Portfolio Allocation with Skewness Risk: A Practical Guide*

Edmond Lezmi
Quantitative Research
Amundi Asset Management, Paris
edmond.lezmi@amundi.com

Hassan Malongo
Quantitative Research
Amundi Asset Management, Paris
hassan.malongo@amundi.com

Thierry Roncalli
Quantitative Research
Amundi Asset Management, Paris
thierry.roncalli@amundi.com

Raphaël Sobotka
Multi-Asset Management
Amundi Asset Management, Paris
raphael.sobotka@amundi.com

February 2019 (First version: June 2018)

Abstract

In this article, we show how to take into account skewness risk in portfolio allocation. Until recently, this issue has been seen as a purely statistical problem, since skewness corresponds to the third statistical moment of a probability distribution. However, in finance, the concept of skewness is more related to extreme events that produce portfolio losses. More precisely, the skewness measures the outcome resulting from bad times and adverse scenarios in financial markets. Based on this interpretation of the skewness risk, we focus on two approaches that are closely connected. The first one is based on the Gaussian mixture model with two regimes: a 'normal' regime and a 'turbulent' regime. The second approach directly incorporates a stress scenario using jump-diffusion modeling. This second approach can be seen as a special case of the first approach. However, it has the advantage of being clearer and more in line with the experience of professionals in financial markets: skewness is due to negative jumps in asset prices. After presenting the mathematical framework, we analyze an investment portfolio that mixes risk premia, more specifically risk parity, momentum and carry strategies. We show that traditional portfolio management based on the volatility risk measure is biased and corresponds to a short-sighted approach to bad times. We then propose to replace the volatility risk measure by a skewness risk measure, which is calculated as an expected shortfall that incorporates a stress scenario. We conclude that constant-mix portfolios may be better adapted than actively managed portfolios, when the investment universe is composed of negatively skewed financial assets.

Keywords: Skewness, volatility, expected shortfall, stress scenario, market regime, drawdown, risk budgeting, equal risk contribution, Gaussian mixture model, jump-diffusion process.

JEL classification: C50, C60, G11.

*The authors are very grateful to Thibault Bourgeron, Alexandre Burgues, Jean-Thomas Heissat, Edouard Knockaert, Didier Maillard and Bruno Taillardat for their helpful comments. We would like to thank Damiano Brigo who pointed out a numerical error in the first version of the paper.

1 Introduction

The concept of skewness risk[1] in portfolio management is an old story (Arditti, 1967; Levy, 1969; Kraus and Litzenberger, 1976). The three-moment CAPM is an extension of the capital asset pricing model of Sharpe (1964) when we apply a Taylor expansion to the utility function. In this case, the expected utility depends on the product of the third derivative of the utility function and the skewness of the portfolio. Since the seminal paper of Arditti (1967), many academic researchers have proposed methods to estimate the moments of asset returns and compute the optimized portfolio (Harvey and Siddique, 2000; Patton, 2004; Jurczenko and Maillet, 2006; Jondeau and Rockinger, 2006; Martellini and Ziemann, 2010; Harvey *et al.*, 2010, Xiong and Idzorek, 2011; Jondeau *et al.*, 2018). However, the three-moment CAPM has never been very popular with professionals, because it involves estimating many unstable parameters like skewness or co-skewness coefficients.

This explains that optimized portfolio models based on the volatility risk measure are the main approaches used by professionals. Indeed, it is extremely difficult to go beyond the mean-variance optimization (Markowitz, 1952) or the equal risk contribution (ERC) portfolio (Maillard *et al.*, 2010) without facing computational and complexity issues. This is why professionals assume that asset returns can be approximated by Gaussian random variables and prefer to use cooking recipes to stress the covariance matrix than directly modeling the skewness risk. With the emergence of alternative risk premia, this approach is flawed since skewness risk is the backbone of these strategies (Lempérière *et al.*, 2017). Indeed, the distinction between market anomalies and skewness risk premia depends on the skewness risk of the strategy (Roncalli, 2017). For instance, a trend-following strategy is considered as a market anomaly since it exhibits a positive skewness. On the contrary, a value strategy has a negative skewness and belongs to skewness risk premia. By making reference to option trading, Burgues *et al.* (2017) illustrate that alternative risk premia exhibit non-linear payoffs that cannot be aggregated in a Gaussian world. Therefore, the 'correlation diversification approach' is less appropriate when managing a portfolio of alternative risk premia.

In this article, we address the challenge of portfolio management with skewness risk by considering a very simple model. First, we acknowledge that professionals are more confident when it comes to estimating and manipulating covariance matrices than co-skewness tensors. Second, utility functions generally depend on risk aversion coefficients that are unknown. This means that optimized portfolios are highly related to the specification of the utility function. Third, portfolio optimization is dynamic, implying that parameter estimation must be straightforward. Therefore, the skewness model must be close to those used by professionals with similar estimated parameters. For that, we consider the Gaussian mixture model (GMM) developed by Bruder *et al.* (2016) to describe the dynamics of asset returns. We impose two market regimes: a normal regime that occurs most frequently and a stress regime that may negatively impact the performance of the portfolio. Each market regime is characterized by its own covariance matrix (and its vector of expected returns). However, the covariance matrix of the stress regime is defined once and for all. It can be calibrated using a long history of data or it can be specified by expert assessment. The Gaussian mixture model is then dynamic since the covariance matrix of the normal regime is updated on a regular basis[2]. Concerning the allocation model, we use the risk budgeting approach based on the expected shortfall risk measure (Roncalli, 2015) because of three main reasons. First,

[1]We use the term skewness risk to name financial events that induce a negative skewness on assets or investment strategies. Positive skewness is less of a concern in investment theory.

[2]The covariance matrix and the vector of expected returns are static for the stress regime, because it models extreme events with a low frequency distribution. If the stress scenario is well calibrated, there is then no reason for the second regime to be dynamic.

the risk budgeting method is now widely accepted by professionals and produces a more realistic asset allocation than the Markowitz model (Roncalli, 2013). Second, the expected shortfall is related to the concept of maximum drawdown, which is a standard risk measure in asset management. Third, we don't face computational issues since the model is very tractable. This last point is very important from a practical point of view, because fund managers don't use complex allocation models that require intensive programming skills.

The choice of a risk budgeting model based on the expected shortfall and a Gaussian mixture model has several advantages. It can be viewed as a disruption of the traditional ERC portfolio based on the volatility risk measure, which becomes a particular case for our allocation model. Indeed, setting the probability of the second regime to zero allows us to retrieve the ERC portfolio. Another special case is to only consider the stress regime. Finally, the Skew-ERC portfolio[3] may be interpreted as the weighted average between the ERC portfolio based on the normal regime and the ERC portfolio based on the stress regime. The weights of each portfolio depend on the risk appetite of the investor. For instance, if the investor is very sensitive to a drawdown risk, he will overweight the second regime. The analysis of Skew-RB portfolios also shows that drawdown management does not necessarily imply an actively-managed dynamic allocation. Of course, we can risk manage each strategy in order to reduce its skewness. However, since skewness events are often sudden and rapid, it is generally too late to reduce ex-post the exposure to the strategy that suffered from a skewness event. It is then better to take into account the drawdown risk from an ex-ante point of view. This is exactly the objective of Skew-RB portfolios, and this is why they do not overreact to skewness events contrary to traditional mean-variance or risk budgeting portfolios.

This paper is organized as follows. Section Two is dedicated to the theoretical framework. It presents the Gaussian mixture model and illustrates that such a model exhibits skewness risk. We also derive the theoretical formula of the expected shortfall for a given portfolio. Then, we can define the risk allocation process and the optimization program to obtain risk budgeting portfolios. Since some jump-diffusion models can be cast into Gaussian mixture models, we highlight the importance of the second regime that can be interpreted as a stressed market regime, while the first regime is similar to a normal market regime. Section Three considers an application of the theoretical model to a universe of traditional and alternative risk premia with the objective of computing a robust asset allocation. This requires a two-step process. We first calibrate the second regime using historical scenarios, and then implement the risk budgeting allocation by considering a dynamic normal regime, but a static stress regime. Finally, Section Four offers some concluding remarks.

2 Risk decomposition and allocation in the presence of skewness risk

In what follows, we consider the Gaussian mixture model (GMM), which has the property to exhibit positive and negative skewness. GMM is a very simple and tractable statistical model because it satisfies many properties of the Gaussian model. In particular, the sum of GMM random variables follows a GMM probability distribution. This stability property is essential when modeling asset returns. Indeed, the fact that asset returns and portfolio returns have the same probability distribution simplifies the risk budgeting problem. In the second part, we will see how the probability distribution of some jump-diffusion processes can be approximated by a Gaussian mixture model. The bridge between these two modeling

[3] In order to distinguish risk budgeting portfolios based on the volatility risk measure and risk budgeting portfolios based on the skewness risk measure, we name these later Skew-RB or Skew-ERC portfolios.

choices is very interesting, since GMM can be viewed as a statistical model that takes into account bad times and adverse scenarios. In this case, skewness risk is just the demonstration of extreme negative asset returns.

2.1 The Gaussian mixture model

2.1.1 Modeling asset returns

Let $R_t = (R_{1,t}, \ldots, R_{n,t})$ be the vector of asset returns. We assume that R_t follows a Gaussian mixture distribution. We have two regimes:

1. In the first regime, the returns are driven by the multivariate normal distribution $\mathcal{N}(\mu_1, \Sigma_1)$;

2. In the second regime, the returns are driven by another multivariate normal distribution $\mathcal{N}(\mu_2, \Sigma_2)$.

The probability of being in the j^{th} regime is denoted by π_j. The corresponding probability density function is then equal to:

$$f(y) = \pi_1 \phi_n(y; \mu_1, \Sigma_1) + \pi_2 \phi_n(y; \mu_2, \Sigma_2)$$

where $\phi_n(y; \mu, \Sigma)$ is the probability density function of the normal distribution $\mathcal{N}(\mu, \Sigma)$. Let $R_t(x)$ be the return of Portfolio $x = (x_1, \ldots, x_n)$:

$$\begin{aligned} R_t(x) &= x^\top R_t \\ &= B_1 Y_1 + B_2 Y_2 \end{aligned}$$

where B_1 and $B_2 = 1 - B_1$ are Bernoulli random variables with probability π_1 and π_2. Y_1 and Y_2 represent the portfolio's return in the first and second regime. We have $Y_j \sim \mathcal{N}\left(\mu_j(x), \sigma_j^2(x)\right)$ where $\mu_j(x) = x^\top \mu_j$ and $\sigma_j(x) = \sqrt{x^\top \Sigma_j x}$. We can show that $R_t(x)$ has a Gaussian mixture distribution[4]:

$$f(y) = \pi_1 \frac{1}{\sigma_1(x)} \phi\left(\frac{y - \mu_1(x)}{\sigma_1(x)}\right) + \pi_2 \frac{1}{\sigma_2(x)} \phi\left(\frac{y - \mu_2(x)}{\sigma_2(x)}\right) \tag{1}$$

2.1.2 Statistical moments of the portfolio's return

Since $R_t(x)$ is a mixture random variable, the k-th moment of $Y = R_t(x)$ is given by:

$$\mathbb{E}\left[Y^k\right] = \pi_1 \mathbb{E}\left[Y_1^k\right] + \pi_2 \mathbb{E}\left[Y_2^k\right]$$

We know that $\mathbb{E}[Y_j] = \mu_j(x)$, $\mathbb{E}\left[Y_j^2\right] = \mu_j^2(x) + \sigma_j^2(x)$ and $\mathbb{E}\left[Y_j^3\right] = \mu_j^3(x) + 3\mu_j(x)\sigma_j^2(x)$. We deduce that:

$$\mathbb{E}[Y] = \pi_1 \mu_1(x) + \pi_2 \mu_2(x)$$

and:

$$\begin{aligned} \sigma^2(Y) &= \pi_1 \mathbb{E}\left[Y_1^2\right] + \pi_2 \mathbb{E}\left[Y_2^2\right] - \mathbb{E}^2[Y] \\ &= \pi_1 \sigma_1^2(x) + \pi_2 \sigma_2^2(x) + \pi_1 \pi_2 (\mu_1(x) - \mu_2(x))^2 \end{aligned}$$

[4] We note $\phi(y)$ and $\Phi(y)$ the probability density function (pdf) and the cumulative distribution function (cdf) of the standard Normal distribution.

because $\pi_2 = 1 - \pi_1$. We also have:

$$\begin{aligned}
\mathbb{E}\left[(Y - \mathbb{E}[Y])^3\right] &= \mathbb{E}\left[Y^3\right] - 3\mathbb{E}[Y]\,\sigma^2(Y) - \mathbb{E}^3[Y] \\
&= \pi_1\pi_2(\pi_2 - \pi_1)(\mu_1(x) - \mu_2(x))^3 + \\
&\quad 3\pi_1\pi_2(\mu_1(x) - \mu_2(x))\left(\sigma_1^2(x) - \sigma_2^2(x)\right)
\end{aligned}$$

The skewness coefficient of Y has then the following expression:

$$\begin{aligned}
\gamma_1(Y) &= \mathbb{E}\left[\left(\frac{Y - \mathbb{E}[Y]}{\sigma(Y)}\right)^3\right] \\
&= \frac{\pi_1\pi_2\left(3(\mu_1(x) - \mu_2(x))\left(\sigma_1^2(x) - \sigma_2^2(x)\right) - (\pi_1 - \pi_2)(\mu_1(x) - \mu_2(x))^3\right)}{\left(\pi_1\sigma_1^2(x) + \pi_2\sigma_2^2(x) + \pi_1\pi_2(\mu_1(x) - \mu_2(x))^2\right)^{3/2}}
\end{aligned} \tag{2}$$

This result is important, because it shows that GMM exhibits skewness. The sign of the skewness coefficient depends on the probability difference $\pi_1 - \pi_2$, the return difference $\mu_1(x) - \mu_2(x)$ and the variance difference $\sigma_1^2(x) - \sigma_2^2(x)$. For instance, the skewness is negative if the second regime has a low probability, a low portfolio return and a high portfolio volatility. In finance, this situation occurs when the second regime corresponds to a financial stress regime. We also notice that the skewness is equal to zero when there is only one regime or when the two regimes have the same expected return.

