How should private investors diversify? - An empirical evaluation of alternative asset allocation policies to construct a "world market portfolio"

Heiko Jacobs, Sebastian Müller and Martin Weber*

November 16, 2010

Abstract

This paper evaluates numerous diversification strategies as a possible remedy against widespread costly investment mistakes of private investors. Our approach allows us to analyze competing policies for the construction of a "world market portfolio". Our results reveal that a very broad range of simple heuristic allocation schemes offers similar diversification gains as even 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 private investors.

Keywords: portfolio theory, household finance, asset allocation, international diversification, heuristics

JEL Classification Code: G11

*Heiko Jacobs is from the Lehrstuhl für Bankbetriebslehre, Universität Mannheim, L 5, 2, 68131 Mannheim. E-Mail: jacobs@bank.BWL.uni-mannheim.de. Sebastian Müller is from the Lehrstuhl für Bankbetriebslehre, Universität Mannheim, L 5, 2, 68131 Mannheim. E-Mail: mueller@bank.BWL.uni-mannheim.de. Martin Weber is from the Lehrstuhl für Bankbetriebslehre, Universität Mannheim, L 5, 2, 68131 Mannheim and CEPR, London. E-Mail: weber@bank.BWL.uni-mannheim.de. We are grateful to Gerd Kommer, Olaf Scherf, Volker Vonhoff, and seminar participants at the Gutmann Center for Portfolio Management (Vienna University of Economics and Business), the portfolio ETF forum, the European Business School, the annual meeting of the German Finance Association (DGF) in Münster, the Campus for Finance Research Conference (CFF) and the University of Mannheim for helpful 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 benefits of diversification as "the only free lunch in investment", private investors seem to sometimes violate even its 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 thus aim to derive easily implementable asset allocation guidelines for individual investors. Our approach allows us to evaluate numerous competing policies for the construction of a "world market portfolio". Specifically, we ask the following questions: From the perspective of private 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 able to realize diversification potential? To what extent do these strategies underperform when benchmarked against sophisticated optimization models?

Empirical studies provide extensive evidence of private investors making portfolio choices that are difficult to reconcile with standard financial theory. As such, households often fail to participate in the stock market at all (see, e.g., Campbell (2006) and Kimball and Shumway (2010)). Given the size of the equity premium over the past, the welfare costs of this behavior are likely to be high. Among those households that do invest in equities, many studies document further costly mistakes. First, individuals tend to prefer domestic over foreign investments thereby forgoing the benefits of international diversification (see 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 (see, 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 (see Odean (1999) and Barber and Odean (2000)). Puzzling investment behavior carries over to diversification over asset classes. Analyzing a large sample of retirement accounts, Agnew et al. (2003) show that most asset allocations are extreme (either 100 percent or zero percent in equities) and there is inertia in asset allocations. Tang et al. (2009) 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 (see e.g., Brinson et al.

(1986) or Ibbotson and Kaplan (2000)). Additionally, recent findings on the correlation structure of international stock markets imply that even worldwide equity market diversification can offer only limited benefits. Increasing return correlations within the stock universe over the last decades (Goetzmann et al. (2005)) lead to decreasing diversification gains (Driessen and Laeven (2007)). Moreover, correlations tend to be particularly high in periods of poor performance (see e.g. Longin and Solnik (2001)). Thus, benefits from global diversification in the stock market tend to be smallest when they are most needed.

To sum up, risk-adjusted portfolios of most private households underperform even standard domestic stock market indices at a significant margin, and thus leave substantial room for improvement. But how should private 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 private investors as most individuals will not have the knowledge and resources to implement sophisticated extensions of the Markowitz model. In addition, Markowitz approaches, while being optimal in theory, suffer from estimation error in expected returns, variances and covariances when implemented in practice. There is a large literature explicitly dealing with how to improve the out-of-sample performance of these strategies - with partly disillusioning results. Recent studies focussing primarily on U.S. stock portfolios show that the estimation error is so severe that various optimization models are oftentimes unable to beat a naive 1/N diversification strategy (see, e.g. DeMiguel et al. (2009b), Tu and Zhou (2009), and Duchin and Levy (2009)).[1]

Hence, it seems insufficient to limit the analysis to these models. In the empirical analysis, we thus analyze the performance of eleven well-established or recently proposed extensions of the Markowitz (1952) framework 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.

[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).

First, we concentrate on global diversification in the stock market. Such an analysis might be considered a complement of the influential study of DeMiguel et al. (2009b). We rely on the bootstrap technique developed in 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), its validity 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 inference even when returns exhibit fat tails or 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 which are not seen from an analysis of traditional performance measures alone.

Second, we extend our analysis to the multi-asset class case 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. To this end, we construct value and momentum factors for bonds and commodities building on recent work of Asness et al. (2009). 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.

We find that none of the Markowitz-based portfolio models is able to significantly outperform simple heuristics out-of-sample. This holds for both international equity diversification and for the asset allocation case. Instead almost any well-balanced fixed-weight proportion of stocks, bonds and commodities is able to realize considerable diversification gains. A number of sensitivity checks assures the robustness of our results. We thus suggest a simple and cost-efficient asset allocation approach for private investors.

