Journal of Yerevan University. Economy
2024, Vol. 15, No. 2(44), December, 112-119
https://doi.org/10.46991/BYSU.G/2024.15.2.112

# COMPARISON OF MARKOWITZ AND HELLINGER-NORMAL PORTFOLIOS

**MESROP MESROPYAN**,
*American University of Armenia*
**VARDAN BARDAKHCHYAN***
*Yerevan State University*

**Abstract:** In this paper we compare the performances of Markowitz portfolio and the portfolio closest to normal in distribution. The latter is obtained by fixing the same desired level of expected returns and optimizing the Hellinger distance to Gaussian distribution with parameters obtained from Markowitz portfolio optimization for the same expected return. We confine ourselves to a long-position portfolio only. We found that in contrast to the expectations, the Hellinger-Normal portfolio does not smooth enough the extreme losses, but it does not do worse in that regard than Markowitz portfolio. We also found that overall in non-long-run passively managed portfolios, the Hellinger-Normal portfolio had better overall realized Sharpe and Kelly ratios.

**Keywords**: *Portfolio analysis, performance measurement, Hellinger's distance*

**JEL codes**: G11

## 1. Introduction

Papers exploring and expanding ways of portfolio construction other than Markowitz mean-variance portfolio (MP), are numerous. While some of them directly incorporate other components in the objective function optimized (like adding skewness (Lai, Lean Yu, Shouyang, 2006; De Athayde, Gustavo, Renato, 2003) or smoothing with entropy (Mercurio, Yuehua, Hong, 2020), or incorporating tail risk instead of variance (Yao, Zhongfei, Yongzeng, 2013)), others take different approaches based on newer ideologies (like in the risk-budgeting approach (Roncalli, 2013)).

The uses of Statistical distances in portfolio theories were primarily motivated by tracking problems (Svetlozar et. al. 2013), (i.e., the desire to keep closer to chosen maybe not-fully-known portfolios).

However, some emerging literature still managed to incorporate a statistical distance approach in portfolio optimization problems, arguing to capture better outcomes (Kim et. al., 2022). For a more thorough analysis of distance-based models, see (Svetlozar et. al. 2008).

* **Mesrop Mesropyan** – Masters' student in the program of Science in Management, American University of Armenia. E-mail: mesropyan.m17@gmail.com, ORCID: https://orcid.org/0000-0001-6673-1867.
**Vardan Bardakhchyan** – Lectuter, Faculty of Mathematics and Mechanics, Chair of Actuarial and Financial Mathematics, YSU. E-mail: vardan.bardakhchyan@ysu.am, ORCID: https://orcid.org/0000-0002-7395-6199.



Received: 27.11.2024
Revised: 09.12.2024
Accepted: 27.12.2024

In this paper we continue our long-lasting research of the portfolio closest to Gaussian in distribution, which was proposed and thoroughly studied in (Mesropyan, Mkrtchyan, 2021; Bardakhchyan, Mesropyan, 2023; Mesropyan, Bardakhchyan, 2023). The distance used was taken to be Hellinger's one, based on purely theoretical and less on practical reasons. Here we call the portfolio thus obtained Hellinger-Normal portfolio (HNP).

Let us first introduce the motivation behind the proposed portfolios.

### *1.1. Market structure as a motivation*

While in most of the literature the proposed alternatives to MP are motivated by the desire to overcome the drawbacks of MP, which we will also elaborate on in this paper, our initial motivation was rather based on a constructive approach to obtain a diagnostic tool for market structure. More precisely we were seeking an indicator the abrupt changes of which will indicate the change of market structure, in a sense of crucially changed correlations between asset returns, or rather a comparable change in vector distribution of asset returns.

According to the central limit theorem (CLT), the more independent random variables are taken into consideration, the closer their rescaled sum is to a Gaussian distribution (Durrett, 2019).

On the other hand, if random variables under consideration are initially normally distributed, this distance is always 0 and does not change. Moreover, in this case non-correlatedness does not play a crucial role, i.e., even the sum of correlated normally distributed random variables is normally distributed.

The converse of this result, Cramer's decomposition theorem, states that if any Gaussian random variable can be decomposed into a sum of two independent ones, then both of the latter's are also Gaussian.

Not contrasting any of the above arguments, still stock markets behave differently in the following more precise sense. While most of the stocks at least in theory have close to normal distribution of their return, or have only slightly heavier tails, the distribution of return of their linear combination is not close to normal, or better to say is rather far from Guassianity (Egan, 2007; Shiryaev, Zuoquan, Xun, 2008; Canan, Meterelliyoz, Tinic. 2016).

