Was 2016 the year of the monkey?

NB: Not to be quoted without the permission of the authors

Andrew Clare, Nick Motson and Stephen Thomas[1]

February 2017

Abstract

According to the Chinese calendar 2016 was the year of the Monkey. In this paper, using a common set of 500 US stocks, we analyse the performance of 1 billion randomly generated stock indices (as if chosen by a monkey) to both a market capitalisation weighted index and several popular smart beta indices. We find that 2016 was indeed a good year for monkeys. In the interests of animal welfare we have since released our monkeys back in to the wild for a well-earned rest and some energy replenishing bananas. It is now the Chinese year of the Rooster. Perhaps the big question is: who will be crowing by the end 2017, monkeys or market cap proponents?

JEL classification: G10; G11; G14; G23
Keywords: Market capitalisation-weighted indices; Alternative equity indices; Smart Beta

[1] Faculty of Finance, The Sir John Cass Business School, City University, London, UK. E-mails, respectively: a.clare@city.ac.uk; n.e.motson@city.ac.uk; and s.h.thomas@city.ac.uk.

1. Introduction

Back in 2011 Clare et al (2011a and 2011b) published research on the then relatively new investment phenomenon of 'smart beta', more appropriately referred to as alternative equity indexing. Using data that spanned 1968 to 2011, their research demonstrated that all of the smart beta approaches to US equity investment that were explored produced superior – and in some cases vastly superior – risk-adjusted returns to those generated by a comparable, market cap-weighted US equity index. They constructed the following alternative indices of US equities: equally-weighted (DeMiguel, Garlappi and Uppal (2009)); diversity-weighted (Fernholz *et al* (1998)); inverse volatility-weighted (inspired by Haugen and Heins (1975); equal risk contribution-weighted (Maillard *et al* (2010)); risk clustering-weighted; minimum variance-weighted (inspired by the MVEF); maximum diversification-weighted (Choueifaty and Coignard (2008)); risk efficient-weighted (Amenc *et al* (2011)); and fundamentally-weighted (Arnott *et al* (2005)).

To try to determine whether the poor performance of the market cap-weighted approach to indexing was due to bad luck or bad index design – Clare et al (2011a, b) undertook an experiment that involved constructing ten million indices, where the weights in each one were chosen randomly on an annual basis. Essentially, the weights in these indices may as well have been chosen by ten million monkeys. The results of the experiment were a little surprising and were a real challenge to index providers and to both the passive and active fund management industry who, respectively, usually track market cap-weighted indices, or try to beat them. Over the period from 1968 to 2011, around 99% of the monkey-constructed indices produced a better risk-adjusted performance than the market cap-weighted approach to index construction. That's around 9,900,000 monkeys. These results led Clare et al to conclude that the poor performance of the market cap-weighted methodology over that 43 year period was due to bad index design rather than to bad luck.

According to the Chinese calendar, 2016 year was the year of the monkey. In this paper we therefore take an impish look at the performance of our monkey index constructors again, to see whether 2016 really was the year of the monkey. The rest of this paper is organised as follows. In section 2 we outline our data and methodology; in section 3 we present our results; and finally in section 4 we draw conclusions.

2. Data and Methodology

The data in our analysis was all obtained from Bloomberg. We began by identifying the largest 500 stocks, by market capitalisation, listed on the NYSE, Amex and NASDAQ stock exchanges, with the added constraint that each stock should have at least 250 weeks of uninterrupted return history. This data is therefore comparable to the data collected by Clare et al (2011a, b) and used as the basis of their study. We take the rules for each index technique that we examine in this paper – that is, the market cap-weighted index, the set of alternative indices and the randomly produced indices – and we identify the relevant weights for the 500 constituent stocks for each index as at 31st December 2015. We then monitor the subsequent performance of each index over 2016.

2.1 The randomisation process

We applied the same Monte Carlo technique used by Clare et al, but here provide more detail on the process which was not provided in their 2011a and b papers.

