Portfolio Construction for Practitioners: Sizing Strategies with Skewed Returns

Farzin Barekat*

May 28, 2019

Abstract

Mean-variance optimization (MVO) is a popular framework for portfolio allocation due to its tractability and intuitive concept. However, MVO has several pitfalls; in particular, it does not take into account characteristics of the strategies' returns beyond their means and volatilities/correlations, which makes it unsuitable for sizing strategies with for example significant left-tail (e.g. volatility selling strategies). In this paper, we propose an extension to the MVO framework that makes it more appropriate for sizing strategies with skewed returns, but at the same time is still intuitive and simple to use in practice. The portfolio manager specifies a penalizing factor for the tail events, and the framework determines the optimal risk allocation based on this specification. The proposed framework has three important features: 1) it is intuitive, interpretable and simple to use in practice, 2) when returns of the underlying strategies are jointly normally distributed, the proposed framework outputs the exact same solution as the MVO, and 3) the formulation is in the form of a convex optimization, which guarantees a unique optimal solution that is easy to compute.

1 Introduction

Mean-variance optimization (MVO) pioneered by Markowitz [4] is the dominant framework for portfolio allocation. The MVO approach is intuitive and tractable, which has led to its popularity. Moreover, it is well known that if strategies' returns are jointly normally distributed, then the optimal allocation given by MVO is the same as maximized expected utility for any utility function (see for example [5]).

On the other hand, in practice many strategies do not have normally distributed returns and their volatilities are not a sufficient measure of their risk. For example, suppose a strategy that returns 1% with probability 0.9 and -9% with probability 0.1. A simple calculation yields that the volatility of the strategy is 3%. Therefore, a junior portfolio

^{*}Assistant Adjunct Professor at UCLA Math Department. Former Senior Vice President at PIMCO.

manager might attribute a return of -9% to a 3x standard deviation move with occurrence probability of 0.13%; however, by definition, the probability of observing this return is 10%!

As the above example shows, using volatility as the sole measure of risk in portfolio construction, especially for strategies with skewed returns, is at best misleading and at worst can result into catastrophic outcomes. Yet, MVO only uses expected returns and volatilities/correlations for portfolio allocation and ignores other characteristics of the return distribution, which can yield counterintuitive risk allocations. Consider the following example: suppose two strategies have the same expected returns and volatilities and are uncorrelated to each other and the rest of the portfolio, but the returns for one of the strategies are normally distributed while the other one has significant left tail. The MVO framework would allocate same amount of risk to both strategies; however, conceivably a prudent portfolio manager would not allocate the same amount of risk to these strategies as their return distribution have very different characteristics.

Therefore, a natural question for portfolio managers is: how to appropriately size strategies whose returns are not normally distributed (in particular are left-skewed)? This question especially becomes relevant for portfolio managers who invest in volatility selling strategies whose returns have significant left tail.

The goal of this paper is to address the question posed above both **qualitatively** and **quantitatively**. To that end, we propose an extension to the MVO framework for portfolio allocation that we call the *mean-variance optimization with expected shortfall penalization (ESP)*. The proposed framework is not too oversimplified and takes into account other characteristics of the strategies' return distribution than their means and volatilities/correlations, but at the same time it is not too overcomplicated that renders it unwieldy to use in practice.

The framework proposed here has practical consequences for portfolio managers: given portfolio manager's aversion to tail events, it determines exactly how much risk should be allocated to left-skewed strategies. As a simple rule of thumb, in typical investment situations, the framework advocates to decrease the size of strategies with significant left tail to around 2/3 of their MVO allocations (see figure 4).

1.1 Related work

The framework proposed here is closely related to other optimization approaches that use expected shortfall (some examples are [2, 6, 7]). Our main contribution is to formulate the problem in such a way that the parameters have clear interpretation and help portfolio managers to build intuition for the results.

1.2 Organization

The rest of the paper is organized as follows: Section 2 introduces several notations and briefly reviews some of the concepts that are used throughout the paper. Section 3 describes the proposed framework that extends MVO. In section 4, we build intuition for the ESP framework by looking at its results in the context of a simple example.

