Downside Risk Optimization vs Mean-Variance Optimization

Andrea Rigamonti

Free University of Bozen-Bolzano

November 22, 2019

Introduction

Classical framework in modern portfolio theory:

- Investors care about **mean and variance** of the returns.
- Investors want to maximize the **Sharpe ratio**:

$$SR = \frac{R - R_f}{\sigma}. \tag{1}$$

Introduction

Classical framework in modern portfolio theory:

- Investors care about **mean and variance** of the returns.
- Investors want to maximize the **Sharpe ratio**:

$$SR = \frac{R - R_f}{\sigma}. \tag{1}$$

Most likely, real investors:

- Only consider **downside deviation** (variability below a certain benchmark) as risk.
- Care about **mean and semivariance** of the returns.
- Want to maximize **the Sortino Ratio**.

Introduction

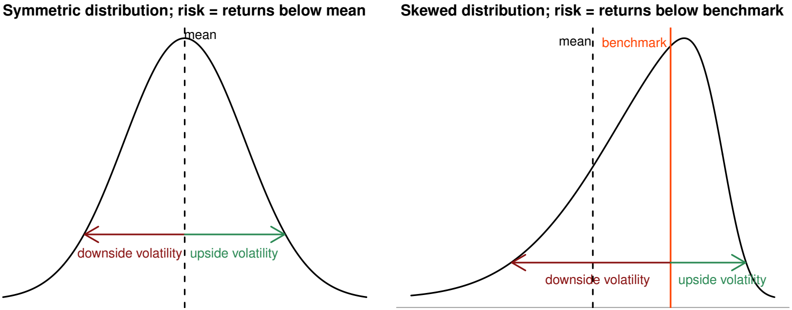
Why is mean-variance optimization more popular?

- Mean-semivariance optimization requires the estimation of a **semicovariance matrix**.
- "Traditional" obstacle: this matrix is **endogenous**.
- Additional issue: more **parameter uncertainty**.
- Other measures of downside risk, like **CVaR**, also perform poorly for the same reason.

Message of the paper: even if using downside risk is feasible, optimizing mean-variance remains preferable in most cases.

Background - Downside volatility

Minimizing variance is equivalent to minimizing downside volatility only if the return distribution is symmetric:



Background - Downside deviation and Sortino ratio

Given T returns r and a benchmark B, downside volatility/risk is measured by **downside deviation**:

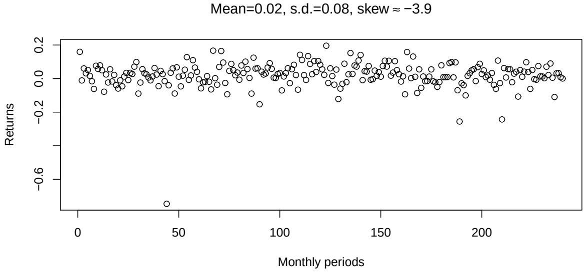
$$\sigma_B = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[Min(r_t - B, 0)]^2}. \tag{2}$$

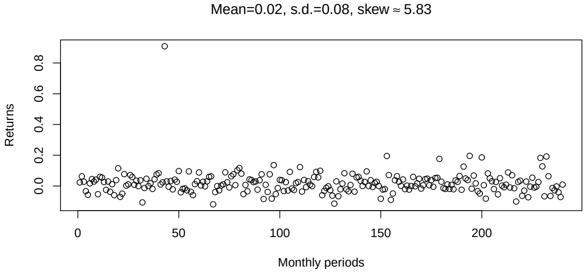
The risk-adjusted return of the investment can then be measured with the **Sortino ratio**:

$$Sortino = \frac{R - B}{\sigma_B}, \tag{3}$$

where R is the average return of the investment.

