How should individual investors diversify?

An empirical evaluation of alternative asset allocation policies

Heiko Jacobs, Sebastian Müller, Martin Weber*

July 2013

Abstract

This paper evaluates numerous diversification strategies as a possible remedy against widespread costly investment mistakes of individual investors. Our results reveal that a very broad range of simple heuristic allocation schemes offers similar diversification gains, as well-established or recently developed portfolio optimization approaches. This holds true for both international diversification in the stock market and diversification over different asset classes. We thus suggest easy-to-implement allocation guidelines for individual investors.

Keywords: Portfolio theory; Household finance; Asset allocation; International diversification; Heuristics

JEL classification: G11

---

* Heiko Jacobs, University of Mannheim, L 5, 2, 68131 Mannheim. E-Mail: jacobs@bank.bwl.uni-mannheim.de. Sebastian Müller, University of Mannheim, L 5, 2, 68131 Mannheim. E-Mail: mueller@bank.bwl.uni-mannheim.de. Martin Weber, University of Mannheim, L 5, 2, 68131 Mannheim. E-Mail: weber@bank.bwl.uni-mannheim.de. We are grateful to Gerd Kommer, Olaf Scherf, Raman Uppal, Volker Vonhoff, and seminar participants at the Gutmann Center for Portfolio Management (Vienna University of Economics and Business), the portfolio ETF forum, the annual meeting of the German Finance Association (DGF), the Campus for Finance Research Conference (CFF), the European Business School, and the University of Mannheim for valuable comments. Furthermore, we would like to thank Andreas Dzemski and Erdal Talay for excellent research assistance. Financial Support from the Deutsche Forschungsgemeinschaft (DFG) and SFB 504 at the University of Mannheim is gratefully acknowledged.

1. Introduction

Despite the recognized benefit of diversification as “the only free lunch in investment,” individual investors seem to sometimes violate even its most basic principles. In fact, “these discrepancies, or investment mistakes, are central to the field of household finance” (Campbell, 2006, p. 1554). In this paper, we derive easily implementable portfolio construction guidelines for individual investors. Our approach allows us to evaluate numerous competing strategies, both for international diversification in stock markets and (additional) diversification across asset classes. Specifically, we ask the following questions: From the perspective of individual investors in real-life situations, what is the most promising way to diversify? Do simple rules of thumb already provide a powerful remedy against widespread investment biases? Which heuristics are particularly efficient at realizing diversification potential? To what extent do these strategies underperform when benchmarked against sophisticated optimization models?

Empirical studies provide extensive evidence of individual investors making portfolio choices, which are difficult to reconcile with standard financial theory. As such, households often fail to participate in the stock market at all (e.g., Campbell, 2006; and Kimball and Shumway, 2010). Among those households that do invest in equities, many studies document further costly mistakes. First, individuals tend to prefer domestic over foreign investments, thereby foregoing the benefits of international diversification (French and Poterba, 1991; Grinblatt and Keloharju, 2001; and Kilka and Weber, 2000). Second, many households own relatively few individual stocks, which may cause a significant exposure to idiosyncratic risk (e.g., Goetzmann and Kumar, 2008; and Polkovnichenko, 2005). Third, data from online brokerage accounts show that many individuals are overconfident and trade too much (Odean, 1999; and Barber and Odean, 2000).

Puzzling investment behavior is also observed when considering diversification over asset classes. Analyzing a large sample of retirement accounts, Agnew et al. (2003) show that most asset allocations are extreme (either 100% or zero percent in equities) and that there is inertia in asset allocations. Tang et al. (2010) conclude that most participants make inefficient portfolio investment choices in retirement plans. The failure of diversifying adequately over asset classes must be considered as particularly problematic as asset allocation has been shown to be the main determinant of portfolio performance (e.g., Brinson et al., 1986; or Ibbotson and Kaplan, 2000).

To put it in a nutshell, risk-adjusted portfolios of most individual investors underperform even standard domestic stock market indices at a significant margin, and thus leave substantial room for improvement. But how should individual investors diversify? While academic research almost exclusively relies on the performance of various extensions of the Markowitz (1952) framework, we also concentrate on the relative investment value of heuristic diversification strategies.

This is particularly relevant for individual investors who typically will not have the knowledge and resources to implement complex optimization models. In addition, Markowitz-based approaches, while being optimal in theory, suffer from estimation error in expected returns, variances, and covariances when implemented in practice. There is a large amount of literature explicitly dealing with methods to improve the out-of-sample performance of these strategies, with partly disillusioning results. Recent studies focusing primarily on U.S. stock portfolios show that the estimation error is so severe that various optimization models are oftentimes unable to beat a naïve $1/N$ diversification strategy (e.g., DeMiguel et al., 2009b; Duchin and Levy, 2009; and Tu and Zhou, 2009).[1]

Hence, it seems insufficient to limit the analysis to extensions of the Markowitz (1952) model. In the empirical analysis, we thus analyze the performance of 11 well-established or recently proposed mathematical optimization methods as opposed to a broad range of plausible heuristics. In doing so, we combine two prominent ways of diversification that are usually analyzed separately: International diversification in the stock market and diversification over different asset classes. To achieve comparability with the previous literature, the following two-step procedure is employed.

First, we concentrate on global diversification in the stock market from the perspective of a eurozone investor. Such an analysis might be considered a complement to the influential study of DeMiguel et al. (2009b). We rely on the bootstrap technique developed Ledoit and Wolf (2008) to assess the significance of differences in Sharpe ratios. In contrast to the standard test statistic of Jobson and Korkie (1981), the validity of the Ledoit and Wolf methodology is not sensitive to the underlying distribution and thus particularly suitable for the analysis of financial time series data. The approach is designed to provide reliable findings even when returns exhibit fat tails or

[1] The out-of-sample performance of an equally-weighted portfolio as compared to the performance of the standard Markowitz approach is in fact a longstanding and controversial debate in portfolio optimization. Early discussions include, for instance, Frankfurter et al. (1971), Brown (1979), or Jobson and Korkie (1981). For a recent study arguing that optimized portfolios do outperform equally-weighted portfolios, see Kritzman et al. (2010).

show typical time series characteristics, such as volatility clustering or autocorrelation. With regard to performance evaluation, we gain additional insights by building on factor models borrowed from the mutual fund literature. We construct a global Carhart (1997) four-factor model using Datastream's stock universe. This allows us to draw inferences that are not possible from an analysis of traditional performance measures alone.

Second, we extend our analysis to the multi-asset class case by additionally incorporating bonds and commodities. In the baseline scenario, we derive simple fixed-weight policies from the academic as well as practitioner literature and compare them to the optimization models. Again, we employ a multi-factor regression framework to identify the underlying drivers of performance. We construct value and momentum factors for bonds and commodities building on the recent work of Asness et al. (2013). Our approach adds to the literature on performance attribution of multi-asset class portfolios. Finally, we analyze the performance of more than 5,000 alternative fixed-weight strategies covering every possible proportion of the asset classes in 1% steps. This enables us to gain deeper insights into the structural composition of promising portfolios.

For the case of international equity diversification, we find that none of the optimization models is able to significantly outperform simple heuristics in an out-of-sample setting. Among the heuristic approaches, the standard approach of a market-weighted stock portfolio appears to be less successful than an equally-weighted portfolio or a fundamentally-weighted portfolio. However, differences between these three heuristic approaches become smaller once the factor loadings to size and value effects are controlled for.

For the case of diversification over different asset classes, we again find no significant differences between optimization models and heuristic approaches. In fact, almost any well-balanced fixed-weight proportion of stocks, bonds, and commodities is able to realize diversification gains that are similar in magnitude to those of the optimization models.

A number of sensitivity checks assure the robustness of our results in both settings. We thus suggest a simple and cost-efficient asset allocation approach for individual investors.

The remainder of this paper is organized as follows. Section 2 describes the data. Section 3 discusses promising optimization models and alternative heuristic strategies. Section 4 contains the empirical analysis. A summary of the results is given in Section 5.

2. Data and descriptive statistics

2.1 Asset classes and data

Given our focus, we pay particular attention to the implementability of our results. We therefore base our study on indices, into which individual investors can easily invest at low costs via exchange-traded funds. We concentrate on eurozone individual investors within a yearly rebalanced buy-and-hold approach. We incorporate stocks, bonds, as well as commodities in the analysis. These asset classes are represented by indices whose selection is based on the criteria transparency, representativeness, investment access, liquidity, and data availability.[2]

[2] We require the index composition and index rules to be disclosed by the index provider (transparency). The index should already cover most of the market within an asset category to reduce complexity (representativeness). In doing so, portfolios can be constructed with only few highly diversified indices. Moreover, low-cost exchange-traded funds tracking these indices should

Based on these requirements, we rely on the Morgan Stanley Capital International (MSCI) index family, which has been widely used in previous studies (e.g., Driessen and Laeven, 2007; De Roon et al., 2001), to cover the global stock universe. In the baseline analysis, stocks are represented by the four regional indices: MSCI Europe, MSCI North America, MSCI Pacific, and MSCI Emerging Markets. These indices currently cover 44 countries and track the performance of several thousands of stocks. The indices are designed to cover 85% of the free float-adjusted market capitalization of the respective equity universe accessible to individual investors.

Bonds are incorporated because of their low correlation with stocks. In the baseline analysis, they are represented by the iBoxx Euro Overall Index, which consists of eurozone bonds of different maturities and credit ratings.[3] The index currently tracks the performance of more than 2,500 bonds. In robustness checks, we also make use of the iBoxx Euro Sovereign Index, which only consists of government bonds, the JPM Global Bond Index, and the JPM Europe Government Bond Index.

Partly due to a lack of investment access, commodities have long been neglected by individual investors. However, many studies provide evidence of the high diversification potential of broad-

---

exist to enable individual investors to actually implement our suggestions (investment access and liquidity). Finally, we require a long return data history to conduct powerful statistical tests (data availability).