Remark 1 *Generally, people use complex probability distributions to model the skewness risk (Fogler and Radcliffe, 1974; McDonald, 1996; Aït-Sahalia and Lo, 1998; Jondeau et al., 2006; Alexander et al., 2012). These statistical models are appealing from a theoretical perspective, but in practice they are difficult to implement. Gaussian mixture models are more known for modeling switching regimes (Frühwirth-Schnatter, 2006). However, they are perfectly adapted to generate skewness risk. Moreover, they are infinitely more tractable than the majority of non-Gaussian probability distributions.*

2.1.3 The expected shortfall risk measure

The expected shortfall of Portfolio x is defined by:

$$\mathrm{ES}_\alpha(x) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(x)\,\mathrm{d}u$$

where α is the confidence level. Bruder *et al.* (2016) shows that:

$$\mathrm{ES}_\alpha(x) = \pi_1\varphi(\mathrm{VaR}_\alpha(x), \mu_1(x), \sigma_1(x)) + \pi_2\varphi(\mathrm{VaR}_\alpha(x), \mu_2(x), \sigma_2(x)) \tag{3}$$

where the function $\varphi(a, b, c)$ is defined by:

$$\varphi(a, b, c) = \frac{c}{1-\alpha}\phi\left(\frac{a+b}{c}\right) - \frac{b}{1-\alpha}\Phi\left(-\frac{a+b}{c}\right)$$

and the value-at-risk $\mathrm{VaR}_\alpha(x)$ is the solution of the following equation:

$$\pi_1\Phi\left(\frac{\mathrm{VaR}_\alpha(x) + \mu_1(x)}{\sigma_1(x)}\right) + \pi_2\Phi\left(\frac{\mathrm{VaR}_\alpha(x) + \mu_2(x)}{\sigma_2(x)}\right) = \alpha \tag{4}$$

2.1.4 Analytical expression of risk contributions

Bruder *et al.* (2016) demonstrate that the vector of risk contributions with respect to the expected shortfall is equal to:

$$\begin{aligned}\mathcal{RC}(x) &= \frac{\varpi_1(x)}{1-\alpha}(x\circ\delta_1(x)) + \frac{\varpi_2(x)}{1-\alpha}(x\circ\delta_2(x)) - \\ &\quad \frac{1}{1-\alpha}\left(\pi_1(x\circ\mu_1)\Phi(-h_1(x)) + \pi_2(x\circ\mu_2)\Phi(-h_2(x))\right)\end{aligned} \tag{5}$$

where $\circ$ is the Hadamard product,

$$\begin{aligned}\delta_1(x) &= \left(1+\frac{h_1(x)}{\sigma_1(x)}\mathrm{VaR}_\alpha(x)\right)\Sigma_1 x - \\ &\quad \mathrm{VaR}_\alpha(x)\frac{\varpi_1(x)\dfrac{h_1(x)}{\sigma_1(x)}\Sigma_1 x + \varpi_2(x)\left(\dfrac{h_2(x)}{\sigma_2(x)}\Sigma_2 x + \mu_1 - \mu_2\right)}{\varpi_1(x)+\varpi_2(x)}\end{aligned}$$

and:

$$\begin{aligned}\delta_2(x) &= \left(1+\frac{h_2(x)}{\sigma_2(x)}\mathrm{VaR}_\alpha(x)\right)\Sigma_2 x - \\ &\quad \mathrm{VaR}_\alpha(x)\frac{\varpi_1(x)\left(\dfrac{h_1(x)}{\sigma_1(x)}\Sigma_1 x + \mu_2 - \mu_1\right) + \varpi_2(x)\dfrac{h_2(x)}{\sigma_2(x)}\Sigma_2 x}{\varpi_1(x)+\varpi_2(x)}\end{aligned}$$

The other notations are $h_j(x) = \sigma_j(x)^{-1}(\mathrm{VaR}_\alpha(x) + \mu_j(x))$, $\varpi_j(x) = \pi_j\sigma_j(x)^{-1}\phi(h_j(x))$.

2.1.5 Risk budgeting portfolios

Roncalli (2013) defines risk budgeting (RB) portfolios using the following non-linear system:

$$\left\{\begin{array}{l}\mathcal{RC}_i(x) = b_i\mathcal{R}(x)\\ b_i > 0\\ x_i \geq 0\\ \sum_{i=1}^n b_i = 1\\ \sum_{i=1}^n x_i = 1\end{array}\right. \tag{6}$$

where b_i is the risk budget of asset i expressed in relative terms. The constraint $b_i > 0$ implies that no risk budget can be set to zero. This restriction is necessary in order to ensure that the RB portfolio is unique (Roncalli, 2013). In the case of the expected shortfall risk measure $\mathrm{ES}_\alpha(x)$, Bruder *et al.* (2016) demonstrate that there exists a value α^- such that the RB portfolio exists and is unique for all $\alpha > \alpha^-$. In this case, the weights of the RB portfolio are equal to:

$$x_i^\star = \frac{y_i^\star}{\sum_{j=1}^n y_j^\star}$$

where $y^\star$ is the solution of the following unconstrained optimization program[5] (Roncalli, 2013):

$$y^\star = \arg\min \mathrm{ES}_\alpha(y) - \lambda\sum_{i=1}^n b_i\ln y_i$$

[5] Here, λ is an arbitrary Lagrange coefficient.

2.1.6 Numerical example

We consider an investment universe of three assets. In the first regime, the expected returns are equal to 5%, 10% and 15%, the volatilities are equal to 8%, 18% and 25%, while the correlation matrix of asset returns is provided by the following matrix:

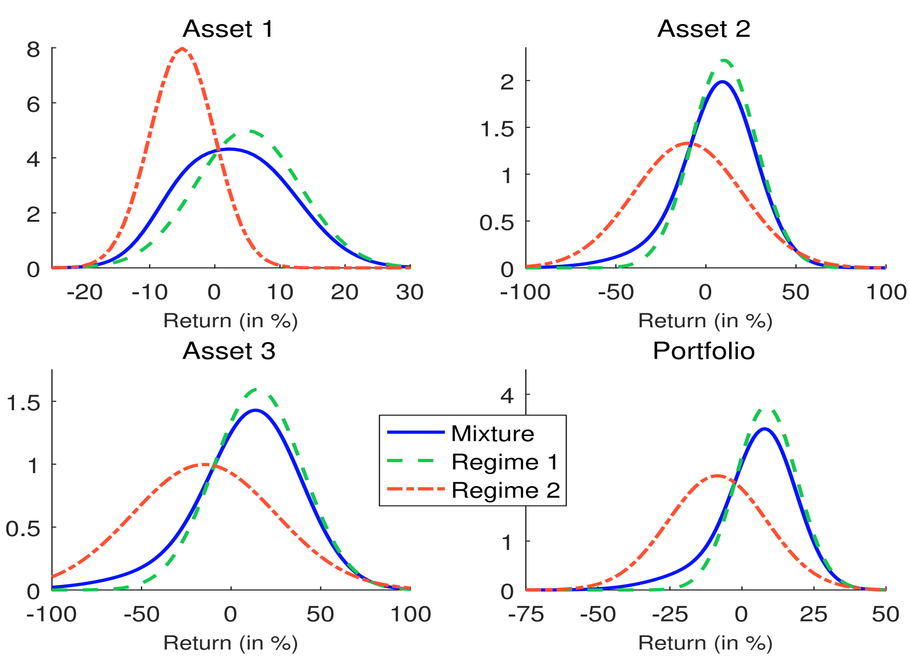
$$C_1 = \begin{pmatrix} 1.00 & & \\ 0.30 & 1.00 & \\ 0.20 & 0.40 & 1.00 \end{pmatrix}$$

In the second regime, the expected returns are equal to -5%, -10% and -15%, the volatilities are equal to 5%, 30% and 40%, while the correlation matrix of asset returns is provided by the following matrix:

$$C_2 = \begin{pmatrix} 1.00 & & \\ 0.30 & 1.00 & \\ 0.30 & 0.80 & 1.00 \end{pmatrix}$$

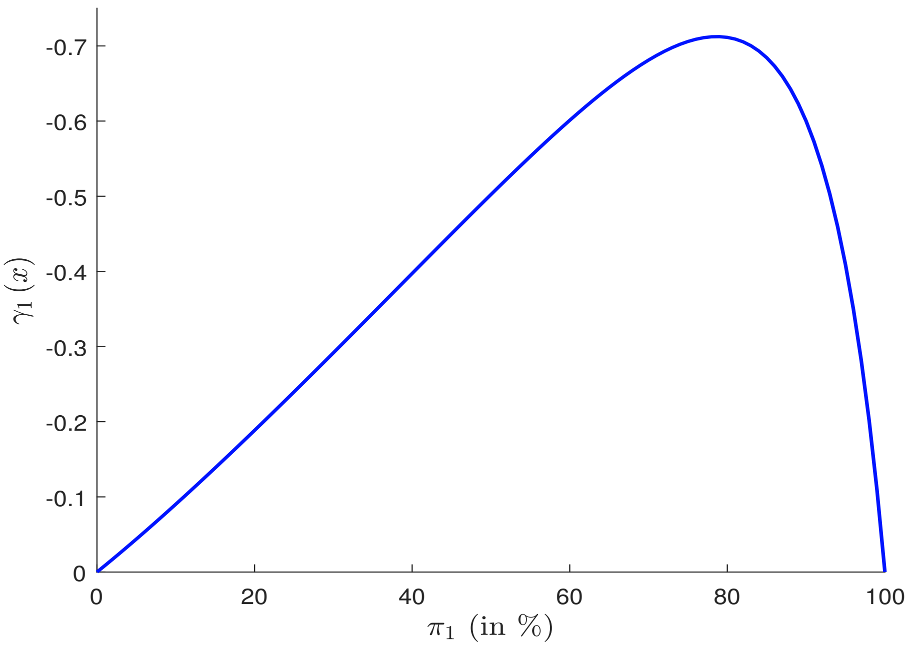
The first regime denotes a 'normal' market regime, while the second regime describes a 'stressed' market regime. The probabilities are $\pi_1 = 80\%$ and $\pi_2 = 20\%$. The composition of the portfolio is respectively 50%, 30% and 20%.

Figure 1: Probability density function of asset and portfolio returns



In Figure 1, we report the probability density functions of asset returns for the two market regimes. The blue line corresponds to the mixture distribution and can be seen as a disruption of the normal regime. The last panel presents the probability distribution of the portfolio, which is left skewed. This result is confirmed by the calculation of the skewness, which is equal to -0.711. The magnitude of the skewness depends on the values taken by π_1 and π_2. For the two extreme cases — $\pi_1 = 1$ and $\pi_1 = 0$, the skewness parameter is equal to zero. In the other cases, the skewness is negative[6] and reaches a peak around $\pi_1 = 80\%$ (see Figure 2). The negative skew is explained by the nature of the second regime. Indeed, it is obvious that it corresponds to a 'bear market' regime because the probability of occurrence is low, expected returns are negative, and volatilities are higher than in the first regime.

Figure 2: Skewness of the portfolio return



In Tables 1 and 2, we report the risk decomposition[7] of the portfolio when the risk measure is the expected shortfall at the 90% confidence level. The Gaussian approximation of asset returns is $\bar{R}_t \sim \mathcal{N}\left(\bar{\mu}, \bar{\Sigma}\right)$ where the expected returns are 3%, 6% and 9%, the volatilities are 8.5%, 22.4% and 31.0% and the correlation matrix is:

$$\bar{C} = \left(\begin{array}{ccc} 1.00 & & \\ 0.40 & 1.00 & \\ 0.35 & 0.62 & 1.00 \end{array}\right)$$

Using this Gaussian approximation, we find that the risk contribution of Asset 1 is 17.28%. This number is greater than the value 15.17%, which is obtained with the Gaussian mixture model. Therefore, the Gaussian approximation overestimates the risk of Asset 1. The issue comes from the correlations, which are higher in the approximated Gaussian model than in the two market regimes. This phenomenon of high correlations is well known by statisticians, since the Gaussian approximation has the tendency to increase inter-regime correlations.

The choice of the model will then have an impact on the computation of RB portfolios. For instance, we have reported the ERC portfolio in Table 3, when the risk budgets are equal and the risk measure is the expected shortfall with a 90% confidence level. If we consider

[6] A Gaussian mixture model can produce a positive skewness, when the stressed regime has low volatilities and the normal regime has high volatilities. However, this situation is not realistic in finance.

[7] x_i indicates the weight of Asset i, $\mathcal{MR}_i$ is the marginal risk, $\mathcal{RC}_i$ is the absolute risk contribution and $\mathcal{RC}_i^\star$ is the normalized risk contribution. All statistics are expressed in %.

