Enough is Enough - Sufficient number of securities in an optimal portfolio

Bachelor's Thesis 15 hp
Department of Business Studies
Uppsala University
Spring Semester of 2016
Date of Submission: 2016-05-30

Iliam Barkino
Marcus Rivera Öman

Supervisor: Adri de Ridder
Abstract

This empirical study has shown that optimal portfolios need approximately 10 securities to diversify away the unsystematic risk. This challenges previous studies of randomly chosen portfolios which states that at least 30 securities are needed. The result of this study sheds light upon the difference in risk diversification between random portfolios and optimal portfolios and is a valuable contribution for investors.

The study suggests that a major part of the unsystematic risk in a portfolio can be diversified away with fewer securities by using portfolio optimization. Individual investors especially, who usually have portfolios consisting of few securities, benefit from these results. There are today multiple user-friendly software applications that can perform the computations of portfolio optimization without the user having to know the mathematics behind the program. Microsoft Excel’s solver function is an example of a well-used software for portfolio optimization. In this study however, MATLAB was used to perform all the optimizations.

The study was executed on data of 140 stocks on NASDAQ Stockholm during 2000-2014. Multiple optimizations were done with varying input in order to yield a result that only depended on the investigated variable, that is, how many different stocks that are needed in order to diversify away the unsystematic risk in a portfolio.

Keywords

Optimal Portfolio, Random Portfolio, Optimization, Modern Portfolio Theory, Variance, Covariance, Diversification, Relative Standard Deviation
Acknowledgments

We, the authors, would like to thank our family and friends for their love, for their continuous support throughout the project, and for their understanding of our absence. We would also like to extend a thank you to our appointed thesis opponents and classmates, Erik Ränkeskog and Anders Vidén, for their valuable feedback and suggestions on how to improve this thesis. Finally, we would like to give a special thanks to our supervisor, Adri de Ridder, for his experienced guidance, for his honest reflections on our ideas and progress, and for being unceasingly positive towards our work. All of your inputs have contributed considerably to this thesis.

Iliam Barkino  Marcus Rivera Öman

30th May 2016
Uppsala University
# Contents

1 Introduction 5
   1.1 Background 5
   1.2 Problem 6
   1.3 Purpose 6
   1.4 Disposition 7

2 Theory 7
   2.1 Modern Portfolio Theory 7
   2.2 Security analysis 8
   2.3 Portfolio optimization 9
   2.4 Portfolio selection 12
   2.5 Benefits of fewer securities 12

3 Method 13
   3.1 Assumptions 13
   3.2 Data 14
   3.3 Optimization Approach 15
   3.4 Compilation of Results 16

4 Results 17

5 Discussion 20
   5.1 Analysis of results 20
   5.2 Conclusion 22
   5.3 Applicability of study 23
   5.4 Suggestions for further research 23

A Interior-point-convex quadprog algorithm (MathWorks, 2016) 26
   A.1 Presolve/Postsolve 26
   A.2 Generate Initial Point 27
   A.3 Predictor-Corrector 28
   A.4 Multiple Corrections 30
A.5 Total Relative Error ........................................ 30

B MATLAB Code ............................................. 31
B.1 Main Code .................................................. 31
B.2 Date performance ......................................... 33
B.3 Portfolio Optimization ..................................... 34
B.4 Portfolio performance ..................................... 36
B.5 Variance-covariance matrix .............................. 36
B.6 The expected rate of return vector ...................... 36
1 Introduction

1.1 Background

Markowitz (1952) introduced the concept of modern portfolio theory (MPT) and he was, in 1990, awarded the Nobel Prize in Economics for his discoveries, together with William Sharpe and Merton Miller (Nobel Media, 1990). The theory focused on the properties of the portfolio rather than the individual assets themselves and stressed the importance of computing the covariance between the assets when evaluating the associated risk of a portfolio. Many investors have since then used the theory to compute the expected risk and the expected rate of return of a portfolio (Kolm, Tüttüncü & Fabozzi, 2013).

Individual investors as well as hedge funds use portfolio optimization when minimizing the risk in their portfolios (Ford, Housel & Mun, 2011). To minimize risk, it has always been conventional wisdom to diversify one’s portfolio - don’t put all your eggs in the same basket. According to definition, a well-diversified portfolio is

"a portfolio that includes a variety of securities so that the weight of any security is small. The risk of a well-diversified portfolio closely approximates the systematic risk of the overall market, and the unsystematic risk of each security has been diversified out of the portfolio.” (Financial Glossary, 2011).

However, a portfolio with a large number of securities may generate a small return in relation to the increased transaction costs. Extremely broad portfolios eventually perform more or less as the overall market. Evans and Archer presented in 1969 a well-cited study that stated that ten different securities would be enough to enjoy the benefits of a well-diversified portfolio. Statman (1987) explicitly opposed this and stated that no less than 30 different securities were needed. Several other studies have been executed to determine how many different securities are sufficient to create a well-diversified portfolio.

What all of these studies have in common is that they conducted their research on a randomly chosen set of securities with a varying number of included securities and repeated the process until the performances of the portfolios converged to a result. To perform the same kind of study but with MPT optimal portfolios rather than random ones has not been practically doable before the introduction of more powerful computers.
and improved optimization algorithms (Chan, Karceski & Lakonishok 1999). The technology of portfolio optimization is commonly used, as there are multiple user-friendly software that can perform the computations. The user of the software does not even have to know the mathematics behind the program, more than it can be used to minimize or maximize values, given certain constraints. Since it today is possible to study the relationship between risk and number of securities in optimal portfolios thanks to technological advancement, it is of high interest to investigate how such a study would stand in relation to previous studies.

Such a study could bring new insights in the discussion about how many securities are needed to create a well-diversified portfolio. The results of such a study would provide valuable guidance for investors using optimization as a tool for creating portfolios. Especially, individual investors could benefit from this study, as the typical individual investor rarely invests in a large number of different securities (Barber & Odeon 2013). If fewer different securities are needed to achieve a well-diversified portfolio, by using portfolio optimization and modern portfolio theory, such portfolios should prove very appealing for the individual investor.

1.2 Problem

The previous results by Evans and Archer (1968) and Statman (1987) suggest that portfolios of randomly chosen securities are well-diversified at approximately 10 and 30 different securities respectively. It is, however, not documented how many different securities an MPT optimal portfolio should have to be well-diversified, in spite the fact that MPT has been a well-established theory for over half a century.