[3] As we aim to derive suggestions for individual investors, we do not consider currency hedging. For internationally-diversified bond portfolios, Black and Litterman (1992) and Eun and Resnick (1994) find that currency risk needs to be controlled for. We thus restrict our analysis to euro-denominated bonds. As the iBoxx index universe is only available from 1999 on, we replace the return of the iBoxx Euro Overall Index with the return of the REXP for the time period before 1999, which is justified by the high correlation of both indices in the period after 1999.

based commodity futures indices.[4] Furthermore, diversification benefits tend to be especially pronounced in times of unexpected inflation and declining stock markets. In the baseline analysis, commodities are represented by the S&P GSCI Commodity Total Return Index. This world production-weighted index currently includes 24 commodity futures contracts that track the performance of energy products, industrial and precious metals, agricultural products, and livestock. In sensitivity checks, commodities are also represented by the Reuters/Jefferies Total Return Index and the DB Commodity Euro Index.

We do not incorporate real estate in our analysis as individual investors are often already heavily exposed to real estate risk (e.g., Calvet et al., 2007; Campbell, 2006). Thus, the additional inclusion of real estate in the overall portfolio might lead to a lack of diversification. Moreover, we do not consider alternative asset classes such as hedge funds and private equity for two reasons. First, their diversification potential in the multi-asset case is often found to be limited (e.g., Amin and Kat, 2003; Ennis and Sebastian, 2005; Paton, 2009; and Phalippou and Gottschalg, 2009). Second, we could not identify indices meeting our selection criteria on a satisfactory level.

Our evaluation period starts in February 1973 and ends in December 2012. This close to 40-year period extends previous studies on international diversification in the stock market (e.g., De Santis and Gerard, 1997; De Roon et al., 2001; Driessen and Laeven, 2007). For all indices, we use euro-denominated total return indices extracted from Thomson Reuters Datastream. Hence,

[4] Historically, these indices delivered equity-like returns and volatilities. At the same time, they provided low and, partly even negative, correlations with stocks and bonds [e.g., Erb and Harvey (2006) and Section 2.2]. Other commodity exposure such as physical trading, individual commodity futures or stocks of companies owning and producing commodities does not offer the specific risk, return, and correlation features of broad-based commodity futures indices (e.g., Erb and Harvey, 2006; and Gorton and Rouwenhorst, 2006). Thus, they are less suitable for our analysis.

our findings refer to an investment without currency hedging, which is a realistic assumption for individual investors.[5]

To implement the heuristic portfolio strategies in the stock universe, we require the gross domestic product (GDP, in current U.S. dollars) and the stock market capitalization of the MSCI index regions. We obtain these data from the World Bank, the International Monetary Fund (IMF), and Thomson Reuters Datastream, respectively. We use the three-month FIBOR (before 1999) and the three-month Euribor (thereafter) as a proxy for the risk-free asset. Historical stock market capitalization data are available from 1973 onwards, which marks the lower bound of our evaluation period.

2.2 Descriptive statistics

Table 1 gives an overview of monthly return parameters of the asset classes, which are represented by the iBoxx Euro Overall Index, the S&P GSCI Commodity Total Return Index, and a number of stock indices. The latter comprise the four regional MSCI indices and, for comparative purposes, a global capitalization-weighted stock index constructed from the four regional indices. The MSCI Emerging Markets are only incorporated from 1988 on, as this is the starting point of the index calculation.

Please insert table 1 here

[5] To convert index levels into euros, we use the synthetic euro/USD exchange rates from Thomson Reuters Datastream. In robustness checks, we redo the analysis using the historical DEM/USD exchange rate as published by Deutsche Bundesbank. The qualitative nature of our results does not change.

Table 1 shows only small differences in the Sharpe ratio of regional stock indices (on average 0.098) compared to the global stock index (0.105). Over the last 25 years, this difference becomes even negative. This result motivates both the analysis of alternative allocation mechanisms for the stock market and the incorporation of additional asset classes.

To assess the additional diversification potential of bonds and commodities, Figures 1 and 2 illustrate the time series behavior of correlations within the stock markets and across asset classes, respectively. Correlation coefficients are computed using a rolling-window approach based on the previous 60 months.

Please insert figures 1 and 2 here

Figure 1 reveals an almost steady increase in the co-movement of international stock markets since the 1980s. However, as Figure 2 illustrates, there is no (in the case of bonds) or, at most, weak (in the case of commodities) evidence of an increase in correlations across asset classes. Nevertheless, correlations vary considerably over time. This finding points to potential estimation errors in optimization methods.

3. Asset allocation models

3.1 Optimization models

The eleven optimization models considered for portfolio selection in the case of both global stock market diversification and diversification over asset classes are briefly summarized in Panel A of Table 2, along with their abbreviations used as reference in the other tables.

Please insert table 2 here

We start with a variety of models that have been suggested in the literature to deal with the well-known problem of estimation error, which is ignored in the traditional mean-variance optimization.[6] These models either impose additional constraints in the optimization process, shrink the estimated input parameters in order to mitigate the impact of estimation error, or both. Short sale constraints prevent the optimization model from taking extreme long and short positions to exploit even small differences in the return structure of assets. Shrinkage models correct the estimated parameters towards a common value. In doing so, they aim at reducing the error-maximizing property of the mean-variance model when historical data are used for parameter estimation (e.g., Jorion, 1986). As shown by Jagannathan and Ma (2003) for U.S. stock portfolios, both approaches work similarly by increasing the number of assets with non-negative weights, which enforces a certain extent of diversification.

The first model we implement is the mean-variance framework with non-negativity condition. The objective of this model is to maximize the Sharpe ratio of the portfolio, which allows us to refrain from considering individual risk preferences in the optimization process. In addition, we employ three extensions of this model that either shrink the sample means, the sample variance-

[6] Consistent with previous empirical evidence, the traditional mean-variance optimization without constraints leads to extreme long and short positions with exorbitant high turnover. Therefore, we refrain from reporting these results.

covariance matrix, or both. The shrinkage estimation of expected returns is based on the work of James and Stein (1961). In our study, we use the estimator proposed by Michaud (1998). We shrink the elements of the variance-covariance matrix employing the constant correlation model developed in Ledoit and Wolf (2004).[7]

In addition to models that try to maximize the Sharpe ratio, we employ several models that aim at constructing minimum variance portfolios, thereby ignoring information about sample mean returns. The superior performance of minimum variance optimization has been demonstrated in various studies, which mostly concentrate on stock markets (e.g., Haugen and Baker, 1991; Chopra et al., 1993; and Jagannathan and Ma, 2003).

We start with the traditional minimum variance approach without and with short sale constraints. We then implement the minimum variance approach with shrinkage estimation of the variance-covariance matrix using the constant correlation model and short sale restriction. Finally, we consider four extensions of the general minimum variance framework that have recently been developed by DeMiguel et al. (2009a). In their empirical analysis, the authors show that this novel class of models often outperforms existing (U.S. stock) portfolio strategies at a significant margin. They impose the additional constraint that the sum of the absolute values of the portfolio weights (known as 1-norm) or the sum of the squared values of the portfolio weights (known as 2-norm) must be smaller than a given parameter threshold δ. Effectively, this constraint allows portfolios to have some short positions, but restricts the total amount of short selling. In order to calibrate the value of the threshold parameter δ, DeMiguel et al. (2009a) use two different

[7] The authors provide the code on their website (http://www.ledoit.net/shrinkCorr.m). We assume a constant correlation equal to the historical correlation average for the stock market indices and a correlation of 0 between different asset classes. Our results are unchanged if we simply use the historical correlation average over all indices irrespective of the asset class underlying the index.

methods: First, they choose the parameter δ, which minimizes the portfolio variance if the sample is cross-validated. Second, they set δ such that the portfolio return is maximized in the last period in order to exploit positive autocorrelation in portfolio returns.[8] In sum, this leads to four norm-constrained (nc) minimum variance approaches.

3.2 Heuristic models

The heuristic models relied on in the baseline analysis are presented in Panel B of Table 2.

3.2.1 International stock market diversification

We consider three different weighting schemes for a global stock portfolio: equal-weighting, market value-weighting, and GDP-weighting.

An equally-weighted portfolio, which is also often referred to as the 1/*N* heuristic, might be considered to be a natural benchmark for more sophisticated methods of portfolio optimization. First, it is very easy to implement. And, second, individual investors have been shown to often rely on this naïve allocation rule (e.g., Benartzi and Thaler, 2007).

Another strategy is to base portfolio weights on the relative market capitalization of the constituents. This concept is at the heart of most major stock market indices and thus easy to follow for individual investors. Liquidity and investment capacity arguments are important benefits of these indices, although of minor relevance for our objective. However, an undisputed

[8] For further information about the derivation of the portfolio models and the motivation of DeMiguel et al. (2009a), we refer the reader to their study. We do not evaluate other portfolio models considered in their paper, because the design of these models is similar to the ones tested in our study. Moreover, all models achieve similar results in terms of out-of-sample portfolio variance, Sharpe ratio, and turnover.

advantage of this approach is its very low turnover as portfolio weights automatically rebalance when security prices fluctuate.

Nevertheless, concerns regarding this weighting scheme have recently been raised. Figure 3 gives the intuition behind these arguments. It shows the time series of portfolio weights of a market value-weighted stock index constructed from the MSCI indices for North America, Europe, the Pacific region, and the Emerging Markets. Figure 3 illustrates that the resulting global stock index tends to be dominated by single regions. Between 1998 and 2007, for example, the weight of North America was on average about 45%. As the MSCI indices themselves are cap-weighted, U.S. large caps substantially drove the performance of the global stock universe during that period. In contrast, the portfolio weights in the previous decade were heavily influenced by the bull and subsequent bear market of the Japanese stock market. The fraction of the Japan-dominated Pacific region was more than 52% in 1989 and dropped heavily to about 15% in 1998. These examples illustrate the pro-cyclical nature of value-weighted indices.

Please insert Figure 3 here

Motivated by many studies arguing that price fluctuations sometimes do not fully reflect changes in company fundamentals (e.g., Shiller, 1981), a growing literature questions the efficiency of value-weighted indices (e.g., Treynor, 2005; Siegel, 2006). Recently, alternative index concepts aimed at better approximating true firm values have been proposed. These indices are often weighted by fundamental measures such as earnings, dividends or book values (Arnott et al., 2005), building on the intuition that this scheme might be less volatile and less driven by sentiment. Consistent with this rationale, back-testing shows that fundamentally-weighted country-specific indices have outperformed standard value-weighted indices (e.g., Arnott et al., 2005).

