Robust Portfolio Optimization with Value-At-Risk Adjusted Sharpe Ratios

Geng Deng, PhD, CFA, FRM[*] Tim Dulaney, PhD, FRM[†]
Craig McCann, PhD, CFA[‡] Olivia Wang, PhD, FRM[§]

November 5, 2013

Abstract

We propose a robust portfolio optimization approach based on Value-at-Risk (VaR) adjusted Sharpe ratios. Traditional Sharpe ratio estimates using a limited series of historical returns are subject to estimation errors. Portfolio optimization based on traditional Sharpe ratios ignores this uncertainty and, as a result, is not robust. In this paper, we propose a robust portfolio optimization model that selects the portfolio with the largest worse-case-scenario Sharpe ratio within a given confidence interval. We show that this framework is equivalent to maximizing the Sharpe ratio reduced by a quantity proportional to the standard deviation in the Sharpe ratio estimator. We highlight the relationship between the VaR-adjusted Sharpe ratios and other modified Sharpe ratios proposed in the literature. In addition, we present both numerical and empirical results comparing optimal portfolios generated by the approach advocated here with those generated by both the traditional and the alternative optimization approaches.

Keywords: Sharpe Ratio, Portfolio Optimization, Robust Optimization, VAR

1 Introduction

In Markowitz's mean-variance framework, optimal portfolios have minimum variance given an expected return, or equivalently maximum expected return given a variance.

[*]Director of Research, Office: (703) 890-0741, Email: GengDeng@slcg.com.
[†]Senior Financial Economist, Office: (703) 539-6777, Email: TimDulaney@slcg.com.
[‡]President, Office: (703) 539-6760, Email: CraigMcCann@slcg.com.
[§]Senior Associate at Ernst & Young, Office: (415) 894-8922, Email:oliviaykwang@gmail.com.

With the ability to borrow and lend at the risk-free rate, the separation property (Bodie et al., 2010) states that the optimal mean-variance portfolio of risky assets is the portfolio with the highest Sharpe ratio, defined as the ratio of the expected excess return to the standard deviation of excess returns. This ratio measures the excess return a portfolio is expected to gain for each unit of risk (or volatility) associated with this excess return.

Practitioners observe a set of excess returns and compute the Sharpe ratio by dividing the sample mean by the sample standard deviation. This procedure is subject to estimation errors including data limitations, negative skewness and positive excess kurtosis of returns. In addition, the ex-post Sharpe ratio implicitly assumes that the returns of the asset under consideration are independent and identically distributed (i.i.d) normal random variables. However, these assumptions are violated by real world financial data. For example, hedge fund return distributions are often negatively skewed and exhibit positive excess kurtosis. Furthermore, the presence of measurable and statistically significant serial autocorrelations indicates that the returns are not independent and heteroskedasticity provides evidence against the identicality of serial returns. The Sharp ratio estimation errors lead to potentially misleading results when the traditional Sharpe ratio is used to determine the optimal portfolio. These observations call into question the appropriateness of using the Sharpe ratio in portfolio allocation decisions.

This paper proposes a robust optimal portfolio allocation approach using a modified version of the traditional Sharpe ratio that we refer to as the "Value-at-Risk (VaR) adjusted Sharpe ratio" (VaRSR). The VaRSR explicitly takes into account the uncertainty involved in estimating the Sharpe ratio, and takes a more conservative view than the traditional Sharpe ratio by including the effects of higher order moments of the return distribution. The approach is specially adapted to assets with non-normal return distributions and limited data. We show that the portfolio allocation approach naturally fits into a max-min robust optimization framework and, as a result, is more reliable than traditional portfolio optimization using Sharpe ratios.

Value-at-Risk (VaR) is widely used in risk management. When applied to a return distribution, VaR estimates the maximum loss on an investment with a prescribed confidence level. In this paper, we apply VaR to the Sharpe ratio by examining the lowest Sharpe ratio consistent with the data in the observation period for a given confidence level. In other words, we use the lower bound of an estimated confidence interval for a Sharpe ratio, instead of the estimated Sharpe ratio itself. By doing so, we limit the probability that the underlying Sharpe ratio estimated using the historical returns is substantially smaller than the measured Sharpe ratio. Intuitively this means portfolio managers should choose portfolios with relatively large estimated Sharpe ratios with a penalty for the uncertainty in such estimated ratios.

Our VaRSR measure contrasts with another alternative to the Sharpe ratio, the "probabilistic Sharpe ratio" (PSR), proposed by Bailey and López de Prado (2011). Using an estimated distribution of the Sharpe ratio, the PSR computes the probability that a Sharpe ratio estimate exceeds a prescribed threshold. Bailey and López de Prado (2011) also introduce the concept of a "Sharpe ratio efficient frontier" which contains combinations of the estimated Sharpe ratio and the standard error of the Sharpe ratio estimate. This is analogous to the traditional efficient frontier which relates the expected excess return to the standard deviation of excess returns. Later we will show that solutions to the traditional Sharpe ratio optimization model, our VaRSR model and the PSR model are all located on the Sharpe ratio efficient frontier. The traditional portfolio optimization approach and the PSR approach are special cases of the more general approach presented here.

We test the effectiveness of our VaRSR approach with a numerical example involving a simple three-asset portfolio and simulated returns. We perform the test on a portfolio consisting of allocations to ten Dow Jones Credit Suisse Hedge Fund Indexes to show the benefits investors could realize by implementing our approach. In particular, we provide evidence that this strategy is effective in mitigating market volatility and volatility of the Sharpe ratio estimator without sacrificing realized returns. In each example, we compare our computed robust portfolio with the traditional and alternative robust Sharpe ratio portfolios.

The sections of the paper are arranged as follows. We begin with a review of the relevant literature in Section 2. In Section 3, we then include a discussion of the statistical properties that Sharpe ratio estimators inherit from the underlying return distribution. This section provides the distribution formulas for the Sharpe ratio estimates and provides a theoretical foundation for the VaRSR measure. In Section 4, we give a detailed introduction of our new measure and compare the measure to alternative approaches present in the literature. Section 5 discusses the details of our tests including simulations and the hedge fund portfolio. The final section is reserved for our conclusions.

2 Existing Literature

This paper stands at the intersection of two strands of literature. On the one hand, it is closely related to the discussion of non-normal return distributions and of data limitations in Sharpe ratio measurements. On the other hand, it naturally fits into the robust optimization framework.

Hodges (1998) and Zakamouline and Koekebakker (2009) define a "Generalized Sharpe Ratio" that takes into account the skewness (third moment) and kurtosis (fourth moment) of the observed historical return distribution. Lo (2002) derives the statistical behavior of observed Sharpe ratios under the assumption that returns are normally distributed. Mertens (2002) extends Lo's result by relaxing the normality assumption. Christie (2005) and Opdyke (2007) further relax the assumption of i.i.d. returns to include stationary and ergodic returns. Christie and Opdyke have shown that the Sharpe ratio estimator is asymptotically normally distributed even when the underlying returns are serially correlated or have time-varying conditional volatilities. These results make the construction of a VaRSR straightforward.

