International Journal of Latest Trends in Engineering and Technology
Vol.(10)Issue(1), pp.009-018
DOI: http://dx.doi.org/10.21172/1.101.03
e-ISSN:2278-621X

INVESTIGATION OF ASSET ALLOCATION PERFORMANCE USING SHRINKAGE ESTIMATORS FOR HIGHER MOMENTS

Yesuk Jung[1], Ian Sutherland[2], Gunhee Lee[3]

Abstract: Post-Modern Portfolio Theory extends the notion that investors' satisfaction is reliant on minimizing downside deviation by incorporating higher order co-moments for optimizing the asset weights in a portfolio to accommodate for the difference in upside and downside deviation. Utilizing the third and fourth order co-moments of asset returns, co-skewness and co-kurtosis respectively, portfolio optimization algorithms are able to provide weights in respect to not only maximizing average return and minimizing risk in terms of covariance, but to maximize co-skewness in order to put higher weights on sets of assets whose returns tend to skew positively while minimizing co-kurtosis in order to curtail outliers. However, the use of higher order moments results in a problem of high dimensionality, where there are many parameters to estimate providing unstable estimators with large standard error. This study assesses the performance of several shrinkage estimators in order to overcome the dimensionality problems and to investigate the performances of several feasible optimization models using S&P 500 data. Our study shows that the optimization method with shrinkage estimators for higher order co-moments is superior to the traditional Markowitz's mean-variance optimization. It can also be concluded that applying shrinkage estimators and higher order co-moments to the problem of portfolio optimization

1. INTRODUCTION

Recently, with the growth of machine learning and artificial intelligence technologies, asset management firms have increasingly shifted focus toward systematic, computer-driven investment management strategies due to a growing accessibility of data, associated low management fees and high performance modeling. Systematic investment management can be administered to automate financial services and recommendations based on quantitative trading rules, often used in robo-advisors. The basic conceptual driver behind many of these systems is Markowitz's Modern Portfolio Theory (MPT). In 1952, Markowitz presented Modern Portfolio Theory, which seeks to find an optimum balance between the trade-off relationship of return and variance of asset returns [15]. However, there are many critics who oppose the underlying assumptions of MPT. Roy argues that investors are more sensitive to downside risk than to upside potential [23]. Moreover, according to behavioral finance, investors tend to give greater emphasis on losses relative to gains in their utility functions - a concept explained by Kahneman and Tversky's loss aversion preferences [12]. Another critique of MPT is that it does not distinguish between upside and downside risk, treating both upward and downward deviation as part of the same concept. To overcome this shortcoming of MPT, Post-Modern Portfolio Theory (PMPT) deals with investors' preferences more realistically by incorporating higher co-moments to minimize the duration and intensity of downside risk exposure. As such, Ang, Chen and Xing examined that there are correlation asymmetries between upside and downside movements using measures of asymmetry such as skewness and co-skewness [2]. Galagedera et al. corroborate the findings that downside risk is explained from the skewed distribution of returns [8]. To handle asymmetries of returns, higher order co-moments need to be considered for portfolio optimization [9-11].

A major problem implementing portfolio optimization using sample estimations for 2^{nd}, 3^{rd} and 4^{th} co-moments is that of estimation error. Due to the curse of high dimensionality, the estimation error problem becomes far worse for the 3^{rd} and 4^{th} co-moments. Ledoit and Wolf appropriately related the use of the sample covariance matrix to be like "error maximization" because of the high estimation error compounded by the tendency of optimizers to hone in on stocks with the highest estimation errors in practice [13]. They suggest a shrinkage estimator for the covariance matrix using a constant correlation utility function (CRRA) approach and they demonstrate how their shrinkage estimator reduces tracking error in real data. Martellini and Ziemann extended Ledoit and Wolf's methodology to the 3^{rd} and 4^{th} co-moments using the constant correlation utility function with a single statistical factor model approach [13, 14, 16]. They demonstrate that the portfolio optimization with improved estimators outperforms portfolios that use sample estimators for optimization. They also show that the single-factor approach out-performs the constant correlation without a statistical factor model in their sample data [16]. Boudt, Lu, and Peeters investigated a multifactor approach to the 3^{rd} and 4^{th} co-moments in portfolio optimization [4]. They revealed that portfolio optimization using higher co-moments with a multifactor approach outperformed their benchmark portfolio and

[1] PhD Candidate, Sogang University, Seoul, South Korea
[2] PhD Candidate, Sogang University, Seoul, South Korea
[3] *Professor, Sogang University, Seoul, South Korea*

made an important reduction in the downside risk of the portfolio. With regard to higher co-moments in portfolio optimization, problems of high dimensionality and high estimation error associated with it are negated using factor models with the constant correlation utility function. Most optimization techniques use the maximization of an expected utility function, and using higher co-moments in optimization is no different in that it commonly utilizes the constant correlation utility function. However, Chen K. accurately commented that the allocation is not sufficiently responsive to higher co-moments within CRRA since not all co-moments should inherently be equal in weight, and there is a need to calibrate the function to get optimal weightings for each co-moment [6].

In this study, we examine how the portfolio optimization with the improved shrinkage estimators for higher co-moments outperform compared to a benchmark portfolio (S&P 500 Index) by overcoming the shortcomings of previous methodologies. We apply Boudt et al.'s multifactor approach for shrinkage estimators on higher co-moments using the CRRA utility function to handle the estimation error and multidimensionality problems [4]. Additionally, we test an adaptation to the CRRA function with heavier weighting on co-skewness in order to determine whether the function can be better calibrated.