One of the reasons behind this is that returns (log-returns) being normal entail the price distribution to be log-normal, and the linear combination of log-normally distributed random variables can substantially deviate from log-normal in tails, being closer at average (Beaulieu, Qiong, 2004).

On the other hand, while dealing with portfolios, one deals with linear combination of returns. Thus the non-Guassianity here cannot be addressed the same way as above. Rather it can be attributed to the combination of two factors: 1) slight deviations from normality in returns of individual stocks and 2) the existing correlation between them.

The first reason solely cannot capture the non-normality of portfolio returns in big enough (meaning having a larger landscape of stocks) portfolios, as thanks to CLT, the return of linear combination would converge in distribution in an absence of correlations.

Thus we argue that the second reason is still more crucial in explanations of the non-Gaussianity of returns, as CLT is not known to hold even in stationary but correlated sequences (Ibragimov, 1975).

While trying to minimize the distance between the normal distribution and the distribution of returns of our portfolio, we found non-negligible and non-vanishing positive distance, indicating the existence of correlations (Mesropyan, Bardakhchyan, 2023). Thus the positive minimum Hellinger distance is an indicator of the latter. Even more, the changes in the magnitude of this distance will indicate the change of correlational structure.

### *1.2. The drawbacks of MP as a motivation.*

Several drawbacks of MP motivated further investigation into other methods of portfolio construction. Some more or less viable drawbacks that may be deemed to be not entirely rational are the following:

(i) In several cases MP gives 0 weights to most of the assets, even for not close to extremes desired expected returns. This feature would not be regarded as a violation of common sense, if we explore all assets given 0 weight to be dominated by the others. While often this is indeed the case, there are situations where the dropped assets are not dominated. This, together with the insufficiency of historical returns, gave rise to expert-based and more smoothed portfolios like entropy-added and Black-Litterman models.

Cevizci, 2016; Michaud, Tongshu, 2015)

(ii) In some cases, MP is still too sensitive to small changes in perceived expected returns, leading to minor changes in expected return estimates that entail some extreme changes of portfolio weights. This can be even combined with the pervious point, suggesting to get rid of some assets which had not so small weights previously. In a setting where assets have an internal bid-ask spread or are tangible and require transportation costs, this may lead to substantial money outflow. Some portfolios like Black-Litterman still overcome even this drawback. (Becker, Marc, 2010; Qian, Gorman, 2001).

(iii) MP does not incorporate all the information. MP portfolios rely heavily on only two descriptive statistics: expectation and variance. While in all-Gaussian framework, these two are enough, in reality there are at least two other features of interest: skewness and kurtosis/VaR/CVaR (as indicators of extreme events). It has been shown that some of the investors extremely value positive skewness and are very averse to extreme losses. These gave rise to several other portfolio theories (Lai, Lean Yu, Shouyang, 2006; Kraus, Litzenberger. 1976; Post, Vliet, Levy, 2008; Kane, 1982).

HNP can handle some of these drawbacks. We had shown previously that it overcomes the (i), may overcome (ii), in that it is less sensitive. And still in most cases fixes (iii). While one should be cautious, that it does not maximize the skewness or minimize the kurtosis, but rather struggles to keep them fixed.

In addition to the above, HNP also brings some predictability with itself; in contrast to what entropy theory would suggest for normal distributions, the predictability here should be understood in a sense that having chosen HNP one can indeed confine himself to two measurements, estimation of expected return and variance.

More precisely, while MP exploits these two estimates, ignoring others, HNP does drive us to the domain, where the latter are of, indeed, lesser importance (as they would not change substantially).

One other crucial feature of HNP, is that it incorporates correlations in a non-linear manner. While the correlational structure in MP is included through the variance, HNP still enables one to keep (but not to observe directly) the track of changes in correlations.

### *1.3. Hellinger-Normal portfolios construction*

Typical construction of HNP portfolio has been described in our previous works [11]. Here we briefly recall main steps. First, the classical MP problem is solved for a given set of assets that will not change during any step of the problem.

$$\begin{cases} E(X) = \overline{e} \\ Var(X) \to min \\ X = \sum_{i=1}^{n} X_i \\ w_i \geq 0 \\ \sum_{i=1}^{n} w_i = 1 \end{cases} \tag{1}$$

Here $X_i$-s are returns of each asset. We assume only a finite number of assets. We confine ourselves to the case of logarithmic return. The weights $w_i$-s represent the percentage of money in each asset. We assume only long positions i.e. $w_i \geq 0$.

