocks in the picked sample of five stocks: $x = 3, 4$ or 5).

2.3 | A sequential selection perspective

Another way of considering the portfolio selection process is that the portfolio manager selects each stock sequentially. That is, the portfolio manager observes a benchmark of stocks and picks the ''best'' stock first, then the next ''best'' stock, until he has selected a total of n stocks for his portfolio. In this particular case, the numbers of good and bad stocks in the portfolio manager's portfolio are distributed as a univariate WNCH distribution (Wallenius, 1963; Fog, 2015). With our previously stated assumptions, the WNCH is given by the mass probability function

$$\Pr(X = x) = \binom{n_g}{x}\binom{N-n_g}{n-x}\int_0^1 (1 - t^{(\omega/d)})^x (1 - t^{1/d})^{n-x} dt, \tag{7}$$

where $d = \omega(n_g - x) + (n_b - (n - x))$.

[10] In Excel, one can use the commands =COMBIN(5,3)*COMBIN(5,2)*1.1^3 to calculate the numerator, which gives 133.10.

The calculation of the WNCH probabilities is more difficult due to the path dependency in that the probability of picking a particular stock depends on what stocks have been picked previously. This recursive dependency leads to a difference equation with the integral being the solution.

To demonstrate how the WNCH distribution works, consider the same simple case of 10 stocks that we used for the FNCH. Let us consider the case of finding the probability of obtaining three good stocks from a selection of five stocks. Table 2 shows 10 possible paths that can occur when picking five stocks from 10 stocks (i.e. $\binom{5}{3}$) The probability that the first stock selected is a good stock is $\frac{1.1\times5}{(1.1\times5)+(1\times5)} = 0.5238$. If a good stock is picked first, then the probability that a second good stock is picked will naturally be lower than 0.5238, and is in fact, 0.4681 (i.e. $\frac{1.1\times4}{(1.1\times4)+(1\times5)} = 0.4681$). These probabilities are shown in Table 2 under the path sequences, with a good stock indicated by a 1 and a bad stock by a 0. Naturally, the probability of the second stock being a bad stock is higher than on the first draw at 0.5319 (i.e. $\frac{1\times5}{(1.1\times4)+(1\times5)}$).[11] This particular sequence can be seen in path 2 of Table 2.

TABLE 2 The different possible paths to picking 5 stocks

This table shows the different paths that can occur when selecting five stocks from a universe of 10 stocks. A ''1'' indicates that a good stock has been picked, while a ''0'' indicates a bad stock was picked. There are 10 possible combinations of picking five stocks consisting of three good stocks from a universe of 10 stocks containing an equal number of good and bad stocks. The probability of picking any given stock in the sequence is shown in the second section of the table, below the paths, and the probability of any individual sequence is shown at the bottom of the table.

Path 1	Path 2	Path 3	Path 4	Path 5	Path 6	Path 7	Path 8	Path 9	Path 10
1	1	1	0	0	0	1	0	1	1
1	0	1	1	1	0	0	1	1	0
1	1	0	1	0	1	1	1	0	0
0	1	1	1	1	1	0	0	0	1
0	0	0	0	1	1	1	1	1	1
Probabilities of Individual Picks									
52.38	52.38	52.38	47.62	47.62	47.62	52.38	47.62	52.38	52.38
46.81	53.19	46.81	57.89	57.89	42.11	53.19	57.89	46.81	53.19
39.76	52.38	60.24	52.38	47.62	64.71	52.38	52.38	60.24	47.62
69.44	45.21	45.21	45.21	59.46	59.46	54.79	54.79	54.79	59.46
64.52	64.52	64.52	64.52	52.38	52.38	52.38	52.38	52.38	52.38
Probabilities of Entire Sequence									
4.37	4.26	4.31	4.21	4.09	4.04	4.19	4.14	4.24	4.13

The entire sequence of probabilities can be calculated in a similar fashion and they are shown in Table 2 for this simple example. The probability of any particular path is given by the product of the probabilities of each individual stock being picked along the path. At the bottom of Table 2, we show the product of individual probabilities giving the probability of the particular path. Thus, path 1 has a 4.37% probability of occurring, path 2 has a 4.26% probability of occurring, and so on. To obtain the probability that three good stocks are picked from the WNCH distribution, one must add

[11] The probability that the first stock picked is a bad stock is $\frac{1\times5}{(1.1\times5)+(1\times5)} = 0.4762$.

the probabilities for all possible paths. For this particular example, if one adds the probabilities at the bottom of the table they equal 41.98%.

A very important observation from this simple 10-stock example is that the Wallenius probability of obtaining three good stocks out of five is higher than the Fisher probability (41.98% compared to 41.49%).

We might think of investment managers as following a stock selection process that falls somewhere in between the two extremes: buying simultaneously a big chunk of their investment universe and holding it for the entire investment horizon versus buying one stock at a time followed by the constant re-ranking of the remaining stocks in the universe. That said, one might also think of a traditional active manager as someone who is constantly updating his portfolio in a sequential-like manner and a ''smart beta'' manager as one who buys in a bulk manner. In addition, managers might bulk-select some stocks and sequentially select other stocks. Thus, in practice, the processes would fall somewhere between the FNCH and WNCH distributions.

The FNCH and WNCH probability functions help us to understand the way in which the probability of selecting good stocks is determined; however, to understand the choice mechanism from the perspective of a portfolio manager, we must consider selection in terms of excess return, tracking error, and the information ratio. This analysis is examined in the following section.

3 | THE OPTIMAL SIZE OF THE PORTFOLIO

3.1 | The information ratio approach

For any number of stocks, n, that the manager decides to draw from a universe of N stocks, there will be a different expected number of good stocks and bad stocks. This expected number of good stocks can be obtained from the FNCH when assuming that stocks are bulk-selected and from the WNCH when assuming that the stocks are sequentially selected. This can be normalized by dividing by n to represent the proportion of good stocks as well as the variance or standard deviation of that proportion of good stocks. Given this, we can express the expected return and standard deviation of the portfolio as

$$E(r_{\mathrm{P}}) = p^* r_g + (1 - p^*) r_b, \tag{8}$$

$$S(r_{\mathrm{P}}) = (r_g - r_b)\frac{\sqrt{\sigma_x^2}}{n}, \tag{9}$$

where p^* represents the expectation from the FNCH or WNCH distribution divided by the sample size (i.e. μ_x/n),[12] σ_x^2 represents the variance of the FNCH or WNCH distribution,[13] and r_g and r_b represent the return of the good and bad stocks respectively, which are constants. Thus, the effective probability, p^*, changes with the number of stocks in the benchmark universe.