We investigate empirically to find optimal portfolios using daily adjusted-close prices of the S&P 500 constituents between 2007 and 2017. To compare the improvements, we use Quadratic Programming via R to solve for the quadratic optimization of incorporating four moments into asset allocation for optimal portfolios. Adding to the evidence of other studies, our results show drastically different weights for the same assets when using higher co-moments. We compare the results of several optimization methods using both lower and higher co-moments, with and without shrinkage estimators, and with and without higher weightings on the 3rd co-moment for applicable models.

2. LITERATURE REVIEW

2.1. Post-Modern Portfolio Theory:

Downside risk has long been a topic of interest, and the development of the Sortino Ratio by Frank Sortino opened the era of the Post-Modern Portfolio Theory (PMPT) [19]. Rom & Ferguson demonstrate that using downside deviation is an effective measure of risk, since it rewards upside potential and punishes downside variance, thus highlighting the following limitations of MPT - (1) the use of portfolio return variance is not the correct measure of risk, and (2) the assumption is false that asset returns follow a symmetric distribution [22]. This commenced the paradigm shift toward Post-Modern Portfolio Theory and determined Modern Portfolio Theory to be a special case of PMPT. Furthermore, Rom defined a Minimum Acceptable Return (MAR) as the target rate of return that meets the important financial objective of the investor. According to Rom, the downside risk is the deviation under MAR and should use the appropriate downside risk statistics of downside probability and average downside magnitude. Moreover, Chen proposes that PMPT is the new approach to portfolio theory that explains market abnormalities and investor behavior, and to take these into consideration, PMPT emphasizes the asymmetry of financial returns via the higher order co-moments of those returns [5].

Investors often have different financial goals - higher return on their principal, a specific benchmark to outperform, or their unique specified financial goal. PMPT allows investors to consider a variety of more specific objectives to meet, and reflects the characteristics of the practical financial world by incorporating higher co-moments of returns.

2.2. Shrinkage estimator on higher order co-moments:

A common problem in sample estimation comes from the estimation error when using sample estimates of the population mean, variance and so on. Best & Grauer investigated mean-variance portfolio weight allocation and concluded that it is likely to be sensitive to parametric changes in the mean and variance [3]. Moreover, Michaud pointed out that as the number of assets increase, so too does the error maximization problem [17]. There is a large body of work that attempts to manage these kinds of estimations errors, and many of the studies focus on trying to find an optimal balance between estimation errors and bias. One of the most common methods to handle this problem is to use shrinkage estimation methods. In a key paper, Ledoit and Wolf introduce a shrinkage estimator on the covariance matrix, which combines the sample covariance with a structured model-based estimator under the constant correlation approach [13]. The estimated shrinkage covariance is $\hat{\Sigma}_{shrinkage} = \delta F + (1-\delta)S$, which minimizes the expected distance between the shrinkage estimator and the true covariance matrix where S is the sample covariance matrix and F is structured estimator. δ is the shrinkage intensity, which is between 0 and 1. The estimated δ derived by Ledoit and Wolf is in the following form [14]:

$$\hat{\delta}^* = \max\left\{0, \min\left(\frac{\hat{k}}{T}, 1\right)\right\}, \quad \text{where } \hat{k} = \frac{\hat{\pi} - \hat{\rho}}{\hat{\gamma}} \tag{1}$$

In Equation 1, π is the asymptotic variance of the sample estimation, γ is the squared error of the structured estimator, and ρ is the asymptotic covariance between the sample and the structured estimator. The covariance parameters are a function of the constant correlation parameter, thus $\widehat{\rho_{ij}} = \hat{r}\sqrt{s_{ii}s_{ij}}$. Martellini and Ziemann extended Ledoit and Wolf's model to the higher order co-moments [13, 14, 16]. Martellini and Ziemann investigated asset weight optimization performance using two different shrinkage methods – (1) one method uses the same constant correlation approach as Ledoit and Wolf [14]; (2) the other method uses a single factor linear model [16]. They showed the improved estimator with the single factor approach outperforms the estimator using the constant correlation approach. Subsequently, Boudt et al. presented an alternative solution to the higher order co-moments using multifactor statistical models, where they effectively reduced the number of parameters to estimate in the 2nd, 3rd and 4th moments [4]. The multifactor statistical model is shown below:

$$r_t = a + Bf_t + e_t \tag{2}$$

where r_t is the asset returns, f_t is the factors in the factor model, B is the factor loadings, and e_t is the asset-specific factors (i.e. residual matrix).

Then, the 2nd, 3rd and 4th moments can be written as:

$$M_2 = \Sigma = BSB' + \Delta \qquad \text{where} \quad S = E[(f_t - \mu_f)(f_t - \mu_f)')] \tag{3}$$

$$M_3 = \Phi = BG(B' \otimes B') + \Omega \qquad \text{where} \quad G = E[(f_t - \mu_f)(f_t - \mu_f)' \otimes (f_t - \mu_f)')] \tag{4}$$

$$M_4 = \Psi = BP(B' \otimes B' \otimes B') + \Upsilon \qquad \text{where} \quad P = E[(f_t - \mu_f)(f_t - \mu_f)' \otimes (f_t - \mu_f)' \otimes (f_t - \mu_f)'] \tag{5}$$

Equation 2 defines the terms as components in a regression of the factor loadings on the asset returns. For the second moment estimation in Equation 3, M_2, is the estimated co-variance matrix where S is a K×K covariance matrix of K number of factors estimated using factor loadings from a statistical linear factor model, B is an N×K matrix of the factor exposures, and Δ is the residual matrix corresponding to the variance of error terms. In Equation 4, the third co-moment estimation, M_3, is the estimated K×K^2 co-skewness matrix where G is the K×K^2 co-skewness matrix of the K factors, Ω is the N×N^2 residual matrix corresponding to the expectation of the idiosyncratic factors' third moment. In Equation 5, the fourth co-moment estimation, M_4, is the estimated K×K^3 co-kurtosis matrix where P is the K×K^3 co-kurtosis matrix of K factors and Υ is the residual matrix based on the complex decomposition of the kurtosis of each asset, the kurtosis of two assets, and the kurtosis between three assets. It should be noted that Kronecker products, ⊗, are required due to the higher dimensionality of the vector spaces involved with higher co-moment calculations, which result in matrices of tensor products [4].