1.3 Purpose

This study aims to investigate the relationship between the number of different securities and risk in optimal portfolios. Furthermore, this relationship will be compared with previous results of random portfolios. Hypothetically, the optimal portfolios converge to the systematic risk faster than random portfolios. This would advocate for portfolio optimization to investors, in particular individual investors.
1.4 Disposition

Section two will describe the theoretical background of this study. It mainly consists of the MPT and optimization mathematics, but will also present some benefits of portfolio optimization. The following section will describe the operationalization of the study, including all assumptions made, data management, optimization approach and result evaluation. Next, section four presents the obtained results with commentary.

Lastly, section five discusses the obtained results and compare them to previous studies. Furthermore, section five will discuss implementations of the results and suggest further studies in the area. The discussion will also include a shorter conclusion that summarizes the study.

2 Theory

2.1 Modern Portfolio Theory

In MPT, the portfolios are the objects of choice. The individual assets that are included in a portfolio are inputs, but they are not the objects of choice on which an investor should focus. The investor should focus on the best possible portfolio that can be created. The portfolio creating process is executed in three steps.

1. **Security analysis** is the step that evaluates the expected return, the variance and the co-variance for each and every investment candidate.

2. **Portfolio optimization** is the step that produces the optimum portfolio possibilities that can be constructed from the available investment candidates.

3. **Portfolio selection** is the step that selects the single most desirable portfolio from the available optimum portfolio possibilities.

Each of these steps are described more thoroughly below.
2.2 Security analysis

The purpose of the security analysis step is to evaluate each investment candidate. Primarily, it should forecast the expected rate of return and the associated risk. The rate of return for a period is computed as follows.

\[ r_{i,1} = \frac{P_{i,1} - P_{i,0}}{P_{i,0}}, \]  

(1)

where \( r_{i,1} \) denotes rate of return of an asset \( i \) for the holding period, \( P_{i,0} \) and \( P_{i,1} \) denote the value of the asset at the beginning of the holding period respectively at the end of the holding period. To compute the expected rate of return from a given historic period, an arithmetic average of each day rate of return is computed for the whole period. The expression is given by

\[ E(r_i) = \frac{1}{T} \sum_{t=1}^{T} r_{i,t}, \]  

(2)

where \( E(r_i) \) denotes the expected rate or return for asset \( i \), \( T \) denotes the number of historic days considered and \( r_{i,t} \) the rate of return for the specific day \( t \) computed with Eq. (1). All the computed expected rates of return are compiled into a vector which index \( i \) points out the expected rate of return for asset \( i \). This vector is referred to as the expected rate of return vector and it is presented in Eq. (3).

\[
\mathbf{r} = \begin{bmatrix}
E(r_1) \\
E(r_2) \\
\vdots \\
E(r_n)
\end{bmatrix}
\]  

(3)

The variance of an asset describes the uncertainty of the expected return is and it is the measurement of the associated risk of an asset. The variance of an asset is computed as

\[ Var(r_i) = \frac{1}{T} \sum_{t=1}^{T} (r_{i,t} - E(r_i))^2 \]  

(4)

where \( Var(r_i) \) denotes the variance for asset \( i \). Furthermore, the co-variance between each asset must be computed. The co-variance describes how similarly the assets behave.
A large positive co-variance between two assets suggests that those two assets behave alike, i.e. the value of asset $i$ increases and so does the value asset $j$. A large negative co-variance suggests that they have opposite behavior, i.e. the value of asset $i$ increases so the value of asset $j$ decreases. If the co-variance between two assets approach zero the assets behave independently from each other, i.e. the value of asset $i$ increases and there is no telling whether the value of asset $j$ increases or decreases. All the possible pairs of assets must be considered when the co-variance is computed. The co-variance between two assets is computed as

$$
\text{Cov}(r_i, r_j) = \frac{1}{T} \sum_{t=1}^{T} (r_{i,t} - E(r_i))(r_{j,t} - E(r_j)),
$$

(5)

where $\text{Cov}(r_i, r_j)$ denotes the co-variance between the rates of return of assets $i$ and asset $j$. In order to more easily use the risk data, the computed co-variances is compiled to a matrix which indices $i$ and $j$ point out the co-variance between asset $i$ and an asset $j$. For the cases where $i$ equals $j$ the input is the variance for that asset $i$. The yielded matrix is presented in Eq. (6).

$$
H = \begin{bmatrix}
\text{Var}(r_1) & \text{Cov}(r_1, r_2) & \ldots & \text{Cov}(r_1, r_n) \\
\text{Cov}(r_2, r_1) & \text{Var}(r_2) & \ldots & \text{Cov}(r_2, r_n) \\
\vdots & \vdots & \ddots & \vdots \\
\text{Cov}(r_n, r_1) & \text{Cov}(r_n, r_2) & \ldots & \text{Var}(r_n)
\end{bmatrix}
$$

(6)

This matrix is called the variance-covariance matrix and is, together with the expected rate of return vector, the primary information needed for performing the portfolio optimization.

### 2.3 Portfolio optimization

The purpose of this step is to produce a set of optimum portfolios from the information obtained in the previous step and some given well defined conditions. The information from the previous step is the expected rate of return vector $\mathbf{r}$ and the variance-covariance matrix $\mathbf{H}$. Define now a new vector, $\mathbf{w}$, which contains the relative weights of the investor’s total capital distributed among the different securities in the portfolio. It is given by
\[ w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{bmatrix}, \]  

(7)

where \( w_i \) denotes the relative weight of the investor’s total capital invested in security \( i \). By definition, the sum of each relative weight is equal to one, meaning the sum of all investments equals the investors total capital. The formula is given by

\[ \sum_{i=1}^{n} w_i = 1, \]  

(8)

where \( n \) denotes the number of available securities. This condition can be rewritten in matrix format for convenience,

\[ e^T \cdot w = 1, \]  

(9)

where

\[ e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \]  

(10)

is the vector consisting of ones equally long as \( w \), the upper letter \( T \) is the notation for transpose, and the mathematical operation performed is the dot product. With \( w \) defined, the total expected rate of return can be calculated as

\[ E(r_{\text{total}}) = \sum_{i=1}^{n} (r_i \cdot w_i), \]  

(11)

which due to the defined vectors \( r \) and \( w \) can be rewritten in matrix format, see Eq. (12).