The portfolio selection problem using the VaRSR also presents an example of a robust portfolio optimization problem. Robust portfolio optimization incorporates the certainty with which the moments of the underlying return distribution are estimated from historical returns. Goldfarb and Iyengar (2003) define the concept of “uncertainty structures” for the estimates of expected returns and variances and show how to efficiently compute robust portfolio allocations with a desired level of confidence. Maximizing the worst-case Sharpe ratio is one of the robust portfolio optimization models presented in Goldfarb and Iyengar (2003). Tütüncü and Koenig (2004) generalize this approach and advocate the conservative portfolio selection program that maximizes the portfolios' returns in the worst-case scenario. These authors typically model the uncertainty sets of input parameters in return construction or use separate uncertainty sets for the distribution of mean and variance estimators, while our approach uses the uncertainty set of Sharpe ratio estimators directly. Our approach is motivated by the observation that Sharpe ratio estimators are approximately normally distributed *even when asset returns are not.* In addition, modeling the Sharpe ratio directly allows us to incorporate skewness and kurtosis information and avoids several key assumptions about the underlying return distribution.

More recently, Zymler et al. (2011) add a portfolio insurance guarantee to optimal portfolios using derivatives to the standard robust portfolio optimization framework as a hedge against catastrophic market events. Instead of using a worst-case scenario approach, DeMiguel and Nogales (2009) use robust estimators, M-estimator and S-estimator, and show their out-of-sample properties. For a recent survey of the contributions of the field of operations research to robust portfolio selection, see Fabozzi et al. (2010). Bertsimas et al. (2011) provide a broad overview of the robust optimization literature, while Ben-Tal and Nemirovski (2007) summarize the status of robust convex optimization in particular.

3 Statistical Properties of Sharpe Ratio Estimators

In this section, we lay out the groundwork for the construction of the VaRSR. We follow Lo (2002) and Mertens (2002) in modeling the distribution of the Sharpe ratio, which is necessary to derive the VaRSR. The Sharpe ratio (SR) of a return distribution is conventionally defined as the ratio of the expected excess return over the risk-free rate (μ) to the standard deviation of the excess returns (σ):

$$SR = \frac{\mu}{\sigma}. \tag{1}$$

Generally speaking, μ and σ are unobservable and have to be estimated from historical data. Given a sample of historical returns $\{R_1, R_2, \ldots, R_n\}$ and risk-free rates $\{R_{f1}, R_{f2}, \ldots, R_{fn}\}$, the estimated Sharpe ratio is

$$\widehat{SR} = \frac{\hat{\mu}}{\hat{\sigma}} \tag{2}$$

where

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}(R_i - R_{fi}) \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(R_i - R_{fi} - \hat{\mu})^2. \tag{3}$$

We begin with the derivation of the distribution of Sharpe ratio estimators assuming i.i.d. normal returns. Assuming that the investment returns $\{R_1, R_2, \ldots, R_n\}$ are i.i.d normal with finite mean μ and variance σ^2, Lo (2002) shows that the following relation holds as a result of the Central Limit Theorem:

$$\sqrt{n}\left(\begin{pmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right) \Rightarrow N\left(0, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}\right) \tag{4}$$

This implies that the variance in the estimators $\hat{\mu}$ and $\hat{\sigma}^2$ take the following asymptotic forms,

$$\mathrm{Var}\left(\hat{\mu}\right) = \frac{\sigma^2}{n} \quad \text{and} \quad \mathrm{Var}\left(\hat{\sigma}^2\right) = \frac{2\sigma^4}{n} \tag{5}$$

and therefore the sampling error of these estimators decreases with increasing sample size. Using Taylor's theorem, for a general function $g(\mu, \sigma^2)$,

$$\sqrt{n}\left(g(\hat{\mu}, \hat{\sigma}^2) - g(\mu, \sigma^2)\right) \Rightarrow N\left(0, \sigma^2\left(\frac{\partial g}{\partial \mu}\right)^2 + 2\sigma^4\left(\frac{\partial g}{\partial \sigma^2}\right)^2\right). \tag{6}$$

If $g(\mu, \sigma^2) = \frac{\mu}{\sigma}$, then

$$\sqrt{n}\left(\frac{\hat{\mu}}{\hat{\sigma}} - \frac{\mu}{\sigma}\right) \Rightarrow N\left(0, 1 + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2\right). \tag{7}$$

As a result, the standard deviation in the estimated Sharpe ratio is then given by

$$\sigma(\widehat{SR}) = \sqrt{\frac{1}{n}\left(1 + \frac{1}{2}SR^2\right)}. \tag{8}$$

After Bessel's correction the estimated standard deviation of $\widehat{SR}$ is given by

$$\hat{\sigma}(\widehat{SR}) = \sqrt{\frac{1}{n-1}\left(1 + \frac{1}{2}\widehat{SR}^2\right)}. \tag{9}$$

Both the unobserved population mean μ and population variance σ^2 contribute to the standard deviation of the Sharpe ratio estimator.

Mertens (2002) shows that the asymptotic distribution of Sharpe ratio estimators for returns drawn from a distribution with finite mean μ, variance σ^2, skewness γ_3, and kurtosis γ_4 is

$$\sqrt{n}\left(\frac{\mu}{\sigma} - \frac{\hat{\mu}}{\hat{\sigma}}\right) \Rightarrow N(0, V) \tag{10}$$

where

$$V = 1 + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 - \left(\frac{\mu}{\sigma}\right)\gamma_3 + \left(\frac{\mu}{\sigma}\right)^2\left(\frac{\gamma_4 - 3}{4}\right). \tag{11}$$

The standard deviation of the Sharpe ratio estimator is then estimated by[1]

$$\hat{\sigma}(\widehat{SR}) = \sqrt{\frac{1}{n-1}\left(1 + \frac{1}{2}\widehat{SR}^2 - \widehat{SR}\hat{\gamma}_3 + \widehat{SR}^2\left(\frac{\hat{\gamma}_4 - 3}{4}\right)\right)}. \tag{12}$$

Comparing Equation (12) to Equation (9) identifies the effect skewness and excess kurtosis have on the errors in the estimation of Sharpe ratios. Return distributions that exhibit negative skewness ($\hat{\gamma}_3 < 0$) and positive excess kurtosis ($\hat{\gamma}_4 > 3$) lead to greater uncertainty in the estimation of the Sharpe ratio than a normal return distribution with the same mean and variance. Given that some assets – such as hedge funds – often exhibit return distributions that are negatively skewed and leptokurtic, it is important to model the variance in the Sharpe ratio estimator explicitly.

