Chapter 3

Sparse Portfolio Allocation

This chapter touches some practical aspects of portfolio allocation and risk assessment from a large pool of financial assets (e.g. stocks)

★ How to assess the risk of a large portfolio?

★ How to estimate a large covariance matrix?

★ Do the errors in estimating large covariance matrix accumulate?
How do they impact on portfolio optimization?

How to select a subset of portfolio to invest?

How to track a portfolio?

Is a given portfolio efficient?

### 3.1 Risks of Large Portfolios

**Risk**: For a given portfolio with allocation \( \mathbf{w} \) on the \( p \) risky assets with returns \( \mathbf{R} \), its risk is

\[
\sqrt{\text{var}(\mathbf{w}^T \mathbf{R})} = \sqrt{R(\mathbf{w})},
\]

where

\[
R(\mathbf{w}) = \mathbf{w}^T \Sigma \mathbf{w}, \quad \Sigma = \text{var}(\mathbf{R}).
\]
**Perceived Risk**: From historical data, one obtains an estimate $\hat{\Sigma}$ and compute $\sqrt{\hat{R}(w)}$, where

$$\hat{R}(w) = w^T \hat{\Sigma} w.$$ 

**How much the difference? Do errors accumulate?**

**Gross exposure**: $c = \|w\|_1 = |w_1| + \cdots + |w_p|$.

**Relation with short position**: Let

$$w^+ = \sum_{w_i \geq 0} w_i, \quad w^- = \sum_{w_i < 0} |w_i|$$

be total long and short positions. Then, $c = w^+ + w^-$. Since the total portfolio is 100%, we have $w^+ - w^- = 1$. Thus, $w^- = (c - 1)/2$, and the minimum gross exposure is $c \geq 1$. 
**No-short portfolios**: All no-short portfolios with $c = 1$, the *minimum* gross exposure.

**Example 3.1**: Consider the following five portfolios of size 4:

- $w_1 = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix}$, $w_2 = \begin{pmatrix} 0 \\ 0.5 \\ 0.25 \\ 0.25 \end{pmatrix}$, $w_3 = \begin{pmatrix} -0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$, $w_4 = \begin{pmatrix} -0.5 \\ 0.5 \\ 1.5 \end{pmatrix}$, $w_5 = \begin{pmatrix} -8 \\ 25 \\ -7 \\ -9 \end{pmatrix}$.

- Both $w_1$ and $w_2$ are no-short portfolios with $c = 1$. Portfolio $w_1$ is more diversified and robust to individual influence.

- $c = \|w_3\|_1 = 2$, and hence the short position is $w^- = (c - 1)/2 = 0.5$ and long position is $w^+ = (c + 1)/2 = 1.5$. It is less stable, since the $L_1$ norm is large (similar to SD).
• $c = \|w_4\|_1 = 3$ and hence the short position is $(c - 1)/2 = 1$. It is even more unstable.

• $c = \|w_5\|_1 = 49$, the least stable portfolio. The short position $w^- = (c - 1)/2 = 24$ and long position is $w^+ = (c + 1)/2 = 25$.

**Unbiasedness**: If $\hat{\Sigma}$ is an unbiased estimate of $\Sigma$, so is $\hat{R}(w)$.

**Proof**: This follows simply from

$$E\hat{R}(w) = w^T(E\hat{\Sigma})w = R(w).$$

However estimation error can be large. The following gives a bound.

**Risk Approximation**: For any portfolio of size $p$ with allocation vector $w$,

$$|R(w) - \hat{R}(w)| \leq e_{\text{max}}c^2,$$
where $e_{\text{max}} = \max_{i,j} |\hat{\sigma}_{ij} - \sigma_{ij}|$, denoted also by $|\hat{\Sigma} - \Sigma|_\infty$.

