Empirical properties of US equities

Alex Shkolnik

Joint winter seminar

January, 2025.

Department of Statistics & Applied Probability
University of California, Santa Barbara

The data

Wharton Research Data Services (WRDS).

https://wrds-www.wharton.upenn.edu

- *Access through UCLA/UCB libraries.*
- *1974-2024 time-series of US equity returns (+ market caps).*
- *The frequency is daily (return).*
- *There is missing data (e.g., acquisition, merger, bankruptcy).*
- *Typical set is 3000 stocks with the largest market cap.*
- *The constituents of this group changes over time.*

We observe a vector $r_j \in \mathbb{R}^p$ on date j.

- *p is the number of stocks/securities/assets.*
- $r_j = (r_{1j}, \ldots, r_{pj})^\top$
- *We observe r_j on n dates.*
- *$(p \times n)$ data matrix $R = (r_{ij})_{1 \le i \le p, 1 \le j \le n}$.*

r_{ij} is the return of stock i on date j.

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pn} \end{pmatrix}.$$

Observing R we construct a portfolio $w \in \mathbb{R}^p$.

$$w = (w_1, \ldots, w_p)$$

- w_i *is the investment is stock* i.
- $\sum_{i=1}^{p} w_i = 1$ *(w.l.o.g)*
- $w_i \geq 0$ *(long position) and* $w_i < 0$ *(short position).*

Construct/invest stock portfolio $w^{(1)}$ based on observing $R^{(1)}$.

$$R^{(1)} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pn} \end{pmatrix}.$$

After m time units a return M_1 on portfolio $w^{(1)}$ is realized.

We also have m new data points and a new data set $R^{(2)}$ based on which we can make the next stock portfolio $w^{(2)}$.

$$R^{(2)} = \begin{pmatrix} r_{1(m+1)} & r_{1(m+2)} & \cdots & r_{1(n+m)} \\ r_{2(m+1)} & r_{2(m+2)} & \cdots & r_{2(n+m)} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ r_{p(m+1)} & r_{p(m+2)} & \cdots & r_{p(n+m)} \end{pmatrix}.$$

Return M_2 on portfolio $w^{(2)}$ is realized, etc ...

Main projects

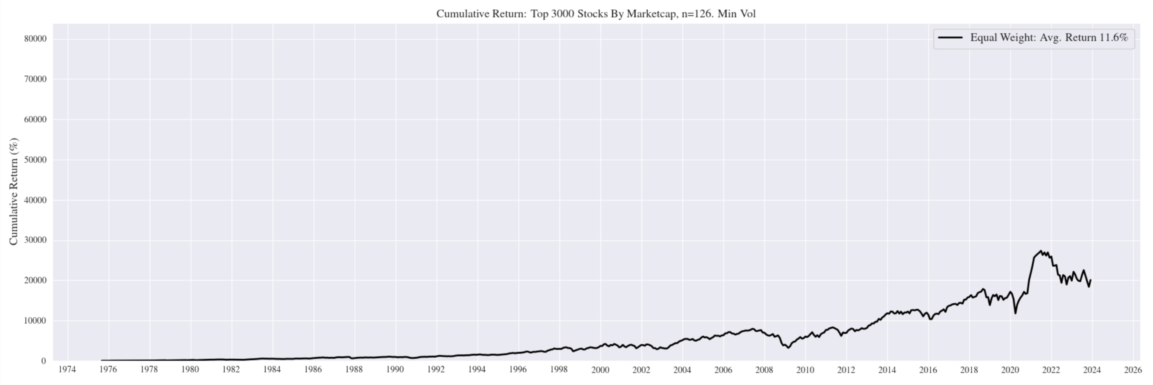
Empirical studies.

- *Compare the portfolio return time series* $M_1, M_2, \ldots,$ *for many different methods of constructing* $w^{(j)}$.
- *Study the spectral properties of the data matrices* $R^{(j)}$ *to identify structure and build better asset return models.*

Benchmarks portfolios.

- *Equally weighted portfolio, i.e.* $w_i = 1/p$.
- *Market cap weighted portfolio, i.e.,* $w_i = cap_i / \sum_{i=1}^{p} cap_i$
- Mean-variance optimized portfolios.

