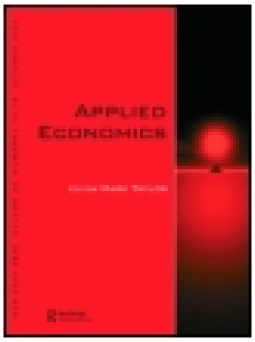

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: http://www.tandfonline.com/loi/raec20

An alternative approach for portfolio performance evaluation: enabling fund evaluation relative to peer group via Malkiel's monkey

Yongjae Lee, Do-Gyun Kwon, Woo Chang Kim & Frank J. Fabozzi

APPLIED ECONOMICS, 2018
https://doi.org/10.1080/00036846.2018.1444263

An alternative approach for portfolio performance evaluation: enabling fund evaluation relative to peer group via Malkiel's monkey

Yongjae Lee[a], Do-Gyun Kwon[b], Woo Chang Kim[b] and Frank J. Fabozzi[c]

[a]Department of Transdisciplinary Studies, Seoul National University, Suwon, Republic of Korea; [b]Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea; [c]EDHEC Business School, Nice, France

ABSTRACT
We propose an implementable portfolio performance evaluation procedure that compares a portfolio with respect to the portfolios constructed by an infinite number of Malkiel's blindfolded monkeys, or equivalently the whole enumeration of all possible portfolios. We argue that this approach exhibits two main advantages. First, it does not require any benchmark portfolios because a portfolio is being compared to an infinite number of portfolios. Second, it is market condition invariant. Since the market conditions are already reflected in the portfolio performances of an infinite blindfolded monkeys, our measure of portfolio performances is invariant to volatile market conditions.

KEYWORDS
Portfolio performance evaluation; blindfolded monkey; random portfolio; Sharpe ratio

JEL CLASSIFICATION
C60; G11

1. Introduction

Although their play styles are widely different, Stephen Curry's 2015–16 Warriors are often compared with Michael Jordan's 1995–96 Bulls. With more than 70 wins, both teams dominated the league. Nevertheless, when it comes to the comparisons of two teams from different era, some argue that the 95–96 Bulls were better as they won the NBA finals. In addition, it is often told that the entire league was better in the '90s. But this argument is without criticism. The 15–16 Warriors certainly broke the Bull's regular season record with only single digit losses, thus some argue that the Warriors should get more credit. Obviously, such debates will continue. The only thing that everyone can agree upon is that these two teams performed the best for their respective season. They competed with other teams, and ended the season with the most wins. Putting it differently, they performed better than anyone in their peer group.

Now let's change our scope to the main topic of this article: fund performance evaluation. Consider a portfolio manager who achieved a −10% return with a Sharpe ratio of −0.4 in the U.S. equity market during 2008. During the year, the U.S. equity market registered a return and Sharpe ratio of about −40% and −1.4, respectively. Conventional portfolio performance evaluation is to determine whether the investment manager outperforms the established benchmark. From this perspective, one can certainly conclude that the manager has performed better than the market.

However, can we tell if the manager dominated the league during the 2008 season like the 95–96 Bulls or the 15–16 Warriors? Or was the manager's performance slightly above the average? Since the return and the Sharpe ratio are merely cardinal numbers, they do not provide a complete picture of fund performance. In order to assess *how* good the fund performance was, we need information on the peer group.

Obviously, there is a very simple way to answer this question. If we have the performance information of *all portfolios*, one can rank any fund among them, just like the NBA season standings. Unfortunately, because there are uncountably many portfolios, it would not be an option to get such information by actually enumerating an infinite number of possible portfolios. Of course, one can collect active fund performance information within the relevant asset universe, and use it as the basis for comparison. In fact, this is a common practice in the

CONTACT Woo Chang Kim wkim@kaist.ac.kr Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

active money management industry. For instance, prime brokers publish fund performance reports containing ranking information.

But there are two critical issues with this approach. First, collecting active fund performance data is not a trivial task, and sometimes the data are not available at all. More importantly, the set of active funds is not necessarily representative of *all portfolios* in the relevant asset universe.

The objective of this study is to propose a novel approach for portfolio performance evaluation that allows us to obtain the relative ranking information without collecting the peer fund group data. More specifically, we want to determine if the performance of a portfolio was at the top 1%, 10%, or bottom 5%, without performance data of *any* other funds. We achieve this task by revisiting the claim of Malkiel (1973). In his best-selling book *A Random Walk Down Wall Street*, Malkiel stated 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'. This statement was based on his belief that stock price movements are basically random walks, so that not only blindfolded monkeys but also fund managers would be unable to predict future stock prices. He thought both monkeys and fund managers would underperform the cap-weighted index.