Table 1: Expected shortfall decomposition (Gaussian approximation) — $\alpha = 90\%$

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	50.00	6.70	3.35	17.28
2	30.00	28.60	8.58	44.23
3	20.00	37.32	7.46	38.49
$\mathrm{ES}_\alpha(x)$			19.39	

Table 2: Expected shortfall decomposition (Gaussian mixture model) — $\alpha = 90\%$

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	50.00	7.06	3.53	15.17
2	30.00	34.71	10.41	44.75
3	20.00	46.64	9.33	40.09
$\mathrm{ES}_\alpha(x)$			23.27	

the ERC portfolio in the normal (or first) regime, the allocation is equal to 60.01% invested in Asset 1, 22.15% in Asset 2, and 17.84% in Asset 3. If we only consider the stressed regime, the allocation becomes 75.45%, 14.15% and 10.40%. In the second regime, the ERC portfolio underweights the second and third assets, because they are riskier. We notice that the Gaussian approximation produces an allocation, which is relatively close to the average of the two previous ERC portfolios. This is not the case for the mixture model. Indeed, the adverse scenario produces bad and extreme events, meaning that the expected shortfall is more sensitive to the second regime than the first regime. This implies that the Skew-RB portfolio calculated with the mixture model converges to the RB portfolio calculated with the second regime when the confidence level α tends to 1. For instance, if α is equal to 99%, we obtain the ERC portfolios reported in Table 4. We verify that the Skew-ERC portfolio calculated with the mixture model is close to the ERC portfolio calculated with the stressed regime. It follows that the confidence level α is an important parameter and may be interpreted as the investor's risk aversion with respect to the stress scenario.

Table 3: ERC portfolio (in %) — 90% expected shortfall

Asset	Regime 1	Regime 2	Regime Avg.	Gaussian Approx.	Mixture model
1	60.01	75.45	63.10	64.36	69.59
2	22.15	14.15	20.55	20.31	17.48
3	17.84	10.40	16.35	15.33	12.93

Table 4: ERC portfolio (in %) — 99% expected shortfall

Asset	Regime 1	Regime 2	Regime Avg.	Gaussian Approx.	Mixture model
1	59.28	76.95	62.82	63.77	75.78
2	22.84	13.25	20.92	20.72	13.94
3	17.87	9.80	16.26	15.50	10.28

2.2 Understanding the skewness risk

The main interest of the Gaussian mixture model is its tractability when calculating risk decomposition and performing risk allocation. It is a way to deal with non-Gaussian probability distributions without the disadvantage of using Monte Carlo simulations for portfolio risk analysis. As seen previously, the Gaussian mixture model can also generate skewness risk, which may be positive or negative. We now consider a specific jump-diffusion model, which is explicitly related to skewness risk. In this case, the skewness results from the existence of jumps in asset prices. It follows that skewness aggregation works differently than volatility aggregation since it is extremely difficult to diversify or hedge jumps.

2.2.1 The jump-diffusion model

We consider n risky assets represented by the vector of prices $S_t = (S_{1,t}, \ldots, S_{n,t})$, the dynamics of which are given by a jump-diffusion process:

$$\frac{\mathrm{d}S_t}{S_t} = \mu \,\mathrm{d}t + \Sigma^{1/2}\,\mathrm{d}W_t + \mathrm{d}Z_t \tag{7}$$

where μ and Σ are the vector of expected returns and the covariance matrix, and W_t is a n-dimensional standard Brownian motion. The n-dimensional jump process Z_t is a compound Poisson process with intensity parameter $\lambda > 0$ and jump sizes following the multivariate Gaussian distribution $\mathcal{N}\left(\tilde{\mu}, \tilde{\Sigma}\right)$. Bruder *et al.* (2016) show that the asset returns $R_t = (R_{1,t}, \ldots, R_{n,t})$ of the jump-diffusion model are approximately distributed as a Gaussian mixture model where[8]:

1. the first regime is characterized by $\pi_1 = 1 - \lambda\,\mathrm{d}t$, $\mu_1 = \mu\,\mathrm{d}t$ and $\Sigma_1 = \Sigma\,\mathrm{d}t$;
2. the second regime is characterized by $\pi_2 = \lambda\,\mathrm{d}t$, $\mu_2 = \mu_1 + \tilde{\mu}$ and $\Sigma_2 = \Sigma_1 + \tilde{\Sigma}$.

Using the jump-diffusion model, we can explicitly characterize the sign of the skewness. Indeed, we have the following property (Bruder *et al.*, 2016):

$$\gamma_1(x) < 0 \Rightarrow x^\top \tilde{\mu} < 0$$

The only way to generate negative skewness is to have jumps that are negative on average. The concept of skewness risk is then highly related to the concept of negative jump. It confirms our previous intuition that negative skewness is due to adverse risk scenarios.

2.2.2 Why does skewness risk differ from volatility risk?

We consider a universe of n homogenous assets with the following characteristics: $\mu_i = 10\%$, $\sigma_i = 20\%$, $\rho_{i,j} = 50\%$, $\tilde{\mu}_i = -10\%$, $\tilde{\sigma}_i = 30\%$ and $\tilde{\rho}_{i,j} = 70\%$. We assume that λ is equal to 20%, meaning that we observe a jump event every five years. In Figure 3, we have reported the statistics of the equally-weighted portfolio. We notice that the expected return is independent from the number of assets n, the volatility is a decreasing function of n whereas the Sharpe ratio is an increasing function of n. We retrieve the results of the Markowitz diversification. However, we also notice that the skewness coefficient increases with the number of assets. Therefore, the relationship between skewness risk and diversification is more complex than expected by the financial theory.

The previous result is disturbing: why does diversification reduce the volatility risk, but not the skewness risk? In order to understand this paradox, we consider the example given

[8] The scalar $\mathrm{d}t$ denotes the time period when measuring asset returns: $R_{i,t} = \ln\left(S_{i,t+\mathrm{d}t}\right) - \ln\left(S_{i,t}\right)$.

Figure 3: Annualized statistics of portfolio's monthly returns

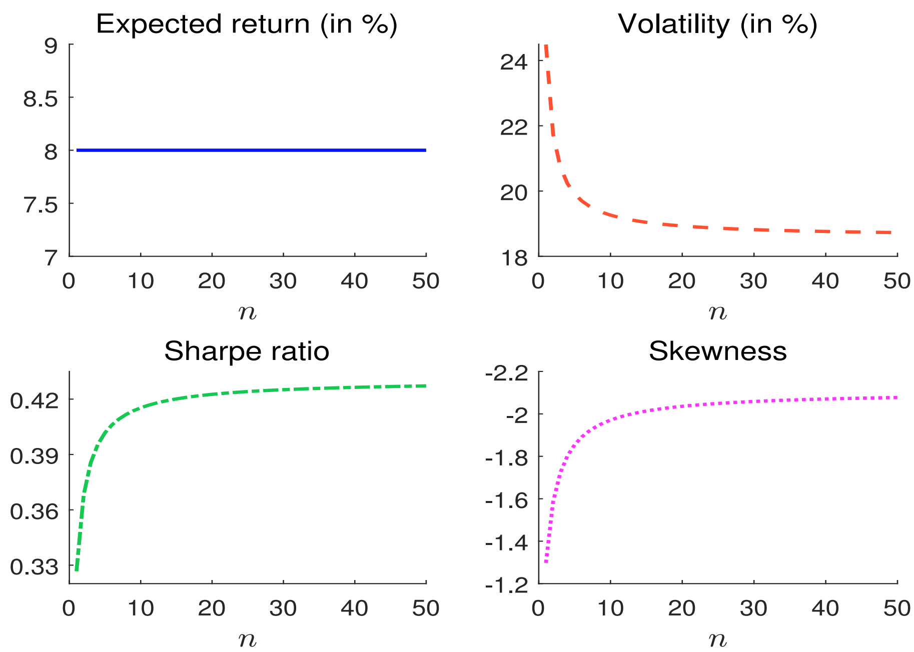
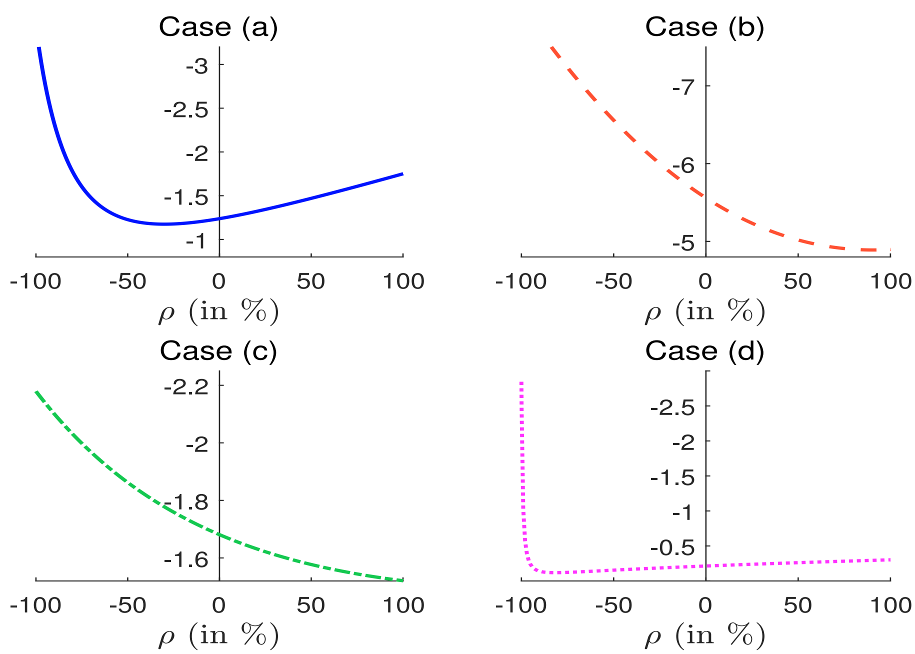


Figure 4: Skewness aggregation of two log-normal random variables



in Hamdan *et al.* (2016) and Roncalli (2017). The authors consider that the opposite of the random vector $X = (X_1, X_2)$ follows a bivariate log-normal distribution:

$$-X \sim \mathcal{LN}_2\left(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho\right)$$

and calculate the skewness of the sum $\gamma_1(X_1 + X_2)$. In Table 5 and Figure 4, we reproduce the numerical illustration of Roncalli (2017). We can make several important comments. First, we notice that the skewness is not a monotone function of the correlation. Second, the skewness risk is maximum when the correlation is equal to -100%. Third, the skewness of a sum of random variables may be larger than the sum of the skewness of each random variable. We conclude that the skewness is not a convex risk measure[9].

Table 5: Skewness aggregation of two log-normal random variables

Case	μ_1	σ_1	μ_2	σ_2	$\gamma_1(X_1)$	$\gamma_1(X_2)$
(a)	0.5	0.5	0.5	0.5	-1.750	-1.750
(b)	0.5	1.0	0.5	0.5	-6.185	-1.750
(c)	1.5	0.1	1.5	0.5	-0.302	-1.750
(d)	1.5	0.1	1.5	0.1	-0.302	-0.302

Case	$\gamma_1(X_1) + \gamma_1(X_2)$	$\gamma_1(X_1 + X_2)$				
		$\rho = -1$	$\rho = -0.5$	$\rho = 0$	$\rho = 0.5$	$\rho = 1$
(a)	-3.500	-3.433	-1.226	-1.238	-1.470	-1.750
(b)	-7.935	-8.030	-6.558	-5.559	-5.018	-4.897
(c)	-2.052	-2.179	-1.862	-1.681	-1.577	-1.521
(d)	-0.604	-2.850	-0.154	-0.213	-0.261	-0.302

From the previous analysis, it follows that diversification helps to reduce the volatility of a portfolio and improves the Sharpe ratio. However, it also increases the skewness risk. How can we explain this inconsistency? As explained by Roncalli (2017), the skewness risk takes the form of a jump risk when asset returns follow a jump-diffusion process. The worst-case scenario concerning skewness aggregation is thus to build a well-diversified portfolio by dramatically reducing the portfolio's volatility. Indeed, it is extremely difficult to diversify the negative jump of an asset. For that, we need to find a second asset that jumps at the same time and has a positive jump. By extending the investment universe, we then increase the volatility diversification and reduce the absolute value of the drawdown, but the drawdown of the portfolio compared to its realized volatility may appear to be very high. This explains that the Sharpe ratio is not the right measure for evaluating the risk/return ratio of a portfolio, when the underlying assets are exposed to skewness risks.

3 Application to a risk premia portfolio

The previous section shows that Gaussian mixture models generate non-zero skewness, and skewness risk is associated with negative jumps in asset returns. We consider here an asset allocation problem, which is very common in asset management. It consists in building a multi-asset risk premia portfolio with non-Gaussian and non-linear strategies. It mixes a

[9] Let $\mathcal{R}(X)$ be the risk measure of the random variable X. $\mathcal{R}$ is a convex risk measure if the following property holds:

$$\mathcal{R}(X_1 + X_2) \leq \mathcal{R}(X_1) + \mathcal{R}(X_2)$$

This property is the pillar of risk diversification.

traditional risk premia exposure captured by an equity/bond risk parity strategy with two alternative risk premia exposures, which are momentum and carry strategies (Burgues *et al.*, 2017).

3.1 Data

We consider a large universe of assets from January 2001 to April 2018. The risk parity strategy is based on several equity asset classes and 10Y sovereign bonds[10]. For the momentum risk premium, we use a lasso-ridge optimized strategy with an EWMA average and with a duration of six months. As proposed by Jusselin *et al.* (2017), we use a multi-asset universe of futures contracts using equity, rates, bonds and currencies. The global carry strategy is based on four individual carry strategies: volatility, FX, bonds and dividends. The volatility carry premium is a systematic short volatility strategy based on SPX variance swaps. For the FX carry premium, we use the ten most liquid DM and EM currencies. The bond carry strategy captures the roll-down premium on DM yield curves (Koijen *et al.*, 2018). Finally, the dividend carry premium is based on dividend futures contracts on the S&P 500, Eurostoxx 50 and Nikkei stock indices. These different strategies are unfunded and their leverages have been calibrated in order to obtain volatility of around 4%.