**Proof:** The risk difference is

$$\left| \sum_{i,j} (\hat{\sigma}_{i,j} - \sigma_{i,j}) w_i w_j \right| \leq \sum_{i,j} |\hat{\sigma}_{i,j} - \sigma_{i,j}| |w_i||w_j|$$

$$\leq e_{\text{max}} \sum_{i,j} |w_i||w_j|$$

$$= e_{\text{max}} \left( \sum_{i=1}^{p} |w_i| \right)^2.$$

**Remarks:**

★ The result holds for any portfolio size even $p \gg T$, and any matrices $\hat{\Sigma}$ (even not semi-definite matrix)

★ Little noise accumulation effect when $c$ is modest ($\leq 2$ or 3), since
$e_{\text{max}}$ increases with $p$ only in logarithmic order (Fan, Zhang and Yu, 2012).

**Example 3.2:** Consider 4 assets all w/ mean returns 10% and volatility 30% per annum. They are equally correlated, with $\rho = 0.5$.

- Data sampled at daily frequency for 3 months;
- Use sample covariance matrix.
- Number of Simulation = 1000.

Then, portfolio variances are estimated *unbiasedly*.

**Results** Risks of five portfolios in Example 3.1 and their mean absolute errors $|R(w)^{1/2} - \hat{R}(w)^{1/2}|$ are summarized in Table 1.1. Unbiasedness of the estimates can also be seen.
Figure 3.1: A realization of simulated daily returns of the four assets over a three-month period and the distribution of the maximum componentwise estimation errors of the sample covariance over 1000 simulations.
Table 3.1: True risks and average estimation Errors for five portfolios.

<table>
<thead>
<tr>
<th>Portfolios</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
<th>$w_4$</th>
<th>$w_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Risks</td>
<td>23.72</td>
<td>24.88</td>
<td>30.00</td>
<td>42.43</td>
<td>607.45</td>
</tr>
<tr>
<td>Average Estimated Risks</td>
<td>23.73</td>
<td>24.86</td>
<td>29.95</td>
<td>42.33</td>
<td>606.04</td>
</tr>
<tr>
<td>Average Estimation Error</td>
<td>1.73</td>
<td>1.75</td>
<td>2.14</td>
<td>2.91</td>
<td>43.00</td>
</tr>
</tbody>
</table>

- True risks increase with the gross exposures and so do the estimation errors.

**Portfolio choice**: When $w$ is chosen by data (e.g. optimization), the resulting allocation is denoted by $\hat{w}$.

- The true risk of the portfolio $\hat{w}$ is $R(\hat{w})$.
- The perceived risk is $\hat{R}(\hat{w})$.
- It is not an unbiased estimate of $R(\hat{w})$, and can be very different.
Example 3.3. Focus on 1000 stocks with least missing data from Russell 3000 in 2003-2007.

- Select randomly 600 stocks, reducing selection biases.
- Use sample co-variance matrix from past 2 years, $T = 504$ (degenerate)

**Perceived Minimum** variance portfolio: $\hat{w}_{opt} = \arg\min \ w^T \hat{\Sigma}^{-1} \hat{\mu} \ 1^T$.

Since the rank of $\hat{\Sigma} \leq T - 1$,

- the perceived risk $\hat{R}(\hat{w}_{opt}) = 0$,
- the actual risk $R(\hat{w}_{opt})$ is far from zero.
- it behaves like a random portfolio, having risk around 20%-30%,
  the average of the market risks.
**Portfolio optimization**: \( \hat{w}_{\text{opt},c} = \arg\min_{w} w^T 1_{1=1}, \|w\|_1 \leq c \hat{R}(w) \), a specific example of portfolio choice. **Noise accumulation** can be seen as follows.

**Computing actual risk**: Optimize portfolio monthly and record daily returns. Compute the SD of portfolio returns. See Fig. 3.2.

**Results**:

★ Optimal no-short sale portfolio. Risk 9.5%

★ Perceived optimal portfolio. Risk 23%.

★ Risk decreases and then increases steadily with gross exposure \( c \).

★ Estimators of large covariance matrix plays also a role.
Figure 3.2: The actual risks of the minimum variance portfolios with gross-exposure constraint $c$, using the 600 randomly selected assets from the 1000 stocks with least missing data in Russel 3000 in the period 2003-2007. The covariance matrices are estimated by the sample covariance, RiskMetrics, and Fama-French 3-factor model.

