Original Article

Further evidence in support of a low-volatility anomaly: Optimizing buy-and-hold portfolios by minimizing historical aggregate volatility

Phil Maguire

is a lecturer in the Department of Computer Science, National University of Ireland, Maynooth. His research explores portfolio optimisation, predictive modelling and the foundations of measurement.

Stephen Kelly

graduated with a BSc in Computational Thinking and an MSc in Mathematics from the National University of Ireland, Maynooth. He is currently a PhD student in City, University of London, working in the area of data science.

Robert Miller

is an undergraduate student at the National University of Ireland, Maynooth taking the BSc degree in Computational Thinking. He has been awarded the Alan Turing and Stephen Cook prizes in Computational Thinking and the Delort and McMahon prizes in mathematics.

Philippe Moser

is a lecturer in the Department of Computer Science, National University of Ireland, Maynooth. His research interests include algorithmic information theory, randomness, computability, complexity theory and computational finance.

Philip Hyland

is a senior lecturer in the School of Business, National College of Ireland. His research focuses on psychological responses to traumatic life events and disorders associated with stress such as PTSD and Complex-PTSD.

Rebecca Maguire

is a lecturer in the School of Business, National College of Ireland. Her area of expertise is the cognitive modelling of uncertainty.

Correspondence: Phil Maguire, Department of Computer Science, National University of Ireland, Maynooth, Ireland
E-mail: pmaguire@cs.nuim.ie

ABSTRACT The 'low-volatility anomaly' is the counter-intuitive observation that portfolios of low-volatility stocks tend to yield higher risk-adjusted returns than portfolios of high-volatility stocks. In this article, we investigate if the anomaly holds, not only for portfolios consisting of individual low-volatility stocks, but for portfolios that have been optimized to minimize aggregate volatility. We exploit patterns in historical price fluctuations to identify optimized portfolios whose aggregate volatility is expected to remain low. These portfolios are evaluated by comparing them against the performance of market capitalization and low-volatility quintile benchmarks out-of-sample. The results reveal that, as well as outperforming

the market, both in terms of returns and risk, optimized low-volatility strategies also out-perform the S&P Low-Volatility Index. These findings provide further support for a low-volatility effect, and imply that the root of the anomaly may lie with a failure to exploit diversification opportunities.

INTRODUCTION

The capital asset pricing model (CAPM) states that the returns of a given stock should be a linear function of its beta (i.e. market risk). In other words, returns should reflect how risky a stock is relative to the market. Surprisingly, this prediction does not match observations: Over the last 50 years, low-volatility portfolios across the world have offered an enviable combination of high average returns and small drawdowns, bucking the intuition that risk should be compensated with higher expected profits (e.g. Haugen and Baker, 1991; Clarke *et al*, 2006; Baker *et al*, 2011). The deviation is so compelling that Baker *et al* (2011) have proposed it as a candidate for "the greatest anomaly in finance".

There remains some debate as to whether the anomaly is real, insofar as it can be exploited in practice (Li *et al*, 2014). It may be the case that the anomaly is small enough that it would be eaten away by transaction costs, or that it reflects some statistical quirk such as future bias. A further question concerning the low-volatility anomaly is whether it occurs at the level of individual stocks, or whether it can be further enhanced by exploiting the relationships between stocks.

Our study seeks specifically to address these three questions. First, we apply a buy-and-hold strategy, which has minimal associated trading costs as no rebalancing is required. Second, we apply our strategy out-of-sample, developing the portfolio based on historical data and thus removing any possibility of future bias. Third, we develop

our portfolios so as to exploit the relationships between stocks, driving volatility even lower in a bid to enhance performance. We track the performance of our portfolios into the future, to see how they perform relative to the market and other low-volatility benchmarks.

Before detailing our strategy, we first review some of the evidence supporting the existence of a low-volatility effect.

EVIDENCE OF A LOW-VOLATILITY ANOMALY

Investigating a broad sample of international developed markets, Ang *et al* (2009) found that stocks with recent past high idiosyncratic volatility had lower future average returns. Across 23 markets, the adjusted difference in average returns between highest and lowest quintile portfolios, sorted by volatility, was −1.31% per month. This effect was found to be individually significant for every G7 country (Canada, France, Germany, Italy, Japan, the United States, and the United Kingdom), suggesting that the relation between high idiosyncratic volatility and low returns is not just a sample-specific or country-specific effect, but a global phenomenon.

Blitz and van Vliet (2007) provided further empirical evidence that stocks with low-volatility earn higher risk-adjusted returns than the market portfolio, even after controlling for well-known effects such as value and size. They found that the annual alpha spread of global low versus high-volatility decile portfolios amounted to 12%

over the 1986–2006 period, observing independent effects in the US, European and Japanese markets.

Baker and Haugen (2012) analysed 33 different markets during the time period from 1990 to 2011, including non-survivors. They computed the volatility of total return for each company in each country over the previous 24 months, ranking stocks by volatility and grouping them into deciles. In each one of the 21 developed countries, the lowest volatility decile had both lower risk *and* higher return, leading to substantial divergence in Sharpe ratios.

