Sparse spanning portfolios and under-diversification with second-order stochastic dominance

Stelios Arvanitis
(AUEB)

Olivier Scaillet
(UNIGE and SFI)

Nikolas Topaloglou
(AUEB and IPAG)

EEA-ESEM Congress, Barcelona, 30 August 2023

The paper: Theoretical contributions

- We develop and implement methods for determining whether relaxing sparsity constraints on portfolios improves the investment opportunity set for risk-averse investors.
- We formulate a new estimation procedure for sparse second-order stochastic spanning based on a greedy algorithm and Linear Programming.
- We show the optimal recovery of the sparse solution asymptotically whether spanning holds or not.

The paper: Empirical findings

- From large equity datasets, we estimate the expected utility loss due to possible under-diversification, and find that there is no benefit from expanding a sparse opportunity set beyond 40 assets.
- The optimal sparse portfolio invests in 10 industry sectors with a larger weighting on small size, high book-to-market, and momentum stocks from the S&P 500 index and cuts tail risk when compared to a sparse mean-variance portfolio.
- On a rolling-window basis, the number of assets shrinks to around 20 assets in crisis periods.

The questions that we address

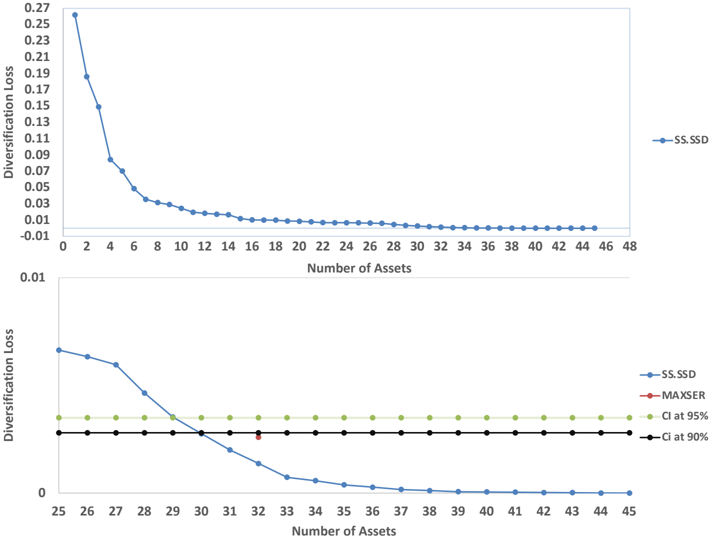
- Is it possible to build a sparse portfolio of dimension q from a large set of assets of dimension p so that we cannot get further improvement from considering additional assets in a second-order stochastic dominance (SSD) paradigm?

- If not, how much do we loose by limiting ourselves to this sparse portfolio in terms of expected utilities compatible with SSD?

- Can we design an optimization algorithm to compute this sparse portfolio from available data?

- Do we have the asymptotic statistical guarantee that we cannot improve on the estimated expected utility loss due to under-diversification by considering another sparse portfolio of the same fixed dimension?

The under-diversification loss in the SSD setting



The under-diversification in the mean-variance (MV) setting

We revisit the late 60s literature for the MV setting (Evans and Archer (1968)) but under the lens of SSD.

Diversification and the Reduction of Dispersion 765

TABLE 1
The following describes in tabular form the portfolio selection process for one computer run:

	.	1	2	3	4	5	. . . . .	40
Securities selected	1	113	44	11	137	55	. . . . .	244
(by identification	2		219	269	27	200	. . . . .	88
number)	3			94	28	74	. . . . .	94
	4				235	71	. . . . .	19
	5					199	. . . . .	37
	.						. . . .	.
	.						. . .	.
	.						. .	.
	.							.
	40							222
Mean Portfolio Return ($\bar{R}_p$)		.117	.109	.087	.070	.071	. . . . .	.0608
Portfolio Standard Deviation (SD_p)		.269	.191	.170	.151	.141	. . . . .	.121

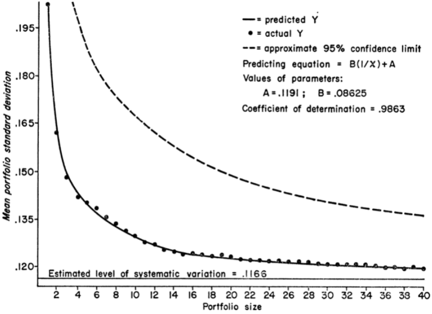


FIGURE 1

Recap on SSD

Utility functions and risk aversion:

U_1 = class of all non-decreasing utilities

U_2 = restriction of U_1 to its concave elements

Remarks:

$u' \geq 0$: increasing (prefer more than less)