In Figure 5, we have reported the cumulative performance of the three generic strategies[11]. Descriptive statistics are given in Table 6. For each strategy, we have calculated the annualized performance $\mu(x)$, the volatility $\sigma(x)$, the Sharpe ratio $\mathrm{SR}(x)$, the maximum drawdown $\mathrm{DD}(x)$ and the skew measure $\xi(x)$, which is the ratio between the drawdown and the volatility. If $\xi(x)$ is greater than 3, this indicates a strategy that presents a high skewness risk. This is the case for the global carry strategy, and most of the carry strategies (volatility, FX and dividend). These empirical results are in line with theoretical risk premia models (Roncalli, 2017). The momentum risk premium may be seen as a market anomaly whereas the performance of the carry strategy may be explained by a skewness risk premium. This is why the carry strategy behaves differently to the momentum strategy or even the risk parity strategy. Indeed, we observe drawdowns in the performance of the carry risk premium that cannot be attributed to a volatility risk. They correspond more to a jump risk. These differences in the origin of losses impacts portfolio management as we will show in the last section of this article.

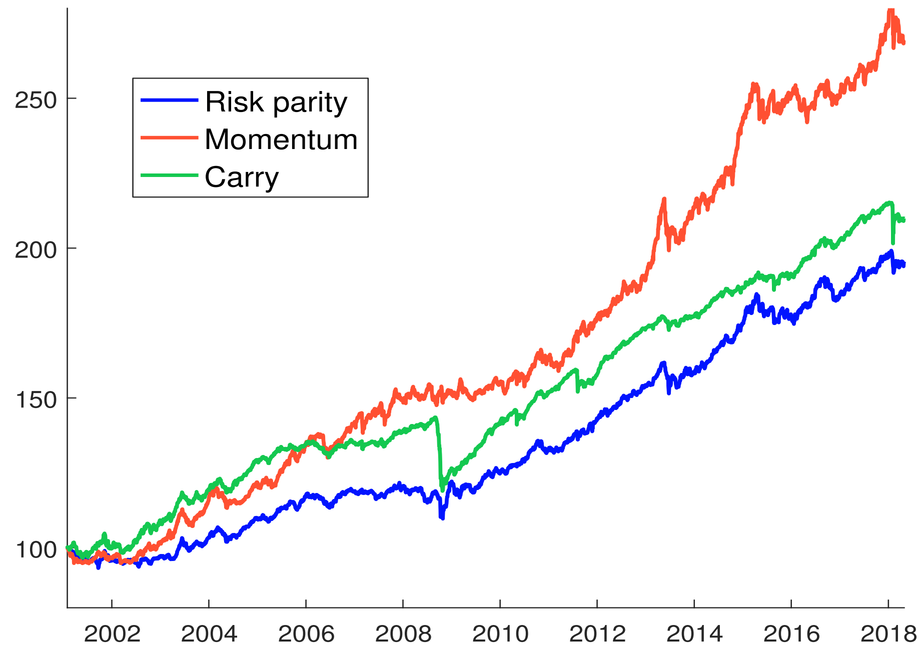
Table 6: Descriptive statistics (in %)

Strategy	$\mu(x)$	$\sigma(x)$	$\mathrm{SR}(x)$	$\mathrm{DD}(x)$	$\xi(x)$
Risk parity	3.97	3.71	1.07	9.94	2.68
Momentum	5.90	4.87	1.21	8.04	1.65
Carry	4.39	3.95	1.11	17.16	4.34
Volatility carry	1.70	4.36	0.39	15.71	3.60
FX carry	4.28	4.63	0.92	14.52	3.13
Bond carry	3.90	4.32	0.90	8.14	1.88
Dividend carry	4.38	4.64	0.94	18.68	4.03

[10] For equities, we use both DM and EM equity indices. For bonds, we restrict the universe to the most liquid DM government bonds.

[11] We also report on page 29 the cumulative performance of the carry strategies. Even if they are not used in our asset allocation exercise, the study of the individual carry risk premia is of primary interest in order to illustrate the aggregation puzzle of the skewness risk. In particular, we will see that it is not obvious to deduce the statistical properties of the global carry strategies from the elementary carry strategies.

Figure 5: Cumulative performance of risk parity, momentum and carry



3.2 Calibrating the stress regime

In order to calibrate the stress regime, we proceed in two steps. First, we determine the univariate stress scenarios for each risk premium strategy. Then, we estimate dependence between these univariate stress scenarios.

3.2.1 Estimating uni-dimensional stress scenarios

To calibrate univariate stress scenarios, we estimate the univariate Gaussian mixture model. Using weekly returns, we report the annualized EM estimates in Table 7. We notice that the regime probabilities are different from one risk premium to another. The issue here is that the results are not comparable. For instance, the occurrence probability of the second regime is 10.86% for the volatility carry strategy, whereas it is equal to 24.99% for the bond carry strategy. If we build a stress scenario based on these estimates, we implicitly underestimate the expected drawdown of momentum, bond carry, currency carry and dividend carry strategies with respect to risk parity, carry and volatility carry strategies. The main reason is that the nature of the second regime differs from one strategy to another, implying that the second regime is not necessarily a stress regime. In order to obtain comparable results, we fix the probabilities π_1 and π_2. This approach is already suggested by Bruder *et al.* (2016), who notice that estimating a two-regime model is generally equivalent to considering two volatility regimes. However, the objective of the Gaussian mixture model in our case is to describe a stressed regime that may produce a sudden drawdown and a high skewness risk.

If we impose that $\pi_1 = 95\%$, we obtain the results[12] given in Table 8. First, we notice that the annualized expected return of the second regime is systematically negative. Second,

[12] They have been estimated using the method of constrained maximum likelihood.

Table 7: EM estimates (in %)

Strategy	$\hat{\pi}_1$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\pi}_2$	$\hat{\mu}_2$	$\hat{\sigma}_2$
Risk parity	88.00	6.51	3.37	12.00	−14.77	7.65
Momentum	82.78	11.35	3.97	17.22	−20.45	8.50
Carry	90.12	6.93	2.56	9.88	−19.07	10.14
Volatility carry	89.14	5.37	1.31	10.86	−28.08	9.82
FX carry	75.01	7.46	3.66	24.99	−5.25	7.35
Bond carry	75.01	7.02	3.99	24.99	−5.41	6.91
Dividend carry	81.76	4.78	1.54	18.24	2.77	12.34

we also verify that the volatility of the second regime is generally between two and eight times the volatility of the first regime. It is then obvious the second regime corresponds to a market stress with negative returns and high volatilities. However, we notice that skewness patterns impact trading strategies differently. For example, the ratio $\hat{\sigma}_2/\hat{\sigma}_1$ is moderate for momentum, currency carry and bond carry strategies. On the contrary, it is very high for volatility carry and dividend carry strategies.

Table 8: Constrained ML estimates (in %)

Strategy	π_1	$\hat{\mu}_1$	$\hat{\sigma}_1$	π_2	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\sigma}_2/\hat{\sigma}_1$
Risk parity	95.00	5.80	3.59	5.00	−28.23	9.21	2.6
Momentum	95.00	9.82	4.45	5.00	−62.75	8.85	2.0
Carry	95.00	6.77	2.74	5.00	−33.52	11.87	4.3
Volatility carry	95.00	5.12	1.46	5.00	−41.94	11.37	7.8
FX carry	95.00	6.18	4.30	5.00	−29.19	10.02	2.3
Bond carry	95.00	5.26	4.49	5.00	−20.32	9.41	2.1
Dividend carry	95.00	5.27	1.98	5.00	−3.48	16.33	8.3

In what follows, we set $\pi_1 = \pi$ and $\pi_2 = 1 - \pi$. In order to measure the sensitivity of the results with respect to the probability π, we calculate the expected drawdown for a return period τ. We assume that the drawdown occurs in the second regime. In this case, we have[13]:

$$\mathbb{E}\left[\mathrm{DD}\left(\pi\right)\right] = -\left(\hat{\mu}_2\left(\pi\right) + \Phi^{-1}\left(\frac{\mathrm{d}t}{\tau}\right)\hat{\sigma}_2\left(\pi\right)\right)$$

where $\mathrm{d}t$ is the sampling frequency[14]. By construction, the drawdown $\mathrm{DD}\left(\pi\right)$ and the estimated parameters $\hat{\mu}_2\left(\pi\right)$ and $\hat{\sigma}_2\left(\pi\right)$ depends on the parameter π, because the model is estimated conditionally to $\pi_1 = \pi$ and $\pi_2 = 1 - \pi$. We also introduce the skew measure as the ratio between the expected drawdown and the annualized volatility of the normal regime:

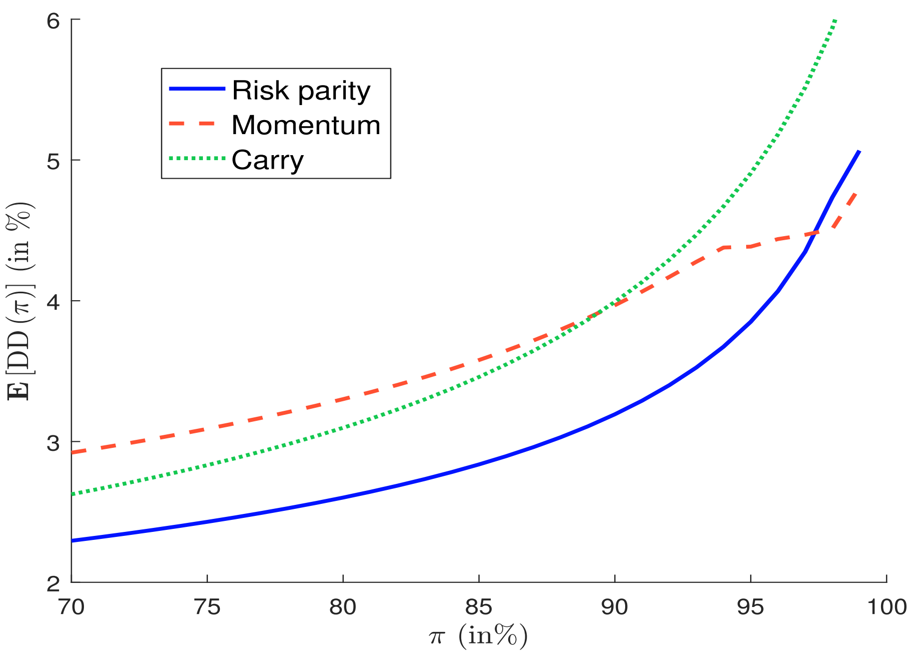
$$\begin{aligned}\zeta\left(\pi\right) &= \frac{\mathbb{E}\left[\mathrm{DD}\left(\pi\right)\right]}{\hat{\sigma}_1} \\ &= -\frac{\hat{\mu}_2\left(\pi\right) + \Phi^{-1}\left(\frac{\mathrm{d}t}{\tau}\right)\hat{\sigma}_2\left(\pi\right)}{\hat{\sigma}_1}\end{aligned}$$

[13]We define the drawdown as a positive risk measure. This is why we take the opposite of the loss.

[14]We reiterate here that we are considering weekly returns. As such, the sampling frequency $\mathrm{d}t$ is set to five trading days.

In Figure 6, we have reported the weekly drawdown in % and the skew ratio for different values of π. As expected, the loss is an increasing function of the probability π, because the second regime is increasingly stressed when its occurrence probability decreases. We notice some differences between the strategies. The dividend and volatility carry strategies are riskier than the other strategies (see Figure 13 on page 29). In particular, the dividend carry strategy presents a high drawdown risk. If we consider the skew measure, we have three types of strategies. The first group is composed of the risk parity, momentum, bond carry and currency carry strategies, and has a low skew measure, which is generally less than 1.5. The global carry strategy has an intermediate skew value. Finally, the riskiest strategies are the dividend and volatility carry strategies. In these two cases, the weekly drawdown may be double the annual volatility if the probability π is very high!

Figure 6: Weekly stress scenario for a four-year return period



In Table 9, we have reported the figures for some specific values of the probability π. All the results are expressed in %, except the ratio $\zeta(\pi)$. The statistic $\bar{\sigma}$ is the annualized long-run volatility. By construction, $\bar{\sigma}$ is greater than the volatility σ_1 of the normal regime. For instance, if we consider the bond carry strategy, 70% of the time we observe a volatility of 3.90%, whereas the long-run volatility is equal to 4.94%. When we increase the probability π, we observe two phenomena that impact the skew ratio: first, the drawdown increases because the stress regime is more severe; second, the volatility of the normal regime increases too, because the first regime contains more and more skewness events. Therefore, it is not obvious whether the skew ratio increases or decreases. However, we notice that the first phenomenon has a larger impact than the second one in Table 9. If we consider a probability π of 95%, we obtain the following conclusions:

- The global carry strategy has a volatility that is lower than the average volatility of individual carry strategies. For the expected drawdown, we observe the opposite effect. These results illustrate that volatility diversification is easier to achieve than