4 VaR-Adjusted Sharpe Ratios

4.1 Definitions

We consider portfolios consisting of k securities, with each portfolio completely characterized by weights $\mathbf{w} \in [l_w, u_w]^k$, where w_i is the percentage of the total portfolio value invested in security $i \in \{1, 2, \ldots, k\}$. Each w_i has a lower bound l_w and an upper bound $u_w > l_w$ and the sum of w_i is 1.[2] Choosing the optimal portfolio using the Sharpe ratio as the objective is equivalent to solving

$$\max_{\mathbf{w} \in \mathbb{R}^k} \left\{ SR(\mathbf{w}) \,\middle|\, \mathbf{w}^T \mathbf{1} = 1, l_w \le w_i \le u_w \right\}. \tag{13}$$

[1] Although the underlying return distribution is not normal, the distribution of Sharpe ratio estimators follows an asymptotically normal distribution.

[2] If there exists short-selling constraints on the n securities, l_w is 0 and u_w is 1. Relaxing such constraints allows for $l_w < 0$ and $u_w > 1$.

In practice, the Sharpe ratio estimator, $\widehat{SR}(\mathbf{w})$, is used in place of the unobservable quantity $SR(\mathbf{w})$.

Since the Sharpe ratio estimator based on the estimated mean and variance in asset returns are subject to significant estimation error, the portfolio weights which maximize the Sharpe ratio estimator are unlikely to maximize the true Sharpe ratio. To mitigate such estimation errors, we introduce a risk-adjusted Sharpe ratio $\widehat{SR} - \gamma\hat{\sigma}(\widehat{SR})$ as the "VaR-adjusted Sharpe ratio" (VaRSR), denoted as $\widehat{SR}_{VaR}(\gamma)$. Here $\widehat{SR}$ is the Sharpe ratio estimator, $\hat{\sigma}(\widehat{SR})$ is its standard deviation, and parameter γ is determined by the confidence level of the VaR estimate. The VaRSR or $\widehat{SR}_{VaR}(\gamma)$, is used as the new objective function for the portfolio allocation problem.[3] Contrasting the traditional formulation in Equation (13), our main portfolio optimization problem becomes

$$\max_{\mathbf{w}\in\mathbb{R}^k} \left\{ \widehat{SR}_{VaR}(\gamma) \,\middle|\, \mathbf{w}^T\mathbf{1} = 1, l_w \leq w_i \leq u_w \right\}. \tag{14}$$

We now show that the formulation fits into a standard robust portfolio optimization framework, which is essentially a max-min problem:

$$\max_{\mathbf{w}\in\mathbb{R}^k} \left\{ \min_{SR\in\Theta_{SR(\mathbf{w})}} SR(\mathbf{w}) \,\middle|\, \mathbf{w}^T\mathbf{1} = 1, l_w \leq w_i \leq u_w \right\}. \tag{15}$$

$\Theta_{SR(\mathbf{w})}$ is an uncertainty set containing the unknown true Sharpe ratio $SR(\mathbf{w})$. The inner-minimization problem $\min_{SR\in\Theta_{SR(\mathbf{w})}} SR(\mathbf{w})$ computes the minimum possible value of $SR(\mathbf{w})$ for each given $\mathbf{w}$ in the uncertainty set $\Theta_{SR(\mathbf{w})}$, and identifies the portfolio with the largest worst-case Sharpe ratio.

To specify the uncertainty set $\Theta_{SR(\mathbf{w})}$ that establishes the equivalence of Equation (14) and Equation (15), recall from Section 3 that SR is a normal random variable with a distribution $N\left(\widehat{SR}, \hat{\sigma}^2(\widehat{SR})\right)$ in the limit of large sample sizes and under quite general assumptions – namely stationarity and ergodicity.[4] Following Zymler et al. (2011), the set Θ_{SR} could be an ellipse with exogenous parameter γ

$$\Theta_{SR} = \{(SR - \widehat{SR})(\hat{\sigma}^2(\widehat{SR}))^{-1}(SR - \widehat{SR}) \leq \gamma^2\}. \tag{16}$$

[3] VaRSR serves as a risk-adjusted Sharpe ratio. The level of VaRSR is strictly less than the Sharpe ratio. Occationally, we may observe that VaRSR is negative while the Sharpe ratio is positive, especially when the risk-adjustment component is large.

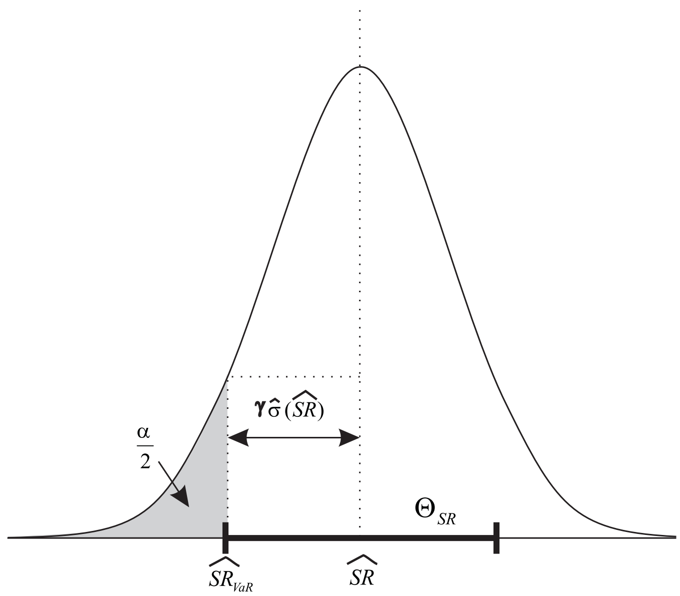
[4] The accuracy of the Sharpe ratio standard deviation estimator increases as the sample size increases. Empirical studies show that the normality result of the Central Limit Theorem is generally a good approximation for sample sizes greater than thirty (Hogg and Tanis, 2009). For a more in-depth discussion concerning the convergence in the limit of large sample sizes, see Greene (2002).

In a one-dimensional setting, the uncertainty set is the interval

$$\Theta_{SR} = \left\{ \left| \frac{SR - \widehat{SR}}{\hat{\sigma}(\widehat{SR})} \right| \leq \gamma \right\}. \tag{17}$$

Therefore, the solution to the inner-minimization problem is exactly $\widehat{SR}_{VaR}(\gamma) = \widehat{SR} - \gamma\hat{\sigma}(\widehat{SR})$.

Figure 1: The Sharpe Ratio Estimator $\widehat{SR}$, the VaR-Adjusted Sharpe Ratio VaRSR ($\widehat{SR}_{VaR}$), and the Uncertainty Set Θ_{SR}.