There are clearly an infinite number of ways of combining a portfolio of 500 stocks. To generate the randomly constructed indices we have adapted an algorithm proposed by Smith and Tromble (2004) in order to generate indices with weights that are uniformly and randomly sampled from the feasible set of weights, so that it is equally likely that we generate an index with a 100% weight in one stock and zero in all others as it is that we generate an index that has an equal weight in all stocks, or even an index that is equivalent to the weights of the Market Cap index. To illustrate the process, consider the following three asset example. Let the weights of these three stocks in the index be represented by w_i. If we add the constraint that $w_i>=0$ and that $\sum_{i=1}^{3} w_i = 1$, then the feasible set of weights (w_1, w_2 and w_3) can be viewed as a hyper-plane which is shown in Figure 1.

In order to reduce the set of possible weights to a finite number we can define a minimum increment for the variation in weights Δw_i. Then, given *n* stocks, the algorithm works in 4 simple steps, as follows:

1. Sample *n-1* numbers uniformly at random from the set {1, 2... $(1/\Delta w)+n-1$)} without replacement.
2. Sort the numbers in ascending order and append a zero to the beginning of the sequence and $(1/\Delta w + n)$ to the end of the sequence.
3. Take the difference between successive numbers in the sample and subtract 1 from each.
4. Multiply these numbers by Δw.

Figure 2 shows a scatter plot of the result of 10,000 repetitions of the above algorithm for n=3 and Δw=0.1%, the uniformity of outcomes illustrates the un-biasedness of the process.

The algorithm can be extended to encompass any number of stocks. For example, consider the example of a portfolio containing 100 stocks (n=100) where the minimum increment is set at 1% (Δw =0.01). The first step would involve selecting 99 random numbers from the set {1, 2 ... 199}. If we suppose that the numbers chosen are {2, 4, 6 … 198} then step 2 would result in the following set of 101 numbers {0, 2, 4, 6 ... 198, 200}. Step 3 would produce 100 identical numbers {1, 1 … 1} and hence step 4 would generate an equally-weighted portfolio with each stock given a weight equal to 1% or 1/n. If instead the 99 random numbers chosen had been {1, 2, 3 ... 99} then the set of weights produced would be zero for the first 99 stocks and 100% for the 100th stock. Since choosing {2, 4, 6 … 198} and choosing {1, 2, 3 …, 99} are equally likely this demonstrates that the randomly generated portfolio weights are unbiased.

Using this algorithm we generate 500 weights that sum to one, with a minimum increment of 0.2%. We then apply these weights to the universe of 500 stocks sampled from Bloomberg in December 2015. We then calculate the performance of the resulting index over the next twelve months. We repeated this process one billion times. This means that we generated one billion indices of 500 US equities, where the weights were chosen at random or as if by a monkey.

2.2 The alternative indices

As well as comparing the performance of the market cap-weighted index against the performance of 1 billion randomly constructed indices, in the interest of comparison, we also constructed and evaluated the performance of the alternative equity index construction processes that were investigated by Clare et al (2011a, b) over 2016 as well. Each alternative index was constructed as described below.

2.2.1 Equal Weights

With the equally-weighted approach, each of the N stocks in the equity universe is assigned an equal weight of 1/N. This is a very simple and perhaps somewhat naive approach to determining weights was examined by DeMiguel, Garlappi and Uppal (2009) and found to outperform many more sophisticated methods due to the avoidance of parameter estimation errors. Equal weighting avoids the concentration risk that might arise from a Market Cap index, however, one of the possible drawbacks of this approach is that, by definition, it gives higher weights to smaller, possibly less liquid stocks than the Market Cap-weighted approach and will also have a higher turnover.

2.2.2 Diversity Weights

The diversity-weighted approach was first proposed by Fernholz et al (1998), effectively it involves raising the market cap weight for stock i (w_i) of each constituent to the power p, that is $(w_i)^p$, where p is bounded between 1 and 0. The weight of each index constituent is then calculated by dividing its $(w_i)^p$ weight by the sum of all $(w_i)^p$'s of all of the constituents in the index. When p is set equal to 1 the constituent weights are equal to market cap weights; when p is set equal to 0 the weights are equivalent to equal weights. We use p=0.76 which is the value used in the original paper.