In another study, Clarke *et al* (2006) found that minimum variance portfolios, based on the 1,000 largest US stocks over the 1968–2005 period, achieved a volatility reduction of about 25%, while delivering comparable, or even higher, average returns than the market portfolio.

Ang *et al* (2009) have pinpointed U.S. markets as the source of the anomaly. Specifically, they found that the low-volatility anomaly in international markets strongly co-moves with the anomaly for U.S. stocks. After controlling for U.S. portfolios, the alphas (i.e. risk-adjusted overperformance relative to the market) of portfolio strategies trading the idiosyncratic volatility effect in various international markets are insignificant. Thus, Ang *et al* (2009) argue that the global idiosyncratic volatility effect is captured by a simple U.S. idiosyncratic volatility factor. In the following sections, we consider some possible explanations for the existence of this U.S.–based anomaly.

Longshot payoffs

One possible explanation is that people are predisposed to express risk-seeking utility towards longshot payoffs, while expressing risk aversion towards lower volatility returns [see prospect theory; Tversky and Kahneman (1992)].

Buying a low-priced, volatile stock is like buying a lottery ticket: There is a small chance of it multiplying significantly in value in a short period, and a much larger chance of it declining in value (Baker *et al*, 2011). Kumar (2009) found that some individual investors do show a clear preference for stocks with lottery-like payoffs, measured as idiosyncratic volatility or skewness. Applying Tversky and Kahneman's (1992) cumulative prospect theory approach, Barberis and Huang (2009) concluded that a positively skewed stock can be overpriced because of its skewness, and thus earn a negative average excess return.

Shefrin and Statman (2000) point out that buying many stocks destroys upside lottery potential, while buying a few volatile stocks leaves upside potential intact. This way of thinking is consistent with the finding that most naive private investors only hold about 1–5 stocks in their portfolio, thereby largely ignoring the diversification benefits that are available within the equity market (Blitz and van Vliet, 2007). This effect may cause high-risk stocks to be overpriced and low-risk stocks to be underpriced (Blitz and van Vliet, 2007). According to Jiang *et al* (2009), the idiosyncratic volatility anomaly is indeed stronger among stocks with a less sophisticated investor base, as would be expected if the anomaly had a behavioural origin.

Lack of leverage

Baker and Haugen (2012) argue that the root of the anomaly lies with the structure of the investment environment. They posit that the compensation structures and internal stock selection processes at asset management firms motivate managers to hold more volatile stocks and shun low-risk stocks.

Leverage is needed to take advantage of the low-volatility anomaly: investors need to hold greater volumes of stock in order to achieve the same level of risk. If a low-risk stock portfolio has a volatility which is, say, two-thirds of that of the market, then 50% leverage needs to be applied in order to obtain the same level of volatility as the market. While this

might seem straightforward, in practice many investors are not in a position to apply leverage, and thus cannot exploit the opportunity (Blitz and van Vliet, 2007). Borrowing restrictions were originally identified by Black (1993) as an argument for the relatively good performance of low-beta stocks.

Although precise statistics are elusive, it appears that relatively few mutual funds use leverage. For example, Baker et al (2011) spot-checked and found that the five largest active domestic equity mutual funds did not use any leverage as of 1st July 2010. Furthermore, the Investment Company Act of 1940 actually prohibits U.S. mutual funds from using more than 33% leverage.

Baker et al (2011) also point out that the typical institutional investor's mandate to beat a fixed benchmark discourages arbitrage activity for both high–alpha, low-beta stocks and low–alpha, high-beta stocks, since these are unlikely to simultaneously track and surpass the benchmark.

As it turns out, such benchmarking is common: 61% of U.S. mutual fund assets are benchmarked to the S&P 500, and 95% are benchmarked to some popular U.S. index (Sensoy, 2009). Under current U.S. SEC rules, all mutual funds must, in their prospectus, select a benchmark and express the returns of the fund relative to that benchmark (Baker et al, 2011).

The focus on these benchmarks means that institutional investment managers are less likely to exploit the low-volatility anomaly, especially when leverage is not available (see Cornell and Roll, 2005). Instead, an investment manager who is seeking to surpass a fixed benchmark without access to leverage is better off exploiting mispricings among stocks with close to market risk (i.e. a beta near 1).

Money tends to be invested in asset classes that have performed well in the past, and with asset managers who have demonstrated above average performance (Blitz and van Vliet, 2007). For this reason, outperformance in up markets may be more desirable than outperformance in down markets (Blitz and van Vliet,

2007). Asset managers may thus be willing to overpay for stocks which are inclined to outperform in up markets; these tend to be high-volatility stocks. At the same time, they are likely to underpay for stocks which outperform in down markets, which tend to be low-volatility stocks (Blitz and van Vliet, 2007).

All of these factors combine to produce a low-volatility anomaly, whose essence is that low risk is undervalued relative to high risk (Baker et al, 2011). In a benchmarked world without access to sufficient leverage, the low-volatility anomaly emerges as a natural consequence.

Betting against beta

The concept of volatility is closely related to that of beta coefficient, which describes the volatility of an asset relative to the market, in essence the correlated relative volatility (by definition the market has a beta of 1). In line with other findings on the low-volatility anomaly, Black (1993) found that, in the period from 1931 to 1965, low–beta stocks in the U.S. did better than the capital asset pricing model (CAPM) predicts, while high-beta stocks did worse.

