Minimum-Variance Stock Picking – A Shift in Preferences for Minimum-Variance Portfolio Constituents

Thomas Dangl Michael Kashofer

first version: July 28, 2013
this version: August 25, 2013

Abstract

Minimum-variance (equity) portfolio selection is increasingly popular among investors. We study a broad set of 63 different, commonly used approaches to build long-only minimum-variance portfolios among US as well as European stocks. We focus on the stock picking characteristics of minimum-variance approaches and find a high degree of consensus across different methods regarding the selection of particular stocks. Hence, the increasing demand for low-volatility equity strategies translates into increasing demand for a rather small subset of stocks, independent of the particular portfolio approach one employs. Further, we analyze the price-to-book multiples of minimum-variance portfolio constituents over the past 23 years (for S&P 500 stocks, and 13 years for STOXX Europe 600 stocks, respectively) and report a gradual shift in preference for these stocks over time. While in the 1990s, minimum-variance constituents traded at a discount (below-average price-to-book), today they trade at a considerable premium, up to 23 % in the US (S&P 500) and up to 40 % in Europe (STOXX Europe 600).

Thomas Dangl and Michael Kashofer are from Vienna University of Technology. Email addresses: thomas.dangl@tuwien.ac.at and michael.kashofer@tuwien.ac.at. Michael Kashofer acknowledges financial support from the Spängler IQAM Research Center.

1 Introduction

Minimum-variance (equity) portfolio selection is increasingly popular among investors.[1] While all kinds of equity strategies saw huge draw downs in the turmoil of 2007-09 and again in 2011, minimum-variance strategies did considerably well and navigated very stable through these crises. Their losses were only a fraction of the losses of pure index investments and they delivered returns over this period that outperformed common benchmarks on a risk-adjusted base and – even more impressive – in absolute terms. For today's risk sensitive investors, minimum-variance equity strategies seem to be a perfect match – providing an exposure to equities as an asset class while preventing extreme tail events.

Advocates of a minimum-variance strategy refer to the so called low-beta and low-volatility anomalies to explain its historical risk adjusted performance. These anomalies denote the empirical finding that stocks with low-volatility (systematic and/or idiosyncratic) earn a higher premium per unit of volatility than high-volatility stocks. Black (1972) and Haugen and Heins (1975) are early papers that document evidence that the relationship between risk and return is much flatter than predicted by the CAPM, or that it is actually inverted as found by Haugen and Heins (1975). Baker, Bradley, and Wurgler (2011) and Frazzini and Pedersen (2013) attest that 40 years later the low-beta anomaly is still present and can profitably be exploited by "betting against beta". The low-volatility anomaly in the cross-section of stock returns is documented by Ang, Hodrick, Xing, and Zhang (2006, 2009) who show that investors get systematically compensated for bearing idiosyncratic risk in US as well as international stock markets.[2]

We take the growing popularity of minimum-variance equity strategies as a motivation to examine the changing valuation of those portfolios over time. More precisely, since

[1] Blackrock Inc. (2013) reports that the global universe of minimum-volatility exchange traded products (ETPs) experienced net asset inflows of more than 1.6 billion USD per month over the first four months in 2013, more than three times the average monthly inflow during 2012.

[2] Black (1972) explains the low-beta anomaly with restricted access to borrowing, a rationale used also in Frazzini and Pedersen (2013). Baker, Bradley, and Wurgler (2011) argue that benchmarking of institutional portfolios impose limits to arbitrage that allow the low-beta, low-volatility anomalies to sustain. While the low-beta anomaly is broadly accepted in the finance community (by academic researchers as well as practitioners), the low-volatility anomaly is still controversially discussed. Bali and Cakici (2008) are skeptical and report, e.g., that data frequency and horizon have a significant influence on whether a low-idiosyncratic volatility anomaly is detected or not. Fu (2009) and Huang, Liu, Rhee, and Zhang (2010) show that after controlling for short-term return reversal and for the time variation in idiosyncratic volatility, the low-volatility anomaly is reversed, i.e., high idiosyncratic volatility is associated with high expected returns, consistent with the effects of imperfect portfolio diversification.

minimum-variance strategies are rather concentrated, we focus on those stocks that tend to be picked by minimum-variance strategies and analyze their price-to-book ratio relative to those which are not candidates for such portfolios. For US stocks as well as European stocks[3] we find – after controlling for various valuation relevant characteristics – that valuation levels of equities which qualify for the inclusion in minimum-variance portfolios have significantly increased over time. Implementing a broad set of different minimum-variance strategies, we find that stocks which have a high propensity of being included in a minimum-variance portfolio show a significantly positive time trend in their price-to-book ratio compared to those stocks which are not candidates for an inclusion. The increase is not only statistically significant but its magnitude is also relevant when assessing future return potential of minimum-variance portfolios. On average, we estimate that today core minimum-variance stocks (those simultaneously selected by many different minimum-variance portfolio approaches) trade at a premium of up to 23 % in the US (S&P 500) and up to 36 % in Europe (STOXX Europe 600).

Since minimum-variance strategies do not include return forecasts in their objective function and use the portfolios' variance as a risk measure, different strategies may only differ with respect to three main model specifications, (i) the estimated diversification potential (the estimate of the covariance matrix), (ii) constraints on portfolio weights, (iii) their portfolio rebalancing policy. The most relevant among these factors is, by far, the estimate of the covariance matrix that is used to optimize risk diversification. Thus, we implement a broad set of different concepts to estimate sample covariance matrices and to establish its positive definiteness.[4] Regarding weight constraints we focus entirely on long only portfolios and use maximum constraints to prevent the strategies from being extremely concentrated.

[3]For US stocks we use the S&P 500 as our universe, for European Stocks the Dow Jownes STOXX Europe 600. For a description of the data set see Section 5.1.

[4]In total we use 21 different methods to determine the variance-covariance structure of individual returns. We compute returns over one and four weeks, respectively, and use horizons (moving window) of one year, two years as well as five years. We apply seven different approaches proposed in the relevant literature to determine estimates of the variance-covariance structure of equity returns (a single-factor market model, shrinkage to the single-factor market model, an average correlation model, shrinkage to average correlation model, shrinkage to average variance model, a three-factor principal component model, a five-factor principal component model, see Table 1) and apply these approaches to three different selections of data frequency / data horizon (weekly returns on a one-year rolling window, weekly returns on a two-years rolling window and returns over four weeks on a five-years rolling window).

Despite the large variety of approaches to set up minimum-variance strategies, there is a strikingly strong agreement across strategies about the attractiveness of certain stocks to be included in a portfolio.

The paper is composed as follows. In Section 2 we introduce the generic framework of a minimum-variance strategy. In Section 3 we discusses the setup of our empirical study to detect valuation levels of minimum-variance portfolio constituents. Section 4 discusses different approaches to estimate the covariance matrix. Section 5 presents the empirical analysis and its results and in Section 6 we present our main result, a shift in risk preferences over the last decades in favor of minimum-variance portfolio constituents. Section 7 discusses robustness checks performed and, finally, Section 8 concludes the study.

2 Minimum-Variance Equity Strategies

Minimum-variance strategies exist in various flavors but they all have the common goal of optimally reducing a portfolio's return variance through portfolio diversification, thereby ignoring any other stock characteristics, most prominently, stock's return expectation. In the light of more than fifty years of portfolio theory it might be surprising that portfolio rules which ignore the relationship between systematic risk and return expectations attract so much interest. Well documented return anomalies like the low-beta and low-volatility anomalies, however, give a rationale for focusing on risk diversification only, see Section 1 for a brief discussion of the empirical evidence on these anomalies.

Forming a minimum-variance portfolio in a given universe of stocks corresponds to solving the following generic problem

$$\begin{aligned} w^* \quad &\in \quad \operatorname{argmin}_w \left\{ \tfrac{1}{2} w' \Sigma w \right\}, \\ &\quad \text{subject to:} \quad w_{\min} \leq w_i \leq w_{\max}, \end{aligned} \tag{1}$$

where $w = (w_1, w_2, \ldots, w_N)'$ is the vector of portfolio weights, Σ denotes the covariance matrix and $w_{\min}, w_{\max}$ are constraints on portfolio weights. We, thus, see that in general different minimum-variance strategies might differ only in

(a) the way they estimate the covariance matrix Σ,

(b) the constraints $w_{\min}, w_{\max}$ they impose on portfolio weights,

(c) their rebalancing policy, i.e., the frequency on which optimization is performed and how the actually invested portfolio is rebalanced towards the new optimum.[5]

While items (a) and (b) determine the target portfolio of a strategy, item (c) characterizes how to approach this target and is mainly driven by transaction cost considerations. Our goal is to focus on the target. We implement a broad variety of different minimum-variance approaches to determine whether a stock is a candidate for inclusion in a minimum-variance portfolio (min-var candidate) at a certain point in time. Then, in the second step, we analyze the valuation of those candidates in search for significant differences between min-var candidates and stocks that are found not to qualify for an inclusion a minimum-variance portfolio. Consequently, we focus entirely on items (a) and (b) in the list above. Rebalancing policies must, apparently, be regarded if one analyzes the performance of such portfolios, but this is not the scope of our work.

In Sections 4 and 5 we will introduce 63 different implementations of minimum-variance approaches that are commonly used in the industry and in academic research as well as ten aggregated target selections that focus on stocks which are picked by several minimum-variance methods simultaneously. These methods of portfolio selection are then applied to US equities (S&P 500) and European equities (STOXX Europe 600).

3 Valuation Levels of Minimum-Variance Portfolio Constituents

Consider a particular approach, A, to determine a long-only minimum-variance portfolio from a universe of stocks at time t. Let $\mathbf{1}_{i,t}^{A}$ denote a min-var indicator that equals 1 if approach A selects stock i into the minimum-variance portfolio at time t and 0 otherwise. For a simpler interpretation of the following regression results, we use the cross-sectionally de-meaned counterpart of $\mathbf{1}_{i,t}^{A}$ and call it $d_{i,t}^{A}$. Thus, all min-var constituents share a common

[5] There might exist further constraints on portfolios which shall not be explicitly regarded in this study, e.g., sector constraints that limit a portfolio's exposure to certain industry sectors, etc. We do, however, control for sector weights and other characteristics of the portfolio in our empirical analysis.

value of the indicator $d^A_{\cdot,t} = d^A_{\text{minvar},t}$ and all non-constituents share a common value of the indictor that satisfies $d^A_{\cdot,t} = d^A_{\neg\text{minvar},t} = d^A_{\text{minvar},t} - 1$.