The remainder of this paper is organized as follows. Section 2 describes our data. Section 3 discusses popular extensions of the Markowitz approach, leading to the selection of promising optimization models for the construction of a "world market portfolio". Subsequently, we derive alternative heuristic asset allocation policies. 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 practicability of our results. We therefore base our study on renowned indices, which are investable for private investors at low costs via exchange-traded funds. We concentrate on Euro zone private investors within a yearly rebalanced buy-and-hold approach.[2] 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, investability, liquidity and data availability.[3]

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 in the "world market portfolio" are represented by the four regional indices MSCI Europe, MSCI North America, MSCI Pacific as well as MSCI Emerging Markets. Taken together, they currently cover 45 countries and track the performance of several thousand stocks. The MSCI indices are designed to cover 85% of the free float-adjusted market capitalization of the respective investable equity universe.

[2] Our motivation to focus on Euro zone investors is twofold. First, many studies takting the special viewpoint of US (stock) investors find the additional benefit from investing abroad to be economically small (e.g. De Roon et al. (2001), Driessen and Laeven (2007)). Non-US perspectives have received far less attention so far. Second, hardly any of the financial products available to Euro zone private investors satisfies our requirements of a transparent, cost-efficient and broadly diversified portfolio. On the one hand, there are passive products which are based on pure stock, bond or commodity indices. Even within the respective asset class, they are often not comprehensively diversified. On the other hand, there are actively managed multi-asset class funds. However, actively managed funds on average underperform passive benchmarks after costs (e.g., Fama and French (2010) and Comer et al. (2009)).

[3] 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, the "world market portfolio" can be constructed with only few highly diversified indices. Moreover, low-cost exchange-traded funds tracking these indices should exist to enable private investors to actually implement our suggestions (investability and liquidity). Finally, we require a long return data history to conduct powerful statistical tests (data availability).

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 Euro zone bonds of different maturities and credit ratings.[4] The index currently tracks the performance of more than 2,200 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 ML European Monetary Union Index.

Partly due to a lack of investability, commodities have long been neglected by private investors. However, many studies provide evidence of the high diversification potential of broad-based commodity futures indices.[5] 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, respectively.

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), Patton (2009) and Phalippou and Gottschalg (2009)). Second, we could not identify indices satisfactorily meeting our selection criteria.

[4]As we aim to derive suggestions for private 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. Our approach is justified by a monthly return correlation of 0.965 between these two indices after 1999.

[5]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 evaluation period starts in February 1973 and ends in December 2009, thus extending previous studies on international diversification in the stock market (e.g., Driessen and Laeven (2007), De Roon et al. (2001) or De Santis and Gerard (1997)). For all indices, we use Euro-denominated total return indices extracted from Thomson Reuters Datastream. Hence, our findings refer to an investment without currency hedging, which is a realistic assumption for private investors.[6]

To implement our 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 as a proxy for the risk-free asset. Historical stock market capitalization data is available from 1973 on, which marks the lower bound of our evaluation period.

2.2 Descriptive Statistics

Table 1 gives an overview of the 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 comparison 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.[7]

Please insert Table 1 here

Table 1 shows only small differences in the average monthly Sharpe ratio of the regional stock indices (0.091) compared to the global stock index (0.098). Over the last 20 years, this difference vanishes completely. This result motivates, first, the analysis of alternative

[6]To convert index levels in Euro we refer to the time series of synthetical Euro/USD exchange rates as calculated by 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.

[7]Driessen and Laeven (2007) emphasize that investment restrictions were imposed on many emerging markets till the mid 80s and that reliable index calculations have only been available since then. Thus, the return of our global stock index can be considered a proxy for the performance of worldwide investable equity.

allocation mechanisms for the stock market and, second, the incorporation of additional asset classes. To assess the diversification potential of a "world market portfolio", figure 1 and figure 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 comovement of international stock markets since the 1980's. However, as Figure 2 illustrates, there is no (in the case of bonds) or at best weak (in the case of commodities) evidence of an increase in correlations across asset classes. Nevertheless, correlations vary considerably through time, which points to potential estimation errors in Markowitz-based optimization methods (see section 3.1). We discuss promising optimization approaches in the next section.

3 Asset Allocation Models

The models considered for portfolio selection in the case of both global stock market diversification and diversification over asset classes are briefly summarized in Table 2. The last column of this Table gives the abbreviation that we use to refer to the model in the results section.

Please insert Table 2 here

3.1 Markowitz-based Optimization Models

We use a variety of model extensions that have been suggested in the existing literature to deal with the well-known problem of estimation error, which is ignored in the traditional mean-variance model of Markowitz (1952).[8] These models either impose additional constraints in the optimization process, shrink the estimated input parameters in order

[8]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.

to mitigate the impact of estimation error, or both. Shortsale 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 toward a common value. In doing so, they aim at reducing the error-maximizing property of the mean-variance model when historical data is used for parameter estimation (e.g., Jorion (1986)). As shown by Jagannathan and Ma (2003), both approaches work similarly by increasing the number of assets with non-negative portfolio weights which enforces a certain extent of diversification.