Even though *the monkey metaphor* may seem provocative, it would be worthwhile to carefully address the concept of comparing fund managers to blindfolded monkeys. The infinite monkey theorem, which also uses the monkey metaphor, states that if there are an infinite number of randomly typing monkeys, one of them will almost surely type an exact version of William Shakespeare's Hamlet. Similarly, if there are infinitely many dart-throwing monkeys, they would construct *all portfolios* in the asset universe. Thus, if we could arrange for an infinite number of dart-throwing monkeys and put the monkey metaphor aside, we could evaluate the performance of a portfolio with respect to the whole enumeration of *all possible portfolios*.

To this end, we employ the uniformly distributed random portfolio (UDRP) methodology developed by Kim and Lee (2016), which can be regarded as an analytical generalization of infinite blindfolded monkeys. We show that the performance distribution of UDRP is equivalent to or better than that of actual active funds. This justifies the use of UDRP as the basis for comparing fund relative performance.

The portfolio performance evaluation methodology that we propose has the following two advantages. First, because a portfolio is evaluated among the whole enumeration of all possible portfolios generated by infinite blindfolded monkeys, it is benchmark-free. Hence, even though asset classes such as fixed-income and commodities are not provided with well-established benchmarks (e.g. cap-weighted index for equities), the proposed method can be easily applied to evaluate portfolios constructed with these investments. Note that infinite blindfolded monkeys can be generated within any compact set of investments. Second, the proposed methodology enables comparisons of portfolios in different asset universes. The proposed portfolio performance measure being market condition invariant implies that, in some sense, the measure is invariant to the market. As the enumeration of all possible portfolios would reflect the market's basic characteristics and the condition at that time, the proposed methodology, with some caution, can conduct a fair comparison between, say, a fixed-income portfolio and an equity portfolio.

While the proposed approach is intuitive and easily implementable, it has certain limitations. Fund managers are subject to various practical matters such as regional and sectoral weight upper/lower bounds, tracking error and turnover constraints, rebalancing frequency, and the like. However, it is difficult to incorporate these constraints in the UDRP methodology since its analytical derivation assumes nothing about the construction of portfolios except the no-shorting constraint. Instead, the UDRP does not require any specific peer information, and, therefore, it measures the performance of fund managers from a broad point of view. On the other hand, Ardia and Boudt (2018) also perform peer evaluation but with actual peer fund information. They were able to carefully address more complicated issues such as false discoveries as pointed out by Barras, Scaillet, and Wermers (2010).

The rest of this article is organized as follows. Section 2 discusses in detail the methodology of evaluating portfolios with respect to an infinite number of blindfolded monkeys. In Section 3, we show how the proposed methodology can be employed to thoroughly verify the conjecture of Malkiel. Section 4 concludes the study.

2. Portfolio performance evaluation with infinite blindfolded monkeys

The portfolio performance evaluation methodology that we propose starts by generating the whole enumeration of all possible portfolios, or equivalently an infinite number of blindfolded monkeys. Then, the performance ranking of a portfolio can be easily derived by ranking the portfolio among infinite blindfolded monkeys. Although analysing the performance of blindfolded monkeys, Arnott et al. (2013) argue that 'it would be time-consuming and costly to arrange for a monkey to throw darts at the *Wall Street Journal*'s stock pages, not to mention tracking down 50 years of their archived stock lists'. Consequently, they simulated the monkeys instead of letting actual live monkeys (or people) throw darts.

However, their simulation approach can also be the subject of their own argument, when it is compared to an analytical derivation. Recall the aforementioned infinite monkey theorem and assume that we are simulating a large population of randomly typing monkeys. Then, how long would it take until one of them actually comes out with a piece of Shakespeare? According to an experiment, only a few word matches could be found even after generating 10^{35} number of pages.[1] On the other hand, we can compute the whole distribution, even including a monkey that types the whole text of Shakespeare's Hamlet, within a second, once provided with analytical expressions.

We first introduce the methodology developed by Kim and Lee (2016), which can be regarded as an analytical generalization of dart-throwing blindfolded monkeys, in Section 2.1. Then, we present in Section 2.2 the portfolio performance evaluation procedure using infinite blindfolded monkeys, and further discuss in detail the issues related to each step.

2.1 An analytical generalization of blindfolded monkeys

Kim and Lee (2016) proposed the concept of a UDRP, which is basically an n-dimensional random vector, or a random portfolio on n number of risky assets, uniformly distributed on an n-dimensional unit hypersphere. They showed that these portfolios on a unit hypersphere represent all feasible portfolios in terms of their Sharpe ratios and analytically derived the Sharpe ratio distribution of all feasible portfolios with a minimal relaxation.[2] For a portfolio $w_s \in \mathbb{R}^n$ with Sharpe ratio s, the probability of w_s to outperform the UDRP in the Sharpe ratio is as follows.

$$\begin{aligned}&\mathbb{P}(w_s \text{ to outperform the UDRP})\\ &= \begin{cases} 1 - \frac{1}{2} I_{\sin^2\theta_s}\left(\frac{n-1}{2}, \frac{1}{2}\right) & \text{if } s \geq 0,\\ \frac{1}{2} I_{\sin^2\theta_s}\left(\frac{n-1}{2}, \frac{1}{2}\right) & \text{else,}\end{cases}\end{aligned}$$

where $I_x(a, b)$ is the regularized incomplete beta function, and θ_s is the angle between $L_\Sigma^T w_s$ and $L_\Sigma^T w^*$. Here, w^* is the optimal tangent portfolio with the maximal Sharpe ratio, and L_Σ^T is a Cholesky decomposition of the covariance matrix Σ of n risky assets. Figure 1 shows the Sharpe ratio distributions of the UDRP with different numbers of assets calculated from the above formula.