Proof. This proof is straightforward. Let

$$r_{\mathrm{P}} = \frac{x}{n} r_g + \left(1 - \frac{x}{n}\right) r_b,$$

[12] We use the same notation for the expected value (μ_x) and variance (σ_x^2) of the FNCH distribution as Liao and Rosen (2001).

[13] The normalized standard deviation of the proportion of good stocks is given by $\sqrt{\sigma_x^2}/n$.

where x represents the number of good stocks present in a sample of n stocks, r_g the return for good stocks, and r_b the return for bad stocks. Then

$$E(r_{\mathrm{P}}) = r_g \frac{E(x)}{n} + r_b \left(1 - \frac{E(x)}{n}\right) = p^* r_g + (1 - p^*) r_b.$$

Similarly,

$$S(r_{\mathrm{P}}) = \sqrt{\mathrm{Var}\left[\frac{x}{n} r_g + \left(1 - \frac{x}{n}\right) r_b\right]}.$$

Because r_g and r_b are constants, this is equal to[14]

$$S(r_{\mathrm{P}}) = \sqrt{(r_g - r_b)^2 \frac{\sigma_x^2}{n^2}} = (r_g - r_b) \frac{\sqrt{\sigma_x^2}}{n}.$$

Let us suppose that the benchmark is all of the stocks in the universe equally weighted. Let us also assume that each good stock has a return of μ and each bad stock has a return of $-\mu$. We can use this to compute the average return and standard deviation of any portfolio chosen by the portfolio manager with a given benchmark universe (N) and selectivity ratio. The selectivity ratio is the percentage of stocks chosen by the portfolio manager as a function of all stocks in the universe (i.e. n/N). The expected return and standard deviation of the portfolio are given by

$$E(r_{\mathrm{P}}) = 2\mu(p^* - 0.5), \tag{10}$$

$$S(r_{\mathrm{P}}) = 2\mu \frac{\sqrt{\sigma_x^2}}{n}. \tag{11}$$

In our simplified universe, for a given, n, n_g, n_b, and ω, we can use these equations to study the behavior of the optimal portfolio. One way to evaluate the optimal enhanced index portfolio is to use the information ratio. For the purposes of this paper, we define the information ratio as

$$IR(n/N, N, \omega, n_g, n_b) = \frac{E(r_{\mathrm{P}}) - E(r_{\mathrm{BM}})}{S(r_{\mathrm{P}} - r_{\mathrm{BM}})}, \tag{12}$$

where $E(r_{\mathrm{BM}})$ is the expected return of the benchmark and $S(r_{\mathrm{P}} - r_{\mathrm{BM}})$ is the standard deviation of the difference in returns from the portfolio and the benchmark.[15]

Given our particular assumptions, the equal-weighted benchmark will have an expected return of $2\mu\left(n_g/N - 0.5\right) = 0$. The standard deviation will also be 0. Thus, the information ratio will be given

[14] In an expanded context, one can think of these returns as the expected returns of these type of stocks.

[15] Most portfolio managers measure the *ex post* information ratio as $IR = (\bar{r}_{\mathrm{P}} - \bar{r}_{\mathrm{BM}})/\sigma_x$, where $\bar{r}_{\mathrm{P}}$ is the average return of the portfolio and $\bar{r}_{\mathrm{BM}}$ is the average return of the benchmark over the measurement period. σ_x is the standard deviation of the difference in returns between the portfolio and the benchmark. This is also known as the tracking error. The more accurate measure is different than this measure and described in detail in Chincarini & Kim, (2006, Chapter 15).

by[16]

$$IR(n/N, N, \omega, n_g, n_b) = \frac{E(r_P)}{S(r_P)}. \tag{13}$$

Given our expression for the information ratio, we can calculate the information ratio as a function of several parameters using the computational values extracted for the FNCH and WNCH distributions.[17]

It has been documented that when security returns are nonnormal, the standard deviation of returns is a less precise measure for the risk of a portfolio (Sortino & Price, 1994; Markowitz, 1959). A more appropriate measure focuses on the losses of the portfolio or the downside risk of the portfolio. Thus, we also use another metric to evaluate the optimal portfolio which is the excess return of the portfolio over the benchmark divided by the semi-tracking error or the tracking error only considering the underperformance of the portfolio,

$$IR(n/N, N, \omega, n_g, n_b) = \frac{E(r_P) - E(r_{BM})}{SS(r_P - r_{BM})} \tag{14}$$

where

$$SS(r_P - r_{BM}) = \sqrt{\psi \sum_{i=1}^{n_l} \left[\min\left(0, r_{P,i} - r_{BM}\right)\right]^2 \cdot f(r_{P,i} - r_{BM})}, \tag{15}$$

in which n_l represents the number of possible returns below the benchmark, $f(r_P - r_{BM})$ represents the density function for each possible return, and ψ is a normalizing variable which is equal to the sum of probabilities of the possible returns below the benchmark (i.e. $\sum_{i=1}^{n_l} f(r_{P,i} - r_{BM})$). For the rest of this paper, we will refer to the normal information ratio as IR and the downside information ratio as DIR.

3.2 | Optimal fraction of benchmark with bulk selection

Figure 1 shows the information ratio as function of the selectivity ratio for various values of ω for a universe of 500 stocks (i.e. $N = 500$, $n_g = 250$, $n_b = 250$). The selectivity ratio is shown from 0.1 or 10% (i.e. 50 stocks in the selected portfolio) to 0.9 or 90% (i.e. 450 stocks in the selected portfolio). The figure shows that regardless of the value of ω, the maximum IR is achieved at a selectivity ratio of 50%. That is, if a portfolio manager wishes to achieve the maximum information ratio, he should only select 50% of the stocks in the benchmark universe. For a benchmark with 500 stocks, this would mean choosing 250 stocks.

Different portfolio managers might have different size benchmarks. For example, some managers might have a 500-stock universe (e.g. the S&P 500), while others might have a 3,000-stock universe (e.g. the Russell 3000). The IR will differ depending on the benchmark chosen. Figure 2 shows the IR

[16] In the simplest case, where the portfolio manager is choosing a one-stock portfolio from a universe of two stocks, the expected return and standard deviation equations are given by $E(r_P) = 2\mu(p^* - 0.5)$ and $S(r_P) = 2\mu\sqrt{p^*(1-p^*)}$. Thus, the information ratio is given by $IR = (p^* - 0.5)/\sqrt{p^*(1-p^*)}$.

[17] For this paper, we used the R program BiasedUrn written by Agner Fog. The distributions are also available from the software Mathematica.

