

# Comparison and Development of Methods for Index Tracking



Dissertation  
zur Erlangung des akademischen Grades  
Doktor der Wirtschaftswissenschaften  
(Dr. rer. pol.)  
der Frankfurt School of Finance & Management

vorgelegt von  
Denis Karlow

Frankfurt am Main  
September 2012

Gutachter:

Prof. Dr. Peter Roßbach (Frankfurt School of Finance & Management)

Prof. Dr. Natalie Packham (Frankfurt School of Finance & Management)

Prof. Dr. Manfred Gilli (University of Geneva)

Tag der Disputation: 18.02.2013

To my parents

## **Acknowledgements**

I would especially like to thank my first adviser, Prof. Dr. Peter Roßbach. He gave me a chance to write my dissertation on a theme that is close to me. Without his support, the completion of the thesis would not have been possible. As Karl Marx said, “there is no royal road to science”, but Peter taught me how “to climb each of its steep paths to have a chance of gaining its luminous summits.” I would also like to thank Prof. Dr. Natalie Packham and Prof. Dr. Manfred Gilli for their time spent reading and evaluating the thesis. I also thank Emily Stapleton for proofreading this thesis.

The time I spent at Frankfurt School was enjoyable and unforgettable. I would like to thank all my colleagues, my fellow PhD students, who made this possible.

I would also like to thank Anna for her patience and love, and my sister Natalie for her unconditional personal support. This doctoral dissertation is dedicated to my parents, who never allowed me to relax and kept on telling me that it is necessary to set goals and reach them.

# Contents

<b>1 INTRODUCTION .....</b>	<b>1</b>
<b>2 FOUNDATIONS OF INDEX TRACKING .....</b>	<b>6</b>
<b>2.1 Index-based investment.....</b>	<b>7</b>
<b>2.2 Index construction .....</b>	<b>12</b>
2.2.1 Equity indices .....	12
2.2.1.1 Construction rules.....	12
2.2.1.2 S&P1500 .....	19
2.2.2 Alternative indices.....	20
<b>2.3 Types of index funds.....</b>	<b>22</b>
<b>2.4 Tracking approaches.....</b>	<b>25</b>
2.4.1 Physical replication .....	25
2.4.1.1 Full replication .....	25
2.4.1.2 Sampling.....	26
2.4.2 Synthetic replication .....	27
2.4.2.1 Instruments .....	27
2.4.2.1.1 Futures .....	27
2.4.2.1.2 Options .....	28
2.4.2.1.3 Swaps.....	29
2.4.2.1.4 Equity-linked notes.....	30
2.4.2.2 Structuring the index fund .....	30
2.4.3 Evaluation and discussion .....	34
<b>2.5 Summary .....</b>	<b>43</b>
<b>3 SAMPLING.....</b>	<b>44</b>
<b>3.1 Related work .....</b>	<b>45</b>
3.1.1 Selection .....	46
3.1.1.1 Simple selection criteria .....	46
3.1.1.2 Coverage of index structure.....	47
3.1.1.3 Optimized selection .....	50
3.1.2 Weighting .....	52
3.1.2.1 Heuristic weighting .....	52
3.1.2.2 Optimized weighting .....	52

<b>3.2 Measures of tracking quality .....</b>	<b>57</b>
3.2.1 Traditional measures of tracking quality .....	57
3.2.2 Ex post tracking quality .....	63
<b>3.3 Selected methods of optimized sampling .....</b>	<b>66</b>
3.3.1 Minimizing traditional measures .....	66
3.3.2 Minimax approach.....	68
3.3.3 Minimization of conditional value-at-risk .....	69
3.3.4 Portfolio structure: weights versus numbers.....	70
<b>3.4 Requirements and performance assessment .....</b>	<b>76</b>
3.4.1 The absolute stability of the tracking quality.....	76
3.4.2 The relative stability of the tracking quality .....	77
3.4.3 Resistance of the tracking portfolio estimation .....	84
<b>3.5 Summary .....</b>	<b>86</b>
<b>4 EMPIRICAL ANALYSIS.....</b>	<b>87</b>
<b>4.1 Portfolio structure: weights versus numbers .....</b>	<b>88</b>
4.1.1 Research method .....	88
4.1.2 Empirical results .....	90
4.1.3 Conclusion.....	98
<b>4.2 The stability of the minimization of traditional measures .....</b>	<b>99</b>
4.2.1 Research method .....	99
4.2.2 Empirical results .....	100
4.2.3 Conclusion.....	113
<b>4.3 Resistance of the tracking portfolio estimation.....</b>	<b>115</b>
4.3.1 Research method .....	115
4.3.2 Empirical results .....	117
4.3.3 Conclusion.....	119
<b>4.4 Summary .....</b>	<b>120</b>
<b>5 A METHOD FOR ROBUST INDEX TRACKING .....</b>	<b>121</b>
<b>5.1 Statistical learning theory .....</b>	<b>122</b>
5.1.1 Overview of statistical learning theory .....	122
5.1.2 General setting of a learning problem and learning consistency .....	123
5.1.3 Support vector (SV) techniques.....	126
5.1.3.1 Support vector machines (SVMs).....	126