Kim and Lee further incorporated the no-shorting constraint to the UDRP to make it represent a more practically sound set of portfolios. Note that the ordinary UDRP is totally indifferent to whether portfolio weights are positive or negative. In practice, however, short positions cannot be taken as freely as long positions. Therefore, we see this non-negative version of the UDRP is better to demonstrate blindfolded monkeys than the ordinary UDRP.[3]

Note that the monkey in Arnott et al. (2013) randomly picked 30 stocks and constructed an equally weighted portfolio on them, whereas the UDRP can pick any number of stocks and the portfolio weights are also completely random. Further, Arnott et al. (2013) simulated dartboard portfolios only 100 times, while the UDRP method deals with an infinite number of random portfolios. Thus, we may well consider the UDRP as an analytical generalization of Malkiel's blindfolded monkeys.

[1]A website 'The Monkey Shakespeare Simulator' actually simulated a large population of randomly typing monkeys and reported that it took 2,737,850 million billion billion billion monkey years to reach a 24 character-long matching sequence from *Henry IV, Part 2*.
[2]They assumed that the portfolios are uniformly distributed on an n-dimensional unit hypersphere in a L_Σ^T-transformed space, instead of the original space. (L_Σ^T denotes a Cholesky decomposition of the covariance matrix Σ of n risky assets. In other words, $\Sigma = L_\Sigma L_\Sigma^T$).
[3]See the Appendix for details about the non-negative UDRP.

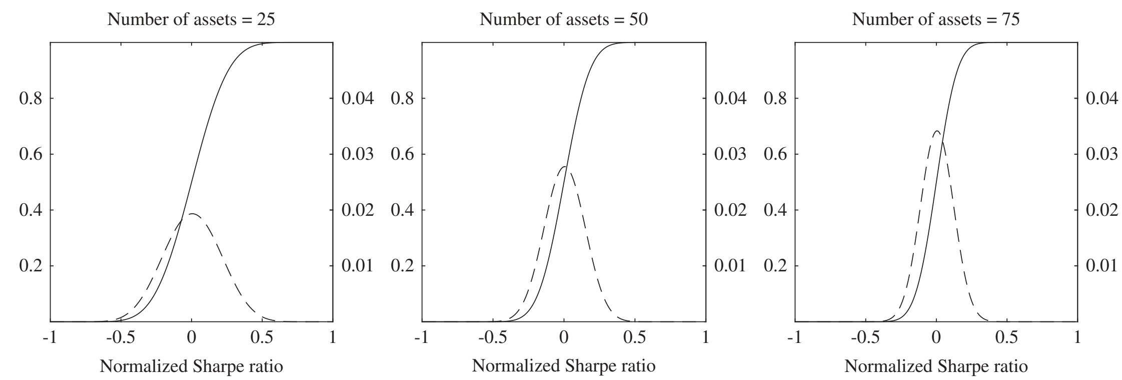

Figure 1. Probability density function (dotted line, right *y*-axis) and cumulative density function (real line, left *y*-axis) of the normalized Sharpe ratio of the UDRP: Normalized Sharpe ratio denotes the Sharpe ratio of a portfolio divided by that of the maximal Sharpe ratio that can be achieved within the market.

2.2 Implementation procedure

Portfolio performance evaluation with infinite blindfolded monkeys can be implemented by using the following three steps:

Step 1. Identify the target asset universe

Step 2. Choose a proxy for the target asset universe

Step 3. Calculate the performance ranking via the UDRP formula

In Step 1, a target asset universe of a portfolio needs to be identified. Of course, a target asset universe should at least include the assets held by the portfolio. For instance, if we wish to evaluate a portfolio of a manager who invests in the U.S. large-cap equities, the target universe for this portfolio should at least be the U.S. large-cap equities. However, one may wish to evaluate the portfolio within a larger asset universe, say, the entire U.S. equity market, depending on the purpose of the performance evaluation. If the target asset universe is set as the U.S. large-cap equities only, then the evaluation excludes the possibility of investments in other asset classes. On the other hand, if the entire U.S. equities are chosen as the target asset universe, then the opportunity costs of not investing in the U.S. mid- or small-cap equities are incorporated into the performance evaluation.