To identify the valuation impact of such a selection, we perform a cross sectional regression of the normalized price-to-book ratio of all time t constituents of the universe on the min-var indicator, plus further variables that control for known valuation relevant characteristics of these stocks.[6]

$$\text{PtBn}_{i,t} = \gamma^A_{\text{minvar},t} d^A_{i,t} + \sum_j \gamma^A_{j,t} c_{j,i,t} + \varepsilon^A_{i,t}, \tag{2}$$

where $\text{PtBn}_{i,t}$ is the price-to-book ratio of stock i at time t, cross-sectionally normalized to a mean of zero and a variance of 1. Variables $c_{j,i,t}$ are de-meaned dummies that control for various characteristics of stock i at time t. We use

- industry indicators according to the ICB industry classification (ten industries),
- size indicators, cross-sectionally sorting stocks into three equally sized groups with respect to their market capitalization,
- dividend indicators, cross-sectionally sorting stocks into three equally sized groups with respect to their dividend-to-book ratio.[7]

The cross sectional normalization of the stocks' price-to-book corrects for different overall valuation levels of the stock market that changes over time. Furthermore, it also corrects for time varying dispersion of the price-to-book across stocks. Thus, the coefficient $\gamma^A_{\text{minvar},t}$ of the min-var indicator picks an eventual valuation anomaly of minimum-variance portfolio constituents that can not be explained by other control characteristics. Its magnitude is in multiples of the cross sectional standard deviation of price-to-book of all index constituents at a given time t. If, e.g., a portfolio approach tends to overweight stocks from a certain industry that shows on average high (or low) price-to-book multiples, the coefficient of the respective industry indicator will compensate for this deviation. Thus, if two stocks share

[6] Since we use de-meaned left-hand-side and right-hand-side variables, the inclusion of a regression constant is not necessary.

[7] We use the dividend-to-book ratio instead of dividend-price ratio or dividend yield because we do not want to have pricing multiples (i.e., ratios that include the market price) on the right-hand-side of our regression.

their industry affiliation, have comparable market valuation and dividend payout (according to our sorts into three subgroups), but only one is selected as a minimum-variance portfolio constituent by approach A, then $\gamma^A_{\text{minvar},t}$ states the expected difference in their price-to-book ratio given in multiples of the standard deviation of all stocks in the universe.

The conditional estimate of a stock's normalized price-to-book, given that the stock is a min-var constituent of approach A at time t (and all other characteristics are at the average level), is then

$$\hat{\text{PtBn}}_t \mid \text{minvar} = \hat{\gamma}^A_{\text{minvar},t} d^A_{\text{minvar},t}. \tag{3}$$

Since these estimates of normalized price-to-book are noisy over time, mainly driven by stock price fluctuations, our analysis follows the estimated coefficients $\hat{\gamma}^A_{\text{minvar},t}$ over time and tries to understand time trends in the valuation levels of minimum-variance constituents. As we will document, these time trends are significant and can be detected in US stocks' as well as in European stocks' valuations independent of the portfolio approach A.

Thus, to understand the linear time trend in the min-var valuation coefficient for a given stock universe and a particular approach A we estimate a GLS model of the form

$$\hat{\gamma}^A_{\text{minvar},t} = a^A + b^A t + \nu^A_t, \tag{4}$$

where the variance of ν^A_t is assumed to be proportional to the squared standard error of $\hat{\gamma}^A_{\text{minvar},t}$ in regression (2). Furthermore, standard errors in a^A and b^A have to be corrected for autocorrelation in the noise term, ν^A.

Our time series estimate of the min-var valuation coefficient is then given by

$$\hat{\hat{\gamma}}^A_{\text{minvar},t} = (\hat{a}^A, \hat{b}^A) \begin{pmatrix} 1 \\ t \end{pmatrix}, \tag{5}$$

with a standard error of

$$\text{se}(\hat{\hat{\gamma}}^A_{\text{minvar},t}) = \left[(1,t) R^A \begin{pmatrix} 1 \\ t \end{pmatrix} \right]^{1/2}, \tag{6}$$

where R^A denotes the autocorrelation-corrected covariance matrix of the estimated coefficients $\hat{a}^A$ and $\hat{b}^A$.

Finally, the estimate of normalized price-to-book, conditional on the fact that a stock is picked by method A, along the estimated time trend is analogous to (3) determined by

$$\hat{\hat{\text{PtBn}}}_t \mid \text{minvar} = \hat{\hat{\gamma}}^A_{\text{minvar},t}\, d^A_{\text{minvar},t}. \tag{7}$$

4 Different Approaches to Estimate the Covariance Matrix

The covariance matrix of (simple) returns determines the diversification potential within a universe of stocks. In determining the sample covariance matrix $\hat{\Sigma}_S$ from observed stock prices (plus dividend payments) at time t there are two basic choices to be made: (i) the increment Δt over which returns are computed (e.g., weekly returns, monthly returns, quarterly returns) and (ii) the history of returns that is regarded in the computation. A particular choice provides a $(N \times K_t)$ Matrix R_t that contains the returns of the K_t constituents of the stock universe over N historical periods, i.e., over the horizon from $t-(N-1)\Delta t$ to t.[8]

The sample covariance matrix $\hat{\Sigma}_S$ is then determined by

$$\hat{\Sigma}_S = \frac{1}{N-1} R'\, M\, R, \tag{8}$$

where the symmetric and idempotent matrix M is the residual maker with respect to a regression onto a constant,

$$M = \mathbb{I} - \mathbf{1}\,(\mathbf{1}'\,\mathbf{1})^{-1}\,\mathbf{1}',$$

with $\mathbb{I}$ the $(N \times N)$ identity matrix and $\mathbf{1}$ a column vector containing N times the constant 1.

When dealing with a broad universe of stocks it is more often than not the case that $N < K$.[9] In such a case, however, the sample covariance matrix $\hat{\Sigma}_S$ is (i) subject to large

[8] All these estimators together with the set of constituent stocks change over time, for brevity, however, we omit the time t subscript from now on.

[9] Consider the S&P 500, then it is common practice to estimate the covariance structure of stock returns from two years of weekly returns. The argument for a restriction of the history to two years is, e.g., an

estimation errors and (ii) it is singular by construction. The weak determination of $\hat{\Sigma}_S$ is evident when recognizing that one has to estimate $K(K+1)/2$ elements of $\hat{\Sigma}_S$ from only $N \cdot K$ observations, see Jobson and Korkie (1980) for a detailed study of the small sample properties of $\hat{\Sigma}_S$. Singularity follows immediately from (8) which implies that the rank of $\hat{\Sigma}_S$ is bounded from above by $\min\{K, N-1\}$.[10]

Thus, the specification of time increment and horizon alone is generally not sufficient to receive a covariance matrix that can be used as an input to the generic minimum-variance problem (1). We need to put additional structure onto the estimation procedure in order to get a properly defined positive definite covariance matrix $\hat{\Sigma}$ to determine a minimum variance portfolio. These structural estimators should not only generate a positive definite covariance matrix but should also deal adequately with estimation errors. I.e., the choice about the particular structural approach constitutes a further degree of freedom in the estimation process.

In the industry as well as in academic research, a set of estimation methods have been established that either put some theory-founded structure onto the estimation, or some simple parametrization that is intended to reduce estimation errors, or both.

Before discussing the different estimation methods we implemented to determine whether a stock is a min-var candidate or not, we briefly introduce the idea of shrinkage estimators.

4.1 Shrinkage Approaches

Structural estimations of the covariance matrix apply very rigid assumptions on the form of the covariance matrix. So valuable information that is contained in the sample covariance matrix is ignored.[11] The idea behind shrinkage is to combine both the sample covariance

implicit assumption of time-variation in the covariance structure which the estimate is able to capture if one restricts the used history.

[10] The residual maker M has at most rank $N-1$ since it generates residuals from a projection onto a one-dimensional subspace of $\mathbb{R}^N$. Since R has at most rank K, we have

$$\mathrm{rank}(\hat{\Sigma}_S) \leq \min\{K, N-1\}.$$

E.g., the sample covariance matrix estimated from two years of weekly returns of the 500 constituents of the S&P 500 at a certain time (104 observations per stock) has at most rank 103. Hence, it is not positive definite and not invertible, because at least 397 of its 500 eigenvalues are exactly equal 0.

[11] Consider the simplest structural estimator, an "average variance" estimator which assumes that the covariance matrix is the identity matrix multiplied by the average variance of stock returns in the sample. This approach reduces the number of coefficients to be estimated from $K(K+1)/2$ to 1 and, hence, it is

matrix $\hat{\Sigma}_S$ and the covariance matrix from a particular structural estimation $\hat{\Sigma}_A$ in a way that allows to control how much structure is imposed onto the estimation. This is usually done in a linear way specified by the weight $\delta \in (0, 1]$ such that

$$\hat{\Sigma} = \delta \hat{\Sigma}_A + (1 - \delta) \hat{\Sigma}_S.$$

This approach is also referred to as shrinking the sample covariance matrix $\hat{\Sigma}_S$ towards a shrinkage target $\hat{\Sigma}_A$.[12] Estimators of this kind are well known in Bayesian statistics dating back to the paradox described in Stein (1956). In the context of portfolio management Jorion (1986) introduced shrinkage to improve estimators of expected asset returns. The choice of the weight δ is always a tradeoff between information contained in the sample estimate and robustness against estimation errors provided by the shrinkage target. Frost and Savarino (1986) used a Bayesian criterion to determine an optimal choice of δ in estimating covariance matrices and recently Ledoit and Wolf (2003, 2004b) determined a more general Byesian framework to optimize the shrinking intensity. We adopt their method in three of our minimum-variance approaches which use the structural estimator as a prior and sample covariance matrix as the signal (thereby taking care of the correlation of these two inputs that are estimated from the same set of observations). Interestingly, Jagannathan and Ma (2003) show that applying constraints on portfolio weights in mean-variance optimization is equivalent to shrinking the covariance matrix. Hence, frequently reported improvements in the out-of-sample characteristics of constrained portfolios relative to unconstrained portfolios can be interpreted as implicit reduction in the estimation error in the covariance matrix.

less prone to estimation errors. This comes at the cost of completely ignoring information about individual riskiness of stocks and the diversification potential that is contained in the data.

[12]Shrinkage of a positive semidefinite matrix $\hat{\Sigma}_S$ towards a positive definite target will result in a positive definite matrix for shrinkage intensities $\delta \in (0, 1]$. Thus, selecting a proper shrinkage target results in a non-singular estimate of $\hat{\Sigma}$. In our implementations, we confine the shrinkage intensity δ not to exceed one as is done in Ledoit and Wolf (2004a) and further apply a lower bound of 0.1 to guarantee well-conditioned covariance estimators.

4.2 Single-Factor Market Model Estimator (MM)

This approach is inspired by the theory of a single-factor market model as represented by the CAPM (see Sharpe (1964) or Treynor (1962)) and is covered by any textbook in corporate finance or portfolio management. The structure that is put on the covariance matrix assumes that pairwise covariance of stocks is solely determined by a single market factor. I.e., the estimated covariance matrix has the form

$$\begin{aligned}\hat{\Sigma}_{\mathrm{MM}} &= \hat{\beta}\hat{\beta}'\sigma^2_{\mathrm{M}} + \hat{\Sigma}_{\mathrm{I,M}} \\ &= \begin{pmatrix} \hat{\beta}_1^2 & \hat{\beta}_1\hat{\beta}_2 & \dots & \hat{\beta}_1\hat{\beta}_N \\ \hat{\beta}_2\hat{\beta}_1 & \hat{\beta}_2^2 & \dots & \hat{\beta}_2\hat{\beta}_N \\ \vdots & & \ddots & \vdots \\ \hat{\beta}_N\hat{\beta}_1 & \beta_N\hat{\beta}_2 & \dots & \hat{\beta}_N^2 \end{pmatrix} \hat{\sigma}^2_{\mathrm{M}} + \begin{pmatrix} \hat{\sigma}^2_{\mathrm{I},1} & 0 & \dots & 0 \\ 0 & \hat{\sigma}^2_{\mathrm{I},2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\sigma}^2_{\mathrm{I},N} \end{pmatrix}\end{aligned} \tag{9}$$

with $\hat{\beta}$ the column vector containing estimated market betas $\hat{\beta}_i$, $i = 1, \dots N$, $\hat{\sigma}^2_{\mathrm{M}}$ is the sample estimate of the market index variance, and $\hat{\sigma}^2_{\mathrm{I},i}$, are the sample estimates of the idiosyncratic variances of the stocks' returns. All these estimates depend on return history up to time t and are, thus, time dependent. Further details on the estimation of $\hat{\beta}$ and $\hat{\sigma}_{\mathrm{I}}$ can be found in Appendix A.1.