The first model we implement is the mean-variance framework with non-negativity condition ($maxsr$). 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 ($js - maxsr$), the sample variance-covariance matrix ($ccm - maxsr$), or both ($js - ccm$). 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).[9]

In addition to models which try to maximize the Sharpe ratio, we employ several models which aim at constructing minimum variance portfolios. The superior performance of minimum variance optimization, in particular compared to models that do not ignore information about sample mean returns, has been demonstrated in various studies (see, e.g., Haugen and Baker (1991), Chopra et al. (1993), and Jagannathan and Ma (2003)). We implement the traditional minimum variance approach with and without short-sale constraints ($minvar$, $minvar - nb$), the minimum variance approach with shrinkage estimation of the variance-covariance matrix using the constant correlation model and short-sale restriction ($ccm - minvar$), and a set of extensions to the general minimum variance framework ($nc1v, nc1r, nc2v, nc2r$) which have recently been developed by DeMiguel et al. (2009a). In their empirical analysis, the authors are able to show that this novel class of

[9]The authors provide the code on their web-site (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.

models often outperforms existing 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 methods. First, they choose the parameter δ which minimizes the portfolio variance if the sample is cross-validated. Second, they set δ to maximize the portfolio return in the last period in order to exploit positive autocorrelation in portfolio returns.[10]

Overall, we believe to use a promising set of scientific portfolio choice models against which we test the heuristic construction rules, which are illustrated in the next subsection.

3.2 Heuristic Models

3.2.1 International Stock-Market Diversification

We consider three different weighting schemes for a global stock portfolio: Equal-weighting (1/N heuristic), market value-weighting and GDP-weighting.

An equally-weighted portfolio might be considered a natural benchmark for more sophisticated methods of portfolio optimization. Firstly, it is very easy to implement. And, secondly, private investors have been shown to often rely on this naive 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 private investors. Liquidity and investment capacity arguments are important benefits of these indices, though of minor relevance for our objective. However, an undisputed advantage of this approach is its very low turnover as portfolio weights

[10] 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 very similar to the ones tested in our study and all models achieve very similar results in terms of out-of-sample portfolio variance, Sharpe ratio and turnover.

automatically rebalance when security prices fluctuate.

Nevertheless, concerns against 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, US 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 heavily dropped 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, backtesting shows that fundamentally-weighted country-specific indices have outperformed standard value-weighted indices in the past (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 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 Diversification over asset classes

The easiest asset allocation policy for private 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 Markowitz-based 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 bankers and brokerage firms as reported in e.g. Annaert et al. (2005) and Arshanapalli et al. (2001) as well as institutional holdings as reported in e.g. Blake et al. (1999), Brinson et al. (1986) 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 Annaert et al. (2005) and Arshanapalli et al. (2001) 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 e.g. Erb and Harvey (2006) and Anson (1999), 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 less stocks, less bonds or less stocks and less bonds. Given this arbitrary choice, we use stocks, bonds and commodities in a fixed proportion of 60%, 25% and 15%. Note again that our objective is just to derive a plausible ex-ante strategy as a starting point for the empirical analysis, not an ex-post optimal portfolio.[11]

[11] In fact, we find that our baseline heuristic performs slightly worse than the other two alternatives. Hence, from an

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 2009. Our implementation of the Markowitz-based models relies on a "rolling-window" approach, i.e. we distinguish between estimation and evaluation period. 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, 1st. The one month lag has the aim of ensuring real-time 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 weights adjustment. 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 s equal to 40 basis points per transaction.[12]

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

ex-post perspective, the benchmark against which we test scientific asset allocation models might be regarded conservative.

[12]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.

what extent competing strategies load on the value, size and momentum premium. The Carhart (1997) 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 strategy and the risk-free asset in period t and MKT_t 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, the global factors are market weighted averages of the country-specific components. Our Internet Appendix provides the reader with 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 $Stocks_t$, $Bonds_t$ and $Commodities_t$ 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 Markowitz-based models, it might be regarded as the impact of the models' market timing on the overall portfolio return. For instance, minimum variances approaches are expected to, on average, heavily rely 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 extent 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 Internet Appendix for details) closely follows recent work by Asness et al. (2009), 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 studied exclusively 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, commodities), three momentum factors (for stocks, bonds, commodities) as well as a size factor (for stocks only).[13]

4.2 Baseline Results

4.2.1 International Stock-Market Diversification

We start the empirical analysis with a comparison of the performance of the eleven Markowitz-based models and the various heuristic models for an internationally diversified stock portfolio. Results are reported in Table 3.

Please insert Table 3 here

Columns 2 and 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

[13]Data required for the construction of bond value and bond momentum factors is only available for the second subperiod (1988-2009) 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 (1973-2009). 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.

reduction in risk, as compared to the standard deviation of the three heuristics, seems small. Consequently, full sample after-costs Sharpe ratios tend to be similar for most approaches. The traditional mean-variance model with short-sale restrictions achieves the highest Sharpe ratio (0.124), but this is only marginally higher than the values of the GDP- and naively-weighted portfolio heuristics (0.120 and 0.122, respectively). The value-weighted heuristic performs somewhat worse with a Sharpe ratio of 0.098, 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, but using any of the other heuristics as benchmark does not change the qualitative nature of our results. We find that none of the scientific 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: 0.08).