In order to generate infinite blindfolded monkeys we need to estimate the expected excess return and the covariance matrix (μ, Σ) of the target asset universe during the evaluation period. This is Step 2. If the target asset universe is the entire U.S. equity market and the evaluation period is year 2015, we need to estimate the expected excess returns and the covariance matrix of all stock in the U.S. equity market in 2015. Of course, this is not a trivial task because there are thousands of stocks in the U.S. equity market. Instead, one can divide the target asset universe into a number of sub-asset classes and use them as a proxy for the target asset universe.[4] For example, the style portfolios or the industry portfolios provided by Kenneth French through his website[5] would be a decent proxy for the U.S. equity market.

Now we are all ready for the performance evaluation, Step 3. The performance ranking of a portfolio among infinite blindfolded monkeys can be calculated by the non-negative UDRP formula given in the Appendix with inputs of the Sharpe ratio of the portfolio to be evaluated and the expected excess returns and the covariance matrix of the target asset universe.

3. Numerical example – U.S. equity mutual funds versus blindfolded monkeys

In this section, we analyse the performance of the U.S. equity mutual funds to provide a comprehensive illustration of how the proposed methodology can be used to evaluate portfolio performance. Note that the intuition for the proposed performance measure came from the claim of Malkiel, and evaluating the U.S. equity mutual funds using the proposed method is equivalent to

[4]The results of Kim, Lee, and Lee (2014) indicate that the Sharpe ratio range of a sub-portfolio proxy is smaller than that of the original target asset universe. Thus, some portfolios might have exceptionally high (or low) Sharpe ratios that exceed the Sharpe ratio range of the proxy asset universe. But, one can easily handle these cases by automatically setting the maximum (or minimum) performance ranking to such portfolios.

[5]Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

directly proving (or disproving) his claim. Specifically, we analyse four datasets, all the U.S. equity mutual funds, and large-cap, mid-cap, and small-cap equity mutual funds, in order to compare the differences in distributional behaviours of mutual fund performances depending on their investment styles. We further test the performance of mutual funds by changing the levels of mutual fund fees to verify how the results are affected by fees.

Even though there is overwhelming empirical evidence reported in the literature about the performance of the U.S. equity mutual funds,[6] our new approach would be able to provide a new perspective. Existing studies could only conduct *points-to-points* comparison analysis, that is, they selected a few points that represented the performance distributions of mutual funds (e.g. average or median), and compared them to the performance of another point that represented the market, such as the cap-weighted index. Even though they had distributions of mutual fund performance, they had to sacrifice the distributional characteristics of the mutual fund industry due to the absence of a full-distributional counterpart. Moreover, results of points-to-points comparison analyses may be vulnerable to the conditions of the benchmark. Note further that Kim and Lee (2016) indicated that the cap-weighted portfolio exhibits large variabilities in historical performance ranking among the enumeration of all possible portfolios. Therefore, as our *distribution-to-distribution* comparison analysis could incorporate the distributional characteristics of mutual funds, its conclusion should be more robust and comprehensive than points-to-points comparison analyses.

3.1 Data description

U.S. equity mutual fund data were collected from the Center for Research in Security Prices (CRSP) survivorship-bias free mutual fund database. We retrieved monthly returns of all mutual funds that are classified as 'domestic' and 'equity' from 1999 to 2014. Table 1 shows the number of U.S. equity mutual funds available from the database. The first column 'Total' represents all the data retrieved, and the other three columns represent U.S. equity mutual funds that stated their strategies as 'Large-cap', 'Mid-cap', and 'Small-cap', respectively. Because there are many other U.S. equity mutual funds that employ many different strategies, the numbers in the first column are not equal to the sum of the numbers in the other columns.

Table 1. Number of U.S. equity mutual funds available from the CRSP database.

	Number of mutual funds			
Years	Total	Large-cap	Mid-cap	Small-cap
1999	6,237	152	641	1,045
2000	7,506	193	766	1,162
2001	8,373	222	897	1,289
2002	9,009	223	1,015	1,400
2003	9,257	223	1,051	1,476
2004	9,404	211	1,097	1,519
2005	10,022	202	1,202	1,606
2006	10,643	200	1,269	1,694
2007	11,409	201	1,341	1,757
2008	13,999	258	1,565	2,006
2009	14,211	250	1,556	1,961
2010	13,897	233	1,407	1,842
2011	14,273	234	1,394	1,828
2012	14,470	232	1,383	1,797
2013	14,652	225	1,346	1,815
2014	13,557	194	1,234	1,675

In order to generate a UDRP that represents the U.S. equity market, we used Kenneth French's 100 style portfolios[7] formed on the market cap and book-to-market ratio, instead of thousands of individual stocks because it is difficult to estimate parameters, such as covariance matrix, of a large number of securities. These 100 style portfolios would be an appropriate proxy for generating a UDRP that represents the U.S. equity market, as they cover all NYSE, AMEX, and NASDAQ stocks. For fund style-based analyses, we divide 100 style portfolios into the top 0–30% (large-cap), 30–70% (mid-cap), and 70–100% (small-cap) in terms of their market capitalization to generate style-based UDRPs. Thus, for example, U.S. large-cap equity mutual funds are analysed with respect to the UDRP generated using the means and covariances of the top 30 portfolios, in terms of their market capitalization, among the 100 style portfolios.