An illustration of the risk-adjusted Sharpe ratio measure $\widehat{SR}_{VaR}$ and the uncertainty set Θ_{SR} is depicted in Figure 1. SR is a normally distributed random variable with a distribution $N\left(\widehat{SR}, \hat{\sigma}^2(\widehat{SR})\right)$. For a given probability α (such as 5%), $\widehat{SR}_{VaR}(\gamma) = \widehat{SR} - \gamma\hat{\sigma}(\widehat{SR})$ defines a lower threshold value such that the likelihood that SR falls below this threshold value is less than or equal to $\frac{\alpha}{2}$

$$Prob\left(SR \leq \widehat{SR}_{VaR}(\gamma)\right) = \frac{\alpha}{2}. \tag{18}$$

The range $\left[\widehat{SR} - \gamma\hat{\sigma}(\widehat{SR}), \widehat{SR} + \gamma\hat{\sigma}(\widehat{SR})\right]$ defines a $(1-\alpha)$ confidence interval for the unobservable quantity SR. Here the parameter γ has a one-to-one correspondence to a

Z value $Z_{\frac{\alpha}{2}}$, given by

$$1-\alpha = P\left(SR \in \Theta_{SR}\right) = P\left(\left|\frac{SR-\widehat{SR}}{\hat{\sigma}(\widehat{SR})}\right| \leq \gamma\right) = P(|Z| \leq \gamma), \tag{19}$$

where Z is a standard normal random variable.

The parameter γ controls the size of the uncertainty set Θ_{SR}. Tradeoffs exist in choosing the appropriate value of γ. On the one hand, a large γ penalizes the estimated Sharpe ratio for remote events, providing a more conservative estimate. On the other hand, if γ is too large, the resulting portfolio might be too conservative.

For a given γ, $\widehat{SR}_{VaR}$ increases with $\widehat{SR}$, the number of sample points n, the skewness estimator $\hat{\gamma}_3$, and decreases with the kurtosis estimator $\hat{\gamma}_4$. Later we will show that this estimator $\widehat{SR}_{VaR}$ is more robust than the simple Sharpe ratio estimator $\widehat{SR}$ and that when the return distribution is non-normal the optimal portfolio based on Value-at-Risk Sharpe ratio and the simple Sharpe ratio can be quite different.

4.2 Comparison to Other Measures

Here we explore the differences as well as the connections between the VaRSR ($\widehat{SR}_{VaR}$) and the "probabilistic Sharpe ratio" (PSR) suggested in Bailey and López de Prado (2011). Although the two approaches are closely related, the two measures differ in how they incorporate uncertainty in Sharpe ratio estimation.

Bailey and López de Prado (2011) define the PSR as the probability that the estimated Sharpe ratio exceeds a benchmark Sharpe ratio (SR^*)

$$f(\widehat{SR}(\mathbf{w}, SR^*)) := \widehat{PSR}(SR^*) = \text{Prob}(\widehat{SR} \geq SR^*) = 1 - \int_{-\infty}^{SR^*} \text{pdf}(\widehat{SR}) d\widehat{SR}. \tag{20}$$

Applying the result that $\widehat{SR}$ is normally distributed, we have

$$\widehat{PSR}(SR^*) = \Phi\left[\frac{(\widehat{SR}-SR^*)\sqrt{n-1}}{\hat{\sigma}(\widehat{SR})}\right] \tag{21}$$

where Φ is the cumulative distribution function (CDF) for the standard normal distribution. A $\widehat{PSR}(SR^*) \geq 95\%$ indicates that the estimated Sharpe ratio is greater than the benchmark Sharpe ratio at a 95% confidence level.

Our approach contrasts with the PSR approach in a number of ways. The VaRSR is motivated by robust portfolio selection with max-min optimization and the PSR is not. The VaRSR ($\widehat{SR}_{VaR}$) computes an adjusted Sharpe ratio based on a prescribed threshold in probability, while the $\widehat{PSR}$ computes probability based on a prescribed threshold in Sharpe ratio. Although the two concepts are closely related, our approach is perhaps more intuitive than that of the PSR.

Figure 2: The Probabilistic Sharpe Ratio $\widehat{PSR}(SR^*)$.

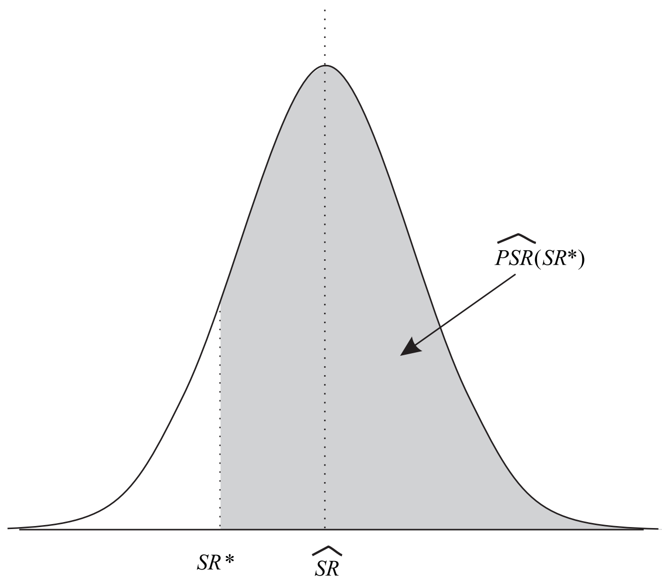


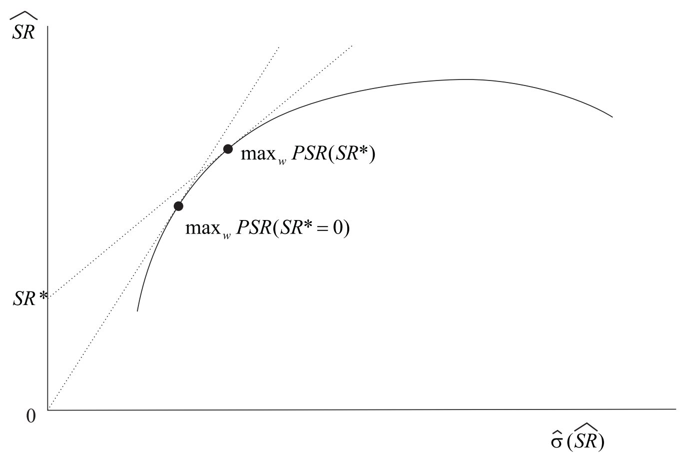
Figure 2 graphically depicts the probabilistic Sharpe ratio $\widehat{PSR}(SR^*)$. The PSR approach selects portfolio weights such that the optimal portfolio's Sharpe ratio distribution has the greatest probability mass in excess of a threshold Sharpe ratio. As illustrated in Figure 1, the traditional optimization framework selects the portfolio weights that maximizes the estimated Sharpe ratio, $\widehat{SR}$ while the VaRSR approach selects portfolio weights that maximize the lower bound $\widehat{SR}_{VaR}$ given a confidence level based on the resulting portfolio's estimated Sharpe ratio $\widehat{SR}$.