★ Explain why wrong constraint helps (Jagannathan and Ma, 03).
3.2 Portfolio Allocation with Gross-Exposure Constraints

**Key Reference**: Fan, Yu and Zhang (2012).

**Markowitz’s mean-variance analysis**:

\[
\min_w w^T \Sigma w, \quad \text{s.t. } w^T \mathbf{1} = 1 \text{ and } w^T \mu = r_0.
\]

**Solution**: \( w_{opt} = c_1 \Sigma^{-1} \mu + c_2 \Sigma^{-1} \mathbf{1} \), where \( c_1 \) and \( c_2 \) are constant depending on quantities such as \( \mu^T \Sigma^{-1} \mu^T \), \( \mu^T \Sigma^{-1} \mathbf{1} \), and \( \mathbf{1}^T \Sigma^{-1} \mathbf{1} \).

★ Cornerstone of modern finance; frequently used in practice.

★ Too sensitive on input vectors and their estimation errors.

★ Minimum variance portfolio depends on the smallest eigenvalue, but is dominated by stochastic noise.
Can result in extreme short positions: $\|w_{opt}\|_1$ can be very large.

( Green and Holdfield, 1992 ).

More severe for large portfolio.

The vector you get $\hat{w}_{opt}$ can be very different from the vector that you want $w_{opt}$.

3.2.1 Portfolio selection with gross-exposure constraint

Utility optimization: Let $c$ be the gross-exposure

$$\max_w E[U(w^T R)]$$

s.t. $w^T 1 = 1, \|w\|_1 \leq c, A^T w = \alpha$. 
■ Equivalent to total short position \( w^- \leq (c - 1)/2 \).

■ Equality constraints: ⭐ on return \((A = \mu)\), ⭐ sector allocations, ⭐ factor-neutral loadings.

■ Portfolio selection: solution is usually sparse.

■ Can be replaced by any risk measures (Artzner et al, 1999)

**Relation with Markowitz**: \( c = \infty \) is Markowitz’s problem.

- When \( c = 1 \), no short-sale allowed, most conservative portfolios.

- Creating a continuous path from most conservative to most aggressive portfolios. See Fig 3.2.
Normal returns and exponential utility: the expected utility is equivalent to

\[ M(\mu, \Sigma) = w^T \mu - A w^T \Sigma w / 2. \]

Utility Approximation:

\[ |M(\hat{\mu}, \hat{\Sigma}) - M(\mu, \Sigma)| \leq \|\hat{\mu} - \mu\|_\infty \|w\|_1 + Ae_{\text{max}} \|w\|_2^2 / 2, \]

- No short-sale portfolios give the tight (accurate) bound;
- Usually not diversified enough, nor efficient enough;
- Facilitate covariance matrix estimation: estimating each element as accurately as possible. Facilitate estimation based on different observation frequencies and models.
Focus on Risks:

- Theory and methods readily extendable to utility case.
- Hard to estimate expected returns; replaced by factor or sector exposure constraints.

Actual and perceived risks: $R(w) = w^T \Sigma w$, $\hat{R}(w) = w^T \hat{\Sigma} w$.

Oracle and empirical allocation vectors:

$$w_{opt} = \text{argmin}_{\|w\|_1 \leq c} R(w), \quad \hat{w}_{opt} = \text{argmin}_{\|w\|_1 \leq c} \hat{R}(w).$$

Optimal Risks:

- **Oracle risk**: $\sqrt{R(w_{opt})}$, best possible actual risk. (unknown)
- **Perceived optimal risk**: $\sqrt{\hat{R}(\hat{w}_{opt})}$, often too small for large $c$. (computable)
Theorem 1: Let $e_{\text{max}} = |\hat{\Sigma} - \Sigma|_{\infty}$. Then, we have (Fan, Zhang, Zhang, 2012)

\[
|R(\hat{\mathbf{w}}_{\text{opt}}) - R(\mathbf{w}_{\text{opt}})| \leq 2e_{\text{max}} c^2
\]

\[
|R(\hat{\mathbf{w}}_{\text{opt}}) - \hat{R}(\hat{\mathbf{w}}_{\text{opt}})| \leq e_{\text{max}} c^2
\]

\[
|R(\mathbf{w}_{\text{opt}}) - \hat{R}(\hat{\mathbf{w}}_{\text{opt}})| \leq e_{\text{max}} c^2.
\]

The bounds are nonasymptotic, tight when $c$ and $e_{\text{max}}$ are small.