These findings justify the inclusion of a fundamentally-oriented global stock market index in our analysis. To transfer the idea from the firm to the regional level, we weight the four MSCI indices based on the relative GDP of their covered countries. As the MSCI indices themselves are market value-weighted, this policy might be considered as a compromise between a cap-weighted and a fundamentally-weighted approach. As can be seen from Figure 4, this procedure indeed results in a less volatile, more balanced allocation.

Please insert Figure 4 here

3.2.2 (Additional) diversification over asset classes

The easiest asset allocation policy for individual investors would arguably be to assign time-invariant weights to stocks, bonds, and commodities. The high number of potential fixed-weight strategies requires the definition of a benchmark against which optimization models can be tested. As selecting any specific strategy is a somewhat arbitrary choice, we employ a two-step procedure. First, we screen the literature to derive a promising baseline policy, which we use in the empirical tests in Section 4.2.2. Second, we analyze the performance of more than 5,000 alternative portfolios with any possible fixed-weights (in 1% steps) in Section 4.3 to assess the robustness of time-invariant allocation policies.

Regarding the ratio of stocks and bonds, we try to determine a best practice solution as a benchmark. Specifically, we study the security market advice of major investment banks and brokerage firms as reported in, for example, Arshanapalli et al. (2001) and Annaert et al. (2005), as well as institutional holdings as reported in, for example, Brinson et al. (1986), Blake et al. (1999), and Ibbotson and Kaplan (2000). Most of these studies analyze the allocation over cash, bonds, and stocks and do not consider other asset classes. We focus on the time series average of the cross-sectional mean of these allocations, as Arshanapalli et al. (2001) and Annaert et al. (2005) document the efficiency of such a strategy. Based on the overall picture, we derive a consensus recommendation of roughly 60% stocks and 40% bonds. Next, we analyze the literature that explicitly deals with commodities in an asset allocation context. Based on, for example, Anson (1999) and Erb and Harvey (2006), we estimate a consensus weight of roughly 15% for commodities.

Constructing an ex ante baseline portfolio from these results leaves us with some degrees of freedom. Specifically, commodities could be incorporated at the expense of fewer stocks, fewer bonds, or both fewer stocks and bonds. Given this arbitrary choice, we use stocks, bonds, and commodities in a fixed proportion of 60%, 25%, and 15%. As in the case of international stock market diversification, the regional MSCI indices in the stock component of these portfolios are either equal-weighted, value-weighted, or fundamentally-weighted. Thus, we have three baseline asset allocation heuristics. It is again noteworthy that our objective is to merely derive plausible ex ante strategies as a starting point for the empirical analysis, and not to derive an ex post optimal portfolio.[9]

4. Empirical analysis

4.1 Performance evaluation methodology

The performance of the portfolio strategies is assessed over the sample period from February 1973 to December 2012. Our implementation of the optimization models relies on a rolling-window approach. Specifically, at the beginning of each February, we use return data of the previous 60 months to calculate the input parameters needed to determine the portfolio weights of each index. Using these weights, we then calculate portfolio returns over the next 12 months without rebalancing. The following February, new portfolio weights are determined by using the updates of the parameter estimates.

We use the resulting time series of out-of-sample returns to compute the Sharpe ratio of each strategy. The ratio is defined as the average monthly excess return over the risk-free rate, divided by the standard deviation of monthly excess returns in the whole sample period. To test for differences in Sharpe ratios, we follow the bootstrap technique recently developed in Ledoit and Wolf (2008).

For the market value-weighting scheme, we calculate the portfolio weights at the rebalancing date using market values as of January, 1. The one-month lag has the aim of ensuring real-time

[9] In fact, we find that the 60/25/15 portfolio performs slightly worse than the other two asset allocation alternatives (i.e., less stocks or less stocks and less bonds). Hence, from an ex post perspective, the benchmark against which we test optimizing asset allocation models might be regarded as conservative.

data availability. The GDP-weighting is based on data from the previous year. We also compute the portfolio turnover of each strategy, which results from the annual weight adjustments. This allows us to calculate the out-of-sample Sharpe ratio after transaction costs. In order to do so, we assume a proportional bid-ask spread equal to 40 basis points per transaction.[10]

For international equity diversification, we also rely on factor models commonly employed in the mutual fund literature. Specifically, in addition to the Jensen (1968) one-factor alpha, we estimate the alpha from a global Carhart (1997) four-factor model to infer to what extent competing strategies load on the value, size, and momentum premium, respectively. The Carhart alpha is estimated from the following model:

$$r_t - r_{f,t} = \alpha^{4F} + \beta_{MKT} \cdot MKT_t + \beta_{SMB} \cdot SMB_t + \beta_{HML} \cdot HML_t + \beta_{WML} \cdot WML_t + \epsilon_t, \quad (1)$$

where r_t and $r_{f,t}$ are the returns of the strategy and the risk-free asset in period t and MKT is the excess return of the market-weighted global equity portfolio. The expressions $SMB, HML,$ and WML denote the returns of the following zero-investment strategies: SMB is the return difference between small and large capitalization stocks, HML is the return difference between stocks with high and low book-to-market ratios, and WML is the return difference between stocks with high and low past stock returns. The Jensen (1968) one-factor alpha is calculated in a similar fashion but uses only the market factor. We construct the global factors using Datastream's world-wide stock universe, following the methodology of Griffin (2002). That is,

[10] The spread is assumed to be the same for each index. It is based on the average bid-ask spread in 2007 for selected exchange-traded funds tracking the indices used in our analysis. Other trading costs and a potential price impact are neglected. These costs should be marginal for broad-based indices, though.

the global factors are market-weighted averages of the country-specific components. The Internet Appendix provides a detailed description of the construction process.

For the asset allocation case, we develop a framework aimed at decomposing the portfolio returns of the competing strategies. In the first step, we run a time series regression of the excess return of each model on the following three factors:

$$r_t - r_{f,t} = \alpha^{3F} + \beta_1 \cdot Stocks_t + \beta_2 \cdot Bonds_t + \beta_3 \cdot Commodities_t + \epsilon_t, \quad (2)$$

where $Stocks,$ $Bonds,$ and $Commodities$ represent the excess after-cost returns of the stock, bond, and commodity market, respectively. The economic interpretation of the coefficients is as follows. The betas represent the linear combination of asset class returns, which best approximates the time series of returns as generated by the model. In this sense, it gives an indication of the fixed-weight strategy that comes closest to the model's performance. For our heuristics, the alpha might be interpreted as the monthly return contribution of the rebalancing approach. For the optimization models, it might be regarded as the impact of the model's market timing on the overall portfolio return. For instance, minimum variance approaches are expected to, on average, rely heavily on bonds and much less on stocks and commodities. However, in some years, they might exhibit a substantially different asset allocation, as the models attempt to profit from uncommon changes in the risk-return structure of the input parameters. The alpha from the regression picks up the success from this market timing strategy.

In the second step, we extend this baseline approach to gain additional insights. To this end, we first construct zero-cost, long-short value and momentum portfolios for both bonds and commodities. Our methodology (see the Internet Appendix for details) closely follows recent work by Asness et al. (2013), who develop simple, intuitive value and momentum measures for these asset classes. The resulting factors can be thought of as proxies for return premia, which, so far, have primarily been exclusively studied in the stock market. They enable us to analyze to what extent portfolio returns are driven by loadings on these common factors. Specifically, we augment the regression specification as given above with three value factors (for stocks, bonds, and commodities), three momentum factors (for stocks, bonds, and commodities), as well as a size factor (for stocks only).[11]

4.2 Baseline results

4.2.1 International stock market diversification

We start the empirical analysis with a comparison of the performance of the 11 optimization models and the three heuristic models for an internationally-diversified stock portfolio. Results are reported in Table 3.

Please insert Table 3 here

[11] Data required for the construction of bond value and bond momentum factors is only available for the second sub-period (1988-2012) of our analysis. For the sake of comparability, we thus report results from both our three- and ten-factor regression only for this period. However, the qualitative nature of our findings from the three-factor model does not change if we rely on the full sample period (1973-2012). Note further that, judging from the inspection of correlations and variance inflation factors, multicollinearity does not seem to be an issue of concern in the case of the ten-factor model.

Columns 2 and 3 of Table 3 show that, after costs, average returns and standard deviations tend to be quite similar for most models. The minimum variance approach and its various extensions exhibit, as expected, the lowest fluctuation in returns. However, in economic terms, the reduction in risk, as compared to the standard deviation of the three heuristics, seems small. Consequently, full sample Sharpe ratios after costs tend to be similar for most approaches. The traditional mean-variance model with short sale restrictions, as well as its extension with constant correlations achieve the highest Sharpe ratio (0.132). However, this value is only marginally higher than the Sharpe ratios of the GDP- and naïvely-weighted portfolio heuristics (0.126 and 0.130, respectively). The value-weighted heuristic performs somewhat worse with a Sharpe ratio of 0.105, suggesting that it might be a less efficient diversification strategy.

To more formally address this issue, we analyze all pairwise differences in Sharpe ratios between the Markowitz models and the three heuristics using the bootstrap technique developed in Ledoit and Wolf (2008). For the sake of brevity, we only report *p*-values for the hypothesis that the Sharpe ratio for each of these models equals the Sharpe ratio of the GDP-weighted stock portfolio in Table 3. We find that none of the optimizing models significantly outperforms any of the three heuristics. Comparing the three heuristics against each other, the outperformance of the GDP-weighted over the popular value-weighted stock portfolio is marginally significant (*p*-value: .08).

To explore potential reasons for the widespread lack of statistical significance, we also examine the performance separately for two sub-periods. Results are reported in Columns 6 and 7 of Table 3. In general, there is no consistency in ranking across sub-periods. For instance, the traditional mean-variance model with short-selling constraints exhibits the highest Sharpe ratio in the second sub-period (1988-2012), but fails to add value over any of the heuristics in the first sub-period. Overall, the analysis suggests that there is no dominating approach.