\[ E(r_{\text{total}}) = r^T \cdot w \]  

(12)
An investor would want to maximize the total expected return as much as possible by rearranging the weights. At the same way as in Eq. (11), the total variance of the portfolio can be computed as follows.

$$\text{Var}(r_{\text{total}}) = \sum_{i=1}^{n} \sum_{j=1}^{n} (\text{Cov}(r_i, r_j) \cdot w_i \cdot w_j),$$

which with the defined matrix $H$ can rewritten by matrices as

$$\text{Var}(r_{\text{total}}) = w^T \cdot H \cdot w.$$  \hspace{1cm} (14)

An investor would want to minimize the total variance as much as possible by rearranging the weights. This causes two inherently different agendas; on one hand, the investor wants to maximize the expected rate of return and on the other hand, the investor wants to minimize the total variance. Mathematically, this can be described as

$$\min(\alpha \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} - (1 - \alpha) \sum_{i=1}^{n} r_i w_i),$$

or in matrix format

$$\min(\alpha w^T \cdot H \cdot w - (1 - \alpha) r^T \cdot w),$$  \hspace{1cm} (16)

where $\alpha$ is referred to as the risk aversion factor. If alpha is equal to one, the investor only cares about minimizing the risk and ignores the rate of return. Respectively, if alpha is equal to zero the investor only cares about maximizing the rate of return and ignores the risk. Consequently, an investor has a risk aversion factor between zero and one.

Given the expression that should be minimized, expression (16), and well defined constraints, Eq. (9), the problem to optimize is given by

$$\min(\alpha w^T \cdot H \cdot w - (1 - \alpha) r^T \cdot w), \hspace{1cm} e^T \cdot w = 1$$  \hspace{1cm} (17)

To optimize the problem above, there are several optimizing algorithms to consider. The algorithm used in this research is the interior-point-convex quadprog algorithm, because it is currently the most used one when it comes to financial optimizing. For further details on how the algorithm works, see Appendix [A].
2.4 Portfolio selection

In the last step, one of the optimal portfolios are chosen. Depending on alpha in the previous step the algorithm will produce different portfolios, all of which are optimal. Each portfolio has the highest possible expected rate of return for any given risk. If the investors allow a higher risk, the investor will expect a higher rate of return. These portfolios all lie on a line called the efficient frontier. Any portfolio beneath the frontier is sub-optimal and any portfolio above the frontier is impossible to create. An illustration of the efficient frontier is presented in Figure 1.

![Efficient Frontier Diagram](Image)

Figure 1: The efficient frontier.

2.5 Benefits of fewer securities

The benefits of having fewer securities in a portfolio are many, with the most obvious one being the decrease of transaction costs. Additionally, an investor that uses fewer securities
is less exposed to security specific taxes, constraints and other market imperfections. Another important benefit of using fewer securities is the decreased cost of managing. Moreover, Statman (1987) points out that portfolios consisting of fewer securities has higher reliability of the standard deviation estimate for returns than portfolios consisting of many securities. However, concentrated portfolios usually contains higher risks due to low diversification.

3 Method

3.1 Assumptions

During this study, some assumptions were made in order to put the study into practice. They are described below.

1. It is assumed that the results are independent from the choice of optimization algorithm,

2. It is assumed that the investor does not use short-selling and negative weights are therefore prevented,

3. It is assumed that the investor only use stocks from NASDAQ Stockholm for investment. That is, there is no risk-free alternative, such as bank saving, available,

4. The investor is risk-neutral, i.e. the investor is equally interested in maximizing the return as in minimizing the risk.

The first assumption is due to the fact that there exist several algorithms that would be able to perform the same task. In rare cases, these algorithms may find different minimum portfolios. However, if the variance-co-variance matrix is symmetric and positive semi definite, which it usually is, there can only exist one minimum, and any algorithm would have found the same point. Because of the rare occurrence of cases were several minimum exists, it is assumed that the average results of this study would not be significantly affected by varying algorithm. The algorithm used for this study is the interior-point-convex quadprog algorithm because it is the standard algorithm in MATLAB’s finance tool.
The second assumption was made because short selling is rarely used by the individual investor, which is the primary audience for this paper. Furthermore, short selling is more regulated by both Swedish and European Union laws (European Commission, 2013). These extra regulations can be difficult to implement mathematically when performing portfolio optimizations. Furthermore, by allowing short selling, the program could produce unrealistic portfolios. For example, short selling an asset for 500% of the total capital and buying another stock for 600% of the total capital would net a 100% investment, which would be feasible theoretically but completely unrealistic. Thus short selling would have to be coupled with some arbitrary constrains that would be dependent of the risk aversion of the investor.

The third assumption, that investors only use stocks from NASDAQ Stockholm is because of the fact that if risk-free options were to be considered, it would be necessary to estimate the real variance and co-variances of such an option. In reality a risk-free option is neither free from risk nor uncorrelated to the market. If these parameters would be set to zero, such a candidate could cause the algorithm to put all the weight in that risk-free option and thus generate zero risk portfolios. Such trivial results would prove little use for an investor. Furthermore, the choice of stock market to import data from is based on accessibility.

The last assumption is made to better reflect the nature of a typical investor. It would be possible to maximize the expected rate of return without considering the expected risk or minimizing the expected risk without considering the expected rate of return. Neither of these extreme portfolios are realistic from an investor’s perspective and all the optimal portfolios created in this research are chosen somewhere in the middle of these extremities.

### 3.2 Data

The data used in this research was collected online from the NASDAQ OMX Nordic web page. For each stock, an excel file containing closing prices for every market day between January 3rd 2000 and December 30th 2014 was downloaded, and all the data was exported to MATLAB. The 140 stocks that were used were listed on NASDAQ Stockholm during that whole period.
3.3 Optimization Approach

There are four input variables to consider before explaining the approach to the problem:

1. \( \text{nStocks} \) - number of different stocks in a portfolio \((1 \leq \text{nStocks} \leq 50)\),

2. \( \text{daysPast} \) - Number of historic data points considered when creating an optimal portfolio,

3. \( \text{daysForward} \) - Number of days ahead of the optimization, where the performance of the created portfolio is measured,

4. \( \text{nOptimization} \) - Number of times between 2000 and 2014 that optimizations were executed.