To further illustrate the relationship between the PSR and VaRSR, it is helpful to define the "Sharpe ratio efficient frontier" (SEF) first proposed by Bailey and López de Prado (2011). In Markowitz (1952), the mean-variance efficient frontier is defined as the set of portfolios which have the largest expected excess return for a given variance of excess returns. In a similar fashion, Bailey and López de Prado (2011) define the Sharpe ratio efficient frontier as the set of portfolios that deliver the greatest Sharpe ratio for a given level of estimation uncertainty. For a given level of uncertainty σ^*, the Sharpe ratio efficient frontier is given by

$$\mathrm{SEF}(\sigma^*) = \max_{\hat{\sigma}\left(\widehat{SR}(\mathbf{w})\right)=\sigma^*} \widehat{SR}(\mathbf{w}). \tag{22}$$

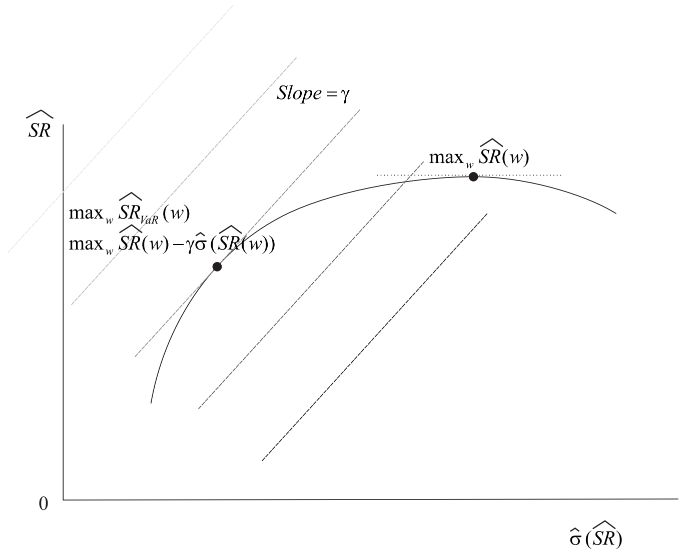
The Sharpe ratio efficient frontier is shown in Figure 3. Each possible portfolio is a data-point on or below the efficient frontier curve. Since the CDF function in Equation (21) is monotonically increasing, maximizing $\widehat{PSR}$ is equivalent to maximizing the argument of the CDF function. In Figure 3, the solutions to the PSR problem are located on the Sharpe ratio efficient frontier where the tangent line intersects the y-axis at SR^*.

Figure 3: PSR Solutions on the Sharpe Ratio Efficient Frontier.



Maximizing the VaRSR for a given probability α is equivalent to finding the tangent line to the Sharpe ratio efficient frontier that has slope γ. See Figure 4 for an illustration. Since the maximum of the Sharpe ratio efficient frontier occurs when the tangent line is zero, the maximum VaRSR is achieved with the optimal portfolio when γ is zero. In that case the VaRSR is the same as the maximum $\widehat{SR}$ which ignores the uncertainty in the measurement of the Sharpe ratio.

Figure 4: $\widehat{SR}_{VaR}$ Solution on the Sharpe Ratio Efficient Frontier.



The solution to the VaR optimization problem for a given γ is the same as the solution to the PSR optimization problem with $SR^* = \widehat{SR}_{VaR} - \gamma\hat{\sigma}(\widehat{SR})$. As a result, there is a correspondence between the solutions to the VaR portfolio optimization advocated here and the PSR portfolio optimization introduced by Bailey and López de Prado (2011). Our approach would suggest choosing a higher γ to more strongly penalize measurement uncertainties. Perhaps counterintuitively, this correspondence suggests that a *lower* benchmark SR^* should be chosen in this case.

5 Numerical Results

5.1 Simulation Results

We first test our model with simulated stock prices. Consider a simple portfolio with three uncorrelated assets and a constant risk-free rate $R_f = 0.01$. We assume that the underlying excess return distribution for each asset is described by the first four central moments given in Table 1.

Table 1: First Four Central Moments of the Simulated Excess Return Distribution for Three Assets.

	Mean (μ)	Volatility (σ)	Skewness (γ_3)	Kurtosis (γ_4)
Asset 1	0.10	0.20	1.00	3.00
Asset 2	0.15	0.30	0.00	7.00
Asset 3	0.20	0.36	-2.5	10.0

For each security, five years of monthly returns are simulated, resulting in a total of 60 data points. The simulation of stock returns is accomplished in MATLAB with the Pearson system distribution.[5] The Pearson distribution facilitates a simple simulation of asset returns when the underlying return distribution exhibits non-zero skewness or excess kurtosis. This simulated data resembles the data typically used by a hedge fund or portfolio manager to compute a five-year Sharpe ratio.

We computed the optimal allocation of portfolio value to the three assets in the traditional approach, the VaRSR approach and the PSR approach ($SR^* = 0$). The optimization model based on the VaRSR is implemented in MATLAB. We use the optimization routine `fmincon` to solve the problem without specifying the gradient or the Hessian matrix for the objective function.[6] We choose our threshold parameter γ to be 1.96, which corresponds to the significance level of 5%. The resulting portfolio weights are summarized in Table 2.

Table 2: Optimal Portfolio Allocations for Each Optimization Approach.

	$\mathbf{w}^*_{SR}$	$\mathbf{w}^*_{PSR}$	$\mathbf{w}^*_{VaR}$
Asset 1	43.32%	64.74%	52.69%
Asset 2	13.35%	8.56%	13.24
Asset 3	43.33%	26.70%	34.08%

The traditional portfolio allocation resulting from maximizing the Sharpe ratio yields portfolio weights given by $\mathbf{w}^*_{SR} = [43.32\%, 13.35\%, 43.33\%]$ and has the highest possible Sharpe ratio 0.6731, but also a high standard deviation – 0.1920. The Sharpe ratio for the

[5] For more details please refer to the MATLAB function `pearsrnd`.

[6] We do not explicitly convert our model into a second-order cone program as many papers in robust optimization do because our inner minimization problem is relatively simple to solve even without this conversion.

weight vector that maximizes the VaRSR is $\mathbf{w}^*_{VaR} = [52.69\%, 13.24\%, 34.08\%]$ is 0.6535 and has a standard deviation of 0.1441.

Table 3 summarizes the optimization results.

Table 3: Estimated Sharpe Ratio, Standard Error of Estimated Sharpe Ratio and Worst Case Sharpe Ratio for Optimal Portfolio Allocations in Each of the Three Optimization Approaches.

Statistic	$\mathbf{w} = \mathbf{w}^*_{SR}$	$\mathbf{w} = \mathbf{w}^*_{PSR}$	$\mathbf{w} = \mathbf{w}^*_{VaR}$
$\hat{\mu}(\mathbf{w})$	0.0281	0.0246	0.0265
$\hat{\sigma}(\mathbf{w})$	0.0418	0.0419	0.0406
$\widehat{SR}(\mathbf{w})$	0.6731	0.5875	0.6535
$\hat{\sigma}(\widehat{SR})(\mathbf{w})$	0.1920	0.1150	0.1441
$\widehat{SR}(\mathbf{w}) - \gamma\hat{\sigma}(\widehat{SR})(\mathbf{w})$	0.2968	0.3621	0.3711

In comparison to the traditional approach, VaRSR reduces the Sharpe ratio standard deviation by 25% while reducing the measured Sharpe ratio by only 3%. The 2-standard deviation lower bound on the traditional Sharpe ratio is 0.2968 and on the VaRSR is 0.3711 – an increase of more than 25%.