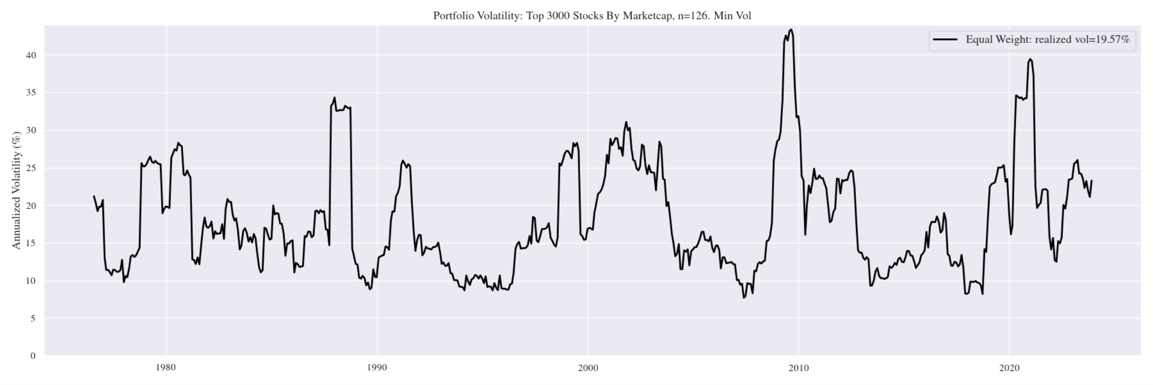
Cumulative returns (%) to the equally weighted portfolio (1974–1924; scaled for comparison with another method).



WRDS data on top 3000 stocks by market cap.

Investments are monthly.

Return volatility (%) to the equally weighted portfolio (1974–1924; scaled for comparison with another method).

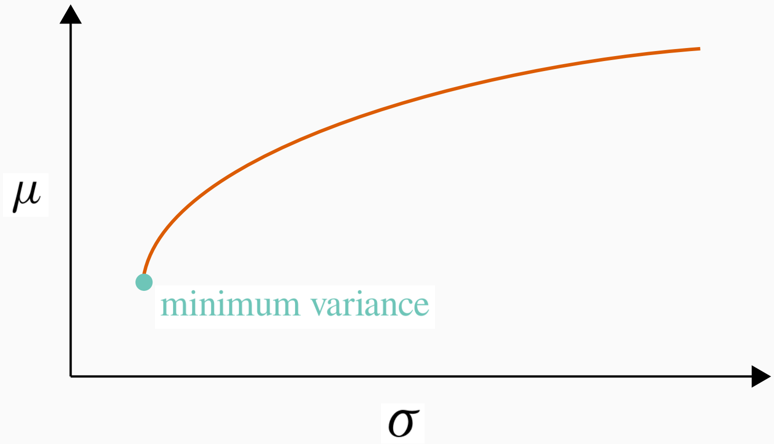


WRDS data on top 3000 stocks by market cap.

Investments are monthly.

Mean-variance optimization

Since Markowitz (1952), quantitative investors have constructed portfolios with mean-variance optimization.



- *A simple quadratic program given a covariance matrix Σ.*
- *We can make two curves (in-sample and out-of-sample).*

The Markowitz quadratic program.

$$\min_{w \in \mathbb{R}^p} \langle w, \Sigma w \rangle$$

subject to:

$$\langle m, w \rangle \geq \alpha \,,$$

$$\langle \mathrm{e}, w \rangle = 1 \,.$$

(every $w_i \geq 0$

$\ldots$ etc.)

- $\langle x, y \rangle = \sum_{i=1}^{p} x_i \, y_i$.
- Σ *is a* $(p \times p)$ *covariance matrix of stock returns.*
- $m \in \mathbb{R}^p$ *is the estimate of expected returns.*
- $\alpha \in \mathbb{R}$ *is the target portfolio return.*
- $\mathrm{e} = (1, \ldots, 1) \in \mathbb{R}^p$

The Markowitz quadratic program.

$$\min_{w \in \mathbb{R}^p} \langle w, \Sigma w \rangle$$

$$\text{subject to:}$$

$$\langle m, w \rangle \geq \alpha\,,$$

$$\langle \mathrm{e}, w \rangle = 1\,.$$

$$(\text{every } w_i \geq 0$$

$$\ldots \text{ etc.})$$

The Markowitz optimization enigma entails the observation that "*mean-variance optimizers are, in a fundamental sense, estimation-error maximizers*" – Michaud (1989).

