

Diversifying the Diversifiers and Tracking the Tracking Error: *Outperforming Cap-Weighted Indices with Limited Risk of Underperformance*

Noël Amenc, Felix Goltz, Ashish Lodh, and Lionel Martellini



The Voices of Influence | iijournals.com

Diversifying the Diversifiers and Tracking the Tracking Error: *Outperforming Cap-Weighted Indices with Limited Risk of Underperformance*

NOËL AMENC, FELIX GOLTZ, ASHISH LODH, AND LIONEL MARTELLINI

NOËL AMENC

is a professor of finance at EDHEC Business School and a director of EDHEC-Risk Institute in London, UK. noel.amenc@edhec-risk.com

Felix Goltz

is head of applied research at the EDHEC-Risk Institute at EDHEC Business School in Nice, France. felix.goltz@edhec-risk.com

ASHISH LODH

is a quantitative analyst with the EDHEC-Risk Institute at EDHEC Business School in Nice, France. ashish.lodh@edhec-risk.com

LIONEL MARTELLINI

is a professor of finance at EDHEC Business School and scientific director of EDHEC-Risk Institute in Nice, France. lionel.martellini@edhec.edu

odern portfolio theory states that investors should allocate their wealth between a tangency portfolio, or maximum Sharpe ratio (MSR) portfolio, and a riskless asset. In practice, when trying to follow this advice, one obviously has to come up with proxies because neither the true tangency portfolio nor a perfectly risk-free asset can be exactly implemented in practice. Traditionally, equity investing has, however, heavily drawn on the idea of the Tobin separation theorem, and cap-weighted equity indices have long been perceived by many practitioners as reasonable proxies for the tangency portfolio. But a consensus is slowly emerging that marketcap-weighted indices tend to be poorly diversified portfolios that are not good proxies for the tangency portfolio. This result is hardly a new finding, because early attempts to provide evidence that cap-weighted portfolios are not well-diversified portfolios and thus lead to an inefficient risk-return trade-off can be traced as far back as Haugen and Baker [1991] or Grinold [1992]. Intuitively, the fact that capweighted indices are inefficient and poorly diversified is perhaps not surprising because they concentrate heavily in the largest market-cap stocks as a result of their one-dimensional construction mechanism that only takes into account a stock's market cap and thus does not allow for any mechanism that can enforce proper diversification.

Following such early criticism of capweighted equity portfolios, more recent papers have documented that cap-weighted portfolios suffer from numerous shortcomings, and various alternative weighting schemes have been proposed to improve on cap weighting; see Amenc et al. [2011], Arnott, Hsu, and Moore [2005], Choueifaty and Coignard [2008], and Maillard, Roncalli, and Teiletche [2008], to name but a few.

Although it is now commonly accepted that moving away from cap weighting tends to enhance diversification and increase riskadjusted performance over long horizons, it has to be recognized that each alternative weighting scheme will expose an investor to two related types of risk, namely, *model selection risk* and *relative performance* risk.

Considering model selection risk, it is clear that choosing a weighting scheme corresponds to choosing a model of optimal portfolio construction. This is the case, in fact, even if a weighting scheme does not explicitly refer to portfolio optimization. In fact, any weighting scheme can be understood as reflecting a set of assumptions under which the resulting portfolio would lead to an optimal portfolio in the sense of modern portfolio theory; see Martellini [forthcoming] or Melas and Kang [2010]. From a pragmatic perspective, it seems reasonable to assume that different market conditions may favor different assumptions, and thus alternative weighting schemes may display different performance depending on market conditions. In fact, when analyzing the performance of alternative weighting schemes, it is indeed the case that performance differences can be pronounced, as evidenced by the results shown in Exhibit 1, which uses data for four popular non-cap-weighted indices in the U.S. universe over a relatively short time horizon for which data for all of these indices are available.

It appears that even in a relatively short span of nine years no single strategy consistently outperforms all other strategies, even though all four strategies show outperformance with respect to cap-weighted indices more often than they show underperformance. Considering half-year returns, depending on the market conditions, each of the four strategies has ex post been the best-performing strategy for *some* subperiod. Also, the worstperforming strategy in one subperiod can be the best performer in the subsequent subperiod, and vice versa. For example, across two half-year periods from July 2008 to June 2009, the equal-weighted strategy went from underperforming the cap-weighted index by -4.53% to overperforming the same index by 11.17%, while the minimum-volatility strategy followed the opposite trend. This kind of behavior points to the fact that each model behaves well only in a certain kind of market condition. Another noteworthy result is that the difference between returns of the best- and worst-performing strategies is substantial and can be as large as 15%. This means that no model can pretend to be uniquely superior. Rather, different models are apparently favored by different market conditions, and there is always a risk that the chosen model may not yield attractive performance in a given period.

For investors who are agnostic about either their capacity to identify the model with superior assumptions or their capacity to take the risk of choosing a particular model in the wrong market conditions, it may be reasonable to assess whether anything can be gained from combining models and thus diversifying model selection risk; see, for example, Kan and Zhou [2007] for a related discussion about combining portfolio strategies. We explore this question in some detail in the next section.

E X H I B I T **1** Relative Return of Different Alternative Weighting Schemes

Subperiod	S&P 500 Equal Weight Index	MSCI USA Minimum Volatility Index	FTSE EDHEC Risk Efficient U.S. Index	FTSE RAFI U.S. 1000 Index	Winning Index– Losing Index
2003 (Jan–Jun)	4.27%	-4.72%	2.46%	0.98%	8.99%
2003 (Jul-Dec)	6.15%	-2.93%	3.46%	3.64%	9.09%
2004 (Jan-Jun)	2.86%	0.05%	3.28%	0.70%	3.24%
2004 (Jul-Dec)	2.76%	2.64%	3.18%	2.91%	0.54%
2005 (Jan-Jun)	0.84%	4.58%	3.95%	1.63%	3.74%
2005 (Jul-Dec)	2.26%	-3.85%	2.04%	-0.49%	6.11%
2006 (Jan-Jun)	1.33%	-0.70%	0.82%	2.84%	3.54%
2006 (Jul-Dec)	-1.44%	-2.54%	-2.30%	0.65%	3.18%
2007 (Jan-Jun)	1.91%	-1.63%	2.06%	1.20%	3.68%
2007 (Jul-Dec)	-5.69%	1.51%	-3.31%	-3.44%	7.20%
2008 (Jan-Jun)	1.67%	0.38%	1.46%	-3.40%	5.07%
2008 (Jul-Dec)	-4.53%	10.10%	-2.40%	-0.70%	14.63%
2009 (Jan-Jun)	11.17%	-4.10%	7.58%	7.22%	15.27%
2009 (Jul-Dec)	7.30%	-3.75%	6.96%	7.22%	11.06%
2010 (Jan-Jun)	3.68%	1.67%	3.73%	3.00%	2.06%
2010 (Jul-Dec)	2.00%	-2.06%	2.26%	0.81%	4.33%
2011 (Jan-Jun)	2.10%	4.01%	3.39%	0.05%	3.96%
2011 (Jul-Dec)	-4.17%	6.36%	-2.32%	-2.06%	10.53%

Notes: The exhibit shows the excess return over the cap-weighted index (S&P 500 Index) of the S&P 500 Equal Weight Index, the MSCI USA Minimum Volatility Index, the FTSE EDHEC Risk Efficient U.S. Index, and the FTSE RAFI U.S. 1000 Index computed over 16 half-year subperiods. For each subperiod, the highest and lowest excess returns across the strategies are highlighted in bold black and bold grey, respectively, and the difference between the best- and worst-performing indices are reported in the last column. Weekly return data from January 3, 2003, to December 30, 2011, are used for the analysis. Considering the second issue, relative performance risk, alternative weighting schemes also lead a priori to an exposure to risk of substantial deviations with respect to cap-weighted benchmarks, because they lead to choices of factor exposure that are different from those of cap-weighted indices. For example, fundamentalsbased indices have been shown to lead to a strong value bias (see, for example, Jun and Malkiel [2007], Kaplan [2008], Blitz and Swinkels [2008], and Amenc, Goltz, and Le Sourd [2009]), which has a positive impact when value stocks underperform growth stocks, but a negative impact otherwise.