2.2.3 Inverse Volatility Weights

Haugen and Heins (1975) demonstrated that low volatility stocks tended to outperform high volatility stocks, since then there has been much research on the "low-volatility anomaly". We calculate the historical return variance of each stock using 250 weeks of weekly data. We then calculate the inverse of this value, so that the stock with the lowest volatility will have the highest inverted volatility. We then simply summed these inverted variances. The weight of stock i is then calculated by dividing the inverse of its return variance by the total inverted return variance. This process therefore assigns the biggest weight to the stock with the lowest volatility, and the lowest weight to the stock with the highest return volatility.

2.2.4 Equal Risk Contribution Weights

Maillard et al (2010) propose weighting each stock such that that the contribution of each stock to the risk of the overall portfolio is equal. We use a covariance matrix based on 250 weeks of weekly history (shrunk using Ledoit and Wolf (2004)) and the algorithm proposed by Maillard et al to calculate equal risk contribution weights.

2.2.5 Minimum Variance Weights

The minimum variance approach uses historical data in an attempt to identify the weights of the global minimum variance portfolio. Authors such as Clarke, de Silva, and Thorley (2006) have identified strong performance of minimum variance portfolios. We use the same shrunk covariance matrix as above and cap individual stock weights at a maximum of 2%.

2.2.6 Maximum Diversification Weights

Choueifaty and Coignard (2008) introduce a measure of portfolio diversification, called the *"Diversification Ratio"*, which is defined as the ratio of a portfolio's weighted average volatility to its overall volatility. Poorly diversified portfolios that have either concentrated weights, highly correlated holdings or even both will exhibit relatively low diversification ratios. Choueifaty and Coignard propose optimisation to identify the 'most diversified portfolio' which is defined as the portfolio with the highest diversification ratio. Intuitively it is apparent that if expected returns are proportional to their volatility, the maximum diversification portfolio will be the same as the maximum Sharpe ratio portfolio which is proven mathematically by Choueifaty *et al* (2013). We use the same shrunk covariance matrix as above and again cap individual stock weights at a maximum of 2%.

2.2.7 Risk Efficient Weights

Amenc *et al* (2010) propose an index construction approach that also embodies estimates of the 'expected return' on the index components. However, instead of assuming or postulating, that the expected return on a stock is directly proportional to its volatility, they instead suggest that it is directly proportional to the downside deviation of the stock's return. Downside deviation focuses attention on negative returns so, other things equal, the greater their frequency and size, the higher will be the stock's measured return downside deviation.

To construct a *Risk Efficient* index Amenc *et al* propose a two stage process. First, the semi-deviation of each stock is calculated. Then on the basis of these estimates, the stocks are grouped into deciles so that the 10% of stocks with the largest downside deviation comprise the first decile; the 10% with the next highest downside deviation comprise the second decile and so on, until ten deciles are identified. The median downside deviation for each decile is then calculated and this value is then assigned to each stock in its decile as the proxy for the expected return of that stock. The second stage then involves finding the portfolio with the maximum expected return (proxied by the median downside deviation of each stock's decile) with the lowest portfolio return standard deviation.

Again, to prevent the optimiser from creating a portfolio with concentrated single stock exposures, Amenc *et al* impose restrictions on the constituent weights that might otherwise be chosen by the optimiser. The weight limits are as follows:

lower limit = 1/(λ x N) x 100%

upper limit = λ/N x 100%

where N represents the total number of stocks under consideration and where λ is a free parameter. If λ is set equal to 1, then all constituent index weights would be equal to 1/N, that is, the index constituents would be equally-weighted. We use the same shrunk covariance matrix as above and set λ equal to 2.