With the multi-factor approach, Boudt et al. consolidated that asset allocation with higher order co-moments can improve performance [4]. Following the same approach, we analyze the importance of higher co-moments on asset allocation and compare the impact that estimators have on portfolio weight allocations on assets.

2.3. Modification of CRRA utility function:

The maximization of an expected utility function is widely used as a criterion for asset allocation, especially when considering higher order co-moments to explain investors' risk preferences. Both Martellini and Ziemann and Boudt et al. use an approximation of the CRRA utility function to deal with higher co-moments in asset allocation [4, 16]. The expected utility function can be approximated using the Taylor series expansion:

$$E[U(X)] \approx U(\bar{X}) + \frac{1}{2}U^{(2)}(\bar{X})\sigma_p^2 + \frac{1}{3!}U^{(3)}(\bar{X})s_p^3 + \frac{1}{4!}U^{(4)}(\bar{X})k_p^4, \tag{6}$$

$$\begin{aligned} \textit{with} \quad & \mathrm{X} = \mu' w \\ & \sigma_p^2 = w' M_2 w, \\ & s_p^3 = w' M_3 (w \otimes w) \\ & k_p^4 = w' M_4 (w \otimes w \otimes w) \end{aligned}$$

In Equation 6, $\bar{X}$ is the expected portfolio return at the end-of-period, w is the weight vector, σ_p^2 is the variance, s_p^2 is the skewness, and k_p^2 is the kurtosis of the end-of-period portfolio return [10].

There are several objections to the use of the Taylor expansion for financial portfolio optimization. Critics argue that the Taylor series expansion may only converge to the expected utility under restrictive conditions [10, 11]. For example, power utility functions, such as CRRA, have a strict convergence region of the Taylor series expansion within 0< $\bar{X}$ <2E($\bar{X}$). Moreover, Chen suggests that a calibration of skewness and kurtosis preferences is needed since the allocation is not sufficiently responsive to higher co-moments in the CRRA utility function [5]. For these reasons, we think that the use of the Taylor expansion in CRRA is not optimal and needs to be modified. Our final model to test is the following expected utility function:

$$E[\ln(X)] = \ln(\bar{X}) - \frac{1}{2\bar{X}^2}\sigma_p^2 + \frac{10}{3\bar{X}^3}s_p^3 - \frac{1}{4\bar{X}^4}k_p^4, \quad (7)$$

with $X = \mu' w$

$\sigma_p^2 = w' M_2 w,$ where M_2 is in (3)

$s_p^3 = w' M_3 (w \otimes w),$ where M_3 is in (4)

$k_p^4 = w' M_4 (w \otimes w \otimes w),$ where M_4 is in (5)

Therefore, using Equation 7, we tested a modified Taylor expansion with a higher weighting on co-skewness in addition to the standard CRRA function, and strictly set the risk aversion parameter, lambda, to 1 as a logarithmic case for CRRA.

3. PERFORMANCE TEST

3.1 Research Design:

In this study, we use the S&P 500 constituents' adjusted-close prices to calculate daily returns from January 1, 2007 to December 30, 2016. Constituents with missing or partial data are removed for a total of 453 constituents used for analysis. To isolate the optimization effect, we randomly select baskets of 25 assets at a time in our constituent list for over 50 random Monte Carlo pulls, for a total of 1,250 randomized asset selections. In other words, since the focus of the study is on the optimization of asset weight allocations, 50 randomized selections of assets are used to avoid selection bias due to non-randomness. Each random basket of 25 assets was optimized over 8 different portfolio optimization techniques and the results were compared.

In addition to (1) the S&P 500 Index (BM), the portfolio optimization techniques utilized include - (2) Markowitz mean-variance (MV) portfolio [15]; (3) a portfolio optimized utilizing the constant correlation function for higher co-moments using the sample estimates for all co-moments (CRRA); (4) a portfolio optimized using the single-factor estimator for higher co-moments ($SFM_{k=1}$) [16]; (5) a portfolio optimized using the single-factor estimator for higher co-moments with a 10× weighting on co-skewness ($SFM_{Modified}$, k=1); (6; 7) two portfolios optimized using multi-factor models for two and three factors (SFM, k=2; SFM, k=3) [4]; and (8; 9) two portfolios optimized using multi-factor models for two and three factors with a 10× weighting on co-skewness ($SFM_{Modified}$, k=2; $SFM_{Modified}$, k=3). The optimization methods are summarized below in Table 1.

Table 1. Summary of 9 models

Abbreviation		Optimization Method	Moment Estimates
1.	BM	S&P 500 Index	NA
2.	MeanVar	Markowitz Mean-Variance optimization based on previous 750 market days (i.e. about 3 years) [15].	Sample Covariance
3.	CRRA	Sample estimates used for all comoments based on previous 750 market days (i.e. about 3 years). CRRA Utility Function used.	Sample Covariance; Sample Co-Skewness; Sample Co-Kurtosis
4.	SFM k=1	Martellini & Ziemann Single-Statistical Factor Model to estimate shrinkage matrices based on previous 750 market days (i.e. about 3 years) [16]. CRRA Utility Function used.	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis
5.	$SFM_{Modified}$ k=1	Martellini & Ziemann Single-Statistical Factor Model to estimate shrinkage matrices based on previous 750 market days (i.e. about 3 years) [16]. Modified CRRA Utility Function used (i.e. 10 times mulitplier on Co-Skewness weight).	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis
6.	SFM k=2	Boudt et al. Multiple Statistical Factor Model using 2 factors to estimate shrinkage matrices based on previous 750 market days (i.e. about 3 years) [4]. CRRA Utility Function used.	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis
7.	$SFM_{Modified}$ k=2	Boudt et al. Multiple Statistical Factor Model using 2 factors to estimate shrinkage matrices based on previous 750 market days (i.e.	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis

		about 3 years) [4]. Modified CRRA Utility Function used (i.e. 10 times mulitplier on Co-Skewness weight).	
8.	SFM k=3	Boudt et al. Multiple Statistical Factor Model using 3 factors to estimate shrinkage matrices based on previous 750 market days (i.e. about 3 years) [4]. CRRA Utility Function used.	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis
9.	$SFM_{Modified}$ k=3	Boudt et al. Multiple Statistical Factor Model using 3 factors to estimate shrinkage matrices based on previous 750 market days (i.e. about 3 years) [4]. Modified CRRA Utility Function used (i.e. 10 times mulitplier on Co-Skewness weight).	Shrinkage Covariance; Shrinkage Co-Skewness; Shrinkage Co-Kurtosis

The S&P 500 Index is used as a baseline portfolio measurement indicating the actual market movements and is capitalization-weighted and maintained by S&P Dow Jones Indices. The mean-variance portfolio is a standard portfolio optimization method using only the first and second order moments of mean and covariance. The higher order moments (CRRA) portfolio uses sample estimates for the covariance, coskewness, and cokurtosis matrices. The CRRA portfolio demonstrates the effectiveness of using higher co-moments without shrinkage estimators and thus, is expected to have low performance due to the noisiness of estimation. The estimated co-moments using statistical factor models (SFM k=1; SFM k=2; SFM k=3) are compared to the higher weight on the 3rd co-moment for modified statistical factor models ($SFM_{Modified}$ k=1; $SFM_{Modified}$ k=2; $SFM_{Modified}$ k=3). The factor models exhibit the effectiveness of shrinkage estimators on higher order co-moments, and the modified CRRA function tests whether a higher weighting on co-skewness can outperform the standard CRRA function. We contend that if the modified function outperforms the standard function, then the standard function is not optimal for portfolio optimization and requires more research in the future to determine an optimal CRRA function.

In the modified CRRA models, the weight multiple of ten on the third order co-moment is set somewhat arbitrarily with some previous knowledge of return distributions, however, it is not representative of an optimized weighting. While the CRRA function incorporates the maximization of investors' expected return and positive skewness with an aversion for variance and kurtosis of returns [24], the equal relative weightings of each moment in the standard model are also somewhat arbitrary in our assessment. We selected only the third order co-moment for the increase in weight because of investors' strong aversion to drawdowns and appeal of upside risk, which is the basis of Postmodern Portfolio Theory. We expect that positive skewness of returns is most likely to result in higher annualized returns with associated lower downside risk compared to other moments, since Chen K. showed that kurtosis has virtually no impact on CRRA [6]. Therefore, the modified portfolio is used to test whether the constant correlation function would be better calibrated with a higher weighting on the co-skewness of returns.

The two constraints used for all models in our study were full investment (i.e. sum of weights is 1), and long-only (i.e. non-negative asset weights). While only using these two constraints may be unrealistic in practice, it demonstrates a more pure effect from the optimization than using additional constraints, such as diversification constraints or box constraints for each asset weight, because additional constraints could interfere by suppressing the effect of optimization. The optimizing method used was evolutionary global optimization via the Differential Evolution algorithm, known as DEoptim [1, 21]. Differential Evolution was used due to its ability to handle non-continuous optimization, noisy data, and its quick adaptability to change, in addition to outperforming several other methods in a small exploratory test prior to this study.

We use a rolling window for annual rebalancing for portfolio weights. Balancing annually on the first market day of January for each year in the dataset, our model utilizes the previous 750 market days, representing roughly 3 years of historical market data, as a training period for optimization. The prior 750 market days for each rebalancing period are used to estimate covariance matrices, shrinkage matrices, sample higher moments, single-factor models and multi-factor models.

In the constant correlation utility function (CRRA), we suppose that the risk aversion parameter, lambda, is equal to 1, i.e. logarithmic utility, as a special case of CRRA. Since our primary purpose is to compare each optimization strategy, we assume that an investor regards that risk and wealth are independent, which lends itself as a simple way to handle risk aversion preference [7].

3.2 Portfolio Performances:

Portfolio optimization methods are compared over several metrics - annualized return, annualized deviation, average drawdown, downside deviation, skewness of returns, kurtosis of returns, and the Sortino Ratio. Since 50 batches of 25 asset baskets were used for optimization, the daily portfolio returns were first averaged over all 50 batches for each optimization method with the exception of the S&P 500 Index. The averaged portfolio returns were then used for the comparative performance metrics.

The annualized return is the percentage change on the portfolio principal per year. Annualized deviation is the deviation of the returns over the period for each portfolio divided by the number of years in the period. Average drawdowns (ADD) is a measure of the magnitude of drawdowns and is formulated ostensibly as the sum of the absolute value of drawdowns over the period divided by the number of drawdowns in the period, i.e. a simple average of the magnitude of drawdowns in the period. Thus, a lower ADD corresponds to less investor dissatisfaction. Downside deviation is a better measure for risk than annualized deviation in terms of Postmodern Portfolio Theory, since downside deviation eliminates positive returns in the calculation of deviation. Downside deviation uses Minimum Acceptable Return (MAR) as a measure from which to determine if returns are positive (i.e. above MAR) or negative (i.e. below MAR). Since MAR is set specifically to each investor's investment objectives, for simplicity, MAR is set to 0 in our metrics. Skewness of returns is the asymmetry measure of the portfolio return distribution. A skewness of 0 describes a symmetric distribution, where a higher skewness is indicative that returns are more skewed to the right, leading to higher investor satisfaction. Kurtosis measures the degree of flatness of the distribution of returns, so a high kurtosis is indicative of more extreme values, where the normal distribution is characterized by a kurtosis of 3. Therefore higher kurtosis is more platykurtic and thus more undesirable to investors.