Figure 7: Skew measure for a four-year return period

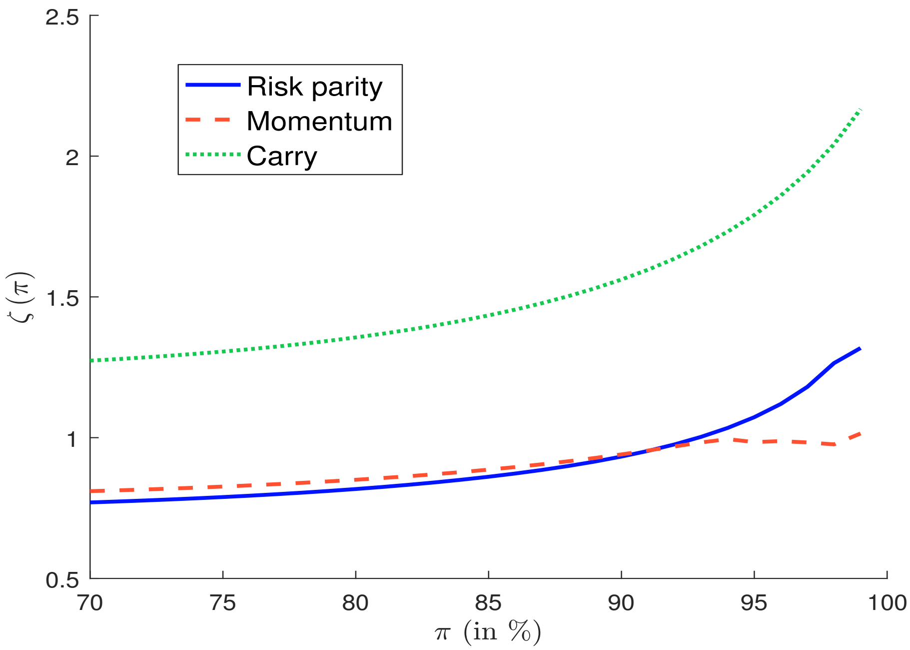


Table 9: Expected drawdown (in %) and skew ratio statistics

Statistic	π (in %)	Risk Parity	Momentum	Carry	Carry Vol.	Carry FX	Carry Bond	Carry Div.
$\bar{\sigma}$		4.23	5.31	4.15	3.75	4.91	4.94	5.46
σ_1	70	2.98	3.61	2.06	1.02	3.54	3.90	1.36
	80	3.18	3.88	2.28	1.16	3.80	4.09	1.51
	90	3.42	4.22	2.56	1.33	4.11	4.33	1.75
	95	3.59	4.45	2.74	1.46	4.30	4.49	1.98
	99	3.84	4.74	3.00	1.67	4.54	4.68	2.39
$\mathbb{E}[\mathrm{DD}(\pi)]$	70	2.29	2.92	2.63	2.86	2.60	2.47	3.77
	80	2.60	3.30	3.10	3.35	2.91	2.73	4.27
	90	3.19	3.97	3.99	4.16	3.50	3.22	5.08
	95	3.85	4.38	4.91	4.89	4.16	3.77	5.93
	99	4.15	4.81	6.51	6.24	4.09	3.58	7.73
$\zeta(\pi)$	70	0.77	0.81	1.27	2.79	0.74	0.63	2.78
	80	0.82	0.85	1.36	2.89	0.77	0.67	2.83
	90	0.93	0.94	1.56	3.13	0.85	0.74	2.91
	95	1.07	0.98	1.79	3.36	0.97	0.84	3.00
	99	1.08	1.01	2.17	3.74	0.90	0.77	3.24

drawdown diversification. While the carry strategy is less volatile than the risk parity and carry strategies, it presents the largest expected drawdown and skew ratio.

- The skewness risk is particularly important for the volatility carry and dividend carry strategies. In the first regime, the volatility of these strategies is low and less than 2%. This implies that at least 50% of the total volatility is explained by skewness risk. For example, in the case of the dividend carry strategy, the volatility of the first regime is equal to 1.98% whereas the total volatility is equal to 5.46%. This means that the realized volatility is not the right risk measure most of the time.

3.2.2 Estimating the dependence between stress scenarios

We reiterate that the probability density function of the asset returns R_t is equal to:

$$f(y) = \pi_1 \phi_n(y; \mu_1, \Sigma_1) + \pi_2 \phi_n(y; \mu_2, \Sigma_2)$$

where n is the number of assets. Since the number of parameters is $n^2 + 3n + 1$, it may be difficult to estimate the model when $n \geq 3$. For instance, we have 41 parameters when n is equal to 5. However, we can reduce the computational issue because we have already estimated the univariate parameters $\hat{\mu}_{1,i}$, $\hat{\sigma}_{1,i}$, $\hat{\mu}_{2,i}$ and $\hat{\sigma}_{2,i}$ and we can fix the probabilities π_1 and π_2. In this case, we have only to estimate the two correlation matrices C_1 and C_2, or $n^2 - n$ parameters[15]. Results are given in Tables 10 and 11. We observe that the correlation matrices of the two regimes are very different. It is then difficult to assume that the dependence between stress scenarios may be approximated by the covariance matrix of the normal regime. For instance, the correlation between risk parity and momentum strategies is 60% in the first regime and 33% in the second regime. Similarly, the correlation between momentum and carry strategies is lower in the second regime than in the first regime. These results are not surprising since momentum generally exhibits positive skewness, and the adverse scenario for momentum strategies is different than for risk parity and carry strategies. On the contrary, the correlation between risk parity and carry strategies remains high in the second regime. Again, this is not surprising because the bad times of the carry risk premium are correlated with the bad times of the risk parity strategy.

Table 10: Correlation matrix $\hat{C}_1$ (in %)

Risk parity	100						
Momentum	60	100					
Carry	69	36	100				
Volatility carry	21	17	65	100			
FX carry	37	5	62	31	100		
Bond carry	62	35	37	-20	-11	100	
Dividend carry	16	4	35	30	33	-25	100

In Table 12, we have also reported the empirical correlation matrix $\hat{C}$ for the entire period. We notice that $\hat{C}$ is closer to $\hat{C}_1$ than $\hat{C}_2$. However, we observe that the empirical correlation is not a weighted average of the normal correlation and the stressed correlation. This implies that the empirical correlation matrix does not help to characterize the dependence in the stressed regime.

[15] When n is large, we can reduce the complexity by estimating each pairwise correlation coefficient independently. Then, we have to regularize the correlation matrix to obtain a positive-definite matrix.

Table 11: Correlation matrix $\hat{C}_2$ (in %)

Risk parity	100						
Momentum	33	100					
Carry	71	22	100				
Volatility carry	53	20	93	100			
FX carry	40	22	51	39	100		
Bond carry	50	11	4	−16	−31	100	
Dividend carry	41	12	64	49	54	−25	100

Table 12: Empirical correlation matrix $\hat{C}$ (in %)

Risk parity	100						
Momentum	53	100					
Carry	66	34	100				
Volatility carry	39	24	88	100			
FX carry	39	13	52	31	100		
Bond carry	58	29	18	−12	−15	100	
Dividend carry	28	9	60	42	38	−19	100

Remark 2 *In Appendix A.1 on page 26, we show that the relationship between the empirical covariance matrix and the covariance matrices of the two regimes is:*

$$\hat{\Sigma} = \pi_1 \hat{\Sigma}_1 + \pi_2 \hat{\Sigma}_2 + \pi_1 \pi_2 (\hat{\mu}_1 - \hat{\mu}_2)(\hat{\mu}_1 - \hat{\mu}_2)^\top$$

It follows $\hat{\Sigma}$ is only related to $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ if and only if the expected returns are the same in the two regimes ($\hat{\mu}_1 = \hat{\mu}_2$). Otherwise, the empirical covariance matrix depends on a modified stressed covariance matrix:

$$\hat{\Sigma} = \pi_1 \hat{\Sigma}_1 + \pi_2 \tilde{\Sigma}_2$$

where $\tilde{\Sigma}_2 = \hat{\Sigma}_2 + \pi_1 (\hat{\mu}_1 - \hat{\mu}_2)(\hat{\mu}_1 - \hat{\mu}_2)^\top$. Since the second regime corresponds to a stress, we have $\hat{\mu}_1 \gg \hat{\mu}_2$, implying that $\tilde{\Sigma}_2$ is larger than $\hat{\Sigma}_2$. It follows that the relationship between the correlation matrices $\hat{C}$, $\hat{C}_1$ and $\hat{C}_2$ is not trivial. For instance, if $\hat{C}_1 = \hat{C}_2$, we obtain that $\hat{C} \succ \hat{C}_1$ in most cases.

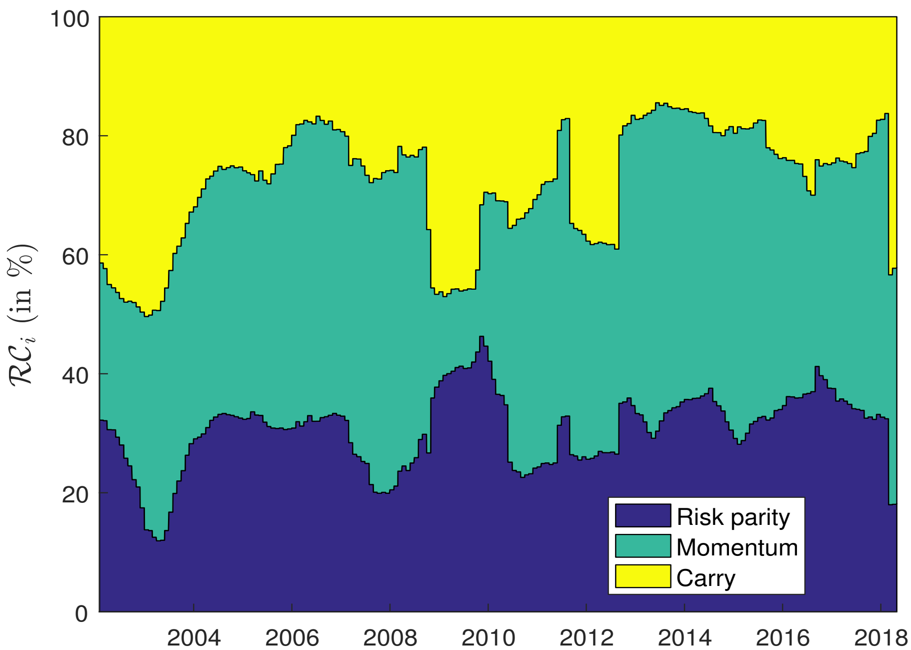
3.3 Designing the portfolio allocation

We first analyze the skewness impact on the risk allocation of constant-mix portfolios. In particular, we show that the traditional approach for estimating volatility and covariance risk is biased. In a second step, we build risk-budgeting portfolios that take into account the skewness risk.

3.3.1 Constant-mix portfolios

We consider an equally-weighted portfolio between the three strategies (risk parity, momentum and carry). We assume that the portfolio is rebalanced at the end of each month, and we calculate the volatility risk using a one-year rolling covariance matrix and daily returns. In Figure 8, we report the evolution of the risk contribution when we consider the traditional volatility risk measure or a Gaussian model. We observe that these risk contributions present some jumps. For example, we observe a sharp increase in the risk contribution of the carry strategy in September/October 2008, August 2011, August 2015 or February 2018. These positive jumps correspond to negative jumps in the cumulative performance of the carry strategy. We also notice that these sharp increases are generally followed by a drop one year later. We explain this phenomenon because the risk measure is based on the one-year covariance matrix, implying that its level depends on whether or not the one-year sample contains the jump event. When the date of the jump is outside the one-year rolling window, the historical volatility is reduced. This is why we observe a second jump in the risk allocation after the initial jump.

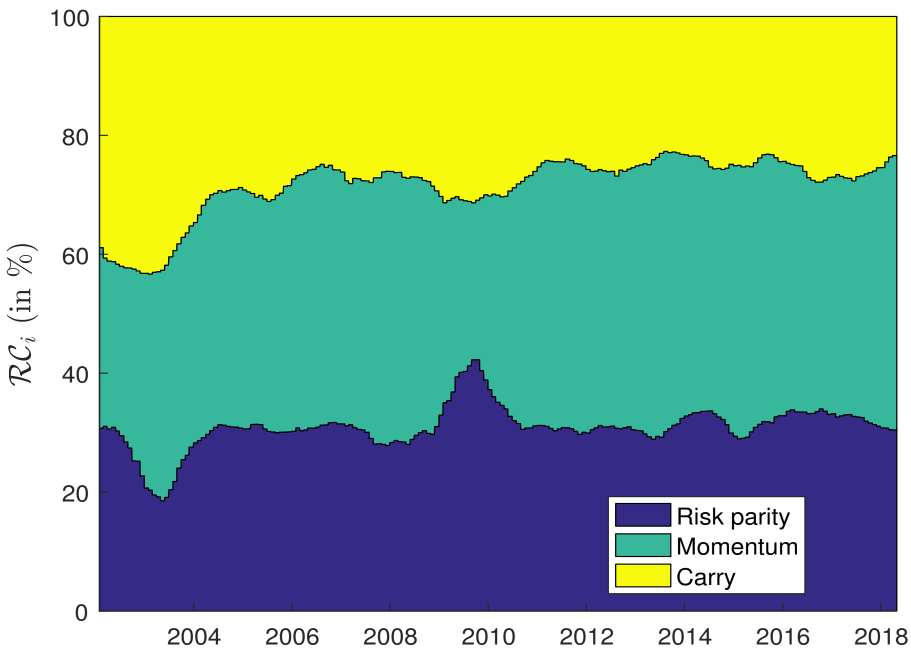
Figure 8: Risk contribution of the EW portfolio (Gaussian model)



We now consider the 90% expected shortfall measure using the jump process. For that, we use the second regime estimated in the previous section. We assume a weekly model for asset returns, meaning that the frequency parameter λ is set to $\pi_2/\mathrm{d}t$, and the time period $\mathrm{d}t$ is equal to $1/52$. For estimating the covariance matrix Σ of the normal regime, we exploit the thresholding algorithm given in Appendix A.2 on page 27. The underlying idea is to split the one-year rolling sample into two subsamples: a subsample that corresponds to the first regime and a subsample that corresponds to the second regime. Therefore, we only consider the first subsample when calculating the empirical covariance matrix $\hat{\Sigma}$. Indeed, if one observation corresponds to a jump event, it must be deleted because it artificially increases the volatility of the first regime. In Figure 9, the thresholding algorithm is implemented with $\pi_2^\star = 50\%$. It is equivalent to assume that the observation t belongs to the first regime if $\hat{\pi}_{2,t}$ is less than 50% and to the jump regime otherwise[16]. We obtain a smoother risk allocation, because the expected shortfall measure is based on the jump model that already incorporates the long-term stress scenario.