2.2.8 Fundamental Weights

Arnott *et al* (2005) argue that alternative measures of the company size or scale may be just as appropriate as a basis for determining constituent weights as the more commonly used metric of market capitalization. We use four alternative measures of the size of a company to construct alternative indices of US equities. The methodology that we use is based upon the one outlined in more detail in Arnott *et al*. We calculate four different indices that weight stocks according the book value of equity as at December 2015 and the 5 year historical average of total dividends, cash-flow, and sales. We then take the average weights of these four indices to form a fundamental composite index.

2.2.9 Scrabble™ Weights

Finally, in this paper we also examine the 2016 performance of an index of US equities where the weights were chosen as at 31st December 2015 according to the rules of Scrabble™. More precisely we use each company's ticker symbol to calculate a Scrabble™ score for each stock, according to the following rules:

1 point – A, E, I, O, U, L, N, S, T, R.
2 points – D, G.
3 points – B, C, M, P.
4 points – F, H, V, W, Y.
5 points – K.
8 points – J, X.
10 points – Q, Z

We then sum the scores of each company's ticker and divide each stock's score by this total to give its weight in the index.

3. Index performance over 2016

3.1 The alternative indices

Table 1 presents the summary statistics of the market cap-weighted index and the performance of each of the alternative indices over 2016, the year of the monkey. All of the alternatives, with the exception of the Minimum Variance index, produce a higher average return than the market cap-weighted index over this period. The Fundamentally-weighted index produces the highest return, 13.67%, and the Scrabble-weighted index also manages to outperform the market cap-weighted index by nearly 2.0%. In risk-adjusted terms, all of the alternatives produce a higher return than the market cap benchmark, with the Maximum Diversification index producing the highest Sharpe ratio of 1.22. The active share[2] statistics in the final column of the table also indicate that the Maximum Diversification index was the most "active" of all the alternative approaches.

These results suggest that the market cap-weighted index would have been a fairly poor index choice for investors over 2016, as it would have been between 1968 and 2011. Also, because there was no rebalancing of these indices throughout the year, the results cannot be attributed to a failure to consider, stock turnover and transactions costs.

3.2 The 1bn randomly constructed indices

Perhaps the market cap-weighted index performed poorly during the year of the money, because it suffered from bad luck. The results in Figure 3, suggest that the result not have been due to bad luck. The black line in the Figure shows the distribution of the returns achieved by the 1 billion monkeys over 2016; the grey line shows the related cumulative frequency. The figure shows that 88% of the monkeys managed to produce a higher return than the market cap benchmark. The average monkey return was 13.4%; the luckiest monkey achieved a return of 27.2%; while the unluckiest monkey managed a return of just 3.83%.

Superimposed on the distribution presented in Figure 1 are the returns achieved by each of the alternative indices. The stand out performer over 2016 is clearly the Fundamentally-weighted index. This index construction technique outperformed 95.7% of the monkeys. Many of the other indices perform poorly against the randomly constructed indices, but the worst of all is the

[2] See Cremers et al, 2009.

minimum variance index which outperforms only 1.5% of the 1bn monkeys. In fact, with the exception of the Fundamentally-weighted index, the monkeys all performed admirably against the so-called smart beta approaches in 2016.

Figure 4 shows the distribution of monkey Sharpe ratios. Here the monkeys do less well against the market cap-weighted index. However, 67% of them still manage to produce a higher risk-adjusted return. The luckiest monkey produced a Sharpe ratio of 1.67, while the unluckiest monkey produced a Sharpe ratio of 0.32. In risk-adjusted terms however, the stand out performer is the Maximum Diversification-weighted index. This index outperforms 99.76% of the Sharpe ratios produced by the randomly constructed indices. The Fundamentally-weighted index also produced an impressive performance, outperforming all but 3.51% of the 1 billion random indices.

4. Conclusions

Taken together, our results suggest that it was a pretty good year for Cass Business School's monkey index constructors, and another poor year for market cap weighted US stock investing. However, it was arguably a much better year for Fundamentally and Maximum Diversification-weighted index approaches, depending upon whether one focuses on raw returns or risk-adjusted returns.