- *The estimation error sits in m and Σ.*

Methods/metrics/parameters

We will use the following portfolio metrics.

- *Portfolio volatility.*
- *Portfolio concentration.*
- *Portfolio return.*

We compare the following 3 methods.

- *Principal component analysis (PCA).*[1]
- *James-Stein-Markowitz (JSM) corrected PCA.*[2]
- *Ledoit-Wolf constant correlation shrinkage (LW).*

[1] We will take a diagonal residual for PCA.

[2] We use the diagonal residual from PCA as weights for JS

Recipe for the covariance model

1. With $\bar{r}$ as above, let $\bar{R}$ be the $(p \times n)$ matrix with $\bar{r}$ in every column, to center the data, i.e,

$$Y = R - \bar{R}. \tag{8}$$

2. For the centered sample covariance matrix $S = YY^\top/n$, write its spectral decomposition as

$$S = \sum_{(\jmath^2, h)} \jmath^2 h h^\top = HH^\top + N \tag{9}$$

   where the sum is over all eigenvalue/eigenvector pairs $(\jmath^2, h)$ of S, H is a $p \times k$ matrix with every column of the form $\jmath h$ sourced from the k largest eigenvalues $\jmath^2$, and $N = S - HH^\top$.

3. The specific risk estimate Δ in (5) sets all the off-diagonal elements of N to zero, i.e.,

$$\Delta = \mathbf{diag}(N). \tag{10}$$

4. The PCA covariance matrix is $\Sigma_{\text{PCA}} = HH^\top + \Delta$.

Recipe for the covariance model

1. For any estimate Δ (e.g., (10)), centering and weighting the data, we set

$$Y = \Delta^{-1/2}\big(R - \bar{R}\big) \tag{15}$$

where $\Delta^{-1/2}$ is diagonal with $\Delta_{ii}^{-1/2} = 1/\sqrt{\Delta_{ii}}$ and $\bar{R}$ is the matrix in (8).

2. Recompute H following (9) but from the re-weighted sample covariance S that uses (15). Set,

$$\bar{H} = \Delta^{1/2} H\,. \tag{16}$$

3. The JSM estimator of the weighted eigenvectors $\bar{H}$ computes a $(k \times k)$-matrix valued shrinkage parameter,

$$C = I - \nu^2 J^{-1}\,, \qquad J = (\bar{H} - M)^\top \Delta^{-1} (\bar{H} - M)\,, \tag{17}$$

where ν^2 is the variance of the noise and $M \neq \bar{H}$ is a $(p \times k)$-matrix shrinkage target.[a]

4. The JSM estimator is analogous to (12) but with matrix valued C and M.

$$H_{\text{JSM}} = \bar{H} C + M(I - C) \tag{18}$$

5. The variance ν^2 is computed per (13) but with N from the reweighted sample covariance S.
6. A shrinkage target M analogous to (14) uses a $(p \times 2)$-matrix $A = \big(\mu_{\text{JS}} \;\; \mathrm{e}\big)$ as

$$M = A(A^\top \Delta^{-1} A)^{-1} A^\top \Delta^{-1} \bar{H}\,. \tag{19}$$

7. The basic JSM covariance model is $\Sigma_{\text{JSM}} = H_{\text{JSM}} H_{\text{JSM}}^\top + \Delta$.

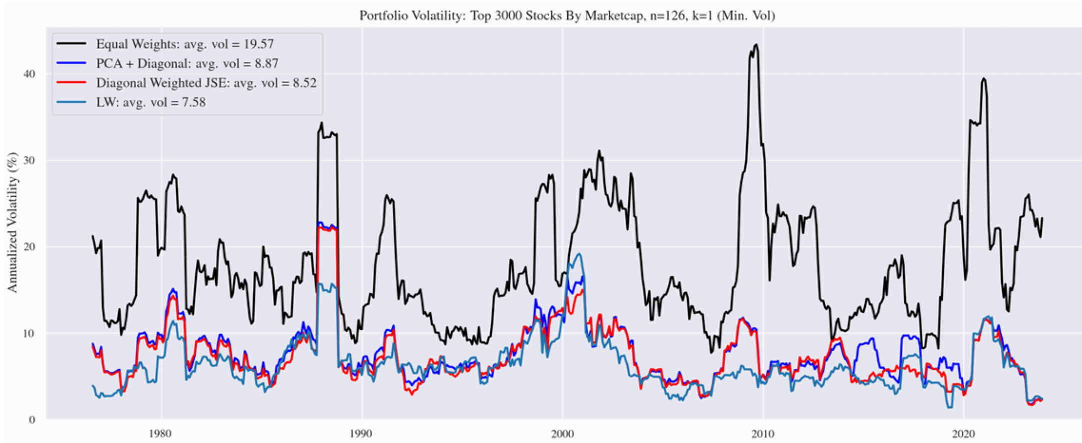
[a] Here, $\neq$ is in the sense that the columns spaces of the two matrices are not identical.