2 Notations and Preliminaries

2.1 Mean-variance optimization

For a given investment horizon, we use the following notation:

- σ_i : the volatility of strategy *i*.
- r_i : return of strategy *i*.
- w_i : the weight of strategy *i* in the portfolio.
- $\tilde{r}_i = \frac{r_i}{\sigma_i}$: volatility-normalized return of strategy *i*.
- $k_i = w_i \sigma_i$: the risk allocated to strategy *i* in the portfolio.
- Ω : correlation matrix among the strategies.
- σ_{Π} : portfolio's risk budget.

Observe that the return of the portfolio over the investment horizon attributed to strategy i is equal to $\tilde{r}_i k_i = r_i w_i$. Bold face letters are used to denote vectors, e.g. **k** denotes vector of portfolio's risk allocations to different strategies.

The MVO risk allocation \mathbf{k}_{MVO} is given by solving the following optimization problem

$$\mathbf{k}_{MVO} = \operatorname*{argmax}_{\mathbf{k}} \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k}] \quad \text{subject to } \mathbf{k}^T \Omega \mathbf{k} \le \sigma_{\Pi}^2.$$
(1)

It is well known that when the correlation matrix is ill-conditioned, \mathbf{k}_{MVO} becomes unstable (see [3] for detailed discussion on this); however, this does not become an issue when the portfolio is comprised of strategies with low correlations.

2.2 Two-point distribution

A two-point distribution is a random variable with the following possible outcomes:

$$\begin{cases} r_d & \text{with probability } p_d, \\ r_u & \text{with probability } p_u = 1 - p_d. \end{cases}$$
(2)

This distribution can be used to represent returns of strategies that are small positive most of the times but have a huge drawdown every once in a while , e.g. volatility selling strategies, shorting natural gas March-April spread, etc.

Appendix A provides some statistics for the two-point distribution. As an example, figure 1 plots the ratio r_d/r_u in terms of the Sharpe ratio of the two-point distribution for several values of p_d .



Figure 1: Ratio of r_d to r_u for the two-point distribution (2).

2.3 Expected shortfall deviation (ESD)

For continuous random variable R, let $Q_R(p)$ denote its quantile function at probability level p, that is $\Pr[R \leq Q_R(p)] = p$ (in words, with probability p, R realizes values less than or equal to $Q_R(p)$). The expected shortfall (ES) of R at probability level p, also known as conditional value at risk¹ (CVaR), is given by²

$$ES_p(R) = \frac{1}{p} \int_0^p Q_R(u) du = \mathbb{E}[R|R \le Q_R(p)].$$
(3)

Observe that when R is the random variable for portfolio's return, $ES_p(R)$ translates into the average return in the worst 100p% cases. Figure 2 shows examples of the difference between mean and expected shortfall for several portfolios (normalized by the volatility of portfolios). Conceivably, given two portfolios with the same expected return and volatility, the portfolio for which this difference is smaller is preferred. To that end, we define *expected shortfall deviation (ESD)*, as the difference between the mean and expected shortfall normalized with corresponding probability:

$$ESD_p(R) = \frac{p}{1-p} \left[\mathbb{E}[R] - ES_p(R) \right].$$
(4)

Figure 3 plots the values of ESD (normalized by the portfolio's volatility) for the same examples as in figure 2. Consequently, given two portfolios with the same expected return and volatility, the portfolio with lower ESD is preferred.

¹The continuity assumption of R is critical here, otherwise this is not true (see for example [1]).

²Some authors use slightly variant definitions for the expected shortfall: with a minus sign in front of the integral or with 1 - p in place of p.



Figure 2: Difference between mean and expected shortfall of portfolio's return in terms of portfolio's volatility (i.e. $(\mathbb{E}[R] - ES_p(R))/\sqrt{\operatorname{Var}[R]})$ for different probability levels p. Here portfolio consist of allocating equal risk to two uncorrelated strategies A and B, where both strategies have Sharpe 1 (i.e. the optimal allocation based on the MVO problem (1)). Returns of strategy B have normal distribution. The line labeled "normal" shows the case where returns of strategy A have normal distribution. The other lines show the cases where strategy A has two-point distribution with corresponding p_d .