2017 is the Chinese year of the rooster. It remains to be seen whether the Cass Business School monkeys will be crowing at the end of next year or not … but all the evidence suggests that they have the beating of market cap-weighted index investing!

5. References

Amenc, N., Goltz, F., Martellini, L., and Retkowsky, P., 2011. Efficient Indexation: An Alternative to Cap-Weighted Indices. *Journal of Investment Management* Vol. 9, 1-23.

Arnott, R.D., Hsu, J., Moore, P., 2005. Fundamental indexation. *Financial Analysts Journal*, 83-99.

Choueifaty, Y., and Coignard. Y., 2008. Toward maximum diversification. *The Journal of Portfolio Management* 35.1, 40-51.

Choueifaty, Y., Froidure, T., and Reynier, J. 2013. Properties of the most diversified portfolio. *Journal of Investment Strategies* Volume, 2(2), 49-70.

Clare, A, N. Motson and S. Thomas, 2011a. An evaluation of alternative equity indices, Part 1: Heuristic and optimised weighting schemes, Cass Business School, London. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2242028.

Clare, A, N. Motson and S. Thomas, 2011b. An evaluation of alternative equity indices, Part 2: Fundamental weighting schemes, Cass Business School, London. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2242034.

Clarke, R.G., De Silva, H., and Thorley. S., 2006. Minimum-variance portfolios in the US equity market. *The Journal of Portfolio Management* 33.1, 10-24.

Cremers, K., Martijn, J., Petajisto, A., 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22.9, 3329-3365.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22.5, 1915-1953.

Fernholz, R., Garvy, R., Hannon, J., 1998. Diversity-Weighted Indexing. *The Journal of Portfolio Management* 24.2, 74-82.

Haugen, R.A., Heins, A.J., 1975. Risk and the rate of return on financial assets: Some old wine in new bottles. *Journal of Financial and Quantitative Analysis* 10.5, 775-784.

Maillard, S., Roncalli, T., Teïletche, J., 2010. The Properties of Equally Weighted Risk Contribution Portfolios. *The Journal of Portfolio Management* 36.4, 60-70.

Smith, N.A., Tromble, R.W., 2004. Sampling uniformly from the unit simplex. *Johns Hopkins University, Tech. Rep*

Figure 1

This figure presents the hyper-plane of feasible weights for a long only portfolio of 3 stocks with weights w_1, w_2 and w_3 where $\sum_{i=1}^{3} w_i = 1$.

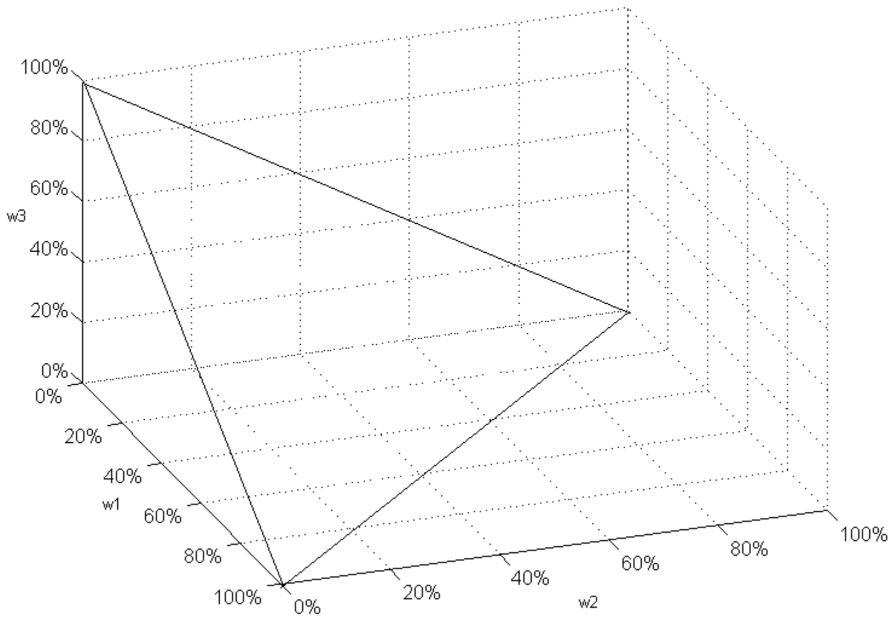

Figure 2

This figure shows a scatter plot of the result of 10,000 applications of the algorithm adapted from Smith and Tromble (2004), with n=3 and Δw=0.1%, overlaid on the hyper-plane presented in Figure 1. The uniformity of outcomes illustrates the un-biasedness of the process.