$u'' \leq 0$: concave (risk aversion)

Risk aversion

Recap on SSD

Stochastic orderings:

Y_1 is said to be smaller than Y_2

- in the stochastic dominance order

  if $Eu(Y_1) \leq Eu(Y_2)$ for $u \in U_1$

denoted as $Y_1 \prec_d Y_2$

- in the increasing concave order

  if $Eu(Y_1) \leq Eu(Y_2)$ for $u \in U_2$

denoted as $Y_1 \prec_{icv} Y_2$

- Y_2 is preferred over Y_1 by all risk averters

  if $Y_1 \prec_{icv} Y_2$ and $EY_1 = EY_2$

denoted as $Y_1 \prec_{cv} Y_2$

Recap on SSD

Main characterizations in terms of cumulative distribution functions (cdf) and quantiles:

$$\text{(i) } Y_1 \prec_d Y_2 \Leftrightarrow F_1(x) \geq F_2(x), \forall x$$

$$\text{(ii) } Y_1 \prec_d Y_2 \Leftrightarrow$$

$$F_1^{-1}(p) \leq F_2^{-1}(p), \forall p \in (0,1)$$

$$\text{(iii) } Y_1 \prec_{icv} Y_2 \Leftrightarrow$$

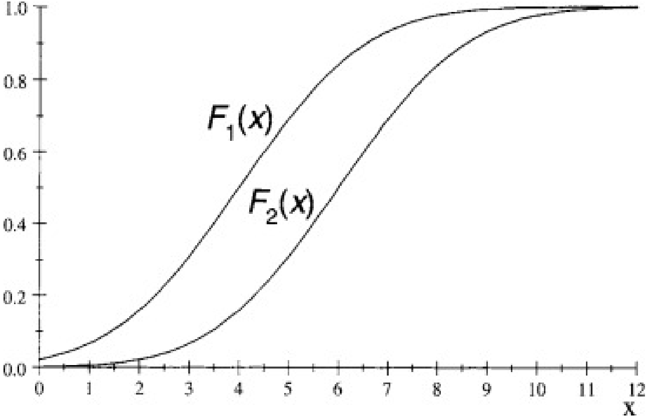
$$\int_{-\infty}^{x} F_1(u)du \geq \int_{-\infty}^{x} F_2(u)du, \forall x$$

$$\text{(iv) } Y_1 \prec_{icv} Y_2 \Leftrightarrow$$

$$\int_0^p F_1^{-1}(u)du \leq \int_0^p F_2^{-1}(u)du, \forall p \in (0,1)$$

Recap on SSD

Graph of characterizations in terms of cdf:



Sparse SSD-Probabilistic Framework

- Financial returns process X^{∞} in $\ell^{\infty}(\mathbb{N},\mathbb{R})$, X is the projection of X^{∞} in the first p coordinates, $\mathbb{P}$ is the distribution of X^{∞}.
- The portfolio weights universe Λ_{∞} a non-empty subset of the $\mathbb{N}$-simplex $\{\lambda \in \mathbb{R}^{\mathbb{N}} : \lambda_i \geq 0, i \in \mathbb{N}, \sum_{i=0}^{\infty} \lambda_i = 1\}$; $\Lambda = \left\{\lambda \in \Lambda_{\infty}, \sum_{i=0}^{p-1} \lambda_i = 1\right\}$, for $p \in \mathbb{N}$, denotes the $p-1$ dimensional unit sub-simplex of Λ_{∞}; K is a non-empty closed subset of Λ.
- λ, κ are generic elements of Λ^{∞}.

Assumption (MS-Moments and Supports)

For some $\varepsilon > 0$, $\max_{0<i\leq+\infty} \mathbb{E}\left[|X_i|^{2+\varepsilon}\right] < +\infty$ and $\inf \bar{\text{co}}\left[\cup_i \text{supp}(X_i)\right] > -\infty$.

- MS implies $D(z,\kappa,\lambda,\mathbb{P}) := \mathbb{E}\left[\left(z - \sum_{i=0}^{\infty} \kappa_i X_i\right)_+\right] - \mathbb{E}\left[\left(z - \sum_{i=0}^{\infty} \lambda_i X_i\right)_+\right]$ is bounded and continuous in z, λ, κ.

Sparse Spanning SSD

Definition (SSD in High Dimensional Setting)

$$\kappa \underset{\text{SSD}}{\succeq} \lambda, \text{ iff } D\left(z, \kappa, \lambda, \mathbb{P}\right) \leq 0 \text{ for all } z \in Z := \bar{\text{co}}\left[\cup_i \text{supp}\left(X_i\right)\right].$$

Definition (Sparse Spanning SSD)

For some fixed q, there exists a $K \subset \Lambda$ with $\text{csupp}\left(K\right) \leq q$ and such that $K \underset{\text{SSD}}{\succeq} \Lambda$, i.e. $\forall \lambda \in \Lambda$, $\exists \kappa \in K : \kappa \underset{\text{SSD}}{\succeq} \lambda$,
where $\text{csupp}\left(K\right) := \#\left\{i : \kappa_i \neq 0, \kappa \in K\right\}$ is the support of a portfolio set.