Example 3.10. Simulate daily returns of $p$ assets over 252 days from a three-factor model with parameters calibrated to the market. Three different risks are computed. Sample covariance matrix is used.

★ Handle any portfolio size even $p \gg T$.

★ Discrepancies increase with $c$. 

■ $\sqrt{R(\hat{\mathbf{w}}_{\text{opt}})}$, actual risk of a selected portfolio. (unknown)
Figure 3.3: The risks of theoretically optimal portfolios, and the actual risks of the empirically optimal portfolios, and the perceived risks of the empirically optimal portfolios under gross-exposure constraints are plotted against the gross-exposure parameter $c$.

★ No noise accumulation effect when $c$ is modest ($\leq 2$ or $3$).

★ Explain why wrong constraint helps (Jagannathan and Ma, 03).

**Theorem 2**: If for a sufficiently large $x$,

$$\max_{i,j} P\{b_T|\sigma_{ij} - \hat{\sigma}_{ij}| > x\} < \exp(-Cx^{1/a}),$$

for some positive constants $a$ and $C$ and rate $b_T$, then

$$|\Sigma - \hat{\Sigma}|_{\infty} = O_P\left(\frac{(\log p)^a}{b_T}\right).$$
Impact of dimensionality is limited.

**Theorem 3:** Suppose that $\|R_t\|_\infty < B$, and the mixing coefficient $\alpha(q) = O(\exp(-Cq^{1/b}))$, $b < 2a - 1$ and \( \log p = o(n^{1/(2b+1)}) \), then conclusion in Theorem 2 holds.

Results holds for unbounded random variables and other dependency (Fan, Zhang and Yu, 2008).

### 3.3 Portfolio Selection and Tracking

#### 3.3.1 Relation with Regression

**Regression problem:** Letting $Y = R_p$ and $X_j = R_p - R_j$,

$$\text{var}(\mathbf{w}^T \mathbf{R}) = \min_b E(\mathbf{w}^T \mathbf{R} - b)^2$$

$$= \min_b E(Y - w_1X_1 - \cdots - w_{p-1}X_{p-1} - b)^2$$
portfolio optimization is equivalent to regression, minimizing with respect to \( b, w_1, \cdots, w_{p-1} \).

**Gross exposure:** \( \|w\|_1 = \|w^*\|_1 + |1 - 1^T w^*| \leq c \),

where \( w^* = (w_1, \cdots, w_{p-1})^T \).

- **not equivalent** to \( \|w^*\|_1 \leq d \) for some \( d \).

- \( d = 0 \) picks \( R_p \), but \( c = 1 \) picks multiple stocks.

**Typical usage:** \( Y = w_0^T R \) a given portfolio with \( w_0^T 1 = 1 \). Then

\[
\text{var}(w^T R) = \min_b E(Y - w^T X - b)^2
\]

where \( X = (w_0^T R) 1 - R \). The problem becomes the constrained least-squares problem: minimizing wrt \( w \) subject to \( w^T 1 = 1 \).

The constraint \( w^T 1 = 1 \) suggests that one needs to drop one of
the $X$-variables with $Y$. Drop the one least correlated with $Y$.

■ A useful example: $Y$ optimal no-short portfolio (sparse).

3.3.2 Portfolio selection and tracking

**Portfolio Tracking:** If $Y$ is the target portfolio and \( \{X_j\}_{j=1}^p \) are candidate assets, the penalized least-squares can be regarded as finding a portfolio to minimize the expected tracking error.

★ Let $w(d)$ be the solution. The selected portfolio allocates $w(d)$ on \( \{X_j\}_{j=1}^p \) and the rest on cash.

★ As $d$ increases, the number of selected portfolio increases and the
tracking errors decrease.