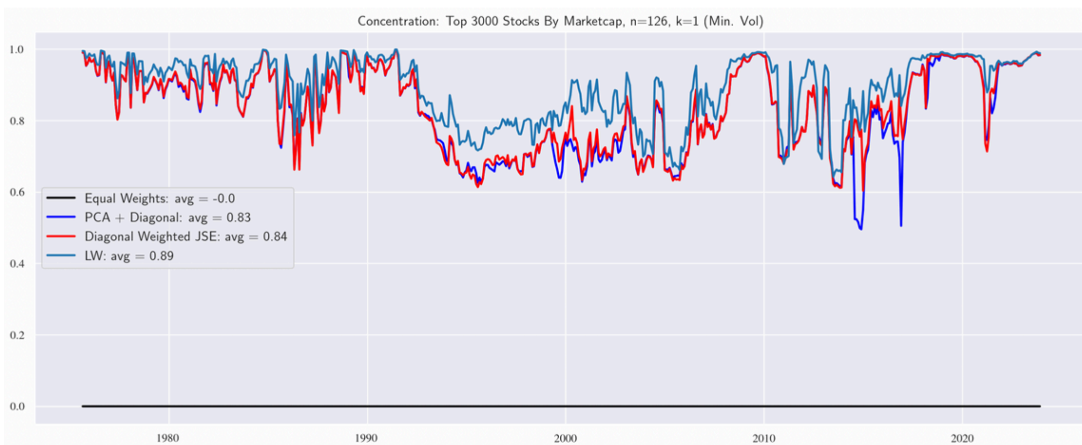
We can make choices for the following parameters.

- *The number of factors* k.
- *The number of return observations* n.