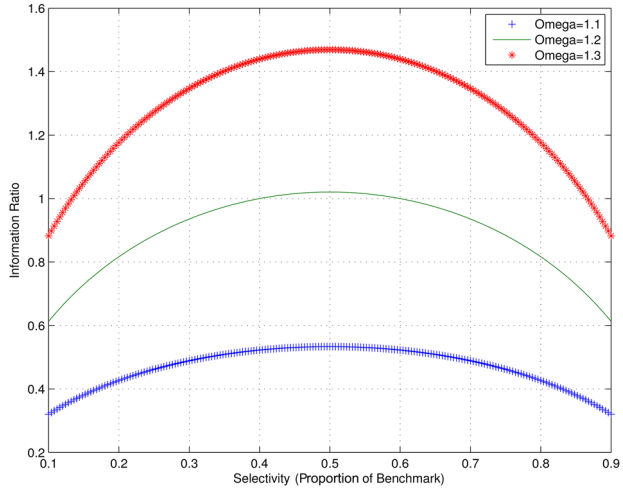

FIGURE 1 Manager selectivity ratio, information ratio, and ω with bulk selection. This figure shows how the information ratio of the portfolio varies as the percentage of stocks chosen from the benchmark increases from 10% to 90%, with $\omega = 1.1,\ 1.2,\ 1.3$.

and selectivity ratio for different sized benchmarks. Although the information ratio is higher for the larger benchmarks, all portfolio managers still have the maximum IR at a selectivity ratio of 50%.

An analytic expression for the ratio of information ratios of different universes is difficult to derive; however, one can get an idea of the relationship by comparing the information ratios for many different universes. In Table 3 we compute the information ratios for various stock-picking abilities, ω, various selectivity ratios, $\phi = n/N$, and various benchmark sizes, N. We compute benchmark sizes from 20 stocks to 5,000 stocks. We also compute the ratio of the information ratios of each one of these to the prior benchmark universe as well as to the square root of their sizes. One can see that the ratio of information ratios from one universe to another is approximately equal to the square root of the ratio

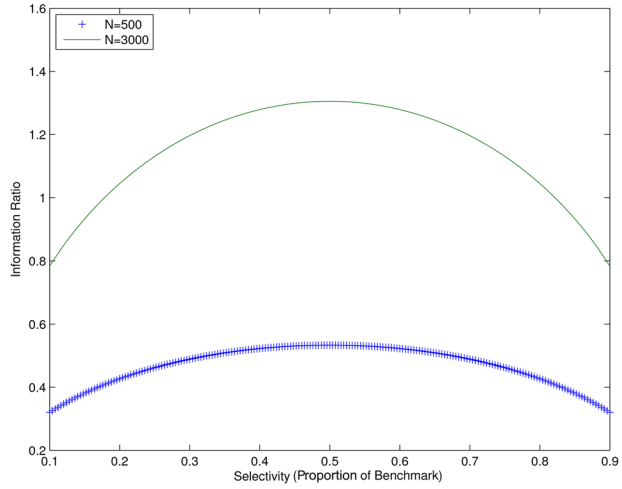

FIGURE 2 Manager selectivity ratio and information ratio with bulk selection. This figure shows how the information ratio of the portfolio varies as the percentage of stocks chosen from the benchmark increases from 10% to 90% for benchmarks of 500 stocks and 3,000 stocks. For this figure, $\omega = 1.1$.

TABLE 3 The information ratio for various parameters and benchmark universes with bulk selection

This table shows the information ratio of the portfolio for the bulk selection method as the size of benchmark universe (N) is varied and the proportion of stocks chosen changes ($\phi = n/N$) for various levels of ω.

N	20	50	100	200	400	1,000	3,000	5,000
$\sqrt{M/N}$		1.58	1.41	1.41	1.41	1.58	1.73	1.29
$\omega = 1.1$								
$\phi = 0.10$	0.07	0.10	0.14	0.20	0.29	0.45	0.78	1.01
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.10	0.16	0.22	0.31	0.44	0.69	1.20	1.54
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.11	0.17	0.24	0.34	0.48	0.75	1.31	1.69
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\omega = 1.3$								
$\phi = 0.10$	0.18	0.28	0.40	0.56	0.79	1.25	2.16	2.79
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.28	0.43	0.60	0.85	1.21	1.90	3.30	4.26
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.30	0.47	0.66	0.93	1.31	2.08	3.60	4.64
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\omega = 1.5$								
$\phi = 0.10$	0.28	0.44	0.61	0.87	1.22	1.93	3.35	4.32
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.43	0.67	0.94	1.32	1.87	2.95	5.10	6.59
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.47	0.73	1.02	1.44	2.03	3.21	5.56	7.18
$IR(M)/IR(N)$		1.56	1.41	1.41	1.41	1.58	1.73	1.29

of the universes. That is,

$$\frac{IR(\phi, M, \omega, M/2, M/2)}{IR(\phi, N, \omega, N/2, N/2)} \approx \frac{\sqrt{M}}{\sqrt{N}}, \tag{16}$$

where $\phi = n/N = m/M$ is the selectivity ratio of the portfolio manager which is chosen to be equal in both benchmark universes.[18]

[18] To get an intuition for this result, one can think of the standard hypergeometric distribution, where an analytical result is available. The mean of the standard hypergeometric is $\frac{nn_g}{N}$ and the standard deviation is

$$\sqrt{\frac{nn_g n_b}{N^2}\frac{N-n}{N-1}}.$$

Substituting for the selectivity ratio of $\phi = n/N = m/M$, for $n_g = n_b = N/2$, results in a normalized information ratio of

$$\frac{\phi}{\sqrt{\frac{\phi(1-\phi)}{N-1}}}.$$

This implies that the ratio of information ratios for a manager that uses a benchmark universe of M versus N, where $M > N$, is given by

We also found an interesting connection between the information ratios of the bulk selection method at 50% selectivity for various N and the respective information ratio for $N = 2$, the case where you just pick one stock (out of two) for the same skill level ω. For example, for $\omega = 1.1$ the probability of picking the first stock as a winner is $p = 0.5238$. The information ratio of that one pick is $IR_1 = (p - 0.5)/\sqrt{p(1-p)} = 0.0477$ according to Eq. (2). In Table 3, you can quickly approximate the information ratios at 50% selectivity using the following FLAM-based formula: $IR = \frac{1}{\sqrt{2}} IR_1 \sqrt{N/2}$. In other words, the maximum information ratio of the bulk selection method at 50% is approximately equal to the FLAM-based information ratio adjusted down by $\frac{1}{\sqrt{2}}$.