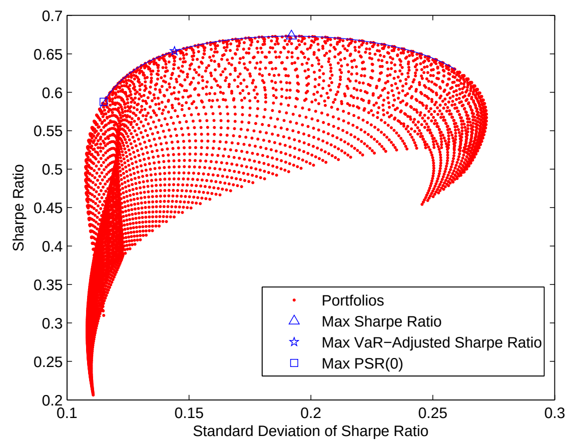
Since the return distribution of Asset 3 is negatively skewed, the VaRSR optimization algorithm penalizes weight on that asset. Compared to the traditional portfolio weights $\mathbf{w}^*_{SR}$, the maximum VaRSR portfolio shifts weight from Asset 3 to Asset 1. The excess kurtosis alters the portfolio optimization algorithm relative to the traditional approach at a higher order due to the excess kurtosis' quadratic coefficient in Equation (12). The weight on Asset 2 remains virtually unchanged between the conventional Sharpe ratio optimization and the VaRSR optimization in this example.

As the number of securities in the portfolio increases, the likelihood that the traditional and VaRSR portfolios will have similar weight vectors decreases. Among the optimization portfolios considered, the PSR approach with the threshold value $SR^* = 0$ yields the lowest Sharpe ratio as well as the lowest Sharpe ratio standard error, resulting from a greater penalty on the uncertainty surrounding $\widehat{SR}$.

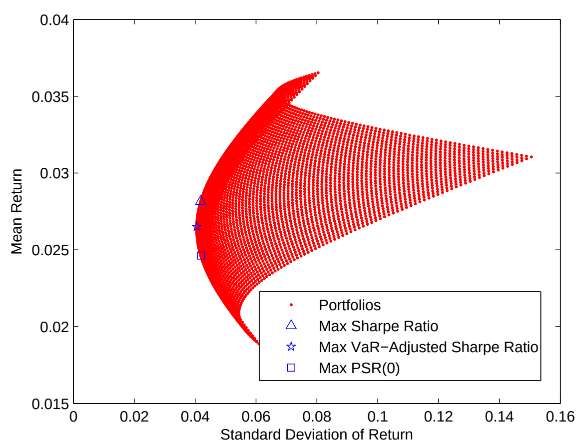
In Figure 5(a) we plot the Sharpe ratio efficient frontier. The red points represent portfolios with weights $w_i = j/100$ for $j \in \{0, 1, \ldots, 100\}$. The possible 5151 portfolios span the space of feasible weights defined by $l_w = 0 \leq w_i \leq u_w = 1$ respecting the portfolio constraint $\mathbf{w}^T\mathbf{1} = 1$. All possible portfolios reside on or below the Sharpe ratio efficient frontier.

The portfolio with the highest Sharpe ratio is denoted with a triangle, the portfolio calculated by maximizing the VaRSR is denoted with a star and the portfolio following the PSR approach is denoted with a square. The standard deviation of the highest Sharpe ratio portfolio is too large for the portfolio to be considered optimal in our model. We plot the Sharpe ratio efficient frontier in a light grey curve. The $\gamma \to 0$ limit of model coincides with the traditional Sharpe ratio optimization model.[7]

Figure 5: Portfolios Formed from Three Independent Assets with Mean Excess Returns of $[0.10, 0.15, 0.20]$ and Volatilities $[0.2, 0.3, 0.36]$.



(a) Sharpe Ratio Efficient Frontier



(b) Mean-Variance Efficient Frontier

[7] For a given value of γ, the optimal portfolio is point on the Sharpe ratio efficient frontier with derivative equal to γ. As a result, varying the parameter γ and determining the optimal portfolio will provide the set of portfolios that comprise the Sharpe ratio efficient frontier.

Figure 5(b) depicts the portfolios on a mean-variance graph (based on excess returns) as well as the efficient frontier of returns. Unlike the portfolio $\mathbf{w}^*_{SR}$, the portfolios $\mathbf{w}^*_{VaR}$ and $\mathbf{w}^*_{PSR}$ are not located on the mean-variance efficient frontier. This is not surprising since the mean-variance frontier only incorporates the information for the first two moments of the observed return distribution (mean and variance), while the portfolios $\mathbf{w}^*_{VaR}$ and $\mathbf{w}^*_{PSR}$ are optimally derived with higher moment information.

Robustness Test As we increase the significance parameter γ, the confidence interval widens. Following Goldfarb and Iyengar (2003), in Figure 6 we plot the Sharpe ratio with weights $\mathbf{w}^*_{SR}$ and the worst-case Sharpe ratio with weights $\mathbf{w}^*_{VaR}$ as a function of the parameter γ. The blue lines are for the Sharpe ratios and the red lines are for the worst-case Sharpe ratios.

Figure 6: Change of Sharpe Ratios with Varying γ.

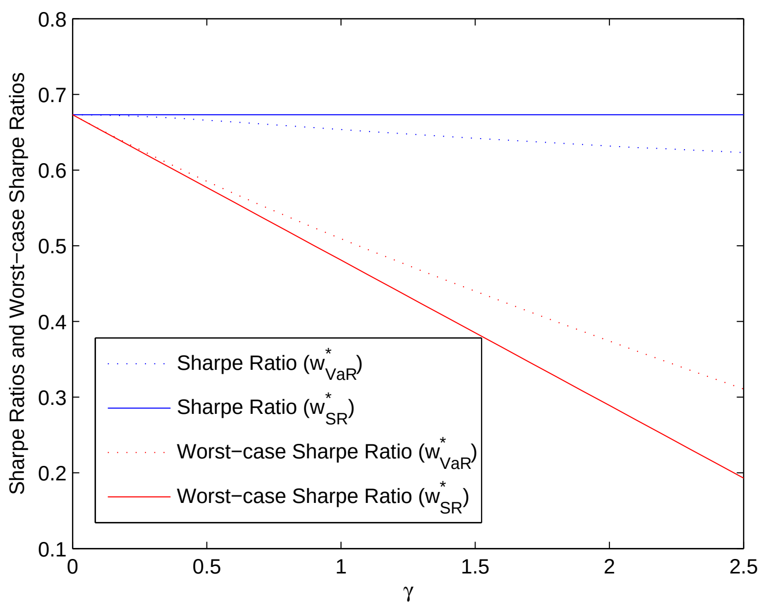


Figure 6 shows that as γ increases, the portfolio becomes more conservative and as a result exhibits a lower Sharpe ratio. Figure 6 also shows the dramatic alteration of the worst-case Sharpe ratio. Our approach drastically increases the worst-case Sharpe ratio but only slightly decreases the mean Sharpe ratio from that of the traditional portfolio optimization approach.