Figure 3: The ESD of portfolio's return in terms of portfolio's volatility (i.e. $(ESD_p(R)/\sqrt{\operatorname{Var}[R]})$ for different probability levels p. Here portfolio consist of allocating equal risk to two uncorrelated strategies A and B, where both strategies have Sharpe 1 (i.e. the optimal allocation based on the MVO problem (1)). Returns of strategy B have normal distribution. The line labeled "normal" shows the case where returns of strategy A have normal distribution. The other lines show the cases where strategy A has two-point distribution with corresponding p_d .

3 Mean Variance Optimization with Expected Shortfall Penalization

This section formulates a simple extension to the MVO framework that has several desirable properties. Again, our main goal is to propose a framework that is more realistic than the MVO framework for investment, but is still straightforward to understand and implement.

As mentioned in the previous section, given two portfolio with the same expected return and volatility, the portfolio with lower ESD is preferred. Motivated by this observation, we modify the MVO problem (1) by penalizing the objective function with ESD:

For $1 \leq \gamma \leq \frac{1}{n}$, let

$$\mathbf{k}_{ESP} = \operatorname*{argmax}_{\mathbf{k}} \left\{ \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k}] - (\gamma - 1) ESD_p[\tilde{\mathbf{r}}^T \mathbf{k}] \right\} \quad \text{subject to } \mathbf{k}^T \Omega \mathbf{k} \le \sigma_{\Pi}^2.$$
(5)

In general, parameters γ and p in (5), which determine the penalization of the tail events, are specified by the portfolio manager based on their level of risk aversion. For example, a portfolio manager with a strong track record has higher tolerance to drawdowns and might use a lower value of γ than a novice portfolio manager who is worried of blow ups.

There are several important features that make the ESP problem (5) ergonomic for risk allocations:

- 1. As shown in theorem 3.1 below, parameter γ has a clear interpretation as the overweighting factor of the tail events. The objective function in (5) is akin to taking expectation of $\tilde{\mathbf{r}}^T \mathbf{k}$ with respect to the probability measure for which the probability of the tail events are increased from p to γp . In particular, setting $\gamma = 1$, reduces the ESP problem to the MVO problem for any distribution of returns $\tilde{\mathbf{r}}$.
- 2. In the case of multivariate normal distribution $\tilde{\mathbf{r}}$, it can be shown that $\mathbf{k}_{MVO} = \mathbf{k}_{ESP}$ for all values of γ and p. Therefore, the ESP coincides with the MVO for jointly normally distributed returns, but when returns are not normally distributed, the ESP enhances the MVO by taking into account the tail characteristics of the returns.
- 3. Note that function $ESD_p(\cdot)$ is a convex function because $ES_p(\cdot)$ is a concave function (see for example [1]). Consequently, the ESP problem (5) is a convex optimization problem³, which guarantees existence and uniqueness of the optimal solution. Moreover, [6] presents a method to write ES optimization (and therefore ESD optimization) as a linear programing problem, which provides a method to efficiently solve (5).
- 4. It is easy to show that the relative risk between strategies in the optimal allocation \mathbf{k}_{ESP} is independents of the overall volatility of the portfolio σ_{Π} (similar to the behavior seen in the MVO framework).

 $^{^{3}}$ We are maximizing a concave objective function over a convex feasible set.

The next theorem motivates the interpretation of γ in the ESP problem (5) as the overweighting factor of the tail events. In some sense, factor γ is analogous to the pricing kernel $\frac{dQ}{dP}$ used to change the probability measure from the physical world to the risk neutral world.

Theorem 3.1 Denote the set of left tail events (with probability p) for portfolio returns $\tilde{\mathbf{r}}^T \mathbf{k}$ by \mathbb{T} and its complement by \mathbb{T}^c . The objective function of the MVO problem (1) is equal to

$$\mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k}] = p \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k} | \mathbb{T}] + (1-p) \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k} | \mathbb{T}^c].$$
(6)

On the other hand, the objective function of the ESP problem (5) is equal to

$$\mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k}] - (\gamma - 1) ESD_p(\tilde{\mathbf{r}}^T \mathbf{k}) = p^* \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k} | \mathbb{T}] + (1 - p^*) \mathbb{E}[\tilde{\mathbf{r}}^T \mathbf{k} | \mathbb{T}^c]$$
(7)

where $p^* = \gamma p$.