The Sortino Ratio is chosen as the primary measure for comparing optimization techniques in this paper for several reasons. The Sortino Ratio is a risk-adjusted return metric that improves upon the Sharpe Ratio by incorporating MAR and downside deviation, thus better reflecting the foundations of PMPT. Sortino Ratio is calculated by subtracting the MAR from the portfolio's return and then dividing by the downside deviation. Hence, a high return with low downside deviation in relation to the MAR will result in a preferably higher Sortino Ratio. As such, the Sortino Ratio more accurately measures the risk of returns failing to meet the investor's desired investment goal [19]. For simplicity and consistency, MAR is set to 0.

Notable portfolio performance metrics are summarized below in Table 2. Mean-Variance outperforms the baseline model with higher returns and lower deviation of returns as expected. The higher moments model using sample estimates, CRRA, also performs poorly as expected relative to other models. Since the CRRA portfolio suffers from the problem of high dimensionality with many parameters to estimate, sample estimates compile high error due to the noisiness of the data, and the result is poor estimates particularly for the higher order co-moments. As a result of this problem, the CRRA portfolio underperformed the Mean-Variance portfolio in every measured metric except for average drawdown. Although the estimates for CRRA are poor, the use of higher order co-moments results in lower average drawdowns in every case of our measured portfolio optimization methods relative to the lower order Mean-Variance optimization.

The statistical factor models notably outperform all other models in terms of annualized returns, with relatively comparable results in other measures. The single-factor model performed best out of the notable standard models summarized in Table 2. There is a meaningful improvement in annualized return, roughly 1 full percentage point increase from the lower order model. While critics may argue that annualized standard deviation has also increased slightly compared to the lower models, the standard deviation measure encompasses both upward and downward deviation and is thereby not an applicable metric to compare the models by. A more apt metric to compare is the average downside deviation (ADD), wherein we see that the magnitude of drawdowns are less intense in the single statistical factor model as compared to the lower order models, albeit with slightly more deviation in those drawdowns. Based on the risk-adjusted return measure, the Sortino Ratio, the SFM k=1 model is the most preferable standard model tested.

The multi-factor models, SFM k=2 and SFM k=3, perform comparably to the single-factor model in that the multi-factor models see a similar improvement in annualized return and average drawdown while experiencing slight increases in the deviation of both the annualized returns and downside deviation when compared to the BM and MV portfolios. Despite the increase in deviations for all statistical factor models, the lower kurtosis is evidence that there are fewer extreme values with more peaked distributions of returns relative to the lower order models. Based on the Sortino Ratio, the appropriate single-measure to evaluate the models, the higher order models using shrinkage estimates on the co-moments are much more preferable than the baseline estimate, Mean-Variance model, and the higher order model without shrinkage estimates.

There are some notable findings on the outcomes of the modified CRRA function. The $SFM_{Modified}$ k=1 optimization outperformed the SFM k=1 model with annualized returns of 12.98% and a Sortino Ratio of 9.991 for the modified single-factor model, thereby outperforming every other model evaluated. Furthermore, $SFM_{Modified}$ k=2 with a Sortino Ratio of 9.754 beat out its base model, SFM k=2. For the more complex $SFM_{Modified}$ k=3 model, a Sortino Ratio of 9.64 was not sufficient to outperform the SFM k=3 base model.

Table 2. Annualized performance of 9 models

Portfolio Optimization Algorithm			Annualized Returns (%)	Annualized Std. Dev. (%)	Average Drawdown (%)	Downside Deviation (%)	Skewness of Returns	Kurtosis of Returns	Sortino Ratio
Method	Estimator	Abbreviation							
S&P 500	NA	1. BM	7.66	15.60	2.051	0.712	-0.525	4.398	4.794

Index									
Mean-Variance	Sample	2. MV	11.50	11.51	1.439	0.511	-0.445	3.536	8.971
Maximize Utility	Sample	3. CRRA	11.22	11.51	1.422	0.512	-0.446	3.471	8.761
	Single Statistical Factor (k=1)	4. SFM k=1	12.80	11.71	1.433	0.517	-0.413	3.258	9.777
		5. $SFM_{Modified}$ k=1	12.98	11.71	1.389	0.516	-0.407	3.166	9.910
	Multiple Statistical Factor (k=2)	6. SFM k=2	12.37	11.69	1.442	0.518	-0.431	3.303	9.469
		7. $SFM_{Modified}$ k=2	12.73	11.68	1.400	0.515	-0.415	3.278	9.754
	Multiple Statistical Factor (k=3)	8. SFM k=3	12.75	11.71	1.387	0.517	-0.421	3.321	9.733
		9. $SFM_{Modified}$ k=3	12.58	11.67	1.440	0.516	-0.415	3.183	9.640

3.3 Cumulative Portfolio Returns:

Figure 1 depicts the cumulative returns over time for the base optimization techniques along with the S&P 500 Index (BM). Analyzing the portfolio cumulative returns over time, it is apparent that the BM appears more erratic over time with more upswings and downswings. Portfolio weight optimizers smooth this noise out with a steadier cumulative return trajectory. For investors, this means that portfolio weight optimization results in more reliable returns over time, and thus smaller drawdowns in terms of both magnitude and duration in relation to the BM portfolio. It is also demonstrated that the statistical factor models outperform all lower order and sample estimate models in terms of cumulative return after about five years of investment at nearly every point in time thereafter.