[16]We have reported the value of the posterior probability $\hat{\pi}_{2,t}$ in Figure 15 on page 30. The frequency of observations that satisfy $\hat{\pi}_{2,t} > 50\%$ is equal to 5.62%. This number is very close to the theoretical probability $\pi_2 = 5\%$.

Figure 9: Risk contribution of the EW portfolio (GMM, $\alpha = 90\%$)



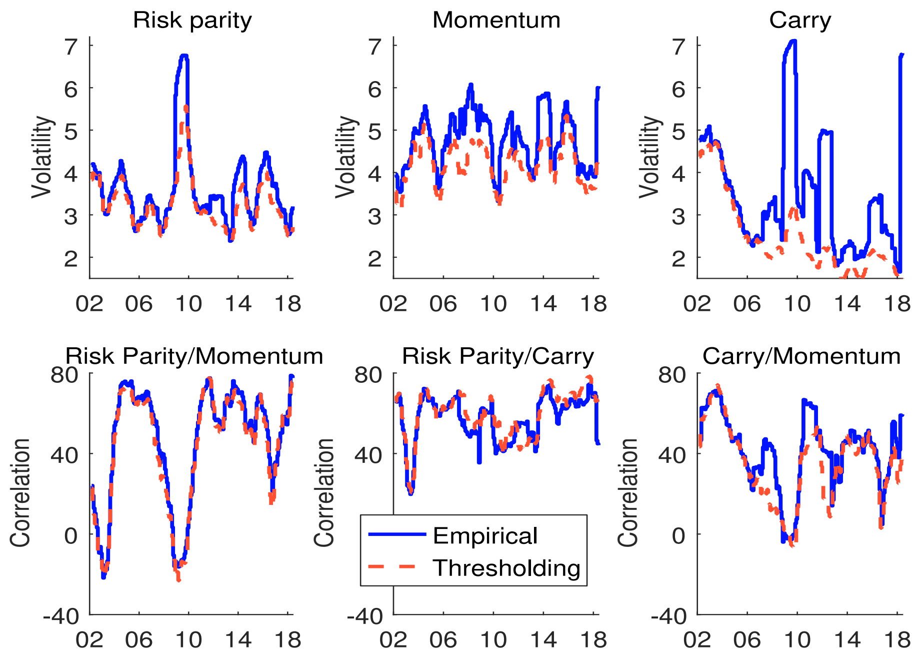
In Table 13, we report the risk allocation of the EW portfolio in the static case when we use the long-term figures. We also report the average risk allocation in the dynamic case, that is when we estimate the covariance matrix using a one-year rolling window. We obtain similar figures using the Gaussian model and the jump model. These results contrast with those obtained in Figures 8 and 9.

Table 13: Risk contribution of the EW portfolio ($\alpha = 90\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	29.53%	30.57%	30.51%	30.89%
Momentum	39.24%	31.01%	41.37%	40.66%
Carry	31.23%	38.42%	28.12%	28.45%

The thresholding method avoids the double counting effect. Indeed, the empirical volatility is generally overestimated because it is calculated with jumps. However, jumps are more an expression of the skewness risk than the volatility risk. Without implementing the thresholding method, we obtain Figure 16 on page 31. We notice that the risk allocation is not smoothed and contains jumps, just as in the case of the risk allocation with the Gaussian model (Figure 8). Figure 10 represents the covariance matrix estimated with empirical and thresholding methods. We notice that the thresholding approach operates a small correction on risk parity and momentum strategies. For instance, the volatility is slightly reduced, but the cross-correlation is equivalent. With the carry strategy, the correction is more important. First, the thresholding approach does not only reduce the volatility, but it eliminates the jumps. Second, it also impacts the cross-correlation. This is why the use of the empirical covariance matrix introduces a bias when estimating the risk allocation of a given portfolio.

Figure 10: Empirical and thresholding estimates of the covariance matrix



Remark 3 *The risk allocation based on the skewness model mainly depends on the two parameters π_2 and α. They both define the severity of the stress scenario. In particular, when the confidence level α of the expected shortfall tends to one, the dynamic allocation becomes flat and is entirely given by the second regime (Figure 17 in page 31). When $\pi_2 \ll 1-\alpha$, the dynamic allocation is mainly driven by the first regime. It follows that the choice of α is important to determine the proportion of each risk (volatility and skewness) to take into account.*

3.3.2 Risk budgeting portfolios

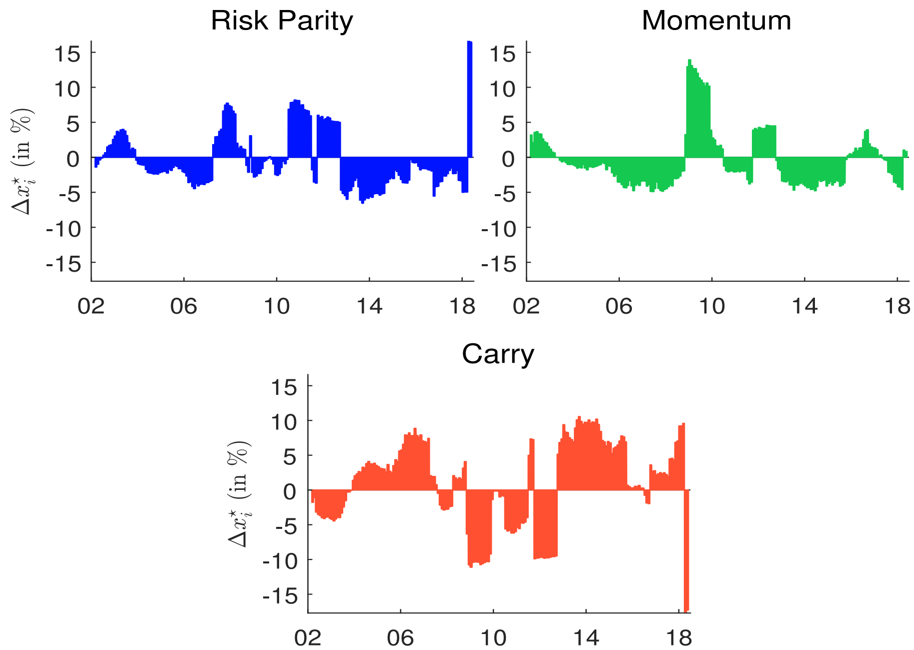
We now implement the risk budgeting approach using the previous framework. We rebalance the portfolio every month and compare the ERC portfolio obtained with the Gaussian model (or the volatility risk measure) with the Skew-ERC portfolio obtained with the Gaussian mixture model (or the skewness risk measure). In Figure 11, we report the weight difference between the ERC portfolio and the Skew-ERC portfolio[17]. As in the previous section, we observe that jumps have a big impact on the allocation. Every time, we observe a negative jump in the carry strategy, the allocation is shifted when we consider the volatility risk measure. This generates a trade-off between the carry strategy and the two other strategies.

Remark 4 *In Figures 20 and 21 on page 33, we consider a higher value for the confidence level α. We deduce that the ERC portfolio overestimates the allocation in the carry strategy with respect to the Skew-ERC portfolio.*

All these results show that skewness risk really matters in portfolio management. If we observe such big differences between ERC and Skew-ERC portfolios, it is because Skew-RB portfolios takes into account the skewness risk from an ex-ante point of view.

[17]The nominal allocation of these portfolios is given in Appendix B.2 on page 32.

Figure 11: Weight difference (in %) between ERC and Skew-ERC portfolios ($\alpha = 90\%$)



4 Conclusion

This paper presents a very simple model of skewness risk, which is tractable from a professional point of view. This model acknowledges that negative skewness is mainly due to stress events in finance. Therefore, the model is just an extension of the Gaussian model that is traditionally used for modeling asset returns but incorporates a second regime for taking into account turbulent periods. It appears that the two market regimes are different in nature. The first regime drives the dynamics of the true volatility, whereas the second regime is responsible for jumps in asset returns. It follows that the normal regime is time-varying and predictable because it describes the short-term volatility, which is a persistent process. On the contrary, the stress regime is a component that is not predictable. This forces us to define the second regime once and for all.

In this framework, skewness risk is then completely different to volatility risk. Generally, volatility risk implies adopting a dynamic allocation in order to respond to market patterns. Skewness risk prefers a static allocation, because jump risk cannot be hedged and should be taken into account from an ex-ante point of view. This result is not necessarily intuitive for professionals, because we generally learn that high risks must be actively managed. However, this basic rule is only valid when the risk is predictable. This is why skewness and jump risks cannot be managed in the same way as volatility risk.

References

[1] Aït-Sahalia, Y., and Lo, A.W. (1998), Nonparametric Estimation of State-price Densities Implicit in Financial Asset Prices, *Journal of Finance*, 53(2), pp. 499-547.

[2] Alexander, C., Cordeiro, G.M., Ortega, E.M., and Sarabia, J.M. (2012), Generalized Beta-generated Distributions, *Computational Statistics & Data Analysis*, 56(6), pp. 1880-1897.

[3] Arditti, F.D. (1967), Risk and the Required Return on Equity, *Journal of Finance*, 22(1), pp. 19-36.

[4] Bruder, B., Kostyuchyk, N., and Roncalli, T. (2016), Risk Parity Portfolios with Skewness Risk: An Application to Factor Investing and Alternative Risk Premia, *SSRN*, `www.ssrn.com/abstract=2813384`.

[5] Burgues, A., Knockaert, A., Lezmi, E., Malongo, H., Roncalli, T. and Sobotka, R. (2017), The Quest for Diversification – Why Does It Make Sense to Mix Risk Parity, Carry and Momentum Risk Premia?, *Amundi Discussion Paper*, 25.

[6] Fogler, H.R., and Radcliffe, R.C. (1974), A Note on Measurement of Skewness, *Journal of Financial and Quantitative Analysis*, 9(3), pp. 485-489.

[7] Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, Springer.

[8] Hamdan, R., Pavlowsky, F., Roncalli, T., and Zheng, B. (2016), A Primer on Alternative Risk Premia, *SSRN*, `www.ssrn.com/abstract=2766850`.

[9] Harvey, C.R., and Siddique, A. (2000), Conditional Skewness in Asset Pricing Tests, *Journal of Finance*, 55(3), pp. 1263-1295.

[10] Harvey, C.R., Liechty, J.C., Liechty, M.W., and Müller, P. (2010), Portfolio Selection with Higher Moments, *Quantitative Finance*, 10(5), pp. 469-485.

[11] Jondeau, E., Jurczenko, E., and Rockinger, M. (2018), Moment Component Analysis: An Illustration with International Stock Markets, *Journal of Business & Economic Statistics*, forthcoming.

[12] Jondeau, E., Poon, S.H., and Rockinger, M. (2007), *Financial Modeling Under Non-Gaussian Distributions*, Springer.

[13] Jondeau, E., and Rockinger, M. (2006), Optimal Portfolio Allocation under Higher Moments, *European Financial Management*, 12(1), pp. 29-55.

[14] Jurczenko, E., and Maillet, B. (2006), *Multi-moment Asset Allocation and Pricing Models*, John Wiley & Sons.

[15] Jusselin, P., Lezmi, E., Malongo, H., Masselin, C., *Roncalli*, T. and Dao, T-L. (2017), Understanding the Momentum Risk premium: An In-Depth Journey Through Trend-Following Strategies, *SSRN*, `www.ssrn.com/abstract=3042173`.

[16] Koijen, R.S.J., Moskowitz, T.J., Pedersen, L.H., and Vrugt, E.B. (2018), Carry, *Journal of Financial Economics*, 127(2), pp. 197-225.

[17] Kraus, A., and Litzenberger, R.H. (1976), Skewness Preference and the Valuation of Risk Assets, *Journal of Finance*, 31(4), pp. 1085-1100.

[18] Lempérière, Y., Deremble, C., Nguyen, T.T., Seager, P., Potters, M., and Bouchaud, J-P. (2017), Risk Premia: Asymmetric Tail Risks and Excess Returns, *Quantitative Finance*, 17(1), pp. 1-14.

[19] Levy, H. (1969), A Utility Function Depending on the First Three Moments, *Journal of Finance*, 24(4), pp. 715-719.

[20] McDonald, J.B. (1996), Probability Distributions for Financial Models, *Handbook of Statistics*, Elsevier, 14, pp. 427-461.

[21] Maillard, S., Roncalli, T., and Teïletche, J. (2010), The Properties of Equally Weighted Risk Contribution Portfolios, *Journal of Portfolio Management*, 36(4), pp. 60-70.

[22] Markowitz, H. (1952), Portfolio Selection, *Journal of Finance*, 7(1), pp. 77-91.

[23] Martellini, L. and Ziemann, V. (2010), Improved Estimates of Higher-Order Comoments and Implications for Portfolio Selection, *Review of Financial Studies*, 23(4), pp. 1467-1502.

[24] Patton, A.J. (2004), On the Out-of-sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, *Journal of Financial Econometrics*, 2(1), pp. 130-168.

[25] Roncalli, T. (2013), *Introduction to Risk Parity and Budgeting*, Chapman & Hall/CRC Financial Mathematics Series.

[26] Roncalli, T. (2015), Introducing Expected Returns into Risk Parity Portfolios: A New Framework for Asset Allocation, *Bankers, Markets & Investors*, 138, pp. 18-28.

[27] Roncalli, T. (2017), Alternative Risk Premia: What Do We Know?, in Jurczenko, E. (Ed.), *Factor Investing and Alternative Risk Premia*, ISTE Press – Elsevier.

[28] Sharpe, W.F. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, 19(3), pp. 425-442.

[29] Xiong, J.X., and Idzorek, T.M. (2011), The Impact of Skewness and Fat Tails on the Asset Allocation Decision, *Financial Analysts Journal*, 67(2), pp. 23-35.