Background - Downside deviation and Sortino ratio

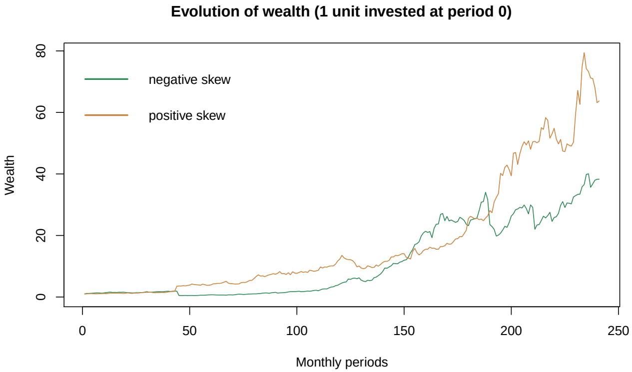




Background - Downside deviation and Sortino ratio

The two series have exactly the same Sharpe Ratio.

With $B = 0.005$: for the negatively skewed one Sortino ratio $=$ 0.242; for the positively skewed one Sortino ratio $= 0.491$.



Background - The semicovariance matrix

The variability below the benchmark B of a set of assets is measured by the semicovariance matrix $\Sigma_{B,p}$.

The semivariance of a portfolio of N assets (with returns r_{it}) and observation window T that underperforms the benchmark in K periods is given by (Markowitz, 1959):

$$S_{B,p} = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \frac{1}{N} \sum_{t=1}^{K} (r_{it} - B)(r_{jt} - B) \tag{4}$$

Background - The semicovariance matrix

The variability below the benchmark B of a set of assets is measured by the semicovariance matrix $\Sigma_{B,p}$.

The semivariance of a portfolio of N assets (with returns r_{it}) and observation window T that underperforms the benchmark in K periods is given by (Markowitz, 1959):

$$S_{B,p} = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \frac{1}{N} \sum_{t=1}^{K} (r_{it} - B)(r_{jt} - B) \tag{4}$$

Problem: using (4) to select the periods that underperform B yields an **endogenous** matrix. A change in the weights w_i and w_j changes K and therefore the elements of $\Sigma_{B,p}$.

Optimization problems that use $\Sigma_{B,p}$ are intractable.

Background - Exogenous approximation of $\Sigma_{B,p}$

Estrada (2008) solution: compute the elements of the semicovariance matrix as:

$$\Sigma_{B,ij} = \frac{1}{T}\sum_{t=1}^{T}[Min(r_{it} - B, 0) \cdot Min(r_{jt} - B, 0)] \qquad (5)$$

- Equation (5) is based on whether individual assets (not the portfolio) underperform B $\rightarrow$ $\boldsymbol{\Sigma}_{B,ij}$ **is exogenous**.
- $\Sigma_{B,ij}$ well approximates $\Sigma_{B,p}$ (and is symmetric).

Mean-semivariance optimization is now feasible!

How good is Estrada's approximation?

Cheremushkin (2009): approximation error is substantial unless assets are positively correlated.

How good is Estrada's approximation?

Cheremushkin (2009): approximation error is substantial unless assets are positively correlated.

I compare the exact (numerical) and Estrada solution.

		DownDev		Sortino	
Corr.	Skew	Estrada	Num.	Estrada	Num.
-0.013	(-0.3,-0.3)	0.0371	0.0363	0.1894	0.2114
	(0.3,0.3)	0.0280	0.0268	0.2499	0.2932
0.003	(-0.3,-0.3)	0.0323	0.0311	0.2779	0.2837
	(0.3,0.3)	0.0265	0.0245	0.3202	0.3371
0.300	(-0.3,-0.3)	0.0328	0.0326	0.3391	0.3560
	(0.3,0.3)	0.0263	0.0261	0.4221	0.4467
0.743	(-0.3,-0.3)	0.0566	0.0563	0.1388	0.1433
	(0.3,0.3)	0.0508	0.0506	0.1445	0.1481

Table 1: Downside deviation (min var/semivar portfolio) and Sortino ratio (mean-var/semivar portfolio) with skew normal returns and 240 months rolling windows

Which strategy works better?

In-sample:

- Zero skew: optimizing mean-semivariance or mean-variance is equivalent.
- Non-zero skew: each strategy is more efficient in terms of its objective function.

Which strategy works better?

In-sample:

- Zero skew: optimizing mean-semivariance or mean-variance is equivalent.
- Non-zero skew: each strategy is more efficient in terms of its objective function.

But what happens out-of-sample?