Frazzini and Pedersen (2014) provided empirical evidence that portfolios of high-beta assets have lower alphas and Sharpe ratios than portfolios of low-beta assets. They found that high beta is associated with low alpha for US equities, 20 international equity markets, treasury bonds, corporate bonds, and futures. They also found that a betting-against-beta (BAB) factor, which is long leveraged low-beta assets and short high-beta assets, produces significant positive risk-adjusted returns, and rivals standard asset pricing factors (e.g. value, momentum, and size) in terms of economic magnitude, statistical significance, and robustness across time periods, subsamples of stocks, and global asset classes.

Frazzinia and Pedersen's (2014) claim that constrained investors stretch for return by increasing their betas is supported by their finding that both mutual funds and individual

investors hold securities with betas that are significantly above one. In contrast, leveraged buyout funds, such as Berkshire Hathaway, tend to buy stocks with betas below one. These investors could be taking advantage of the BAB effect by applying leverage to safe assets and being compensated by investors facing borrowing constraints who take the other side. According to Frazzini and Pedersen (2014), Warren Buffett gets rich by betting against beta, that is, buying stocks with betas significantly below one and applying leverage.

Low-volatility stocks, or low-volatility portfolios?

Given the outperformance of low-volatility stocks, the question arises of whether it is possible to beat the performance of low-risk quintile portfolios by taking advantage of the benefits of diversification (Baker *et al*, 2011). For example, a portfolio of two uncorrelated but slightly more individually volatile stocks can be even less volatile than a portfolio of two correlated stocks with lower volatility. Does optimizing volatility in this way amplify the low-volatility effect?

Investigating this possibility, Baker *et al* (2011) constructed a minimum-variance portfolio that took advantage of finer detail in the covariance matrix. Following the estimated security method of Clarke *et al* (2006), they compared the returns of an optimized low-volatility portfolio against the performance of the lowest volatility quintile for the period 1968–2008. The total volatility was reduced from 12.7% for the bottom volatility quintile to 11.5% for the optimized portfolio. Because this reduction in volatility comes at no expense in terms of average returns, the Sharpe ratios are substantially better in the optimized case (0.47 versus 0.38 for bottom quintile).

Is the effect real?

On the other side of the argument, Li *et al* (2014) have claimed that portfolio-rebalancing requirements, and the impact of associated transaction costs on low-volatility portfolios, eliminates the benefits of the purported anomaly.

They suggested that, over the period 1963–2010, the existence and trading efficacy of the low-volatility stock anomaly was more limited than widely believed, and could not be exploited in practice. For example, they found that abnormal returns are concentrated among smaller stocks, which suffer from a lack of liquidity and thus attract high transaction costs. When low-priced stocks are omitted from value-weighted long/short portfolios, alpha is largely eliminated.

In summary, there is strong evidence of a low-volatility anomaly, perhaps caused by a combination of risk-seeking utility, inappropriate benchmarking, and a lack of access to leverage. Nevertheless, the question has been raised by Li *et al* (2014) as to whether this anomaly can be easily exploited. This is the question addressed by our study.

THE STUDY: OPTIMIZING PORTFOLIOS BY MINIMIZING HISTORICAL VOLATILITY

In the following study, we investigate whether the low-volatility can be easily exploited to generate profit. First of all, a long-term buy-and-hold strategy is adopted, thus avoiding the issues of low liquidity and high transaction costs identified by Li *et al* (2014). Secondly, in contrast to previous studies (e.g. Baker *et al*, 2011), our analysis is guaranteed out-of-sample, insofar as the test data were not available when the training data were collected.

In-sample analyses, whereby portfolios are developed and tested on the same dataset, are problematic because they invite overfitting. It is often easy to find some manipulation of the data that leads to outperformance by exploiting random variation in the dataset, variation which is unlikely to persist into the future (see Hawkins, 2004; Maguire *et al*, 2012). Accordingly, it becomes difficult to

differentiate between models which have been overfitted to the data versus those whose development has not been influenced by random quirks. One way to control for overfitting is to take into account the length of the model relative to the amount of data compression it achieves (see Li and Vityányi, 2008). Another, more straight-forward, method is to separate the training set and the test set.

A separate issue which arises when delineating training and test sets of historical returns is that the training set will inevitably be older then the test set, raising the possibility of contamination by future bias. For example, if we look at the S&P 500 stocks that are in the index today, and track them back 20 years, we will see that these stocks considerably outperform the market. The reason is that they are future biased and fail-proof: These companies are the ones that are guaranteed to survive and grow to take their place in the S&P index 20 years into the future. This information was not available to historical investors, thus the selection of current S&P 500 stocks leaks information from the future into the past. In order for a study of this type to be fully reliable, its training data must have been identified before the test data were *knowable*.

Our study is truly out–of–sample. We collected the training data, namely daily stock market price fluctuations for the S&P 500 companies, before the test data had been generated by the stock market. Then we waited for two years. The training period was the 2.5-year period from 1st January 2008 to 30th June 2011. The test period was the 2-year period from 1st July 2011 to 30th June 2013.