The difference between cumulative returns over the various optimization techniques becomes more apparent over longer time horizons. In our study, a clearer divergence begins to appear in cumulative returns between three and four years. The outperformance of higher co-moment models using estimates via statistical factor models becomes clear around five years after investment. It should also be noted that the MV portfolio performs very similarly to the CRRA model. While the BM portfolio performs erratically over time, the performance of other optimization techniques remains consistent with one another in their performance rankings over time.

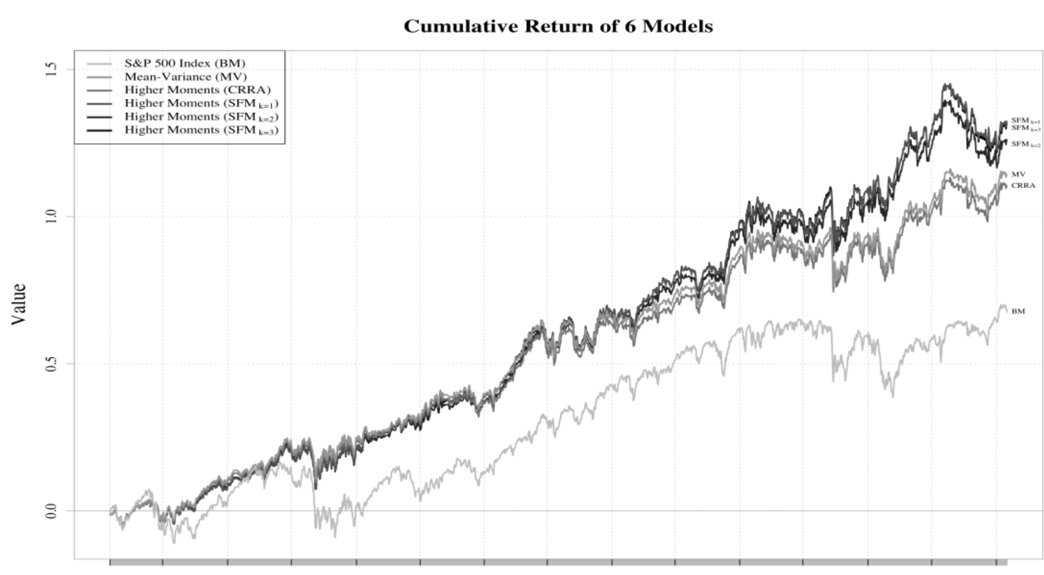

Fig. 1. Cumulative returns of 6 models since first market day of 2010

Figure 2 summarizes the model performances by displaying the cumulative return annually. It is more clearly shown that CRRA and MV perform similarly to one another, while SFM models outperform likewise with one another. All models outperform the benchmark. SFM k=1 model is shown because it was the highest performing model, however, it is edged out by $SFM_{Modified}$ k=1. This lends itself as strong evidence to the importance of the third co-moment and the suitability of Post modern Portfolio Theory for practical application. See Table 3 for detailed statistics on cumulative annual return. All cumulative return statistics are calculated using the first market day in 2010 as the origin date, and numbers represent the percentage increase in return at period end from the origin date.