To explore potential reasons for the widespread lack of statistical significance, we also examine the performance separately for two subperiods. Results are reported in columns 6 and 7 of Table 3. In general, there is no consistency in ranking across subperiods. For instance, the traditional mean-variance model exhibits the highest Sharpe ratio in the second subperiod (1988-2009), but fails to add value over any of the heuristics in the first subperiod. 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. Four models (two minimum variance models and the naively- and GDP-weighted heuristic) exhibit a positive, statistically significant and economically meaningful one factor alpha, but this vanishes once one controls for global momentum, value and size effects. 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 one factor alpha alone. For instance, we find that the GDP-weighted global stock portfolio loads significantly on 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 Markowitz models add value under realistic conditions. It is a natural question to ask whether higher transaction costs prevent the Markowitz models from achieving a better performance, in particular as these models are only optimal under the assumption of no transaction costs. If so, it might still be worthwhile to set up a Markowitz approach to manage an equity portfolio, but to impose certain trading restrictions. As Table 3 shows, the mean turnover of all Markowitz-based 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 Markowitz models is able to significantly outperform any of the heuristics. Nevertheless, assuming higher transaction costs (than 40 bp) and more frequent (than yearly) rebalancing generally works in favor of the heuristic models.

4.2.2 Diversification over asset classes

In the following, we include bonds and commodities in the baseline analysis. Again, we compare the performance of eleven scientific 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 models. The minimum variance approach with short-sale constraints and shrunk covariance matrix ($ccm - minvar$) achieves the highest Sharpe ratio (0.161). In contrast, other Markowitz-based strategies exhibit poor risk-adjusted returns. For instance, the

Sharpe ratio of the traditional mean-variance model with short-sale restrictions ($maxsr$) is only 0.110, which is even lower than in the case of international equity diversification. Hence, not all Markowitz 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 Markowitz models. However, 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 benchmark.

The evidence supports the conclusion that scientific portfolio choice models are not able to outperform a passive benchmark, irrespective of whether we focus on international equity diversification or on diversification over asset classes. However, the heterogeneity in Sharpe ratios among the Markowitz models raises the intriguing possibility that some models are better suited to the asset allocation context than other optimization procedures. To investigate this issue, we implement our three and ten factor regression models. The intuition is to decompose the portfolio weights induced by Markowitz-based approaches in a fixed-weight and a time-varying component. In that sense, Markowitz 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 Markowitz 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. With the exception of one Markowitz model, which has a significant negative three factor alpha, the alphas of all other models are economically close to and statistically not significantly different from zero. Interestingly, the three and ten factor alphas of the fixed-weight heuristics with GDP- and equal-weighting in the stock domain are positive. This result provides further evidence that the value-weighted stock portfolio has not been a particulary successful diversification strategy over the past compared to other potential heuristics.

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 different Markowitz models. One potential concern about this approach may be that the good performance of our baseline heuristic results from backward optimization. To examine whether other possible heuristic strategies perform much worse than our baseline, 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.[14] The stock component of the portfolios is based on the GDP-weighting approach. Figure 5 displays our 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 42% perform better or equal than our baseline heuristic and 58% perform

[14] 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.

worse. Those that perform worse are very often heavily tilted towards only one asset class. If we subdivide the sample period into the subperiods from 1973-1988 and 1988-2009, the resulting figures look very similar. It follows that the 60-25-15 asset allocation policy is only one out of many different fixed-weight asset allocation schemes which achieve a good performance and which are not dominated by sophisticated academic portfolio models. This is good news for private 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.

Illustration of economic significance: Return gap

Since differences in Sharpe ratios are hard 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, $Return\,Gap_t$, in month t is obtained from the following equation:

$$Return\,Gap_t = r_{bm,t} - [\frac{\sigma_{bm}}{\sigma} r_t + (1 - \frac{\sigma_{bm}}{\sigma}) r_{f,t}], \qquad (3)$$

where $r_{f,t}$ 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, our baseline heuristics, 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 naive 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, the GDP-weighted strategy outperforms every stock index as well as bonds and commodities in terms of risk-adjusted return. Including additional asset classes, as implemented in the 60-25-15 portfolio, strengthens these results. The outperformance ranges here from 8.4 to 28.1 basis points per month (or roughly 100 to well more than 300 basis points per year) and thus is economically meaningful. 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

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 G7 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. 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 conclusions drawn in this paper.

Rebalancing frequency

Monthly instead of annual rebalancing does not lead to significantly better results before costs for both the scientific portfolio models and the heuristics. After transactions costs, performance tends to deteriorate for most approaches. In general, the performance drop is more severe for the Markowitz models. This is rooted in their higher turnover in combination with their poor market timing abilities, as analyzed in section 4.2.2. For the heuristics, the rather minor importance of the rebalancing frequency can also be inferred from the insignficant 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 too much tilted towards only one asset. In this regard, the major benefit of portfolio rebalancing is to avoid extreme portfolios consisting of mainly only one asset.

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.

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 IMF. The performance of the GDP-weighting scheme is virtually unchanged in the first check and slightly improves in the second check.