Adjusted close values for each stock were obtained from the Datastream Worldscope Fundamentals online historical financial database. We only considered companies that were in the S&P 500 at the moment the data were drawn down, namely 30th June 2011, and which had continuous data available throughout the training period. For the

analysis of the test period, we continued to track the selected companies in the various portfolios, even if they later slipped out of the S&P 500. Companies that disappeared due to mergers or acquisitions during the test period were simply removed from the portfolios on the date their trading data ceased to be available.

The first question we addressed using this dataset was that of whether low volatility persists: Does the fact that a stock was low volatility in the past imply that it will be low volatility in the future? We found that the correlation between S&P 500 stock volatilities for the training and test periods was 0.65, $p < .001$. Clearly, volatility is something that persists. Unlike returns, it does not fluctuate randomly between periods. The next question to be answered was whether volatility-optimized portfolios tend to outperform the market into the future using a simple buy-and-hold strategy.

Optimization algorithms

It has previously been claimed that the low-volatility anomaly is specifically related to high-beta stocks, and that profits can be earned by betting against beta, that is, using leverage to buy low-beta stocks, and building up a low-beta portfolio (e.g. Black, 1993; Frazzini and Pedersen, 2014). However, if we optimize a portfolio to minimize beta, then, without exception, all of the weight is placed on the single lowest beta stock available. Intuitively, this is not desirable, because an undiversified portfolio stands a greater risk of failing to maintain its fitted features (e.g. becoming high beta). This leads us to suggest that, while often associated with low beta, the low-volatility anomaly more specifically concerns diversification, and the failure of investors to exploit it. If we optimize a portfolio to minimize volatility rather than beta, then multiple stocks will be selected, matching the intuition that diversification among several companies is a superior strategy.

In total, we implemented four different algorithms for minimizing volatility, and compared these against the S&P 500, the lowest volatility quintile of the S&P 500, and, for the test period, the S&P Low-Volatility Index. This index takes the 100 least volatile stocks in the S&P 500, and weights constituents relative to the inverse of their corresponding volatility, with the least volatile stocks receiving the highest weights. The four algorithms, which we implemented in Java, are described below in pseudocode.

far. The weighting of the first stock does not influence volatility, so it will automatically be set to the maximum weighting. The weighting of the second stock will then be amplified from zero upwards, as long as the overall volatility of the portfolio keeps dropping. Then the same process will be applied to the third stock and so forth, until all stocks have been 'tuned in' sequentially. In the pseudocode below, *WeightVec* is the weighted vector of stocks and *Index* is the S&P500 index of stocks.

Algorithm 1 Tuner

1: **procedure** GetPortfolio
2:     $WeightVec \leftarrow Repeat(0, Length(Index))$
3:     **for** i in $Range(0, Length(Index))$ **do**
4:         **while** $SD(WeightVec[i] \leftarrow WeightVec[i] + 1) <= SD(WeightVec)$ **do**
5:             $WeightVec[i] \leftarrow WeightVec[i] + 1$
6:     **return** $WeightVec$

Sequential tuner (Tuner)

This algorithm runs once for each stock in the index. It moves along the stocks in order, and the weights of each stock are incremented as long as the standard deviation of the portfolio continues to fall. If including a stock fails to lower the standard deviation, the weight is set to zero and the algorithm moves to the next stock until they all have been attempted.

For example, all of the stocks initially start off with zero investment. The algorithm considers stocks one at a time and, using a greedy strategy, will choose the weighting for each which minimizes the overall volatility so

Low correlation portfolio (Low Corr)

This algorithm runs once for each stock in the index. Each run picks a different stock to start with, and then adds stocks into the portfolio whose correlation with that starting stock is lower than some specified threshold (we used a value of 0.35). The amount of each stock added is chosen randomly. The process iterates over all stocks in the S&P 500 index, adding in those that have a low correlation with one of the stocks already in the portfolio. Once the standard deviation of the portfolio falls below some specified value (we used 1.45%), a mutator (see SD-Prob) is run to optimize its weightings.

Algorithm 2 Low Corr

1: **procedure** GetPortfolio
2: *loop*:
3:     $Portfolio.add(RandInt(0,10)$ of $Rand(Stock))$
4:     **for** $Stock$ in $Index$ **do**
5:         **for** $Items$ in $Portfolio$ **do**
6:             **if** $CORREL(Stock, Items) < Threshold$ **then**
7:                 $Portfolio.add(RandInt(0,10)$ of $Stock)$
8:     **if** $SD(Basket) < Value$ **then**
9:         **return** $Optimize(Portfolio)$
10: *GOTO loop*

Selective fixed weights (Fixed)

The numeric value of each stock's weight is fixed by its standard deviation. As per the S&P Low-Volatility Index, the higher the standard deviation, the lower the weight (given by $\frac{1}{\sigma}$). For each stock, the algorithm runs through the other stocks, and if adding them to the portfolio lowers the overall standard deviation, the portfolio is updated. In contrast to the S&P Low-Volatility Index, which includes the full lowest volatility quintile, stocks are added selectively so as to minimize the overall historical volatility.