The risk premia associated with the factors that alternative weighting schemes are exposed to, and which differ from the exposures of cap-weighted indices, may at times lead to less reward than the risk premia associated with cap weighting. Therefore, although many alternative weighting schemes have been shown to outperform cap-weighted indices over long time periods, no guarantee exists that in the short term these alternative weighting schemes always outperform cap weighting, as the negative numbers that we report in columns 2–4 of Exhibit 1 show. Exhibit 2 complements this analysis and shows the maximum drawdown relative to the capweighted index for the aforementioned popular alternatively weighted equity indices in the U.S. universe. Even over the relatively short time frame for which data are available for all of these indices, periods of very pronounced underperformance occur for each of the indices, with an extreme tracking error that can be as high as almost 12% for some of the alternative indices.

The presence of significant levels of relative risk may not be surprising, because all of these strategies need to deviate from the default cap-weighting scheme to generate outperformance. But that this risk, with respect to the performance of cap-weighted indices, is not controlled at all seems problematic for investors who maintain some reference or comparison with cap weighting at some stage of their investment process. In particular, this relative risk is a severe concern for a chief investment officer (CIO) who has made the choice of adopting an alternative weighting scheme. When such underperformance occurs with active managers, the failure of a third-party manager-a risk inherent in the very logic of the delegation process of portfolio managementtypically translates into the termination of the manager; see Goyal and Wahal [2008] for a description of the manager selection and termination process. In the case of underperformance of an alternative equity index, however, it would be difficult for the CIO to blame anyone but himself for the selection of the index.

Thus, CIOs who deviate from cap-weighted indices take on considerable reputational risk, because capweighted indices represent a common reference for the

ЕХНІВІТ **2**

Relative Risk of Alternative Weighting Schemes

Risk Measures	S&P 500 Equal Weight Index	MSCI USA Minimum Volatility Index	FTSE EDHEC Risk Efficient U.S. Index	FTSE RAFI U.S. 1000 Index
Max Relative Drawdown	13.60%	12.23%	8.43%	12.67%
Start Date	02/23/07	11/21/08	04/21/06	06/29/07
End Date	11/21/08	04/23/10	11/21/08	03/06/09
Annualized Excess Return over				
Cap-Weighted Index	3.24%	1.21%	3.75%	2.18%
Tracking Error (TE)	4.64%	5.69%	3.82%	4.78%
Extreme Tracking Error (95th				
percentile of rolling one-year TE)	9.57%	7.42%	6.72%	11.92%

Notes: The exhibit shows the historical extremes of underperformance, annualized excess return over the cap-weighted index (S&P 500 Index), annualized tracking error and extreme tracking error of the equal-weighted index (S&P 500 Equal Weight Index), MSCI USA Minimum Volatility Index, FTSE EDHEC Risk Efficient U.S. Index, and FTSE RAFI U.S. 1000 Index with respect to the S&P 500 Index. Maximum Relative Drawdown is the maximum drawdown of the long–short index whose return is given by the fractional change in the ratio of the strategy index to the benchmark index. Extreme Tracking Error corresponds to the 95th percentile of rolling one-year tracking error (i.e., the annualized standard deviation of a portfolio long in the alternatively weighted index and short in the S&P 500 Index) over the entire horizon. All statistics are annualized and are based on weekly data from January 3, 2003, to December 30, 2011. peer group of CIOs. Interestingly, all alternative index providers refer to cap weighting when analyzing the performance and risk of their method. This is perhaps the clearest indication that cap-weighted indices remain the ultimate reference and are likely to remain so for the foreseeable future. If cap-weighted indices are the ultimate reference, a strong case should be made for an explicit focus on relative risk, even though relative risk is rarely assessed by providers of alternatively weighted indices. In fact, active managers have been managing their portfolios subject to relative risk budgets for decades, being constrained to respect a level of tracking error relative to the cap-weighted index. The absence of such relative risk budgets for alternative weighting schemes means that they somewhat resemble active management with an unlimited relative risk budget. Likewise, comparing the excess returns of different non-cap-weighted indices resembles a comparison among active managers without any reference to the risk budgets which have been allocated to them.

From this perspective, it is misleading that providers of alternative indices promote their indices on the basis of the alleged overperformance with respect to cap-weighted indices, when no systematic process is in place to ensure that the relative risk associated with such potential for outperformance is being explicitly managed. In fact, the need to ensure proper tracking error control is all the more acute because the relatively short time span for which live performance of the alternative indices is available makes it impossible for an investor to statistically assess whether the reported performance emanates from skill, which is required to ensure better diversification, or luck, which can allow for the risk exposure to risk factors at the right time. Only the introduction of an explicit process for measuring and managing the difference in factor exposure between the alternative index and the cap-weighted index can ensure that the outperformance is based on skill and not luck and also ensure that downside relative risk is limited.

Ultimately, when a manager is faced with the objective of generating attractive risk-adjusted returns, taking into account both absolute risk (volatility or value at risk) and relative risk (tracking error or, perhaps more importantly, value at tracking-error risk, which is the risk of substantially underperforming the reference), the manager should adopt an effective approach for diversification as well as for relative risk control. Diversification, which is the only free lunch in finance, allows investors—for a given level of volatility—to maximize

expected returns. Relative risk control allows investors to reap these benefits of diversification without straying too far from their peer group represented by standard cap-weighted indices. Improved diversification and relative risk control are nothing but ways of trying to achieve efficient spending of the investor's risk budget.¹

In this article, we explore two noncompeting approaches aimed at achieving robust outperformance over cap-weighted indices when explicitly recognizing the presence of model risk and relative risk. The first approach consists of "diversifying the diversifiers." This approach allocates across alternative weighting schemes in order to diversify away the risk of making incorrect assumptions about the conditions of optimality behind different strategies in different market conditions. We find, in particular, that a combination of a minimum-volatility portfolio and a maximum Sharpe ratio portfolio, which differ from each other in performance in various market conditions, has a higher probability of outperformance compared to the cap-weighted index than either of its component strategies. The combination portfolio also has a lower level of average tracking error risk. Extreme tracking error risk, however, remains substantial, as expected, because diversification is well known to fail at managing extreme risk, which manifests itself as a situation when all available options tend to fail simultaneously (Amenc and Martellini [2011]).

The second approach can be characterized as "tracking the tracking error." That is, the approach implements an explicit relative risk control mechanism as a means of reducing the consequences of a strategy's severe short-term underperformance relative to the standard cap-weighted index; this appears to be a minimum requirement for any strategy that will be subject to peer group comparisons. We find that the worst annual performance difference between the cap-weighted index and a diversified portfolio is dramatically improved when integrating a suitably designed relative risk control process; however no proportional decrease in expected outperformance is observed. When put together, these two ingredients allow for robust access to outperformance within well-defined relative risk budgets.

DIVERSIFICATION ACROSS EQUITY PORTFOLIO OPTIMIZATION STRATEGIES

In practice, there can be numerous proxies for optimal portfolios that rely on more or less restrictive conditions of optimality. In this section, we focus on empirical results for two remarkable proxies of efficient frontier portfolios and leave aside ad hoc weighting schemes that weight stocks by more or less arbitrary characteristics, such as dividends or revenues, because the conditions for optimality of such schemes are unclear. A first proxy is to try to obtain the minimum-risk portfolio on the efficient frontier. This global minimumvariance (GMV) portfolio, which is suboptimal from an ex ante perspective, could well be ex post a good proxy for the maximum Sharpe ratio portfolio, because it avoids estimation error from expected return estimates. The other remarkable portfolio is an explicit MSR proxy that capitalizes on differences in expected returns across stocks in addition to differences in risk parameters. Before discussing the actual results, we first describe the data underlying this analysis, as well as the construction of improved equity portfolios.