3.3 | Optimal fraction of benchmark with sequential selection

The optimal selectivity is more complicated when a manager's selection process is sequential. Figure 3 shows the information ratio as function of the selectivity ratio for various values of ω for a universe of 500 stocks (i.e. $N = 500$, $n_g = 250$, $n_b = 250$) for the Wallenius distribution assuming sequential selection. The selectivity ratio is shown from 0.1 or 10% (i.e. 50 stocks in the selected portfolio) to 0.9 or 90% (i.e. 450 stocks in the selected portfolio).

The figure shows that, unlike the bulk selection method, depending on the value of ω, the maximum IR is achieved at similar, but different selectivity ratios. That is, if a portfolio manager wishes to achieve the maximum information ratio, they should only select 79.6%, 80%, or 80.8% of the universe depending on whether the ω is 1.1, 1.2, or 1.3. In addition, as ω increases, that is, the portfolio manager becomes better and better at picking good stocks, the optimal selectivity moves closer to pure benchmarking (i.e. 100%). For example, at $\omega = 2.1$, the maximum information ratio occurs at a selectivity of 98.4%.[19] As ω increases even further, the selectivity moves closer to 100% benchmarking.

Just as with the bulk selection method, the IR will differ depending on the benchmark chosen. Figure 4 shows the IR and selectivity ratio for different sized benchmarks. The information ratio is higher for the 3,000-stock benchmark at 2.10 versus 0.857, but also the selectivity level is slightly different at 79.8% versus 79.6%, respectively.

Similar to the case of bulk selection, as the benchmark universe gets larger, the information ratio improves approximately as $\sqrt{M}/\sqrt{N}$. The values are shown in Table 4 for the sequential selection case.

Using the downside information ratio, DIR, the results are quite similar, as can be seen in Figure 5. This generally will be true for reasonable values of ω. For unrealistically high values of ω, however, the DIR is the most appropriate metric. We will elaborate further on this in the next subsection.

$$\frac{IR(M)}{IR(N)} = \frac{\sqrt{M-1}}{\sqrt{N-1}}.$$

With this distribution, the normalized mean is constant, yet the standard deviation of the distribution declines as the square root of the sample size, thus the information ratio grows approximately according to the square root of the universe, which, due to the constant selectivity ratio, is also the square root of the sample size.

[19] The probability of choosing a good stock on the first draw is $\frac{\omega}{1+\omega}$. This comes from the formulation that

$$\Pr(X = 1) = \frac{n_g}{n_g + (\frac{\omega_2}{\omega_1})n_b},$$

where $n_g = N/2 = n_b$, and $\omega = \frac{\omega_2}{\omega_1}$. In this particular case, with $\omega = 2.1$, the probability of picking a good stock on the first draw is 67%.

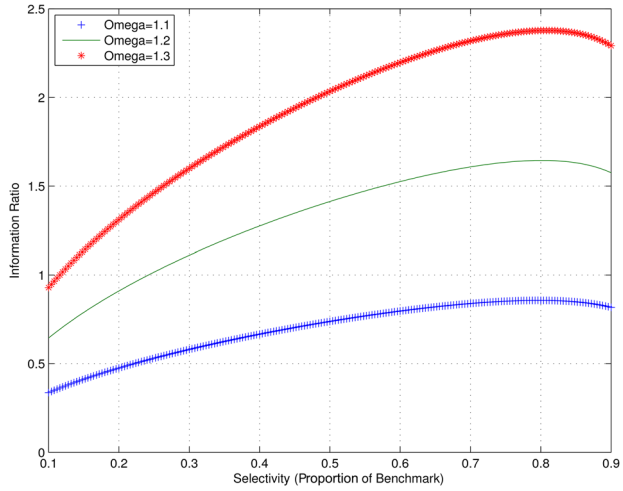

FIGURE 3 Manager selectivity ratio, information ratio, and ω with sequential selection. This figure shows how the information ratio of the portfolio varies as the proportion of stocks chosen from the benchmark increases from 0.1 or 10% to 0.9 or 90%, with $\omega = 1.1,\ 1.2,\ 1.3$.

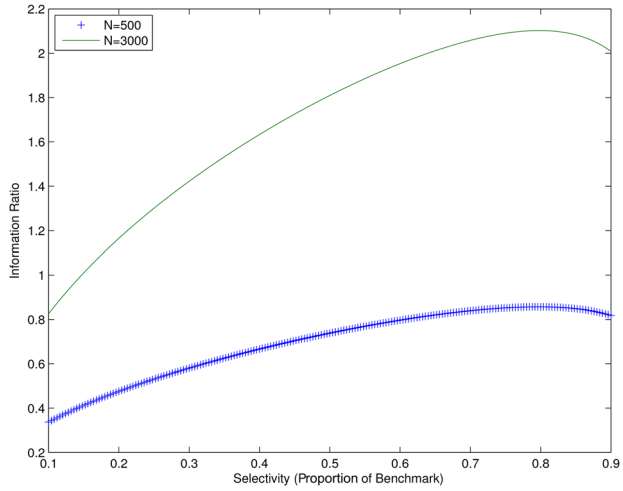

FIGURE 4 Manager selectivity ratio and information ratio with sequential selection. This figure shows how the information ratio of the portfolio varies as the proportion of stocks chosen from the benchmark increases from 0.1 or 10% to 0.9 or 90% for benchmarks of 500 stocks and 3,000 stocks. For this figure, $\omega = 1.1$.

3.4 | Intuition and characteristics about the optimal selectivity

Given our assumptions, we find the following characteristics of our enhanced index framework.

Characteristic 1. *Given a benchmark universe of stocks, N, the highest information ratio for a manager with skill level ω is obtained at a selectivity ratio (n/N) between 50% and 80% (i.e. percentage of stocks chosen from a benchmark universe). For the bulk selection method, it is always at 50%. For the sequential selection method, it is near 80% for reasonable values of ω.*

Characteristic 2. *Given a manager with skill level ω that stays constant as the universe increases, a larger universe, M, will result in a larger information ratio, which is approximately $\sqrt{M/N}$ larger.*

TABLE 4 The information ratio for various parameters and benchmark universes with sequential selection

The table shows the information ratio of the portfolio for the bulk selection method as the size of benchmark universe (N) is varied and the percentage of stocks chosen changes ($\phi = n/N$) for various levels of ω.