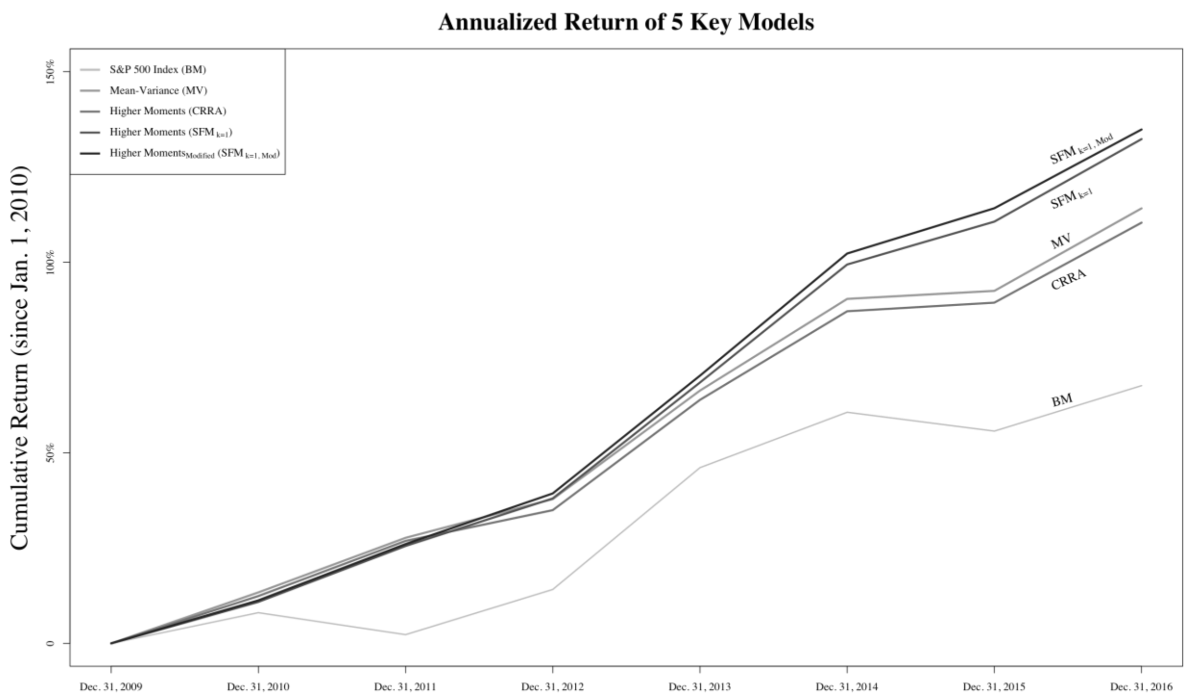

Fig. 2. Annualized cumulative returns of 5 models since first market day of 2010

Table 3. Annualized cumulative return statistics of 12 models

Optimization Model	Cumulative Return as of Period End						
	2010	2011	2012	2013	2014	2015	2016
1. BM	8.06	2.29	14.15	46.10	60.65	55.69	67.63
2. MeanVar	13.37	27.70	37.86	66.33	90.38	92.49	114.14
3. CRRA	12.41	26.88	34.96	63.89	87.14	89.40	110.39
4. SFM k=1	10.89	25.56	38.04	68.45	99.41	110.66	132.31
5. $SFM_{Modified}$ k=1	11.29	26.05	39.35	70.24	102.29	114.18	134.84
6. SFM k=2	10.68	24.41	36.07	66.88	96.80	106.45	126.11
7. $SFM_{Modified}$ k=2	11.35	25.60	37.77	68.89	99.66	110.16	131.27
8. SFM k=3	11.12	25.04	37.73	68.18	100.12	110.48	131.46
9. $SFM_{Modified}$ k=3	10.42	23.75	36.90	67.47	99.02	108.84	129.11

4. DISCUSSION

4.1 Conclusion:

In this study, we demonstrate that utilizing the higher order co-moments of asset returns outperforms against the S&P 500 Index (BM) and Mean-Variance portfolio (MV), but only when the higher order co-moments are estimated using shrinkage techniques. We go on to show that the often used constant correlation utility function (CRRA) is ill-equipped to optimize higher order co-moments and that a modified version with a ten times weighting of the 3^{rd} co-moment outperformed the base

CRRA optimization models, calling into question whether the CRRA has optimally tuned weights for each moment.
Higher order co-moment models are only beneficial when using shrinkage estimates for the co-moments due to the noisiness of the sample estimates. The single and multi-factor statistical factor models provided by Martellini and Ziemann and Boudt et al. are shown to be valid estimators for the higher co-moments [4, 16]. It is noted that a model using sample estimates of higher moments, the CRRA model, performs similarly to the Mean-Variance model as a result of the noisiness of moment estimations.
Furthermore, the modified higher order co-moment models, which have a ten times multiplier on co-skewness, consistently outperform the standard higher order co-moment models over time. When putting a higher weighting on the 3^{rd} order co-moment in the quadratic optimization, the skewness of portfolio returns is higher and kurtosis is lower resulting in better performance in annual returns with nearly identical average drawdowns as the non-weighted models, giving preferable risk-adjusted returns.
The expected improvements in the average drawdowns from the higher order co-moment optimizations were corroborated in our results. The improvements in annual returns, however, exceeded our expectations. We also show that investors should expect the benefit of using higher co-moment optimization in their cumulative returns to start after several years, about five years in our dataset.
More research should be done to find a better calibration for the CRRA utility function. Our study shows that a higher weighting on the third moment is beneficial, and we expect that a higher relative weighting on the mean and skewness with a lower relative weighting on the covariance and kurtosis might be ideal. However, without further study, there is no empirical support for this hypothesis yet.
We hope that the impact of our study will draw more interest into Post-Modern Portfolio Theory and prove to be a stepping stone for further research into the optimization of weights that, not only improve portfolio asset management practices, but are better for investors well-being.

4.2 Limitations and future study:

In addition to the nine optimization models presented above, we also tested an equally-weighted portfolio (EW) with a 1/N strategy, a Mean-Variance portfolio using Ledoit & Wolf's (2004) shrinkage estimator on the covariance ($MV_{Shrinkage}$), and a modified higher moments model using sample estimations for all moments, but with a ten times weight multiplier on the third comoment ($CRRA_{Modified}$). The S&P 500 Index has an extremely high correlation to the Equal Weight portfolio at 0.986. While there are significantly higher annualized returns of 11.81% in the EW portfolio as compared to the S&P 500 Index (BM), the deviation of those returns is also high at 16.83%. This may be in large part due to the survivorship bias, a common issue in similar such studies, and in some small part, this difference may be attributable to the stocks selected in the random batches. While critics may argue that the EW portfolio makes a strong baseline, the cumulative returns are far too erratic over time, despite the high annualized returns, and in terms of our single-measure comparison metric, all models outperformed the EW in terms of its Sortino Ratio of 6.563, except the BM portfolio.
Ledoit & Wolf (2004) demonstrated that the shrinkage estimator outperforms the sample estimation, and our results corroborated these findings in that $MV_{Shrinkage}$ outperformed the MV portfolio. To illustrate, the $MV_{Shrinkage}$ has a Sortino Ratio of 9.020 while the MV has a Sortino Ratio of 8.971 due to the fact that the $MV_{Shrinkage}$ edged out the MV portfolio with slightly higher annualized returns (roughly 0.02% improvement) while maintaining a slightly lower downside deviation (about 0.003% less). These results substantiate our claims that shrinkage estimators outperform sample estimators, even when dealing with lower order models.
The modified higher order moments ($CRRA_{Modified}$) portfolio uses the same sample estimates but weights the co-skewness by 10 times. The $CRRA_{Modified}$ portfolio outperformed the CRRA by slightly edging out the annualized returns while maintaining nearly identical average drawdowns. The outcome is a better Sortino Ratio of 9.5 for the $CRRA_{Modified}$ compared to just 8.76 for CRRA. The $CRRA_{Modified}$ model performs more similarly to the statistical factor models, but with more deviation over time due to the noisy estimations of higher order co-moments. The improvements when putting a higher weighting on the 3^{rd} co-moment is ample evidence that the constant correlation utility function is not optimal and is in need of better calibration for weighting of the co-moments portfolio optimization. It is demonstrated that the simpler models, such as the CRRA, SFM k=1 and SFM k=2 models benefit from a heavier weighting on co-skewness. While this does not hold true for the SFM k=3 model, we hypothesize that this may be the case because the complexity of the model already takes into account the 3^{rd} order co-moment better than the simpler models do, however more study is required to confirm this hypothesis.
Further insight is required into the affects of the number of assets being optimized. Our study uses 25 assets for 50 random draws of assets, however, a higher number of assets may result in more accurate estimates for shrinkage and factor models. A problem with using higher numbers of assets in each random draw is that it significantly heightens computation time. For financial organizations with world-class computing power, this may not be a problem, but for the average home-investor or organization with antiquated hardware, computation power is a limitation that directly affects the number of assets to be considered for higher order optimizations.