- Equation (5) only uses a subset of the T observations. **More parameter uncertainty** in the semicovariance matrix.
- Zero skew: using variance is always preferable.
- Non-zero skew: using variance can still be superior even in terms of downside risk/Sortino ratio.

<u>Using the wrong objective function can be rational!</u>

Empirical study - One asset case

I consider one asset and want to estimate the semivariance σ_B^2.

Two possible estimators:

1. sample semivariance using sample downside deviation $\rightarrow$ always unbiased, but high variance;
2. sample variance/2 $\rightarrow$ biased if return distribution is not symmetric, but lower variance.

Empirical study - One asset case

I consider one asset and want to estimate the semivariance σ_B^2.

Two possible estimators:

1. sample semivariance using sample downside deviation $\rightarrow$ always unbiased, but high variance;
2. sample variance/2 $\rightarrow$ biased if return distribution is not symmetric, but lower variance.

The best estimator has the lowest Mean Squared Error:

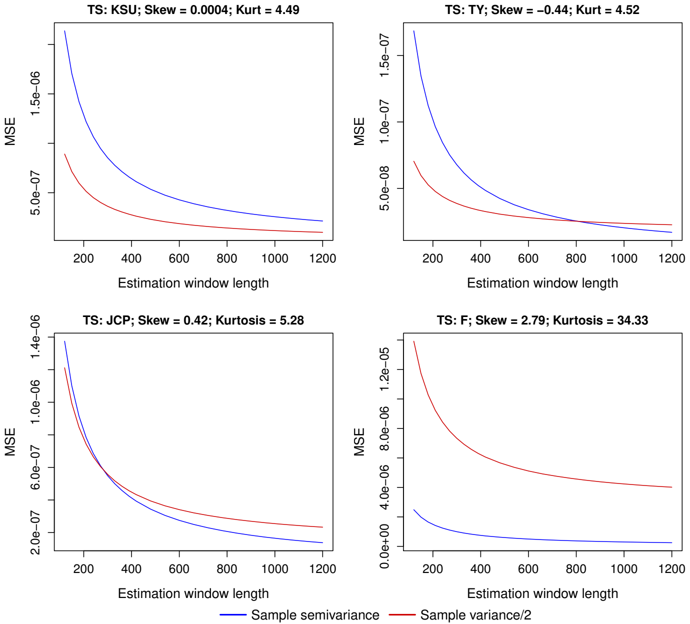
$$MSE(\hat{\sigma}_B^2) = \mathrm{Var}(\hat{\sigma}_B^2) + (\mathrm{Bias}(\hat{\sigma}_B^2, \sigma_B^2))^2, \tag{6}$$

where $\mathrm{Bias}(\hat{\sigma}_B^2, \sigma_B^2) = \mathrm{E}(\hat{\sigma}_B^2) - \sigma_B^2$.

Empirical study - One asset case

- The more asymmetric the distribution is, the higher is the bias of the second estimator.
- The longer the estimation window is, the lower the variance of the estimators.
- I consider four different series with different skewness (KSU, TY, JCP, F) and several different sample sizes.
- For each series and sample size I draw with replacement 500000 samples of monthly returns.
- For each series the benchmark B is set equal to the mean.
- I estimate semivariance with both estimators and compare the MSE.

Empirical study - One asset case



Empirical study - Portfolio with two risky assets

Analogously to the case with one asset:

- For a given T, mean-semivariance should work better than mean-variance if returns are sufficiently skewed.
- Given a non-zero skew, mean-semivariance should work better than mean-variance with a sufficiently high T.