These probabilities control the likelihood that the algorithm will select a given stock for the portfolio. Portfolios are randomly selected according to these probabilities and tested. If the resulting standard deviation is below a set threshold (we used 1.45%), then a mutator is run on the portfolio to fine-tune the weights. This mutator randomly increments or decrements the stock weights in the portfolio and saves the changes if the standard deviation is lowered.

Algorithm 3 Fixed

1: **procedure** GETPORTFOLIO
2:     $WeightVec \leftarrow [1/SD(Stock)$ **for** $Stock\ in\ Index]$
3:     **for** i in $Range(0, Length(Index))$ **do**
4:         **if** $SD(TestVec[i] \leftarrow WeightVec[i]) < SD(TestVec)$ **then**
5:             $TestVec[i] \leftarrow WeightVec[i]$
6:     **return** $TestVec$

Algorithm 4 SD-Prob

1: **procedure** GETPORTFOLIO
2:     $SDVector \leftarrow SD(Stocks)$
3:     $Prob \leftarrow [1/x^3$ for x in $SDVector]$
4: $loop:$
5:     $SamplePort \leftarrow [RandInt(0,50)$ **if** $RandReal(0,1) < x$ **for** $x\ in\ Prob]$
6:     **if** $SD(SamplePort) < SD(Lowest)$ **then** $Lowest \leftarrow SamplePort$
7:     **if** $Lowest < Threshold$ **then return** $Optimize(Lowest)$
8: $GOTO\ loop$
9: **procedure** OPTIMIZE
10:     **while** $Attempts < 50$ **do**
11:         $temp \leftarrow [x + 1$ **if** $RandReal(0,1) < 0.1$ **else** $x - 1$ **if** $RandReal(0,1) > 0.9$ **else** x **for** $x\ in\ Lowest]$
12:         **if** $SD(temp) < SD(NewVec)$ **then** $NewVec \leftarrow temp$
13:         $Attempts\ +1$
14:     **return** $NewVec$

Standard deviation probabilistic (SD-Prob)

This algorithm assigns probabilities to stocks based on their standard deviations.

The following tables show how the four algorithms compared against the market and low-volatility quintile portfolios in the training and test periods. A range of standard

Table 1: Performance of in-sample portfolios optimized to training data

	Tuner	Low Corr	Fixed	SD-Prob	Market	Quintile
# stocks	38	13	43	8	500	100
vol (annual)	15.5%	15.4%	16.7%	16.1%	28.4%	19.2%
return (ann.)	12.9%	12.8%	16.6%	20.6%	−4.5%	13.8%
Sharpe ratio	0.83	0.83	0.99	1.28	−0.16	0.72
beta	0.5	0.48	0.52	0.46	1	0.64
Treynor's ratio	0.26	0.26	0.32	0.44	−0.04	0.21
Jensen's alpha	0.32	0.32	0.41	0.49	0	0.35
tracking error	1.14	1.2	0.99	1.25	0	0.8
inform. ratio	0.32	0.3	0.45	0.43	NA	0.48
RDC	2.47	2.46	3.05	4.66	−1.15	2.11

metrics is given including annualized volatility, annualized return, Sharpe ratio (excess return to standard deviation), beta coefficient (correlated relative volatility with the market), Treynor's ratio (excess return to beta), Jensen's alpha (abnormal return over beta expectation), tracking error (standard deviation of divergence from market) and information ratio (excess return to standard deviation of tracking error).

Finally, we include the randomness deficiency coefficient metric (RDC; see Maguire *et al*, 2013), which provides an accessible overview of the extent to which excess returns deviate from chance. RDC is essentially a significance value for time series. An RDC of 0 indicates that the overall level of return fails to surpass that of the risk-free rate. A positive RDC is the average number of randomly generated time series with the same characteristics as this one that would have to be considered before coming across one with excess returns as good as this one. For example, an RDC of 3 implies that, on average, we expect to discard three similar randomly generated time series exceeding the risk-free rate before finding one better than this (slightly above chance performance). The higher the RDC, the more convincing the case that the portfolio is outperforming, not by chance, but because of some consistent edge. A negative RDC indicates that a portfolio has performed so badly that shorting it would have yielded performance above the risk-free rate.

As can be seen from Table 1, each of our four algorithms successfully identified portfolios that had lower volatility than the market, and even lower than the bottom volatility quintile. In addition, these portfolios involved far fewer stocks, ranging from 43 to only 8. There was also a strong correlation between returns and volatility, with the algorithm-derived portfolios obtaining higher RDCs than the benchmark portfolios. Figure 1 plots the time series of the market and bottom quintile versus the top-performing strategy, namely Selective Fixed Weights. Of course, the key question was whether these portfolios, trained on historical data, were either overfitted, or would maintain their outperformance into the future.