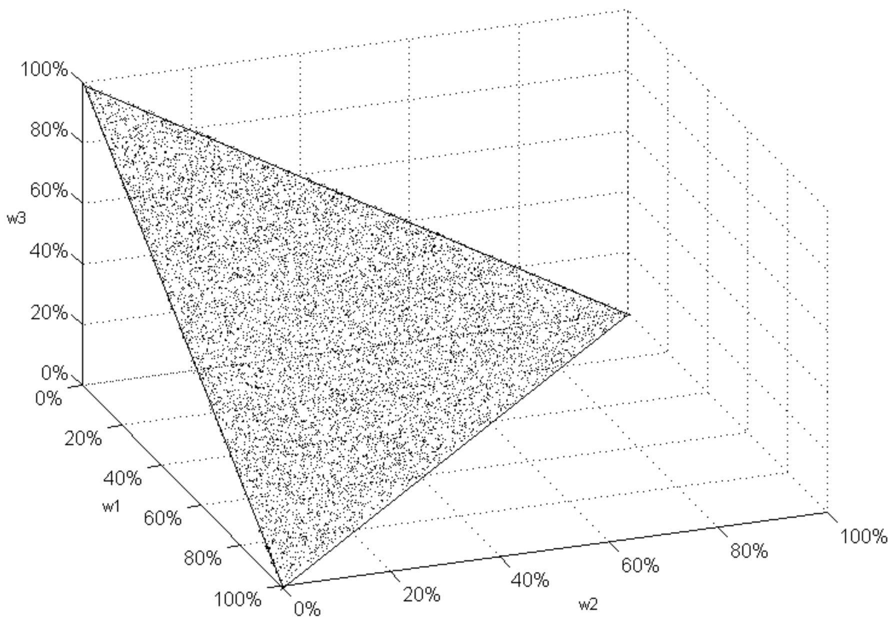

Figure 3: Total Return 1 Billion Random Portfolios versus Indices

This figure presents the total return for 1 billion indices produced using random weights generated by the algorithm described in appendix 2 with n=500 and Δw=0.02%. The frequency distribution is shown in dark grey (left axis) and the cumulative frequency in light grey (right axis). The returns of the Market Capitalisation weighted index and the alternatives are overlaid for ease of reference