Proof: See appendix B.

4 Numerical Results

This section builds intuition for the results of the ESP framework by applying it to a simple example. Suppose we want to allocate risk to strategy A with left-skewed returns. We use the two-point distribution (2) with $p_d < p_u$ to represent return distribution of A (think of A as a strategy that returns a small positive most of the time, but occasionally has a huge drawdown, e.g. vol selling strategies, shorting natural gas March-April spread, etc.). Note that we could use any distribution with a heavy left tail in the ESP framework; here we use the two-point distribution for ease of comprehension. We use B to represent the rest of the strategies in the portfolio, which is assumed to be well diversified and therefore it is reasonable to assume B has normally distributed returns. For ease of exposition, we assume that strategy B has Sharpe ratio of 1, and strategies A and B are uncorrelated. In typical investment scenarios, strategy A would have a lower Sharpe ratio than strategy B, since the latter is made up of many well-diversified strategies. Therefore, we only consider cases where Sharpe ratio of A is smaller or equal than one⁴.

Figure 4 shows the ratio of the risk that the ESP framework allocates to strategy A and B as Sharpe of A varies. Here we plot the results for four different types of return distributions for strategy A (i.e. normal distribution, as well as two-point distributions with $p_d = 20\%, 10\%, 5\%$). We highlight the following observations in figure 4:

• As noted in section 3, when both A and B have normally distributed returns, the ESP framework makes the same allocation as the MVO framework (hence the solid black line with slope 1).

⁴In appendix C, we consider what happens when strategy A has higher Sharpe than B even though this is not a typical investment scenario.

• When strategy A has left-skewed return, the ESP framework underweights it in comparison to the MVO framework. This should appear very intuitive to portfolio managers: left-skewed assets have more inherent risk. The more left-skewed the returns (characterized by lower p_d ; see figure 1), the lower the allocated risk. The advantage of using the ESP framework is to quantitatively determine the appropriate amount of underweighting required.



Figure 4: Ratio of risk allocation between strategies A and B using the ESP portfolio optimization (5) with parameters $\gamma = 4$ and p = 0.05 for different values of the Sharpe ratio of strategy A. Here portfolio consists of two uncorrelated strategies A and B. Returns of strategy B have normal distribution with Sharpe 1. The line labeled "normal" shows the case where returns of strategy A have normal distribution. The other lines show the cases where strategy A has two-point distribution with corresponding p_d .

As an example, figure 5 shows the profile of returns for MVO and ESP allocations. It is noteworthy that the ESP portfolio does not sacrifice too much on the Sharpe ratio (1.37 versus MVO's Sharpe ratio of 1.41), but manages to improve the expected shortfall of the portfolio by more than half of volatility (i.e. ESP's expected shortfall of -1.1 versus MVO's expected shortfall of -1.7).

Figure 6 shows the sensitivity of allocation between strategies A and B to different values of the overweighting factor of tail events γ . When $\gamma = 1$, ESP reduces to the MVO framework; therefore, regardless of the return distribution of strategy A, the risk allocated to A is half of the risk allocated to B (since Sharpe of A is half of Sharpe of B and the two strategies are uncorrelated). As noted in section 3, in the case when returns of strategy Aare normally distributed, the ESP allocation of risk is independent of the value of γ (hence the black solid horizontal line in figure 6). Figure 7 shows the sensitivity of allocation between strategies A and B to different values of the ESD probability level p. When p = 0, ESP reduces to the MVO framework; therefore, regardless of the return distribution of strategy A, the risk allocated to A is half of the risk allocated to B (since Sharpe of A is half of Sharpe of B and the two strategies are uncorrelated). As noted in section 3, in the case when returns of strategy A are normally distributed, the ESP allocation of risk is independent of the value of p (hence the black solid horizontal line in figure 7).



