Risk-Based Allocation of Principal Portfolios

Christoph KIND

Preliminary version
April 10, 2013

Abstract: Risk-based asset allocation strategies are mainly used to diversify nominal asset weights. In this paper, we discuss the diversification of risk factors. The analysis is based on the idea of Partovi and Caputo (2004), who use principal component analysis to transform a portfolio into a set of uncorrelated principal portfolios. Risk-based asset allocation strategies can be applied to these uncorrelated sources of risk. A similar route has been taken by Meucci (2009) with his idea of a maximum entropy portfolio. We discuss the relation of this approach with the concept of principal risk parity. Both strategies are backtested against nominal diversification strategies in a multi-asset portfolio. We find no evidence that risk diversification does outperform nominal diversification and discuss possible reasons for this.

Keywords: Asset allocation; portfolio optimization.

JEL Classification Numbers: G11.

*FRANKFURT-TRUST Investmentgesellschaft mbH, Bockenheimer Landstr. 10, D-60323 Frankfurt am Main, Germany, telephone: 0049-69-92050266, e-mail: christoph.kind@frankfurt-trust.de. The author likes to thank Corinna Ament and Andreas Detering for helpful comments.
1 Introduction

In his analysis of sovereign wealth funds, Andrew Ang makes an analogy between risk factors and nutrients. “Factor risk is reflected in different assets just as nutrients are obtained by eating different foods. (...) Assets are bundles of different types of factors just as foods contain different combinations of nutrients” (Ang, 2012, p. 101). The important conclusion to be drawn from this analogy is that nominal diversification may not equal true diversification. A meal consisting of different foods with very similar nutrients might not be considered balanced. In the same way, a nominally diversified portfolio that is dominated by a single risk factor will not be considered well diversified.

There is, however, an important difference between foods and assets. The nutritional content of foods can be measured by accepted standard methods. But there is no consensus on how to identify risk factors. In this paper, we follow Partovi and Caputo (2004) who use principal component analysis to extract uncorrelated synthetic portfolios. These ‘principal portfolios’ as they are called by Partovi and Caputo may be understood as uncorrelated sources of risk.

This paper applies risk-based asset allocation strategies to principal portfolios. We start with an analysis of basic strategies like equal weights, minimum variance and risk parity in the second section. The third section discusses the application of these strategies to principal portfolios. If a full investment constraint on principal weights is imposed, general solutions can be derived for equal weights, minimum variance and risk parity principal portfolios. The optimal principal weights can be transferred back into the original asset space, but this will not lead to fully invested long-only portfolios in general. The fourth section shows how to include constraints on asset weights into the construction of diversified principal portfolios. We also discuss the relation of principal risk parity to the approach proposed by Meucci (2009). In his paper, Meucci measures the diversification of risk contributions with the exponential of the Shannon entropy. This measure is minimized when all portfolio risk is due to one single principal portfolio. The maximum value is reached when risk is homogeneously spread among all principal portfolios. Different opinions about the outperformance potential of risk-based asset allocation strategies can be found in the literature. Qian (2005) argues that
risk parity portfolios should lead to superior risk-adjusted returns because they are not based on forecasts and therefore avoid the often criticized ‘error-maximizing’ nature of mean-variance optimization (see for example Michaud, 1989). But there are also more critical accounts. Chaves et al. (2011) show that the risk parity strategy’s performance can be highly dependent on the investment universe. Lee (2011) questions whether a strategy that is completely ignorant of future returns should consistently outperform the market portfolio.

There is only scarce and controversial evidence on the empirical performance of risk diversification strategies so far. The fifth section of this paper documents a backtest for a multi-asset portfolio. None of the tested diversification strategies was able to outperform the simple equal weights strategy over the whole backtest period. We conclude the paper with a discussion of possible reasons for this underperformance.

2 Risk-Based Asset Allocation

We consider a portfolio of \( N \) assets. The portfolio weight of asset \( i \) is given by \( w_i \). The \( N \)-dimensional column vector \( \mathbf{w} = \{w_1, \ldots, w_N\} \) contains the portfolio weights \( w_i \) of all assets \( i \). Each asset \( i \) is characterized by the standard deviation of its returns \( \sigma_i \). The correlation between the returns of two assets \( i \) and \( j \) is given by the correlation coefficient \( \rho_{ij} \in [-1, 1] \).

The symmetric \( N \)-dimensional matrix \( \Omega \) summarizes the risk structure of the system. The main diagonal of \( \Omega \) contains the variances \( \sigma_i^2 \), the off-diagonal elements correspond to the covariances between assets \( i \) and \( j \) given by \( \sigma_{ij} = \rho_{ij} \sigma_i \sigma_j \). We assume standard deviations and correlations to be given, so that portfolio volatility \( \sigma \) can be expressed as a function of the weights vector:

\[
\sigma(\mathbf{w}) = \sqrt{\mathbf{w}^T \Omega \mathbf{w}} = \sqrt{\sum_i w_i^2 \sigma_i^2 + \sum_{i \neq j} w_i w_j \rho_{ij} \sigma_i \sigma_j}.
\]