For each \( \text{nStocks} \), an optimal portfolio was constructed given a specific value of \( \text{daysPast} \). The suggested optimal portfolio was then evaluated at \( \text{daysForward} \) days ahead of the optimization date. This procedure was repeated multiple times over the historic data and the performance of every optimal portfolio was stored. The variables were set to:

\[
\text{daysPast} = 2 \text{ years} \quad \text{(18)}
\]
\[
\text{daysForward} = 1 \text{ year} \quad \text{(19)}
\]

Several other \( \text{daysPast} \) and \( \text{daysForward} \) were also tested to see if there would be a significant sensitivity in the choice of these inputs and thus evaluate the robustness of the results. More thoroughly, the results for different combinations between the two variables were saved and the mean performance was measured. This was done for:

\[
\text{daysPast} = 1, 1.5, 2, 2.5, 3 \text{ years} \quad \text{(20)}
\]
\[
\text{daysForward} = 0.5, 1, 1.5, 2, 2.5 \text{ years} \quad \text{(21)}
\]

When determining how frequent to optimize a portfolio and valuate it, there was a problem consisting of two conflicting issues:

1. The optimization had to be repeated often enough to gain reliable results,
2. Too many optimizations would demand too long computing run time.

To determine this trade off, several test sessions were done until a satisfying reliability and run time were reached. More specifically, the optimization frequency was set to one every 20 trading days. This yielded 188 optimization occasions for every portfolio size over the whole data set, and a run time of four hours. That is:

\[ n_{\text{Optimization}} = 188 \]  \hspace{1cm} (22)

However, the number of optimizations does not imply how often an investor has to reshape his or her portfolio but rather an indication of how much new data is used for every optimization occasion. I.e., with a new optimization occasion every 20 trading days and \( days_{\text{Past}} = 1 \text{ year} = 250 \text{ days} \), 8% new data will be used every new occasion. It is \( days_{\text{Forward}} \) that determines how often the portfolio is reshaped.

### 3.4 Compilation of Results

At the end of the optimizations, the performances of all portfolios containing \( n \) different stocks was calculated. The performances of the optimal portfolios were then compared to the performances of the randomly chosen portfolios, obtained from Statman (1987). Since the stocks used in this study differ from the ones used by Statman, the random portfolios and the optimal portfolios do not converge to the same asymptote line.

While the optimal portfolios converge to a risk corresponding to the systematic risk of NASDAQ Stockholm, the random portfolios created by Statman converge to a risk corresponding to the systematic risk of another market, at another time. In order to make the obtained results comparable to Statman’s, the obtained standard deviations of the optimal portfolios were calibrated to equalize the asymptotes of the two different market risks.

Finally, the risk of the optimal and random portfolios were compared by the ratio between the portfolio risks and the risk of one random stock, during the examined period of 2000-2014.
4 Results

The study showed that the optimal portfolios yielded lower risks than the random portfolios. The risk of an optimal portfolio decreases as the number of stocks increases. Initially the risk decreases rapidly but as the number of stocks becomes large, the risk converges to an asymptote. The risk of random portfolios decreases likewise as the number of stocks increases. This is presented in Figure 2.

Two major differences can be observed: Firstly, the risk of optimal portfolios containing one single stock is approximately 54% of the risk of a single random stock. Secondly, the risk of optimal portfolios converge faster to the asymptote. These two observations together suggests that optimal portfolios requires less number of stocks in order to di-

Figure 2: The ratio between the risk of a portfolio and the risk of a single random stock in the case where daysPast is 2 years and daysForward is 1 year.

*Data collected from Statman(1987)
versify away the unsystematic risk. The exact values used in Figure 2 are presented in Table 1.

Table 1: Portfolio risks, both actual and in relation to the risk of one random security.

Parameter values: \( \text{daysPast} = 2 \) years, \( \text{daysForward} = 1 \) year.

<table>
<thead>
<tr>
<th>Number of securities in portfolio</th>
<th>Optimal Portfolio Standard Deviation [%]</th>
<th>Optimal Portfolio relative Standard Deviation</th>
<th>Random Portfolio* Standard Deviation [%]</th>
<th>Random Portfolio* relative Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44.69</td>
<td>0.5436</td>
<td>49.24</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>39.92</td>
<td>0.4856</td>
<td>37.36</td>
<td>0.7588</td>
</tr>
<tr>
<td>4</td>
<td>37.19</td>
<td>0.4544</td>
<td>29.69</td>
<td>0.6030</td>
</tr>
<tr>
<td>6</td>
<td>35.89</td>
<td>0.4317</td>
<td>26.64</td>
<td>0.5411</td>
</tr>
<tr>
<td>8</td>
<td>35.08</td>
<td>0.4268</td>
<td>24.98</td>
<td>0.5074</td>
</tr>
<tr>
<td>10</td>
<td>34.69</td>
<td>0.4220</td>
<td>23.93</td>
<td>0.4861</td>
</tr>
<tr>
<td>12</td>
<td>34.46</td>
<td>0.4192</td>
<td>23.20</td>
<td>0.4713</td>
</tr>
<tr>
<td>14</td>
<td>34.32</td>
<td>0.4175</td>
<td>22.67</td>
<td>0.4604</td>
</tr>
<tr>
<td>16</td>
<td>34.22</td>
<td>0.4162</td>
<td>22.26</td>
<td>0.4521</td>
</tr>
<tr>
<td>18</td>
<td>34.16</td>
<td>0.4155</td>
<td>21.94</td>
<td>0.4456</td>
</tr>
<tr>
<td>20</td>
<td>34.13</td>
<td>0.4152</td>
<td>21.68</td>
<td>0.4403</td>
</tr>
<tr>
<td>25</td>
<td>34.12</td>
<td>0.4150</td>
<td>21.20</td>
<td>0.4305</td>
</tr>
<tr>
<td>30</td>
<td>34.07</td>
<td>0.4144</td>
<td>20.87</td>
<td>0.4239</td>
</tr>
<tr>
<td>35</td>
<td>33.96</td>
<td>0.4131</td>
<td>20.63</td>
<td>0.4191</td>
</tr>
<tr>
<td>40</td>
<td>33.83</td>
<td>0.4116</td>
<td>20.46</td>
<td>0.4155</td>
</tr>
<tr>
<td>45</td>
<td>33.80</td>
<td>0.4111</td>
<td>20.32</td>
<td>0.4126</td>
</tr>
<tr>
<td>50</td>
<td>33.73</td>
<td>0.4103</td>
<td>20.20</td>
<td>0.4103</td>
</tr>
</tbody>
</table>

*Data collected from Statman(1987)

It is notable that the standard deviations of the portfolios do not converge to the same value, whilst the relative standard deviation of the portfolios converge towards the same asymptotic value. This is explained in 3.4 Compilation of Results.