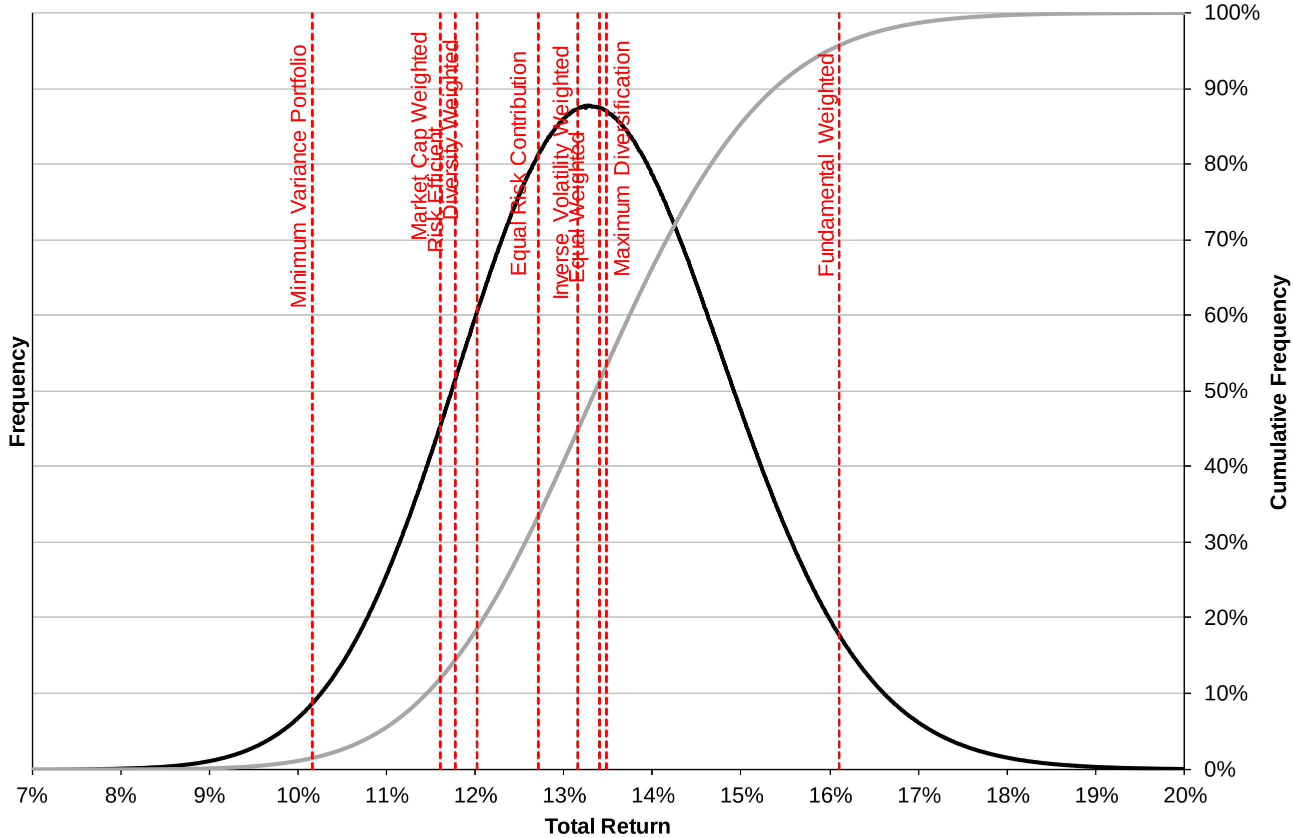

Figure 4: Sharpe Ratio 1 Billion Random Portfolios versus Indices

This figure presents the Sharpe ratio for 1 billion indices produced using random weights generated by the algorithm described in appendix 2 with n=500 and Δw=0.02%. The frequency distribution is shown in dark grey (left axis) and the cumulative frequency in light grey (right axis). The Sharpe ratios of the Market Capitalisation weighted index and the alternatives are overlaid for ease of reference