While there is no concern about differing transaction costs between models in our research design since all 25 assets are rebalanced annually, there is still a question of how rebalancing frequency might affect the performance of the models. While more frequent rebalancing may be ideal, the computation time and higher transaction costs associated with more frequent portfolio rebalancing make it a complex issue that demands future study. To increase this complexity, market-shifting events make an impact that can be amplified by using the prior three years as a training period, and should also be a consideration when choosing the timing of rebalancing.

We chose a training period of 750 market days prior to the rebalance date, equating to about three years. However, there is insufficient research into the performance of the length of training periods. A longer training period yields more robust estimates for shrinkage estimators, however, using further past data also biases those estimates, as the market changes over time. While a shorter training period may be ideal theoretically based on market, industry and company changes, computationally it is problematic due to the need for many observations to calculate shrinkage estimators, particularly for higher order estimators.

While the CRRA utility function is most commonly used in practice regarding higher co-moment models, our study does not evaluate the viability of other utility functions in the model. Handling downside risk more directly with Expected Shortfall or Value at Risk (VaR) in the utility function may lead to better results in terms of lowering the associated portfolio risk and requires more study in the future. Additionally, our use of the CRRA utility function in this study is constrained to the assumption that investors only have a logarithmic utility, since the risk preference parameter of lambda equals 1. There needs to be more research into risk-seeking and risk-averse investors' utility to understand how optimizations differ when lambda is set to different values.

Furthermore, we demonstrate that the CRRA utility function is not optimized for the higher order co-moment optimizations by showing that the weighting of the 3rd order co-moment outperforms the standard models. However, we do not suggest that this is the optimal weighting on the CRRA utility function. We only seek to demonstrate that the function is not optimal as it is. Further study is required to assess more optimal weights for the CRRA utility function. We expect that risk attitude is likely a key component to a better optimization of weightings for each moment in the CRRA utility function.

Disclosure: This research was supported by the Ministry of Science and ICT (MSIT) Korea, under the Information Technology Research Center (ITRC) support program (IITP-2017-2015-0-00369) supervised by the Institute for Information & Communications Technology Promotion (IITP).

REFERENCES

[1] Ardia, D., Boudt, K., Carl, P., Mullen, K. M., & Peterson, B. G. (2011). Differential Evolution with DEoptim. R Journal, 3(1).
[2] Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. The Review of Financial Studies, 19(4), 1191-1239.
[3] Best, M. J., & Grauer, R. R. (1991). Sensitivity analysis for mean-variance portfolio problems. Management Science, 37(8), 980-989.
[4] Boudt, K., Lu, W., & Peeters, B. (2015). Higher order comoments of multifactor models and asset allocation. Finance Research Letters, 13, 225-233.
[5] Chen, J. M. (2016). Postmodern Portfolio Theory. In Postmodern Portfolio Theory (pp. 27-38). Palgrave Macmillan US.
[6] Chen, K. (2003). Calibrating the Skewness and Kurtosis Preference of Investors. Switzerland: HEC Lausanne.
[7] Cotter, J., & Hanly, J. (2012). A utility based approach to energy hedging. Energy Economics, 34(3), 817-827.
[8] Galagedera, D. U. (2007). An alternative perspective on the relationship between downside beta and CAPM beta. Emerging Markets Review, 8(1), 4-19.
[9] Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. The Journal of Finance, 55(3), 1263-1295.
[10] Jondeau, E., & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. European Financial Management, 12(1), 29-55.cdr
[11] Jondeau, E., & Rockinger, M. (2012). On the importance of time variability in higher moments for asset allocation. Journal of Financial Econometrics, 10(1), 84-123.
[12] Kahneman, D. (1979). Prospect theory: An analysis of decisions under risk. Econometrica, 47, 278.
[13] Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of empirical finance, 10(5), 603-621.
[14] Ledoit, O., & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. The Review of Financial Studies, 30(12), 4349-4388.
[15] Markowitz, H. (1952). Portfolio selection. The journal of finance, 7(1), 77-91.
[16] Martellini, L., & Ziemann, V. (2009). Improved estimates of higher-order comoments and implications for portfolio selection. The Review of Financial Studies, 23(4), 1467-1502.
[17] Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized'optimal?. Financial Analysts Journal, 45(1), 31-42.
[18] Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. The Journal of Portfolio Management, 30(4), 110-119.
[19] Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. the Journal of Investing, 3(3), 59-64.
[20] Swisher, P., & Kasten, G. W. (2005). Post-modern portfolio theory. Journal of Financial Planning-Denver, 18(9), 74.
[21] Price, K., Storn, R. M., & Lampinen, J. A. (2006). Differential evolution: a practical approach to global optimization. Springer Science & Business Media.
[22] Rom, B. M., & Ferguson, K. W. (1994). Post-modern portfolio theory comes of age. The Journal of Investing, 3(3), 11-17.
[23] Roy, A. D. (1952). Safety first and the holding of assets. Econometrica: Journal of the Econometric Society, 431-449.
[24] Rubinstein, M. (2006). Multi-moment asset allocation and pricing models (Vol. 399). John Wiley & Sons.