4.3 Shrinkage towards a Single-Factor Market Model Estimator (SMM)

The assumptions of structural estimates like the single-factor market model approach are very rigid so that one ignores a lot of information that is present (though noisy) in the sample covariance matrix. As discussed in Section 4.1, shrinking the sample covariance matrix $\hat{\Sigma}_S$ towards the structural estimator $\hat{\Sigma}_{\mathrm{MM}}$ applying the approach of Ledoit and Wolf (2003) is equivalent to using $\hat{\Sigma}_{\mathrm{MM}}$ as a prior estimate and optimally adapting this prior to the signal obtained through $\hat{\Sigma}_S$ (fully regarding the noisiness of $\hat{\Sigma}_S$). The resulting estimate of this method is then

$$\hat{\Sigma}_{\mathrm{SMM}} = \delta_{\mathrm{SMM}}\hat{\Sigma}_{\mathrm{MM}} + (1 - \delta_{\mathrm{SMM}})\hat{\Sigma}_S \tag{10}$$

Please note that the shrinkage intensity, δ_{SMM}, differs depending on the shrinkage target that is used and it is, of course, changing over time.

4.4 The Average-Correlation Estimator (AC)

The idea to avoid estimation errors in the covariance matrix by assuming that the pair-wise correlation of asset returns is identical over all pairs of stocks dates back to Elton and Gruber (1973). They propound the idea that deviations of the individual correlation coefficients from their average "are random or sufficiently unstable, so that zero is a better estimate". In a way this is a step back from the single-factor market model in assuming that even the co-movement with a single factor cannot be estimated reliably and better is replaced by an average correlation between all the stocks.

The structural estimator is then

$$\hat{\Sigma}_{\text{AC}} = \Delta C \Delta, \quad \text{with } \Delta = \operatorname{diag}(\hat{\sigma}_i), C = \begin{pmatrix} 1 & \hat{\rho} & \cdots & \hat{\rho} \\ \hat{\rho} & 1 & \cdots & \hat{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho} & \hat{\rho} & \cdots & 1 \end{pmatrix}, \tag{11}$$

where $\hat{\sigma}_i$ is the estimated standard deviation of stock i's return and $\hat{\rho}$ is the sample estimate of the mean of the pair-wise correlation, see Section A.2 for details.

While the average pair-wise correlation can apparently be estimated with more accuracy than individual pair-wise correlation, it might be an oversimplification to average out all sample information about correlations. Recent work by Elton, Gruber, and Spitzer (2006) documents competitive performance of portfolio strategies based on the average-correlation estimator of the covariance matrix.

4.5 Shrinkage Towards the Average-Correlation Estimator (SAC)

A further approach is to use the average-correlation estimator $\hat{\Sigma}_{\text{AC}}$ as a shrinkage target in a Ledoit and Wolf (2003) shrinkage estimator

$$\hat{\Sigma}_{\text{SAC}} = \delta_{\text{SAC}} \hat{\Sigma}_{\text{AC}} + (1 - \delta_{\text{SAC}}) \hat{\Sigma}_S \tag{12}$$

4.6 Shrinkage Towards the Average-Variance Estimator (SAV)

A reaction to concerns in the estimation quality of the average pair-wise correlation of returns as well as in individual return variances is the average-variance estimator. For this structural estimator it is assumed that asset returns are pair-wise uncorrelated and, furthermore, the individual return variance is estimated by an estimate of the mean of the return variances. This reduces the number of elements of the covariance matrix to be estimated from $K(K+1)/2$ in the sample covariance matrix to only 1, the average-variance estimator. We do not regard this extreme estimator as a generator of minimum-variance portfolios because it will not provide a stock selection but result simply in the equally-weighted portfolio. Rather we use the average-variance estimator as a Ledoit and Wolf (2003) shrinkage target

$$\hat{\Sigma}_{\mathrm{SAV}} = \delta_{\mathrm{SAV}} \hat{\Sigma}_{\mathrm{AV}} + (1 - \delta_{\mathrm{SAV}}) \hat{\Sigma}_S, \tag{13}$$

with

$$\hat{\Sigma}_{\mathrm{AV}} = \hat{\bar{\sigma}}^2 \mathbb{I},$$

with $\hat{\bar{\sigma}}^2$ the average over all estimated individual return variances and $\mathbb{I}$ the $(K \times K)$ identity matrix, see Section A.3 for details.

4.7 The Principal Component Estimator (PC)

A common practice to reduce the noisiness of covariance matrix estimators and to receive a positive definite matrix is to apply an eigenvalue decomposition of the sample covariance matrix. Usually only a small number of eigenvectors explain almost the entire sample variance. So the structural principal component estimator PCn selects the set of n eigenvectors of $\hat{\Sigma}_S$ corresponding to the n largest eigenvalues and constructs an structural estimator in the following form

$$\hat{\Sigma}_{\mathrm{PC}n} = U' \Lambda_n U + \Sigma_{\mathrm{I,PC},n},$$

where U is the matrix that contains the orthogonal eigenvectors of $\hat{\Sigma}_S$ in its rows, sorted by the magnitude of the corresponding eigenvalues. Λ_n is $(K \times K)$ diagonal, with the n largest eigenvalues of $\hat{\Sigma}_S$ as first n elements of the diagonal, the rest replaced by 0. $\hat{\Sigma}_{\mathrm{PC}n}$ is positive definite if $n < \mathrm{rank}(R)$. The matrix $\Sigma_{\mathrm{I,PC},n}$ is a diagonal matrix containing the

Table 1: Notation: Covariance estimation methods

method	description
MM	single-factor market model estimator
SMM	shrinkage towards the single-factor market model estimator
AC	average-correlation estimator
SAC	shrinkage towards the average-correlation estimator
SAV	shrinkage towards the average-variance estimator
PC3	three-factor principal components model
PC5	five-factor principal components model

residual variances from projecting the individual returns onto the subspace spanned by the first n eigenvectors.

In our analysis we use principal component estimators with n equal to 3 as well as with n equal to 5 components.

Table 1 summarizes the different approaches in estimating the covariance matrix of returns that we use throughout this study.

5 The Empirical Analysis

5.1 The Data

We do our analysis for US equities and European equities, respectively. The universe of stocks in the markets which are considered in the study is defined by the constituent lists of indexes representative of large and mid-cap, liquid stocks, S&P 500 Index for the US and STOXX Europe 600 Index for Europe. All data is retrieved from Thomson Reuters Datastream.

The overall history that can be researched in these stock universes is determined by the availability of the constituent lists of the indexes. Table 2 gives an overview of the data availability. For the US, available history is 23.7 years and for Europe 13.7 years.

Base data for estimating covariance matrices are simple returns calculated from Datastream's total return index data field (RI, assuming full reinvestment of distributed dividends). Returns are in USD for the S&P 500 constituents and unhedged EUR returns in Europe. We consider two different data frequencies, one observation per week as well as

Table 2: Data: Description and availability

region	index	availability of index constituents	avg. nr. constituents	currency
USA	S&P 500	10-1989 to 05-2013	500	USD
Europe	STOXX Europe 600	09-1999 to 05-2013	600	EUR

one observation every four weeks. To avoid survivorship bias, at each point in time t only those stocks are considered which are active time-t index constituents.

5.2 The Setup

In our empirical study, we aim to implement a broad set of different approaches in constructing minimum-variance portfolios in order to understand common features of min-var candidates, i.e., those stocks which are picked by one or more portfolio methods. According to the discussion in Section 2, in each of the markets we analyze minimum-variance approaches that differ with respect to the following dimensions:

(a) The way of estimating the covariance matrix

- Data frequency and history: weekly returns on a one-year rolling window (**1w1a**), weekly returns on a two-years rolling window (**1w2a**), returns calculated over four weeks on a five-years rolling window (**4w5a**),
- Estimation method: each of the seven methods listed in Table 1

(b) Constraints on portfolio weights.

- Long-only with unconstrained maximum ($w_{\max} = \mathbf{100\%}$), long-only with 8% maximum ($w_{\max} = \mathbf{8\%}$), long-only with 4% maximum ($w_{\max} = \mathbf{4\%}$).

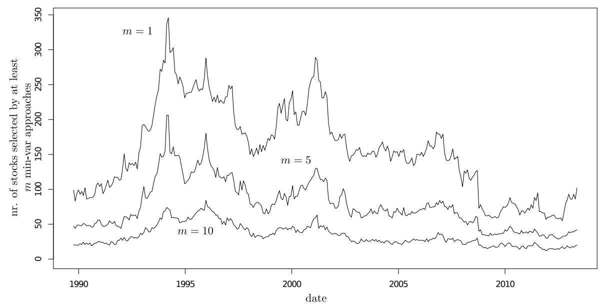
At a given point in time we produce 21 different estimates of the covariance matrix, $\hat{\Sigma}$. Each of these matrices is then input to three minimum-variance portfolio optimizations which differ in the applied constraint on the maximum weight. Thus, we compare 63 different portfolio selections for each equity universe.

We conduct this cross sectional study of 63 minimum-variance approaches every four weeks from the beginning of the data availability until May 2013.

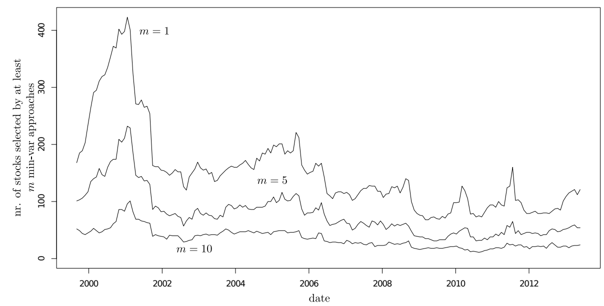
Within this setup we confirm a set of common features of minimum-variance portfolios, which are already well-documented (see Appendix A.4 for more details): Long-only minimum-variance portfolios tend to be concentrated portfolios, they pick stocks which have low beta and low idiosyncratic risk. Moreover, we observe that different methods to construct minimum-variance portfolios show a high level of agreement when deciding which stocks should be picked. Table 3 documents the overlap. Consider the results for S&P 500 in panel a) of Table 3. Over the most recent five years, an average of only 74 out of 500 stocks qualifies as minimum-variance candidate in the broad sense that it is selected by at least one out of 21 minimum-variance approaches that apply a maximum weight of 100%. If we require larger diversification and use a maximum of only 4% of portfolio weight, this number increases to 93 which is still less than 20% of stocks in the universe.