A Mathematical results

A.1 Gaussian approximation of asset returns

Let $R_t = (R_{1,t}, \ldots, R_{n,t})$ be the asset returns. We remind that:

$$R_t = B_1 Y_1 + B_2 Y_2$$

where B_1 and $B_2 = 1 - B_1$ are Bernoulli random variable with probability π_1 and π_2, $Y_1 \sim \mathcal{N}(\mu_1, \Sigma_1)$ and $Y_2 \sim \mathcal{N}(\mu_2, \Sigma_2)$. It follows that:

$$\mathbb{E}[R_t] = \pi_1 \mu_1 + \pi_2 \mu_2$$

and:

$$\operatorname{cov}(R_t) = \mathbb{E}\left[R_t R_t^\top\right] - \mathbb{E}[R_t]\,\mathbb{E}[R_t]^\top$$

We have:

$$\begin{array}{rcl}
\mathbb{E}\left[R_t R_t^\top\right] & = & \mathbb{E}\left[B_1^2 Y_1 Y_1^\top + B_2^2 Y_2 Y_2^\top + 2 B_1 B_2 Y_1 Y_2^\top\right] \\
& = & \pi_1\left(\Sigma_1 + \mu_1 \mu_1^\top\right) + \pi_2\left(\Sigma_2 + \mu_2 \mu_2^\top\right) \\
& = & \pi_1 \Sigma_1 + \pi_2 \Sigma_2 + \pi_1 \mu_1 \mu_1^\top + \pi_2 \mu_2 \mu_2^\top
\end{array}$$

and:

$$\begin{array}{rcl}
\mathbb{E}[R_t]\,\mathbb{E}[R_t]^\top & = & (\pi_1 \mu_1 + \pi_2 \mu_2)(\pi_1 \mu_1 + \pi_2 \mu_2)^\top \\
& = & \pi_1^2 \mu_1 \mu_1^\top + \pi_2^2 \mu_2 \mu_2^\top + \pi_1 \pi_2\left(\mu_1 \mu_2^\top + \mu_2 \mu_1^\top\right)
\end{array}$$

We deduce that[18]:

$$\begin{array}{rcl}
\operatorname{cov}(R_t) & = & \pi_1 \Sigma_1 + \pi_2 \Sigma_2 + \pi_1 \mu_1 \mu_1^\top + \pi_2 \mu_2 \mu_2^\top - \\
& & \left(\pi_1^2 \mu_1 \mu_1^\top + \pi_2^2 \mu_2 \mu_2^\top + \pi_1 \pi_2\left(\mu_1 \mu_2^\top + \mu_2 \mu_1^\top\right)\right) \\
& = & \pi_1 \Sigma_1 + \pi_2 \Sigma_2 + \pi_1 \pi_2\left(\mu_1 \mu_1^\top + \mu_2 \mu_2^\top\right) - \\
& & \pi_1 \pi_2\left(\mu_1 \mu_2^\top + \mu_2 \mu_1^\top\right)
\end{array}$$

We deduce that the Gaussian approximation of asset returns is:

$$R_t \approx \bar{R}_t \sim \mathcal{N}\left(\bar{\mu}, \bar{\Sigma}\right)$$

where:

$$\bar{\mu} = \pi_1 \mu_1 + \pi_2 \mu_2$$

and:

$$\bar{\Sigma} = \pi_1 \Sigma_1 + \pi_2 \Sigma_2 + \pi_1 \pi_2 (\mu_1 - \mu_2)(\mu_1 - \mu_2)^\top$$

[18] We verify that the variance of portfolio returns is equal to:

$$\begin{array}{rcl}
\sigma^2(R_t(x)) & = & x^\top \operatorname{cov}(R_t)\, x \\
& = & \pi_1 \sigma_1^2(x) + \pi_2 \sigma_1^2(x) + \pi_1 \pi_2\left(\mu_1^2(x) + \mu_2^2(x)\right) - \\
& & \pi_1 \pi_2\left(\mu_1(x)\,\mu_2(x) + \mu_2(x)\,\mu_1(x)\right) \\
& = & \pi_1 \sigma_1^2(x) + \pi_2 \sigma_1^2(x) + \pi_1 \pi_2 (\mu_1(x) - \mu_2(x))^2
\end{array}$$

A.2 The thresholding approach

Let R_t be the vector of asset returns at time t. If R_t follows a Gaussian mixture model, the posterior probability of observing a stress at time t is equal to:

$$\hat{\pi}_{2,t} = \frac{\pi_2 \phi_n (R_t; \mu_2, \Sigma_2)}{\pi_1 \phi_n (R_t; \mu_1, \Sigma_1) + \pi_2 \phi_n (R_t; \mu_2, \Sigma_2)}$$

We assume that there is a stress at time t if the posterior probability $\hat{\pi}_{2,t}$ is greater than a given threshold $\pi_2^\star$:

$$\mathcal{S}_t = 1 \Leftrightarrow \hat{\pi}_{2,t} \geq \pi_2^\star$$

Knowing the parameters π_1, μ_1, Σ_1, π_2, μ_2, Σ_2, we can then detect all the dates t that correspond to a stress. However, in practice, the Gaussian mixture model takes the following form:

$$f(y) = \pi_1 \phi_n (y; \mu_{1,t}, \Sigma_{1,t}) + \pi_2 \phi_n (y; \mu_2, \Sigma_2)$$

where $\mu_{1,t}$ and $\Sigma_{1,t}$ are time-varying parameters. Most of the times, $\mu_{1,t}$ and $\Sigma_{1,t}$ correspond to rolling estimates using a one-year window. The algorithm for estimating the parameters $\mu_{1,t}$ and $\Sigma_{1,t}$ is then given by the following steps:

1. Let n_{rw} be the length of the rolling window. We initialize the filtering procedure at time $n_{rw} + 1$ with the Gaussian estimates.

2. At time t, we calculate the posterior jump probabilities:

$$\hat{\pi}_{2,t-s} = \frac{\pi_2 \phi_n (R_{t-s}; \mu_2, \Sigma_2)}{\pi_1 \phi_n \left(R_{t-s}, \hat{\mu}_{1,t-1}, \hat{\Sigma}_{1,t-1}\right) + \pi_2 \phi_n (R_{t-s}; \mu_2, \Sigma_2)}$$

   for $s = 0, \ldots, n_{rw} - 1$. The probabilities $\hat{\pi}_{2,t-s}$ are then based on the estimates $\hat{\mu}_{1,t-1}$ and $\hat{\Sigma}_{1,t-1}$ calculated at time $t-1$.

3. We update $\hat{\mu}_{t-1}$ and $\hat{\Sigma}_{t-1}$ by calculating the empirical mean and covariance matrix with the subsample of $\{R_{t-n_{rw}+1}, \ldots, R_t\}$. The observation $t-s$ is included in the subsample if it satisfies the rule $\hat{\pi}_{2,t-s} < \pi_2^\star$. These new estimates $\hat{\mu}_t$ and $\hat{\Sigma}_t$ will be valid in order to calculate the posterior probabilities $\hat{\pi}_{t+1-s}$ at time $t+1$.

4. We go back to Step 2.

B Additional results

B.1 Tables

Here are some additional results concerning the risk budgeting approach.

Table 14: Risk contribution of the EW portfolio ($\alpha = 95\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	29.53%	30.93%	30.51%	30.04%
Momentum	39.24%	30.12%	41.37%	38.98%
Carry	31.23%	38.95%	28.12%	30.98%

Table 15: Risk contribution of the EW portfolio ($\alpha = 99\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	29.53%	31.45%	30.51%	30.12%
Momentum	39.24%	28.85%	41.37%	33.38%
Carry	31.23%	39.70%	28.12%	36.51%

Table 16: Weights of ERC and Skew-ERC portfolios ($\alpha = 90\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	36.23%	35.86%	34.23%	34.53%
Momentum	29.24%	34.84%	27.52%	28.02%
Carry	34.53%	29.30%	38.25%	37.45%

Table 17: Weights of ERC and Skew-ERC portfolios ($\alpha = 95\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	36.23%	35.48%	34.23%	35.90%
Momentum	29.24%	35.53%	27.52%	29.10%
Carry	34.53%	28.99%	38.25%	35.00%

Table 18: Weights of ERC and Skew-ERC portfolios ($\alpha = 99\%$)

Strategy	Static		Dynamic	
	Gaussian	GMM	Gaussian	GMM
Risk parity	36.23%	34.93%	34.23%	36.35%
Momentum	29.24%	36.52%	27.52%	33.05%
Carry	34.53%	28.55%	38.25%	30.60%

In the case of the EW portfolio, we notice that the risk allocation based on the Gaussian model may significantly differ from the risk allocation based on the Gaussian mixture model (GMM). Moreover, we also observe some large differences between static and dynamic results in the GMM case. However, these differences reduce as the confidence level α increases.

The weights obtained for the ERC portfolio are in line with the results obtained for the EW portfolio. Generally, the ERC portfolio overweights the carry strategy with respect to the Skew-ERC portfolio. It is remarkable that the trade-off is between the carry strategy and the momentum strategy, that is between a strategy with a negative skewness and a strategy with a positive skewness.

B.2 Figures

We report below some additional figures concerning the individual carry strategies, the risk allocation of the EW portfolio and the weight allocation of the ERC portfolio.

Figure 12: Cumulative performance of carry strategies

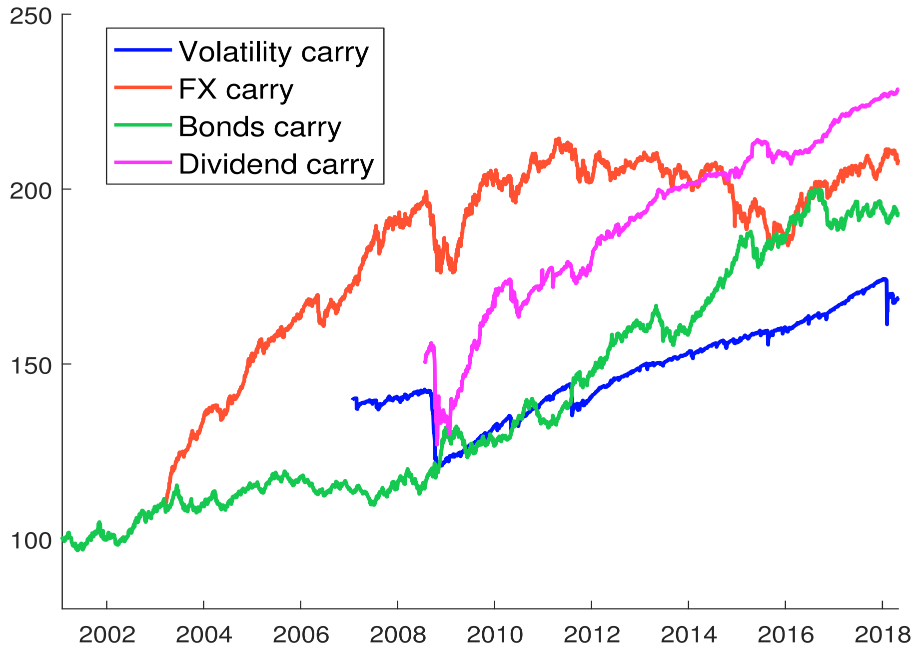


Figure 13: Weekly stress scenario for a four-year return period (carry strategies)

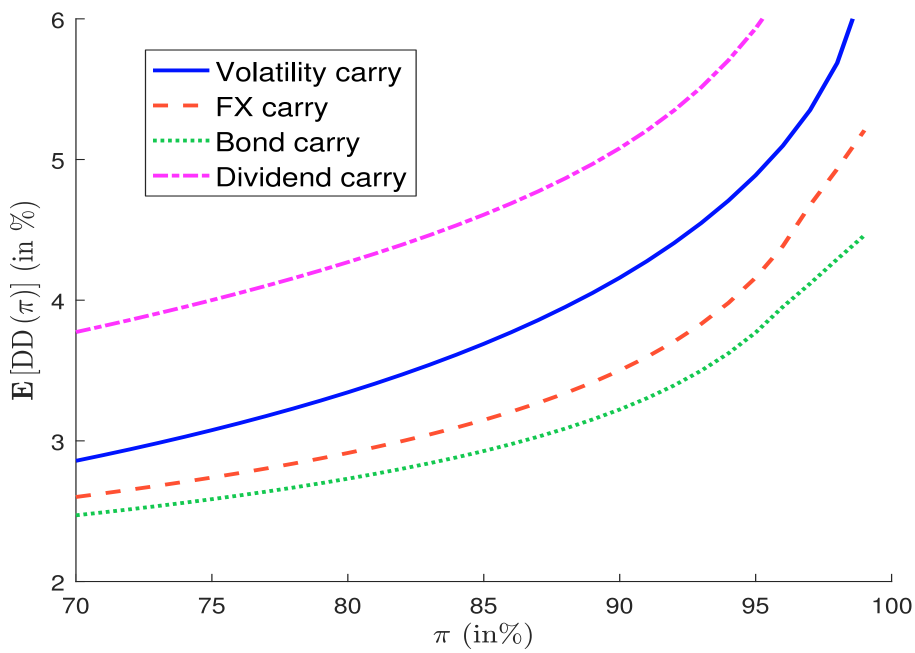


Figure 14: Skew measure for a four-year return period (carry strategies)