5.1.3.1.1	Separable case .....	126
5.1.3.1.2	Non-separable case .....	129
5.1.3.1.3	Kernel SVM.....	132
5.1.3.2	Support vector regression (SVR).....	135
5.1.4	Model selection .....	138
5.1.4.1	Bias-variance trade-off, model complexity, and overfitting .....	138
5.1.4.2	Analytic method of model selection for classification problems.....	142
5.1.4.3	Analytic methods of model selection for regression problems.....	145
5.1.4.4	Measuring model complexity .....	148
5.1.4.4.1	Measuring the effective degrees of freedom.....	149
5.1.4.4.2	Measuring the effective VC dimension .....	149
5.1.4.5	Resampling methods of model selection .....	151
5.1.4.6	Model selection based on the number of support vectors.....	152
5.1.4.7	Model selection process .....	152
5.1.5	Extended version of SVR .....	153
5.1.6	The transformation of SVR into SVM.....	154
<b>5.2</b>	<b>Tracking portfolio selection using SV techniques.....</b>	<b>159</b>
5.2.1	SVR for index tracking.....	159
5.2.2	Selecting the size of the insensitive region .....	161
5.2.2.1	Minimax approach.....	161
5.2.2.2	$\nu$ -SVR.....	163
5.2.3	Preliminary informational structures for the tracking portfolio .....	163
5.2.4	Structural tracking error minimization (STEM) .....	167
5.2.5	Further modifications of the SVR.....	168
<b>5.3</b>	<b>Comparison of STEM-based methods with other approaches .....</b>	<b>171</b>
5.3.1	Mean-variance portfolio optimization .....	171
5.3.2	Generalized framework for portfolio optimization.....	172
5.3.3	Portfolio optimization with a chance constraint .....	174
5.3.4	Conditional value-at-risk .....	176
5.3.5	Box constraints .....	176
<b>5.4</b>	<b>Empirical analysis of STEM.....</b>	<b>178</b>
5.4.1	Resistance of tracking portfolio estimation using STEM .....	178
5.4.1.1	Research method .....	178
5.4.1.2	Empirical results .....	179
5.4.1.3	Conclusion.....	181
5.4.2	Selecting the model of optimal complexity .....	181
5.4.2.1	Research method .....	181
5.4.2.2	Empirical results .....	185

5.4.2.3 Conclusion.....	188
5.4.3 Stability of methods using the STEM approach .....	189
5.4.3.1 Research method .....	189
5.4.3.2 Empirical results.....	189
5.4.3.3 Conclusion.....	199
5.5 Summary .....	200
<b>6 CONCLUDING REMARKS .....</b>	<b>202</b>
<b>APPENDIX 1 .....</b>	<b>204</b>
<b>APPENDIX 2 .....</b>	<b>207</b>
<b>REFERENCES .....</b>	<b>210</b>
<b>ABOUT THE AUTHOR.....</b>	<b>220</b>