- The Sparse Spanning Definition generalizes Arvanitis et al. (2018) since i) it allows for a high dimensional setting, ii) it only prescribes the existence of a "low-dimensional" spanning subset of Λ.
- Any procedure designed to test whether SS-SSD holds, would have to search for a spanning set inside the collection of "low-dimensional" subsets of Λ.

Sparse Spanning SSD

Lemma (Sparse Spanning SSD Characterization)

Under Assumption MS then $K \underset{\text{SSD}}{\succeq} \Lambda$ iff $\inf_{\mathcal{L}_{p,q}} \sup_{\Lambda} \inf_{K} \sup_{z \in Z} D(z, \kappa, \lambda, \mathbb{P}) \leq 0$, with $\mathcal{L}_{p,q} := \{K \subset \Lambda : K \text{ closed}, 0 < \text{csupp}(K) \leq q\}$.

Lemma (Numerically Useful Characterization)

Under Assumption MS, and if Λ is closed in the Euclidean topology. Then, for all p,

$$\inf_{\mathcal{L}_{p,q}} \sup_{\Lambda} \inf_{K} \sup_{z \in Z} D(z, \kappa, \lambda, \mathbb{P}) = \sup_{z \in Z} \sup_{\Lambda} \inf_{\mathcal{L}_{p,q}} \inf_{K} D(z, \kappa, \lambda, \mathbb{P}).$$

- Allows for separation of the optimizations w.r.t. Λ and $\mathcal{L}_{p,q} \times K$, for any z.
- Useful especially when the outer optimization over Z is approximated by some discretization as in our empirical numerical implementation.

Empirical Sparse Optimization: Greedy Algorithm

- We design a procedure that evaluates $\sup_{z\in Z}\sup_{\Lambda}\inf_{\mathcal{L}_{p,q}}\inf_{K} D\left(z,\kappa,\lambda,\mathbb{P}_T\right)$; $\mathbb{P}_T$ is the empirical distribution.
- We use the Forward Stepwise Selection Algorithm for the $\inf_{\mathcal{L}_{p,q}}\inf_{K}$ part.

Algorithm (Elenberg et al. (2018))

Inputs: the sparsity Parameter $q<p$, the # of iterations $r_T(q)$, for a given set S the set function $2^p\to\mathbb{R}$ defined as

$$func\left(S\right):=\inf_{\mathrm{csupp}(S)\leq q}\frac{1}{T}\sum_{t=0}^{T}\left(z-\sum_{i=0}^{\infty}\kappa_i X_{i,t}\right)_{+}.$$

a Choose the initial set S_0,

b for $i=1,\ldots,r_T(q)$ do,

c $s:=\arg\max_{j\in[p]/S_{i-1}} func\left(S_{i-1}\cup\{j\}\right)-func\left(S_{i-1}\right)$,

d $S_i:=S_{i-1}\cup\{s\}$.

Empirical Sparse Optimization: Statistical Guarantees

Assumption (Restricted Strong Convexity-Restricted Smoothness)

X has a continuous density f. $\mathbb{E}\left(z - \sum_{i=0}^{\infty} \kappa_i X_{0,i}\right)_+$ is twice differentiable w.r.t. any κ appearing in some pair of $\Lambda_{(\lfloor q(\ln(T+1))\rfloor)}$ for all $z \in Z$. For $m_{\lfloor q(\ln(T+1))\rfloor}$ denoting the supremum and $M_{\lfloor q(\ln(T+1))\rfloor}$ the infimum over $\Lambda_{(\lfloor q(\ln(T+1))\rfloor)}$, of the smallest and the largest eigenvalues of the Hessian matrix of $\mathbb{E}\left(z - \sum_{i=0}^{\infty} \kappa_i X_{0,i}\right)_+$, then as $T \to \infty$, $\frac{m_{\lfloor q(\ln(T+1))\rfloor}}{M_{\lfloor q(\ln(T+1))\rfloor}} \ln T \to +\infty$ uniformly in Z.

- E.g. f is a Gaussian density and V is the second moment matrix. The assumption holds if $\frac{\text{Condition Number of } V}{\ln T} \to 0$.