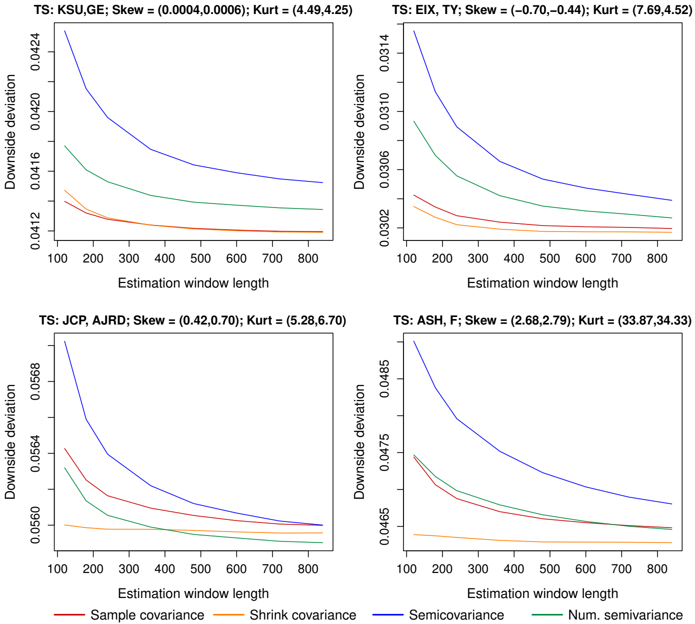
Advantage of mean-variance: shrinkage estimators for the covariance matrix (Ledoit & Wolf, 2004).

Does mean-semivariance optimization ever work well in realistic conditions?

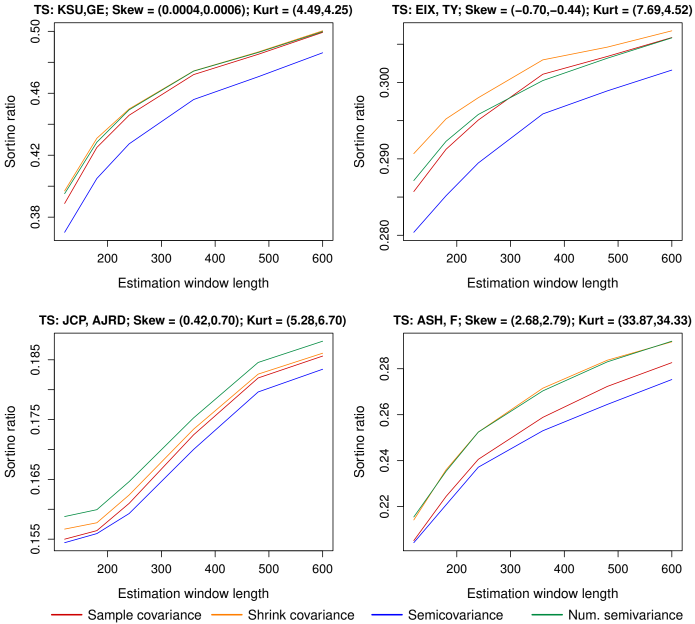
Empirical study - Portfolio with two risky assets

- I consider four portfolios with two risky assets and one riskless asset (risk-free rate from Fama-French).
- The pairs of assets have different degrees of skewness.
- I generate 500000 returns for each asset by drawing with replacement from real empirical monthly returns series.
- I set the benchmark equal to the risk-free rate.
- Using the sample and shrunk covariance matrix, and the Estrada (2008) and numerical solution, I compute first the minimum variance/semivariance portfolios, and then the mean-variance/semivariance portfolios.
- I use different rolling windows, and compare the downside deviation and Sortino ratio.

Minimum variance/semivariance portfolios



Mean-variance/semivariance portfolios



Results from two assets case

- Estrada (2008) never consistently outperforms the variance-based portfolios.
- Minimizing the semivariance numerically is only competitive with a positive skew and low kurtosis.
- High kurtosis seems to hurt semivariance-based optimization more than variance-based optimization.