Data and Portfolio Construction Methodologies

The data we use in this article are the weekly returns of all S&P 500 stocks obtained from the CRSP database. The constitution of the S&P 500 has been available on CRSP since 1959, and all equity return time series are weekly total returns (including reinvestment of periodic payments such as dividends). We avoid any survivorship bias in our optimized portfolios, and we ensure a match of the constitution with the cap-weighted index by selecting at the start of each quarter the stocks that are present in the S&P 500 at that time.² The time horizon for all portfolio returns extends from January 2, 1959, to December 31, 2010. Quarterly rebalancing is performed for the construction of minimum-volatility and maximum Sharpe ratio portfolios. At each rebalancing, weekly returns over the last two years are used for optimization. After the weightings are determined, we use stock returns for the next quarter to calculate out-ofsample portfolio returns. The cap-weighted portfolio is the S&P 500 Index as published by CRSP.

This section describes portfolio optimization and constraints of the optimized strategies. As explained in the introduction, we focus on two notable efficient frontier portfolios: the global minimum-variance (GMV) minimum-volatility portfolio and the maximum Sharpe ratio portfolio. We do not seek to include ad hoc weighting schemes, such as equal weighted or fundamentals weighted, in the analysis, because under reasonable assumptions they are unlikely to be good proxies for optimal portfolios.

GMV portfolio. The concept of minimum-variance investing gained popularity after the early work of Haugen and Baker [1991] who argued that the capweighting scheme is highly inefficient and techniques exist to obtain similar expected returns with significantly lower volatility; also see related results by Schwartz [2000] and Jagannathan and Ma [2003]. The minimum-volatility portfolio is, in fact, the optimal portfolio (MSR portfolio) under the restrictive assumption that all stocks have identical expected returns. The mean–variance optimization problem under such assumption reduces to

min $(w^{T} \times \sum \times w)$ subject to $1^{T} \times w = 1$

where Σ is the covariance matrix of stock returns and w is the weight vector. In the absence of constraints, the solution for the preceding minimization problem can be derived analytically as

$$w^* = \frac{\sum^T \times 1}{1^T \times \sum^T \times 1}$$

The two most important factors in the practical construction of a minimum-volatility portfolio are the choice of a method to estimate the covariance matrix and the set of constraints on weights. The standard sample-based covariance matrix estimate is extremely noisy; see, for example, Chan, Karceski, and Lakonishok [1999]. As a result, researchers have assessed different ways of improving covariance matrix estimation, such as increasing the data frequency, imposing structure on the estimator by imposing, for example, factor structure (Chan, Karceski, and Lakonishok), or using optimal shrinkage approaches to mix a sample estimator with a structural estimator; see Ledoit and Wolf [2003, 2004]. Another line of thought is to impose rigid restrictions directly on the asset weights, such as short-sale constraints (see Frost and Savarino [1988] and Jagannathan and Ma [2003]), or more *flexible* "norm constraints" as proposed by DeMiguel et al. [2009]. Jagannathan and Ma [2003], among others, showed that applying such constraints to the outputs of the variance minimization (i.e., the weights) is equivalent to shrinking the inputs (i.e., the covariance matrix).

In the empirical exercises that follow, we set an upper bound of λ/N and a lower bound of $1/\lambda N$ for all stock weights to ensure that all constituents are represented without an extreme concentration in any single stock.³ λ acts as a parameter that defines the degree of similarity with the equal-weighted scheme (with $\lambda = 1$ corresponding to an equal-weighted portfolio). High values of λ correspond to looser constraints. Next, we will test the constraint level $\lambda = 6$. In tests of tighter constraints corresponding to $\lambda = 4$ we obtain qualitatively similar results; we do not report these results.⁴

When imposing a factor structure on the covariance matrix, an attractive candidate for a factor model is to extract factors using principal components analysis (PCA); see Fujiwara et al. [2006] for evidence that using PCA-based covariance matrix estimates in portfolio optimization leads to attractive out-of-sample performance. PCA generates the factors from the data that have the most explanatory power. The advantage of using such statistical factors is that the factor selection risk inherent in choosing explicit factors can be avoided. Also, the PCA method allows for a considerable reduction in the number of factors, because they are by construction orthogonal to each other. In this article, we use PCAbased covariance matrix estimators that limit the number of factors through an optimal selection criterion based on random matrix theory; see Amenc et al. [2011].

Risk-based MSR portfolio. Based on the aforementioned elements, we define the minimum-volatility portfolios that we test in this article. An alternative to minimizing volatility is to aim explicitly at maximizing the risk-adjusted performance (Sharpe ratio). This leads to the following optimization problem: to

$$\max\left(\mathbf{W} \times \frac{\boldsymbol{\mu}}{(\mathbf{W}^{\mathrm{T}} \times \boldsymbol{\Sigma} \times \boldsymbol{W})}\right) \quad \text{such that } \mathbf{1}^{\mathrm{T}} \times \mathbf{W} = \mathbf{1}$$

The difference with respect to the minimum-volatility portfolio is that Sharpe ratio maximization requires inputs for expected returns. To create a proxy for the tangency portfolio, we follow the *efficient indexation approach* of Amenc et al. [2011]. Because direct return estimation is prone to large errors (see, for example, Britten-Jones [1999]), Amenc et al. proposed a parsimonious risk-based return estimation procedure. This procedure, based on the fundamental principle of a risk-return trade-off, advocates the use of a downside risk measure (we use downside volatility or semi-deviation, that is, the standard deviation of negative returns) as a proxy for expected return. In order to ensure robustness, all stocks are sorted by semi-deviation and attributed to decile portfolios. The median value of semi-deviation in each decile is then assigned to all stocks in that decile portfolio.

Combination of strategies. Although we use all resources at hand to reduce estimation error while constructing the tangency portfolio, a fair amount of parameter uncertainty exists. Kan and Zhou [2007] argued that, in the presence of parameter uncertainty, a two-fund theorem, which states a combination of the tangency portfolio and the risk-free asset results in the best risk-reward ratio, does not hold. If the investor wants to improve out-of-sample performance, he is better off holding a three-fund portfolio: a combination of the risk-free asset, the tangency portfolio, and the global minimum-variance portfolio. The intuition behind this approach is that estimation errors of the two optimized portfolios are not perfectly correlated. Thus, there is room for diversifying the model risk inherent in choosing either the GMV or the MSR portfolio. In other words, because the minimum-volatility and risk-based maximum Sharpe ratio portfolios are, in fact, quite different strategies by construction and may display differences in performance properties, we can combine the strategies and perhaps obtain a portfolio with even better diversification. Drawing on this insight, we also test a combination of our minimum-volatility and efficient portfolios. The combination is a simple equal-weighted portfolio across these two strategies, for which we set the weights of the GMV or MSR portion of the overall portfolio equal to 50% at each quarterly rebalancing date.

Liquidity and turnover constraints. In addition to achieving robust parameter estimation, the optimized portfolios also have to be implementable in practice in order to be relevant. We address implementation issues through two elements: liquidity rules and turnover control.

We apply two liquidity rules to ensure that the portfolios have sufficient liquidity. After the optimal weights are obtained, a cap on the weights of certain stocks is applied. At each rebalancing, the change in weight of each stock is capped at its market-cap weight. This allows us to avoid large rebalancings in the smallest stocks without having to limit the total weight that can be invested in the less liquid stocks. Second, the weight of each stock is capped at a multiple of 10 of its market-cap weight. We do this to avoid a large investment in the smallest stocks, which would be impractical. In practice, these capping rules concern less than 5% of stocks in the S&P 500 universe, and their application allows us to maintain portfolios that are very similar to the original optimized portfolios while improving portfolio liquidity.

To achieve low turnover, we set a rebalancing threshold. More precisely, we impose a rule that the portfolio is not rebalanced until two-way turnover reaches a threshold of 70% in a given quarter. The idea behind this is to avoid having to rebalance when newly optimized weights deviate from the current weights by a relatively small amount. This technique is inspired by Leland [1999], who formulated a "no-trade interval" of weights to reduce portfolio turnover in the presence of transaction costs and showed that this method can substantially reduce transaction costs when compared to a simple periodic rebalancing strategy.⁵

Empirical results: Descriptive statistics. We compare standard measures of risk and return, as well as risk-adjusted performance ratios of different optimized portfolios and a cap-weighted portfolio of S&P 500 constituents. In particular, for the optimized strategies, we include the minimum volatility and the MSR portfolios, as well as the diversified strategy that combines both approaches.