Table 2 reveals the performance of the same set of portfolios out-of-sample. Every low-volatility strategy remained lower volatility than the market into the future. Furthermore, all strategies bar one yielded greater returns than the market, and positive alphas, leading to better Sharpe ratios and RDCs. The S&P Low-Volatility Index outperformed the bottom volatility quintile, but was itself outperformed by two of our four algorithms. The best performer of all was the Selective Fixed Weights algorithm, which is very similar to the S&P Low-Volatility Index, insofar as stocks are inversely weighted relative to their volatility. However, the key difference is that in our algorithm the stocks are added selectively so as to minimize historical volatility, while the Low-Volatility
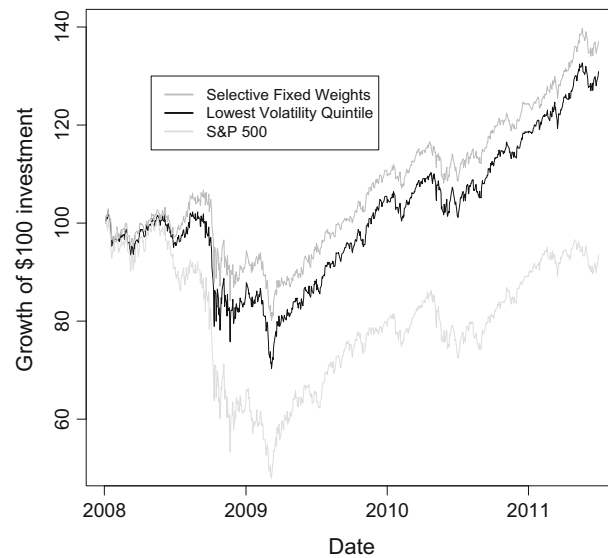
Figure 1: Performance of various portfolios for training period.

Table 2: Performance of out-of-sample portfolios optimized to training data and applied to test data

	Tuner	Low Corr	Fixed	SD-Prob	Market	Quintile	S&P Low
# stocks	38	13	43	8	500	100	100
vol (annual)	14.3%	14.7%	12.1%	13.6%	18.6%	13.4%	13.4%
return (ann.)	10.4%	9.4%	13.8%	14.9%	9.7%	13.9%	14.5%
Sharpe ratio	0.73	0.64	1.15	1.1	0.52	1.04	1.08
beta	0.66	0.66	0.59	0.58	1	0.68	0.66
Treynor's ratio	0.16	0.14	0.23	0.26	0.1	0.2	0.22
Jensen's alpha	0.08	0.06	0.17	0.2	0	0.16	0.18
tracking error	0.61	0.64	0.56	0.71	0	0.46	0.51
inform. ratio	0.02	-0.01	0.16	0.16	NA	0.2	0.21
RDC	3.02	2.57	7.62	6.64	2.07	5.89	6.56

Index includes the full lowest volatility quintile. The 43 stocks in our algorithm-derived portfolio achieved an annualized volatility of 12.1% and an RDC of 7.62, compared with 13.4% and 6.56 for the Low-Volatility Index.

Although the performance advantage of our top-performing Selective Fixed Weights algorithm over the S&P Low-Volatility Index was only marginal, this superior performance was achieved with far fewer stocks (43 versus 100), and without any updates to composition over the 2-year period, leading to significant savings as regards transaction costs. Even our SD-Prob portfolio, including only 8 stocks, managed to achieve lower volatility than the market out-of-sample, as

well as superior Sharpe ratio and RDC. The fact that the S&P Low-Volatility Index is rebalanced each quarter no doubt gives it an advantage over our buy-and-hold strategy. Further investigations are required to determine if performance can be pushed even higher through regular rebalancing.

Figure 2 plots the time series of the market and the S&P Low-Volatility Index versus our top-performing strategy. As can be seen, the Low-Volatility Index boasts slightly higher growth (14.5% versus 13.8%) but it does so as the expense of higher volatility, leading to a lower Sharpe ratio and RDC. In support of Baker *et al*'s (2011) proposal, this result suggests that the low-volatility effect can indeed be enhanced by diversifying the selection of

Figure 2: Performance of various portfolios for test period.

Table 3: Stocks in lowest volatility quintile

	Tuner	Low Corr	Fixed	SD-Prob
% Stocks in lowest quintile	78.9%	92.3%	90.6%	100%

stocks *within* a low-volatility portfolio, although further evidence is needed to bolster this conclusion.

While Baker *et al* (2011) restricted their optimization to allow only the bottom quintile of low-volatility stocks, our algorithms were able to select any stock, no matter how volatile. Table 3 shows the percentage of stocks selected by each of the algorithms that were drawn from the bottom volatility quintile. The fact that these portfolios involved some high-volatility stocks raises the question of whether overfitting could be reduced, and portfolio performance enhanced, by restricting the domain to the bottom volatility quintile. Accordingly, we recomputed portfolios for the four algorithms based only on these 100 lowest volatility stocks. The results are shown in Tables 4 and 5.

As can be seen in Table 4, our algorithms, as expected, performed less well in-sample when restricted in their choices, with most of them registering higher volatilities and lower RDCs. However, this drop in performance is not necessarily a bad thing: The reduction in choice may curtail the ability of the algorithms to overfit, with positive implications for future performance.