Alphas from time series regressions of portfolio returns on a global one-factor Jensen (1968) or four-factor Carhart (1997) model do not lead to a different conclusion. Six optimization models as well as the naïvely- and GDP-weighted heuristic exhibit a positive, significant, and economically meaningful one-factor alpha. The highest level of statistical significance is, in fact, achieved for the two heuristic approaches. However, for the majority of approaches, the alpha becomes smaller once one controls for global momentum, value, and size effects. Only three extensions of the mean-variance framework still yield a marginally significant alpha.

This result highlights the importance of well-known risk premia for global index construction and portfolio optimization, which is not seen from an analysis of the Sharpe ratio or the one-factor alpha alone. For instance, we find that the GDP-weighted global stock portfolio loads significantly onto the premia associated with the international value and size factor, which prevents its excess return from remaining statistically significant. With regard to the value factor, we find a similar behavior also for the equal-weighted portfolio, as well as for all minimum variance approaches. A complete overview of the factor loadings associated with the portfolio models is given in the Internet Appendix.

Our analysis is based on after-cost returns because we are interested in whether optimization models add value under realistic conditions. It is a natural question to ask whether higher transaction costs prevent these models from achieving a better performance, in particular as they are only optimal under the assumption of no transaction costs. If so, it might still be worthwhile to set up an optimization approach to manage an equity portfolio, but to impose certain trading restrictions. As Table 3 shows, the mean turnover of all optimization models is indeed substantially larger than the turnover of the heuristics. However, its economic impact on our results is weak. Even before costs, none of the optimization models are able to significantly outperform the heuristics. Nevertheless, assuming higher transaction costs (> 40 bps) and more frequent (opposed to yearly) rebalancing generally works in favor of the heuristic models.

4.2.2 Diversification over asset classes

In the following, we additionally include bonds and commodities in the baseline analysis. Again, we compare the performance of 11 optimizing portfolio choice models with three heuristics. The latter only differ in their stock weighting scheme (value-weighted, equal-weighted, GDP-weighted). The proportion invested in bonds (25%) and commodities (15%) is the same across heuristics and motivated by the literature survey in Section 3.2.2. In Section 4.3, we extensively vary these portfolio weights to assess the sensitivity of our findings.

Please insert Table 4 here

Table 4 shows the main results. Compared to the international diversification in the stock market, there is less homogeneity in mean returns, standard deviations, and Sharpe ratios across optimization models. The minimum variance approach with short sale constraints and shrunk covariance matrix achieves the highest Sharpe ratio (0.186). In contrast, other strategies exhibit poor risk-adjusted returns. For instance, the Sharpe ratio of the traditional mean-variance model with short sale restrictions is only 0.102. This value is even lower than its Sharpe ratio in the case of international equity diversification (0.132), where it turned out to be among the most successful models. Hence, not all optimization approaches are able to realize the diversification potential of additional asset classes.

The performance of the fixed-weight heuristics is between the best and worst performing optimization models. However, the *p*-values reported in Table 4 reveal that we cannot reject the hypothesis of equal Sharpe ratios for the 60%-25%-15% asset allocation policy with GDP-weighting and any of the optimization models. In unreported results, we find that the same holds true when using the other heuristics as a benchmark.

The evidence from the asset allocation thus again supports the conclusion that optimizing portfolio choice models are not able to outperform a passive benchmark. However, the heterogeneity in Sharpe ratios among the optimization models reveals the intriguing possibility that some models are better suited to the asset allocation context than others. To investigate this issue, we implement our three- and ten-factor regression models. The intuition behind this approach is to decompose the portfolio weights induced by optimization approaches into a fixed-weight and a time-varying component. In that sense, these models are similar to the heuristic portfolio strategies. In contrast to the latter, however, the time-varying component does not reflect the contribution from simple rebalancing back to the original asset allocation, but the attempt to exploit recent changes in the return and risk characteristics of the asset classes in order to optimize the portfolio. Our regression framework picks up both the fixed-weight and the time-varying contribution to portfolio performance. The betas give an indication of which linear combination of fixed-weight asset allocation schemes would give a similar return time series as the optimization models themselves. The alphas might be interpreted as the additional value stemming from the time variation in portfolio weights.

However, as shown in the rightmost columns of Table 4, there is no additional value of the optimization procedures in the asset allocation context. In fact, two models yield a significant negative three-factor alpha and three models generate a significant negative ten-factor alpha. All these models aim at maximizing the Sharpe ratio. The alphas of all other (mostly variance-minimizing) optimization models are economically close to and statistically not significantly different from zero. Interestingly, the three-factor alphas of the fixed-weight heuristics with GDP- and equal-weighting in the stock domain are highly significantly positive. The latter approach is, moreover, the only strategy that also generates a significantly positive ten-factor alpha.

4.3 Variations in the fixed weight asset allocation strategy

We derive the 60%-25%-15% asset allocation strategy from the existing literature and use it as a benchmark for the optimization models. One potential concern regarding this approach may be that the good performance of the baseline heuristic results at least in part from backward optimization. To examine whether other possible heuristic strategies perform much worse, we calculate the Sharpe ratio after costs for a variety of different fixed-weight asset allocation schemes as well. In constructing the portfolios, we increase the portfolio weight of each asset class in steps of 1% from 0% to 100%, reduce the weight of the second class by the same amount, and hold the weight of the third portfolio constituent constant. Imposing a non-negativity constraint for portfolio weights, this approach yields 5,151 different portfolios.[12] The stock component of the portfolios is based on the GDP-weighting approach.

Figure 5 displays the results. In order to interpret the figure, note that the portfolio weight of the commodity component indirectly follows from the weights of the two other asset classes. For instance, the portfolio with 0% in stocks and 0% in bonds is completely invested in the commodity index.

Please insert Figure 5 here

Figure 5 shows a substantial increase in Sharpe ratios when moving away from portfolios with an extreme portfolio allocation (e.g., 100% of only one asset class). And, furthermore, the slope in the Sharpe ratio becomes flat as we move to the middle of the graph. This pattern suggests that a wide range of well-balanced allocation approaches over asset classes are able to offer substantial diversification gains. In fact, of the 5,151 tested portfolios, approximately 40% perform better or equal than our baseline heuristic and 60% perform worse. Those that perform worse are very often heavily tilted towards only one asset class. Inferences are similar if we consider the sub-periods from 1973-1988 and 1988-2012.[13] It follows that the 60%-25%-15% asset allocation policy is only one out of many different fixed-weight asset allocation schemes that achieve a good performance and are not dominated by sophisticated academic portfolio

[12] The number of portfolios can be explained as follows. Ignoring short sale restrictions yields a $N \times N$ matrix of different portfolios, where N equals the number of steps. However, $N \cdot (N-1)/2$ of these portfolios would lead to a short position in one asset class. In our case with 101 steps we have 10,201 portfolios of which 5,050 imply a short position. The difference of 5,151 is the number of portfolios analyzed.

[13] In terms of the Sharpe ratio, a large fraction of bonds in the portfolio tends to be more beneficial in the second sub-period than in the first sub-period. This appears to be mainly driven by the low average level of the risk-free rate in the recent past in combination with the low volatility of bonds.

models. This is good news for individual investors. Although it is not possible to identify the best performing portfolio ex ante, almost any form of well-balanced allocation of asset classes already offers Sharpe ratios similar to the best performing strategy.

4.4 Further results and robustness checks

In this section, we illustrate the economic meaningfulness of our results and verify their robustness in a number of sensitivity checks. These tests differ with respect to the data set, the rebalancing frequency, the input parameter estimation method for the Markowitz models, the implementation of the GDP-weighting heuristic, and the performance measure used. Where applicable, these tests are performed both for international diversification in the stock market and for additional diversification across asset classes.

4.4.1 Illustration of economic significance: Return gap

Since differences in Sharpe ratios are difficult to interpret from an economic point of view, we also rely on the return gap as a more intuitive performance measure, which is rooted in the risk-matching procedure suggested by Modigliani and Modigliani (1997). By combining the portfolio under consideration with the risk-free asset, Modigliani and Modigliani (1997) adjust the volatility of the portfolio to the volatility of the benchmark portfolio. Afterwards, the returns of the combined portfolio can be compared to the returns of the benchmark. More specifically, the return gap, $ReturnGap_t$, in month t is obtained from the following equation:

$$ReturnGap_t = r_{bm,t} - \left[\frac{\sigma_{bm}}{\sigma} r_t + \left(1 - \frac{\sigma_{bm}}{\sigma}\right) r_{f,t}\right], \tag{3}$$

where $r_{\mathrm{f,t}}$ is the risk-free rate in $t, r_{bm,t}$ stands for the return of the benchmark, and σ and σ_{bm} denote the monthly standard deviation of the portfolio and benchmark return over the sample period. We choose the GDP-weighted stock portfolio or the 60%-25%-15% asset allocation portfolio (with GPD-weighting in the stock component) as benchmarks. Using the GDP-weighted strategy as a benchmark allows us to assess the benefit of heuristic diversification in the stock universe. Relying on the 60%-25%-15% strategy as a benchmark is intended to exemplarily quantify the additional benefits obtained from a naïve fixed-weight allocation over different asset classes.

Table 5 verifies that heuristic diversification, both in the stock market and in the asset allocation case, adds value. With the exception of the MSCI Emerging Markets Index, the GDP-weighted strategy outperforms each of the 11 (national or regional) stock indices in terms of risk-adjusted return. Including additional asset classes, as implemented in the 60%-25%-15% portfolio, strengthens these results. Its outperformance, when benchmarked against the stock indices, ranges from 1.88 to 30.42 bps per month (or roughly 20 bps to well more than 350 bps per year) and thus is economically meaningful. The 60%-25%-15% portfolio also outperforms each single asset class (i.e., an aggregated stock, bond, or commodity portfolio) by roughly 5 to 20 bps per month. Collectively, the findings in Table 5 might be interpreted as exemplified evidence that relying on simple rules of thumb in diversifying substantially improves the risk-return profile of the overall portfolio.