After solving (1), we make use of found minimal variances ($\overline{\sigma}^2$) for each level of expected return[1].

$$\begin{cases} E(X) = \overline{e} \\ H^2\left(X, N\left(\overline{e}, \overline{\sigma}^2\right)\right) \to min \\ X = \sum_{i=1}^{n} X_i \\ w_i \geq 0 \\ \sum_{i=1}^{n} w_i = 1 \end{cases} \tag{2}$$

Where $H(X, Y)$ – is Hellinger's distance between X, Y r.v.-s[2].

Here, the parameter $\overline{\sigma}^2$ is just short hand notation for the found minimal variance corresponding to given level of $\overline{e}$.

Let us keep track of the main calculations, by adding the definition of squared Hellinger distance. For a continuous case

$$H^2(f, g) = 1 - \int \sqrt{f(x)g(x)}dx \tag{3}$$

While with one of the random variables being continuous but made constant over bins in histogram construction, we have the following formula

$$H^2(X, Y) = 1 - \sum_{j=0}^{n-1} O_j \int_{a_j}^{a_{j+1}} \sqrt{f_Y(x)}dx \tag{4}$$

Generally $O_j$-s are any numbers, not bound to be different. We take the interval cut enough fine to have one value for each interval.

Whenever the counterpart's (Y's) distribution is normal each element of (4) will take the following form:

$$\begin{aligned} \int_{a_j}^{a_{j+1}} \sqrt{f_Y(x)} &= \frac{1}{\sigma^{\frac{1}{2}}(2\pi)^{\frac{1}{4}}} \int_{a_j}^{a_{j+1}} e^{-\frac{(x-\mu)^2}{4\sigma^2}} dx \\ &\xrightarrow{\sigma_1 = \sigma\sqrt{2}} \sqrt{2}\sqrt{\sigma}(2\pi)^{\frac{1}{4}} \left(F_N\left(a_{j+1}|\mu, \sigma_1^2\right) - F_N\left(a_j|\mu, \sigma_1^2\right)\right) \end{aligned} \tag{5}$$

Hereafter we are less interested in the values of squared Hellinger distances than in the optimizing portfolio, i.e., the vector of optimal weights. We strive to compare the portfolios obtained by solving (1) and (2) respectively. We denote

---

[1] Solving (1) one gets two things for each fixed level of expected return: the weights' vector (i.e. the risk-minimizing portfolio) and the precise value of variance (i.e. some correspondence $\overline{e} \to \overline{\sigma^2}$).

[2] Here we use $(X, Y)$ , $H(F_X, F_Y)$ and in absolutely continuous case $H(f_X, f_Y)$ interchangeably.

the solving vectors by $w_{MP}$ and $w_{HN}$.

In the next section we describe the empirical procedure and data used, and compare performances of both portfolios in a not active management setting, in a sense made explicit further.

We argue that HN performs better than MP, with a small possible cost of higher variance and, in half of the situations, slightly more extreme loss. We show that by comparing 10 days' performance Sharpe ratio and Kelly ratio. Also we made it explicit that our results are not due to random effects.

## 2. Empirical analysis

For an illustrative example we have taken 3 years' (the period from 01/2021-01/2024) data of 13 assets (10 of which are included in S&P500 index)[3].

We added three other assets to have negative correlations for at least some period of time (see Table 1). For example, the 12-month log-returns correlation matrix has the following form for the first year.