N	20	50	100	200	400	1,000	3,000	5,000
$\sqrt{M/N}$		1.58	1.41	1.41	1.41	1.58	1.73	1.29
$\omega = 1.1$								
$\phi = 0.10$	0.07	0.11	0.15	0.21	0.30	0.48	0.83	1.07
$IR(M)/IR(N)$		1.58	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.12	0.18	0.26	0.37	0.52	0.82	1.42	1.84
$IR(M)/IR(N)$		1.58	1.42	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.15	0.23	0.33	0.47	0.66	1.04	1.81	2.34
$IR(M)/IR(N)$		1.59	1.42	1.42	1.41	1.58	1.73	1.29
$\omega = 1.3$								
$\phi = 0.10$	0.19	0.29	0.42	0.59	0.83	1.31	2.28	2.94
$IR(M)/IR(N)$		1.58	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.32	0.50	0.71	1.01	1.43	2.26	3.92	5.06
$IR(M)/IR(N)$		1.58	1.42	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.40	0.64	0.91	1.29	1.82	2.88	4.98	6.43
$IR(M)/IR(N)$		1.59	1.42	1.42	1.41	1.58	1.73	1.29
$\omega = 1.5$								
$\phi = 0.10$	0.29	0.46	0.64	0.91	1.29	2.04	3.53	4.55
$IR(M)/IR(N)$		1.58	1.41	1.41	1.41	1.58	1.73	1.29
$\phi = 0.30$	0.49	0.78	1.11	1.56	2.21	3.50	6.06	7.83
$IR(M)/IR(N)$		1.58	1.42	1.41	1.41	1.58	1.73	1.29
$\phi = 0.50$	0.62	0.99	1.40	1.99	2.81	4.45	7.71	9.96
$IR(M)/IR(N)$		1.59	1.42	1.42	1.41	1.58	1.73	1.29

Characteristic 3. *Given a certain selectivity ratio, the information ratio for the sequential selection method will always be higher than the information ratio for the bulk selection method given a constant skill level, ω.*

Our framework shows that the reasonable selectivity ratio for enhanced indexing varies between 50% and 80%. This is related to the assumptions of our model as well as the way in which managers select stocks: bulk versus sequential selection methods. To develop a better intuition for why bulk and sequential selection methods maximize information ratios at 50% and around 80% respectively, we juxtaposed their expected return and tracking error lines in Figure 6.

Let us look at the bulk selection lines first. The expected return for bulk selection is approximately a straight downward-sloping line.[20] Naturally, the line would start at

[20] The expected return declines as more stocks are selected for both the sequential and bulk methods. The intuition is that when one picks the first stock, given a certain skill, the likelihood of having only one good stock is greater than 50% (given positive skill). However, if one picks 501 stocks (out of a total of 1,000), it is certain that there will be at least one bad stock in the

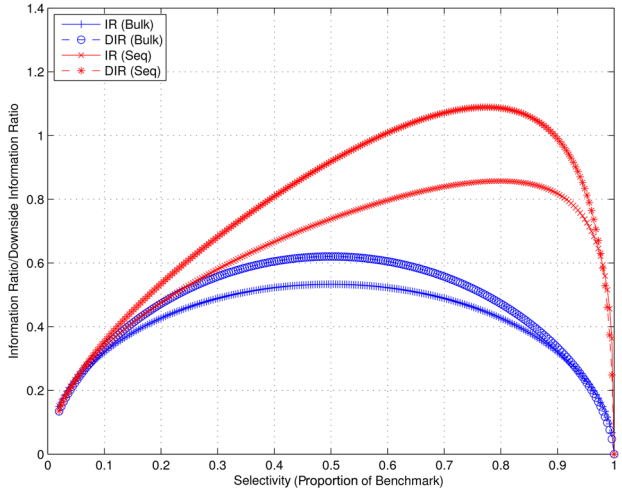

FIGURE 5 Information ratio, downside information ratio and selectivity for sequential and bulk selection. This figure shows how both the information ratio (IR) and the downside information ratio (DIR) vary with selectivity. For this figure, $\omega = 1.1$ and $N = 500$.

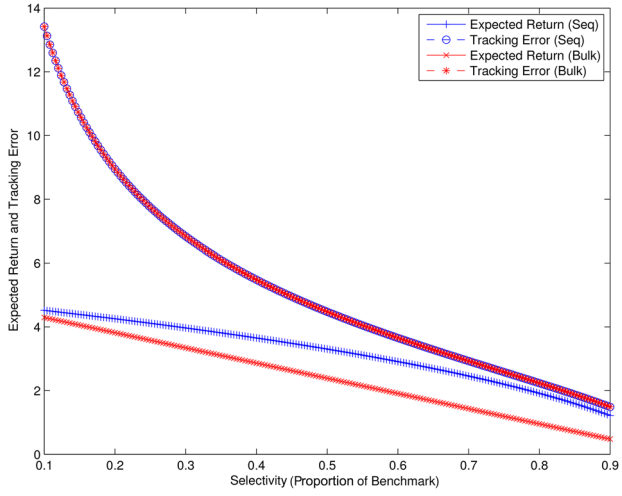

FIGURE 6 Manager selectivity ratio, expected return, and tracking error with sequential and bulk selection. This figure shows how the expected return of the portfolio and the tracking error vary with the proportion of stocks chosen from the benchmark. For this figure, $N = 500$ and $\omega = 1.1$.

4.76%,[21] which is the excess expected return of the first and only stock to be a winner in a one-stock portfolio, and it would end at 0 with 100% of the universe (500 stocks) selected into the portfolio. The other line represents the tracking error (which is extremely close to the tracking error of the sequential selection method). Due to the hypergeometric nature of the distribution,

portfolio regardless of whether one picks sequentially or bulk. Thus, as one selects more stocks, regardless of the method, the probability of having all good stocks in the portfolio declines.

[21] The probability of obtaining a good stock in a one-stock portfolio is 0.5238. Thus, the expected excess return over the benchmark would be $E(r_P - r_{BM}) = 0.5238 - (1 - 0.5238) = 0.0476$. We arbitrarily chose an expected return for good stocks of 1 and for bad stocks of -1. Any other value could be used and would not change the results of our analysis.

the tracking error is decreasing in such a manner that the rate of growth (i.e. deceleration) is similar across the selectivity continuum especially in the 30% to 70% range (0.3 to 0.7 in the figure). For example, for a value of $\omega = 1.1$, the percentage reductions in tracking error between selection intervals 30–40%, 40–50%, 50–60%, 60–70% are 19.81%, 18.35%, 18.35%, 19.84% for the bulk selection method and 19.81%, 18.36%, 18.35%, 19.87% for the sequential selection method.

For the bulk selection method, as one approaches the 50% selectivity point from the left, the percentage decrease in expected return is smaller than the percentage decrease in the tracking error. For example, between 40% and 50% selectivity, the expected return decreases by 16.7% but the tracking error decreases by 18.35%. As a result the information ratio is still increasing. Between 50% and 60% selectivity, expected return decreases by 20.0% but the tracking error decreases by 18.35%. Beyond the 50% selectivity point, the expected return decreases at a relatively faster rate than the tracking error, thus contributing to the decline in the information ratio to the right of the 50% selectivity point.