5.2 Empirical Results

As an empirical example of our framework, we determine the optimal allocation for portfolios constructed from 10 Dow Jones Credit Suisse Hedge Fund Indexes from January 1996 to December 2011.[8] We use hedge fund indexes because these funds typically exhibit negative skewness and positive excess kurtosis. As a result, the probability that the estimated Sharpe ratio for such investments accurately reflects the true underlying Sharpe ratio is smaller than that of an analogous asset with normally distributed returns and identical mean and variance (Bailey and López de Prado, 2011). Table 4 summarizes the hedge fund strategies and first four central moments of their historical return distributions.[9]

Table 4: Dow Jones Credit Suisse Hedge Fund Indexes: Moments of the Return Distribution Reflect Historical Monthly Excess Returns Observed from January 1996 to December 2011.

#	Strategies	Mean (μ)	Standard Deviation (σ)	Skewness (γ_3)	Kurtosis (γ_4)	$\widehat{SR}$	$\hat{\sigma}(\widehat{SR})$
1	Convertible Arbitrage	0.38%	2.09%	-2.66	18.39	0.183	0.092
2	Dedicated Short Bias	-0.51%	5.02%	0.67	4.30	-0.102	0.075
3	Emerging Markets	0.50%	4.06%	-1.31	9.70	0.123	0.079
4	Equity Market Neutral	0.20%	3.15%	-11.34	148.57	0.065	0.099
5	Event Driven	0.47%	1.89%	-2.31	13.78	0.249	0.096
6	Fixed Income Arbitrage	0.15%	1.76%	-4.16	30.04	0.085	0.086
7	Global Macro	0.72%	2.75%	-0.31	7.24	0.263	0.079
8	Long/Short Equity	0.52%	2.94%	-0.10	6.13	0.178	0.074
9	Managed Futures	0.30%	3.37%	0.074	2.62	0.089	0.072
10	Multi-Strategy	0.39%	1.50%	-1.90	10.64	0.257	0.093

In Figure 7, we show the value of a January 1996 initial investment of $100 in each of the Dow Jones Credit Suisse Hedge Fund Indexes over time. While some of the indexes were significantly affected by this financial crisis in late 2008 (e.g. Equity Market Neutral), others emerged from the crisis relatively unscathed (e.g. Managed Futures).

[8] For details, please see http://www.hedgeindex.com/. As a proxy for the risk-free rate of interest, we use one month LIBOR rates.

[9] In Section 5.1, we assumed the assets are uncorrelated. In this section, we implicitly use the historically accurate correlations between the various hedge fund indexes.

Figure 7: Value of a Hypothetical January 1996 Initial Investment of $100 in Each of the Dow Jones Credit Suisse Hedge Fund Indexes

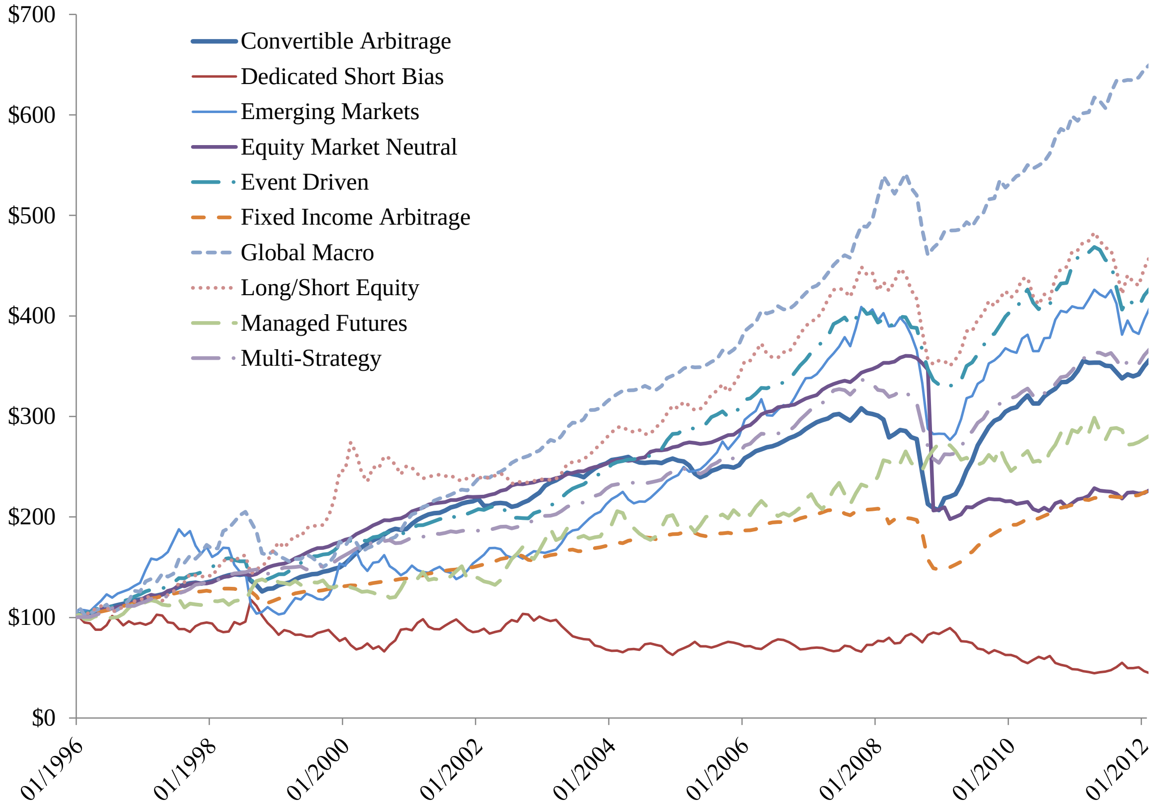


Table 5 presents the optimal weights obtained by maximizing the VaRSR and the traditional Sharpe ratio. Weights on indexes not listed in Table 5 are zero for the optimization approaches considered therein.

Table 5: Optimal Portfolio Allocation Given Historical Monthly Returns of the Dow Jones Credit Suisse Hedge Fund Indexes. $\gamma = 0$ Corresponds to the Traditional Portfolio Optimization.