## List of Tables

Table 2.1: Index additions and deletions. Source: Standard & Poor's, Deutsche Börse. ....	17
Table 2.2: S&P1500 Methodology.....	19
Table 2.3: Correlation of Substitute Basket / Collateral Basket with the Index .....	32
Table 3.1: Literature Review: Weights or Numbers .....	54
Table 3.2: Literature Review: Length of Estimation Period and Data Frequency.....	55
Table 3.3: Approximations of Tracking Portfolio Returns .....	75
Table 4.1: The Number of Possible Combinations of Assets for the Tracking Portfolios.....	89
Table 4.2: Portfolio Structure: Results for the Estimation Period (RMSE).....	91
Table 4.3: Portfolio Structure: Results for the Estimation Period (ETQ) .....	92
Table 4.4: Portfolio Structure: Results for the Investment Period (ETQ).....	93
Table 4.5: Best Methods by Median (ETQ).....	94
Table 4.6: Results for the Investment Period (RMSE).....	95
Table 4.7: Best Methods by Median (RMSE).....	96
Table 4.8: Wilcoxon Test Summary Table for Period 1, Tracking Portfolio Size 240.....	97
Table 4.9: Best Methods, with Significance Level 0.05 (ETQ) .....	97
Table 4.10: Best Methods, with Significance Level 0.05 (RMSE).....	97
Table 4.11: Comparison of Levels of Tracking Quality over Different Time Periods .....	100
Table 4.12: Comparison of Ratios: Investment/Estimation .....	102
Table 4.13: Stability of Relative Tracking Quality (MAD, TP size: 120) .....	103
Table 4.14: Stability Ratio and Correlation Coefficients for Traditional Measures .....	105
Table 4.15: Levels of Tracking Quality over Different Time Periods (Best 10%).....	106
Table 4.16: Ratios of Tracking Qualities (5000/Best 10%) .....	107
Table 4.17: Comparison of Levels of Tracking Quality in the Investment Period (ETQ).....	109
Table 4.18: Stability Ratios, and Kendall and Spearman Rank Correlation Coefficients (ETQ) .....	110
Table 4.19: Comparison of Levels of Tracking Quality (ETQ, Best 10%) .....	111
Table 4.20: Ratios of Tracking Qualities (ETQ, 5,000/Best 10%) .....	112
Table 4.21: Resistance of the Tracking Portfolio Estimation .....	117
Table 5.1: STEM-Based Methods of Optimized Sampling .....	178
Table 5.2: Reduction in the Variability of the Tracking Portfolio Estimation in the Investment Period, with Decreasing C (Period 5).....	180
Table 5.3: Minimal Values of the Median ETQs for the Five Time Periods .....	181
Table 5.4: Analytic Model Selection Methods for the SVM_CW Method (Period 1) .....	183
Table 5.5: Resampling Methods for SVM_CW Strategy (Period 1) .....	185
Table 5.6: Tracking Quality in the Estimation and Investment Periods (Traditional Methods) .....	186
Table 5.7: Tracking Qualities in the Estimation Period (STEM).....	186

Table 5.8: Tracking Qualities in the Investment Period (STEM) .....	187
Table 5.9: Aggregated Statistics for Traditional and STEM-Based Methods in the Investment Period (ETQ) .....	188
Table 5.10: Comparison of Levels of Tracking Quality over Different Time Periods .....	190
Table 5.11: Comparison of Ratios of Investment Period to Estimation Period Tracking Quality .....	191
Table 5.12: Levels of Tracking Quality over the Five Time Periods (Best 10% of 5,000) ...	192
Table 5.13: Ratios of Tracking Qualities (5,000/Best 10%) .....	192
Table 5.14: Stability Ratios for the STEM-Based Methods.....	193
Table 5.15: Kendall Rank Correlation Coefficients for the STEM-Based Methods .....	193
Table 5.16: Spearman Correlation Coefficients for the STEM-Based Methods.....	194
Table 5.17: Superior Methods, According to Relative Stability .....	194
Table 5.18: Ex Post Tracking Quality for the STEM-Based Methods.....	195
Table 5.19: Comparison of Ex Post Tracking Quality (Best 10%) .....	195
Table 5.20: Ratios of Ex Post Tracking Quality (5,000 / Best 10%) .....	196
Table 5.21: Stability Ratios for ETQ .....	196
Table 5.22: Kendall Rank Correlation Coefficients for ETQ .....	197
Table 5.23: Spearman Rank Correlation Coefficients for ETQ .....	197
Table 5.24: Best Methods Based on a Comparison of Median ETQs.....	198
Table 5.25: Comparison Table for the Significance Tests .....	198
Table 5.26: Best Methods as Indicated by Significance Tests Comparing ETQs.....	198
Table 5.27: Best Methods According to a Comparison of Median RMSEs. ....	199
Table 5.28: Best Methods as Indicated by Significance Tests Comparing RMSEs .....	199