Long-term risk and return analysis. Panel A of Exhibit 3 reports the longterm performance statistics for the different strategies. Panel B summarizes the long-term risk–reward differences over the cap-weighted index and indicates their significance levels.⁶

Clearly, all the optimized portfolios outperform the cap-weighted index in terms of average annual return, volatility, and Sharpe ratio. The minimum-volatility technique provides less volatility and less annualized return than the max Sharpe ratio proxy, but both strategies deliver almost the same

Sharpe ratio (0.39), a value substantially higher than the Sharpe ratio for the cap-weighted index (0.27). This finding confirms that both portfolios are reasonably good proxies for the optimal portfolio, which dominates the

E X H I B I T 3 Performance Statistics and Risk–Reward Difference Compared to S&P 500

	Cap Weighted	Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)
Panel A: Performance Statistic	cs			
Ann. Return	9.60%	10.89%	11.02%	10.95%
Ann. Std	15.46%	14.10%	14.53%	14.31%
Ann. Semi-Dev	11.28%	10.39%	10.73%	10.55%
Tracking Error	0.00%	2.76%	2.77%	2.72%
Market β	1.00	0.90	0.93	0.91
Sharpe Ratio	0.27	0.39	0.39	0.39
Sortino Ratio	0.39	0.55	0.55	0.55
Information Ratio	—	0.47	0.51	0.50
Treynor Ratio	0.04	0.06	0.06	0.06
Max Drawdown	53.83%	50.42%	51.72%	51.07%
95% VaR	3.31%	3.02%	3.14%	3.08%
99% VaR	7.77%	7.64%	7.77%	7.70%
Skewness	-0.34	-0.50	-0.52	-0.51
Kurtosis	8.06	9.09	8.82	8.95
95% Value at Tracking-Error				
Risk (VaTER)	-	3.89%	3.96%	3.87%
Panel B: Difference over the C	Cap-Weighted	Index		
Diff. in Returns		1.29%	1.42%	1.35%
P-value		1.14%	0.27%	0.49%
Diff. in Volatility	—	-1.36%	-0.93%	-1.15%
P-value	—	0.00%	0.13%	0.01%
Diff. in SR	_	0.12	0.11	0.12
P-value	-	0.40%	0.16%	0.14%

Note: Panel A—Performance statistics shows summary statistics of different equity portfolios, namely, cap-weighted (S&P 500 Index), minimum-volatility, maximum Sharpe ratio, and diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios. Annualized Semi-Deviation is below-mean semi-deviation and the minimum acceptable return to compute the Sortino ratio is the risk-free rate. A Cornish–Fisher expansion is used to compute value at risk. The 95% value at tracking-error risk statistics are annualized by multiplying the weekly statistics by $52^{0.5}$.

Panel B—The panel shows differences in average returns, in volatility, and in Sharpe ratios between each index and the cap-weighted S&P 500 Index and the associated P-values. All differences are computed from annualized statistics and the average returns are a geometric average. The P-values for differences are computed using a paired t-test for the average returns, a Fisher test for the volatility, and bootstrap method for the Sharpe ratio. Differences that are significantly different from zero at a 5% confidence level are indicated in bold. The period of analysis is from January 2, 1959, to December 31, 2010. All statistics are annualized, and performance ratios that involve average returns are based on the geometric average.

cap-weighted indices over long periods of time. A closer analysis unveils that the difference in return compared to the cap-weighted index is statistically significant for both strategies and their 50–50 diversified portfolio, but the difference is more significant (P-values closer to zero) for the MSR proxy than for the minimum-volatility proxy. Similarly, the minimum-volatility strategy results in a more significant lowering of volatility (lower P-values) than the MSR portfolio, which lowers volatility, but with less magnitude and less statistical significance (higher P-values). This phenomenon is explained by the fact that the minimum-volatility strategy explicitly aims to lower portfolio volatility. In the end, the comparable improvement in Sharpe ratio is achieved by the two proxies in two different ways: the MSR portfolio has a greater focus on improving performance, and the minimum-volatility portfolio has a greater focus on decreasing volatility.⁷

The benefits of optimized weighting schemes are obvious from not only the Sharpe Ratio but also other reward-to-risk ratios. For example, all alternatively weighted strategies have a considerably higher Sortino ratio (0.55) than the cap-weighted index (0.39). Another interesting finding is that the optimized portfolios do not increase the maximum drawdown levels compared to the cap-weighted indices. Although the results suggest that the proxies for the GMV and MSR strategies that we test generate similar improvements in Sharpe ratios, these indices may well show pronounced differences in terms of how returns behave over time, which is what

we turn to now. This is, in essence, what motivates the introduction of the diversified approach (50% GMV + 50% MSR), whose performance will be discussed in the next section.

Short-term differences between optimized strategies and the performance of a diversified approach. Our proxies for the GMV and MSR strategies rely on the same covariance matrix estimate, but they have different assumptions on expected returns and show differences in terms of their return behavior, in particular, portfolio beta. In this subsection, we analyze in greater detail the performance of minimum-volatility and maximum Sharpe ratio portfolio proxies to qualify how different these two strategies can be in different market conditions.

To achieve this goal, we compare the performance statistics of all alternative strategies, conditional on market environment. In particular, we report the excess returns and volatilities of the MSR and GMV portfolio proxies compared to the cap-weighted portfolio, using equity market excess performance (based on the capweighted proxy) as a conditioning variable. Using weekly data, annualized returns and annualized volatilities are computed for all portfolios for each quarter, resulting in 208 quarterly values. We sort these values by market excess returns and divide them into quintiles. We compute the geometric mean of returns and the arithmetic mean of volatilities for each quintile for all strategies. In addition to the minimum-volatility and maximum Sharpe ratio portfolio proxies, we also report the results for the diversified portfolio, which is rebalanced quarterly to an equal-weighted mix of minimum-volatility and maximum Sharpe ratio portfolio proxies.

Exhibit 4 helps us draw a clear conclusion on the sensitivity to bull and bear markets of the optimized strategies. In conditions characterized by low (meaning strongly negative) market returns, minimum volatility outperforms the market by a wide margin but underperforms dramatically when the market is extremely bullish, as is expected in both cases. The maximum Sharpe ratio, although it has a similar tendency to add more value in bear markets, shows more stable outperformance. In quintiles sorted by market volatility, minimum volatility performs better, in

E X H I B I T **4** Performance in Different Market Conditions: Return Difference with Cap-Weighted Index

		Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)
Conditioning Variable	Regime		Excess Return over S&P 500 Index	
Annualized Equity Market	Low	3.20%	2.06%	2.63%
Excess Returns (Market-	2	2.62%	1.95%	2.29%
Risk-Free Rate)	3	0.78%	0.98%	0.88%
	4	-0.50%	0.94%	0.22%
	High	-2.62%	-0.10%	-1.36%
Annualized Market	Low	1.40%	2.53%	1.96%
Volatility (of Excess	2	1.01%	1.03%	1.02%
Market Return)	3	-0.10%	-0.07%	-0.08%
	4	1.64%	2.05%	1.85%
	High	2.24%	1.58%	1.91%

Notes: This exhibit shows the excess returns of the minimum-volatility, maximum Sharpe ratio, and diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios over the S&P 500 Index in different market conditions characterized by quintiles sorted by annualized market returns and annualized market volatility. Quintiles are formed based on quarterly returns that are computed using weekly returns from January 2, 1959, to December 31, 2010. general, in high-volatility periods, whereas the maximum Sharpe ratio portfolio proxy adds most value in quarters with low volatility.