Figure 3 displays the average risk trends when the inputs \( \text{daysForward} \) and \( \text{daysPast} \) were varied. This was to examine the robustness of the results. The average risk trend behaves similarly to the specific case of Figure 2. Table 2 shows the specifics of Figure 3.
Figure 3: The ratio between the risk of a portfolio and the risk of a single random stock, in the case where daysPast and daysForward were varied.

*Data collected from Statman(1987)
Table 2: Mean portfolio risks, both actual and in relation to the risk of one random stock, with varying $days_{Past}$ and $days_{Forward}$.

<table>
<thead>
<tr>
<th>Number of securities in portfolio</th>
<th>Optimal Portfolio Standard Deviation [%]</th>
<th>Optimal Portfolio relative Standard Deviation</th>
<th>Random Portfolio* Standard Deviation [%]</th>
<th>Random Portfolio* relative Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43.71</td>
<td>0.5857</td>
<td>49.24</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>41.53</td>
<td>0.5039</td>
<td>37.36</td>
<td>0.7588</td>
</tr>
<tr>
<td>4</td>
<td>39.50</td>
<td>0.4542</td>
<td>29.69</td>
<td>0.6030</td>
</tr>
<tr>
<td>6</td>
<td>37.76</td>
<td>0.4354</td>
<td>26.64</td>
<td>0.5411</td>
</tr>
<tr>
<td>8</td>
<td>36.87</td>
<td>0.4268</td>
<td>24.98</td>
<td>0.5074</td>
</tr>
<tr>
<td>10</td>
<td>36.60</td>
<td>0.4262</td>
<td>23.93</td>
<td>0.4861</td>
</tr>
<tr>
<td>12</td>
<td>36.50</td>
<td>0.4256</td>
<td>23.20</td>
<td>0.4713</td>
</tr>
<tr>
<td>14</td>
<td>36.45</td>
<td>0.4223</td>
<td>22.67</td>
<td>0.4604</td>
</tr>
<tr>
<td>16</td>
<td>36.35</td>
<td>0.4193</td>
<td>22.26</td>
<td>0.4521</td>
</tr>
<tr>
<td>18</td>
<td>36.18</td>
<td>0.4173</td>
<td>21.94</td>
<td>0.4456</td>
</tr>
<tr>
<td>20</td>
<td>36.13</td>
<td>0.4162</td>
<td>21.68</td>
<td>0.4403</td>
</tr>
<tr>
<td>25</td>
<td>36.02</td>
<td>0.4143</td>
<td>21.20</td>
<td>0.4305</td>
</tr>
<tr>
<td>30</td>
<td>35.97</td>
<td>0.4135</td>
<td>20.87</td>
<td>0.4239</td>
</tr>
<tr>
<td>35</td>
<td>36.05</td>
<td>0.4132</td>
<td>20.63</td>
<td>0.4191</td>
</tr>
<tr>
<td>40</td>
<td>35.95</td>
<td>0.4116</td>
<td>20.46</td>
<td>0.4155</td>
</tr>
<tr>
<td>45</td>
<td>35.93</td>
<td>0.4114</td>
<td>20.32</td>
<td>0.4126</td>
</tr>
<tr>
<td>50</td>
<td>35.94</td>
<td>0.4103</td>
<td>20.20</td>
<td>0.4103</td>
</tr>
</tbody>
</table>

*Data collected from Statman(1987)

Once again, the actual portfolio standard deviations do not converge to the same value, but the relative standard deviations do, in accordance with the calibration of the obtained results for differences in market risks. In the case when $days_{Past}$ and $days_{Forward}$ were varied and a mean of their standard deviations was obtained, the relative standard deviation was higher than in the case with fixed parameters, which can be observed by comparing the results in Table 1 and Table 2. It is also shown in Figure 3.

## 5 Discussion

### 5.1 Analysis of results

The value to which the curves are converging to in Figure 2 signifies the systematic risk of the market and cannot be eliminated by diversification. Statman argued that, after
30 random securities in a portfolio, the marginal risk decrease was not large enough to compensate for the increase of transaction costs and the cost-benefit relationship was at an equilibrium. Thus 30 securities would be sufficient to diversify a portfolio. The same marginal risk decrease can be observed at approximately 10 securities in the case of optimal portfolios. Statman’s argument can therefore also be used to state that 10 securities would be sufficient to diversify optimal portfolios.

However, there are some crucial differences between this study and Statman’s that should be considered when interpreting these results. The data Statman base his study on is from the US stock market S&P500 from before year 1987 while this study was based on data from the Swedish stock market NASDAQ Stockholm ahead of the year 2000. The risk behavior and other stock market characteristics may be inherently different in that set of data. An observable example of this is the different systematic market risks that the results converge to. Other less noticeable differences could also exists such as co-variances intensity between the stocks in the market. Furthermore, Statman calculated the cost-benefit equilibrium considering the average transaction costs of his time and place which are not necessarily the same as of today in Sweden. With this in mind, it is important to interpret the results with caution. The information derived from this comparison is more to be regarded as strong indications rather than absolute facts.

The reason why less securities are needed to diversify optimal portfolios are the major two differences between the risk trends stated in the results. The observation that the risk of a single optimal stock portfolios is 54% of the risk of a single random stock can be explained by looking into the algorithm. If the optimization algorithm only may choose one security, it will choose the security with the least historic volatility. Such security will, on average, be less volatile than the average security.

The second observation is that the risk converges to the systematic risk faster. This is due to the fact that the algorithm considers the co-variances of every investment candidate and chooses the securities so that the expected risk will become as small as possible. Both of these differences are intuitive and the results are consistent with what was expected.

The results suggests that the optimal portfolios diversify at fewer securities compared to random portfolios, in accordance to expectations. It is however important to be aware of the possible occurrence of survivorship bias, which could have caused the optimal
portfolios to contain less risk than they normally would. Since the optimization procedure required intact data from all available stocks to work, it was required to sort out many companies that were not registered for the whole studied period. Some of these companies could have been unregistered because of bankruptcy and some companies were new. These unaccounted companies could very well have been included in the created portfolios and affect the risk behavior, perhaps only slightly, but still worth mentioning.