***Table 1***

**Correlation matrix for assets.**

| | ADBE | CAT | DIS | FDX | GS | IBM | JNJ | KO | NKE | XOM | SPOT | BRN | FET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ADBE | 1.00 | -0.04 | 0.14 | 0.25 | -0.01 | -0.09 | 0.09 | 0.21 | 0.33 | -0.02 | 0.41 | 0.03 | 0.03 |
| CAT | -0.04 | 1.00 | 0.43 | 0.31 | 0.67 | 0.32 | 0.18 | 0.19 | 0.14 | 0.58 | 0.08 | 0.17 | 0.29 |
| DIS | 0.14 | 0.43 | 1.00 | 0.21 | 0.43 | 0.17 | 0.11 | 0.30 | 0.24 | 0.36 | 0.17 | 0.23 | 0.20 |
| FDX | 0.25 | 0.31 | 0.21 | 1.00 | 0.31 | 0.17 | 0.14 | 0.22 | 0.11 | 0.29 | 0.22 | 0.13 | 0.12 |
| GS | -0.01 | 0.67 | 0.43 | 0.31 | 1.00 | 0.27 | 0.14 | 0.17 | 0.24 | 0.57 | 0.16 | 0.15 | 0.31 |
| IBM | -0.09 | 0.32 | 0.17 | 0.17 | 0.27 | 1.00 | 0.26 | 0.29 | 0.07 | 0.35 | -0.03 | 0.09 | 0.10 |
| JNJ | 0.09 | 0.18 | 0.11 | 0.14 | 0.14 | 0.26 | 1.00 | 0.43 | 0.10 | 0.14 | -0.04 | 0.18 | 0.01 |
| KO | 0.21 | 0.19 | 0.30 | 0.22 | 0.17 | 0.29 | 0.43 | 1.00 | 0.16 | 0.20 | 0.04 | 0.19 | 0.05 |
| NKE | 0.33 | 0.14 | 0.24 | 0.11 | 0.24 | 0.07 | 0.10 | 0.16 | 1.00 | 0.14 | 0.23 | 0.09 | 0.19 |
| XOM | -0.02 | 0.58 | 0.36 | 0.29 | 0.57 | 0.35 | 0.14 | 0.20 | 0.14 | 1.00 | 0.01 | 0.23 | 0.41 |
| SPOT | 0.41 | 0.08 | 0.17 | 0.22 | 0.16 | -0.03 | -0.04 | 0.04 | 0.23 | 0.01 | 1.00 | 0.07 | 0.09 |
| BRN | 0.03 | 0.17 | 0.23 | 0.13 | 0.15 | 0.09 | 0.18 | 0.19 | 0.09 | 0.23 | 0.07 | 1.00 | 0.10 |
| FET | 0.03 | 0.29 | 0.20 | 0.12 | 0.31 | 0.10 | 0.01 | 0.05 | 0.19 | 0.41 | 0.09 | 0.10 | 1.00 |

Source: Authors' calculations.

We took the following procedure when constructing MP and HN. We used 1 year of data before the taken date as a historical input. We then calculated log returns of each of the assets and evaluated the mean returns based on the 1-year historical data.

Next we used the range of feasible expected returns of the portfolio (i.e., the range between the highest expected return out of 13 and the smallest). Then we divided the interval into 10 equal parts, and on the split point of each interval, we solved the MP problem (1) to find both the MP portfolio weights and the minimal variances for each level of expected return. The latter we used in HN portfolio construction according to (2).

To obtain the solution of (2) we employed formulas (4) and (5), for which we used the binning procedure to obtain a histogram of the portfolio return distribution of the vector of weights. We used the optimal number of bins according to Sturges's formula.

After solving and obtaining the optimal portfolios, we keep the weights for another 10 trading days. And after the period we evaluate the performance.

[3] ADBE; CAT; DIS; FDX; GS; IBM; JNJ; KO; NKE; XOM; SPOT; BRN; FET.

After the period we shifted our historical data by these 10 days, or put more simply, we moved to the next date where we are allowed to change the weights. There we did the same procedure again and so forth. Thus we got 43 cumulative descriptive statistics for a period observed.

Based on the observations we obtained the following results.

(1) At an average of 11.6% (5 out of 43) of cases, MP showed better realized performance in all measures. While in all other cases HNP showed better performance on at least 1 of the two measures considered.

(2) As for tail risk, in most of the cases HNP did worse compared to MP. However, the difference still may be considered insignificant, as in many of these cases the extreme losses do not deviate even by 3% compared to each other.

(3) In most of the cases, HNP did significantly better compared to more extreme worst-case losses. More precisely, on average, HNP showed 21.4% better performance [1] (i.e., 1,214 times more average return), while showing, on average, only 4.3% more extreme case losses (so -1.043 times more losses). Cumulatively, if we kept only HNP portfolios most deviant from MP, we would, on average, get up to 5.68% more income, with the risk of losing up to 1.25% more money in each 10 days.

The standard deviation differences are up 0.18% different, so that HNP has on average a 1.018 higher standard deviation than MP. While cumulatively in 3 years it may bring a 7.6% difference in standard deviation.

(4) In periods with explicitly dominant assets HNP showed generally worse performance, than MP.

Some results are given in the following Table 2.