Figure 1 displays the number of stocks that are picked by at least m different approaches using $w_{\max} = 100\%$. We will later refer to those selections as aggregate min-var targets. It is apparent that among S&P 500 stocks as well as among STOXX Europe 600 stocks, min-var approaches become more and more concentrated and the overlaps of different methods increase. While in the 1990s, up to 346 out of 500 S&P 500 stocks are picked by at least one of those 21 approaches and in the early 2000s their number still reaches 250 to 300, the number of stocks that qualify for at least one inclusion in a min-var portfolio is below 60 in 2009 and again in 2011 and 2012, with an average of 75 stocks in the most recent five years, see Table 3. The aggregate min-var target which contains all stocks that are picked by at least 10 out of 21 approaches with $w_{\max} = 100\%$ counts 84 stocks at the end of 1995 and reaches its minimum size of only 12 stocks in 2011. The average over the last five years is 18 stocks.

Aggregate min-var targets among STOXX Europe 600 stocks which contain all stocks picked by at least one min-var approach with $w_{\max} = 100\%$ have its maximum size of 423 stocks at the beginning of 2001 and its minimum size of 68 in 2009, with an average over the most recent five years of 94. The more concentrated aggregate target with stocks that are picked by at least 10 out of 21 min-var methods have a maximum of 101 constituents in early 2001 and a minimum of 11 constituents mid 2010. The average over the most recent five years is 20 stocks.



(a) S&P 500



(b) STOXX Europe 600

Figure 1: Number of stocks that are simultaneously picked by at least m different min-var approaches with $w_{\max} = 100\%$, for $m = 1$, $m = 5$, and $m = 10$.

To get a feeling for the extent to which the selections agree on stock picking, consider the following thought experiment. Let us assume that all 21 portfolio approaches are completely independent and select at random 22 stocks out of 500 (from Table 11 we see that on average a $w_{\max} = 100\%$ portfolio has 22 constituents). Then the number of stocks that are selected at least by one out of 21 portfolio approaches is expected to be 306. The probability that a stock is selected by at least 5 approaches is then only 0.19% (or approx. 1 out of 500 stocks). The probability of being selected by at least 10 or 15 independent approaches is virtually zero (6.1×10^{-9} and 1.89×10^{-16}, respectively). Thus, we conclude that different minimum variance approaches have a rather high degree of coincidence in their stock selection.

6 A Shift in Risk Preferences: The Increasing Price-to-Book Ratio of Minimum-Variance Portfolio Constituents

We conduct the cross sectional regression (2) every four weeks for all 63 different minimum-variance approaches and collect the estimates $\hat{\gamma}^A_{\text{minvar},t}$ of the relative valuation coefficients. We then perform a GLS time-series regression (4) of $\hat{\gamma}^A_{\text{minvar},t}$ on a linear time trend and consider the estimated slope coefficients $\hat{b}^A$ of each of the approaches. Standard errors of these slope coefficients are corrected according to Newey and West (1987, 1994) for the

Table 3: Overlap in portfolio selection: The average number of stocks picked by at least m different minimum-variance selections is documented for various maximum weights and different regions. Each line compares 21 models. Averages are calculated over the five-year period from 2008-05-31 to 2013-05-31.

a) S&P 500

	$m = 1$	$m = 5$	$m = 10$	$m = 15$
$w_{\max} = 100\%$	74.43	37.83	18.18	8.34
$w_{\max} = 8\%$	78.40	41.52	21.62	10.49
$w_{\max} = 4\%$	92.94	52.95	29.94	18.11

b) STOXX Europe 600

	$n = 1$	$n = 5$	$n = 10$	$n = 15$
$w_{\max} = 100\%$	93.40	44.63	19.95	8.82
$w_{\max} = 8\%$	98.74	48.52	23.40	10.83
$w_{\max} = 4\%$	115.63	61.68	32.60	15.75

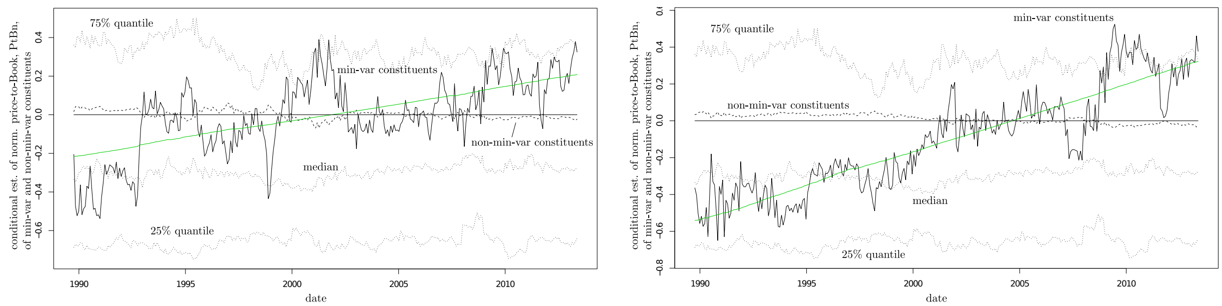
autocorrelation in $\hat{\gamma}^{A}_{\text{minvar},t}$ which comes from an overlap of historical return data from which the covariance matrices are estimated.

Figure 2 shows two graphs that are representative for the time trend in the relative valuation of minimum-variance portfolio constituents for the S&P 500. It plots the conditional estimate of the normalized price-to-book, PtBn, of min-var constituents as a solid line, the unconditional PtBn being zero due to normalization. This conditional estimate is given by the estimated coefficient $\hat{\gamma}^{A}_{\text{minvar},t}$ of the minimum-variance indicator times the level of the indicator $d^{A}_{\text{minvar},t}$ shared by all min-var constituents at t. We also plot the conditional estimate of PtBn for non-minimum-variance constituents as a dashed line. Due to cross sectional normalization of price-to-book and the fact that long-only minimum-variance approaches pick only a small number of stocks, the conditional estimate for non-constituents is very close to the unconditional estimate of zero. Panel a) plots conditional valuation estimates for portfolios based on the single-factor market model estimator, and Panel b) is based on the average-correlation estimator. Both approaches use returns at a four-weeks frequency over a history of five years (4w5y) and a maximum weight of 4% per stock ($w_{\max} = 4\%$). The graphs in Figure 2 also indicate the position of the 25% quantile, the median, and the 75% quantile of the cross-sectionally normalized price-to-book, PtBn, relative to the unconditional estimate of PtBn which is, by construction, equal zero. Independent of the particular portfolio approach, we find that the valuation of minimum-variance portfolio constituents increases steadily over time.

While in the 1990s, minimum-variance constituents traded at a discount relative to the unconditional mean among S&P 500 stocks, this discount turned into a premium somewhen between the years 2000 and 2005. At present, the average PtBn premium of minimum-variance constituents is considerably above the unconditional mean. Or in other words, we document a shift in market preferences for min-var constituents over the past 23 years which is relevant given the dispersion of price-to-book of S&P 500 stocks. Even after controlling for other pricing relevant characteristics like size, industry affiliation, and dividend payout, we see that min-var constituents today trade at a premium comparable to the difference between the 75% quantile of the price-to-book among S&P 500 stocks and its mean. Section 6.2 provides an estimate of the price impact of the detected shift in preferences.

Figure 3 shows the relative valuation of minimum-variance constituents in the universe of STOXX Europe 600 stocks. Although the available history of STOXX Europe 600 index compositions is ten years shorter compared to our US data, we identify a similar time trend in the relative valuation of min-var constituents. Again, between 2000 and 2005, the conditional estimate of min-var constituents' valuation turns from a discount into a premium, increasing steadily over the past 13 years. Today, minimum-variance constituents trade at a considerable premium. The magnitude of this premium is again comparable to the difference between the 75% quantile of the price-to-book distribution among STOXX Europe 600 stocks and its mean.

To quantify the magnitude of this shift in preferences for stocks that qualify as minimum-variance portfolio constituents, we report in Table 4 estimates of the annualized slope coefficients $\hat{b}^A$ in regression (4) for all 63 different minimum-variance approaches employed for the S&P 500 (Panel a) as well as for STOXX Europe 600 (Panel b) together with Newey-West corrected t-values (in parentheses). For the S&P 500, the estimated drift coefficients are between 1.0% and 4.1% per year (in multiples of cross sectional standard deviations of price-to-book) with an average of 2.0%. That means that over the past 23.7 years, on average the price-to-book coefficient $\gamma^A_{\text{minvar},t}$ of minimum-variance candidate stocks increased about 0.48 standard deviations of the the cross-sectional price-to-book distribution. Regarding the significance of these slope coefficients, we report t-values between 2.9 and 12.1.



(a) Single-factor market model estimator, 4w5y, $w_{\text{max}} = 4\%$

(b) Average-correlation estimator, 4w5y, $w_{\text{max}} = 4\%$

Figure 2: Time trend in the relative valuation of minimum-variance constituents for two different portfolio approaches on S&P 500 stocks, using returns at returns over four weeks on a 5 years rolling window (4w5y) and a maximum weight of 4% per stock ($w_{\text{max}} = 4\%$). The solid line shows the conditional estimate of the normalized price-to-book of min-var portfolio constituents, the dashed line that of non-minimum-variance constituents. Dotted lines show the 25% quantile, the median, and the 75% quantile of the normalized prize-to-book relative to the unconditional mean represented by the abscissa. The green line plots the time trend of the conditional estimate of min-var constituent's normalized price-to-book from GLS-regression (4). While the conditional estimate of PtBn of minimum-variance constituents was negative in the 1990s, relative valuation in recent years is considerably positive.

The least significant slope in the valuation coefficient is observed when using a shrinkage to average variance estimator with weekly returns on a one-year rolling window (1w1y) and a maximum weight of 4% ($w_{\text{max}} = 4\%$).

For STOXX Europe 600 stocks, the slope coefficients of all models are positive, between 1.5% and 5.8% per year with an average of 3.9%. I.e., over the 13.7 years which are covered by our data, on average the price-to-book coefficient $\gamma^{A}_{\text{minvar},t}$ increased by 0.54 standard deviations of the cross sectional distribution of price-to-book. Calculated t-values are between 0.53 and 8.9. There are 9 out of 63 approaches that show a t-value below 1.645, the least significant is the shrinkage towards average correlation estimator approach when using returns over 4 weeks on a five years rolling window (4w5y) and a maximum weight of 4% ($w_{\text{max}} = 4\%$).