3.2 Experiments

In order to investigate Malkiel's conjecture thoroughly, we checked the existence of stochastic dominance between the Sharpe ratio distributions of the U.S. equity

[6] See, for example, Treynor (1965), Sharpe (1966), Jensen (1968), Malkiel (1995), Carhart (1997), Wermers (2000), and Fama and French (2010). Berk and Van Binsbergen (2015) provide a review of mutual fund performance but their conclusion is different from most of the literature; that is, they argue that managers do have skills.

[7] Source: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html .

mutual funds and blindfolded monkeys (i.e. the UDRP). Stochastic dominance basically determines a partial order between two random variables. Thus, if we regard the U.S. equity mutual funds and blindfolded monkeys as two entities that produce random Sharpe ratios based on their own probability distributions, we now can literally determine a *stochastic ordering* between the two entities by examining the existence of stochastic dominance between the two probability distributions.

The qth order stochastic dominance is defined as follows. Consider two random variables X_A and X_B. X_A is qth order stochastically dominant over X_B if

$$\mathcal{I}^q[f_{X_A}(x)] \leq \mathcal{I}^q[f_{X_B}(x)], \text{ for all } x \in (-\infty, \infty),$$

where $\mathcal{I}$ is the integration operator, and f_X is the probability density function of X. Note that a lower-order dominance implies a higher-order dominance. Thus, the lower the order is, the stronger the dominance. We check until the second order to determine whether there is a stochastic dominance between two entities, as higher-order dominances do not exhibit many practical implications.

Experiments are conducted as follows. For four investment universes (total, large-cap, mid-cap, and small-cap) from year 1999 to 2014,

Step 1: Calculate the Sharpe ratios of mutual funds

Step 2: Calculate the Sharpe ratio distribution of blindfolded monkeys (non-negative UDRP) (using the means and covariances of style portfolios)

Step 3: Check the existence of first- or second-order stochastic dominance between the two Sharpe ratio distributions

Step 4: Repeat Steps 1–3 with various mutual fund fee adjustments[8]

(−5%, −4%, −3%, −2%, −1%, 0%, 1%, 2%)

The following subsections present the results of the experiments above. First, we discuss whether or not mutual funds outperform blindfolded monkeys after fees. Second, we further investigate mutual fund performances against blindfolded monkeys to see how the after-fee results are affected by mutual fund fees.

3.2.1 Are the fund managers better than blindfolded monkeys after fees?

Table 2 presents the existence of first- or second-order stochastic dominance between the Sharpe ratio distributions of the U.S. equity mutual funds and blindfolded monkeys. We marked '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys, and '0' if there is no entity that dominates the other.

Table 2 shows that there is no stochastic ordering between the two subjects in most cases (68.75–87.50%, depending on the investment universe). In other words, risk-adjusted performances of mutual funds are indistinguishable from those of blindfolded monkeys in most cases. Although the results slightly vary depending on the investment style universe, the number of monkey-dominating cases (11) is more than double the number of the opposite cases (5). Most importantly, when the all the mutual fund data are employed, blindfolded monkeys dominate mutual funds in more than 30% of the cases, whereas there are no opposite cases. Hence, these results clearly show that mutual funds do not systemically outperform blindfolded monkeys after fees. Arguably, mutual funds appear to be doing worse than the monkeys after fees, except for when the analysis is within the large-cap investments. However, we note that the results for the large-cap investments may not be as solid as the others because the number of large-cap funds is quite small compared to other fund styles.

Table 2. Stochastic dominance between the U.S. equity mutual funds and blindfolded monkeys (after-fees).

Year	Total	Large-cap	Mid-cap	Small-cap
1999	1	−1	0	1
2000	1	0	0	0
2001	1	0	1	1
2002	1	0	1	0
2003	0	0	0	1
2004	0	0	0	0
2005	0	0	0	0
2006	0	0	0	1
2007	0	0	0	0
2008	1	0	0	0
2009	0	0	0	0
2010	0	0	0	0
2011	0	0	0	0
2012	0	0	−1	−1
2013	0	0	0	0
2014	0	−1	−1	0
UDRP	31.25%	0.00%	12.50%	25.00%
MF	0.00%	12.50%	12.50%	6.25%
None	68.75%	87.50%	75.00%	68.75%

The existence of dominance is represented as '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys and '0' if there is no dominance. The last three rows represent the proportions of each case. In other words, 'UDRP', 'MF', and 'None' show the percentages of monkey-dominating, fund-dominating, and no-dominance cases, respectively.

[8]We adjust mutual fund fees by simply adding −5% to 2% annual returns to mutual fund after-fee returns. For example, we reduce mutual fund fees by 1% by adding −1% annual returns to all mutual fund after-fee returns. Even though it does not precisely incorporate various fee structures, it is an effective way to see the effect of fee adjustments.