Figure 5: Distribution of returns for portfolios constructed with the MVO portfolio optimization (1) and the ESP portfolio optimization (5). Here portfolios have unit volatility and consist of two uncorrelated strategies A and B. Returns of strategy B have normal distribution with Sharpe 1. Returns of strategy A have two-point distribution with Sharpe 1 and $p_d = 0.05$. The parameters used for the ESP optimization are p = 0.05 and $\gamma = 4$. The black dashed vertical line corresponds to 0.05 quantile for the MVO portfolio. The expected shortfalls with probability 5% for the MVO and ESP portfolios are -1.7 and -1.1, respectively (denoted by the dotted lines). The Sharpe ratios of the MVO and ESP portfolios are 1.41 and 1.37, respectively.

Acknowledgement

The author benefited greatly from constructive feedback and insightful discussions with Yuri Garbuzov for this project.



Figure 6: Ratio of risk allocation between strategies A and B using the ESP portfolio optimization (5) for different values of γ and with p = 0.05. Here portfolio consist of two uncorrelated strategies A and B. Returns of strategy B have normal distribution with Sharpe 1. Strategy A has Sharpe 0.5. The line labeled "normal" shows the case where returns of strategy A have normal distribution. The other lines show the cases where strategy A has two-point distribution with corresponding p_d .



Figure 7: Ratio of risk allocation between strategies A and B using the ESP portfolio optimization (5) for different values of p and with $\gamma = 4$. Here portfolio consist of two uncorrelated strategies A and B. Returns of strategy B have normal distribution with Sharpe 1. Strategy A has Sharpe 0.5. The line labeled "normal" shows the case where returns of strategy A have normal distribution. The other lines show the cases where strategy A has two-point distribution with corresponding p_d .

A Some statistics for the two-point distribution

Consider the two-point distribution

$$\begin{cases} r_d & \text{with probability } p_d, \\ r_u & \text{with probability } p_u = 1 - p_d. \end{cases}$$

The mean μ of the distribution is given by

$$\mu = p_u r_u + p_d r_d,$$

and the variance σ^2 is given by

$$\sigma^2 = p_d p_u (r_u - r_d)^2.$$

Given ratio r_d/r_u , Sharpe S of the two-point distribution is given by

$$S = \frac{1 + (r_d/r_u)(p_d/p_u)}{\sqrt{p_d/p_u}(1 - r_d/r_u)}$$

Conversely, given Sharpe S, we can back-out the ratio r_d/r_u for the two-point distribution:

$$\frac{r_d}{r_u} = \frac{S\sqrt{p_d/p_u} - 1}{S\sqrt{p_d/p_u} + p_d/p_u}$$

B Interpretation of factor γ

In this appendix, we provide proof for theorem 3.1, which justifies interpretation of the parameter γ as the overweighting factor of the tail events.

Proof of Theorem 3.1: Set $R = \tilde{\mathbf{r}}^T \mathbf{k}$. Observe that

$$\mathbb{T} = \{ R \le Q_R(p) \}, \quad \text{and} \quad \mathbb{T}^c = \{ R > Q_R(p) \}.$$

By the law of iterated expectation,

$$\mathbb{E}[R] = p \mathbb{E}[R|\mathbb{T}] + (1-p) \mathbb{E}[R|\mathbb{T}^c], \qquad (B.1)$$

which yields identity (6). It remains to show (7). Observe that

$$\mathbb{E}[R] - (\gamma - 1)ESD_p(R)$$

= $\mathbb{E}[R] - \frac{(\gamma - 1)p}{1 - p} [\mathbb{E}[R] - ES_p(R)]$
= $p \mathbb{E}[R|\mathbb{T}] + (1 - p) \mathbb{E}[R|\mathbb{T}^c] - \frac{(\gamma - 1)p}{1 - p} [(1 - p) \mathbb{E}[R|\mathbb{T}^c] - (1 - p) \mathbb{E}[R|\mathbb{T}]]$
= $\gamma p \mathbb{E}[R|\mathbb{T}] + (1 - \gamma p) \mathbb{E}[R|\mathbb{T}^c],$

where we used (4) for the first equality, and used (B.1) and (3) for the second equality. Finally, setting $p^* = \gamma p$ in the above equation yields

$$\mathbb{E}[R] - (\gamma - 1)ESD_p(R) = p^* \mathbb{E}[R|\mathbb{T}] + (1 - p^*) \mathbb{E}[R|\mathbb{T}^c]$$
(B.2)
atity (7).

which yields identity (7).