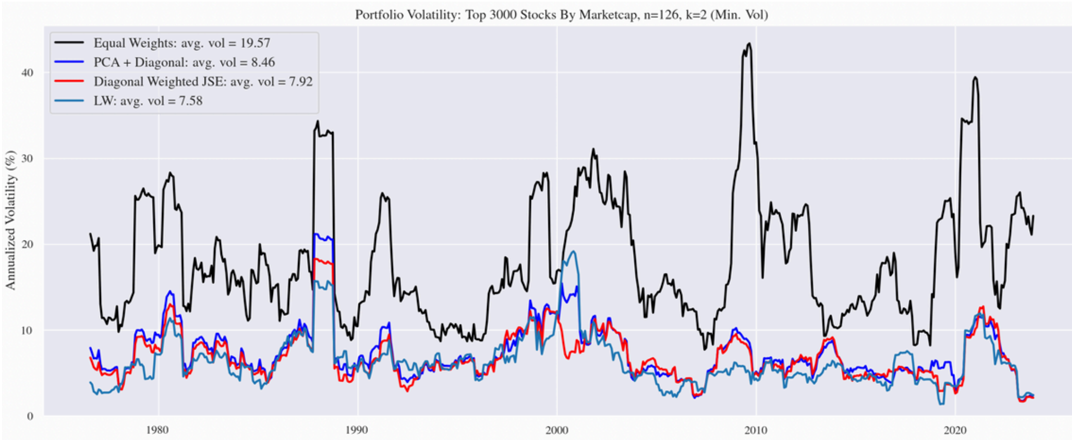
Minimum volatility

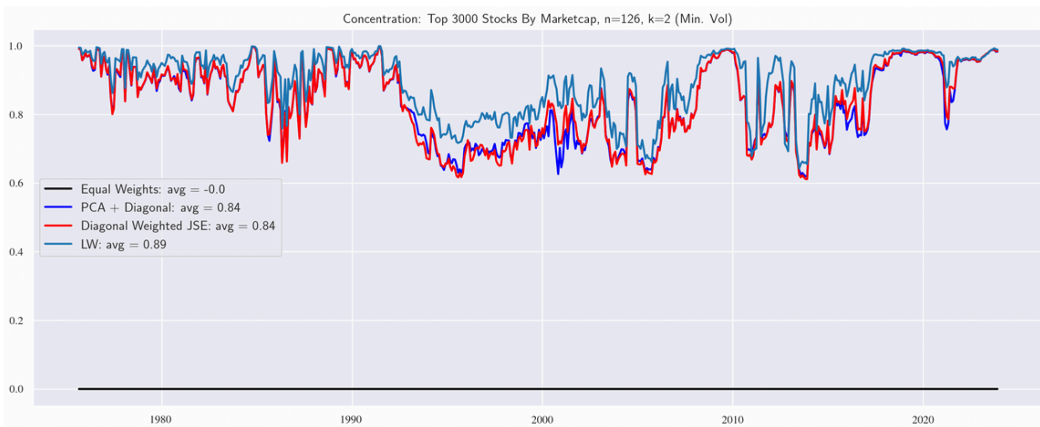
Analysis of k=1



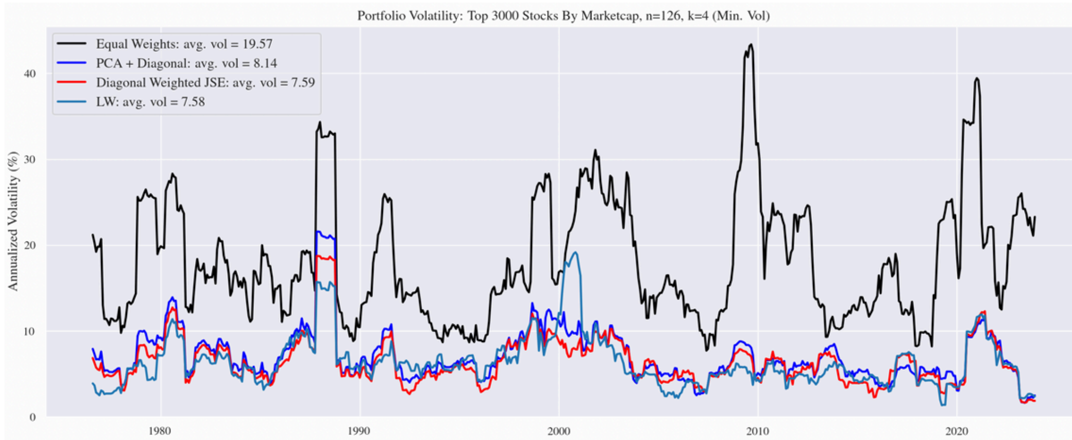