The change in risk by a marginal variation in portfolio weights is given by the derivative

\[
\frac{\partial \sigma}{\partial \mathbf{w}} = \frac{\Omega \mathbf{w}}{\sigma}.
\]
The elements of this vector are the marginal contributions to risk (MCR) of each asset $i$:

$$MCR_i = \frac{\partial \sigma}{\partial w_i} = \frac{w_i \sigma_i^2 + \sum_{j \neq i} w_j \rho_{ij} \sigma_i \sigma_j}{\sigma}. \quad (3)$$

The total contribution to risk (TCR) of asset $i$ equals the marginal contribution to risk times the portfolio weight:

$$TCR_i = w_i \frac{\partial \sigma}{\partial w_i} = \frac{w_i^2 \sigma_i^2 + \sum_{j \neq i} w_i w_j \rho_{ij} \sigma_i \sigma_j}{\sigma}. \quad (4)$$

Taking equations 1 and 2 into account, the sum of the TCRs is equal to portfolio risk:

$$\sum_{i=1}^{N} TCR_i = \mathbf{w}' \frac{\partial \sigma}{\partial \mathbf{w}} = \sigma. \quad (5)$$

Equation 5 also follows from Euler’s theorem for homogeneous functions of degree 1.

Finally, we define the percentage contribution to risk $PCR$ of an asset $i$ as

$$PCR_i \equiv \frac{TCR_i}{\sigma} \frac{\sigma}{\sigma_i} = \frac{w_i}{\sigma} \frac{\partial \sigma}{\partial w_i}. \quad (6)$$

$PCR_i$ is the elasticity of portfolio volatility with respect to a small change in the weight of asset $i$. The sum of all percentage contributions to risk is equal to one.

The equal weights (EW) portfolio is defined by identical weights $w_i = w_j$ for all $i, j$. With $\mathbf{1}$ as a $N$-dimensional column vector of ones, the sum of portfolio weights can be normalized to one with the budget restriction $\mathbf{w}' \mathbf{1} = 1$. The weights of the EW-portfolio are then given by $w_i = 1/N$ for all $i$.

It is easy to show that the variance of the EW-portfolio is equal to

$$\sigma_{EW}^2 = \frac{1}{N} \bar{\sigma}^2 + \frac{N-1}{N} \bar{\sigma}_{ij}, \quad (7)$$

where $\bar{\sigma}^2$ and $\bar{\sigma}_{ij}$ are the arithmetic averages of all variances and all covariances respectively.

The weights of the minimum variance (MV) portfolio can be derived by solving

$$\min_w \sigma^2(w) = \mathbf{w}' \Omega \mathbf{w}$$

s.t. $\mathbf{w}' \mathbf{1} = 1. \quad (8)$$
With unlimited short-sales, the weights of the global minimum variance portfolio are

\[ w_{MV} = \frac{\Omega^{-1}1}{1'\Omega^{-1}1}. \]  \hspace{1cm} (9)

and the variance of the MV-portfolio is equal to

\[ \sigma^2_{MV} = w_{MV}'\Omega w_{MV} = \frac{1}{1'\Omega^{-1}1}. \]  \hspace{1cm} (10)

Evaluating the derivative given in equation 2 at minimum variance and using equations 9 and 10, we get the following relation:

\[ \left| \frac{\partial \sigma}{\partial w} \right|_{\sigma = \sigma_{MV}} = \sigma_{MV}1. \]  \hspace{1cm} (11)

The minimum variance portfolio is characterized by the fact that marginal contributions to risk are identical across assets. The marginal contribution to risk for each asset equals the portfolio risk.

Minimum variance optimization forces the MCRs to be identical, but not the TCRs. If an investor thinks in terms of risk budgets, it might be more natural to force the TCRs to be equal across assets. This is the idea of a risk parity (RP) portfolio, which is defined by

\[ w_i \frac{\partial \sigma(w)}{\partial w_i} = w_j \frac{\partial \sigma(w)}{\partial w_j} \text{ for all } i, j. \]  \hspace{1cm} (12)

Under the budget constraint, it follows from equation 5 that the TCR for every asset \( i \) is equal to \( \sigma/N \).

Even with unlimited short-sales, there is no general solution for the weights of a risk parity portfolio. With respect to portfolio risk, Maillard et al. (2010) show that the risk of a RP-portfolio is bounded from below by the risk of a MV-portfolio and from above by the risk of an EW-portfolio:

\[ \sigma_{MV} \leq \sigma_{RP} \leq \sigma_{EW}. \]  \hspace{1cm} (13)

The risk parity portfolio may be viewed as an attempt to reconcile variance minimization and nominal diversification. The EW-portfolio maximizes the nominal diversification, but pays no attention to volatility. The MV-portfolio minimizes volatility, but ignores nominal diversification. The RP-portfolio may be understood as a form of variance minimization subject to a constraint of sufficient nominal diversification (see again Maillard et al., 2010).
3 Principal Portfolios

Real-life portfolios are usually characterized by non-zero covariances $\sigma_{ij}$ for at least some $i$ and $j$. The idea of a principal portfolio is to find a change of base of the original asset space that generates a new set of synthetic assets with zero covariances $s_{\mu\nu}$ for all combinations of $\mu$ and $\nu$. We know from basic linear algebra that any real-valued symmetric $N \times N$ matrix is orthogonally diagonalizable and possesses a complete set of orthogonal eigenvectors. If the matrix is positive semi-definite, all eigenvalues will be non-negative.