Table 5 describes the performance of the same portfolios out-of-sample. Again, every low-volatility strategy outperformed the market. The Selective Fixed Weights algorithm continued to outperform the S&P Low-Volatility Index fund (RDC of 7.68 versus 6.56). However, although the Tuner algorithm fared slightly better when restricted, the other algorithms had higher volatilities than before. This suggests that, in most cases, allowing higher volatility stocks to be added into a portfolio can actually be beneficial. The results do not offer support for artificially restricting the selection of stocks to the bottom volatility quintile. It is worth noting that the portfolios involving greater number of stocks (e.g. >30) outperformed those with fewer stocks (e.g. <10), suggesting that a minimum-weighting constraint might reduce overfitting for those algorithms that

Table 4: Performance of in-sample portfolios optimized to training data using low-volatility quintile

	Tuner	Low Corr	Fixed	SD-Prob	Market	Quintile
# stocks	27	9	35	10	500	100
vol (annual)	15.4%	16.2%	16.7%	15.7%	28.4%	19.2%
return (ann.)	16.4%	11.2%	14.1%	13.1%	−4.5%	13.8%
Sharpe ratio	1.06	0.69	0.84	0.83	−0.16	0.72
beta	0.48	0.49	0.55	0.5	1	0.64
Treynor's ratio	0.34	0.23	0.26	0.26	−0.04	0.21
Jensen's alpha	0.4	0.28	0.35	0.32	0	0.35
tracking error	1.18	1.23	1.01	1.2	0	0.8
inform. ratio	0.37	0.26	0.39	0.31	NA	0.48
RDC	3.39	2.07	2.5	2.48	−1.16	2.11

Table 5: Performance of out-of-sample portfolios optimized to training data using low-volatility quintile and applied to test data

	Tuner	Low Corr	Fixed	SD-Prob	Market	Quintile	S&P Low
# stocks	27	9	35	10	500	100	100
vol (annual)	13.1%	16.2%	12.7%	15.4%	18.6%	13.4%	13.4%
return (ann.)	14.3%	9.2%	14.6%	10.6%	9.7%	13.9%	14.5%
Sharpe ratio	1.1	0.57	1.15	0.68	0.52	1.04	1.08
beta	0.60	0.72	0.62	0.70	1	0.68	0.66
Treynor's ratio	0.24	0.13	0.24	0.15	0.1	0.2	0.22
Jensen's alpha	0.18	0.05	0.19	0.08	0	0.16	0.18
tracking error	0.63	0.66	0.55	0.64	0	0.46	0.51
inform. ratio	0.17	−0.02	0.2	0.03	NA	0.2	0.21
RDC	6.68	2.25	7.68	2.78	2.07	5.89	6.56

tend to identify portfolios with fewer components (see Maguire *et al*, 2014).

In sum, the results as a whole offer strong support for the idea that low-volatility portfolios can outperform the market even when following a simple buy-and-hold strategy. Every low-volatility portfolio in our study achieved higher Sharpe ratios and RDCs than the market. We performed a correlation between annualized volatility and RDC for the 25 unique portfolios involved in the study, across both training and test periods. The correlation between them was −0.79, indicating a very strong negative relationship between volatility and risk-adjusted portfolio performance: The lowest volatility portfolios performed best, and the highest volatility portfolios performed worst.

CONCLUSION

Li *et al* (2014) questioned whether the low-volatility anomaly could be exploited in practice, claiming that rebalancing requirements and transaction costs would wipe out any potential benefits. We have provided evidence that this is not the case: It is possible to develop a low-volatility buy-and-hold strategy based on historical daily returns that continues to outperform the market out-of-sample, with no need for frequent rebalancing.

We have also provided some evidence that optimized portfolios can outperform the S&P Low-Volatility Index by exploiting the different relationships that exist between stocks and reducing the overall volatility of the portfolio. Rather than relying on a covariance matrix (see Baker *et al*, 2011), we employed a variety of techniques to minimize the historical volatility of the portfolio itself. This performance was maintained into the future. Further investigations are required to find whether these results hold up over different time periods, to find the optimal training and rebalancing periods, and to find strategies which minimize the potential for overfitting.

In practice, these results imply that the capitalization-weighted S&P 500 is not the most diversified index achievable. If all investors invested optimally with full access to leverage so as to maximize their expected risk-to-reward ratio, then, as implied by CAPM, return would indeed be a linear function of beta: every possible source of diversification would be fully exploited, and long-term passive investing would be the ideal form of long-term investment. However, our results imply that it is possible to develop portfolios that are more diversified than the market, and hence outperform the market.

The efficient market hypothesis (EMH) asserts that financial markets are informationally efficient (e.g. Malkiel and Fama, 1970). Informational efficiency entails that current prices reflect all publicly available information, and that prices instantly change to reflect new public information. Accordingly, all stocks, and indeed all portfolios of stocks, should follow a random walk, where knowledge of past events has no value for predicting future prices changes (Maguire *et al*, 2014).

However, the notion of short-term individual stock efficiency described by the EMH is a separate concept to that of long-term market diversification efficiency. Even if a market is prediction-efficient it may not be diversification-efficient: Sophisticated diversifiers may be drawing down a larger proportion of the long-term risk premium the market provides, leaving other naive diversifiers to shoulder risk without the expected rewards (Maguire *et al*, 2014). The central tenet of CAPM, namely that return is linearly related to beta, assumes that the market represents the most diversified index. But if that's not true then investors who simply buy the market index may be transferring their risk premiums to sophisticated diversifiers who hold lower volatility portfolios (Maguire *et al*, 2014).