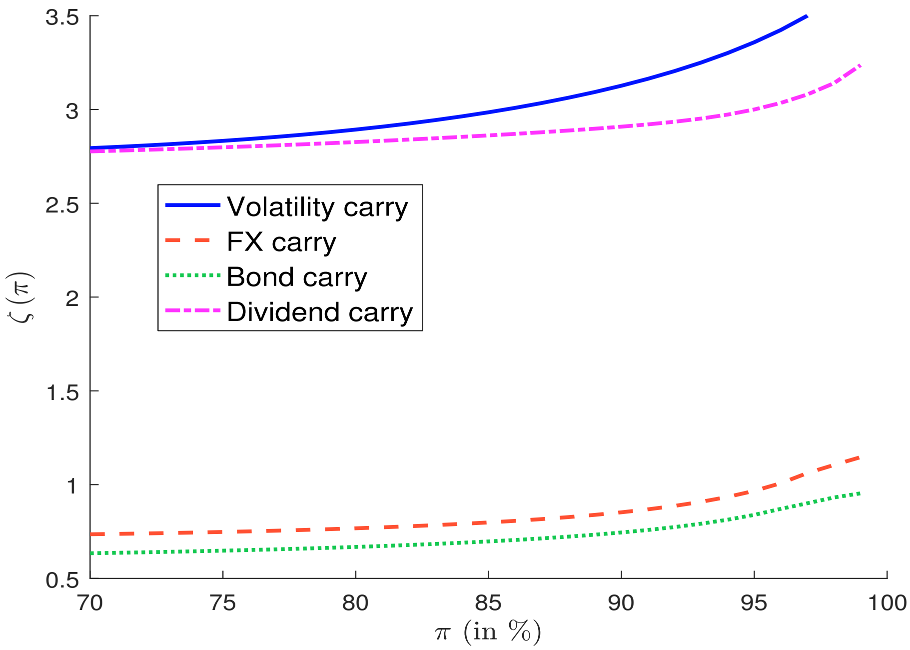


Figure 15: Posterior probability $\hat{\pi}_{2,t}$ (in %)

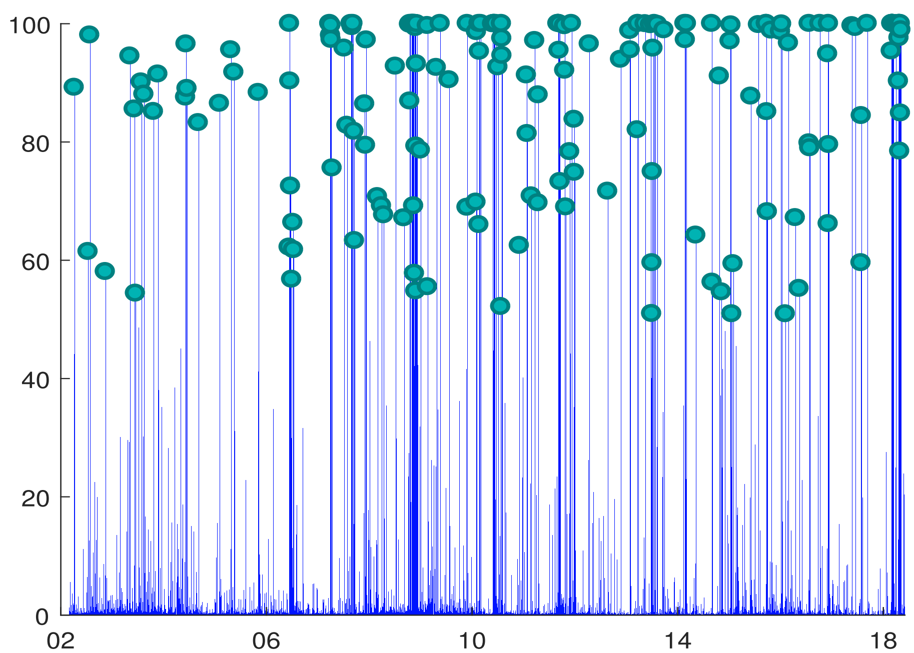


Figure 16: Risk contribution of the EW portfolio (GMM, no thresholding, $\alpha = 90\%$)

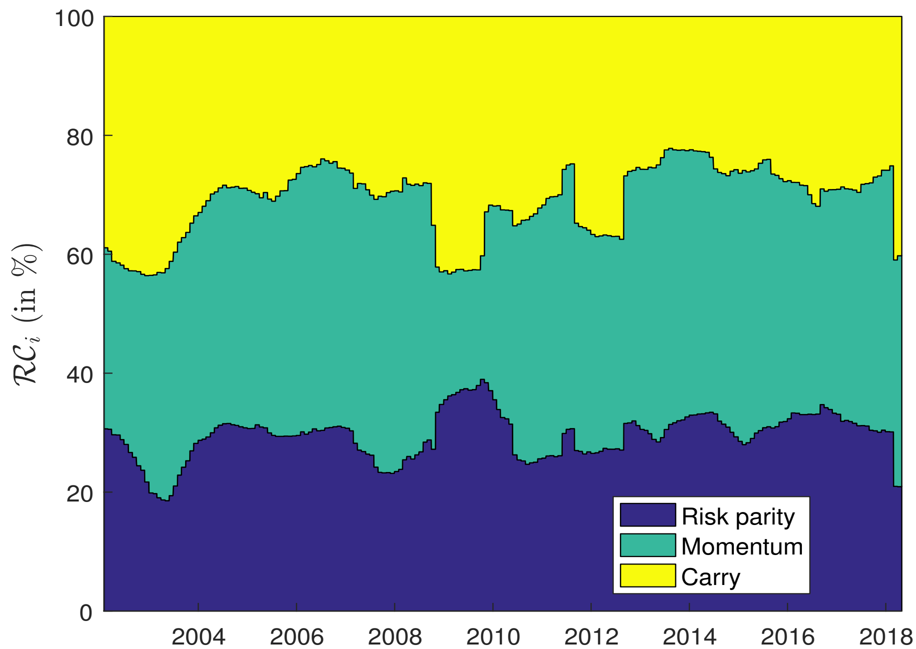


Figure 17: Risk contribution of the EW portfolio (GMMl, $\alpha = 99\%$)

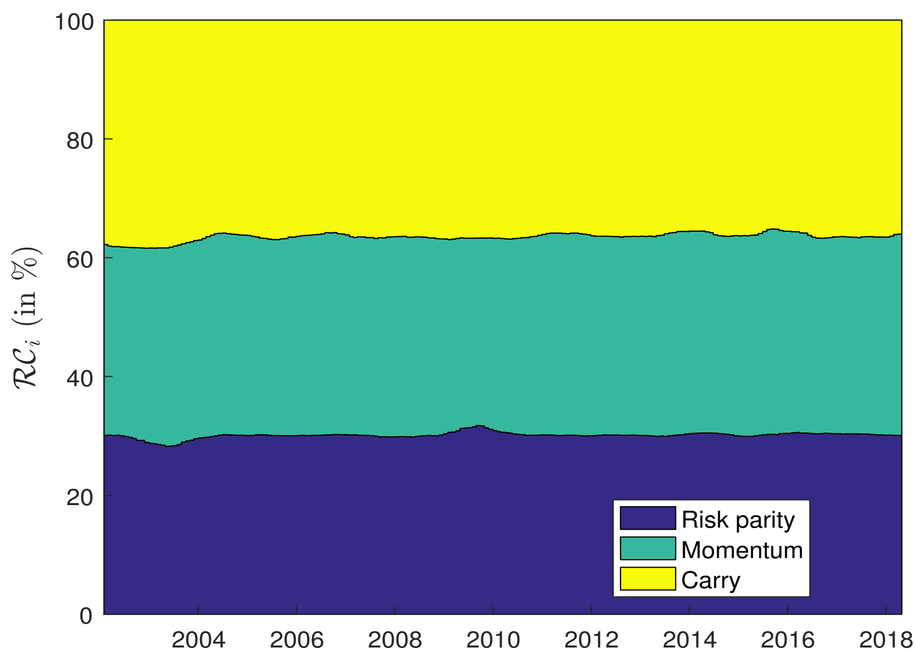


Figure 18: Weight allocation of the ERC portfolio

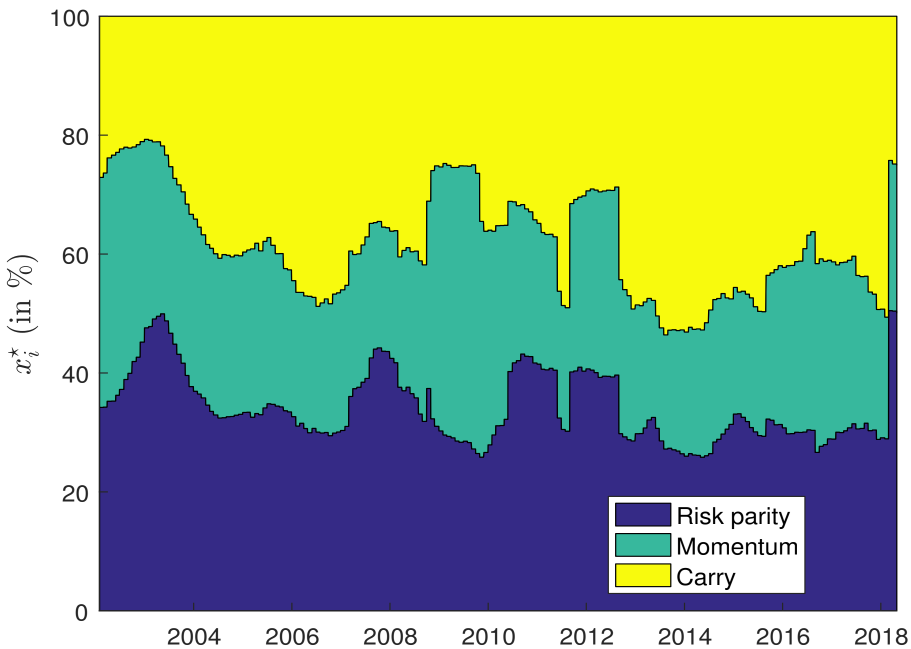


Figure 19: Weight allocation of the Skew-ERC portfolio ($\alpha = 90\%$)

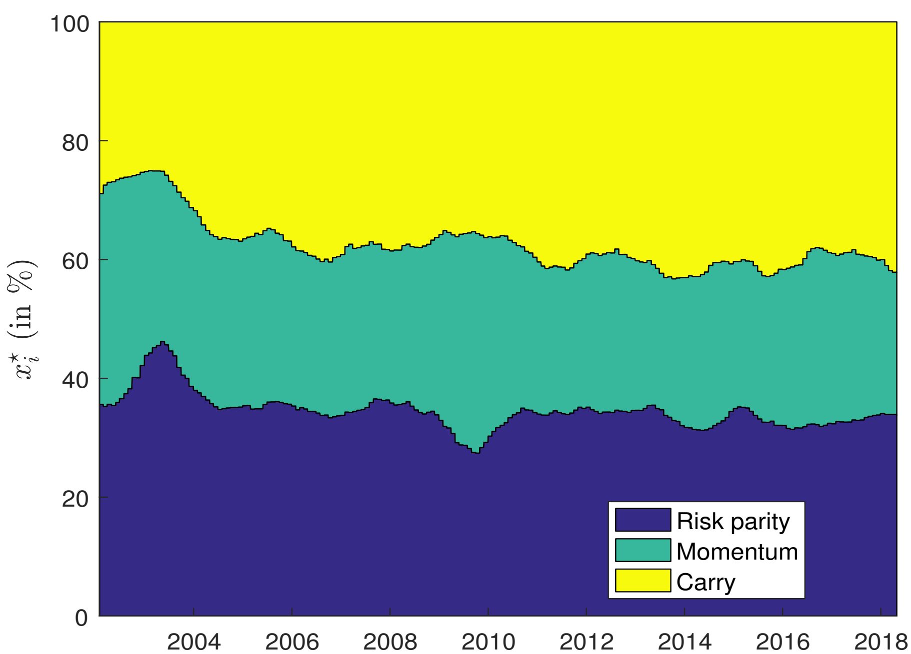


Figure 20: Weight difference between ERC and Skew-ERC portfolios ($\alpha = 95\%$)

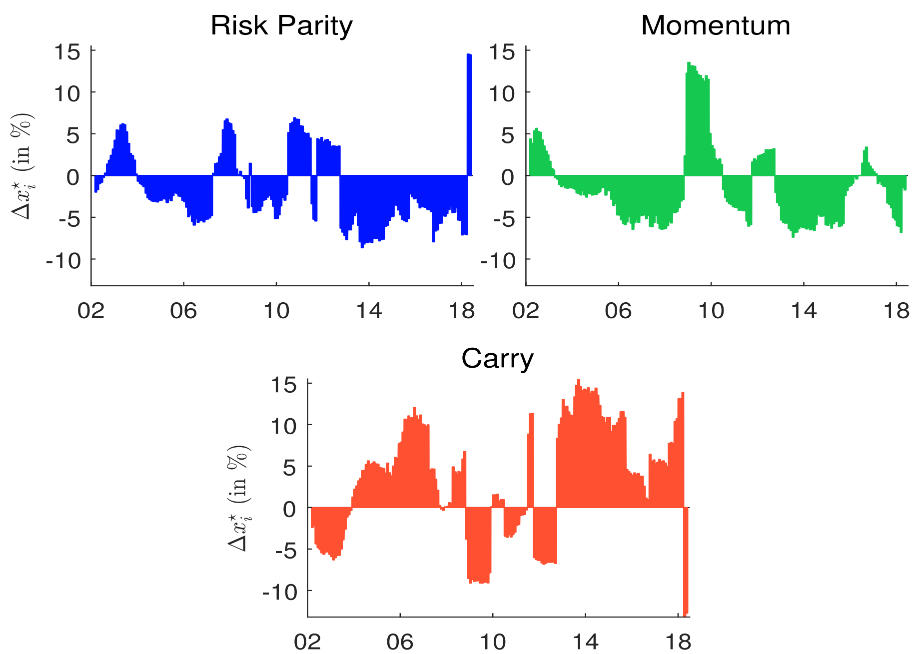


Figure 21: Weight difference between ERC and Skew-ERC portfolios ($\alpha = 99\%$)