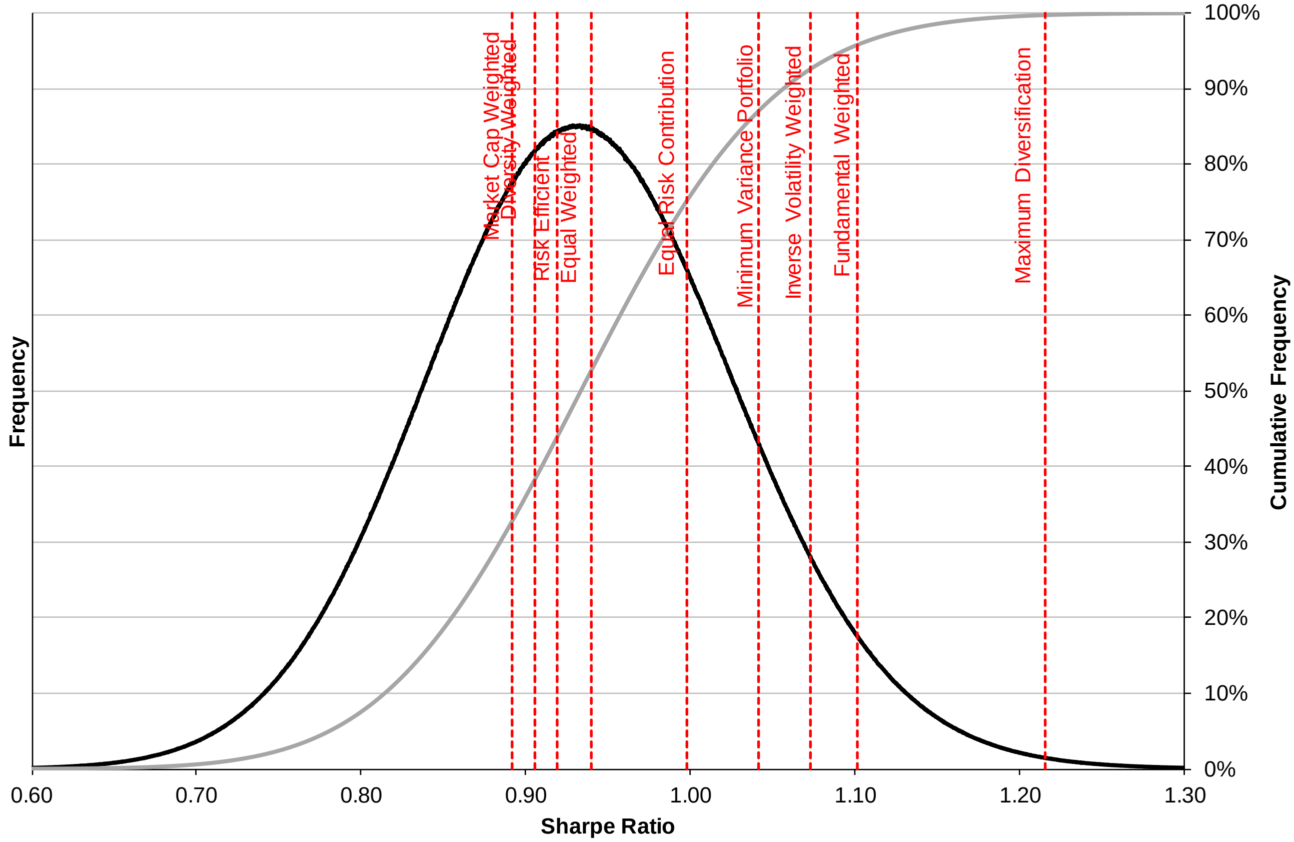

Table 1: Return and Risk Characteristics of a range of US Equity Indices over 2016

This table presents the performance of a range of indices comprising the largest 500 stocks, by market capitalisation, as at 31st December 2015, listed on the NYSE, Amex and NASDAQ stock exchanges, with the added constraint that each stock had to have had at least 250 weeks of uninterrupted return history. See the text, and Clare et al (2011a, b) for information about the construction of each index.

	Terminal Wealth	Total Return	Mean Return (Annualised)	Standard Deviation (Annualised)	Sharpe Ratio	Active Share
Market Cap-Weighted	111.6	11.6%	11.8%	13.1%	0.89	0%
Equally-Weighted	113.4	13.4%	13.6%	14.3%	0.94	45%
Diversity-Weighted	112.0	12.0%	12.2%	13.3%	0.91	12%
Inverse Volatility-Weighted	113.2	13.2%	13.1%	12.0%	1.07	44%
Equal Risk Contribution	112.7	12.7%	12.8%	12.6%	1.00	44%
Minimum Variance Portfolio (2% Cap)	110.2	10.2%	10.1%	9.5%	1.04	85%
Maximum Diversification (2% Cap)	113.5	13.5%	13.2%	10.7%	1.22	90%
Risk Efficient (Lambda = 2)	111.8	11.8%	11.9%	12.8%	0.92	52%
Fundamentally-Weighted	116.1	16.1%	16.0%	14.3%	1.10	25%
Scrabble	113.4	13.4%	13.7%	14.5%	0.93	47%