Assumption (Mixing)

$(X_t^{\infty})_{t\in\mathbb{Z}}$ is strictly stationary and absolutely regular with mixing coefficients $(\beta_m)_{m\in\mathbb{N}}$ that satisfy $\beta_m \sim b^m$ for some $b \in (0,1)$, as $m \to \infty$.

- E.g. linear and GARCH type models.

Empirical Sparse Optimization: Statistical Guarantees

Theorem (Consistency, Rates and Asymptotic Distribution)

Under the previous assumptions, if Λ is closed and for large enough p it is also convex, and if $\frac{\ln p}{\sqrt{T}} \to 0$, then for fixed q,

$$M^{\mathrm{FS}}\left(\Lambda, \mathcal{L}_{p,q}, \mathbb{P}_T, q\left(\ln T\right)^2\right) \rightsquigarrow M\left(\Lambda_\infty, \mathcal{L}_{\infty,q}, \mathbb{P}\right).$$

If furthermore: i) (Condition CO) $z \to \mathbb{E}\left[\left(z - \sum_{i=0}^{\infty} \kappa_i X_{t,i}\right)_+\right]$ is strictly concave for any κ with $\mathrm{csupp}\left(\kappa\right) \leq q$, and ii) (Condition CM) for any $z > \inf Z$, $\mathbb{E}\left[\left(z - \sum_{i=0}^{\infty} \lambda_i X_{t,i}\right)_+\right]$ has a compact subset of minimizers over Λ_∞ and $\mathbb{E}\left[\left(z - \sum_{i=0}^{\infty} \kappa_i X_{t,i}\right)_+\right]$ has a compact set of minimizers of support $\leq q$, then

$$\sqrt{T}\left(M^{\mathrm{FS}}\left(\Lambda, \mathcal{L}_{p,q}, \mathbb{P}_T, q\left(\ln T\right)^2\right) - M\left(\Lambda_\infty, \mathcal{L}_{\infty,q}, \mathbb{P}\right)\right) \rightsquigarrow \sup \inf_{(z,\lambda,\kappa)\in\Gamma} \mathcal{G}\left(z, \lambda, \kappa\right);$$

$\mathcal{G}\left(z, \lambda, \kappa\right)$ zero mean Gaussian, $\Gamma := \arg\max_{z\in Z, \lambda\in\Lambda_\infty} \min_{\mathrm{csupp}(\kappa)\leq q} D\left(z, \lambda, \kappa, \mathbb{P}\right)$.

Empirical Sparse Optimization: Statistical Guarantees

- CO: under Normality, the condition holds if $\frac{\text{Condition Number of } V}{\ln T} \to 0$.

- CM: by Theorem 4.5 of Beer and Lucchetti (1991), compactness of the set of minimizers is a generic property in the sense of Baire category. It is expected to hold at least for a dense subset of Z.

- $\kappa_{z,T}$ is the solution of $\inf_{\text{csupp}(\kappa)\leq q} \frac{1}{T}\sum_{t=0}^{T}\left(z_t - \sum_{i=0}^{\infty}\kappa_i X_{t,i}\right)_+$ over $\mathcal{L}_{p,q}$. $\Gamma^\star$ is the subset of Γ that contains the triplets at which some accumulation point of $\kappa_{z,T}$ appears. For $0 < b_T \leq T$, consider the subsamples $(X_j)_{j=t,\dots t+b_T-1}$ for all $t = 1, 2, \dots, T - b_T + 1$. For $\alpha \in (0,1)$, $q_{T,B_T}(1-\alpha)$ is the $1-\alpha$ quantile of the empirical distribution of $\left(\sqrt{b_T}\left(\sup_{Z\times\Lambda_p} D\left(z, \kappa_{z,T}, \lambda, \mathbb{P}_{t,b_T}\right) - M^{\text{FS}}\left(\Lambda, \mathcal{L}_{p,q}, \mathbb{P}_T, q\left(\ln T\right)^2\right)\right)\right)_{t=1,\dots,T-b_T+1}$, where $\mathbb{P}_{t,b_T}$ is the empirical distribution of $(X_j)_{j=t,\dots,t+b_T-1}$ and we use the same $\kappa_{z,T}$ across subsamples. Hence, we get a **fast subsampling** method.