The eigenvector equations for the covariance matrix $\Sigma$ are given by

$$\Omega e^\mu = \lambda_\mu e^\mu \text{ for } \mu = 1, \ldots, N,$$

(14)

where $e$ denotes the $N$-dimensional eigenvector and $\lambda$ the scalar eigenvalue.

We can now define a $N$-dimensional square matrix $E \equiv \{e^1, \ldots, e^N\}$ whose columns are the $N$ eigenvectors $e^\mu$. With $\Lambda \equiv \text{diag} \{\lambda_1, \ldots, \lambda_N\}$ and $\circ$ symbolizing the Hadamard product, the following relations hold:

$$E^{-1}\Omega E = \Lambda,$$

(15)

$$\begin{bmatrix}
\sigma_1^2 \\
\sigma_2^2 \\
\vdots \\
\sigma_N^2 \\
\end{bmatrix} = (E \circ (E^{-1})') \begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\vdots \\
\lambda_N \\
\end{bmatrix}.$$

(16)

Eigenvectors are only defined up to a non-zero scalar multiple. It is common practice to standardize the eigenvectors to unit length so that $(e^\mu)' e^{\nu} = \delta_{\mu\nu}$, where $\delta_{\mu\nu}$ is the Kronecker delta. With this standardization, the matrix of eigenvectors $E$ is said to be orthonormal. In comparison to an orthogonal matrix, $E$ then has the additional characteristics that $E^{-1} = E'$ and the sum of squares of any row and column is equal to unity (Jackson, 2005, p. 439). With an orthonormal base, it follows from equation 16 that each asset’s variance equals a weighted sum of the eigenvalues:

$$\sigma_i^2 = \sum_{\mu=1}^{N} \lambda_\mu (e_i^\mu)^2, \quad \sum_{\mu=1}^{N} (e_i^\mu)^2 = 1 \text{ for all } i.$$

(17)

We now define a new $N$-dimensional column vector $p$ as

$$p \equiv E^{-1}w.$$

(18)
Substituting \( w \) by \( Ep \) in equation 1 gives

\[
\sigma^2 = (Ep)' \Omega Ep = p'E' \Omega Ep = p' \Lambda p.
\]

(19)

Portfolio variance can either be described by the quadratic form \( w' \Omega w \) or the quadratic form \( p' \Lambda p \). We can use the transformations shown in equations 15 and 19 to translate the original system characterized by a weights vector \( w \) and a covariance matrix \( \Omega \) into a principal system characterized by a principal weights vector \( p \) and a diagonal covariance matrix of the principal portfolios \( \Lambda \). This change of base simplifies the covariance structure considerably. On the other hand, much of the complexity of \( \Omega \) is now shifted into the principal weights vector \( p \) whose elements can be negative and will in general not sum to one.

Next, we discuss the allocation of principal portfolios when a full investment constraint \( p' 1 = 1 \) is imposed. The resulting optimal weights \( p^* \) will in general not be equal to the transposed optimal weights \( p(w^*) \). Since principal portfolios are characterized by zero covariances, the results are simplifications of the results derived in section 2.

The portfolio given by \( p_{EW} \equiv \{1/N, ..., 1/N\} \) is the equal weights portfolio of fully invested principals. An equal weighting of principals amounts to a very bad diversification of risk factors. With orthonormal eigenvectors, the variance of each principal portfolio \( s^2_\mu \) is given by the corresponding eigenvalue \( \lambda_\mu \). The eigensystem of a typical covariance matrix is dominated by a very large eigenvalue, and the dispersion of principal volatilities is much larger than the dispersion of the original volatilities.

The variance of the principal equal weights portfolio can be derived from equation 7 as

\[
s^2_{EW} = \frac{1}{N} \overline{s^2},
\]

(20)

where \( \overline{s^2} \) is the arithmetic average of the principal portfolio variances.

We now turn to the minimum variance portfolio. The weights of the MV-principal portfolio can be easily derived from equation 9 as

\[
p_{\mu, MV} = \frac{1/s^2_\mu}{\sum_\mu 1/s^2_\mu}.
\]

(21)

The variance of this portfolio follows from equation 10 as

\[
s^2_{MV} = \frac{1}{\sum_\mu 1/s^2_\mu}.
\]