For the sequential selection method, the analysis is slightly more complicated. The expected return is higher at every point compared to the bulk selection method and the line is concave over selectivity. This helps to explain why the information ratio is larger for the sequential selection method along every selectivity point, given that their tracking errors are similar. Thus, in the sequential selection case, the expected return initially declines at a relatively slower pace and then its decline begins accelerating somewhere past the 70% selectivity point. Its tracking error continues to decline very similarly to the bulk selection method, and hence its information ratio continues to increase beyond the 50% selectivity point, reaching a maximum at around 80%. For reasonable values of ω, the selectivity ratio with the maximum information ratio tends to be similar to those discussed in this section for $\omega = 1.1$.

As the stock-picking skill of the manager increases to very high levels, such as $\omega > 2$, then the behavior of the sequential selection method's highest information ratio converges to a selectivity level of 100%, while the bulk selection method continues to peak at 50%. Although these skill levels are highly unrealistic in practice, the behavior is interesting enough to warrant an example and a discussion. Figure 7 illustrates the situation when $\omega = 10$, which implies that the manager is so skilled that the probability of his picking a good stock on the first draw is 90.91%. The figure shows that as one approaches roughly 60% selectivity, each additional stock pick accelerates the decrease in both the expected return and the tracking error of the portfolio. However, the percentage decline in the tracking error of the portfolio is still relatively greater than that of the decline in the expected return and hence the information ratio becomes larger as one approaches full indexing (i.e. 100% selectivity). This can be seen more easily in Figure 8, which graphs the information ratios of the two methods at this selection ability.

The analysis suggests that as the portfolio manager becomes extremely skilled, the optimal choice for the sequential selection method is perfect indexing. This is a very counterintuitive result. One would think that as the skill increases to very high levels, the manager would engage in very active management. The mathematics of why the optimal choice is pure indexing is simple; the decline in expected return from additional stock picking is not as rapid as the decline in the tracking error. However, it also implies something about using the information ratio. At very high levels of skill, the information ratio is not the appropriate measure since the distribution of returns is not as centered around zero.[22]

To illustrate this concept, Figure 9 shows the density function for the parameters $N = 500$, $n_g = 250$, $n_b = 250$, and a selectivity level of 64.8%. For reasonable values of skill (i.e. $\omega = 1.1$), the distribution of excess returns is more centered around zero; however, for very high values of skill (i.e. $\omega = 10$), the distribution is very far from zero. In fact, in general the distribution of returns is not

[22] The measured skewness also rises and deviates from zero as the ω rises to high levels.

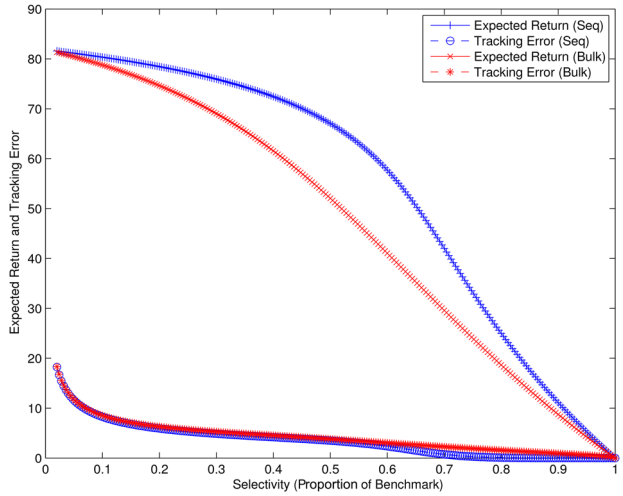

FIGURE 7 Manager selectivity ratio, expected return, and tracking error with sequential and bulk selection. This figure shows how the expected return of the portfolio and the tracking error vary with the proportion of stocks chosen from the benchmark. For this figure, $N = 500$ and $\omega = 10$.

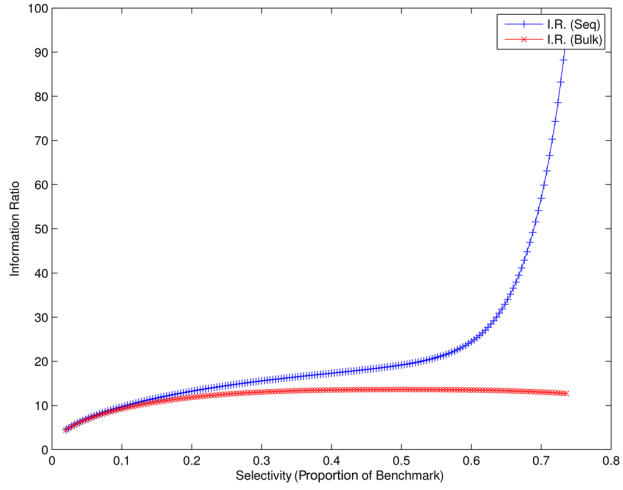

FIGURE 8 Manager selectivity ratio and information ratio for sequential and bulk selection. This figure shows how the information ratio varies with the proportion of stocks chosen from the benchmark. For this figure, $N = 500$ and $\omega = 10$.

centered around zero and suggests using a downside risk measure, such as semi-tracking error. In particular, the cumulative probability of returns being less than the benchmark is 18.03% and 7.30×10^{-75}% respectively. The result is much more severe as ω rises.

At very high ω levels (and for that matter, at all ω levels), the DIR is a better measure of optimality. Figure 10 depicts the IR and DIR for $\omega = 10$ and $\omega = 100$.[23] The figure illustrates that while the IR breaks down, the DIR gives us very intuitive results. As the portfolio manager's skill increases to unrealistically high levels, the optimal portfolio converges from around 80% selectivity towards

[23] The acute reader will note that there are some missing values for DIR at certain return levels. This is because the probabilities of any draw other than one specific draw are essentially zero. Thus, there is no variation in returns and the DIR is undefined.