Please insert Table 5 here

4.4.2 Variation in the data set

We extensively vary the data set to examine whether our findings are robust with respect to the indices used to represent the asset classes. First, we exclude the MSCI Emerging Markets Index, which is not available prior to 1988 from the calculations. Second, we rely on the country-specific MSCI indices for the G-7 states instead of the regional MSCI indices. Third, we redo our analysis in the asset allocation context using only the MSCI world as the stock market component. Fourth, we also use alternative indices for bonds and commodities as outlined in Section 2.1. This procedure often leads to a reduction in the sample size, since most index alternatives have a shorter return data history. However, in the overall picture, we find that the variation in the data set does not alter any of our conclusions.

4.4.3 Rebalancing frequency

Monthly instead of annual rebalancing does not lead to significantly better results for both the optimizing portfolio models and the heuristics. While the Sharpe ratios increase for some optimization models, they decrease for others. At the same time, we observe not surprisingly a substantial increase in turnover rates, which casts further doubt on the benefits of a higher rebalancing frequency. For the heuristics, the rather minor importance of the rebalancing frequency can also be inferred from the insignificant alphas in Table 4, as well as from Figure 5. The latter shows that shifts in the portfolio weights are not harmful as long as the portfolio is not tilted too extremely towards only one asset. In this regard, the major benefit of portfolio rebalancing is to avoid extreme portfolios consisting of mainly only one asset.

4.4.4 Parametrization

In the baseline analysis, we use a time window of 60 months to estimate the input parameters for the Markowitz-based models. To examine whether the performance of these models improves when a longer time-series of historical returns is used for parametrization, we base the estimation method also on a rolling-window approach with 1) 120 months and with 2) all historical data available in a particular month. We do not observe a consistent improvement in the results of the Markowitz models in the additional tests. Furthermore, the out-of-sample Sharpe ratios are still not significantly different from those of the heuristic models. This holds for both international equity diversification and diversification over asset classes.

4.4.5 Implementation of the GDP-weighting heuristic

We change the methodology of the GDP-weighting scheme in two ways. First, we base portfolio weights on the relative GDP of the next year to proxy for rational expectations. Second, we use GDP weights derived from purchasing power parity (PPP) valuations as provided by the World Bank and the International Monetary Fund (IMF). The performance of the GDP-weighting scheme is virtually unchanged in the first check and slightly improves in the second check.

4.4.6 Other performance measures

The recent literature has proposed a number of alternative performance ratios. Therefore, we repeat our analysis utilizing asymmetrical performance measures that have been shown to be particularly suited for non-normal return distributions (e.g., Biglova et al., 2004; Farinelli et al., 2008; Farinelli et al., 2009). Specifically, we employ the Sortino ratio, the Rachev ratio, and the Generalized Rachev ratio.[14] The Sortino ratio is computed as the average excess return over the risk-free rate divided by the downside volatility of the excess return. The Rachev ratio relies on the conditional value at risk of the excess return. Portfolios with the highest Rachev ratios are the ones that best manage to simultaneously deliver high returns and get insurance for high losses.

[14] For a detailed description of these ratios, we refer the reader to Biglova et al. (2004) and Rachev et al. (2007). To implement the ratios, we apply the parametrization described in Biglova et al. (2004) and Farinelli et al. (2008).

The General Rachev ratio additionally takes investors degree of risk aversion into account. Utilizing these alternative measures does not change the qualitative nature of our results. Heuristic portfolio allocation mechanisms still yield similar results compared to optimizing portfolio choice models. Furthermore, there is no consistency in ranking across performance ratios, which again indicates that there is no overall dominating approach.

5. Conclusion

We examine the investment value of heuristic diversification strategies as a possible remedy against widespread costly investment mistakes. The field of household finance suggests that many individual investors do not fully exploit the benefits of diversification and incur non-trivial welfare costs as a consequence. Given this context, we ask whether and which simplistic guidelines offer a promising way for investors to diversify. We compare 11 optimization methods favored or recently proposed in the literature with a broad range of heuristic allocation strategies, both for international stock market diversification and in the asset allocation case.

Our main results can be summarized as follows. First, for global equity diversification, prominent optimization models do not outperform heuristic stock weighting schemes. Global value, momentum, and size premiums are important drivers of the portfolio performance of many strategies, both scientific and heuristic. Second, the inclusion of additional asset classes is, in general, highly beneficial. Diversification gains are mainly driven by a well-balanced allocation over different asset classes. As long as the portfolio is not heavily tilted towards one asset class, almost any form of naïve fixed-weight allocation strategy realizes diversification potential. Third, in the asset allocation context, optimization methods again do not add substantial value.

Our findings are good news for individual investors: relying on simple rules of thumb in asset allocation significantly improves the performance of any single asset class portfolio. Moreover, following these easily implementable strategies does not lead to lower risk-adjusted returns as compared to even very sophisticated and recently proposed portfolio choice models.

Our study suggests several directions for further research. First, provided the availability of reliable data, the analysis could be extended to other asset classes. Eun et al. (2008) and Petrella (2005), for example, argue that investors can gain additional diversification benefits from small and mid caps. Second, instead of relying on historical data to estimate input parameters, alternative approaches (such as implied return estimates from analyst forecasts or option prizes) could be analyzed. Third, future research should explore whether combining portfolio optimization concepts with heuristic allocation schemes is a fruitful direction. Within a bottom-up approach, for example, minimum variance models could be implemented on an individual asset level (e.g., Jagannathan and Ma, 2003), while plausible heuristics might be used on an index or asset class level.

References

Agnew, J., P. Balduzzi, and A. Sunden, 2003, "Portfolio Choice and Trading in a Large 401(k) Plan," *American Economic Review*, 93, 193–215.

Amin, G. S., and H. M. Kat, 2003, "Stocks, Bonds and Hedge Funds: Not a Free Lunch!", *Journal of Portfolio Management*, 29, 113–120.

Annaert, J., M. J. D. Ceuster, and W. V. Hyfte, 2005, "The Value of Asset Allocation Advice: Evidence from The Economist's Quarterly Portfolio Poll," *Journal of Banking and Finance*, 29, 661–680.

Anson, M. J., 1999, "Maximizing Utility with Commodity Futures Diversification," *Journal of Portfolio Management*, 25, 86–94.

Arnott, R. D., J. Hsu, and P. Moore, 2005, "Fundamental Indexation," *Financial Analysts Journal*, 61, 83–99.

Arshanapalli, B., T. D. Coggin, and W. Nelson, 2001, "Is Fixed-Weight Asset Allocation Really Better?," *Journal of Portfolio Management*, 27, 27–38.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen, 2013, "Value and Momentum Everywhere," *Journal of Finance forthcoming*.

Barber, B. M., and T. Odean, 2000, "Trading is Hazardous to Your Wealth," *Journal of Finance*, 55, 773–806.

Benartzi, S., and R. Thaler, 2007, "Heuristics and Biases in Retirement Savings Behavior", *Journal of Economic Perspectives*, 21, 81–104.

Biglova, A., S. Ortobelli, S. Rachev, and S. Stoyanov, 2004, "Different Approaches to Risk Estimation in Portfolio Theory," *Journal of Portfolio Management*, 31, 103–112.

Black, F., and R. Litterman, 1992, "Global Portfolio Optimization," *Financial Analysts Journal*, 48, 28–43.

Blake, D., B. N. Lehmann, and A. Timmermann, 1999, "Asset Allocation Dynamics and Pension Fund Performance," *Journal of Business*, 72, 429–461.

Brinson, G. P., L. R. Hood, and G. L. Beebower, 1986, "Determinants of Portfolio Performance," *Financial Analysts Journal*, 42, 39–44.

Brown, S. J., 1979, "Optimal Portfolio Choice Under Uncertainty: A Bayesian Approach," *Estimation Risk and Optimal Portfolio Choice, ed., Bawa, Brown and Klein*.

Calvet, L., J. Y. Campbell, and P. Sodini, 2007, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115, 707–747.

Campbell, J. Y., 2006, "Household Finance," *Journal of Finance*, 61, 1553–1604.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," Journal of Finance, 52, 57–82.

Chopra, V. K., C. R. Hensel, and A. L. Turner, 1993, "Massaging Mean Variance Inputs: Returns from Alternative Global Investment Strategies in the 1980s," *Management Science*, 39, 845–855.

Comer, G., N. Larrymore, and J. Rodriguez, 2009, "Controlling for Fixed Income Exposure in Portfolio Evaluation: Evidence from Hybrid Mutual Funds," *Review of Financial Studies*, 22, 481–507.

De Roon, F. A., T. E. Nijman, and B. J. M. Werker, 2001, "Testing for Mean-Variance spanning with Short Sales Constraints and Transaction Costs: The Case of Emerging Markets," *Journal of Finance*, 56, 721–742.

De Santis, G., and B. Gerard, 1997, "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *Journal of Finance*, 52, 1881–1912.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal, 2009a, "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science*, 55, 798–812.

DeMiguel, V., L. Garlappi, and R. Uppal, 2009b, "Optimal versus Naïve Diversification How Inefficient is the 1/N Portfolio Strategy?," *Review of Financial Studies*, 22, 1915–1953.

Driessen, J., and L. Laeven, 2007, International Portfolio Diversification Benefits: Cross-Country Evidence from a Local Perspective," *Journal of Banking and Finance*, 31, 1693–1712.

Duchin, R., and H. Levy, 2009, "Markowitz Versus the Talmudic Portfolio Diversification Strategies," *Journal of Portfolio Management*, 35, 71–74.

Ennis, R. M., and M. D. Sebastian, 2005, "Asset Allocation with Private Equity," *Journal of Private Equity*, 8, 81–87.

Erb, C. B., and C. R. Harvey, 2006, "The Strategic and Tactical Value of Commodity Futures," *Financial Analysts Journal*, 62, 69–97.

Eun, C. S., W. Huang, and S. Lai, 2008, "International Diversification with Large- and Small-Cap Stocks," *Journal of Financial and Quantitative Analysis*, 43, 489–524.