(22)
Table 1: Notations in asset and principal space

<table>
<thead>
<tr>
<th></th>
<th>Assets</th>
<th>Principals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights and indices</td>
<td>$w_i, w_j$</td>
<td>$p_\mu, p_\nu$</td>
</tr>
<tr>
<td>Portfolio variance</td>
<td>$\sigma^2 = \sum_i \sum_j w_i w_j \rho_{i,j} \sigma_i \sigma_j$</td>
<td>$s^2 = \sum_{\mu} p^2_\mu s_\mu^2$</td>
</tr>
<tr>
<td>Marginal contribution to risk</td>
<td>$\frac{w_i \sigma_i^2 + \sum_j w_j \rho_{i,j} \sigma_i \sigma_j}{\sigma^2}$</td>
<td>$\frac{p^2_\mu s_\mu^2}{s^2}$</td>
</tr>
<tr>
<td>Total contribution to risk</td>
<td>$\frac{w_i \sigma_i^2 + \sum_j w_j \rho_{i,j} \sigma_i \sigma_j}{\sigma^2}$</td>
<td>$\frac{p^2_\mu s_\mu^2}{s^2}$</td>
</tr>
<tr>
<td>Percentage contribution to risk</td>
<td>$\frac{w_i \sigma_i^2 + \sum_j w_j \rho_{i,j} \sigma_i \sigma_j}{\sigma^2}$</td>
<td>$\frac{p^2_\mu s_\mu^2}{s^2}$</td>
</tr>
</tbody>
</table>

Next, we look at risk parity portfolios. We noted earlier that no general solution for the weights of a risk parity portfolio exists. Maillard et al. (2010) derive a solution for the special case when correlations for every couple of variables are equal, that is $\rho_{i,j} = \rho$ for all $i, j$. Since $\rho_{\mu \nu} = 0$ for all $\mu, \nu$, we can apply this solution to the case of principal portfolios to get

$$p_{\mu,RP} = \frac{1/s_\mu}{\sum_\mu 1/s_\mu}. \quad (23)$$

Plugging these weights into the equation that defines portfolio variance, we are now able to derive a general solution for the variance of a RP-portfolio:

$$s_{RP}^2 = N \left( \sum_\mu \frac{1}{s_\mu} \right)^2. \quad (24)$$

The variances derived in equations 20 and 22 are special cases of the variances derived in the second section. The inequality relation shown in equation 13 therefore also holds for principal portfolios, that is

$$s_{MV} \leq s_{RP} \leq s_{EW}. \quad (25)$$

We show in the appendix how this relation can be proved using the Cauchy-Schwarz inequality.

A synopsis of the basic concepts for the risk-based allocation of assets and principals concludes this section. Table 1 summarizes the different notations and identities.
4 Principal Risk Parity and Maximum Entropy

The analysis of principal portfolios is motivated by the fact that principals represent risk factors. If the asset allocation process is based on risk budgets, it makes sense to apply risk-based portfolio construction techniques not to the original assets, but to principals. The general solutions $\mathbf{p}^*$ for principal MV and principal RP portfolios given in equations 21 and 23 can be transferred back into the asset space by the change of base $\mathbf{w} = \mathbf{E}\mathbf{p}^*$, but this will in general not result in fully invested portfolios with non-negative weights for every asset. For that reason, we have to build optimization problems that allow for non-negativity and full-investment constraints on the original asset weights.

The general form of this optimization problem is

$$
\begin{align*}
\min_{\mathbf{w}} & \quad f(\mathbf{p}(\mathbf{w})) \\
\text{s.t.} & \quad \mathbf{w}'1 = 1.
\end{align*}
$$

(26)

Short positions may be ruled out by adding non-negativity constraints $w_i \geq 0$ for $i = 1, ..., N$ to this general problem.

In case of the equal-weights principal portfolio, the objective function can be stated as

$$
\begin{align*}
\text{f}_{\text{EW}} &= \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} (p_{\mu} - p_{\nu})^2.
\end{align*}
$$

(27)

It is obvious that the solution to this problem will in general deviate from the unrestricted equal-weights vector $\mathbf{w}_{\text{EW}} \equiv \{1/N, ..., 1/N\}$.

Next, we look at the principal mean variance portfolio where the objective function is given by

$$
\begin{align*}
\text{f}_{\text{MV}} &= \mathbf{p}'\mathbf{\Lambda}\mathbf{p}.
\end{align*}
$$

(28)

Because of equation 19, this problem is equivalent to 8 and has the same solution vector $\mathbf{w}_{\text{MV}}$ given in equation 9.

The principal risk parity portfolio is characterized by equal total contributions to risk $TCR_\mu(\mathbf{w})$ for all principals $\mu$. Following Maillard et al. (2010), we can state the objective function of this optimization problem as

$$
\begin{align*}
\text{f}_{\text{RP}} &= \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} (TCR_\mu - TCR_\nu)^2.
\end{align*}
$$

(29)

An alternative approach to the problem of diversifying the risks of principal portfolios has been proposed by Meucci (2009). He uses the dispersion of
PCR_\mu$, which he calls *diversification distribution*, as a measure of diversification. In principal space, percentage contributions to risk are non-negative, and sum to one by definition. So they can be treated as probabilities with an unknown distribution. The maximum entropy principle suggests that, from among all probability distributions satisfying given constraints, we choose the one of maximum entropy, i.e., the one that is closest to uniform. Meucci (2009) uses the exponential of the Shannon entropy, which can be understood as the number of effective (i.e., independent) bets in a portfolio.\(^1\) The objective function of the *maximum entropy* portfolio is then given by

\[
 f_{ME} = - \exp \left( - \sum_{\mu=1}^{N} PCR_\mu \ln(PCR_\mu) \right). \tag{30}
\]

A uniform diversification distribution will minimize equation 30, meaning that total and percentage contributions to risk are identical across principals. Without any additional constraints, the risk parity problem described in equation 29 and the maximum entropy problem described in equation 30 have the same solution. On theoretical grounds, the maximum entropy approach might be preferred because it is widely used in information theory and natural sciences. In biology, the exponential of the Shannon entropy is used to measure the effective number of species in a population, which is obviously similar to the notion of effective bets in a portfolio.