The obtained results make the conclusion of Evans and Archer interesting; 10 securities are sufficient to diversify away the unsystematic risk. However, while 10 securities can be sufficient to diversify away the unsystematic risk when using portfolio optimization, it will not be sufficient when creating any portfolio of that size. The vast majority of portfolios are not considered optimal according to modern portfolio theory. So Statman’s conclusions still hold true when using portfolios that have not been created through portfolio optimization. This difference is crucial and can not be over stated; 10 securities can be sufficient if the portfolio is optimal but 30 securities will be sufficient regardless portfolio type.

The results from this study shed light upon the difference between optimal portfolios and random portfolios, and how optimization can be used to effectively diversify away the unsystematic risk of a portfolio. Moreover, they suggest that individual investors could use optimization as a tool to diversify their portfolios. Even though not all the unsystematic risk can be eliminated, optimization is yet a powerful tool that can be used on relatively small portfolios, consisting of a handful securities, to diversify away a considerable amount of the portfolio risk, as presented in Figure 2.

5.2 Conclusion

When performing portfolio optimization with modern portfolio theory, it is sufficient to use 10 securities to create a well-diversified portfolio. This is due to the fact that an optimal portfolio deliberately is chosen to achieve maximum diversification given the number of securities allowed, and such portfolios on average have less associated risks than a random portfolio. This insight can be used by various investors when using portfolio optimization with modern portfolio theory.
5.3 Applicability of study

This study has shown that, by using portfolio optimization when creating a portfolio, it is not necessary to use more than 10 securities. Using 10 securities instead of 30, as Statman suggested, yields multiple advantages, as mentioned in section 2.5 Benefits of fewer securities. Since the factors of knowledge asymmetry and company news is not accounted for in modern portfolio theory, it is still advisable to keep track on the companies in the portfolio. Having this in mind, the time and work spent on managing 10 securities instead of 30 would be significantly less, as would the transaction costs, potential tax costs and costs of other market imperfections.

This study will also be a firm source of information for anyone curious on optimization theory and its benefits today. Moreover, this paper and its algorithms can be used as guidance or tools for similar studies.

5.4 Suggestions for further research

Since the possibility of computing optimal portfolios on big amounts of data has increased vastly during recent years, a lot of research opportunities have arisen in this field. Here follow a few suggestions on how to take this research further and contribute to the evaluation portfolio optimization from an investor perspective:

- A study could be done investigating what set of constraints on the optimization would yield lowest portfolio risk, highest portfolio return or highest portfolio Sharpe ratio. Examples of relevant constraints to investigate could be:
  
  1. Combination of daysPast and daysForward,
  2. Risk aversion, i.e. priority between risk minimization and return maximization,
  3. Allowance of short selling.

- Return of optimal portfolios in relation to randomly chosen portfolios.

- Sharpe ratio of optimal portfolios in relation to randomly chosen portfolios.

- How often a portfolio should be updated.
References


No. 1 (Mar., 1952), pp. 77-91.


A Interior-point-convex quadprog algorithm (Math-Works, 2016)

Quadratic programming is the problem of finding a vector $x$ that minimizes a quadratic function, possibly subject to linear constraints:

$$\min \left( \frac{1}{2} x^T \cdot H \cdot x + c^T \cdot x \right)$$

(23)

such that

$$A \cdot x \leq b, \quad A_{eq} \cdot x = b_{eq}$$

(24)

The interior-point-convex algorithm performs the following steps:

- Presolve/Postsolve
- Generate Initial Point
- Predictor-Corrector
- Multiple Corrections
- Total Relative Error

A.1 Presolve/Postsolve

The algorithm begins by attempting to simplify the problem by removing redundancies and simplifying constraints. The tasks performed during the presolve step include:

- Check if any variables have equal upper and lower bounds. If so, check for feasibility, and then fix and remove the variables.
- Check if any linear inequality constraint involves just one variable. If so, check for feasibility, and change the linear constraint to a bound.
- Check if any linear equality constraint involves just one variable. If so, check for feasibility, and then fix and remove the variable.
• Check if any linear constraint matrix has zero rows. If so, check for feasibility, and delete the rows.

• Check if the bounds and linear constraints are consistent.

• Check if any variables appear only as linear terms in the objective function and do not appear in any linear constraint. If so, check for feasibility and boundedness, and fix the variables at their appropriate bounds.

• Change any linear inequality constraints to linear equality constraints by adding slack variables.

If algorithm detects an infeasible or unbounded problem, it halts and issues an appropriate exit message. The algorithm might arrive at a single feasible point, which represents the solution. If the algorithm does not detect an infeasible or unbounded problem in the presolve step, it continues, if necessary, with the other steps. At the end, the algorithm reconstructs the original problem, undoing any presolve transformations. This final step is the postsolve step.

A.2 Generate Initial Point

The initial point \( x_0 \) for the algorithm is:

1. Initialize \( x_0 \) to \( \text{ones}(n,1) \), where \( n \) is the number of rows in \( H \).

2. For components that have both an upper bound \( \text{ub} \) and a lower bound \( \text{lb} \), if a component of \( x_0 \) is not strictly inside the bounds, the component is set to \( \frac{\text{ub} + \text{lb}}{2} \).

3. For components that have only one bound, modify the component if necessary to lie strictly inside the bound.
A.3 Predictor-Corrector

Similar to the fmincon interior-point algorithm, the interior-point-convex algorithm tries to find a point where the Karush-Kuhn-Tucker (KKT) conditions hold. For the quadratic programming problem described in Quadratic Programming Definition, these conditions are:

\[
H \cdot x + c - A_{eq}^T \cdot y - \overline{A}^T \cdot z = 0 \tag{25}
\]
\[
\overline{A} \cdot x - \overline{b} - s = 0 \tag{26}
\]
\[
A_{eq} \cdot x - b_{eq} = 0 \tag{27}
\]
\[
s_i \cdot z_i = 0, \ i = 1, 2, \ldots, m \tag{28}
\]
\[
s \geq 0 \tag{29}
\]
\[
z \geq 0 \tag{30}
\]

Here

- \( \overline{A} \) is the extended linear inequality matrix that includes bounds written as linear inequalities. \( \overline{b} \) is the corresponding linear inequality vector, including bounds.