Why should diversification inefficiency persist in the market? As pointed out by Baker *et al* (2011) and Blitz and van Vliet (2007), adherence to benchmarks, lack of leverage and upmarket outperformance are all likely to play a role. While there is much focus by investment companies on exploiting short-term pricing inefficiencies, there is less awareness of long-term algorithmic diversification strategies (e.g. Maguire *et al*, 2014; Choueifaty *et al*, 2013). While pricing inefficiency can be exploited to generate quick short-term profits, diversification inefficiency only delivers over the longer term, and requires greater tenacity to exploit.

If sophisticated diversifiers are winning, then somebody else must be losing. The growing exploitation of diversification inefficiency may explain why stock market returns have fallen short of expectations in the 21st century (Maguire *et al*, 2014). Sophisticated diversifiers who exploit the low-volatility anomaly may be gaining more reward per unit of risk than naive diversifiers, who are taking on risk by buying the stock market index, yet failing to achieve the expected historical rewards (Maguire *et al*, 2014).

In conclusion, these results provide further evidence in support of the low-volatility anomaly. They reinforce the consensus that the phenomenon is due to a failure by investors to exploit diversification and reduce risk. Although it has been argued that such failure is linked to the inability to avail of leverage (e.g. Baker and Haugen, 2012; Blitz and van Vliet, 2007; Black, 1993), it is worth noting that the current study managed to outperform the returns of the market without any use of leverage. Thus, it may be that a combination of factors gives rise to the anomaly, including the emphasis on market-linked benchmarks and the motivation of asset managers to overpay for stocks that outperform in up markets.

NOTE

1. The material on diversification inefficiency presented in the concluding section draws from research presented at the IEEE CIFEr 2014 conference in London (Maguire *et al*, 2014).

REFERENCES

Ang, A., Hodrick, R., Xing, Y., Zhang, X. (2009) High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics* 91: 1–23.

Baker, N. and Haugen, R. (2012) Low risk stocks outperform within all observable markets of the world. SSRN Electronic Journal 2055431.

Baker, M., Bradley, B. and Wurgler, J. (2011) Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal* 67: 40–54.

Barberis, N. and Huang, M. (2009) Preferences with frames: A new utility specification that allows for the framing of risks. *Journal of Economic Dynamics and Control* 33: 1555–1576.

Black, F. (1993) Beta and return. *The Journal of Portfolio Management* 20: 8–18

Blitz, D. and van Vliet, P. (2007) The volatility effect. *The Journal of Portfolio Management* 34: 102–113.

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

Clarke, R., de Silva, H. and Thorley, S. (2006) Minimum-variance portfolios in the U.S. equity market. *The Journal of Portfolio Management* 33: 10–24.

Cornell, B. and Roll, R. (2005) A delegated-agent asset-pricing model. *Financial Analysts Journal* 61: 57–69.

Frazzini, A. and Pedersen, L. (2014) Betting against beta. *Journal of Financial Economics* 111: 1–25.

Haugen, R. and Baker, N. (1991) The efficient market inefficiency of capitalization-weighted stock portfolios. *The Journal of Portfolio Management* 17: 35–40.

Hawkins, D. (2004) The problem of overfitting. *Journal of Chemical Information and Modeling*. 44: 1–12.

Jiang, G., Xu, D. and Yao, T. (2009) The information content of idiosyncratic volatility. *JFQ* 44: 1.

Kumar, A. (2009) Who gambles in the stock market?. *The Journal of Finance* 64: 1889–1933.

Li, M. and Vityányi, P. (2008) *An Introduction to Kolmogorov Complexity and its Applications*. Springer, Berlin.

Li, X., Sullivan, R. and Garcia-Feijóo, L. (2014) The limits to arbitrage and the low-volatility anomaly. *Financial Analysts Journal* 70: 52–63.

Maguire, P., Moser, P., McDonnell, J., Kelly, R., Fuller, S., and Maguire, R. (2013) A probabilistic risk-to-reward measure for evaluating the performance of financial securities. In: *Computational Intelligence for Financial Engineering & Economics (CIFEr). IEEE CIFEr Conference*, pp. 102–109.

Maguire, P., O'Sullivan, D., Moser, P., and Dunne, G. (2012) Risk-adjusted portfolio optimisation using a parallel multi-objective evolutionary algorithm. In: *Computational Intelligence for Financial Engineering & Economics (CIFEr). IEEE CIFEr Conference*, pp. 1–8.

Maguire, P., Moser, P., O'Reilly, K., McMenamin, C., Kelly, R., and Maguire, R. (2014) Maximizing positive portfolio diversification. In: *Computational Intelligence for Financial Engineering & Economics (CIFEr). IEEE CIFEr Conference*, pp. 174–181.

Malkiel, B. G. and Fama, E. F. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25(2): 383–417.

Sensoy, B. (2009) Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics* 92: 25–39.

Shefrin, H. and Statman, M. (2000) Behavioral portfolio theory. *The Journal of Financial and Quantitative Analysis* 35: 127.

Tversky, A. and Kahneman, D. (1992) Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.