Other performance measures

The recent literature has proposed a number of alternative performance ratios. So, we repeat our analysis utilizing asymmetrical performance measures which 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.[15] The Sortino ratio is computed as

[15] For a detailed description of these ratios, we refer the reader to Biglova et al. (2004) and Rachev et al. (2007). To

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 which best manage to simultaneously deliver high returns and get insurance for high losses. The General Rachev ratio additionaly takes investors degree of risk aversion into account. Utilizing these alternative measures does not change the qualitative nature of our results. A broad spectrum of heuristic portfolio allocation mechanisms still yields similar results as scientific 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

In this study, 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 private 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. To this end, we compare eleven Markowitz-based 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 Markowitz extensions 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 towards a "world market portfolio" 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 naive fixed-weight allocation strategy realizes diversification potential. Third, Markowitz-based optimization methods again do not add substantial value.

implement the ratios, we apply the parametrization described in Biglova et al. (2004) and Farinelli et al. (2008).

Our findings is good news for private investors: Relying on simple rules of thumb in asset allocation significantly improves upon 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, alternatives to the estimation of input parameters from historical data 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 (see 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, 2009, "Value and Momentum Everywhere," *University of Chicago Working Paper*.

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 Naive 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," *forthcoming Journal of Finance*.

Farinelli, S., M. Ferreira, D. Rossello, M. Thoeny, and L. Tibiletti, 2008, "Beyond Sharpe ratio: Optimal asset allocation using different performance ratio," *Journal of Banking and Finance*, 32, 2057–2063.

——— , 2009, "Optimal asset allocation aid system: From one-sizevs 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," 12, 433463.

Goetzmann, W. N., L. Li, and K. G. Rouwenhorst, 2005, "Long-Term Global Market Correlations," *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 trades," *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," Working paper.

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 international 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 rewardrisk 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, 2009, "The Efficiency of Sponsor and Participant Portfolio Choices in 401(K) Plans," Working paper.

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 Naive Diversification Strategies," *Journal of Financial Economics*, forthcoming.

Table 1: Descriptive Statistics for the Different Indices

This table reports the return distribution of the various indices which we consider for portfolio construction. Returns are calculated using Datastream's total return index (code: RI) 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. Return	VaR 95%
Stocks: Regional Indices					
Emerging Markets	88-09	0.122	1.29%	7.45%	-12.13%
Europe	73-09	0.116	1.01%	4.86%	-7.75%
North America	73-09	0.093	0.95%	5.39%	-8.11%
Pacific	73-09	0.065	0.83%	5.91%	-8.61%
Average	**73-09**	**0.091**	**0.93%**	**5.39%**	**-8.16%**
Average	**88-09**	**0.072**	**0.82%**	**5.88%**	**-9.57%**
Global Stock Index	**73-09**	**0.098**	**0.92%**	**4.79%**	**-8.44%**
Global Stock Index	**88-09**	**0.060**	**0.68%**	**4.92%**	**-8.73%**
Other Asset Classes					
Bonds	73-09	0.108	0.57%	1.12%	-1.27%
Commodities	73-09	0.076	0.92%	6.28%	-9.65%

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) which 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: Markowitz-based portfolio optimization models from the existing literature	
1	Maximum Sharpe ratio approach with shortsale constraints	maxsr
2	Minimum variance approach without shortsale constraints	minvar-nb
3	Minimum variance approach with shortsale constraints	minvar
4	James/Stein estimator of expected returns with shortsale constraints	js
5	James/Stein estimator of expected returns plus Ledoit/Wolf constant correlation model with shortsale constraints	js-ccm
6	Maximum Sharpe ratio approach plus Ledoit/Wolf constant correlation model with shortsale constraints	ccm-maxsr
7	Minimum variance approach plus Ledoit/Wolf constant correlation model with shortsale constraints	ccm-minvar
8	1-norm constrained minimum variance portfolio with δ calibrated using cross-validation over portfolio variance	nc1v
9	1-norm constrained minimum variance portfolio with δ calibrated by maximizing portfolio return in previous period	nc1r
10	2-norm constrained minimum variance portfolio with δ calibrated using cross-validation over portfolio variance	nc2v
11	2-norm constrained minimum variance portfolio with δ calibrated by maximizing portfolio return in previous period	nc2r
	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	naiv
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; naiv

Table 3: Markowitz vs. Heuristics: International Stock Market Diversification Results

This table reports means, standard deviations, Sharpe ratios and alphas of monthly out-of-sample returns after costs as well as average turnover for the international equity portfolios which are constructed using the various Markowitz-based optimization models and heuristic models. Sharpe ratios are reported for the total sample period (1973-2009) and two sub-sample periods (February 1973-January 1988 and February 1988-December 2009). P-values that the Sharpe ratio for each of these models is different from that for the GDP-weighted stock portfolio, our baseline heuristic, are calculated using the bootstrap technique developed in Ledoit and Wolf (2008). We assume a bid-ask spread of 40 basis points 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, ** indicates significance at the 5% level and *** indicates significance at the 1% level. See section 3 and table 3 for a description of the models. Details on the construction of the factors used in the regression framework are given in the Internet Appendix.