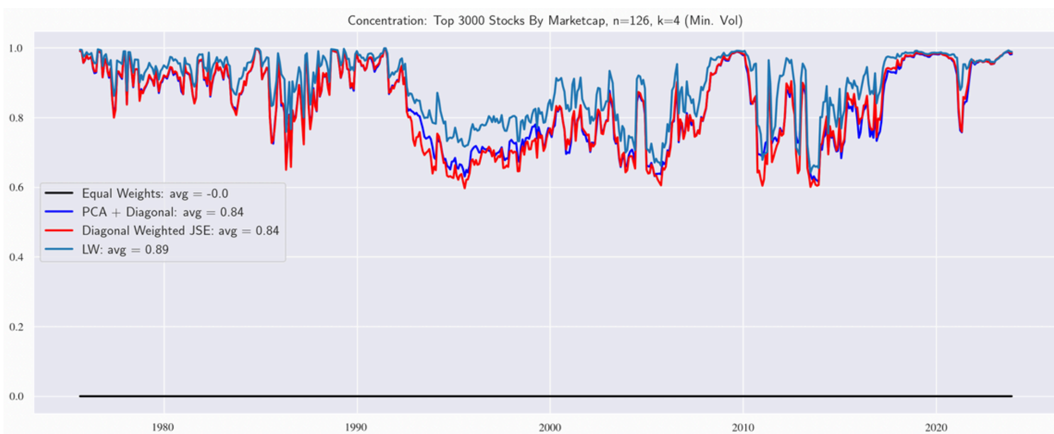
Analysis of k=2



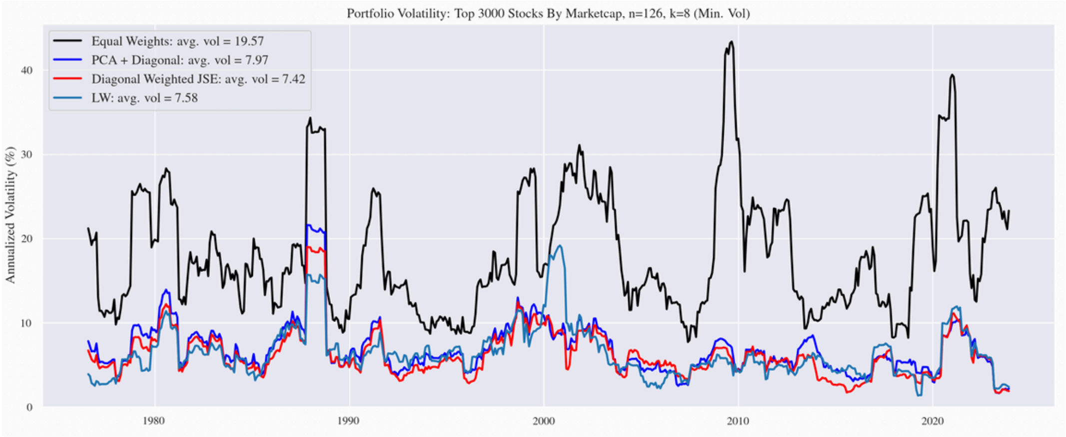


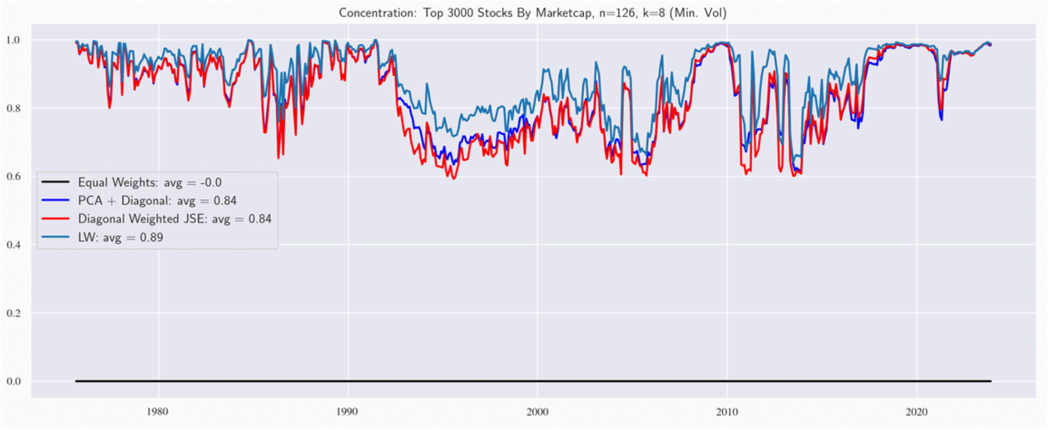
Analysis of k=4





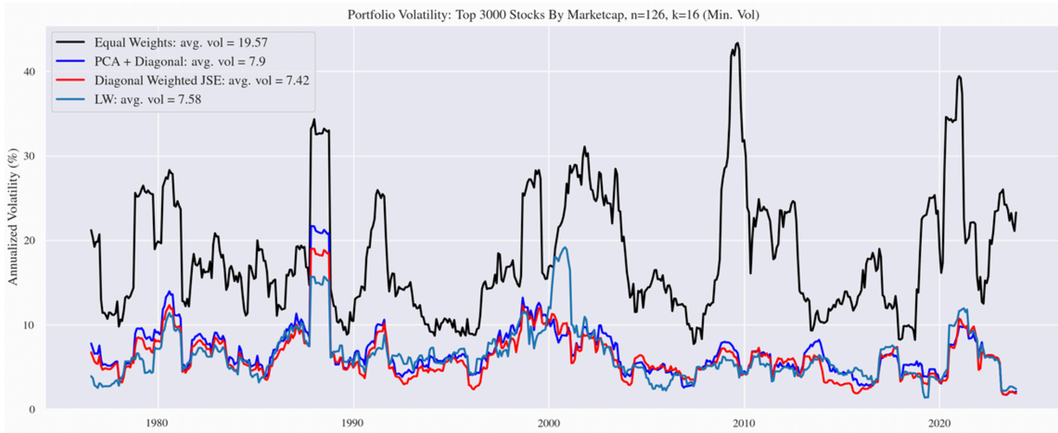