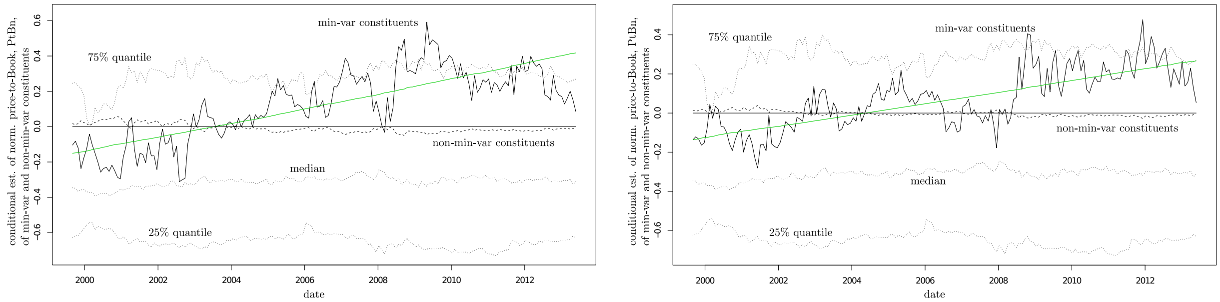
Table 4: Estimates $\hat{b}^A$ of the annualized drift coefficient in (4) together with Newey-West corrected t-values (in parentheses).

a) S&P 500

		MM	SMM	AC	SAC	SAV	PC3	PC5
1w1y								
	$w_{\max} = 100\%$	0.0137	0.0131	0.0222	0.0180	0.0102	0.0131	0.0135
		(4.10)	(4.17)	(5.98)	(5.90)	(3.01)	(4.03)	(3.98)
	$w_{\max} = 8\%$	0.014	0.0133	0.0223	0.0179	0.0103	0.0128	0.0125
		(4.23)	(4.22)	(5.42)	(5.98)	(3.28)	(4.05)	(3.83)
	$w_{\max} = 4\%$	0.013	0.0123	0.0200	0.0183	0.0099	0.0124	0.0118
		(3.84)	(4.05)	(5.63)	(5.25)	(2.97)	(3.87)	(3.57)
1w2y								
	$w_{\max} = 100\%$	0.0192	0.0189	0.0245	0.0207	0.0155	0.0168	0.0182
		(6.59)	(6.16)	(5.20)	(4.60)	(5.27)	(5.34)	(6.57)
	$w_{\max} = 8\%$	0.0181	0.0177	0.0265	0.0204	0.0156	0.0166	0.0175
		(6.58)	(6.08)	(5.47)	(5.30)	(6.35)	(5.53)	(7.25)
	$w_{\max} = 4\%$	0.0160	0.0156	0.0233	0.0219	0.0143	0.0160	0.0170
		(6.00)	(4.86)	(4.79)	(6.03)	(5.99)	(4.30)	(6.73)
4w5y								
	$w_{\max} = 100\%$	0.0206	0.0232	0.0400	0.0356	0.0237	0.0223	0.0269
		(5.97)	(4.85)	(10.24)	(8.96)	(7.07)	(6.03)	(8.11)
	$w_{\max} = 8\%$	0.0204	0.0225	0.0387	0.0337	0.0236	0.0219	0.0265
		(5.72)	(4.70)	(9.64)	(8.18)	(7.09)	(6.63)	(8.82)
	$w_{\max} = 4\%$	0.0205	0.022	0.0405	0.0377	0.0224	0.0215	0.0272
		(5.94)	(4.93)	(12.11)	(11.61)	(6.90)	(7.31)	(10.35)

b) STOXX Europe 600

		MM	SMM	AC	SAC	SAV	PC3	PC5
1w1y								
	$w_{\max} = 100\%$	0.0583	0.0424	0.0481	0.0478	0.0337	0.0359	0.0363
		(3.92)	(6.04)	(8.49)	(8.91)	(5.88)	(5.55)	(4.95)
	$w_{\max} = 8\%$	0.0578	0.0422	0.044	0.0414	0.0338	0.0371	0.0367
		(3.79)	(5.92)	(7.51)	(6.24)	(5.55)	(5.40)	(5.10)
	$w_{\max} = 4\%$	0.0574	0.0427	0.0412	0.0369	0.0326	0.0382	0.0363
		(3.89)	(5.61)	(4.94)	(5.00)	(3.44)	(5.36)	(4.82)
1w2y								
	$w_{\max} = 100\%$	0.0568	0.0507	0.0437	0.0422	0.0390	0.0440	0.0435
		(7.24)	(6.22)	(6.24)	(7.08)	(6.51)	(3.67)	(6.15)
	$w_{\max} = 8\%$	0.0567	0.0506	0.0418	0.0384	0.0384	0.0426	0.0429
		(7.51)	(6.32)	(3.49)	(6.13)	(6.23)	(3.41)	(5.74)
	$w_{\max} = 4\%$	0.0570	0.0507	0.0439	0.0425	0.0384	0.0417	0.0425
		(6.20)	(5.55)	(5.01)	(5.36)	(5.82)	(3.61)	(5.61)
4w5y								
	$w_{\max} = 100\%$	0.0519	0.0409	0.0434	0.0331	0.0188	0.0150	0.0203
		(1.85)	(1.62)	(3.68)	(4.16)	(1.65)	(0.95)	(1.57)
	$w_{\max} = 8\%$	0.0513	0.0423	0.0404	0.0236	0.0174	0.0169	0.0211
		(2.05)	(1.80)	(4.83)	(1.71)	(1.62)	(1.14)	(1.55)
	$w_{\max} = 4\%$	0.0526	0.0421	0.0356	0.0245	0.0158	0.0215	0.0254
		(2.14)	(1.82)	(3.95)	(0.53)	(1.03)	(1.42)	(1.88)



(a) Single-factor market model estimator, 4w5y, $w_{\max} = 4\%$

(b) Average-correlation estimator, 4w5y, $w_{\max} = 4\%$

Figure 3: Time trend in the relative valuation of minimum-variance constituents for two different portfolio approaches on STOXX Europe 600 stocks, using returns at returns over four weeks on a 5 years rolling window (4w5y) and a maximum weight of 4% per stock ($w_{\max} = 4\%$). The solid line shows the conditional estimate of the normalized price-to-book of min-var portfolio constituents, the dashed line that of non-minimum-variance constituents. Dotted lines show the 25% quantile, the median, and the 75% quantile of the normalized prize-to-book relative to the unconditional mean represented by the abscissa. The green line plots the time trend of the conditional estimate of min-var constituent's normalized price-to-book from GLS-regression (4). While the conditional estimate of PtBn of minimum-variance constituents was below neutral around the year 2000, relative valuation in recent years is considerably positive.

6.1 Valuation Trends in Aggregated Min-Var Targets

For individual min-var portfolio approaches we find a time trend in the valuation coefficient which is positive without exception (though some t-values are rather low for European equities). Our interest, however, does not focus on the individual min-var approach but on the commonalities among all implemented approaches. For this reason we study aggregate min-var portfolio selections which are already discussed above. These target portfolios are aggregated in the sense that they contain stocks which are included in at least m individual min-var portfolios, for m in 1, 2, $\ldots$ 10. Table 5 displays the estimated drift $\hat{b}$ in valuation coefficients. The idea behind this aggregation is the argument that a stock which is considered a min-var constituent by a high number of individual approaches is somehow closer to the core of a min-var target portfolio aspired by minimum-variance investors compared to a stock that is only picked by one or two approaches.

When inspecting drifts in the coefficient $\hat{b}$ in the time-series regression (4) we see that Newey-West corrected t-values are high throughout all values of m with a minimum of 4.8

Table 5: Estimates $\hat{b}^A$ of the annualized drift coefficient in (4) together with Newey-West corrected t-values (in parentheses) for aggregate min-var targets, i.e., for stocks that are picked by at least m different portfolio approaches.

a) S&P 500

	$w_{\max} = 100\%$	$w_{\max} = 8\%$	$w_{\max} = 4\%$
$m = 1$	0.0160	0.0158	0.0151
	(4.78)	(5.09)	(5.17)
$m = 2$	0.0170	0.0165	0.0166
	(5.33)	(5.21)	(5.53)
$m = 3$	0.0188	0.0182	0.0177
	(6.20)	(6.69)	(6.00)
$m = 4$	0.0203	0.0197	0.0176
	(8.30)	(7.81)	(6.74)
$m = 5$	0.0212	0.0208	0.0193
	(9.00)	(9.08)	(8.38)
$m = 6$	0.0235	0.0225	0.0214
	(8.86)	(9.06)	(8.43)
$m = 7$	0.0251	0.0246	0.0227
	(8.48)	(9.15)	(7.96)
$m = 8$	0.0250	0.0247	0.0234
	(9.01)	(9.93)	(8.09)
$m = 9$	0.0244	0.0243	0.0231
	(7.73)	(8.55)	(8.21)
$m = 10$	0.0250	0.0240	0.0231
	(8.43)	(8.55)	(8.29)

b) STOXX Europe 600

	$w_{\max} = 100\%$	$w_{\max} = 8\%$	$w_{\max} = 4\%$
$m = 1$	0.0445	0.0441	0.0419
	(6.62)	(6.65)	(5.90)
$m = 2$	0.0432	0.0428	0.0432
	(5.42)	(6.03)	(7.76)
$m = 3$	0.0426	0.0428	0.0443
	(6.09)	(6.20)	(7.94)
$m = 4$	0.0416	0.0418	0.0464
	(5.39)	(5.90)	(6.80)
$m = 5$	0.0433	0.0432	0.0451
	(5.11)	(4.94)	(6.41)
$m = 6$	0.0483	0.0457	0.0455
	(5.17)	(5.26)	(3.78)
$m = 7$	0.0496	0.0466	0.0462
	(4.75)	(5.33)	(4.43)
$m = 8$	0.0509	0.0480	0.0481
	(7.03)	(5.76)	(5.71)
$m = 9$	0.0514	0.0481	0.0489
	(6.16)	(5.63)	(6.74)
$m = 10$	0.0532	0.0501	0.0481
	(6.76)	(6.47)	(5.60)

for S&P 500 stocks and 3.8 for STOXX Europe 600 stocks. We see the following intuitive behavior of $\hat{b}$: It increases with growing m, independent of $w_{\max}$. This means, the shift in valuation is stronger for stocks that are picked by a larger number of different models. I.e., the stronger the agreement of different models about the min-var qualification, the stronger is the shift in investors' preferences. For a given m, the shift is stronger for higher $w_{\max}$. This makes sense since a higher value of $w_{\max}$ means that portfolio approaches considered for the aggregation are allowed to focus more on min-var favorites, while with decreasing $w_{\max}$, approaches are forced to diversify into stocks less preferred. For S&P 500, focussing on the aggregation of $m = 10$ implies $\hat{b}$ of 2.5% per year, which means that for stocks that are selected by at least 10 out of 21 approaches, the coefficient of the min-var indicator, $\hat{\gamma}$ increased by 0.6 standard deviations of the cross-sectional price-to-book distribution over the last 23.7 years. For European stocks, this drift is even more pronounced and reaches 5.3% per year for stocks that are selected by at least 10 approaches, implying an increase of $\hat{\gamma}$ by 0.73 standard deviations over the last 13.7 years.

6.2 Implied Price Impact

To estimate the implicit price impact of such a shift in price-to-book, we first evaluate the estimate of the pricing coefficient $\hat{\hat{\gamma}}^A_{\text{minvar},t}$ at the begin of the data set (10-1989 for S&P 500 and 09-1999 for STOXX Europe 600) as well as at today (05-2013). We do this for the aggregate min-var targets $m = 1$ to $m = 10$ and list the results and the respective standard errors according to (6) in Table 6.