***Table 2***

**Comparison of some realized numerical characteristics of MP and HNP, when adjustments are only allowed biweekly (each 10 market days).**

| | MP | HNP | Comment |
|---|---|---|---|
| Daily mean realized return | -0.0011 | -0.0002 | So in each of the cases adjustments to portfolios are made every week |
| Biweekly mean realized return | 0.0015 | 0.0067 | This result differ from the previous ones, most probably due to seasonality |
| Daily mean realized standard deviation | 0.0122 (1.22%) | 0.0128 (1.28%) | Note that in average, standard deviation is only slightly higher. |
| Mean biweekly Sharpe ratio | 0.063 | 0.115 | These results are approximated to 3 digits, after the floating point. |
| Mean biweekly Kelly ratio | 18.5 | 23.04 | |
| Mean Extreme values | -0.0012 (1.2% loss) | -0.0014 (1.4% loss) | In extreme cases losses could be as high as 14% |

**Discussion and conclusion.**

Each of the four main results found has different but not contradictory compelling reasons.

(1) On average HNP showed better realized performance. While for Sharpe ratio it is not obvious why, the Kelly criteria case has more theoretical reasoning.

The Kelly criterion assesses a better fit for dynamic (i.e., long-run) investment choices. As we had chosen somewhat not long, but also not daily management of the portfolio, the Kelly criterion should be the more crucial one. And Kelly criterion favored HNP.

Initially HNP was proposed as a more robust alternative to MP. In that sense it should have been a better choice for long-run investment.

One more precise reason is that the Kelly criterion is based on maximizing the geometric average of returns, while MP only considers average and variance. It is known that the geometric average can be modeled as the logarithm of the portfolio for non-negative returns, which, when decomposed into Taylor series, includes all initial moments (cumulants), not only the first two. This means that the portfolio that fixes all moments (like taking the portfolio to be close to a normal distribution) will perform better than the portfolio only considering the first few moments.

Things differ for the Sharpe ratio case. The MP is the exact portfolio that maximizes the Sharpe ratio; still, HNP performs approximately twice as good as MP. The reason may lie in the changing environment, i.e., the changing multivariate distribution of returns. While with HNP one tries to "fix" the distribution as close to normal as it is possible, MP works with this more "complex" dynamic and still ignores most of the information.

(2) The reason why HNP showed more extreme losses than MP may be due to the fact that the tails are not entirely close to that of normal distribution. This is the fact that HNP tries to bypass. But it may be the case that thus it confines to the cases that may have "worse" tails (i.e., adjusting to heavier tails) than in MP. In other words, it may be that MP solutions may have less kurtosis than 3 (the kurtosis of normal distribution). So, it is reasonable to use HNP when the normality assumption is countered by the heavy tails hypothesis. Still, this difference is not drastically extreme.

(3) This part is self-obvious and is partly discussed above. We can't explain the magnitude of the difference. One should still be cautious, as extreme losses with percentage differences, though small, are in fact bigger in magnitude (bigger in pure amount). Standard deviation differences still do not mean that HNP cumulatively did lose more than MP. Rather, it means that general fluctuations around the better mean are slightly higher than in MP around its' mean.

(4) It is obvious that the one thing HNP does is try to overcome corner solutions. When one has obvious dominant assets, it is not reasonable to try to forcefully deviate from corner solutions. Thus the HNP procedure becomes redundant. In that case, even uniform weights do better than HNP.

In conclusion, we propose to use HNP only when there are no obviously dominant assets. However, in all other cases, HNP will most probably show better performance. Still, several other comparisons should be made.

## References

Bardakhchyan, Vardan, and Mesrop Mesropyan. 2023. "Hellinger's distance to normal distribution as market invariant." *Bulletin of Yerevan University G: Economics* 14, no. 1 (40): 64-71. https://doi.org/10.46991/BYSU:G/2023.14.1.064

Beaulieu, Norman C., and Qiong Xie. 2004. "An optimal lognormal approximation to lognormal sum distributions." *IEEE Transactions on vehicular technology* 53.2: 479-489. DOI: 10.1109/TVT.2004.823494

Becker, Franziska, and Marc Gürtler. 2010. "Quantitative forecast model for the application of the Black-Litterman approach." *Paris December 2009 Finance International Meeting AFFI-EUROFIDAI*. Available at SSRN: https://ssrn.com/abstract=1483571 or http://dx.doi.org/10.2139/ssrn.1483571

Cevizci, Aylin. 2016. "A comparison of optimal portfolio performances of three optimization methods." *International Journal of Commerce and Finance* 2.1: 137-146.