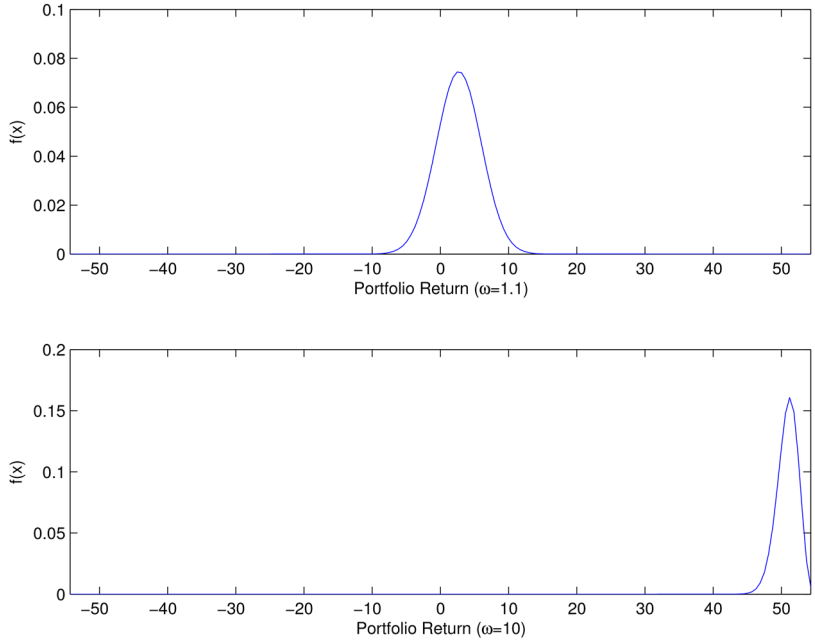

FIGURE 9 **Probability density function of excess returns for the sequential selection method for $\omega = 1.1$ and $\omega = 10$.**

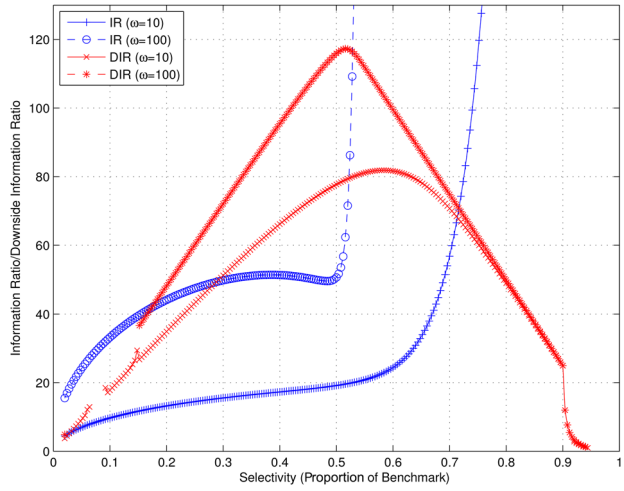

FIGURE 10 **Information ratio (IR), downside information ratio (DIR) and selectivity for sequential selection for $\omega = 10$, $\omega = 100$, and $N = 500$.**

just slightly above 50% selectivity. This makes intuitive sense. An extremely skilled manager will be able to pick all the best stocks by just choosing around 50% of the stocks. Another important feature our research highlights is that there are return distributions where the downside risk measures can be very useful when standard risk measures break down. This is important, since several studies have emphasized that standard risk measures, such as the information ratio, are sufficiently adequate versus downside risk measures (Eling & Schuhmacher, 2007).

As already mentioned, one of the consequences of our analysis is that the sequential selection method is marginally superior to the bulk selection method from an information ratio and downside

information ratio perspective. In reality, investment managers would always find themselves following a stock selection process that falls somewhere in between the two extremes: buying simultaneously a big chunk of their investment universe and holding it for the entire investment horizon versus buying one stock at a time followed by the constant re-ranking of the remaining stocks in the universe. Thus, in practice, the processes would fall somewhere between the FNCH and WNCH distributions. However, if our assumptions are valid, it would be beneficial for an investment manager to shift the process as much as possible towards the sequential selection method, which in turn would result in both a higher optimal selectivity and a higher information ratio.

4 | DISCUSSION

The purpose of this paper is to propose a different way of determining how to manage an enhanced index portfolio. In order to do this, we made certain assumptions about how the investment universe works. One of our assumptions was that 50% of the stocks in the index would outperform and 50% would underperform in a given period. This assumption is simplistic, and does not take into account the positive skewness of stock returns, since the upside potential of a stock is unlimited while the downside loss of a stock is limited. To the extent that the difference between the number of outperforming and underperforming stocks is not too far from 50%, the gist of our main conclusions would be similar.

Our second assumption was that the magnitude of the returns is not important. This is surely not a fair assumption, since a very good stock pick with a very high return can afford a manager many bad picks of lower-magnitude poor returns. This symmetry is unrealistic, but we hope that further work can investigate how much the results change when more realistic distributions are allowed to occur. For example, one might simulate returns based on the actual historical distribution of an individual stock return.[24] Having said this, our assumptions are not much different than existing assumptions in the literature; for example, Constable and Armitage (2006) make a similar assumption for their basic model derivations.

Our third assumption is that a portfolio manager knows his ability to pick good stocks over bad stocks and that this ability remains constant with the number of stocks selected in the portfolio. This assumption is also strong, since one might imagine that as a portfolio manager selects more stocks, his ability to pick winners declines. That is, it is easier to pick a few really good stocks than a lot of really good stocks. A modification to the work in this paper would be to create a function of ω that declines with the number of stocks in the portfolio. Another potentially useful approach would be to take the average ω over a selection range and use that as the appropriate ω for the analysis.[25] Having said this, much of the research in this area has used similar assumptions. That is, the probability of picking winners stays constant (Constable & Armitage, 2006; Van Loon, 2018; Hallerbach, 2014).[26]

Our fourth assumption is that the benchmark is an equal-weighted benchmark. This is not a strong assumption, in the sense that many benchmarks are equally weighted. Also, this analysis could be modified to consider the implications when the benchmark is market-cap weighted or any other type of weighting scheme. It would simply make the analysis slightly more complicated.

Finally, we constructed our entire analysis without considering the variance–covariance matrix of stock returns. This seems very unrealistic, even though it was important to think about the selectivity

[24] One could also think about modeling a really good stock as the equivalent of many stocks in an expanded version of our model.

[25] An expanded version of our model might also describe the implications and dynamics of such a formulation.

[26] In addition, the original FLAM has many extreme assumptions, such as that all the bets by the portfolio manager on stocks are independent and not related to each other (Grinold, 1989). Buckle (2004) and others have tried to address some of these unrealistic assumptions.

issue in a new light. Having said that, one can think about the selectivity issue in the following way. Imagine that a portfolio manager must pick an optimal portfolio and hold it until maturity. Ignoring shorting and leverage, the manager really does not care about fluctuations until the portfolio liquidation day at some time in the future. In that sense, on liquidation day, it may be that about 50% of the stocks will turn out to be winners and 50% will turn out to be losers. In some sense, that will be the only thing that matters and intertemporal volatility may have really not been a concern.