**Portfolio improvement**: Given a portfolio \( Y \) constructed from the financial assets \( \{X_j\}_{j=1}^p \), is it efficient? Can it be improved?

- Run LASSO with variables \( \{X_j\}_{j=1}^p \) and \( Y \). PLS can be interpreted as modifying weights to improve the performance of \( Y \).
- If the portfolio \( Y \) is efficient, the reg coefficients are close to zero.

**Empirical risk path** \( \hat{R}(d) \) helps decision making with ★limited number of stocks ★limited exposure.

**Example 3.11**: Take \( Y = \text{CRSP} \) and \( X = 10 \) industrial portfolios, downloaded from Data Library of Kenneth French: e.g.

\[ 1 = \text{“Consumer Non-durables”}, \quad 9 = \text{“Utilities”}, \]
2 = “Consumer Durables”, 5 = “Hi-tech equipment”.

Suppose that our goal is to improve the risk of CRSP by using those 10 industrial portfolios on 1/8/05 (hard, of course).

- Use one-year daily returns before the date.
- Run LASSO for various $d$.
- Each portfolio (corresponding to a given $d$) is held one year.

**Results:**

- Total at any vertical line in the left panel of Fig. 3.4 is one.
- Ex-post risk is computed based on the daily returns from Jan. 8, 2005 to Jan. 8, 2006.
Figure 3.4: Illustration of risk improvement by using the penalized least-squares. (a) The solution paths as a function of \(d\). The numbers on the top shows the number of assets recruited for a given \(d\). (b) The ex-ante and ex-post risks (annualized volatility) of the selected portfolios.

★ Ex-ante (perceived risk) and Ex-post have the same decreasing pattern until 6 stocks are added.

★ Using Ex-ante as guide, one would have selected 4-6 portfolios and the risk is improved.
3.4 Empirical Applications

We now illustrate the methods by using two data sets:

♦ 600 stocks randomly selected 1000 stocks with least missing data in Russell 3000 during 2003-2007 (5 years);

♦ 100 Fama-Portfolios (less sensitive to allocation vectors).

**Evaluation**: Portfolio optimized monthly, and returns recorded daily.

3.4.1 Fame-French 100 portfolios

**Data**: 100 portfolios formed by size and BE ratios \((10 \times 10)\) from 1998–2007 (10 years), downloaded Data Library of Kenneth French.
**Covariance matrix**: Estimated by sample covariance matrix, factor model using last **twelve** months daily data, and RiskMetrics ($\lambda = .97$).

**Testing Period**: 9 years since 1999 and evaluation on a rolling basis.

**Results**:
- Optimal exposure $c$ is around 2. Sharpe ratios peak there too.
- Optimal no-short portfolios have smaller risk than equally weighted portfolio, as expected.
- RiskMetrics performs the best due to shorter time-domain smoothing and hence less biased (recalling portfolios are less volatile)
- Factor model performs the worst due to modeling biases (can not
Figure 3.5: Characteristics of invested portfolios as a function of exposure constraints $c$ from the Fama-French 100 industrial portfolios formed by the size and book to market from Jan 2, 1998 to December 31, 2007. (a) Annualized risk of portfolios. (b) Sharpe ratio of portfolios. (c) Max weight of allocations. (d) Annualized return of portfolios.
Table 3.2: Ex-post risks and other characteristics of constrained optimal portfolios for 100 Fama-French portfolios

<table>
<thead>
<tr>
<th>Methods</th>
<th>Mean</th>
<th>Std</th>
<th>Sharpe-R</th>
<th>Max-W</th>
<th>Min-W</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Covariance Matrix Estimator</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No short (c = 1)</td>
<td>19.51</td>
<td>10.14</td>
<td>1.60</td>
<td>0.27</td>
<td>-0.00</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Exact (c = 1.5)</td>
<td>21.04</td>
<td>8.41</td>
<td>2.11</td>
<td>0.25</td>
<td>-0.07</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Exact (c = 2)</td>
<td>20.55</td>
<td>7.56</td>
<td>2.28</td>
<td>0.24</td>
<td>-0.09</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Exact (c = 3)</td>
<td>18.26</td>
<td>7.13</td>
<td>2.09</td>
<td>0.24</td>
<td>-0.11</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>GMV</td>
<td>17.55</td>
<td>7.82</td>
<td>1.82</td>
<td>0.66</td>
<td>-0.32</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td><strong>Unmanaged Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal-W</td>
<td>10.86</td>
<td>16.33</td>
<td>0.46</td>
<td>0.01</td>
<td>0.01</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>CRSP</td>
<td>8.2</td>
<td>17.9</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

expect factor models to hold approximately due to the constructions, but it is most stable.