Empirical Sparse Optimization: Statistical Guarantees

Proposition (Subsampling Confidence Intervals)

Suppose that (Condition ND) for the given q, $\Gamma^{\star}$ contains at least one non trivial triplet. Under the premises of the previous theorem, if $b_T \to \infty$, $\frac{b_T}{T} \to 0$ and $\alpha < \frac{1}{2}$, then we get the conservative result:

$$\limsup_{T\to\infty} \mathbb{P}\left[M\left(\Lambda_\infty, \mathcal{L}_{\infty,q}, \mathbb{P}\right) \in \left(M^{\mathsf{FS}}\left(\Lambda, \mathcal{L}_{p,q}, \mathbb{P}_T, q\left(\ln T\right)^2\right) \mp q_{T,B_T}\left(1-\alpha\right)\right)\right] \geq 1-\alpha.$$

If moreover there exists a unique q-sparse element of Λ that appears in every triple in $\Gamma^{\star}$, then we get the exact result:

$$\lim_{T\to\infty} \mathbb{P}\left[M\left(\Lambda_\infty, \mathcal{L}_{\infty,q}, \mathbb{P}\right) \in \left[M^{\mathsf{FS}}\left(\Lambda, \mathcal{L}_{p,q}, \mathbb{P}_T, q\left(\ln T\right)^2\right) \mp q_{T,B_T}\left(1-\alpha\right)\right]\right] = 1-\alpha.$$

Empirical Sparse Optimization: Statistical Guarantees

- Under linear independence, ND holds whenever every q-sparse efficient element is matched by an efficient element of appropriately large support compared to the maximum desired level of q for the underlying analysis.

- The evaluation of the quantile by subsampling has small computational burden since we avoid the costly sparse optimization w.r.t. κ inside each subsample. We only need to compute once $\kappa_{z,T}$ on the full sample and keep it fixed across subsamples.

- Usually, Z is approximated by some finite discretization and optimization w.r.t. λ is performed via linearization of the SD conditions and the use of LP methods. Then, the computational cost of sparse optimization is avoided and the asymptotic results above hold as long as the discretized set converges to a dense subset of Z.

Numerical Implementation: Optimization criterion

For $q < p$, we consider the following empirical optimization problem:

$$\sup_{u \in \mathcal{U}} \inf_{\mathcal{L}_{p,q}} \left(\sup_{\boldsymbol{\lambda} \in \tilde{}} \mathbb{E}_{\mathbb{P}_T} \left[u \left(X^{\mathrm{T}} \boldsymbol{\lambda} \right) \right] - \sup_{\boldsymbol{\kappa} \in \mathrm{K}} \mathbb{E}_{\mathbb{P}_T} \left[u \left(X^{\mathrm{T}} \boldsymbol{\kappa} \right) \right] \right). \qquad (1)$$

We approximate every element of $\mathcal{U}$ with arbitrary prescribed accuracy using a finite set of increasing and concave piecewise-linear functions.

Numerical Implementation: Greedy algorithm

We start with an empty set, and then we gradually increase the number of assets w in K (we add 1 asset at a time) until we find a set $K \subset \Lambda$ with $K \underset{\text{SSD}}{\succeq} \Lambda$.

In each iteration, we search for the asset that increases (1) the most.

Steps of Greedy algorithm:

For $w = 1$ to q:

1. If $w = 1$, we search for the single asset that increases the value of (1) the most, thus $\mathcal{L}_{p,q}$ is a singleton.
2. For $1 < w < q$, we solve (1) for each additional asset, and we keep the subset K with dimension w, that maximizes (1) the most.
3. If we find a spanning set K inside the collection of all possible subsets of Λ with dimension w, then the algorithm stops.
4. Else, if $w = q$, we end up with a sparse portfolio set K that "comes as close as possible" to SSD spanning its high dimensional universe of portfolios, and we evaluate the utility loss. If non zero, we continue for $w > q$.

Monte Carlo

Table 1: Monte Carlo Experiment 1. The experiment is based on a problem with N=49, 500 normally distributed assets and T=300, 500, 1000 time series observations. We compute the average number and standard deviation of assets selected. We also measure the variability of the loss, by computing the average and the standard error of the loss.

Sample size T	300	500	1000
Case 1: N=49, q=13			
Assets selected:			
Average number	11.45	12.04	12.54
St Deviation	1.18	1.12	1.13
Variability of the Loss:			
Average Loss	0.0002	0.0002	0.0001
Standard Error	0	0	0
Case 2: N=500, q=45			
Assets selected:			
Average number	42.3	42.85	43.34
St Deviation	1.68	1.57	1.54
Variability of the Loss:			
Average Loss	0.0001	0.0001	0.0001
Standard Error	0	0	0

Monte Carlo

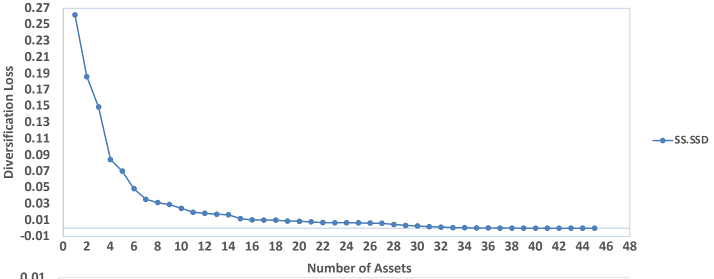
Table 2: Monte Carlo Experiment 2. The experiment is based on a problem with N=50 normally distributed assets and T=300, 500, 1000 time series observations. We compute the average number and standard deviation of assets selected. We also measure the variability of the loss, by computing the average loss and the standard error of the loss.