3.2.2 Are the fund managers better than blindfolded monkeys before fees?

Now we look at whether the conclusion in the previous subsection is reversed when the effects of mutual fund fees are considered. As described at the beginning of Section 3.2, we investigated the effects of mutual fund fee levels by simply adding or subtracting 1–5% of annual returns to all mutual fund returns. For example, we simulate the effect of a 1% mutual fund fee reduction by adding 1% of the annual returns to the mutual fund returns. Even though this approach cannot incorporate all different fund fee structures, we believe that it would well simulate the overall effects of various levels of fee adjustments on fund performances.

Stochastic dominances between all the U.S. equity mutual funds and blindfolded monkeys with various levels of mutual fund fee adjustments are shown in Table 3. The results in the column with 0% fee adjustment are the same as those presented in the 'Total' column of Table 2. Because the average of U.S. mutual fund fees are said to be around 2% (see Khorana, Servaes, and Tufano 2009), it would be reasonable to consider that −1% to −3% fee adjustments would demonstrate before-fee results.

While the dominance results are almost unaffected by the 1% fee reduction, almost all stochastic orderings between mutual funds and random portfolios disappear when mutual fund fees are reduced by 2–3%. That is, mutual funds and blindfolded monkeys are not outperforming each other before fees. It is, however, interesting to note that mutual funds do not dominate monkeys with further reductions in mutual fund fees, whereas monkeys become largely dominant with fee increments. Therefore, the overall results indicate that blindfolded monkeys are slightly dominant over mutual funds even before fees, although there is almost no actual before-fee dominance between the two.

The test results of large-cap, mid-cap, and small-cap mutual funds are presented in Tables 4, 5, and 6, respectively. It can be seen from the tables that the no-dominance case is still dominant (56.25–81.25%) around −1% to −3% fee adjustments for all fund styles. In these cases, however, fund managers do dominate blindfolded monkeys in around 12.50–43.75% of the total cases, whereas there were no fund-dominating cases when the whole equity market was considered. In this analysis, therefore, the skills of fund managers become more prominent when the target investment universe gets more specific. Therefore, we may argue from these results that it would become harder for fund managers to beat the market when larger investment universes are considered, or equivalently, a market should consist of a sufficiently large number of assets in order for the market to be efficient.

Table 3. Stochastic dominance between the U.S. equity mutual funds (total) and blindfolded monkeys with mutual fund fee adjustments.

	Mutual fund fee adjustment							
Year	−5%	−4%	−3%	−2%	−1%	0%	1%	2%
1999	0	0	0	0	1	1	1	1
2000	0	0	0	0	1	1	1	1
2001	0	0	0	1	1	1	1	1
2002	0	0	0	0	1	1	1	1
2003	0	0	0	0	0	0	0	1
2004	0	0	0	0	0	0	1	1
2005	0	0	0	0	0	0	0	1
2006	0	0	0	0	0	0	0	1
2007	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	1	1	1
2009	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	1	1
2011	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0
2013	0	0	0	0	0	0	1	1
2014	0	0	0	0	0	0	0	0
UDRP	0.00%	0.00%	0.00%	6.25%	25.00%	31.25%	50.00%	68.75%
MF	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
None	100.00%	100.00%	100.00%	93.75%	75.00%	68.75%	50.00%	31.25%

The existence of dominance is represented as '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys and '0' if there is no dominance. The last three rows represent the proportions of each case. In other words, 'UDRP', 'MF', and 'None' show the percentages of monkey-dominating, fund-dominating, and no-dominance cases, respectively.

Table 4. Stochastic dominance between the U.S. equity mutual funds (large-cap) and blindfolded monkeys with mutual fund fee adjustments.

	Mutual fund fee adjustment							
Year	−5%	−4%	−3%	−2%	−1%	0%	1%	2%
1999	−1	−1	−1	−1	−1	−1	0	0
2000	0	0	0	0	0	0	0	0
2001	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	1
2003	0	0	0	0	0	0	0	0
2004	−1	−1	0	0	0	0	0	0
2005	−1	−1	0	0	0	0	0	0
2006	−1	−1	−1	−1	0	0	0	0
2007	−1	−1	−1	0	0	0	0	0
2008	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0
2010	−1	−1	0	0	0	0	0	0
2011	−1	−1	−1	−1	−1	0	0	0
2012	−1	−1	−1	−1	−1	0	0	0
2013	−1	−1	−1	−1	0	0	0	0
2014	−1	−1	−1	−1	−1	−1	−1	−1
UDRP	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.25%
MF	62.50%	62.50%	43.75%	37.50%	25.00%	12.50%	6.25%	6.25%
None	37.50%	37.50%	56.25%	62.50%	75.00%	87.50%	93.75%	87.50%

The existence of dominance is represented as '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys and '0' if there is no dominance. The last three rows represent the proportions of each case. In other words, 'UDRP', 'MF', and 'None' show the percentages of monkey-dominating, fund-dominating, and no-dominance cases, respectively.

Table 5. Stochastic dominance between the U.S. equity mutual funds (mid-cap) and blindfolded monkeys with mutual fund fee adjustments.