Results from this analysis can be useful from a tactical perspective for investors who want to switch dynamically between different strategies depending on their views about the likely market regime. More important, perhaps, in the likely absence of strong predictive power over future market performance, these results can be used from a strategic perspective to support the idea of holding a mixture of both proxies, as opposed to selecting either one. "Diversifying the diversifiers" appears as an efficient alternative to "timing the diversifiers," which would result in a wider range of outcomes depending on timing skills and uncertainty in market conditions. To illustrate this, we simulate the performance of a dynamic switch between the GMV and MSR portfolios over the period January 1959–December 2010, considering two extreme versions: the *always-right timer*, who always selects each quarter the strategy that outperforms and holds it for the respective quarter, and the always-wrong timer, who systematically selects the strategy that underperforms. The results show a Sharpe ratio of 0.46 for the perfect timer versus 0.31 for the alwayswrong timer, suggesting a wide dispersion around the

Ехнівіт 5

Re	lative	Average	Risk in	Different	Market	Conditions:	Tracking	Error

		Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)
Conditioning Variable	Regime		Tracking Error with S&P 500 Index	
Annualized Equity Market	Low	3.28%	3.30%	3.23%
Excess Returns (Market-	2	2.12%	2.17%	2.10%
Risk-Free Rate)	3	1.81%	1.82%	1.77%
	4	2.21%	2.18%	2.14%
	High	2.37%	2.54%	2.41%
Annualized Market	Low	1.55%	1.72%	1.59%
Volatility (of Excess	2	1.78%	1.78%	1.73%
Market Return)	3	1.99%	2.08%	1.99%
	4	2.45%	2.50%	2.42%
	High	3.97%	3.89%	3.87%

Notes: This exhibit shows the mean annualized tracking error with respect to the S&P 500 Index of the minimum-volatility, maximum Sharpe ratio, and diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios in different market conditions characterized by quintiles sorted by annualized excess market returns. Quintiles are formed based on quarterly returns that are computed using weekly returns from January 2, 1959, to December 31, 2010.

0.39 Sharpe ratio achieved with the diversified approach that holds 50% of each strategy at all times.

In fact, we find that while the mixture of GMV and MSR has an expected excess performance over the cap-weighted benchmark given by the average of the two, its relative risk is improved with respect to both. This result holds over the long term, as Exhibit 3 shows, because the tracking error of the mixture is lower than the tracking error of both the MSR and GMV portfolios. The result also holds across various market conditions, as confirmed by the results reported in Exhibit 5, that is, the volatility of the return differences over the market cap-weighted portfolio (i.e., the tracking error) across the various regimes. In most cases we again find that the tracking error of the diversified portfolio is lower than that of both GMV and MSR portfolios. This hardly surprising result confirms that some diversification occurs with respect to the uncertainty in outperformance by each of the two schemes. In unreported results, we also find that the diversified portfolio also enjoys a lower probability of underperformance with respect to the cap-weighted index compared to each of its components taken individually.

The combined portfolio, as Exhibit 3 shows, hardly exhibits any improvement in terms of extreme tracking

error. The value at tracking-error risk (VaTER) for the diversified strategy and its two components is comparable when analyzing the entire time period. This result is confirmed by Exhibit 6, which reports VaTER numbers assessed as conditional on market returns and market volatility. The results reported in the exhibit show that in 9 of 10 cases, the diversified strategy does not lower VaTER compared to the individual strategy with the lowest VaTER.

Overall, the results for the conditional analysis (reported in Exhibits 5 and 6) confirm the findings based on the entire time period (reported in Exhibit 3): diversifying across MSR and GMV effectively reduces the tracking error, a measure of average deviation from the benchmark, but it is ineffective in reducing the extreme relative risk in terms of value at tracking-error risk. This result simply suggests that diversification may not be a

E X H I B I T **6** Relative Extreme Risk in Different Market Conditions: 95% Value at Tracking-Error Risk

		Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)
Conditioning Variable	Regime		95% Value at Tracking-Error Risk	
Annualized Equity Market	Low	5.13%	4.99%	5.13%
Excess Returns (Market-	2	2.92%	3.11%	2.89%
Risk-Free Rate)	3	3.18%	3.03%	2.99%
	4	3.65%	3.50%	3.52%
	High	4.95%	4.66%	4.71%
Annualized Market	Low	2.39%	2.83%	2.54%
Volatility (of Excess	2	3.02%	2.86%	2.86%
Market Return)	3	3.33%	3.59%	3.50%
	4	4.06%	4.05%	3.90%
	High	7.10%	6.77%	7.11%

Notes: This exhibit shows annualized 95% value at tracking-error risk (based on weekly relative returns) with respect to the S&P 500 Index of the minimum-volatility, maximum Sharpe ratio, and diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios in different market conditions characterized by quintiles sorted by annualized excess market returns and annualized volatility of excess market returns. All statistics are annualized by multiplying the weekly statistics by $52^{0.5}$. Quintiles are formed based on quarterly returns that are computed using weekly returns from January 2, 1959, to December 31, 2010.

sufficient approach with respect to relative risk management. In fact, a strong case can be made that hedging, as opposed to diversification, is the proper approach for managing the risk of severe deviation with respect to a given benchmark. In a relative context, diversification consists of allocating to different assets or strategies in the hope that they will not underperform the benchmark in different market conditions. Hedging, however, consists of aligning the risk exposure of the portfolio to the risk exposure of the benchmark to ensure that extreme downside deviations as measured against the benchmark will not occur. In the next section, we elaborate on this comment and introduce a methodology specifically intended to achieve a better control of tracking error risk, which will prove particularly effective at managing extreme levels of downside risk.

RELATIVE RISK-CONTROLLED VERSIONS OF OPTIMIZED PORTFOLIOS

In the previous section, we showed that mixing various proxies for optimal portfolios versus investing solely in one proxy (GMV or MSR) leads to increasing

the probability of outperformance and decreasing the average tracking error with respect to a cap-weighted strategy. We find that this diversification approach has little impact, if any, on the risk of severe underperformance.

In order to better understand the difference between diversification and hedging, we propose an analogy that should usefully illustrate the point. Let us move away from the problem at hand for a moment and consider instead a more general problem of asset allocation under liability constraints from a pension fund perspective. The goal in the asset-liability management context is to outperform a liability-driven benchmark without excessive downside risk with respect to the benchmark. Using a well-diversified performance-seeking portfolio is certainly useful for generating outperformance with respect to the liability benchmark, but it does not allow for proper management of liability (downside) risk. In other

words, one does not diversify away liability risk, one hedges it away. The exact same line of reasoning applies here. In fact, the analogy is profound and not merely superficial, because one can formally consider the cap-weighted index as a kind of liability target for an investor facing benchmark constraints. The goal in our asset-only context is therefore to outperform the benchmark (here an asset benchmark) within a given relative downside risk budget. Diversification is useful for engineering and designing a performance engine that is likely to outperform the cap-weighted "liability" benchmark with the highest possible probability, but it does not allow for the management of extreme tracking error risk. Only hedging, that is, aligning the factor exposure of the performance engine with that of the cap-weighted benchmark, can allow for a proper management of that risk. In what follows, we provide a detailed analysis of how to implement an effective methodology for relative risk control. As such, we believe that the material in the next section may be of interest outside the particular context of the present analysis, and that it generally applies to portfolio construction problems involving a tracking error constraint.

Methodology for Relative Risk Control

As discussed in the introduction, any alternatively weighted portfolio may lead to significant relative risk with respect to the performance of the cap-weighted index. The objective of the present section is to analyze in detail how the relative risk of the optimization-based strategies discussed in this article can be managed within a suitably designed risk-control framework.