To receive the estimate of normalized price-to-book at the begin and at the end of the time horizon, we multiply $\hat{\hat{\gamma}}^A_{\text{minvar},t}$ with the min-var indicator $d^A_{\text{minvar},t}$ (the de-meaned min-var dummy), according to (3). In a further step, we translate the estimate of normalized price-to-book into an estimate of the raw price-to-book using sample statistics of price-to-book of S&P 500 and STOXX Europe 600 stocks from Table 7.[13]

Finally, we estimate the impact on non-normalized price-to-book and the implied impact on the pricing on constituents of aggregate min-var targets.

$$\hat{\hat{\text{PtB}}}_t \,|\, \text{minvar} = \text{avg. PtB}_t + (\hat{\hat{\text{PtBn}}}_t | \, \text{minvar}) \times \text{stdev.}(\text{PtB})_t \tag{14}$$

[13] Data are trimmed at the 2% and at the 98% quantile respectively.

Table 6: Estimate $\hat{\hat{\gamma}}^{A}_{\text{minvar},t}$ and standard errors of the min-var valuation coefficient from time-series regression (4). Estimates are calculated at the begin and at the end of the data set for constituents of min-var target portfolios, i.e., portfolios of stocks that are picked by at least m different approaches with $w_{\max} = 100\%$.

a) S&P 500

	$t = $ 10-1989		$t = $ 05-2013	
	$\hat{\hat{\gamma}}^{A}_{\text{minvar},t}$	$\text{se}(\hat{\hat{\gamma}}^{A}_{\text{minvar},t})$	$\hat{\hat{\gamma}}^{A}_{\text{minvar},t}$	$\text{se}(\hat{\hat{\gamma}}^{A}_{\text{minvar},t})$
$m = 1$	-0.1531	0.0410	0.2266	0.0551
$m = 2$	-0.1733	0.0388	0.2297	0.0534
$m = 3$	-0.1940	0.0387	0.2509	0.0517
$m = 4$	-0.2182	0.0327	0.2620	0.0393
$m = 5$	-0.2316	0.0335	0.2699	0.0360
$m = 6$	-0.2659	0.0364	0.2920	0.0422
$m = 7$	-0.2864	0.0363	0.3077	0.0479
$m = 8$	-0.2894	0.0311	0.3032	0.0473
$m = 9$	-0.2818	0.0354	0.2971	0.0531
$m = 10$	-0.2875	0.0365	0.3044	0.0492

b) STOXX Europe 600

	$t = $ 09-1999		$t = $ 05-2013	
	$\hat{\hat{\gamma}}^{A}_{\text{minvar},t}$	$\text{se}(\hat{\hat{\gamma}}^{A}_{\text{minvar},t})$	$\hat{\hat{\gamma}}^{A}_{\text{minvar},t}$	$\text{se}(\hat{\hat{\gamma}}^{A}_{\text{minvar},t})$
m=1	-0.1720	0.0487	0.4402	0.0553
m=2	-0.1547	0.0443	0.4401	0.0743
m=3	-0.1626	0.0414	0.4237	0.0665
m=4	-0.1552	0.0427	0.4171	0.0769
m=5	-0.1677	0.0420	0.4292	0.0868
m=6	-0.2106	0.0406	0.4551	0.1008
m=7	-0.2193	0.0414	0.4640	0.1124
m=8	-0.2230	0.0287	0.4782	0.0794
m=9	-0.2304	0.0325	0.4767	0.0899
m=10	-0.2473	0.0302	0.4848	0.0846

Table 7: Sample statistics of price-to-book of S&P 500 and STOXX Europe 600 stocks at the begin and at the end of our data sample.

	avg. PtB	stdev. PtB
S&P 500: 10-1989	2.34	1.36
S&P 500: 05-2013	3.57	2.78
STOXX Europe 600: 09-1999	3.48	3.11
STOXX Europe 600: 05-2013	2.70	2.34

where avg. PtB and stdev. PtB are the cross sectional average and standard deviation of price-to-book (see Table 7), respectively. The conditional estimate of normalized price-to-book, given that the stock is a min-var candidate, $\hat{\hat{\text{PtBn}}}_t|\,\text{minvar}$, is determined by (7).

Table 8 shows the final result. For US stocks as well as for European stocks, we see a discount in price-to-book (compared to the unconditional average stated in Table 7) as well as in the implied pricing at the begin of our data set. This discount turns into a premium today. With increasing m, aggregate targets are more and more focused on stocks that are picked by a large number of different min-var approaches. The intuition that valuation effects should be more emphasized for higher aggregation is perfectly supported by our results.

In the US, for the broadest definition of an aggregate minimum-variance target portfolio, i.e., stocks that are selected by at lest one of the approaches, $m = 1$, the estimated price discount in 10-1989 is $-6.91\,\%$. This broad target selection contains 99 out of 500 stocks. For stocks that qualify for at least 10 different approaches (21 out of 500), this discount grows to $-15.98\,\%$. Today, the estimated premium is $13.89\,\%$ for the broad $m = 1$ selection and grows to $22.72\,\%$ today. The size of the targets is 102 and 20, respectively, see Figure 1.

In Europe, the trend in preferences for min-var constituent stocks is more pronounced. The estimated price discount in 09-1999 is $-10.52\,\%$ for the broadest min-var target ($m = 1$) and $-19.96\,\%$ for the concentrated target which includes only those stocks that are selected by at least 10 individual approaches ($m = 10$). The size of these targets is 168 and 52 out of 600 stocks, respectively. The estimated premium at which min-var candidates trade today is $30.10\,\%$ for the broadest target (121 stocks) and as much as $40.28\,\%$ for the concentrated target ($m = 10$, 24 stocks).

Our results demonstrate that over the past decades, there has been a shift in preferences for stocks that are suitable for an inclusion in a minimum variance portfolio both in the US as well as in Europe. This shift can be detected independently of the minimum-variance portfolio approach we use and it is stronger for stocks that are selected by different min-var approaches simultaneously.

Table 8: The estimated of the conditional price-to-book ratio, $\hat{\hat{\text{PtB}}}_t|\text{minvar}$, for constituents of min-var targets for $m = 1$ to $m = 10$. The column *price impact* is the discount or premium in price-to-book relative to the cross sectional average. Standard errors of these estimates, se(.), are also stated.

a) S&P 500

$t = 10\text{-}1989$

	$\hat{\hat{\text{PtB}}}$|minvar	se($\hat{\hat{\text{PtB}}}$|minvar)	price impact	se(price impact)
m=1	2.1755	0.0433	-6.91 %	1.85 %
m=2	2.1422	0.0436	-8.34 %	1.87 %
m=3	2.1111	0.0450	-9.67 %	1.93 %
m=4	2.0768	0.0390	-11.14 %	1.67 %
m=5	2.0551	0.0408	-12.07 %	1.74 %
m=6	2.0042	0.0456	-14.24 %	1.95 %
m=7	1.9712	0.0464	-15.65 %	1.99 %
m=8	1.9639	0.0402	-15.97 %	1.72 %
m=9	1.9728	0.0458	-15.59 %	1.96 %
m=10	1.9636	0.0474	-15.98 %	2.03 %

$t = 05\text{-}2013$

	$\hat{\hat{\text{PtB}}}$|minvar	se($\hat{\hat{\text{PtB}}}$|minvar)	price impact	se(price impact)
m=1	4.0642	0.1204	13.89 %	3.37 %
m=2	4.1096	0.1259	15.16 %	3.53 %
m=3	4.1814	0.1263	17.17 %	3.54 %
m=4	4.2237	0.0983	18.36 %	2.75 %
m=5	4.2530	0.0913	19.18 %	2.56 %
m=6	4.3328	0.1105	21.41 %	3.10 %
m=7	4.3756	0.1257	22.61 %	3.52 %
m=8	4.3673	0.1245	22.38 %	3.49 %
m=9	4.3565	0.1408	22.08 %	3.94 %
m=10	4.3793	0.1311	22.72 %	3.67 %

b) STOXX Europe 600

$t = 09\text{-}1999$

	$\hat{\hat{\text{PtB}}}$|minvar	se($\hat{\hat{\text{PtB}}}$|minvar)	price impact	se(price impact)
m=1	3.1121	0.1036	-10.52 %	2.98 %
m=2	3.1252	0.101	-10.14 %	2.90 %
m=3	3.0872	0.0995	-11.24 %	2.86 %
m=4	3.0969	0.1049	-10.96 %	3.02 %
m=5	3.0555	0.1059	-12.15 %	3.05 %
m=6	2.9263	0.1065	-15.86 %	3.06 %
m=7	2.8944	0.1102	-16.78 %	3.17 %
m=8	2.8689	0.0785	-17.51 %	2.26 %
m=9	2.8379	0.0904	-18.41 %	2.6 %
m=10	2.7837	0.0849	-19.96 %	2.44 %

$t = 05\text{-}2013$

	$\hat{\hat{\text{PtB}}}$|minvar	se($\hat{\hat{\text{PtB}}}$|minvar)	price impact	se(price impact)
m=1	3.5092	0.1021	30.10 %	3.78 %
m=2	3.5595	0.1455	31.97 %	5.40 %
m=3	3.5584	0.1352	31.93 %	5.01 %
m=4	3.5740	0.1617	32.51 %	5.99 %
m=5	3.6065	0.1839	33.71 %	6.82 %
m=6	3.6781	0.2172	36.36 %	8.05 %
m=7	3.7087	0.2451	37.50 %	9.09 %
m=8	3.7534	0.1753	39.16 %	6.50 %
m=9	3.7579	0.2001	39.33 %	7.42 %
m=10	3.7838	0.1896	40.28 %	7.03 %

7 Robustness Checks

We perform a set of robustness checks to provide evidence that our findings are robust to changes in the specific setup of the study. First, one might be concerned about autocorrelation in the estimated coefficients $\hat{\gamma}^{A}_{\text{minvar},t}$ that might not be properly treated by a simple Newey-West correction. Second, estimation errors in the sample covariance matrix $\hat{\Sigma}_s$ might influence all portfolio approaches at the same time and lead to artifacts in all portfolio selections. Third, we control for collinearity in variables, since it might be possible that the selection into a minimum-variance portfolio is strongly correlated to other characteristics (like high dividend payout) and that valuation effects are attributed to the min-var indicator simply because of this collinearity.

7.1 Avoiding Autocorrelation in Valuation Coefficients

When estimating coefficients $\hat{\gamma}^{A}_{\text{minvar},t}$ on a four-weeks frequency, the overlap in return data from which covariance matrices are estimated creates serial autocorrelation. To avoid this, we decrease the evaluation frequency in two steps. First, we calculate portfolios only year by year (having no data overlap for approaches that work with a one year history, i.e., 1w1y). Second, we analyze minimum-variance portfolios on a biannual frequency (then 1w1y and 1w2y are without overlap). Table 9 shows the estimated pricing impact when cross sectional valuation coefficients are determined annually and biannually. Though biannual evaluation leads to only 12 observations per time series regression for S&P 500 stocks and only 7 observations for STOXX Europe 600 stocks, trends in valuation are robust with respect to changes in the valuation frequency and significance is sustained.