♦ The factor model provides most diversified portfolio, since the max-
Table 3.3: Ex-post risks and other characteristics of constrained optimal portfolios for 100 Fama-French portfolios

<table>
<thead>
<tr>
<th>Methods</th>
<th>Mean</th>
<th>Std</th>
<th>Sharpe-R</th>
<th>Max-W</th>
<th>Min-W</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor-Based Covariance Matrix Estimator</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No short ($c = 1$)</td>
<td>20.40</td>
<td>10.19</td>
<td>1.67</td>
<td>0.21</td>
<td>-0.00</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Exact ($c = 1.5$)</td>
<td>22.05</td>
<td>8.56</td>
<td>2.19</td>
<td>0.19</td>
<td>-0.05</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Exact ($c = 2$)</td>
<td>21.11</td>
<td>7.96</td>
<td>2.23</td>
<td>0.18</td>
<td>-0.05</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Exact ($c = 3$)</td>
<td>19.95</td>
<td>7.77</td>
<td>2.14</td>
<td>0.17</td>
<td>-0.05</td>
<td>35</td>
<td>41</td>
</tr>
<tr>
<td>GMV</td>
<td>19.90</td>
<td>7.93</td>
<td>2.09</td>
<td>0.43</td>
<td>-0.14</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td><strong>Covariance Estimation from Risk Metrics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No short ($c = 1$)</td>
<td>15.45</td>
<td>9.27</td>
<td>1.31</td>
<td>0.30</td>
<td>-0.00</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Exact ($c = 1.5$)</td>
<td>15.96</td>
<td>7.81</td>
<td>1.61</td>
<td>0.29</td>
<td>-0.07</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Exact ($c = 2$)</td>
<td>14.99</td>
<td>7.38</td>
<td>1.58</td>
<td>0.29</td>
<td>-0.10</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Exact ($c = 3$)</td>
<td>14.03</td>
<td>7.34</td>
<td>1.46</td>
<td>0.29</td>
<td>-0.13</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>GMV Portfolio</td>
<td>13.99</td>
<td>9.47</td>
<td>1.12</td>
<td>0.78</td>
<td>-0.54</td>
<td>53</td>
<td>47</td>
</tr>
</tbody>
</table>
imum allocations on individual portfolios have the smallest weights.

♦ GMV exists when covariance matrix is estimated by the sample covariance matrix or factor model, but it is outperformed by the regularized method.
3.4.2 Russell 3000 stocks

**Covariance matrix**: Estimated by sample covariance, and factor model used last **twenty-four** months daily data, and RiskMetrics ($\lambda = .97$). This is due to higher volatilities of stocks than portfolios.

**Results**: Shown in Fig. 3.2. Tables 1.5–1.7 show additional details.

- Optimal no-short portfolio is not diversified enough, picking 53 stocks on average.
- Its risk improved by the choice with $c = 2$.
- GMV is not unique, since $T = 518$ and $p = 600$. It is a random
portfolio. \( c = 8 \) provides a proxy.