Analysis of k=8

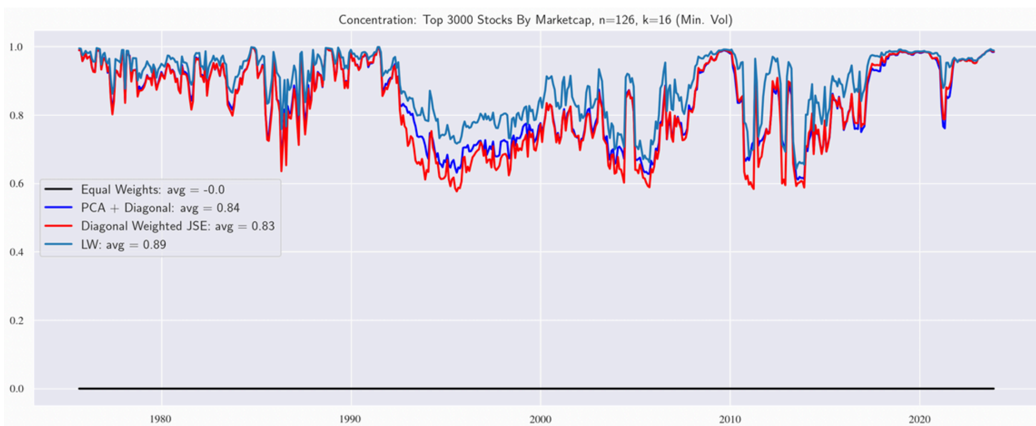




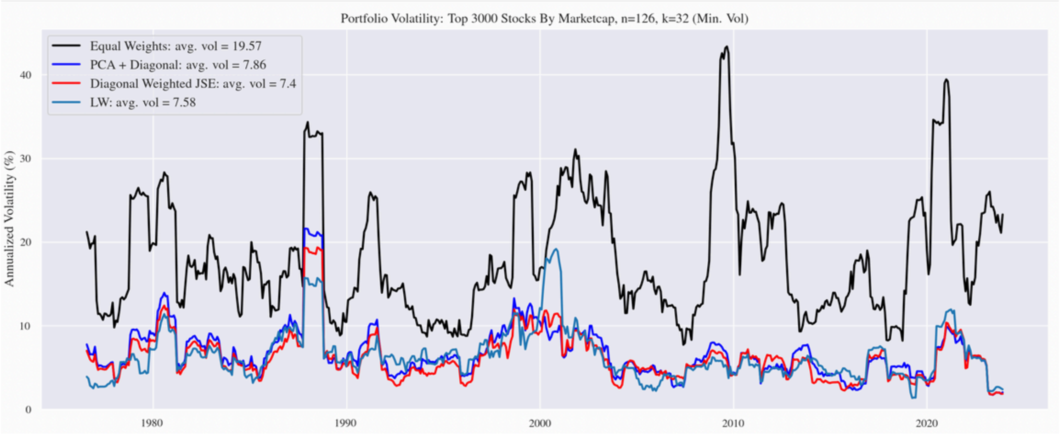
Cumulative Return: Top 3000 Stocks By Marketcap, n=126, k=8 (Min. Vol)