7.2 Resampling

Estimation errors in the sample covariance matrix might influence all portfolio approaches in the same way, such that a common bias affects portfolio selection. To test the robustness of our results, we use a resampling approach. Each time t we determine the sample covariance matrix $\hat{\Sigma}_t$ and determine eigenvectors for all nonzero eigenvalues. We resample the entire return history up to time t for 1000 times under the assumption that

Table 9: Robustness check. Estimated conditional price-to-book and pricing impact for aggregate targets when evaluation is only done annually and biannually, respectively. Though biannual evaluation leads to only 12 observations per time series regression for S&P 500 stocks and only 7 observations for STOXX Europe 600 stocks, trends in valuation are robust with respect to changes in the valuation frequency.

a) S&P 500, annual evaluation

$t = 10\text{-}1989$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	2.0035	0.0615	-7.27 %	2.84 %
m=10	1.8325	0.0812	-15.18 %	3.76 %

$t = 05\text{-}2013$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	4.1535	0.1772	15.78 %	4.94 %
m=10	4.5302	0.2698	26.27 %	7.52 %

b) S&P 500, biannual evaluation

$t = 10\text{-}1989$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	2.0464	0.0871	-5.69 %	4.01 %
m=10	1.9225	0.1173	-11.4 %	5.40 %

$t = 05\text{-}2013$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	4.0357	0.2779	13.59 %	7.82 %
m=10	4.5466	0.4078	27.97 %	11.48 %

c) STOXX Europe 600, annual evaluation

$t = 09\text{-}1999$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	3.1684	0.0469	-5.60 %	1.40 %
m=10	2.7179	0.1961	-19.03 %	5.84 %

$t = 05\text{-}2013$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	3.0760	0.0680	25.02 %	2.76 %
m=10	3.3595	0.1608	36.55 %	6.54 %

d) STOXX Europe 600, biannual evaluation

$t = 09\text{-}1999$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	2.7159	0.0625	-4.38 %	2.20 %
m=10	2.4954	0.1277	-12.14 %	4.49 %

$t = 05\text{-}2013$

	$\hat{\hat{\mathrm{PtB}}}$|minvar	se($\hat{\hat{\mathrm{PtB}}}$|minvar)	price impact	se(price impact)
m=1	3.0817	0.1196	23.46 %	4.79 %
m=10	3.1946	0.1830	27.98 %	7.33 %

Table 10: Robustness check: Estimates $\hat{b}^A$ of the drift annualized coefficient in (4) when return histories are resampled 1000 times every 4 weeks and the coefficients of the minimum-variance indicator $f_{i,t}^A$ are determined in regression (15). Newey-West corrected t-values in parentheses

S&P 500		MM	SMM	AC	SAC	SAV	PC3	PC5
4w5y								
	$w_{\max} = 100\%$	0.0543	0.0544	0.0603	0.0665	0.0493	0.0482	0.0583
		(8.83)	(8.69)	(8.05)	(6.08)	(8.4)	(5.81)	(9.28)
	$w_{\max} = 8\%$	0.0533	0.0525	0.0582	0.0611	0.0487	0.0467	0.0560
		(8.93)	(9.09)	(7.77)	(6.4)	(8.67)	(5.72)	(9.68)
	$w_{\max} = 4\%$	0.0500	0.0483	0.0536	0.0551	0.0449	0.0424	0.0500
		(8.85)	(8.83)	(7.13)	(7.1)	(8.82)	(4.86)	(8.51)

returns are jointly normally distributed and total variation in returns is fully determined by those principal components. For each of the 1000 resampled return histories, we again use our min-var procedures to estimate a covariance matrix and then form the corresponding minimum-variance portfolio. So rather than having an indicator at time t that denotes whether a stock is selected by a particular portfolio approach or not, we compute now a fraction $f_{i,t}^A$ that tells us how many times a particular stock is selected by an approach (again cross-sectionally de-meaned). Then we change regression (2) in the following way,

$$\text{PtBn}_{i,t} = \gamma_{\text{minvar},t}^A f_{i,t}^A + \sum_j \gamma_{j,t}^A c_{j,i,t} + \varepsilon_{i,t}^A. \tag{15}$$

The time-series regression (4) is then performed without change. Due to space restrictions we only report results for S&P 500, returns over four weeks and a history of 5 years in Table 10. We see that results from resampling are highly significant and show a drift in price-to-book which is slightly higher than in the empirical counterpart. Hence, we conclude that our results are robust against estimation errors in the sample covariance matrix $\hat{\Sigma}_S$ and their simultaneous spillover to all portfolio approaches.

7.3 Collinearity in Variables

Valuation effects that are attributed to the min-var indicator might originate from collinearity of the min-var indicator with other variables. So consider the case where minimum-variance portfolio constituents tend to be stocks with high dividend payout. Then we might argue that in fact only high dividend payout is valuation relevant but due to high correlation of the variables, an ordinary least squares method might wrongly impute part of the effect to the min-var indicator. To check that our findings are not driven by collinearity problems, we use the most conservative treatment of this problem. We orthogonalize the min-var indicator with respect to all other variables and, hence, attribute valuation effects in the first place to the control variables and only the strict orthogonal effect that cannot be explained by the controls is assigned to the min-var indicator.

As expected, with orthogonalized min-var indicator, the valuation effects are reduced. For S&P 500 stocks, estimated discount in 10-1989 for aggregate min-var target portfolios with $m = 1, \ldots, 10$ is between -5.18% and -8.58% whereas the original effect is between -6.91% and -15.98%, see Table 8. The premium at 05-2013 is with orthogonalized indicator between 11.76% and 19.39% compared to a range of 13.76% and 22.72% in Table 8. For STOXX Europe 600 stocks and orthogonalized min-var indicator, the estimated discount in 09-1999 for $m = 1, \ldots, 10$ is between -8.88% and -17.23% (compared to -10.52% to -19.96% without orthogonalization). The estimated premium in 05-2013 is between 26.12% and 38.26% (compared to 30.10% to 40.28% without orthogonalization). Standard errors have the same order of magnitude as those reported in Table 8. Thus, even under the most conservative treatment of the collinearity problem, we detect a huge and significant increase in preferences for minimum-variance portfolio constituents.

8 Conclusions

We study a set of 63 different long-only minimum-variance portfolio approaches on US as well as on European stocks, S&P 500 and STOXX Europe 600, respectively. Beside other well known features, like high concentration and low-beta stock selection, we find that different minimum-variance strategies show a high degree of consensus regarding which stocks should be picked for a minimum-variance portfolio. We study valuation anomalies of those minimum-variance stocks, in particular price-to-book anomalies that cannot be explained by industry affiliation, size, or dividend payout policy and find that over the past decades (our history is 23.7 years for US stocks and 13.7 years for Europena stocks), minimum-variance socks' valuation turned from a discount in the 1990s into a rather large premium today. Thus, we infer a shift in preferences in favor of minimum-variance stocks during the last decade(s). We further study aggregate minimum-variance target constituents, i.e., stocks which are picked by several minimum-variance approaches simultaneously. Stocks that are picked by a larger number of different approaches exhibit a more pronounced shift in their valuation and show a larger premium today. All these facts are indications that the high popularity of minimum-variance strategies among investors and the huge inflow into products that offer minimum-variance equity returns show impact on the market valuation of stocks which are favored by those strategies. Investors should take these results into account when assessing future return potential to minimum-variance equity strategies.

Appendix

A.1 Single-Factor Market Model Estimator

As laid out in Section 4.2 we require an estimate of the vector $\hat{\beta}$ being the projection of all stock returns onto the market index returns. At time t, we perform a regression of stock i's most recent N returns on the contemporaneous index returns $r_{M,s}$, where N is given according to the combination of data frequency and rolling window length:

$$r_{i,s} = \alpha_i + \beta_i r_{M,s} + \varepsilon_{i,s}, \qquad s = t, \dots, t - N + 1, \tag{A.1}$$

and repeat this for all stocks currently in the relevant universe ($1 \leq i \leq K$).

Estimating $\hat{\beta}_i$ and $\hat{\sigma}^2_{I,i} = \frac{1}{N-2} \sum \hat{\varepsilon}^2_i$ from (A.1), the underlying assumption that the market is indeed the only common source of risk (in the sense that it causes all of the covariance between stocks) allows the simple representation of the covariance matrix as in (9).

Using the Woodbury identity, the single-factor structure also permits analytical inversion of the covariance matrix (and thus analytical calculation of the weights for long-short portfolios, see e.g. Scherer (2011)). In the long-only case, only stocks with $\hat{\beta}_i$ below a certain critical threshold receive positive weight, see Clarke, de Silva, and Thorley (2011).

A.2 The Average-Correlation Estimator

The construction of the covariance structure in the average-correlation model in Equation (11) is entirely based on the sample covariance matrix $\hat{\Sigma}_S$. The matrix Δ holds the volatilities of each stock, which can be obtained as the square root of the diagonal elements of $\hat{\Sigma}_S$,

$$\hat{\sigma}_i = \sqrt{[\hat{\Sigma}_S]_{ii}}. \tag{A.2}$$

Now we can estimate all pairwise correlations as

$$\hat{\rho}_{ij} = \frac{[\hat{\Sigma}_S]_{ij}}{\hat{\sigma}_i \hat{\sigma}_j}, \qquad 1 \le i \le j \le K, \tag{A.3}$$

and calculate their average, $\hat{\rho}$, to construct the matrix C.

The simple structure of $\hat{\Sigma}_{\text{AC}}$ carries over to a simple representation of its inverse and, thus, allows analytical calculation of the minimum-variance portfolio weights,

$$\hat{\Sigma}_{\text{AC}}^{-1} = \Delta^{-1} C^{-1} \Delta^{-1}, \qquad \Delta^{-1} = \text{diag}(\hat{\sigma}_i^{-1}).$$

As long as $\frac{-1}{K-1} < \hat{\rho} < 1$, C is positive definite and invertible. The inverse of C is then given by (the overall structure is defined by symmetry):

$$C^{-1} = \begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{bmatrix}, \qquad C^{-1} \in \mathbb{R}^{K \times K},$$

$$b = \frac{1}{\hat{\rho}(K-1) - (K-2) - \frac{1}{\hat{\rho}}}, \qquad a = 1 - b\hat{\rho}(K-1).$$

In the unconstrained (long-short) case, the minimum-variance weights result in

$$w^* \propto \mathbf{\Sigma}^{-1}\mathbf{1},$$

and thus with $\hat{\mathbf{\Sigma}}_{\mathrm{AC}}^{-1}$ as

$$w^* = k_1 \hat{\mathbf{\Sigma}}_{\mathrm{AC}}^{-1} \mathbf{1} = k_1 \Delta^{-1} C^{-1} \Delta^{-1} \mathbf{1} = k_1 \Delta^{-1} C^{-1} \boldsymbol{\sigma}^{-1}, \qquad \boldsymbol{\sigma}^{-1} = (\sigma_1^{-1}, \ldots, \sigma_K^{-1})'$$

$$w_i^* = k_1 \sigma_i^{-1} (a \sigma_i^{-1} + \sum_{j \neq i} b \sigma_j^{-1}) = k_2 \sigma_i^{-1} (\sigma_i^{-1} + \sum_{j \neq i} \frac{b}{a} \sigma_j^{-1}), \tag{A.4}$$

with k_1 and $k_2 = k_1 a$ being scaling factors to ensure the weights sum up to one.