Portfolio Model	Mean Return	Std. Dev. Return	Mean Annual Turnover	Sharpe Ratio 1973-2009	Sharpe Ratio 1973-1988	Sharpe Ratio 1988-2009	p-value $H_0 : SR = SR_{gdp}$	α_{1F}	t-stat α_{1F}	α_{4F}	t-stat α_{4F}
Panel A: Markowitz-based Optimization Models											
maxsr	1.18%	5.87%	57.69%	0.124	0.121	0.126	0.87	0.23%	1.61	0.23%	1.48
minvar-nb	0.92%	4.56%	31.29%	0.102	0.166	0.060	0.51	0.06%	0.64	-0.04%	-0.46
minvar	1.00%	4.62%	23.41%	0.118	0.166	0.087	0.93	0.11%	1.73*	0.03%	0.48
js	1.01%	5.15%	75.44%	0.108	0.121	0.099	0.59	0.11%	0.94	0.11%	0.93
js-ccm	0.99%	5.15%	70.64%	0.105	0.125	0.093	0.53	0.10%	0.84	0.13%	1.05
ccm-maxsr	1.08%	5.64%	58.97%	0.112	0.125	0.106	0.77	0.15%	1.13	0.18%	1.23
ccm-minvar	0.98%	4.62%	19.19%	0.114	0.158	0.084	0.73	0.09%	1.40	0.04%	0.58
nc1v	0.93%	4.57%	31.42%	0.105	0.166	0.065	0.58	0.07%	0.79	-0.03%	-0.29
nc1r	1.00%	4.59%	27.32%	0.119	0.166	0.088	0.96	0.12%	1.66*	0.03%	0.35
nc2v	0.94%	4.58%	28.28%	0.107	0.166	0.068	0.55	0.07%	0.93	-0.02%	-0.33
nc2r	0.98%	4.60%	25.94%	0.115	0.166	0.082	0.81	0.10%	1.50	0.01%	0.17
Panel B: Heuristic Models											
gdp	1.03%	4.84%	11.08%	0.120	0.155	0.098	.	0.12%	2.36**	0.08%	1.54
macap	0.92%	4.79%	4.80%	0.098	0.157	0.060	0.08	0.00%	0.00	0.00%	0.00
naiv	1.04%	4.83%	12.85%	0.122	0.176	0.090	0.77	0.13%	2.42**	0.08%	1.42

Table 4: Markowitz vs. Heuristics: Asset Allocation Results

This table reports means, standard deviations and Sharpe ratios of monthly out-of-sample returns after costs as well as average turnover for the asset allocation portfolios which are constructed using the various Markowitz-based optimization models and heuristic models. Sharpe ratios are reported for the total sample period (1973-2009) and two sub-sample periods (February 1973-January 1988 and February 1988-December 2009). P-values that the Sharpe ratio for each of these models is different from that for the 60-25-15 asset allocation portfolio with GDP-weighting in the stock market, our baseline heuristic, are calculated using the bootstrap technique developed in Ledoit and Wolf (2008). We assume a bid-ask spread of 40 basis points 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 is the intercept from a ten factor model, augmented by value, size and momentum factors. Data required for the construction of bond value and bond momentum factors is only available for the second subperiod (1988-2009) of our analysis. For the sake of comparability, we thus report results from both our three and ten factor regression only for this period. For the t-statistics, * indicates significance at the 10% level, ** indicates significance at the 5% level and *** indicates significance at the 1% level. See section 3 and table 3 for a description of the models. Details on the construction of the factors used in the regression framework are given in the Internet Appendix.

Portfolio Model	Mean Return	Std. Dev. Return	Mean Annual Turnover	Sharpe Ratio 1973-2009	Sharpe Ratio 1973-1988	Sharpe Ratio 1988-2009	p-value $H_0 : SR = SR_{gdp}$	α_{3F}	t-stat α_{3F}	α_{10F}	t-stat α_{10F}
Panel A: Markowitz-based Optimization Models											
maxsr	0.86%	3.74%	49.01%	0.110	0.179	0.052	0.54	-0.09%	-0.64	-0.16%	-1.05
minvar-nb	0.59%	1.13%	13.23%	0.128	0.167	0.097	0.86	-0.02%	-0.75	-0.03%	-1.07
minvar	0.63%	1.11%	7.29%	0.160	0.185	0.139	0.79	0.01%	0.43	-0.01%	-0.32
js	0.75%	2.79%	45.29%	0.109	0.179	0.044	0.55	-0.09%	-0.86	-0.10%	-0.95
js-ccm	0.83%	3.03%	44.91%	0.126	0.186	0.080	0.76	-0.02%	-0.12	-0.02%	-0.15
ccm-maxsr	0.73%	3.70%	53.11%	0.077	0.180	-0.005	0.21	-0.28%	-2.06**	-0.35%	-2.41**
ccm-minvar	0.63%	1.13%	6.11%	0.161	0.182	0.144	0.77	0.00%	0.44	-0.01%	-1.01
nc1v	0.59%	1.13%	13.42%	0.130	0.168	0.099	0.88	-0.02%	-0.67	-0.03%	-1.02
nc1r	0.62%	1.11%	10.07%	0.150	0.163	0.141	0.90	0.01%	0.54	0.00%	-0.26
nc2v	0.61%	1.12%	9.50%	0.142	0.185	0.107	0.98	-0.02%	-0.83	-0.03%	-1.29
nc2r	0.63%	1.11%	7.29%	0.160	0.185	0.139	0.78	0.01%	0.43	-0.01%	-0.32
Panel B: Heuristic Models											
60-25-15; gdp	0.92%	3.29%	12.95%	0.141	0.184	0.115	.	0.12%	2.86***	0.08%	1.79*
60-25-15; macap	0.85%	3.24%	10.25%	0.123	0.185	0.083	0.11	0.02%	1.18	0.00%	-0.23
60-25-15; naiv	0.92%	3.28%	13.61%	0.143	0.205	0.107	0.67	0.10%	2.45**	0.05%	1.18