Index	$\mathbf{w}^*_{VaR}$ given γ (Probability $\alpha/2$)						
#	0	1.282 (90%)	1.645 (95%)	1.960 (97.5%)	2.326 (99%)	3.090 (99.9%)	$\mathbf{w}^*_{PSR}$
2	2.7%	0	0	0	0	0	0
4	0	0	0	0	0	0.8%	1.6%
5	26.7%	28.3%	29.7%	30.3%	30.7%	31.4%	31.7%
7	26.6%	29.8%	30.3%	30.7%	30.9%	31.8%	32.8%
9	3.9%	9.3%	10.7%	11.9%	13.3%	15.6%	17.6%
10	40.3%	32.6%	29.4%	27.1%	25.1%	20.5%	16.2%

At first our results may seem counterintuitive since the VaRSR portfolios do not increase the weight on the index with the highest skewness and a low kurtosis - the Dedicated Short Bias Index. On the other hand, the measured Sharpe ratio is negative and as a result the positive skewness *increases* the standard deviation of the Sharpe ratio estimator – see Equation (12).[10]

The traditional portfolio optimization approach allocates over 40% of the portfolio to the Multi-Strategy index and essentially splits the remainder evenly between the Event Driven index and the Global Macro index. Although the more conservative (larger γ) optimal portfolios have larger allocations to a similar subset of indexes, the Event Driven and Global Macro indexes each have larger weight than the Multi-Strategy index. Again, the PSR approach with $SR^* = 0$ represents a more conservative approach with a higher penalty on the uncertainty of the estimated Sharpe ratio. Setting $SR^* = 0$ corresponds to our approach with $\gamma = 5.91$.

In Table 6, we summarize the Sharpe ratio, the Sharpe ratio standard error and the worst-case Sharpe ratio for the traditional Sharpe ratio optimization, the VaRSR optimization ($\gamma = 1.96$) and the PSR optimization ($SR^* = 0$). Although the mean Sharpe ratio is larger for the traditional approach, the worst-case Sharpe ratio is lower than both the PSR optimization and the VaRSR optimization with the latter being the highest.

[10] It is, in principle, possible that the optimization procedure advocated here would have non-zero weight on an asset with a negative mean excess return if such an allocation would have diversification benefits to sufficiently alter the return distribution to increase the VaRSR.

Table 6: Estimated Sharpe Ratio, Standard Error of Estimated Sharpe Ratio and Worst Case Sharpe Ratio for Optimal Portfolio Allocations for Portfolio Allocation to the Dow Jones Credit Suisse Hedge Fund Indexes.

Statistic	$\mathbf{w} = \mathbf{w}^*_{SR}$	$\mathbf{w} = \mathbf{w}^*_{PSR}$	$\mathbf{w} = \mathbf{w}^*_{VaR}$
$\hat{\mu}(\mathbf{w})$	0.0047	0.0051	0.0051
$\hat{\sigma}(\mathbf{w})$	0.0144	0.0165	0.0158
$\widehat{SR}(\mathbf{w})$	0.3286	0.3068	0.3205
$\hat{\sigma}(\widehat{SR})(\mathbf{w})$	0.0904	0.0763	0.0811
$\widehat{SR}(\mathbf{w}) - \gamma\hat{\sigma}(\widehat{SR})(\mathbf{w})$	0.1514	0.1572	0.1615

6 Conclusion

In this paper, we proposed a robust alternative to the traditional portfolio optimization problem using the concept of Value-at-Risk (VaR). Our approach is motivated by the observation that even if asset returns exhibit higher moments which are inconsistent with the normal distribution, the distribution of Sharpe ratio estimators follows is normally distributed. We call this new measure “VaR-adjusted Sharpe ratio” (VaRSR). The approach advocated here is a natural generalization to the standard portfolio optimization and intuitively connects to other alternatives proposed in the literature.

An ancillary benefit of the approach taken here is that it incorporates the higher order central moments of a portfolio’s excess return distributions. Although the standard portfolio optimization approach would allocate equal portion of a portfolio to two uncorrelated assets with the same mean, standard deviation, kurtosis but opposite skewness, the optimal portfolio based on the VaRSR has a larger investment in the asset with positively skewed excess returns.

We showed that this alternative measure limits the probability that the underlying Sharpe ratio estimated using the historical returns is substantially smaller than the computed Sharpe ratio. Furthermore, solutions to the traditional Sharpe ratio optimization model, our VaRSR model and the probabilistic Sharpe ratio model are all located on the Sharpe ratio efficient frontier introduced by Bailey and López de Prado (2011). The Sharpe ratio efficient frontier exhibits a second level of optimality beyond the mean-variance efficient frontier, which only uses information in the first two moments of returns. While the optimal portfolio in our framework is slightly shifted away from the mean-variance efficient frontier, the portfolio is enhanced by greater robustness.

Using numerical examples, we showed the superiority of our approach over both the traditional portfolio optimization as well as the probabilistic Sharpe ratio. We presented evidence that our approach is effective in mitigating realized volatility without sacrificing realized returns.

References

Bailey, D. and López de Prado, M. (2012) The Sharpe ratio efficient frontier. *Journal of Risk* 15(2):3–44.

Ben-Tal, A. and Nemirovski, A. (2007) Selected topics in robust convex optimization. *Mathematical Programming Series B* 112(1):125–158.

Bertsimas, D., Brown, D. B., and Caramanis, C. (2011) Theory and applications of robust optimization. *SIAM Review* 53(3):464–501.

Bodie, A., Kane, A., and Marcus, A. (2010) *Investments*. McGraw-Hill/Irwin, 9th edition.

Christie, S. (2005) Is the Sharpe ratio useful in asset allocation? MAFC Research Papers No. 31, Applied Finance Centre, Macquarie University.

DeMiguel, V. and Nogales, F. J. (2009) Portfolio selection with robust estimation. *Operations Research* 57(3):560–577.

Fabozzi, F. J. , Huang, D., and Zhou, G. (2010) Robust portfolios: contributions from operations research and finance. *Annals of Operations Research* 176(1):191–220.

Goldfarb, D. and Iyengar, G. (2003) Robust portfolio selection problems. *Mathematics of Operations Research* 28(1):1–38.

Greene, W. H. (2002) *Econometic Analysis*. Prentice Hall.

Hodges, S. (1998) A generalization of the sharpe ratio and its applications to valuations bounds and risk measures. Working paper, Finanical Options Research Centre, University of Warwick.

Hogg, R. and Tanis, E. (2009) *Probability and Statistical Inference*. Prentice Hall, 9th edition.

Lo, A. W. (2002) Statistics of Sharpe ratios. *Financial Analysts Journal* 58(4):36–52.

Markowitz, H. (1952) Portfolio selection. *Journal of Finance* 7(1):77–91.

Mertens, E. (2002) Comments on variance of the IID estimator in Lo (2002). Working paper.

Opdyke, J. (2007) Comparing Sharpe ratios: so where are the p-values? *Journal of Asset Management* 8(5):308–336.

Tütüncü, R. H. and Koenig, M. (2004) Robust asset allocation. *Annals of Operations Research* 132(1):157–187.

Zakamouline, V. and Koekebakker, S. (2009) Portfolio performance evaluatioin with generalized Sharpe ratios: Beyond the mean and variance. *Journal of Banking and Finance* 33(7):1242–1254.

Zymler, S., Rustem, B., and Kuhn, D. (2011) Robust portfolio optimization with derivative insurance guarantees. *European Journal of Operations Research* 210(2):410–424.