To determine which stocks are long and short (i.e., whether $w_i \gtrless 0$) only the sign of the term in the brackets relating stock i's volatility to that of all the other stocks is relevant (note that contrary to the first sum the second includes σ_i^{-1})

$$(\sigma_i^{-1} + \sum_{j \neq i} \frac{b}{a} \sigma_j^{-1}) \propto (\sigma_i^{-1} - \frac{1}{K - 1 + \frac{1}{\hat{\rho}}} \sum_{j=1}^{K} \sigma_j^{-1}). \tag{A.5}$$

One can therefore identify a long-short threshold σ_{LS} for the volatility in the average-correlation model similar to the threshold for β determined by Clarke, de Silva, and Thorley (2011) in the single-factor market model,

$$\sigma_i^{-1} \gtrless \sigma_{LS}^{-1} = \frac{1}{K - 1 + \frac{1}{\hat{\rho}}} \sum_{j=1}^{K} \sigma_j^{-1} \Leftrightarrow w_i^* \gtrless 0$$

.

Furthermore, one sees that the threshold for the (non-inverted) volatility is approximately ($K >> \frac{1}{\hat{\rho}}$) the harmonic mean of all the volatilities.

As Equations (A.4) and (A.5) show, the distance of the inverse volatility to the threshold as well as the value of the inverse volatility itself are the factors determining the weights, given $\hat{\rho}$ and K.

Again as in Clarke, de Silva, and Thorley (2011) a similar argument can be made for the long-only case: Suppose one has already found the set of stocks eligible for inclusion in the long-only minimum variance portfolio. Applying (A.4) to this restricted set with $K' \leq K$ will yield the corresponding weights and have the above threshold criterion fulfilled, i.e.,

$$\sigma_i^{-1} > \sigma_L^{-1} = \frac{1}{K' - 1 + \frac{1}{\hat{\rho}}} \sum_{\sigma_j^{-1} > \sigma_L^{-1}} \sigma_j^{-1}, \qquad K' = \#(\sigma_j^{-1} > \sigma_L^{-1}).$$

To find this set in the first place it is enough to sort the inverse volatilities in a descending order and compare each one with the (scaled) cumulative sum until the condition is no longer met.

A solution for the more general mean-variance portfolio is given by Kwan (2006).

A.3 The Average-Variance Estimator

Similar to the average-correlation estimator, the average-variance estimator can be derived from the sample covariance matrix $\hat{\Sigma}_S$ as

$$\hat{\bar{\sigma}}^2 = \frac{\sum_{i=1}^{K} [\hat{\Sigma}_S]_{ii}}{K}.$$

A.4 Common Features of Minimum-Variance Portfolios

Minimum-variance equity portfolios share a set of common features which are well documented in the literature. We briefly discuss some of the most important common characteristics and document empirical evidence from our different portfolio implementations on the four data sets.

Concentrated portfolios. Long only minimum-variance portfolios tend to be very concentrated. Without constraining the maximum weight of stocks in the portfolio, minimum-variance optimization assigns positive weights only to a low number of stocks. This feature was reported, e.g., by DeMiguel, Garlappi, Nogales, and Uppal (2009) and Clarke, de Silva, and Thorley (2011). Moreover, among selected stocks, weight is concentrated only in a small number of constituents. Our minimum-variance portfolios definitely show this feature on all four datasets. See Table 11 for the sample statistics.

Thus, we clearly see that weights allocated by minimum-variance strategies are not well diversified among all stocks available in the universe and weights are by far not evenly distributed among selected stocks. For S&P 500 and STOXX Europe 600 we also see that over time the concentration of minimum-variance portfolios has definitely increased. Since professional portfolio managers will not agree with such concentrated portfolios, we impose maximum constraints on the portfolios (8% and 4% respectively), so we force the selection mechanism into a broader selection as well as into more uniform weights. Table 11

Table 11: Concentration of long-only minimum-variance portfolios. For long-only portfolios without constraint on the maximum weight we document statistics over the last five years, starting 2008-05-31 until 2013-05-31. **avgNr** is the average number of constituents. **avgNrEff** is the average effective number of stocks in the portfolio, calculated as the average of the inverse Herfindahl index. **avgNr75** is the average of the minimum number of constituents that are needed to represent at least 75% of the portfolio weight.

a) S&P 500

		MM	SMM	AC	SAC	SAV	PC3	PC5
1w1y								
	avgNr	25.22	25.06	14.00	15.63	30.03	26.98	24.09
	avgNrEff	13.75	13.67	8.01	8.18	17.42	13.7	12.34
	avgNr75	10.45	10.38	6.38	6.42	12.95	10.55	9.54
1w2y								
	avgNr	23.29	23.22	14.83	16.32	30.46	24.94	23.31
	avgNrEff	13.03	13	9.01	9.73	17.25	13.68	12.11
	avgNr75	10.09	10.06	6.6	7.29	12.97	10.35	9.35
4w5y								
	avgNr	20.37	20.62	15.42	15.91	34.17	21.08	21.86
	avgNrEff	12.59	12.69	11.09	11.06	19.92	11.88	12.03
	avgNr75	9.09	9.17	8.29	8.25	14.58	8.46	8.65

b) STOXX Europe 600

		MM	SMM	AC	SAC	SAV	PC3	PC5
1w1y								
	avgNr	35.51	35.46	13.29	15.62	34.69	33.97	32.6
	avgNrEff	16.14	16.14	6.32	7.08	19.03	14.93	14.03
	avgNr75	13.34	13.28	4.98	5.74	14.6	12.43	11.97
1w2y								
	avgNr	29.03	28.98	14.03	16.23	31.29	28.18	27.45
	avgNrEff	13.75	13.75	7.00	7.46	16.12	12.86	12.33
	avgNr75	10.91	10.92	5.49	5.74	12.58	10.37	9.98
4w5y								
	avgNr	30.51	27.22	13.71	18.03	33.45	26.68	24.22
	avgNrEff	13.13	12.67	6.40	8.64	16.33	13.60	13.17
	avgNr75	10.38	10.38	5.15	6.88	12.91	10.48	10.06

documents portfolio concentration also for these portfolios.

Low Beta Portfolios As reported by Clarke, de Silva, and Thorley (2011) and Scherer (2011), long only minimum-variance strategies tend to pick low-beta stocks. If the covariance matrix is a single-factor market model estimator (see Section 4.2) Clarke, de Silva, and Thorley (2011) show analytically the existence of a critical upper beta-threshold that discriminates stocks that are included in the long-only minimum-variance portfolio from those which are not. Whereas such a proof does not exist for general estimators of the covariance matrix, we find that this low-beta characteristic is present in a broad set of minimum-variance portfolios.

Low Idiosyncratic Risk

Scherer (2011) reports that long-only minimum-variance portfolios tend to pick stocks that show not only low-beta but also low idiosyncratic risk. We confirm this finding in the broad set of different minimum-variance approaches which we implement.

References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, *The Journal of Finance* 61, 259–299.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009, High idiosyncratic volatility and low returns: International and further U.S. evidence, *Journal of Financial Economics* 91, 1–23.

Baker, Malcom, Brendan Bradley, and Jeffrey Wurgler, 2011, Benchmarks as Limits to Arbitrage: Understanding the Low Volatility Anomaly., *Financial Analysts Journal* 67, 40–54.

Bali, Turan G., and Nusret Cakici, 2008, Idiosyncratic Volatility and the Cross Section of Expected Returns, *Journal of Financial and Quantitative Analysis* 43, 29–58.

Black, Fischer, 1972, Capital Market Equilibrium with Restricted Borrowing, *The Journal of Business* 45, 444–455.

Blackrock Inc., 2013, Monthly Industry Highlight, ETF Landscape-Publications, April 2013.

Clarke, Roger, Harindra de Silva, and Steven Thorley, 2011, Minimum-Variance Portfolio Composition, *Journal of Portfolio Management* 37 (2), 31–45.

DeMiguel, Victor, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal, 2009, A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms, *Management Science* 55, 798–812.

Elton, Edwin J., and Martin J. Gruber, 1973, Estimating the Dependence Structure of Share Prices-Implications for Portfolio Selection, *The Journal of Finance* 28, 1203–1232.

Elton, Edwin J, Martin J Gruber, and Jonathan Spitzer, 2006, Improved estimates of correlation coefficients and their impact on optimum portfolios, *European Financial Management* 12, 303–318.

Frazzini, Andrea, and Lasse Heje Pedersen, 2013, Betting Against Beta, AQR/CBS working paper.

Frost, Peter A, and James E Savarino, 1986, An empirical Bayes approach to efficient portfolio selection, *Journal of Financial and Quantitative Analysis* 21, 293–305.

Fu, Fangjian, 2009, Idiosyncratic risk and the cross-section of expected stock returns, *Journal of Financial Economics* pp. 24–37.

Haugen, Robert A., and A. James Heins, 1975, Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles, *The Journal of Financial and Quantitative Analysis* 10, 775–784.

Huang, Wei, Qianqiu Liu, S. Ghon Rhee, and Liang Zhang, 2010, Return Reversals, Idiosyncratic Risk, and Expected Returns, *The Review of Financial Studies* 23, 147–168.

Jagannathan, Ravi, and Tongshu Ma, 2003, Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps, *The Journal of Finance* 58, 1652–1683.

Jobson, J.D., and B. Korkie, 1980, Estimation for Markowitz Efficient Portfolios, *Journal of the American Statistical Association* 75, 544–554.

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

Kwan, Clarence CY, 2006, Some further analytical properties of the constant correlation model for portfolio selection, *International Journal of Theoretical and Applied Finance* 9, 1071–1091.

Ledoit, Olivier, and Michael Wolf, 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* 10, 603–621.

Ledoit, Olivier, and Michael Wolf, 2004a, Honey, I Shrunk the Sample Covariance Matrix, *Journal of Portfolio Management* pp. 110–119.

Ledoit, Olivier, and Michael Wolf, 2004b, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.

Newey, W.K., and K.D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Newey, W.K., and K.D. West, 1994, Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies* 61, 631–653.

Scherer, Bernd, 2011, A note on the returns from minimum variance investing, *Journal of Empirical Finance* 18, 652–660.

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance* 19, 425–442.

Stein, Charles, 1956, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in *Proceedings of the Third Berkeley symposium on mathematical statistics and probability* vol. 1 pp. 197–206.

Treynor, Jack L., 1962, Toward a Theory of Market Value of Risky Assets, Unpublished manuscript. [”Rough Draft” dated by Treynor to the fall of 1962. A final version was published by Robert Korajczyk (editor) in 1999, in Asset Pricing and Portfolio Performance. A reprint of the text is also included in Treynor (2008)].

Treynor, Jack L., 2008, *Treynor on Institutional Investing.* (John Wiley).