Corlu, Canan G., Melike Meterelliyoz, and Murat Tinic. 2016. "Empirical distributions of daily equity index returns: A comparison." *Expert systems with applications* 54: 170-192; https://doi.org/10.1016/j.eswa.2015.12.048

De Athayde, Gustavo M., and Renato G. Flôres. 2003. "Incorporating skewness and kurtosis in portfolio optimization: A multidimensional efficient set." In *Advances in portfolio construction and implementation*, pp. 243-257. Butterworth-Heinemann. https://doi.org/10.1016/B978-075065448-7.50011-2

Durrett, Rick. 2019. *Probability: theory and examples*. Vol. 49. Cambridge University Press.

Egan, William J., 2007. "The distribution of S&P 500 index returns." *Available at SSRN 955639*. Available at SSRN https://ssrn.com/abstract=955639 or http://dx.doi.org/10.2139/ssrn.955639

Ibragimov, I. 1975."Independent and stationary sequences of random variables." *Wolters, Noordhoff Pub.*

Kane, Alex. 1982. "Skewness preference and portfolio choice." *Journal of Financial and Quantitative Analysis* 17, no. 1: 15-25. https://doi.org/10.2307/2330926

Kim, Hongseon, Soonbong Lee, Seung Bum Soh, and Seongmoon Kim. 2022. "Improving portfolio investment performance with distance-based portfolio-combining algorithms." *Journal of Financial Research* 45, no. 4: 941-959. https://doi.org/10.1111/jfir.12303

Kraus, Alan, and Robert H. Litzenberger. 1976. "Skewness preference and the valuation of risk assets." *The Journal of finance* 31, no. 4: 1085-1100. https://doi.org/10.2307/2326275

Lai, Kin Keung, Lean Yu, and Shouyang Wang. 2006. "Mean-variance-skewness-kurtosis-based portfolio optimization." In *First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, vol. 2, pp. 292-297. IEEE. DOI: 10.1109/IMSCCS.2006.239

Mercurio, Peter Joseph, Yuehua Wu, and Hong Xie. 2020. "An entropy-based approach to portfolio optimization." *Entropy* 22, no. 3: 332. https://doi.org/10.3390/e22030332

Mesropyan, M. T., and V. G. Bardakhchyan. 2023. "Hellinger's Distance and Correlation for a Subclass of Stable Distributions." *Journal of Contemporary Mathematical Analysis (Armenian Academy of Sciences)* 58, no. 3: 191-195. https://doi.org/10.3103/S106836232303007X

M. Mesropyan and V. Mkrtchyan. 2021. "Assessing normality of group of assets based on portfolio construction", *Alternative scientific journal, 3,* pp. 14-21.

Michaud, Richard O., and Tongshu Ma. 2015. "Efficient asset management: a practical guide to stock portfolio optimization and asset allocation.": 901-904.https://doi.org/10.1093/rfs/14.3.901

Post, Thierry, Pim Van Vliet, and Haim Levy. 2008. "Risk aversion and skewness preference." *Journal of Banking & Finance* 32, no. 7: 1178-1187.

Qian, Edward, and Stephen Gorman. 2001. "Conditional distribution in portfolio theory." *Financial Analysts Journal* 57.2: 44-51. https://doi.org/10.2469/faj.v57.n2.2432

Rachev, Svetlozar T., Lev B. Klebanov, Stoyan V. Stoyanov, Frank Fabozzi. 2013. "Main Directions in the Theory of Probability Metrics." *The Methods of Distances in the Theory of Probability and Statistics*: 1-7. https://doi.org/10.1007/978-1-4614-4869-3 1

Rachev, Svetlozar, Sergio Ortobelli, Stoyan Stoyanov, Frank J. Fabozzi, and Almira Biglova. 2008. "Desirable properties of an ideal risk measure in portfolio theory." *International Journal of Theoretical and Applied Finance* 11, no. 01: 19-54. https://doi.org/10.1142/S0219024908004713

Roncalli, Thierry. 2013. *Introduction to risk parity and budgeting*. CRC Press.

Shiryaev, Albert, Zuoquan Xu, and Xun Yu Zhou. 2008. "Thou shalt buy and hold." *Quantitative finance* 8.8: 765-776. https://doi.org/10.1080/14697680802563732

Yao, Haixiang, Zhongfei Li, and Yongzeng Lai. 2013. "Mean–CVaR portfolio selection: A nonparametric estimation framework." *Computers & Operations Research* 40, no. 4: 1014-1022. https://doi.org/10.1016/j.cor.2012.11.007