It should be noted that Meucci (2009) uses conditional principal component analysis to compute the diversification distribution. This method allows to incorporate rebalancing constraints. We do not follow this approach here but calculate the diversification distribution with standard principal component analysis.

The two ways of diversifying principal portfolio risk shown in equations 29 and 30 may be understood as special cases of a more general approach. The risk parity principle is based on equal risk budgets for every asset or risk factor. In the case of principal portfolios, risk budgets \(b_\mu\) are identical for all principals \(\mu\). But it is also possible to allocate a specific risk budget to each principal portfolio as in Bruder and Roncalli (2012). The objective function of the more general *risk budgeting* optimization is given by

\[
 f_{RB} = \sum_{\mu=1}^{N} (TCR_\mu - b_\mu)^2. \tag{31}
\]

\(^1\)The exponential of the Shannon entropy \(\exp(-\sum_i x_i \log x_i)\), where \(0 \leq x_i \leq 1\) and \(\sum_i x_i = 1\), should not be confused with the exponential entropy that is given by \(\sum_i x_i \exp (1 - x_i)\) as in Pal and Pal (1992).
The same argument can be made for the maximum entropy approach. The Shannon entropy tries to fit an unknown probability distribution satisfying given constraints as close as possible to the uniform distribution. Bera and Park (2008) argue that this is a special case of the more general cross-entropy approach where an unknown distribution can be fitted to any desired reference distribution.

5 From Theory to Practice

The original risk parity approach aims at a nominal diversification of asset weights. Since assets may be understood as bundles of risk factors, nominal diversification is no guarantee for a diversification of risk factors. Applying the principles of risk parity portfolios to risk factors leads to risk diversification. We outlined the theoretical basis for this kind of diversification in the last section.

The question remains why an investor should care about diversification anyway. The classical Markowitz portfolio theory is often associated with the idea of diversification, but mean-variance optimization is in fact not based on any measure of diversification. In addition, there is no consensus definition of diversification in financial research.

The case for nominal diversification can be based on the well-known problems of classical Markowitz optimization. Many studies have shown that mean-variance portfolios based on sample estimates of means and variances perform badly. One way of dealing with that problem is to constrain nominal asset weights (see for example DeMiguel et al., 2009a). These approaches try to find the optimal nominal diversification in contrast to the maximum nominal diversification given by the equal-weights portfolio. The evidence that optimal nominal diversification outperforms maximum nominal diversification is mixed (see for example DeMiguel et al., 2009b). In addition, risk parity and mean variance strategies are also sensitive to covariance estimation errors.

When we turn to risk diversification, the argument is different. Building on the approach by Meucci (2009), we have shown in the last section that diversifying risk factors is equal to maximizing the number of independent bets in a portfolio. We can connect this idea with the fundamental law of active management (Grinold and Kahn, 2000), which expresses the information ratio as a positive function of the information coefficient and the number of independent bets. Increasing risk diversification should lead to
higher information ratios and better performance.
The risk diversification approach can be backtested across asset classes and within an asset class (for the latter, please refer to Lohre et al., 2012a). The performance of risk diversification strategies across asset classes has been tested by Poddig and Unger (2012) and Lohre et al. (2012b), and they come to different conclusions. Lohre et al. (2012b) analyze a multi-asset portfolio consisting of government bonds, developed markets equities, emerging market equities, corporate bonds and a commodity index with monthly data covering the period from December 1987 to September 2011. In a rolling window backtest, the risk-adjusted performance of the maximum entropy approach dominates nominal diversification strategies. Poddig and Unger (2012) use a Monte Carlo simulation and a global portfolio of major asset classes to study the maximum entropy approach against equal weights, minimum variance, risk parity and classical Markowitz strategies. The backtest is based on daily data and monthly rebalancing for a period from January 1995 to July 2010. In this study, the maximum entropy approach “delivers the worst performance in all circumstances” (Poddig and Unger, 2012, p. 393).