- \( s \) is the vector of slacks that convert inequality constraints to equalities. \( s \) has length \( m \), the number of linear inequalities and bounds.

- \( z \) is the vector of Lagrange multipliers corresponding to \( s \).

- \( y \) is the vector of Lagrange multipliers associated with the equality constraints.

The algorithm first predicts a step from the Newton-Raphson formula, then computes a corrector step. The corrector attempts to better enforce the nonlinear constraint \( s_i \cdot z_i = 0 \). Definitions for the predictor step:

- \( r_d \), the dual residual:

\[
r_d = H \cdot x + c - A_{eq}^T \cdot y - \overline{A}^T \cdot z \tag{31}
\]
• \( \mathbf{r}_{eq} \), the primal equality constraint residual:

\[
\mathbf{r}_{eq} = \mathbf{A}_{eq} \cdot \mathbf{x} - \mathbf{b}_{eq}
\]  

(32)

• \( \mathbf{r}_{ineq} \), the primal inequality constraint residual, which includes bounds and slacks:

\[
\mathbf{r}_{ineq} = \overline{\mathbf{A}} \cdot \mathbf{x} - \overline{\mathbf{b}} - \mathbf{s}
\]  

(33)

• \( \mathbf{r}_{sz} \), the complementarity residual:

\[
\mathbf{r}_{sz} = \mathbf{S}z
\]  

(34)

\( \mathbf{S} \) is the diagonal matrix of slack terms, \( \mathbf{z} \) is the column matrix of Lagrange multipliers.

• \( \mathbf{r}_c \), the average complementarity:

\[
\mathbf{r}_c = \frac{\mathbf{s}^T \cdot \mathbf{z}}{m}
\]  

(35)

In a Newton step, the changes in \( \mathbf{x}, \mathbf{s}, \mathbf{y}, \) and \( \mathbf{z} \), are given by:

\[
\begin{pmatrix}
\mathbf{H} & 0 & -\mathbf{A}_{eq}^T & -\overline{\mathbf{A}}^T \\
\mathbf{A}_{eq} & 0 & 0 & 0 \\
\overline{\mathbf{A}} & -\mathbf{I} & 0 & 0 \\
0 & \mathbf{Z} & 0 & \mathbf{S}
\end{pmatrix}
\begin{pmatrix}
\Delta \mathbf{x} \\
\Delta \mathbf{s} \\
\Delta \mathbf{y} \\
\Delta \mathbf{z}
\end{pmatrix}
= -
\begin{pmatrix}
\mathbf{r}_d \\
\mathbf{r}_{eq} \\
\mathbf{r}_{ineq} \\
\mathbf{r}_{sz}
\end{pmatrix}
\]  

(36)

However, a full Newton step might be infeasible, because of the positivity constraints on \( \mathbf{s} \) and \( \mathbf{z} \). Therefore, quadprog shortens the step, if necessary, to maintain positivity. Additionally, to maintain a "centered" position in the interior, instead of trying to solve \( \mathbf{s}_i \cdot \mathbf{z}_i = 0 \), the algorithm takes a positive parameter \( \sigma \), and tries to solve

\[
\mathbf{s}_i \mathbf{z}_i = \sigma \mathbf{r}_c
\]  

(37)

quadprog replaces \( \mathbf{r}_{sz} \) in the Newton step equation with \( \mathbf{r}_{sz} + \Delta \mathbf{s} \Delta \mathbf{z} - \sigma \mathbf{r}_c \mathbf{1} \), where \( \mathbf{1} \) is the vector of ones. Also, quadprog reorders the Newton equations to obtain a symmetric, more numerically stable system for the predictor step calculation.
A.4 Multiple Corrections

After calculating the corrected Newton step, quadprog can perform more calculations to get both a longer current step, and to prepare for better subsequent steps. These multiple correction calculations can improve both performance and robustness.

A.5 Total Relative Error

quadprog calculates a merit function $\phi$ at every iteration. The merit function is a measure of feasibility, and is also called total relative error. quadprog stops if the merit function grows too large. In this case, quadprog declares the problem to be infeasible. The merit function is related to the KKT conditions. Use the following definitions:

$$\rho = \max(1, ||H||, ||A||, ||A_{eq}||, ||c||, ||b||, ||b_{eq}||)$$  \hfill (38)
$$r_{eq} = A_{eq} \cdot c - b_{eq}$$  \hfill (39)
$$r_{ineq} = A \cdot x - \bar{b} + s$$  \hfill (40)
$$r_d = H \cdot x + c^T \lambda_{eq} + A_{eq}^T \bar{\lambda}_{ineq}$$  \hfill (41)
$$g = x^T \cdot H \cdot x + f^T \cdot x - \bar{b}^T \bar{\lambda}_{ineq} - b_{eq}^T \lambda_{eq}.$$  \hfill (42)

The notation $A$ and $b$ means the linear inequality coefficients, augmented with terms to represent bounds. The notation $\bar{\lambda}_{ineq}$ similarly represents Lagrange multipliers for the linear inequality constraints, including bound constraints. This was called $z$ in Predictor-Corrector, and $\lambda_{eq}$ was called $y$. The merit function $\phi$ is

$$\frac{1}{\rho}(\max ||r_{eq}||_\infty, ||r_{ineq}||_\infty, ||r_d||_\infty) + g).$$  \hfill (43)

(MathWorks, 2016)
B MATLAB Code

B.1 Main Code

%This is the main code for the computing sequences of the study
%It calls another functions, which in turn calls other functions
%(see below in appendix)

%% Preparing
clear
close all

load PricesAndDatesAndComparision.mat; %Loading company data

%% Initial Parameters

%Numbers of past data points for each optimization
DaysPast = [250 375 500 625 750];

%Days after optimization when performance is measured
DaysForward = [125 250 375 500 625];

minN = 1; %Smallest optimal portfolio
maxN = 50; %Largest optimal portfolio
intervals = 25; %Days between optimization processes

%Number of optimization processes
nOptimizations = floor(length(dateRef)/intervals);

%Pre-allocating matrix for saving standard deviations
Stdevs = zeros(maxN-minN+1,length(DaysPast),length(DaysForward));