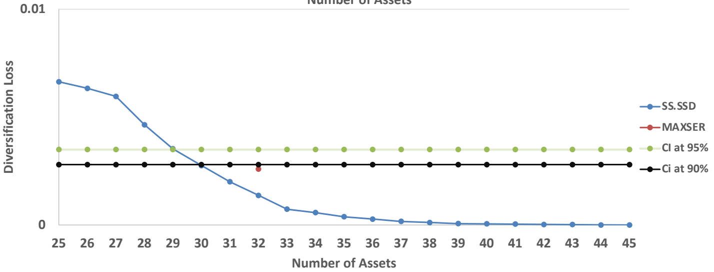
Sample size T	300	500	1000
Case 1: q=5			
Assets selected:			
Average number	5	5	5
St Deviation	0.0	0.0	0.0
Variability of the Loss:			
Average Loss	0.01	0.009	0.008
Standard Error	0.0003	0.0003	0.0002
Case 2: q=10			
Assets selected:			
Average number	10	10	10
St Deviation	0.0	0.0	0.0
Variability of the Loss:			
Average Loss	0.0	0.0	0.0
Standard Error	0.0	0.0	0.0

Empirical application

- In the empirical application, we analyze large data sets of equity returns to study whether sparse SSD holds or not.
- We investigate the performance of our strategy based on the S&P 500 index constituents, and we compare the results with the sparse mean-variance efficient portfolios of Ao, Li, and Zheng (2019).
- We consider the period from January 1981 to December 2020, a total of 480 monthly return observations.

The under-diversification loss in the SSD setting





Empirical application: In-Sample Analysis

Table 3: In-sample performance: risk and performance measures

	MAXSER	SS-SSD	$1/N$
Measures			
Average return	0.0126	0.0129	0.0133
Standard Deviation	0.0314	0.0331	0.0458
Sharpe ratio	0.4013	0.3904	0.2899
Skewness	-0.2122	-0.1986	-0.2689
Kurtosis	1.2521	1.7595	2.9690
Value at Risk	-0.0430	-0.0396	-0.0615
Expected Shortfall	-0.0651	-0.0617	-0.0959

Entries report the risk and performance measures (Sharpe ratio, Skewness, Kurtosis, VaR, ES) for the MAXSER, the SS-SSD optimal portfolios as well as the $1/N$ portfolio. The data cover the period from January, 1980 to December, 2020.

Empirical application: In-Sample

Table 4: In-sample analysis: Average S&P500 Industry weights

	MAXSER	SS-SSD
Weights		
Capital Goods	4.50%	3.43%
Consumer Services	8.39%	6.57%
Financial	4.73%	3.60%
Consumer Staples	0.0%	3.24%
Food	3.21%	2.70%
Health care	7.43%	8.31%
Household	5.58%	4.37%
IMedia	4.58%	4.34%
Pharm	6.89%	5.69%
Retailing	17.21%	19.43%
Software	14.51%	16.21%
Technology	11.45%	12.79%
Transportation	5.62%	4.81%

Entries report the average Industry weights of the MAXSER and the SS-SSD portfolios in the major Industries of the S&P500 Index.

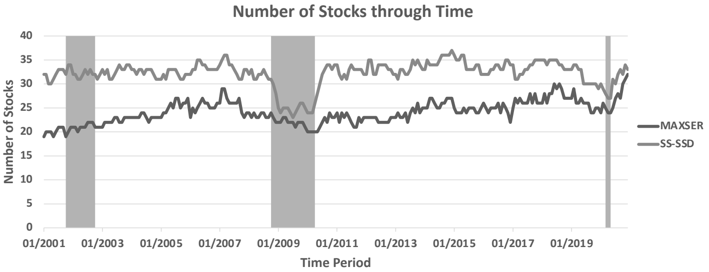
Empirical application: Out-Of-Sample

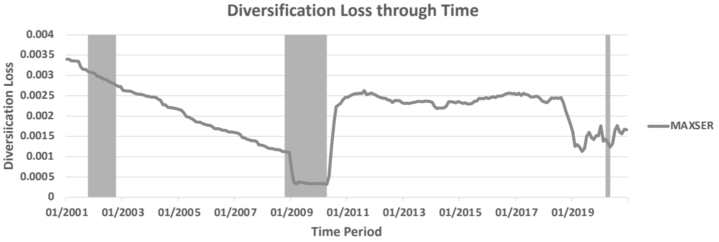
- We conduct out-of-sample backtesting experiments and we evaluate the optimal SS-SSD portfolios achieving a zero diversification loss in a rolling-window scheme.