C The skewed-return strategy is the main source of alpha in the portfolio

In this appendix we investigate the result of the ESP framework when the Sharpe ratio of A is greater than the Sharpe ratio of B (i.e. when the strategy with the skewed-return has superior performance than the rest of the portfolio combined). Note that this is **not** a typical investment situation; nevertheless, the result may seem counterintuitive at first, which prompted us to mention it here. Figure 8 plots the same result as figure 4; however, with the range of the Sharpe ratio of strategy A extended to 2.5. As figure 8 shows, the underweighting of the skewed-return strategy becomes smaller as its Sharpe increases. In some cases, the ESP framework would even overweight the skewed-return strategy in comparison to the MVO allocation, which might seem surprising and counterintuitive at first glance.

To elucidate this behavior, we show in figure 9 the return distribution of the combined portfolio in a scenario where the ESP framework would slightly overweight the skewedreturn strategy. Compare this figure with figure 5: In figure 5, the performance of the strategies are comparable, which results into the ESP framework underweighting the skewreturn strategy to attenuates the tail of the ESP portfolio in comparison to the MVO portfolio. On the other hand, figure 9 shows an example where the skewed-return strategy has much better performance than the strategy with normally distributed returns. Comparing the distributions of the MVO and the ESP portfolios in figure 9 suggest that the ESP portfolio prefers to slightly overweight the skewed-return strategy as it does not significantly change the tail behavior of the distribution, but increases the probability of the two modes of the return distribution. **Intuitively, when the skewed-return strategy has much better performance than the strategy with normally distributed returns, the ESP framework underweights the latter strategy to not dilute the strong performance of the skewed-return strategy.**

References

- C. Acerbi and D. Tasche. On the coherence of expected shortfall. Journal of Banking and Fiance, 26:1487–1503, 2002.
- [2] L.R. Goldberg, M.Y. Hayes, and O. Mahmoud. Minimizing shortfall. Quantitative Fiance, 13(10):1533-1545, 2013.
- [3] M. Lopez de Prado. Building diversified portfolio that outperform out-of-sample. *Jour*nal of Portfolio Management, 42(4):59–69, 2016.
- [4] H.M. Markowitz. Portfolio selection. Journal of Finance, 7(1):77–91, 1952.
- [5] A. Meucci. Risk and Asset Allocation. Springer, 2007.
- [6] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:493–517, 2000.



Figure 8: Same setting as in figure 4 with different range for the x-axis. Note that Sharpe A greater than 1 is **not** a typical investment situation, because we are assuming strategy B, which is made up of the remaining well-diversified strategies in the portfolio, has Sharpe 1. This figure is included to highlight the point mentioned in appendix C.

[7] R.T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. Conditional Value-at-Risk for General Loss Distributions, 26:1443–1471, 2002.



Figure 9: Distribution of returns for portfolios constructed with the MVO portfolio optimization (1) and the ESP portfolio optimization (5). Here portfolios have unit volatility and consist of two uncorrelated strategies A and B. Returns of strategy B have normal distribution with Sharpe 1. Returns of strategy A have two-point distribution with Sharpe 2.5 and $p_d = 0.2$. The parameters used for the ESP optimization are p = 0.05 and $\gamma = 4$. The black dashed vertical line corresponds to 0.05 quantile for the MVO portfolio. The expected shortfalls with probability 5% for the MVO and ESP portfolios are 0.36 and 0.38, respectively (denoted by the dotted lines). The Sharpe ratios of the MVO and ESP portfolios are 2.692 and 2.691, respectively. **Note that this is not a typical investment situation**, because we are assuming strategy A has much greater Sharpe than strategy Bthat is made up of the remaining well-diversified strategies in the portfolio. This figure is included to highlight the point mentioned in appendix C.