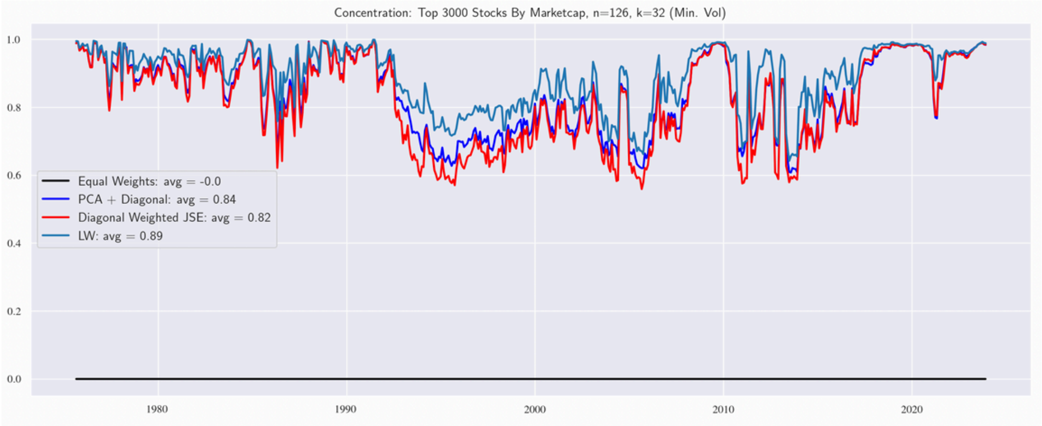
Analysis of k=16





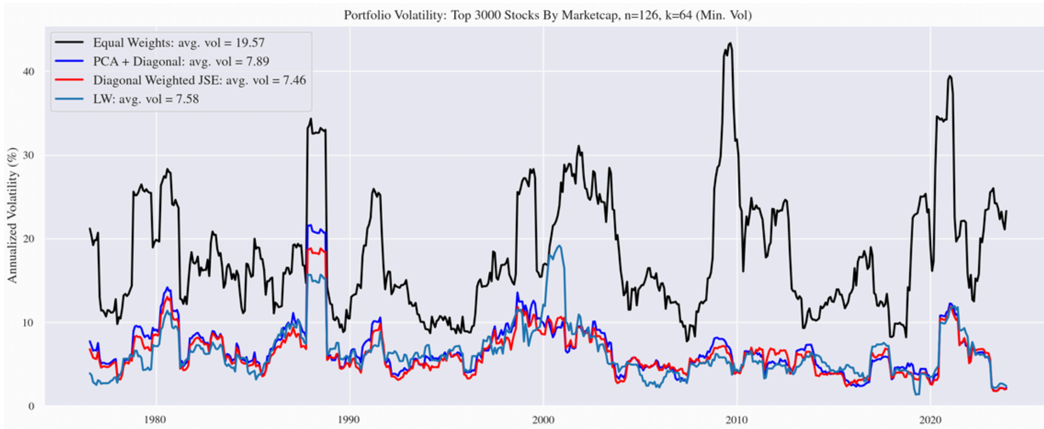
Analysis of k=32

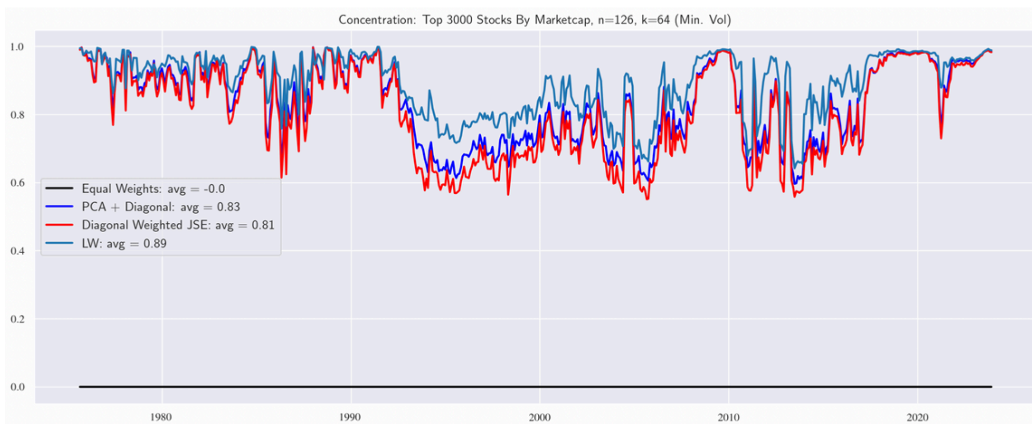


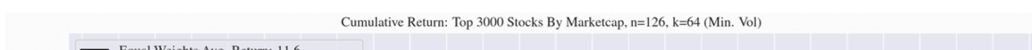


Cumulative Return: Top 3000 Stocks By Marketcap, n=126, k=32 (Min. Vol)

Analysis of k=64







References

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

Michaud, R. O. (1989), 'The markowitz optimization enigma: Is 'optimized'optimal?', *Financial analysts journal* 45(1), 31–42.