Eun, C. S., and B. G. Resnick, 1994, "International Diversification of Investment Portfolios: U.S. and Japanese Perspectives," *Management Science*, 40, 140–161.

Fama, E., and K. R. French, 2010, "Luck versus Skill in the Cross Section of Mutual Fund Returns," J*ournal of Finance*, 65, 1915-1947.

Farinelli, S., M. Ferreira, D. Rossello, M. Thoeny, and L. Tibiletti, 2008, "Beyond Sharpe Ratio: Optimal Asset Allocation Using Different Performance Ratios," *Journal of Banking and Finance*, 32, 2057–2063.

———, 2009, "Optimal Asset Allocation Aid System: From One-Size vs Tailor-Made Performance Ratio," *European Journal of Operational Research*, 192, 209215.

Frankfurter, G. M., H. E. Phillips, and J. P. Seagle, 1971, "Portfolio Selection: The Effects of Uncertain Means, Variances and Covariances," *Journal of Financial and Quantitative Analysis*, 7, 1251–1262.

French, K. R., and J. M. Poterba, 1991, "Investor Diversification and International Equity Markets," *American Economic Review*, 81, 222–226.

Goetzmann, W. N., and A. Kumar, 2008, "Equity Portfolio Diversification," *Review of Finance*, Vol. 12, No. 3, 433-463, 2008.

Goetzmann, W. N., L. Li, and K. G. Rouwenhorst, 2005, "Long-Term Global Market Corre lations," *Journal of Business*, 78, 1–38.

Gorton, G., and K. G. Rouwenhorst, 2006, "Facts and Fantasies about Commodity Futures," *Financial Analysts Journal*, 62, 47–68.

Griffin, J. M., 2002, "Are the Fama and French Factors Global or Country Specific?," *Review of Financial Studies*, 15, 783–803

Grinblatt, M., and M. Keloharju, 2001, "How Distance, Language and Culture Influence Stockholdings and Ttrades," *Journal of Finance*, 56, 1053–1073.

Haugen, R. A., and N. L. Baker, 1991, "The Efficient Market Inefficiency of Capitalization-Weighted Stock Portfolios," *Journal of Portfolio Management*, 17, 35–40.

Ibbotson, R. G., and P. D. Kaplan, 2000, "Does Asset Allocation Policy Explain 40, 90 or 100 Percent of Performance?," *Financial Analysts Journal*, 56, 26–33.

Jagannathan, R., and T. Ma, 2003, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1683.

James, W., and C. Stein, 1961, "Estimation with Quadratic Loss," *Procedings of the*

Fourth Berkeley Symposium on Probability and Statistics, 1, 361–379.

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

Jobson, J. D., and B. M. Korkie, 1981, “Performance Hypothesis Testing with the Sharpe and Treynor Measures,” *Journal of Finance*, 36, 889–908.

Jorion, P., 1986, “Bayes-Stein Estimation for Portfolio Analysis,” *Journal of Financial and Quantitative Analysis*, 21, 279–292.

Kilka, M., and M. Weber, 2000, “Home Bias in International Stock Return Expectations”, *Journal of Psychology and Financial Markets*, 1, 176–192.

Kimball, M. S., and T. Shumway, 2010, “Investor Sophistication and the Home Bias, Diversification, and Employer Stock Puzzles,” Unpublished working paper, University of Michigan.

Kritzman, M., S. Page, and D. Turkington, 2010, “In Defense of Optimization: The Fallacy of 1/N,” *Financial Analysts Journal*, 66, 31–39.

Ledoit, O., and M. Wolf, 2004, “Honey, I Shrunk the Sample Covariance Matrix,” *Journal of Portfolio Management*, 31, 110–119.

———, 2008, “Robust Performance Hypothesis Testing with the Sharpe Ratio,” *Journal of Empirical Finance*, 15, 850–859.

Longin, F., and B. Solnik, 2001, “Extreme Correlation of Tnternational Equity Markets,” *Journal of Finance*, 56, 649–676.

Markowitz, H., 1952, “Portfolio Selection,” *Journal of Finance*, 7, 77–91.

Michaud, R. O., 1998, *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. HBS Press.

Modigliani, F., and L. Modigliani, 1997, “Risk-Adjusted Performance,” *Journal of Portfolio Management*, 23, 45–54.

Odean, T., 1999, “Do Investors Trade Too Much?,” *American Economic* Review, 89, 1279–1298.

Patton, A. J., 2009, “Are ”Market Neutral” Hedge Funds Really Market Neutral?,” *Review of Financial Studies*, 22, 2495–2530.

Petrella, G., 2005, “Are Euro Area Small Cap Stocks an Asset class? Evidence from Mean-Variance Spanning Tests,” *European Financial Management*, 11, 229–253.

Phalippou, L., and O. Gottschalg, 2009, "The Performance of Private Equity Funds," *Review of Financial Studies*, 22, 1747–1776.

Polkovnichenko, V., 2005, "Household Portfolio Diversification: A Case for Rank Dependent Preferences," *Review of Financial Studies*, 18, 1467–1502.

Rachev, S. T., T. Jaic, S. Stoyanov, and F. J. Fabozzi, 2007, "Momentum Strategies Based on Reward-Risk Stock Selection Criteria," *Journal of Banking and Finance*, 31, 2325–2346.

Shiller, R. J., 1981, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–436.

Siegel, J. J., 2006, "The 'Noisy Market' Hypothesis," *The Wall Street* Journal, June 14 2006.

Tang, N., O. S. Mitchell, G. R. Mottola, and S. P. Utkus, 2010, "The Efficiency of Sponsor and Participant Portfolio Choices in 401(K) Plans," Journal of Public Economics, 94, 1073-1085.

Treynor, J., 2005, "Why Market-Valuation-Indifferent Indexing Works," *Financial Analysts Journal*, 61, 65–69.

Tu, J., and G. Zhou, 2009, "Markowitz Meets Talmud: A Combination of Sophisticated and Naïve Diversification Strategies," *Journal of Financial Economics*, 99, 204-215.

Table 1: Descriptive statistics for the different indices

This table reports monthly return parameters of the various indices that we consider for portfolio construction. Returns are calculated using Datastream's total return index and denominated in euro. *Global Stock Index* is a market-weighted stock index comprising the four different regional stock indices: MSCI Europe, MSCI North America, MSCI Pacific, and MSCI Emerging Markets.

Asset Class/ Region	Sample Period	Sharpe Ratio	Mean Return	Std. Dev.	VaR 95%
Stocks: Regional Indices					
Emerging Markets	88-12	0.132	1.28%	7.12%	-12.11%
Europe	73-12	0.120	0.98%	4.67%	-7.76%
North America	73-12	0.105	0.96%	5.20%	-7.69%
Pacific	73-12	0.070	0.82%	5.68%	-8.82%
Average (excl. Emerging Markets)	73-12	0.098	0.92%	5.18%	-8.09%
Average	88-12	0.089	0.84%	5.60%	-9.29%
Global Stock Index	73-12	0.105	0.90%	4.61%	-7.84%
Global Stock Index	88-12	0.081	0.71%	4.67%	-8.74%
Other Asset Classes					
Bonds	73-12	0.132	0.56%	1.09%	-1.24%
Commodities	73-12	0.078	0.89%	6.11%	-9.16%

Table 2: List of portfolio models

This table lists the various Markowitz-based optimization models from the existing literature (Panel A) and heuristic models (Panel B) that we consider for portfolio construction. δ is the threshold parameter developed in DeMiguel et al. (2009a) to limit the norm of the portfolio weight vector. The last column gives the abbreviation that we use to refer to the model.

No.	Portfolio Model	Abbreviation
	Panel A: Portfolio optimization models from the existing literature	
1	Maximum Sharpe ratio approach with short sale constraints	maxsr-noshort
2	James/Stein estimator of expected returns with short sale constraints	maxsr-noshort-js
3	Maximum Sharpe ratio approach plus Ledoit/Wolf constant correlation model with short sale constraints	maxsr–noshort-ccm
4	James/Stein estimator of expected returns plus Ledoit/Wolf constant correlation model with short sale constraints	maxsr-noshort-js-ccm
5	Minimum variance approach without short sale constraints	minvar
6	Minimum variance approach with short sale constraints	minvar-noshort
7	Minimum variance approach plus Ledoit/Wolf constant correlation model with short sale constraints	minvar-noshort-ccm
8	1-norm constrained minimum variance portfolio with δ calibrated using cross-validation over portfolio variance	minvar-nc-1v
9	1-norm constrained minimum variance portfolio with δ calibrated by maximizing portfolio return in previous period	minvar-nc-1r
10	2-norm constrained minimum variance portfolio with δ calibrated using cross-validation over portfolio variance	minvar-nc-2v
11	2-norm constrained minimum variance portfolio with δ calibrated by maximizing portfolio return in previous period	minvar-nc-2r
	Panel B: Heuristic portfolio models considered in this paper	
12	GDP-weighted stock portfolio	GDP
13	Market-weighted stock portfolio	macap
14	Equally-weighted stock portfolio (1/N)	naïve
15	Asset Allocation Model with the following weights: 60% stocks, 25% bonds, and 15% commodities; stock portfolio is GDP-weighted	60-25-15; GDP
16	Asset Allocation Model with the following weights: 60% stocks, 25% bonds, and 15% commodities; stock portfolio is market-weighted	60-25-15; macap
17	Asset Allocation Model with the following weights: 60% stocks, 25% bonds, and 15% commodities; stock portfolio is equally-weighted	60-25-15; naïve

Table 3: Optimizing vs. heuristics: results for international stock market diversification

This table reports mean returns, return standard deviations, Sharpe ratios, and alphas of monthly out-of sample returns after costs, as well as average turnover for international equity portfolios that are constructed using eleven optimization models and three heuristics. Sharpe ratios are reported for the total sample period (February 1973-December 2012) and two sub-sample periods (February 1973-January 1988 and February 1988-December 2012). *P*-values for the hypothesis that the Sharpe ratio of a given model is equal to the Sharpe ratio of the GDP-weighted stock portfolio are calculated using the bootstrap technique developed in Ledoit and Wolf (2008). We assume a bid-ask spread of 40 bps to calculate after-cost returns. α^{1F} is the Jensen (1968) one-factor alpha; α^{4F} is the Carhart (1997) four-factor alpha. For the *t*-statistics, * (**/***) indicates significance at the 10% level (5% level / 1% level). See Section 3 and Table 2 for a description of portfolio construction models. Details on the construction of the factors used in the regression framework are provided in the Internet Appendix.