- Each month, portfolios are constructed using the monthly returns during the prior 240 months. The clock is advanced and the realized returns of the optimal portfolios are determined from the actual returns of the various assets. The same procedure is then repeated for the next time period and the ex post realized returns over the period from 01/2001 to 12/2020 (240 months) are computed.

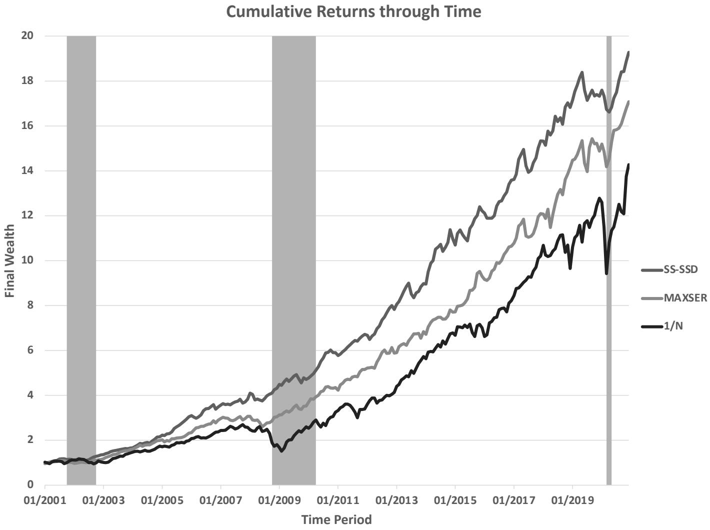
- We again compare the performance of the optimal SS-SSD portfolios with that of the MAXSER portfolios of Ao, Li, and Zheng (2019).

Empirical application: Out-Of-Sample





Empirical application: Out-Of-Sample



Empirical application: Out-Of-Sample

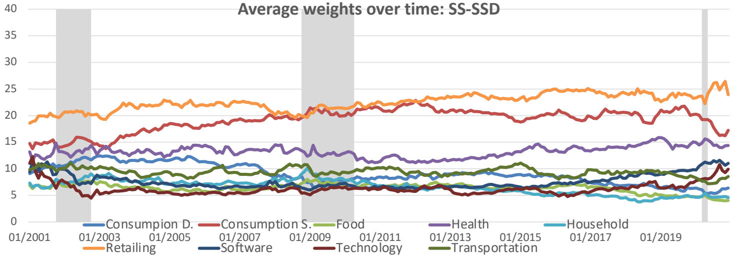
Table 5: Out-of-sample performance: risk and performance measures

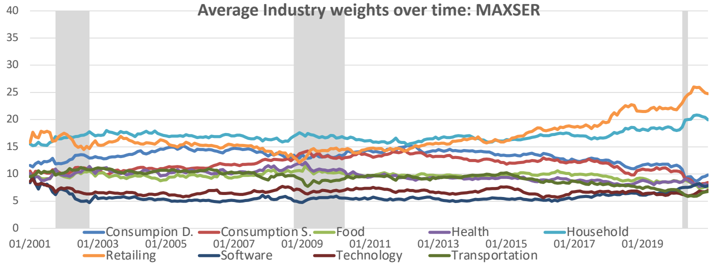
	MAXSER	SS-SSD	$1/N$
Measures			
Average return	0.0122	0.0127	0.0121
Standard Deviation	0.0258	0.0239	0.0450
Sharpe ratio	0.4056	0.4571	0.2313
Downside Sharpe Ratio	0.8614	1.1188	0.9311
Value at Risk	-0.0403	-0.0295	-0.0744
Expected Shortfall	-0.0532	-0.0476	-0.1004
UP ratio	1.0864	1.2014	0.7704
Portfolio Turnover	8.477%	8.835%	0.0
Return Loss	0.087%		0.156%
Opportunity Cost			
Exponential Utility			
ARA=2	0.073%		0.126%
ARA=4	0.081%		0.139%
ARA=6	0.092%		0.152%