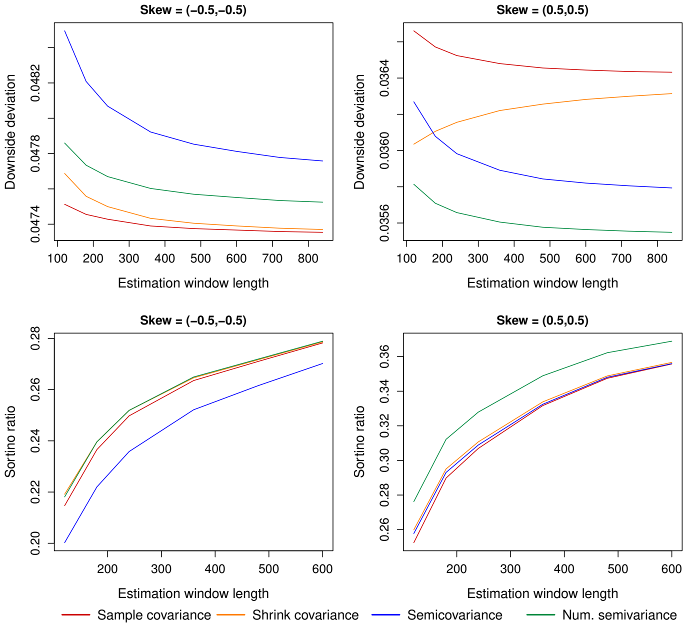
Results from two assets case

- Estrada (2008) never consistently outperforms the variance-based portfolios.
- Minimizing the semivariance numerically is only competitive with a positive skew and low kurtosis.
- High kurtosis seems to hurt semivariance-based optimization more than variance-based optimization.

Results might be better when the distribution is positively skewed because more observations fall below the benchmark.

I check if this holds with skew normal distributed returns with the same mean and covariance but different skew.

Negative vs. positive skew



Alternative approach: minimizing CVaR

CVaR is a more popular measure of downside risk.

Alternative approach: minimizing CVaR

CVaR is a more popular measure of downside risk.

Lim et al. (2011) find that it suffers from similar problems.

Alternative approach: minimizing CVaR

CVaR is a more popular measure of downside risk.

Lim et al. (2011) find that it suffers from similar problems.

I test 1/N, minimum variance (using Ledoit & Wolf, 2004), minimum semivariance (using Estrada, 2008) and minimum CVaR (95% CI) portfolios on real data:

- dataset: weekly DowJones returns (Feb 1990 - Apr 2016) from Bruni (2016);
- 5 and 10 years rolling window estimation;
- 843 out-of-sample returns;
- with and without short-selling (Jagannathan, 2003).

Alternative approach: minimizing CVaR

Strategy	**Dw. Dev.**	**Sortino**	**CVaR**
1/N	0.0172	0.0791	-0.0555
Min Var	0.0140	0.0552	-0.0449
Min Var Long	0.0139	0.0713	-0.0454
Min Semivar	0.0163	0.0014	-0.0521
Min Semivar Long	0.0144	0.0485	-0.0459
Min CVaR	0.0168	0.0047	-0.0539
Min CVaR Long	0.0148	0.0388	-0.0484

Table 2: 5 years rolling window estimation

Alternative approach: minimizing CVaR

Strategy	**Dw. Dev.**	**Sortino**	**CVaR**
1/N	0.0172	0.0791	-0.0555
Min Var	0.0138	0.0689	-0.0454
Min Var Long	0.0141	0.0795	-0.0465
Min Semivar	0.0150	0.0132	-0.0483
Min Semivar Long	0.0145	0.0537	-0.0478
Min CVaR	0.0148	0.0078	-0.0467
Min CVaR Long	0.0143	0.0546	-0.0469

Table 3: 10 years rolling window estimation

Conclusions

- Precise estimates of the semicovariance matrix are very difficult to obtain.
- In a simulation study, minimizing the semivariance never consistently yields a lower downside deviation or higher Sortino ratio than portfolios that minimize the variance.

Conclusions

- Precise estimates of the semicovariance matrix are very difficult to obtain.
- In a simulation study, minimizing the semivariance never consistently yields a lower downside deviation or higher Sortino ratio than portfolios that minimize the variance.
- The CVaR has similar estimation issues.
- Both minimum semivariance and minimum CVaR portfolios fail to beat the minimum variance portfolio in an empirical application.

Conclusions

- Precise estimates of the semicovariance matrix are very difficult to obtain.
- In a simulation study, minimizing the semivariance never consistently yields a lower downside deviation or higher Sortino ratio than portfolios that minimize the variance.
- The CVaR has similar estimation issues.
- Both minimum semivariance and minimum CVaR portfolios fail to beat the minimum variance portfolio in an empirical application.
- **Problem is in the concept of downside risk**, not in the way we measure it.
- The popularity of variance as a measure of risk is rational and empirically justified.

Thank You!