In fact, effective tracking control can be achieved as a two step-process. In a first step, one should combine the optimized portfolio with a suitably designed, timevarying quantity of the benchmark portfolio so as to ensure that relative risk is kept under the budgeted limits. Here the suitably designed quantity is taken to be a function of the tracking error budget, as well as time-varying parameters of the optimized portfolio. This approach, which is sometimes referred to as core-satellite approach in a benchmark portfolio management context, will be shown to be an effective way of respecting ex post a target level of tracking error risk budget by combining the optimized weights with the market-cap weights of the reference index.⁸ The problem, however, is that if the optimized portfolio is originally endowed with an ill-behaved tracking error process (i.e., a tracking error that may ex post deviate substantially from the average tracking error level), the desire to keep the overall tracking error under control at all times would result in investing too small an amount in the optimized portfolio and, conversely, too large an amount in the benchmark portfolio. As a result, the combined core-satellite portfolio may have a very limited outperformance potential.

To alleviate this concern, the natural procedure consists of making sure ex ante that the optimized portfolio risk exposures are sufficiently well aligned with the benchmark portfolio risk exposure so that a substantial allocation to the optimized portfolio can be allowed within the core–satellite scheme.

In this last step, explicit tracking error constraints may be used in the optimization procedure (see, for example, Jorion [2003]), and the risk factor exposures of the optimized portfolio may be aligned with those of the cap-weighted portfolio through explicit constraints on factor exposure. The purpose of such explicit constraints in portfolio optimization is to set an anchor at a predefined level of tracking error that the investor wants to respect.

We now turn to a more detailed description of the methodology for relative risk control before we analyze

the empirical results in terms of the relative risk of both relative risk-controlled and uncontrolled strategies. In a nutshell, the standard fund separation theorems state that the optimal solution to an investment problem in the presence of a benchmark is a mix of a performanceseeking portfolio (PSP) and the benchmark portfolio. The performance-seeking portfolio is a proxy for the tangency portfolio (i.e., the portfolio with the highest risk-reward ratio). In our case, the proxies we use for the tangency portfolio are those introduced in the previous section, namely, the global minimum-volatility portfolio, the maximum Sharpe ratio portfolio, and the diversified portfolio allocated 50% to the GMV and MSR proxies. The separation theorem states that the solution in terms of optimal weights is nothing but a mix of these unconstrained performance-seeking strategies, which is represented by a non-cap-weighted weight vector, denoted by w^{PSP} , and a cap-weighted weight vector, denoted by w^{CW} (Amenc, Goltz, and Le Sourd [2010] provided more details on the fund separation theorem in a benchmarked money management context),

$$W_t^* = \frac{\lambda_t}{\gamma \times \sigma_t} \times W_t^{PSP} + \left(1 - \frac{1}{\gamma}\right) \times W_t^{CW}$$

Here, γ is a risk aversion parameter that is used to characterize the tracking error risk budget, λ is the Sharpe ratio of the efficient portfolio proxy (the GMV, MSR, or mixture portfolio), and σ_{t} is its volatility at time t. The allocation to the "risky" component (the component with weights w_t^{PSP}) is an increasing function of the PSP's Sharpe ratio and a decreasing function of the investor's risk aversion. The allocation to the safe component (i.e., the cap-weighted benchmark) for which "safe" and "risky" are defined relative to the capweighted benchmark increases with risk aversion. If risk aversion goes to infinity, which is consistent with a zero tracking error risk budget, then 100% of the weight is allocated to the cap-weighted benchmark, as it should be. Note that in theory the weights allocated to the risky and safe building blocks do not add up to one, so in theory a positive or negative amount should also be allocated to cash, the risk-free asset that is used to make up the balance of the portfolio. In the implementation that follows, we normalize the weights allocated to the "risky" and "safe" building blocks so as to avoid any investment in cash because the focus is on an equityonly portfolio.

In practice, the risk aversion parameter γ is calibrated to reach a target tracking error level. Hence, for a given γ the weight allocated to the improved versus cap-weighted portfolio should rationally be a function of market conditions and, in particular, a (decreasing) function of volatility levels. Of course, uncertainty in parameter estimates may distort the optimality of this scheme, but the main insight, namely, that the weight allocated to the risky asset should be a function of market conditions, is a robust finding.

As explained before, one can intuitively expect to derive benefits by combining a core-satellite approach to tracking error control with creating a "well-behaved" satellite portfolio (i.e., a satellite portfolio with built-in relative risk controls that allow the achievement of a reliable tracking error out of sample). Therefore, in practice this approach works better when the performanceseeking portfolio is optimized under a reasonable tracking-error-at-risk objective. The tracking error objective allows us to anchor the level of tracking error of the performance-seeking portfolio on firm grounds and consequently allows us to maintain a relatively stable level of tracking error. Such a constraint on tracking error should be set relatively loosely to achieve portfolios with a high information ratio. In the strategies we analyze next, we set the tracking error constraint level at 5% annualized tracking error.

An important practical issue with such tracking error constraints is that they are likely to be exceeded out of sample. Therefore, it is important to impose additional constraints to achieve an out-of-sample reliability of the target level of tracking error. In particular, in the strategies we test, the tracking error at risk is controlled by putting additional constraints on the performance-seeking portfolio's exposure to implicit risk factors to make them equal to those of the core portfolio. The implicit risk factors are identical to the risk factors used in covariance estimation of the optimized strategies. These constraints on factor exposures are aimed at allowing for robust out-of-sample tracking error at risk.

A simple example may help illustrate the sense of such constraints on risk factor exposures. If we take the example of a minimum-variance strategy that uses a tracking error constraint of 5%, the optimal solution in a given quarter may result in a portfolio with a strong difference in exposure to risk factors compared to cap weighting, say a much heavier loading on the utilities sector. If, over the calibration period, stocks that were heavily exposed to this risk factor did not behave very differently from the cap-weighted index (which may happen if the return of the factor is similar to the return of the benchmark in the calibration period), then the tracking error constraint used in the optimization will not prevent a large difference in loading on this factor to occur with respect to the cap-weighted index. If, out of sample, a substantial shock to the factor occurs (e.g., if the utilities sector experiences strong underperformance), then the tracking error target will be exceeded. But if the loading of the optimized portfolio on this risk factor has been aligned to be identical to that of the cap-weighted portfolio, then the out-of-sample tracking error can be expected to be much more reliable.

Subsequently, once one has constructed an attractive satellite portfolio with a well-defined and reliable level of tracking error, the core-satellite approach can be used to further lower the tracking error to whatever target is consistent with investor preferences and/ or constraints. In the empirical analysis that follows, we combine both explicit constraints on tracking error and on risk factor loadings with a core-satellite approach. We report results for a target tracking error level of 3%, which is achieved starting with a 5% tracking error objective within the satellite, and a suitable mixture of the 5% tracking error satellite portfolio and the capweighted core portfolio.

Using this method for relative risk control, we construct relative risk–controlled versions of the minimum-volatility, maximum Sharpe ratio, and diversified portfolios.

Analyzing Performance and Relative Risk of Optimized Strategies with and without Relative Risk Control

In this section, we focus on the relative return, tracking error, and relative drawdown of alternative strategies with respect to the cap-weighted index (i.e., the standard S&P 500 Index). We assess the minimum-variance, maximum Sharpe ratio, and diversified portfolios, as well as relative risk–controlled versions of these portfolios. Note that liquidity and turnover constraints remain the same as described for the portfolios without relative risk control.

In Exhibit 5, we compare the different strategies based on their annualized expected excess returns over the cap-weighted index, their information ratios, and their modified information ratios (modified IRs), defined as the ratio of the annualized excess returns of the strategy over the S&P 500 to 95% of one-year trailing tracking error with respect to the S&P 500. The modified information ratio works just like the original information ratio, but rather than adjusting excess returns for average tracking error, it adjusts for the extreme tracking error. This adjustment reflects a concern of CIOs over extreme deviations from their peer group. To compute the oneyear trailing tracking error, we construct a portfolio long in the optimized portfolio and short in the S&P 500 and compute annualized standard deviation of weekly returns of the portfolio over the past year (52 weeks) and repeat the process using this one-year rolling window over the entire dataset.⁹ Similarly, a one-year trailing relative return series is also computed, and distributional statistics such as median, 5th percentile, and minimum values are reported.¹⁰ For a relative drawdown analysis, rather than constructing a portfolio long the alternative strategy and short the S&P 500, we construct a portfolio using the ratio of the indices.¹¹ The probability of outperformance is the probability of positive excess returns of the test portfolio compared to the benchmark's returns for rolling-period time horizons. We look at the trailing returns each week. The rolling-window lengths used are one year and three years.