- Factor-model provides most stable estimates.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Std</th>
<th>Max-W</th>
<th>Min-W</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>No short</td>
<td>9.28</td>
<td>0.14</td>
<td>0.00</td>
<td>53</td>
<td>0</td>
</tr>
<tr>
<td>( c = 2 )</td>
<td>8.20</td>
<td>0.11</td>
<td>-0.06</td>
<td>123</td>
<td>67</td>
</tr>
<tr>
<td>( c = 3 )</td>
<td>8.43</td>
<td>0.09</td>
<td>-0.07</td>
<td>169</td>
<td>117</td>
</tr>
<tr>
<td>( c = 4 )</td>
<td>8.94</td>
<td>0.10</td>
<td>-0.08</td>
<td>201</td>
<td>154</td>
</tr>
<tr>
<td>( c = 5 )</td>
<td>9.66</td>
<td>0.12</td>
<td>-0.10</td>
<td>225</td>
<td>181</td>
</tr>
<tr>
<td>( c = 6 )</td>
<td>10.51</td>
<td>0.13</td>
<td>-0.10</td>
<td>242</td>
<td>201</td>
</tr>
<tr>
<td>( c = 7 )</td>
<td>11.34</td>
<td>0.14</td>
<td>-0.11</td>
<td>255</td>
<td>219</td>
</tr>
<tr>
<td>( c = 8 )</td>
<td>12.20</td>
<td>0.17</td>
<td>-0.12</td>
<td>267</td>
<td>235</td>
</tr>
</tbody>
</table>
Table 3.5: Ex-post risks and other characteristics of constrained optimal portfolios

<table>
<thead>
<tr>
<th>Factor-Based Covariance Matrix Estimator</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No short</td>
<td>9.08</td>
<td>0.12</td>
<td>0.00</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td>c = 2</td>
<td>8.31</td>
<td>0.06</td>
<td>-0.03</td>
<td>188</td>
<td>120</td>
</tr>
<tr>
<td>c = 3</td>
<td>8.65</td>
<td>0.05</td>
<td>-0.03</td>
<td>314</td>
<td>272</td>
</tr>
<tr>
<td>c = 4</td>
<td>8.66</td>
<td>0.05</td>
<td>-0.03</td>
<td>315</td>
<td>273</td>
</tr>
<tr>
<td>c = 5</td>
<td>8.66</td>
<td>0.05</td>
<td>-0.03</td>
<td>315</td>
<td>273</td>
</tr>
<tr>
<td>c = 6</td>
<td>8.66</td>
<td>0.05</td>
<td>-0.03</td>
<td>315</td>
<td>273</td>
</tr>
<tr>
<td>c = 7</td>
<td>8.66</td>
<td>0.05</td>
<td>-0.03</td>
<td>315</td>
<td>273</td>
</tr>
<tr>
<td>c = 8</td>
<td>8.66</td>
<td>0.05</td>
<td>-0.03</td>
<td>315</td>
<td>273</td>
</tr>
</tbody>
</table>
Table 3.6: Ex-post risks and other characteristics of constrained optimal portfolios

<table>
<thead>
<tr>
<th>Methods</th>
<th>Std</th>
<th>Max-W</th>
<th>Min-W</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Covariance Estimation from Risk Metrics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No short</td>
<td>9.78</td>
<td>0.40</td>
<td>0.00</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>c = 2</td>
<td>8.44</td>
<td>0.12</td>
<td>-0.06</td>
<td>119</td>
<td>63</td>
</tr>
<tr>
<td>c = 3</td>
<td>8.95</td>
<td>0.11</td>
<td>-0.07</td>
<td>191</td>
<td>133</td>
</tr>
<tr>
<td>c = 4</td>
<td>9.43</td>
<td>0.12</td>
<td>-0.09</td>
<td>246</td>
<td>192</td>
</tr>
<tr>
<td>c = 5</td>
<td>10.04</td>
<td>0.12</td>
<td>-0.10</td>
<td>279</td>
<td>233</td>
</tr>
<tr>
<td>c = 6</td>
<td>10.53</td>
<td>0.12</td>
<td>-0.11</td>
<td>300</td>
<td>258</td>
</tr>
<tr>
<td>c = 7</td>
<td>10.92</td>
<td>0.13</td>
<td>-0.11</td>
<td>311</td>
<td>272</td>
</tr>
<tr>
<td><strong>c = 8</strong></td>
<td><strong>11.06</strong></td>
<td><strong>0.13</strong></td>
<td><strong>-0.10</strong></td>
<td><strong>315</strong></td>
<td><strong>277</strong></td>
</tr>
</tbody>
</table>