	Mutual fund fee adjustment							
Year	−5%	−4%	−3%	−2%	−1%	0%	1%	2%
1999	0	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	1	1
2001	1	1	1	1	1	1	1	1
2002	0	0	0	0	0	1	1	1
2003	0	0	0	0	0	0	1	1
2004	0	0	0	0	0	0	1	1
2005	0	0	0	0	0	0	0	1
2006	0	0	0	0	0	0	1	1
2007	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	1
2009	0	0	0	0	0	0	0	0
2010	−1	−1	−1	−1	−1	0	0	0
2011	−1	−1	−1	−1	0	0	0	0
2012	−1	−1	−1	−1	−1	−1	0	0
2013	−1	−1	−1	−1	0	0	0	0
2014	−1	−1	−1	−1	−1	−1	−1	−1
UDRP	6.25%	6.25%	6.25%	6.25%	6.25%	12.50%	37.50%	50.00%
MF	31.25%	31.25%	31.25%	31.25%	18.75%	12.50%	6.25%	6.25%
None	62.50%	62.50%	62.50%	62.50%	75.00%	75.00%	56.25%	43.75%

The existence of dominance is represented as '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys and '0' if there is no dominance. The last three rows represent the proportions of each case. In other words, 'UDRP', 'MF', and 'None' show the percentages of monkey-dominating, fund-dominating, and no-dominance cases, respectively.

4. Conclusions

In this study, we revisited the claim of Malkiel (1973) which states that fund managers would not outperform a dart-throwing blindfolded monkey, to propose an alternative approach to portfolio performance evaluation that compares a portfolio with respect to an infinite number of blindfolded monkeys. We used the method of a UDRP developed by Kim and Lee (2016) to actually implement infinite blindfolded monkeys.

The proposed method exhibits two main advantages. First, it is benchmark-free. In other words, we do not have to specify the benchmark because a portfolio is already being compared to an infinite number of portfolios. Thus, the methodology can be easily applied to markets other than equities where well-established benchmarks rarely exist. Second, it is market condition invariant. As the market condition is already reflected in the performances

Table 6. Stochastic dominance between the U.S. equity mutual funds (small-cap) and blindfolded monkeys with mutual fund fee adjustments.

	Mutual fund fee adjustment							
Year	−5%	−4%	−3%	−2%	−1%	0%	1%	2%
1999	1	1	1	1	1	1	1	1
2000	−1	0	0	0	0	0	0	0
2001	0	0	0	0	0	1	1	1
2002	−1	−1	−1	0	0	0	0	0
2003	0	0	0	0	0	1	1	1
2004	−1	−1	−1	−1	−1	0	0	0
2005	−1	−1	0	0	0	0	0	0
2006	−1	0	0	0	0	1	1	1
2007	−1	−1	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0
2010	−1	−1	−1	−1	0	0	0	0
2011	0	0	0	0	0	0	0	0
2012	−1	−1	−1	−1	−1	−1	0	0
2013	0	0	0	0	0	0	0	1
2014	−1	−1	−1	0	0	0	0	0
UDRP	6.25%	6.25%	6.25%	6.25%	6.25%	25.00%	25.00%	31.25%
MF	56.25%	43.75%	31.25%	18.75%	12.50%	6.25%	0.00%	0.00%
None	37.50%	50.00%	62.50%	75.00%	81.25%	68.75%	75.00%	68.75%

The existence of dominance is represented as '1' if blindfolded monkeys dominate mutual funds, '−1' if mutual funds dominate blindfolded monkeys and '0' if there is no dominance. The last three rows represent the proportions of each case. In other words, 'UDRP', 'MF', and 'None' show the percentages of monkey-dominating, fund-dominating, and no-dominance cases, respectively.

of blindfolded monkeys, our method can provide a consistent measure of portfolio performance regardless of the ever-changing market conditions. In other words, we now have a simple procedure that allows us to obtain something similar to the NBA season standing for portfolios.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Ardia, D., and K. Boudt. 2018. "The Peer Performance Ratios of Hedge Funds." *Journal of Banking and Finance* 87: 351–368.

Arnott, R. D., J. Hsu, V. Kalesnik, and P. Tindall. 2013. "The Surprising Alpha from Malkiel's Monkey and Upside-Down Strategies." *The Journal of Portfolio Management* 39 (4): 91–105.

Barras, L., O. Scaillet, and R. Wermers. 2010. "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas." *The Journal of Finance* 65 (1): 179–216.

Berk, J. B., and J. H. van Binsbergen. 2015. "Measuring Skill in the Mutual Fund Industry." *Journal of Financial Economics* 118 (1): 1–20.

Carhart, M. 1997. "On Persistence in Mutual Fund Performance." *The Journal of Finance* 52 (1): 57–82.

Fama, E. F., and K. R. French. 2010. "Luck versus Skill in the Cross-Section of Mutual Fund Returns." *The Journal of Finance* 65 (5): 1915–1947.