Results reported in Exhibit 7 show a clear difference between the relative risk–controlled strategies and strategies without relative risk control. In particular, the extreme tracking error (i.e., the 95th percentile of the trailing one-year tracking error) is markedly lower for the relative risk–controlled portfolios. A remarkable

Ехнівіт 7

Comparing Portfolios v	with and	without Relative	Risk Co	ontrol
------------------------	----------	------------------	---------	--------

	Without Relative Risk Control			With Relative Risk Control		
	Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)	Minimum Volatility	Maximum Sharpe Ratio	Diversified (50% Minimum Volatility + 50% Maximum Sharpe Ratio)
Expected Return over CW	1.29%	1.42%	1.35%	0.79%	1.18%	0.99%
Information Ratio	0.47	0.51	0.50	0.44	0.52	0.54
Modified IR	0.23	0.28	0.27	0.25	0.29	0.30
Median Relative Return	1.06%	1.46%	1.35%	0.67%	1.14%	1.12%
5% Relative Return	-4.49%	-4.35%	-4.32%	-2.94%	-3.60%	-3.15%
Min Relative Return	-12.01%	-10.57%	-11.29%	-8.45%	-7.05%	-7.68%
Median Tracking Error	2.08%	2.15%	2.05%	1.53%	1.87%	1.51%
95% Tracking Error	5.60%	5.09%	5.07%	3.09%	4.03%	3.26%
Max Tracking Error	8.69%	8.36%	8.51%	4.48%	5.20%	4.77%
Max Relative Drawdown	22.96%	21.05%	21.83%	15.78%	14.65%	14.88%
Start Date	09/16/93	03/24/94	10/07/93	09/16/93	09/22/94	03/24/94
End Date	03/24/00	03/24/00	03/24/00	03/24/00	03/24/00	03/24/00
Weekly Value at Tracking-						
Error Risk (VaTER)	-0.54%	-0.55%	-0.54%	-0.38%	-0.45%	-0.38%
Max TUW (weeks)	340	313	337	340	287	313
Prob. of Outperformance						
(One-Yr. Rel. Return)	66.00%	65.03%	67.09%	64.65%	64.95%	67.47%
Prob. of Outperformance						
(Three-Yr. Rel. Return)	75.96%	75.84%	76.86%	76.31%	70.84%	76.39%

Notes: This exhibit shows a comparison of the minimum-volatility, maximum Sharpe ratio, and diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios with and without relative risk control in terms of annualized expected return over the S&P 500 Index, information ratio, and modified information ratio. The 95% value at tracking-error risk with respect to the S&P 500 Index is based on weekly relative returns over the entire period. Tracking error and relative return statistics are computed over a one-year rolling window, with weekly assessment. The probability of outperformance corresponds to the frequency of one-year (three-year) relative returns being positive when assessed at the beginning of each year. All values except for VaTER numbers are annualized and are based on weekly data from January 2, 1959, to December 31, 2010.

result from Exhibits 5 and 6 is the fact that the ex post measure of extreme tracking error risk, measured by the 95% tracking error, is barely higher than the 3% ex ante target level for the GMV and diversified portfolios, and only reaches 4.03% for the MSR proxy. This stands in sharp contrast to the results obtained without the explicit tracking error control procedure, where all reported ex post 95% tracking error levels are in excess of 5%.

Moreover, although the introduction of tracking error constraints obviously involves an opportunity cost in terms of outperformance, the outperformance of the relative risk–controlled portfolios over the cap-weighted index is higher per unit of extreme tracking error risk than it is for portfolios without relative risk control. This can be seen from the modified information ratio, in which we modify the standard information ratio by dividing excess returns by extreme tracking error rather than by simple average tracking error. It is also interesting to look at even more extreme levels of tracking error, notably the maximum level over a rolling one-year

window during the 52-year period that we analyze in this article. With the explicit tracking error constraints, the maximum tracking error of the diversified portfolio is brought down by 44% (from 8.51% to 4.77%), whereas its median relative return is reduced by only 17% (from 1.35% to 1.12%). Also, the maximum relative drawdown shows an improvement of about 30% for all three optimized strategies when using the risk control methodology. For a three-year investment horizon, the probability of outperformance in terms of relative return is around 76% for the diversified portfolio, which is higher than that of equal-weighted strategies (close to 65%). It should also be noted that although diversifying across strategies alone was not able to reduce value at tracking-error risk, the hedging approach (the relative risk-control procedure) effectively reduces VaTER.

Another remarkable result from Exhibit 7 is that among relative risk-controlled portfolios, the diversified strategy possesses a higher information ratio than both minimum volatility and maximum Sharpe ratio

EXHIBIT 8



Dynamic Analysis of Relative Risk Indicators for Diversified Portfolios

Notes: Panel A—The plot shows the relative drawdown of diversified portfolios (50% minimum volatility + 50% maximum Sharpe ratio) with and without relative risk control as compared to the S&P 500 Index.

Panel B—The plot shows one-year trailing tracking error for diversified (50% minimum volatility + 50% maximum Sharpe ratio) portfolios with and without relative risk control as compared to the S&P 500 Index.

The period of analysis ranges from January 2, 1959, to December 31, 2010.

proxies. With and without relative risk control, the diversified portfolio has slightly higher outperformance probability than both maximum Sharpe ratio and minimum volatility portfolios due to the fact that the two strategies' relative performance differs depending on market conditions. These results support the fact that diversifying across strategies results in better risk-adjusted performance and overall better chances of effective outperformance than selecting either one of them.

In order to provide a sense for the time periods over which the heaviest relative draw downs occurred, we report the results of two additional statistics. For the sake of brevity we only show results for diversified portfolios; the plots for minimum-volatility and maximum Sharpe ratio portfolio proxies show a similar decease in relative risk when using the relative risk–control framework. Panel A of Exhibit 8 compares the time evolution of the relative drawdown of relative risk–controlled and nonrelative-risk-controlled, loosely constrained diversified portfolios. Panel B shows one-year rolling tracking error for the same portfolios.

CONCLUSION

In this article, we analyze the properties of two remarkable proxies for optimally diversified portfolios in the equity universe, notably the maximum Sharpe ratio portfolio and the minimum-volatility portfolio. Although both portfolios have the merit of approximating efficient portfolios, they rely on different assumptions to be truly optimal. As a result, they exhibit differences in performance characteristics. We assess how such differences can be exploited to diversify across various forms of efficient portfolios. Furthermore, because any deviation from a default market-cap-weighting scheme introduces relative performance risk, we explore how such relative risk can be mitigated in the construction of investable proxies for efficient portfolios.

Our empirical results show that minimum-volatility portfolios provide defensive exposure to equity that does well in adverse market conditions, whereas maximum Sharpe ratio portfolios provide a higher access to the upside of equity markets. On the one hand, combining both approaches naturally leads to a smoother conditional performance and would be a reasonable approach for all investors except those explicitly endowed with a trustworthy forward-looking view about equity market performance. On the other hand, implementing such strategies without relative risk control leads to high levels of extreme tracking-error risk. Thus, we argue in favor of the introduction of a suitably designed relative riskcontrol process in order to achieve more access to outperformance per unit of extreme relative risk taken. Overall, the results we present in this article suggest that it is possible to diversify model risk as well as control relative risk compared to cap-weighted indices, while respecting practical liquidity and implementation constraints. One of our key messages is that diversification of proxies for efficient portfolios can enhance the probability of outperforming the cap-weighted benchmark, whereas hedging can allow for effective relative downside risk protection. Of course, these choices are not mutually exclusive because investors would rationally aim at maximizing their chances to outperform the benchmark, while setting strict risk limits/budgets in terms of downside risk.

ENDNOTES

¹In this respect, one needs to make a clear distinction between ad hoc weighting schemes, which weight stocks by more or less arbitrary characteristics such as dividends or revenues, and weighting schemes that have an explicit objective in terms of efficient spending of a risk budget.