Table 5: Return Gaps of Various Indices compared to the GDP-weighted stock portfolio and the 60-25-15 asset allocation portfolio

This table reports the Sharpe ratio and Value-at-Risk at the 95% confidence level of monthly returns for various indices as well as the GDP-weighted stock portfolio and the 60-25-15 asset allocation portfolio with GDP-weighting in the stock market, which are our baseline heuristic models for portfolio construction. Moreover, the table presents the Return Gap of these indices in basis points (bp) per month compared to our baseline heuristics. Portfolio weights are readjusted every February each year. See section 3 and table 3 for a description of the models and subsection 4.1 for a description of the computation of the Return Gap.

Asset Class/ Region	Sample Period	Sharpe Ratio	VaR 95%	Return Gap (bp per month) GDP-stock portfolio	Return Gap (bp per month) 60-25-15 portfolio
			Panel A: Stock Indices		
MSCI Germany	73-09	0.101	-8.95%	7.8	12.4
MSCI France	73-09	0.104	-9.35%	6.2	11.3
MSCI Italy	73-09	0.064	-10.37%	28.1	26.2
MSCI United Kingdom	73-09	0.096	-9.02%	12.5	15.6
MSCI United States	73-09	0.089	-8.30%	15.4	17.6
MSCI Canada	73-09	0.087	-8.65%	16.0	18.0
MSCI Japan	73-09	0.055	-9.04%	30.9	28.1
MSCI Europe	73-09	0.116	-7.75%	1.8	8.4
MSCI North America	73-09	0.093	-8.11%	13.3	16.2
MSCI Pacific	73-09	0.065	-8.61%	26.0	24.8
MSCI Emerging Markets	88-09	0.122	-12.13%	-12.0	-2.2
			Panel B: Asset Classes		
GDP-stock portfolio	73-09	0.120	-7.90%	.	7.1
Bonds	73-09	0.108	-1.27%	5.7	11.0
Commodities	73-09	0.076	-9.65%	20.7	21.2

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.

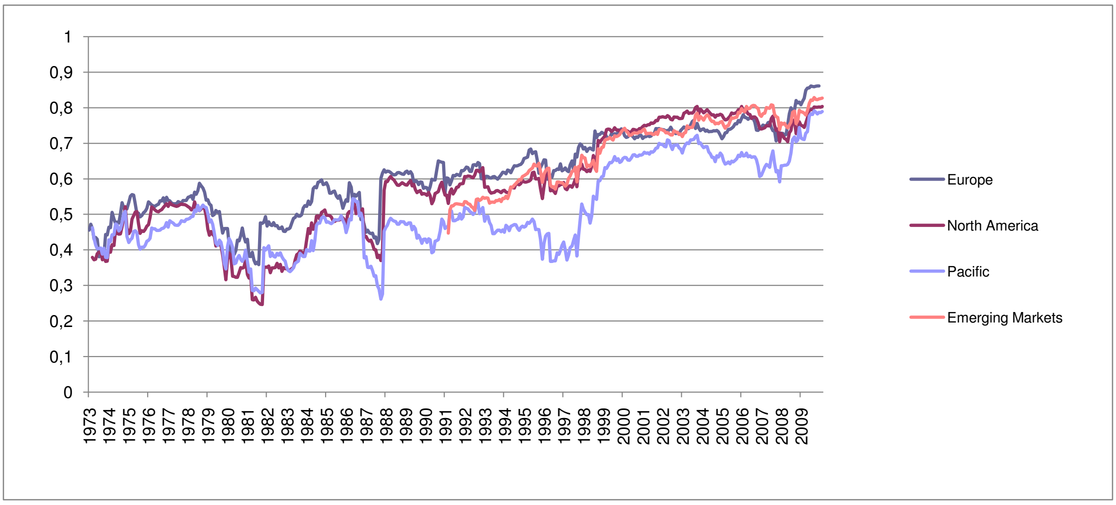

Figure 2: Time-Series Behavior of Correlations between Asset Classes

This figure depicts the movement in the average correlation over the sample period for the iBoxx Euro Overall Index and the S&P GSCI Commodity Total Return Index with respect to the regional MSCI stock indices. Correlation coefficients are computed using a rolling window approach based on the previous 60 months.