Our backtest universe comprises eight generic future contracts covering four asset classes (see table 2). We use monthly US-Dollar returns over the period February 1990 to January 2013. Using futures to cover asset classes has the advantage that short positions can be implemented without additional costs. Apart from margin payments, no funding is needed for a futures portfolio. Under the assumption that the underlying cash of a funded futures strategy is invested at the risk-free rate, the Sharpe ratio of such a strategy simply equals the return of the futures portfolio divided by its standard deviation. These numbers are shown in table 2. It should be noted that the risk-return relations over the whole sample period do not show any kind of ‘volatility
Table 3: Correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>TY</th>
<th>SP</th>
<th>Z</th>
<th>C</th>
<th>QS</th>
<th>JY</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>1.00</td>
<td>0.95</td>
<td>-0.08</td>
<td>-0.09</td>
<td>0.01</td>
<td>-0.15</td>
<td>0.16</td>
<td>-0.10</td>
</tr>
<tr>
<td>TY</td>
<td>0.95</td>
<td>1.00</td>
<td>-0.08</td>
<td>-0.10</td>
<td>0.02</td>
<td>-0.12</td>
<td>0.21</td>
<td>-0.08</td>
</tr>
<tr>
<td>SP</td>
<td>-0.08</td>
<td>-0.08</td>
<td>1.00</td>
<td>0.79</td>
<td>0.25</td>
<td>0.04</td>
<td>0.00</td>
<td>0.51</td>
</tr>
<tr>
<td>Z</td>
<td>-0.09</td>
<td>-0.10</td>
<td>0.79</td>
<td>1.00</td>
<td>0.19</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.39</td>
</tr>
<tr>
<td>C</td>
<td>0.01</td>
<td>0.02</td>
<td>0.25</td>
<td>0.19</td>
<td>1.00</td>
<td>0.02</td>
<td>0.04</td>
<td>0.20</td>
</tr>
<tr>
<td>QS</td>
<td>-0.15</td>
<td>-0.12</td>
<td>0.04</td>
<td>0.04</td>
<td>0.02</td>
<td>1.00</td>
<td>0.05</td>
<td>0.29</td>
</tr>
<tr>
<td>JY</td>
<td>0.16</td>
<td>0.21</td>
<td>0.00</td>
<td>-0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td>CD</td>
<td>-0.10</td>
<td>-0.08</td>
<td>0.51</td>
<td>0.39</td>
<td>0.20</td>
<td>0.29</td>
<td>0.01</td>
<td>1.00</td>
</tr>
</tbody>
</table>

anomaly’. Assets with higher risks are in general associated with higher returns. Table 3 summarizes the correlation structure over the complete sample period. Correlations are generally low across asset classes. Significant positive correlations can be observed between the two equity indices and between the two bond indices.

We backtest the strategies over a rolling time window of 60 months as in Lohre et al. (2012b). Since five years of history are needed, the back test period starts in February 1995. Weights are adjusted monthly. To prevent transaction costs from becoming too high, weights are smoothed using an exponential moving average with a decay parameter set to 3 months. The principal portfolios are calculated on the basis of the standard eigensystem of the covariance matrix.

We evaluate the performance of each strategy with a set of statistics. The Sharpe ratio is computed from annualized risks and returns. As measures of downside risks we report maximum drawdown (MDD), historical value-at-risk (VaR) and historical conditional value-at-risk (CVaR). The VaR is calculated for a 95 percent confidence level. To get an impression of the transaction costs involved, we calculate the average monthly turnover over $T$ periods as defined in DeMiguel et al. (2009b) as

$$\text{Turnover} = \frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{i=1}^{N} |w_{i,t+1} - w_{i,t}|,$$  \hspace{1cm} (32)

in which $w_{i,t+1}$ is the portfolio weight of asset $i$ at time $t$, and $w_{i,t}$ is the portfolio weight before rebalancing. The value of this statistic equals the average monthly amount of buy and sell transactions as a percentage of the portfolio value. It is clear that even for the equal-weights portfolio, turnover will be positive because movements of market prices lead to a deviation of
Table 4: Backtest results

<table>
<thead>
<tr>
<th></th>
<th>Assets</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EW</td>
<td>RP</td>
<td>PRP</td>
<td>ME</td>
<td>PRPLS</td>
<td>MELS</td>
<td>PRB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return (%)</td>
<td>5.75</td>
<td>3.61</td>
<td>3.75</td>
<td>3.97</td>
<td>2.37</td>
<td>0.17</td>
<td>6.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk (%)</td>
<td>8.44</td>
<td>6.07</td>
<td>6.49</td>
<td>6.98</td>
<td>5.70</td>
<td>7.47</td>
<td>10.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe</td>
<td>0.68</td>
<td>0.59</td>
<td>0.58</td>
<td>0.57</td>
<td>0.42</td>
<td>0.02</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD (%)</td>
<td>28.3</td>
<td>17.5</td>
<td>19.5</td>
<td>19.7</td>
<td>11.4</td>
<td>21.4</td>
<td>22.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VaR (95%)</td>
<td>-3.5</td>
<td>-2.7</td>
<td>-2.7</td>
<td>-3.0</td>
<td>-2.6</td>
<td>-3.6</td>
<td>-4.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVaR (95%)</td>
<td>-5.3</td>
<td>-3.8</td>
<td>-4.2</td>
<td>-4.4</td>
<td>-3.6</td>
<td>-5.4</td>
<td>-6.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turnover (%)</td>
<td>0.40</td>
<td>1.51</td>
<td>9.49</td>
<td>10.47</td>
<td>10.83</td>
<td>35.72</td>
<td>9.20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

portfolio weights from equal weights.