Although this is a theoretical paper, we believe future empirical work can be examined. Take for example, a quantitative enhanced portfolio manager who uses a model to select winners and losers from a universe of securities. In a typical factor ranking, the manager might rank stocks by the factor into deciles. It is likely that, conditional on a factor, the probability of a stock being a winner is higher, the higher its ratio.[27] However, one can relate this to our current theory by imagining that the ranking method will be such that stocks with a z-score greater than 0 will have a probability of being a winner that is greater than 50%, while a stock with a z-score less than 0 will have a probability of being a winner that is less than 50%. Thus, the following experiment could be performed. Rank stocks in every portfolio formation period (either through a one-time bulk selection method or by accumulating the portfolio gradually over several quarters, more similarly to the sequential selection method) and compute the performance of a portfolio that holds 10%, 20%, 50%, and 70% of the universe according to this ranking method. Then compare the information ratio of the different selection criteria and formation methods out-of-sample for various horizons, such as 1, 3, or 5 years. Of course, there are many variants to this empirical strategy that will equally shed light on our theory.

5 | CONCLUSION

Many portfolio managers manage portfolios in the enhanced index universe. Their goal is to choose stocks from a universe of stocks in the index in order to outperform the index while minimizing the risk, where risk is typically defined as deviation from benchmark returns. The traditional method to solve this problem is to maximize the *ex ante* information ratio of the portfolio using some forecast of expected returns and the variance–covariance matrix of stock returns. In this paper, we propose a simplified model of investing to ask how many stocks the portfolio manager should choose when managing the portfolio, or alternatively what percentage of the stocks in the investment universe should the portfolio manager select. We call this the optimal selectivity of the enhanced index portfolio.

In order to complete our analysis, we introduced two useful tools to study the enhanced index portfolio analysis that, to our knowledge, have never been applied to this problem. The first tool is Fisher's noncentral hypergeometric distribution, and the second tool is Wallenius's noncentral hypergeometric distribution.

When deciding on the optimal selectivity for an enhanced index portfolio, we found that the optimal selectivity is somewhere between 50% and 80% for reasonable parameter values using the information ratio as a selection criterion. For the bulk selection method, regardless of how many stocks are in the benchmark universe, the optimal selectivity is at 50%. However, as the investment universe increases, the information ratio of the portfolio increases. In particular, the information ratio increases at a rate $\sqrt{M/N}$, where M is the number of stocks in the larger benchmark and N is the number of stocks in the smaller benchmark. These results are related to the law of large numbers, in the sense that as the benchmark universe becomes larger, but the skill level of the manager remains the same (i.e. ω is a constant regardless of N), then the accuracy of outperformance becomes more precise (i.e. lower standard deviation of portfolio returns for a given size portfolio). For the sequential selection method,

[27] For example, if the factor model says that returns of stocks are related to the factor as $r_{it} = \alpha + \beta z_{it} + \epsilon_t$, then among a universe of stocks the probability of a stock having a winner return (i.e. $r_{it} - \bar{r}) > 0$ will be larger for stocks with a higher z_{it}.

although the maximum information ratio is not always at 80% and does vary, the square-root law is also true. Also for high levels of skill, the appropriate metric for evaluating the optimal portfolio is the downside information ratio.

In this paper, we analyzed the enhanced index portfolio selection process employing a simplified model of investing, with the hope that it provides a new way of looking at the enhanced investment selectivity decision. We also hope that further research sheds more light on this approach, including simulation work using real stock market data and theoretical work that relaxes some of our assumptions.

REFERENCES

Bertrand, P. (2010). Another look at portfolio optimization under tracking-error constraints. *Financial Analysts Journal*, *66*(3), 78–90.

Buckle, D. (2004). How to calculate breadth: An evolution of the fundamental law of active portfolio management. *Journal of Asset Management 4*(6), 393–405.

Chincarini, L. B., & Kim, D. (2006). *Quantitative Equity Portfolio Management. An Active Approach to Portfolio Construction and Management*. New York: McGraw-Hill.

Chincarini, L. B., & Kim, D. (2007). Another look at the information ratio. *Journal of Asset Management*, *8*, 284–295.

Clarke, R., de Silva H., & Thorley S. (2002). Portfolio constraints and the fundamental law of active management. *Financial Analysts Journal*, *58*, 48–66.

Constable, N., & Armitage, J. (2006). Information ratios and batting averages. *Financial Analyst Journal*, *62*, 24–31.

Cox, J. C., Ross, S. A., & Rubinstein M. (1997). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*(3), 229–263.

Eling, M., & Schuhmacher F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance*, *31*, 2632–2647.

Fog, A. (2013). *Biased urn theory*. *Working Paper*, November 11. Retrieved from https://bit.ly/2Dc19xV.

Grinold, R. C. (1989). The fundamental law of active management. *Journal of Portfolio Management*, *15*, 30–37.

Hallerbach, W. G. (2014). On the expected performance of market timing strategies. *Journal of Portfolio Management*,*40*, 42–51.

Jorion, P. (2003). Portfolio optimization with tracking-error constraints. *Financial Analysts Journal*, *59*(5), 70–82.

Lam, K., & Li, W. (2004). Is the ‘‘perfect’’ timing strategy truly perfect? *Review of Quantitative Finance and Accounting*, *22*, 39–51.

Liao, J. G., & Rosen, O. (2001). Fast and stable algorithms for computing and sampling from the noncentral hypergeometric distribution. *American Statistician*, *55*, 366–369.

Markowitz, H. M. (1959). *Portfolio Selection*. New York: Wiley.

Qian, E., & Hua, R. (2004). Active risk and information ratio. *Journal of Investment Management*, *2*(3).

Roll, R. (1992). A mean-variance analysis of tracking error. *Journal of Portfolio Management*, *18*, 13–22.

Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *Journal of Investing*, *3*, 59–64.

Van Loon, R. J. M. (2018). Timing versus sizing skill in the investment process. *Journal of Portfolio Management*, *44*, 25–32.

Wallenius, K. T. (1963). *Biased sampling: The noncentral hypergeometric probability distribution* (Technical Reort No. 70). Department of Statistics, Stanford University.

Ye, J. (2008). How variation in signal quality affects performance. *Financial Analysts Journal*, *64*, 48–61.

Zhou, G. (2008a). On the fundamental law of active portfolio management: What happens if our estimates are wrong?. *Journal of Portfolio Management*, *34*(4), 26–33.

Zhou, G. (2008b). On the fundamental law of active portfolio management: How to make conditional investments unconditionally optimal. *Journal of Portfolio Management*, *35*(1), 12–21.

How to cite this article: Bolshakov A, Chincarini LB. Manager skill and portfolio size with respect to a benchmark. *Eur Financ Manag*. 2019;1–22. https://doi.org/10.1111/eufm.12210