%% Loop
for k = 1:length(DaysForward) %Looping for different daysForward
    daysForward = DaysForward(k);
for j = 1:length(DaysPast) %Looping for different daysPast
daysPast = DaysPast(j);

if daysPast>=daysForward %Checking that daysPast>daysForward

%Preallocation
PerformanceOfN = zeros(nOptimizations,maxN-minN+1);
Start = ceil(daysPast/intervals);
End = floor(nOptimizations-daysForward/intervals)-1;
for i = Start:End

%Calling the optimization programs
[nPerformance, nStocks] = datePerformance(priceData,...
   relData, dateRef, dateRef(intervals*i + 1),...
   daysPast, daysForward, minN, maxN);
%Saving relevant values
PerformanceOfN(i,:) = nPerformance;

end

%Saving the standard deviations of the optimizations for
%every combination of n, daysPast and daysForward
Stdevs(:,j,k) = std(PerformanceOfN);

end
end
end

%% Result Management

%Annualizing the results for all daysForward
S = Stdevs;
S(S==0) = nan;
for norms = 1:length(DaysForward)
   S(:,:,norms) = (S(:,:,norms))*sqrt(250/DaysForward(norms));
end

%Mean prestanda of all combinations of daysPast and daysForward
S2 = zeros(maxN,1);
for i=1:maxN
    S2(i,1) = nanmean(nanmean(S(i,:,:)));
end

% Calibrating results for comparison
S2=S2/(S2(end)/statResults(end));
S2=S2/statResults(1);
statResults = statResults/statResults(1);

% Presenting Results
x = minN:maxN;
x = x';
f = fit (x,S2,'exp2'); % Obtaining a trend for the results
figure()
plot(1:length(S2),S2,'k .',statStocks,statResults);
hold on
plot (f)
axis([1 50 0.37 1.05])
xlabel ('Number of Stocks in Portfolio','FontSize',12)
ylabel ('Standard Deviation Ratio','FontSize',12)
legend({'Risk of Optimal Portfolios', 'Risk of Random Portfolios*',...
    'Risk trend of Optimal portfolios'}, 'Position',...
    [0.48,0.72,0.34,0.12], 'FontSize',12)

B.2 Date performance

function [np, n] = datePerformance(priceData, relData, period, date,...
    daysPast, daysForward, minN, maxN)

% This function returns the performance of each portfolio
% optimizations performed on given date. Provided days past the use as
% optimization data, days forward - days ahead of portfolio creation date
% to evaluate portfolio performance on. Pricedata is the price data of the
% the whole stock history, relData is the same but converted to relative
% differences rather that absolute. Period is the whole stock history in
% dates. minN and maxN tell us between which portfolio sizes the program
% should work with. In the study we go between 1 and 50 stocks.

[W,Wi,~,~] = portopt(priceData, relData, period,...
    period(find(period==date)-daysPast), date, minN, maxN);

[~,nports]=size(W);

n = (minN:maxN)';
np = zeros(nports,1);

for i=1:nports

    np(i,1) = portperformance(priceData, period, date,...
        period(find(period==date)+daysForward), W(1:i,i), Wi(1:i,i));
end

end

B.3 Portfolio Optimization

function [W,Wi,exvar,exret,C,r] = portopt(priceData, relData,...
    period, start, ends,minN,maxN)

%This function creates a number of optimal portfolios on the date given by
%input end, using data from the input start. Pricedata is the price data of
%the whole stock history, relData is the same but converted to relative
%differences rather that absolute. Period is the whole stock history in
%dates. minN and maxN tell us between which portfolio sizes the program
%should work with. In the study we go between 1 and 50 stocks.

C = covar(relData, period, start, ends); % creates variance-covariance matrix
r = OAROT(priceData, period, start-1, ends-1); % creates expected rate of
% return vector.
[~,nstocks] = size(priceData);
b = zeros(nstocks,1);
A = -eye(nstocks); %Disallowing negative weights - no short selling
Aeq = ones(1,nstocks);
beq = 1; %Make sure all weights add up to one.

options = optimset('Display','off');

wall=quadprog(100*C,-r,A,b,Aeq,beq,[],[],[],options);

%Perform preliminary optimization using interior-point convex algorithm.

[~, walli ] = sort( wall, 'descend' );
% sort stocks by weights

n=minN:maxN;
exvar = zeros(1,maxN-minN+1);
exret = zeros(1,maxN-minN+1);
W=zeros(maxN,maxN-minN+1);
Wi=zeros(maxN,maxN-minN+1);

for i = 1:maxN-minN+1
    wi = walli(1:i); %creating portfolio choices for each portfolio size
    Wi(1:i,i) = wi;

    cC = covar(relData(:,wi), period, start, ends); %create cov matrix
    cr = r(wi); %and rate of return vector for the choosen stocks.

    cb = zeros(n(i),1);
    cA = -eye(n(i));

    cAeq = ones(1,n(i));
    cbeq = 1;

    %using the same constraints as in the preliminary optimization
\begin{verbatim}
W(1:n(i),i)=quadprog(100*cC,-cr,cA,cb,cAeq,cbeq,[],[],[],options);

%determine new weheights for each portfolio size.

exret(i) = cr*W(1:n(i),i);
exvar(i) = W(1:n(i),i)'*cC*W(1:n(i),i);

end
end

B.4 Portfolio performance

function portReturn = portperformance(priceData, period,...
    currentDate, futureDate,w,wi)
    %Evaluates the performance for a portfolio wi, with the weights w.
currentDateIndex = find(period==currentDate);
futureDateIndex = find(period==futureDate);
r = (priceData(futureDateIndex,wi)./priceData(currentDateIndex,wi));
portReturn = r*w-1;
end

B.5 Variance-covariance matrix

function C = covar(relData, period, start, ends)
    %creates the variance-covariance matrix within given dates and given data.
    startIndex = find(period==start);
    endIndex = find(period==ends);
    subRelData = relData(startIndex:endIndex,:);
    C = cov(subRelData);
    %matlab function cov performs the calculations stated in the theory
    %section.
end

B.6 The expected rate of return vector
\end{verbatim}

36
function r = OAROT(priceData, period, start, ends)
%Returns the ordinary arithmetic rate of return.
startIndex = find(period==start);
endIndex = find(period==ends);
subPriceData = priceData(startIndex:endIndex,:);

[rSize,cSize] = size(subPriceData);

r = zeros(1,cSize);

for i = 2: rSize
    r(1,:) = r(1,:)+(subPriceData(i,:)/subPriceData(i-1,:)-1)/(rSize-1);
end