Empirical application: Out-Of-Sample

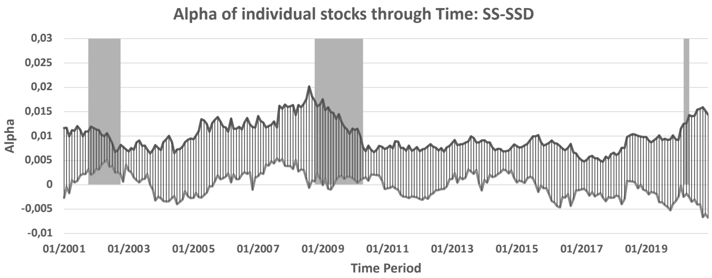
- We analyze the composition of the SS-SSD and the MAXSER portfolios through time.
- We observe that both portfolios are well diversified and invest in almost the same Industries, with different overall weights.
- The optimal SS-SSD portfolio invests mainly in 9 industry sectors with a larger weighting on small size, high book-to-market, and momentum stocks from the S&P 500 index.
- We additionally estimate the Alpha and Beta coefficients of the individual stocks of these portfolios during the out-of-sample period. For the estimations, the previous 5 years of individual monthly returns have been used (60 monthly returns).

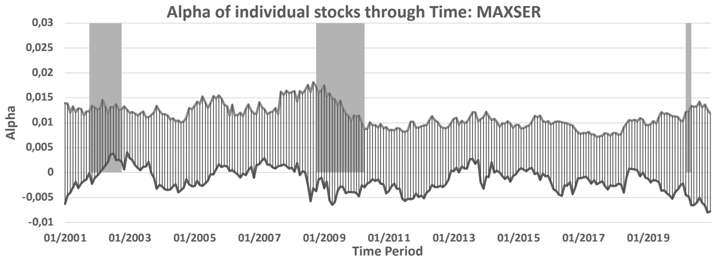
Empirical application: Out-Of-Sample



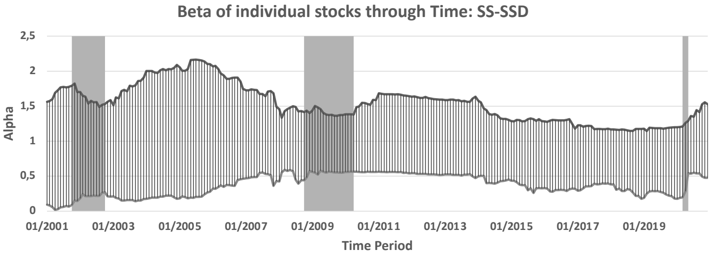


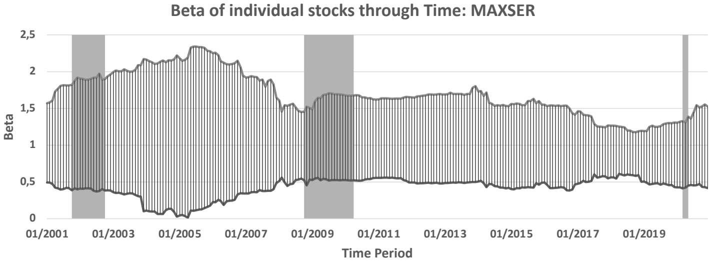
Empirical application: Out-Of-Sample





Empirical application: Out-Of-Sample





Empirical application: Out-Of-Sample

Finally, we investigate which factors explain the returns of the active investors with SSD preferences.

We use CAPM, the Fama-French 6-factor model (2016), the q-factor model of Hou, Xue and Zhang, (2015), the M4 factor model of Stambaugh and Yuan, (2017), the Barillas and Shanken 6-factor model (2018) and the 3-factor model of Daniel, Hirshleifer, and Sun (2020).

We consider linear regression models of the following form:

$$R_{p,t} - R_{f,t} = a_i + \sum_i b_i R_{i,t} + e_{i,t},$$

where $R_{p,t}$ is the return of either the MAXSER or SS-SSD optimal portfolio at period t, and, $R_{i,t}$ is the return on the i_{th} factor.

Empirical application: Out-Of-Sample

- The results indicate that none of the factor models could fully explain the performance of the two strategies.
- The intercept a_i is statistically different from zero in all cases.
- We also observe that the only factors that are significant for the MAXSER returns are the FIN factor of the 3-factor model of Daniel, Hirshleifer, and Sun (2020), and the MGMT1 factor of the Stambaugh and Yuan(2017), four-factor model.
- On the other hand, there is no statistically significant factor that explains the returns of the SS-SSD portfolios even if we face a defensive tilt: MKT: <1, SMB: < 0, HML: > 0.
- The results indicate that perhaps other factors drive the performance of the these portfolios.

Conclusions

- Our new methodology shows that we can often limit ourselves to a subset of a large investment opportunity set without sacrifying expected utility because of under-diversification.
- It also reveals that a sparse mean-variance portfolio selection yields under-diversification w.r.t. an optimal sparse spanning portfolio.
- The paper focuses on second-order stochastic dominance but could be modified to accommodate higher-order stochastic dominance.
- We could then check whether the empirical findings extend in such settings as well.