Table 4 reports the backtest statistics for asset and principal strategies. Two nominal diversification strategies have been tested: Equal weights (EW) and risk parity (RP). Compared to equal weights, risk parity lowers portfolio volatility and downside risks, but leads to a lower Sharpe ratio and higher turnover.

Also two risk diversification strategies have been tested: principal risk parity (PRP) and maximum entropy (ME). PRP leads to higher volatility and a slight deterioration in downside risks. The Sharpe ratio is almost identical to RP, but turnover is much higher. ME raises volatility further and increases downside risks and turnover. The Sharpe ratio declines slightly. PRP slightly dominates ME, but both strategies fail to improve on the results of nominal diversification with respect to risk and risk-adjusted return.

What are the reasons for the disappointing performance of risk diversification strategies? Poddig and Unger (2012) assume that the calculation of principal portfolios suffers from estimation errors. We now discuss two alternative explanations.

First, it could be argued that long-only constraints have a negative impact on performance. Long-only constrained portfolios will not only fail to exactly replicate principals, but will also lead to a set of principals that is not completely uncorrelated. The solid blue and red lines in figure 1 show the percentage degree of diversification over the backtest period for the PRP and ME strategies. Diversification does indeed never reach the maximum level and declines over time. When we allow the individual asset weights to fluctuate between -100 and +100%, both risk diversification strategies can be backtested as long/short-strategies. We name these strategies PRPLS and MELS and show their diversification over time as dashed red and blue lines in figure 1. Diversification now indeed reaches the maximum level almost all
of the time. But, as can be seen from table 4, performance does not improve at all. PRPLS reduces volatility and downside risks, but risk-adjusted returns fall markedly and turnover rises further. The performance of MELS is dismal: Turnover rises massively, and the Sharpe ratio collapses. The bad performance of long-short strategies can be seen as further evidence for the estimation problem that is magnified when short sales are allowed. But it also raises the question whether maximum diversification is really desirable. The second possible explanation for the bad performance of risk diversification could be based on the claim that an optimal level of diversification leads to better results. Diversification is neither a sufficient nor a necessary condition for performance. We know from the ‘fundamental law of active management’ that a diversification of bets will not improve performance when these bets are ‘bad’, i.e. have a low information coefficient. On the other hand, concentrated bets may lead to good performance if those few bets have attractive pay-offs. To get an impression of the bets associated with principal portfolios, we take up the idea of risk budgets developed at the end of section 4. The risk budgeting approach allows for individual risk budgets. By setting the risk budget for one principal close to one and for all other principals close
Table 5: Principal portfolios

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return (%)</td>
<td>14.0</td>
<td>3.5</td>
<td>4.7</td>
<td>1.8</td>
<td>3.2</td>
<td>3.5</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Risk (%)</td>
<td>24.5</td>
<td>18.7</td>
<td>12.1</td>
<td>8.2</td>
<td>7.8</td>
<td>7.2</td>
<td>7.2</td>
<td>7.4</td>
</tr>
<tr>
<td>Sharpe</td>
<td>0.57</td>
<td>0.19</td>
<td>0.39</td>
<td>0.21</td>
<td>0.41</td>
<td>0.48</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td>MDD (%)</td>
<td>42.7</td>
<td>55.1</td>
<td>41.1</td>
<td>24.9</td>
<td>13.1</td>
<td>22.3</td>
<td>17.8</td>
<td>17.1</td>
</tr>
<tr>
<td>VaR (95%)</td>
<td>-10.2</td>
<td>-8.4</td>
<td>-5.7</td>
<td>-3.8</td>
<td>-3.1</td>
<td>-3.3</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>CVaR (95%)</td>
<td>-13.3</td>
<td>-12.7</td>
<td>-8.4</td>
<td>-4.5</td>
<td>-4.8</td>
<td>-4.3</td>
<td>-4.5</td>
<td>-4.5</td>
</tr>
<tr>
<td>Turnover (%)</td>
<td>6.0</td>
<td>6.9</td>
<td>7.6</td>
<td>7.6</td>
<td>15.2</td>
<td>19.3</td>
<td>18.1</td>
<td>24.4</td>
</tr>
</tbody>
</table>

to zero, we can use equation 31 to construct ‘principal-mimicking portfolios’ (see Bruder and Roncalli (2012) for the subtleties of the risk budgeting optimization process). The resulting portfolios are reported in table 5.

Since eigenvalues are ordered according to size, it is not surprising that principal volatilities are generally decreasing. Note that turnover rises significantly for low volatility principals. While high volatility principals are concentrated in the most volatile assets, low volatility principal portfolios are more diversified and need dynamic rebalancing. There is also a marked difference between the Sharpe ratios of the eight principal portfolios. It is worth noting that none of these Sharpe ratios beats the one of the equal weights portfolio.