²Because return series from CRSP data are incomplete at times for some stocks, we remove those stocks from the optimization process for which more than 10 weeks of data are unavailable. These stocks are, however, added post optimization with a minimal weight, which is $1/\lambda N$ for minimum-volatility and maximum Sharpe ratio portfolios. This is to ensure that no stocks are excluded, meaning that the constitution of the optimized portfolios matches the constitution of the cap-weighted index.

³DeMiguel et al. [2009] showed that using norm constraints on portfolio weights (i.e., setting an upper bound on the sum of squares of portfolio weights) allows for a better use of the risk budget compared to imposing a limit to the weight on each stock. In this article, we use basic weight for simplicity and comparability.

⁴It should be noted, however, that the chosen levels of constraints allow for significant deviations from equal weighting. For instance, using the value $\lambda = 6$, a stock that is weighted at the upper bound will have 36 times the weight of a stock weighted at the lower bound, thus allowing for considerable differences in the weights attributed to different stocks.

⁵But to avoid any risk of long periods without rebalancing, a threshold of two years is set (i.e., if the portfolio will be rebalanced irrespective of its threshold being reached if it has continued to remain unrebalanced for the last two years). ⁶All significance statistics are performed at a 5% confidence level. A paired *t*-test is performed for returns and a Fisher Test is performed for volatility. We use a nonparametric bootstrap method (Ledoit and Wolf [2008]) for the hypothesis testing of the Sharpe ratio.

⁷Another notable and related difference is that for a given level of weight constraints the MSR portfolio has a higher market beta than the minimum-volatility portfolio. The lowbeta nature of minimum volatility is a commonly recognized feature of minimum-volatility portfolios; see, for example, Chan, Karceski, and Lakonishok [1999].

⁸To pursue the aforementioned analogy, this approach is formally similar to the liability-driven investing (LDI) paradigm in the asset–liability management context.

⁹We start this process at the 53rd week and repeat the process for each week until the data are exhausted. Thus, annualized tracking error is computed over a one-year rolling window at each week (except the first 52 weeks for which the one-year window ceases to exist). Hence, we obtain a trailing tracking-error series for each strategy, and we report median, 95th percentile, and maximum values of this series.

¹⁰For a one-year relative return series, we look at past oneyear weekly returns starting at the 53rd week and compute the annualized excess return of the strategy over the S&P 500 for this one year. The process is repeated for each of the remaining weeks to obtain one-year trailing relative return series.

¹¹We construct two indices to start with: the first uses the weekly returns of an alternative strategy (called the Long Index) and the second uses the same for the S&P 500 (called the Short Index). We construct a third portfolio (modified Long–Short portfolio) whose weekly returns are given by the fractional increase in the ratio of Long Index value to Short Index value over the given week. Using this return series, we obtain a modified Long–Short Index and conduct drawdown analysis on it. Subsequently, we report statistics such as maximum relative drawdown, maximum time under water (the period of time the maximum relative drawdown lasted), and the start and end dates of time under water.

REFERENCES

Amenc, N., F. Goltz, and V. Le Sourd. "The Performance of Characteristics-Based Indices." *European Financial Management*, Vol. 15, No. 2 (2009), pp. 241-278.

Amenc, N., F. Goltz, L. Martellini, and V. Milhau. "New Frontiers in Investment Benchmarking and Liability-Driven Investing." EDHEC-Risk publication, 2010.

Amenc, N., F. Goltz, L. Martellini, and P. Retkowsky. "Efficient Indexation: An Alternative to Cap-Weighted Indices." *Journal of Investment Management*, Vol. 9, No. 4 (2011), pp. 1-23. Amenc, N., and L. Martellini. "In Diversification We Trust?" *The Journal of Portfolio Management*, Vol. 7, No. 2 (2011), pp. 1-2.

Arnott, R., J. Hsu, and J. Moore. "Fundamental Indexation." *Financial Analysts Journal*, Vol. 60, No. 2 (2005), pp. 83-99.

Blitz, D., and L. Swinkels. "Fundamental Indexation: An Active Value Strategy in Disguise." *Journal of Asset Management*, Vol. 9, No. 4 (2008), pp. 264–269.

Britten-Jones, M. "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights." *Journal of Finance*, Vol. 54, No. 2 (1999), pp. 655-671.

Chan, L.K.C., J. Karceski, and J. Lakonishok. "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model." *Review of Financial Studies*, Vol. 12, No. 5 (1999), pp. 937-974.

Choueifaty, Y., and Y. Coignard. "Toward Maximum Diversification." *The Journal of Portfolio Management*, Vol. 35, No. 1 (2008), pp. 40–51.

DeMiguel, V., L. Garlappi, F. Nogales, and R. Uppal. "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms." *Management Science*, Vol. 55, No. 5 (2009), pp. 798–812.

Frost, P., and J. Savarino. "For Better Performance: Constrain Portfolio Weights." *The Journal of Portfolio Management*, Vol. 15, No. 1 (1988), pp. 29–34.

Fujiwara, Y., W. Souma, H. Murasato, and H. Yoon. "Application of PCA and Random Matrix Theory to Passive Fund Management." In *Practical Fruits of Econophysics, Proceedings of the Third Nikkei Econophysics Symposium*, edited by Hideki Takayasu. Tokyo: Springer, 2006, pp. 226-230.

Goyal, A., and S. Wahal. "The Selection and Termination of Investment Management Firms by Plan Sponsors." *Journal of Finance*, Vol. 63, No. 4 (2008), pp. 1805–1847.

Grinold, R. "Are Benchmark Portfolios Efficient?" *The Journal of Portfolio Management*, Vol. 19, No. 1 (1992), pp. 34-40.

Haugen, R., and N. Baker. "The Efficient Market Inefficiency of Capitalization-Weighted Stock Portfolios." *The Journal of Portfolio Management*, Vol. 17, No. 3 (1991), pp. 35-40. Jagannathan, R., and T. Ma. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *Journal of Finance*, Vol. 54, No. 4 (2003), 1651-1683.

Jorion, P. "Portfolio Optimization with Tracking-Error Constraints." *Financial Analysts Journal*, Vol. 59, No. 5 (2003), pp. 70-82.

Jun, D., and B. Malkiel. "New Paradigms in Stock Market Indexing." *European Financial Management*, Vol. 14, No. 1 (2007), pp. 118-126.

Kan, R., and G. Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3 (2007), pp. 621-656.

Kaplan, P. "Why Fundamental Indexation Might—or Might Not—Work." *Financial Analysts Journal*, Vol. 64, No. 1 (2008), pp. 32-39.

Ledoit, O., and M. Wolf. "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection." *Journal of Empirical Finance*, Vol. 10, No. 5 (2003), pp. 603–621.

——. "Honey, I Shrunk the Sample Covariance Matrix." *The Journal of Portfolio Management*, Vol. 30, No. 4 (2004), pp. 110–119.

——. "Robust Performance Hypothesis Testing with the Sharpe Ratio." *Journal of Empirical Finance*, Vol. 15, No. 5 (2008), pp. 850-859.

Leland, H. "Optimal Portfolio Management with Transactions Costs and Capital Gains Taxes." Working paper, University of California Berkeley, 1999.

Maillard, S., T. Roncalli, and J. Teiletche. "The Properties of Equally Weighted Risk Contributions Portfolios." *The Journal of Portfolio Management*, Vol. 36, No. 4 (2010), pp. 60–70.

Martellini, L. "Alternatives to Cap-Weighted Indices: Towards the Design of Improved Equity Benchmarks." Working paper, EDHEC Risk Institute, *Pension System and Reform*, forthcoming.

Melas, D., and X. Kang. "Applications of Systematic Indexes in the Investment Process." *Journal of Indexes*, Vol. 13, No. 5 (2010), pp. 10–18.

Schwartz, T. "How to Beat the S&P 500 with Portfolio Optimization." Working paper, DePaul University, 2000.

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@iijournals.com or 212-224-3675.