Portfolio Model	Mean Return	Std. Dev. Return	Mean Annual Turnover	Sharpe Ratio 1973-2012	Sharpe Ratio 1973-1988	Sharpe Ratio 1988-2012	*P*-value H_0: $SR=SR_{gdp}$	$\alpha_1 F$	*t*-stat α_1F	$\alpha_4 F$	*t*-stat α_4F
Panel A: Optimization Models											
maxsr-noshort	1.16%	5.62%	51.45%	0.132	0.123	0.138	0.71	0.24%*	1.77	0.27%	1.84*
maxsr-noshort-js	0.98%	4.84%	73.72%	0.116	0.123	0.112	0.80	0.11%	1.07	0.17%	1.53
maxsr–noshort-ccm	1.16%	5.59%	46.48%	0.132	0.128	0.135	0.73	0.24%*	1.76	0.27%	1.89*
maxsr-noshort-js-ccm	0.99%	4.87%	68.82%	0.117	0.128	0.110	0.80	0.12%	1.09	0.19%	1.70*
minvar	0.91%	4.37%	32.63%	0.113	0.165	0.081	0.62	0.07%	0.88	0.03%	0.35
minvar-noshort	0.98%	4.41%	23.18%	0.127	0.165	0.105	0.87	0.12%*	1.91	0.09%	1.38
minvar-noshort-ccm	0.97%	4.42%	20.51%	0.124	0.157	0.105	0.96	0.11%*	1.65	0.09%	1.35
minvar-nc-1v	0.97%	4.38%	26.80%	0.127	0.165	0.105	0.89	0.12%*	1.80	0.09%	1.20
minvar-nc-1r	0.92%	4.37%	31.66%	0.115	0.165	0.086	0.70	0.08%	1.02	0.04%	0.48
minvar-nc-2v	0.97%	4.39%	25.77%	0.125	0.165	0.101	0.98	0.11%*	1.72	0.08%	1.15
minvar-nc-1r	0.92%	4.37%	29.75%	0.116	0.165	0.086	0.62	0.07%	1.06	0.04%	0.55
Panel B: Heuristic Models											
gdp	1.01%	4.68%	11.54%	0.126	0.155	0.109	.	0.11%**	2.28	0.08%	1.52
macap	0.90%	4.61%	4.45%	0.105	0.157	0.074	0.08*	.	.	.	.
naïve	1.02%	4.66%	12.24%	0.130	0.182	0.102	0.46	0.13%***	2.65	0.08%	1.55

Table 4: Optimizing vs. heuristics: Results for asset allocation

This table reports mean returns, return standard deviations, Sharpe ratios and alphas of monthly out-of sample returns after costs as well as average turnover for asset allocation portfolios which are constructed using eleven optimization models and three heuristics. Sharpe ratios are reported for the total sample period (February 1973-December 2012) and two sub-sample periods (February 1973-January 1988 and February 1988-December 2012). *P*-values for the hypothesis that the Sharpe ratio of a given model is equal to the Sharpe ratio of the "60%25%15%; GDP" heuristic are calculated using the bootstrap technique developed in Ledoit and Wolf (2008). We assume a bid-ask spread of 40 bps to calculate after-cost returns. α^{3F} is the intercept from a three-factor model including the market, bond, and commodity factor; α^{10F} is the intercept from a ten-factor model, augmented by value, size, and momentum factors. Data required for the construction of bond, value, and momentum factors is only available for the second sub-period (1988-2012). For the sake of comparability, we thus report results from both the three-factor and the ten-factor regression only for this period. For the *t*-statistics, * (**/***) indicates significance at the 10% level (5% level / 1% level). See Section 3 and Table 2 for a description of portfolio construction models. Details on the construction of the factors used in the regression framework are provided in the Internet Appendix.

Portfolio Model	Mean Return	Std. Dev. Return	Mean Annual Turnover	Sharpe Ratio 1973-2012	Sharpe Ratio 1973-1988	Sharpe Ratio 1988-2012	*P*-value H_0: SR=SR_{gdp}	$\alpha_3 F$	*t*-stat $\alpha_3 F$	$\alpha_{10} F$	*t*-stat $\alpha_{10} F$
Panel A: Optimization Models											
maxsr-noshort	0.77%	3.51%	47.63%	0.102	0.176	0.049	0.37	-0.19%	-1.37	-0.24%	-1.64
maxsr-noshort-js	0.72%	2.57%	38.92%	0.121	0.180	0.069	0.63	-0.11%	-1.30	-0.14%	-1.58
maxsr–noshort-ccm	0.75%	3.60%	50.23%	0.093	0.180	0.034	0.27	-0.24%	-1.69*	-0.29%	-1.83*
maxsr-noshort-js-ccm	0.69%	2.64%	43.27%	0.106	0.183	0.042	0.40	-0.17%	-1.93*	-0.21%	-2.15**
minvar	0.58%	1.09%	12.65%	0.155	0.159	0.153	0.91	-0.01%	-0.30	0.00%	0.09
minvar-noshort	0.61%	1.07%	6.84%	0.179	0.176	0.183	0.63	0.01%	0.56	0.01%	0.43
minvar-noshort-ccm	0.62%	1.10%	5.89%	0.186	0.181	0.193	0.54	0.01%	1.54	0.00%	0.21
minvar-nc-1v	0.61%	1.07%	7.49%	0.179	0.177	0.183	0.62	0.01%	0.56	0.01%	0.43
minvar-nc-1r	0.58%	1.08%	12.57%	0.157	0.159	0.156	0.89	-0.01%	-0.25	0.00%	0.07
minvar-nc-2v	0.61%	1.07%	6.84%	0.179	0.176	0.183	0.63	0.01%	0.56	0.01%	0.43
minvar-nc-1r	0.60%	1.07%	8.06%	0.171	0.176	0.171	0.71	0.00%	0.04	0.01%	0.35
Panel B: Heuristic Models											
60-25-15; gdp	0.89%	3.17%	12.92%	0.150	0.186	0.130	.	0.12%	3.21***	0.08%	1.84*
60-25-15; macap	0.83%	3.11%	10.11%	0.133	0.188	0.101	0.11	.	.	.	.
60-25-15; naïve	0.91%	3.16%	13.23%	0.155	0.212	0.124	0.39	0.10%	2.70***	0.04%	1.07

Table 5: Return gaps relative to GDP-weighted stock portfolio and 60%-25%-15% asset allocation

This table reports the Sharpe ratio and the Value-at-Risk (at the 95% confidence level) of monthly returns for various indices, as well as for the GDP-weighted stock portfolio and the 60%-25%-15% asset allocation strategy (with GDP-weighting the stock market). Moreover, the table presents the return gap of these indices in basis points (bps) per month compared to the GDP-weighted stock portfolio and the 60%-25%-15% asset allocation strategy (with GDP-weighting the stock market). Portfolio weights are readjusted every February. See Section 3 and Table 2 for a description of the models and Subsection 4.4 for a description of the return gap.

Asset Class/ Region	Sample Period	Sharpe Ratio	VaR 95%	Return Gap (bps per month) GDP-stock portfolio	Return Gap (bps per month) 60-25-15 portfolio
Panel A: Stock Indices					
MSCI Germany	73-12	0.106	-9.01%	8.18	13.26
MSCI France	73-12	0.102	-9.03%	9.75	14.32
MSCI Italy	73-12	0.056	-10.45%	33.53	30.42
MSCI United Kingdom	73-12	0.102	-9.15%	12.11	15.92
MSCI United States	73-12	0.101	-7.73%	11.93	15.80
MSCI Canada	73-12	0.091	-8.81%	16.09	18.62
MSCI Japan	73-12	0.059	-9.20%	31.16	28.82
MSCI Europe	73-12	0.120	-7.76%	2.54	9.44
MSCI North America	73-12	0.105	-7.69%	10.20	14.63
MSCI Pacific	73-12	0.070	-8.82%	25.28	24.84
MSCI Emerging Markets	88-12	0.132	-12.11%	-7.39	1.88
Panel B: Asset Classes					
GDP-stock portfolio	73-12	0.126	-7.66%	.	7.73
Bonds	73-12	0.132	-1.24%	-3.20	5.55
Commodities	73-12	0.078	-9.16%	21.88	22.53

Figure 1: Time series behavior of correlations within the stock market

This figure depicts the movement in the average correlation over the sample period for the regional stock indices MSCI Europe, MSCI North America, MSCI Pacific, and MSCI Emerging Markets with respect to all other stock indices. Correlation coefficients are computed using a rolling-window approach based on the previous 60 months.