Table 5 suggests that a ‘barbell-strategy’ which combines the most and the least volatile principals might improve backtest results. We tested a portfolio where the first factor is given a risk budget of 75% and the last factor a risk budget of 25%. The performance of this principal risk budgeting portfolio PRB is reported in table 4, and the diversification over time is shown by the green line in figure 1. The Sharpe ratio of the PRB portfolio is better than the one of the standard risk parity approach, but downside risk measures are even higher than the ones for equal weights. The diversification of the risk budget portfolio is around 25% most of the time. Compared to the diversification of the equal weights-portfolio that is given by the black line in figure 1, this is a rather low level of diversification.

What are the conclusions? There is no reason to expect a general outperformance from risk diversification strategies. These strategies are vulnerable to estimation errors of the principal portfolios. In addition, diversifying bad bets will not lead to outperformance. It also seems that it is better to look for optimal diversification instead of maximum diversification. There is not much to be gained from diversification in an already well-diversified universe.
6 Conclusions

An unbalanced diet has too much of some nutrients and not enough of others. In the same way, a badly diversified portfolio has too much of some risks and not enough of others. Risk-based asset allocation strategies try to tackle that problem, but they usually focus on nominal diversification. In this paper, we applied risk-based asset allocation strategies to principal portfolios. Principal portfolios reflect the underlying risk factors that make up a portfolio.

Using principal component analysis to construct uncorrelated principal portfolios, general solutions for equal weights, minimum variance and risk parity principal portfolios were derived. If full investment constraints are imposed not on principal weights but on asset weights, no general solutions exist. We showed how to construct optimization problems in these cases and discussed the relation of principal risk parity with the maximum entropy approach proposed by Meucci (2009).

The last section of the paper documented a backtest of nominal and risk diversification strategies applied to a multi-asset portfolio. No diversification strategy was able to outperform the simple equal weights strategy over the whole backtest period. We hinted at possible explanations for this result. Apart from difficulties with the estimation of principle portfolios, the quality of the diversified bets is important. Diversifying bad bets will not lead to outperformance.
References


A Appendix

The inequalities in equation 25 may also be stated in terms of variances, that is

\[ s_{MV}^2 \leq s_{RP}^2 \leq s_{EW}^2. \]  

To prove this relation, we use two versions of the Cauchy-Schwarz inequality. If \( x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{R} \), the Cauchy-Schwarz inequality states that

\[
\left( \sum_{\mu=1}^{N} x_\mu y_\mu \right)^2 \leq \left( \sum_{\mu=1}^{N} x_\mu^2 \right) \left( \sum_{\mu=1}^{N} y_\mu^2 \right). \]  

With \( a \in \mathbb{R} \) and \( x_1 = s_1^a, \ldots, x_N = s_N^a, y_1 = \frac{1}{s_1^a}, \ldots, y_N = \frac{1}{s_N^a} \), this inequality becomes

\[
N^2 \leq \left( \sum_{\mu=1}^{N} s_\mu^{2a} \right) \left( \sum_{\mu=1}^{N} \frac{1}{s_\mu^{2a}} \right). \]  

With \( x_1 = s_1^a, \ldots, x_N = s_N^a, y_1 = 1, \ldots, y_N = 1 \), this inequality becomes

\[
\left( \sum_{\mu=1}^{N} s_\mu^a \right)^2 \leq \left( \sum_{\mu=1}^{N} s_\mu^{2a} \right) N. \]  

We first want to prove that \( s_{RP}^2 \geq s_{MV}^2 \). From equations 22 and 24, this amounts to showing that

\[
N \left( \sum_{\mu} s_\mu^{-1} \right)^2 \geq \frac{1}{\sum_{\mu} s_\mu^2}. \]  

The proof of this inequality follows immediately from equation 36 with \( a = -1 \).

Next, we want to prove the inequality \( s_{EW}^2 \geq s_{MV}^2 \). From equations 20 and 22, one has to show that

\[
\sum_{\mu} s_\mu^2 \frac{1}{N^2} \geq \frac{1}{\sum_{\mu} s_\mu^2}. \]  

The proof of this inequality follows immediately from equation 35 with \( a = 1 \).

Finally, we want to prove that \( s_{EW}^2 \geq s_{RP}^2 \). Using equations 20 and 24, we can state this relation as

\[
\sum_{\mu} s_\mu^2 \frac{1}{N^2} \geq \frac{N}{(\sum_{\mu} s_\mu^{-1})^2}. \]  

19
which is equivalent to the inequality

$$\sum_{\mu} s_{\mu}^2 \left( \sum_{\mu} \frac{1}{s_{\mu}} \right)^2 \geq N^3. \quad (40)$$

By expanding the left-hand side of 40, one gets

$$\frac{\sum_{\mu} s_{\mu}^2}{\left( \sum_{\mu} s_{\mu} \right)^2} \left( \sum_{\mu} \frac{1}{s_{\mu}} \sum_{\mu} s_{\mu} \right)^2 \geq N^3. \quad (41)$$

The first term on the left-hand side of equation 41 is greater or equal to $N^{-1}$ according to equation 36 with $a = 1$. The term in brackets on the left-hand side of equation 41 is greater or equal to $N^2$ according to equation 35 with $a = 1/2$. Therefore, the left-hand side of equation 41 is greater or equal to

$$N^{-1} (N^2)^2 = N^3,$$

which completes the proof.