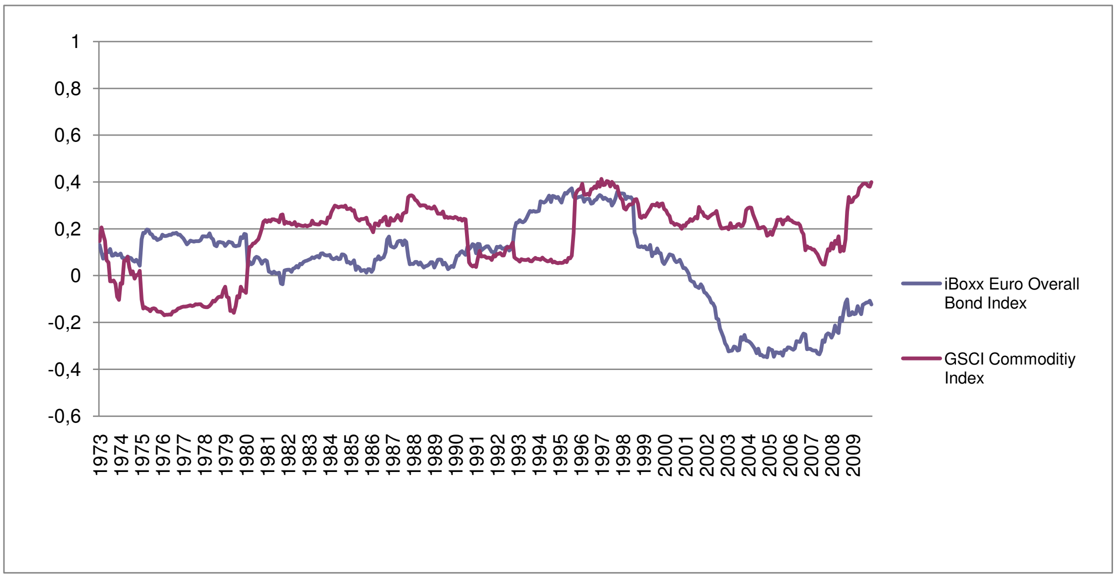

Figure 3: Time-Series Evolution of Portfolio Weights of a Cap-weighted Stock Index

This figure depicts the portfolio weights of a market-value weighted stock index constructed from the MSCI indices for North America, Europe, the Pacific region and the Emerging Markets over the sample period. The data source is Thomson Reuters Datastream.

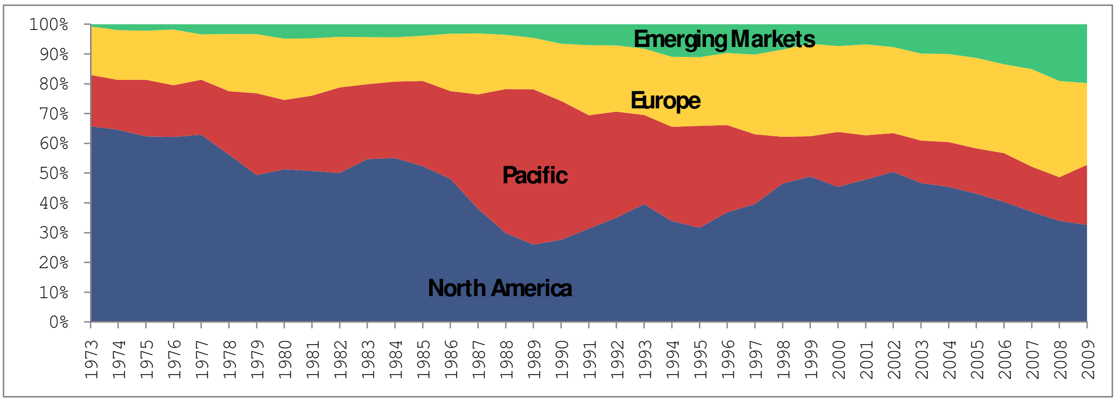

Figure 4: Time-Series Evolution of Portfolio Weights of a GDP-weighted Stock Index

This figure depicts the portfolio weights of a GDP-weighted stock index constructed from the MSCI indices for North America, Europe, the Pacific region and the Emerging Markets over the sample period. Data sources are the World Bank for the period 1973-2005 and the International Monetary Fund for the period 2006-2008.

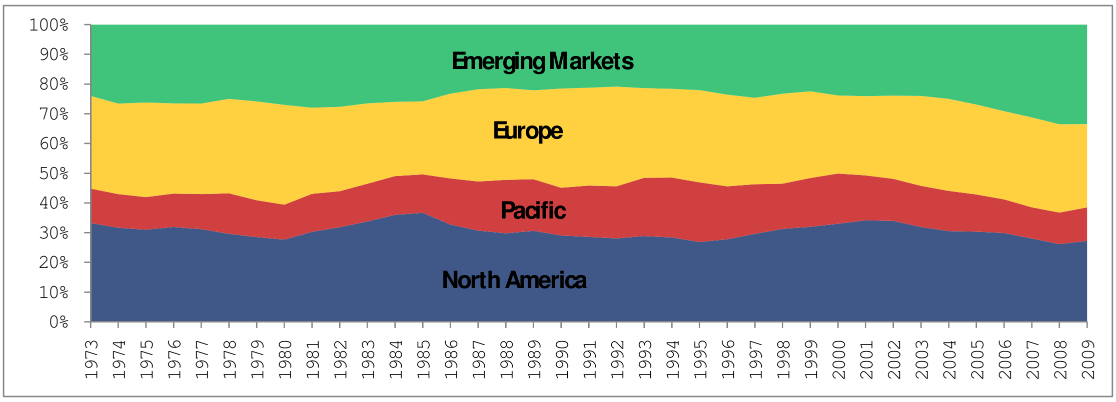

Figure 5: Graphical Presentation of the Performance of Alternative Fixed-Weight Asset Allocation Strategies

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 2 classes appropriately. This approach yields 5,151 different portfolios. The stock component of the portfolios comprises the four regional MSCI indices and is GDP-weighted.