Jensen, M. C. 1968. "The Performance of Mutual Funds in the Period 1945-1964." *The Journal of Finance* 23: 389–416.

Khorana, A., H. Servaes, and P. Tufano. 2009. "Mutual Fund Fees around the World." *Review of Financial Studies* 22 (3): 1279–1310.

Kim, W. C., and Y. Lee. 2016. "A Uniformly Distributed Random Portfolio." *Quantitative Finance* 16 (2): 297–307.

Kim, W. C., Y. Lee, and Y. H. Lee. 2014. "Cost of Asset Allocation in Equity Market: How Much Do Investors Lose Due to Bad Asset Class Design?" *The Journal of Portfolio Management* 41 (1): 34–44.

Lee, Y., and W. C. Kim. 2014. Concise Formulas for the Surface Area of the Intersection of Two Hyperspherical Caps. *KAIST Technical Report*, IE-TR-2014-01. http://ie.kaist.ac.kr/uploads/professor/tech_file/Concise+Formulas+for+the+Surface+Area+of+the+Intersection+of+Two+Hyperspherical+Caps.pdf

Li, S. 2011. "Concise Formulas for the Area and Volume of a Hyperspherical Cap." *Asian Journal of Mathematics and Statistics* 4 (1): 66–70.

Malkiel, B. G. 1973. *A Random Walk Down Wall Street*. New York: Norton.

Malkiel, B. G. 1995. "Returns from Investing in Equity Mutual Funds, 1971-1991." *Journal of Finance* 55 (4): 1705–1770.

Sharpe, W. F. 1966. "Mutual Fund Performance." *The Journal of Business* 39 (1): 119–138.

Treynor, J. L. 1965. "How to Rate Management of Investment Funds." *Harvard Business Review* 43 (1): 63–75.

Wermers, R. 2000. "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses." *The Journal of Finance* 55 (4): 1655–1695.

Appendix. Implementation of non-negative UDRP

A.1. Assumptions

In Kim and Lee (2016), an ordinary UDRP is a random variable that is uniformly distributed on the surface of an n-dimensional unit hypersphere in a L_Σ^T-transformed space, where L_Σ^T is a Cholesky decomposition of the covariance matrix Σ of n risky securities (i.e. $\Sigma = L_\Sigma L_\Sigma^T$). Then, the non-negative UDRP would be a random variable that is uniformly distributed on the non-negative part of the surface of an n-dimensional unit hypersphere.

However, it is difficult to analytically derive the performance distribution of all possible portfolios within the exact non-negative area on the surface of an n-dimensional unit hypersphere. Therefore, Kim and Lee (2016) approximate the non-negative area based on the following observations. The non-negative area on the surface of a unit hypersphere is centred at 1, which denotes an n-dimensional vector of ones, and has the surface area equal to $A_n/2^n$, where A_n represents the surface area of an n-dimensional unit hypersphere. Consequently, they approximate the non-negative area with a hypersphreical cap FR^+ that is centred at $L_\Sigma^T 1$ and has the surface area equal to $A_n/2^n$. Then, the approximated non-negative area FR^+ should have the colatitude angle of

$$\theta_{FR^+} = \arcsin\sqrt{I^{-1}_{1/2^{n-1}}\left(\frac{n-1}{2}, \frac{1}{2}\right)}.$$

A.2. Formula

For a portfolio $w_s \in \mathbb{R}^n$ with Sharpe ratio s, the probability of w_s to outperform the non-negative UDRP in the Sharpe ratio can be calculated as follows.

$$\mathbb{P}(w_s \text{ to outperform the non}-\text{negative UDRP}) \\ = 1 - A_n^C \frac{\left(L_\Sigma^T 1, \theta_{\mathrm{FR}^+}\right) \cap C\left(L_\Sigma^T w^*, \theta_s\right)}{A_n^C\left(L_\Sigma^T 1, \theta_{\mathrm{FR}^+}\right)}$$

Here, w^* is the optimal tangent portfolio, and θ_s and $\theta_v \in [0, \pi]$ are the angle between $L_\Sigma^T w_s$ and $L_\Sigma^T w^*$ and the angle between $L_\Sigma^T 1$ and $L_\Sigma^T w^*$, respectively. $C(v, \theta)$ represents the unit hyperspherical cap whose axis is v and colatitude angle is θ, and $A_n^{C\left(L_\Sigma^T 1, \theta_{FR^+}\right) \cap C\left(L_\Sigma^T w^*, \theta_s\right)}$ and $A_n^{C\left(L_\Sigma^T 1, \theta_{FR^+}\right)}$ are the surface areas of $C\left(L_\Sigma^T 1, \theta_{FR^+}\right) \cap C\left(L_\Sigma^T w^*, \theta_s\right)$ and $C\left(L_\Sigma^T 1, \theta_{FR^+}\right)$, respectively. The surface area of a hyperspherical cap and the intersection of two hyperspherical caps can be calculated by the formula given by Li (2011) and Lee and Kim (2014).