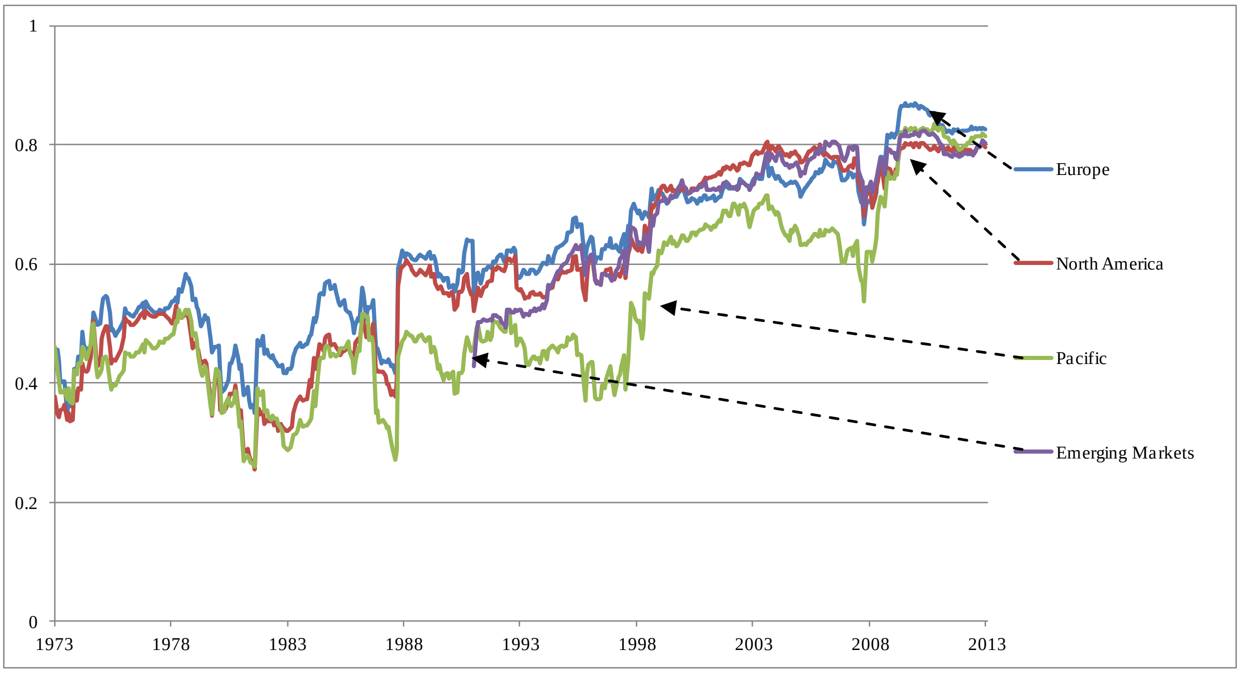

Figure 2: Time series behavior of correlations between asset classes

This figure depicts the movement in the average correlation over the sample period between the iBoxx Euro Overall Index (REXP before 1999) and the regional MSCI stock indices, as well as between the S&P GSCI Commodity Total Return Index and the regional MSCI stock indices. Correlations are computed separately for each index pair and then averaged using a rolling-window approach based on the previous 60 months.

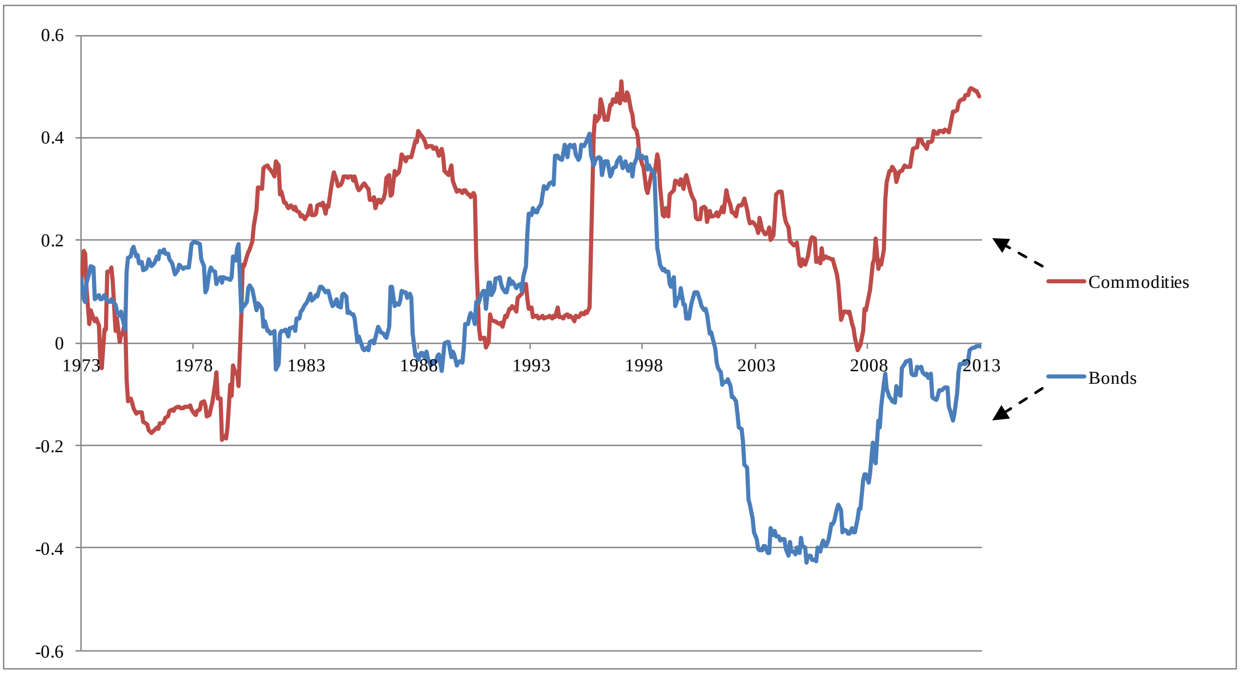

Figure 3: Time series evolution of portfolio weights of a cap-weighted stock index

This figure depicts the portfolio weights of a market value-weighted global stock index comprised of stocks from North America, Europe, the Pacific region, and the Emerging Markets over the sample period from 1973 to 2012. The data source is Thomson Reuters Datastream.

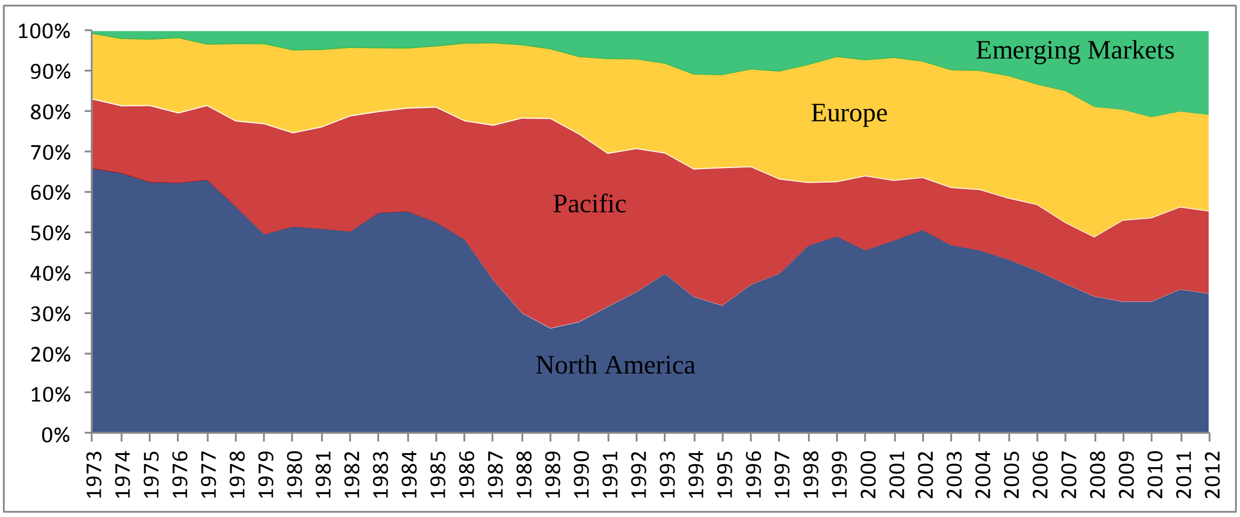

Figure 4: Time series evolution of portfolio weights of a GDP-weighted stock index

This figure depicts the portfolio weights of a GDP-weighted stock index comprised of stocks from North America, Europe, the Pacific region, and the Emerging Markets over the sample period from 1973 to 2012. Data sources are the World Bank for the period 1973-2005 and the International Monetary Fund for the period 2006-2012.

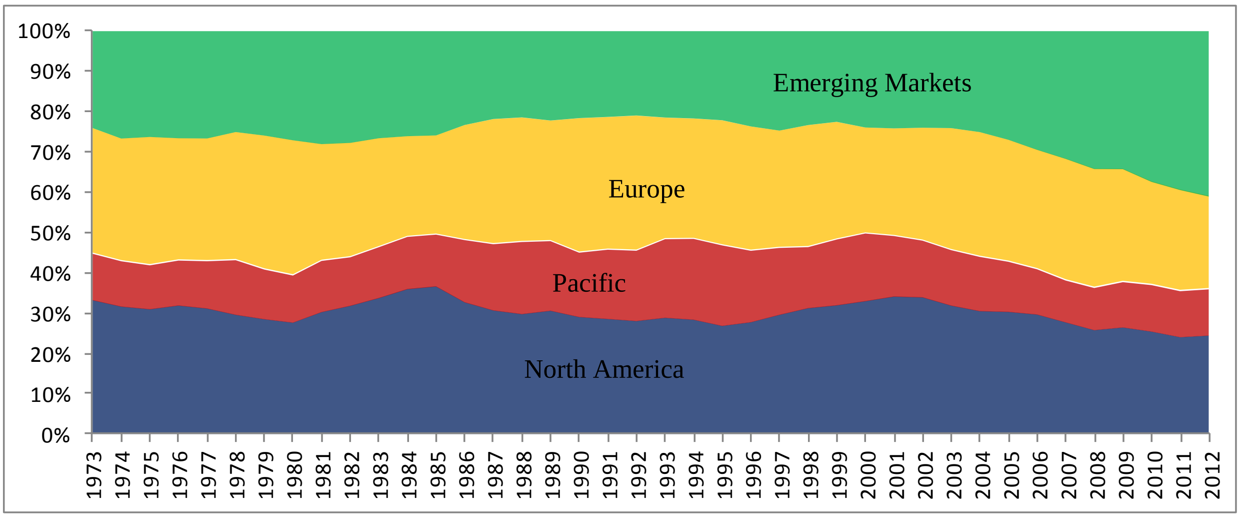

Figure 5: Performance of alternative fixed-weight asset allocation strategies

This figure depicts the Sharpe ratios of alternative heuristic portfolio strategies in the asset allocation context. In constructing the portfolios, we increase the portfolio weight of each asset class at the rebalancing date in steps of 1% from 0% to 100% and adjust the portfolio weights of the other two classes accordingly. This approach yields 5,151 different portfolios. The stock component of the portfolio is comprised of the four regional MSCI indices and is GDP-weighted.