## List of Figures

Figure 2.1: Index universe.....	12
Figure 2.2: Weighting schemas.....	14
Figure 2.3: S&P1500 Index divisor. Source: Datastream .....	17
Figure 2.4: Index type by reinvestment policy.....	18
Figure 2.5: Operation of a Mutual Fund .....	22
Figure 2.6: The ETF Share Creation Process.....	23
Figure 2.7: Overlapping of Assets from the Tracking Portfolio and the Index .....	26
Figure 2.8: Options Strategy for Tracking an Index (Synthetic Long Index).....	29
Figure 2.9: Options Strategy for Tracking an Index (Long Combo) .....	29
Figure 2.10: Index Swap .....	30
Figure 2.11: Equity-linked note .....	30
Figure 2.12: Unfunded Swap .....	31
Figure 2.13: Funded Swap I.....	31
Figure 2.14: Funded Swap II.....	32
Figure 2.15: Combination of Substitute Portfolio and Swap .....	33
Figure 2.16: Index and Index Inverse.....	33
Figure 2.17: White Box and Black Box Approaches to Index Tracking .....	40
Figure 2.18: Which Method of Index Tracking should be used?.....	42
Figure 3.1: Overview of Sampling Methods.....	45
Figure 3.2: Stratified Sampling .....	48
Figure 3.3: Hierarchical Clustering.....	48
Figure 3.4: Matching the Factor Structure of the Index .....	49
Figure 3.5: Difference between MAD and MSE.....	59
Figure 3.6: Weakness of the TEV .....	60
Figure 3.7: Returns of the DAX and Two Tracking Portfolios with Equal MADs .....	61
Figure 3.8: Values of a DAX and Two Tracking Portfolios with Equal MADs.....	62
Figure 3.9: Returns of the DAX and Two Tracking Portfolios with Equal MSEs .....	62
Figure 3.10: Values of the DAX and Two Tracking Portfolios with Equal MSEs.....	63
Figure 3.11: Ex Post Tracking Quality Shown as the Area between Two Curves .....	64
Figure 3.12: MAXD and CVAR .....	69
Figure 3.13: Weights versus Numbers .....	71
Figure 3.14: Prediction Errors .....	77
Figure 3.15: Prediction Errors Based on Relative Quality in the Investment and Estimation Periods.....	77
Figure 3.16: Illustration of the Two Types of Error.....	78
Figure 3.17: The Two Types of Error Depicted in a Dispersion Table .....	78
Figure 3.18: Three Cases of Relative Stability .....	80

Figure 3.19: Type I and II Errors Together .....	80
Figure 3.20: Stability Ratio .....	83
Figure 4.1: S&P1500.....	88
Figure 4.2: Comparison of Methods for the Approximation of Portfolio Returns using Wilcoxon Test .....	90
Figure 4.3: Histograms for each of the Approximation Procedures (Period 1, Size 240, ETQ) .....	94
Figure 4.4: Stability of the Median (ETQ).....	96
Figure 4.5: Scatter Plots for Traditional Measures, for Tracking Portfolios of Size 120 .....	108
Figure 4.6: Comparing Traditional Measures with ETQs.....	113
Figure 4.7: Stock Selection for the Tracking Portfolio .....	115
Figure 4.8: Design of the Experiment .....	116
Figure 4.9: Example of the Sensitivity of the Tracking Portfolios .....	118
Figure 5.1: Consistency of the learning process .....	125
Figure 5.2: a) Separating Hyperplane b) Separating Hyperplane is not Unique c) Optimal Separating Hyperplane .....	127
Figure 5.3: Optimal Separating Hyperplane and Support Vectors.....	129
Figure 5.4: SVM for Non-Separable Case .....	130
Figure 5.5: a) Non-linear Separating in Input Space b) Linear Separating in Feature Space .....	132
Figure 5.6: $\epsilon$ -Insensitive Loss Function .....	135
Figure 5.7: Support Vector Regression .....	136
Figure 5.8: Training Set and Test Set.....	138
Figure 5.9: True Value and Estimated Values .....	139
Figure 5.10: VC Dimension for Indicator Functions .....	142
Figure 5.11: Margin Hyperplane with a Lower VC Dimension .....	143
Figure 5.12: SRM and Bounds on Test Error.....	145
Figure 5.13: Correction Factors for Complexity .....	147
Figure 5.14: Reducing Complexity for the Hyperplane.....	148
Figure 5.15: Measuring the VC Dimension .....	150
Figure 5.16: Hold Out .....	151
Figure 5.17: K-Fold Cross-Validation .....	151
Figure 5.18: Model Selection Process.....	152
Figure 5.19: Conversion of SVR into SVM: a) Original Data and Regression Line b) Shifted Data and Separating Plane.....	155
Figure 5.20: Precision and Margin for SVR .....	158
Figure 5.21: Comparison of Different Measures of Tracking Quality.....	159
Figure 5.22: SVR as a Relaxation of the Minimax Approach .....	162
Figure 5.23: Ellipsoidal and Spherical Constraints in 2-Dimensional Space .....	164
Figure 5.24: Spherical Constraint in 3-Dimensional Space .....	164
Figure 5.25: The Nested Structure for STEM .....	167

Figure 5.26: $\varepsilon$ -Insensitive Loss Function with a Quadratic Loss .....	169
Figure 5.27: $\varepsilon$ -Insensitive Loss Function with a Downside Loss.....	170
Figure 5.28: 1-Norm Constraint on the Portfolio Weights.....	172
Figure 5.29: Estimation Error for the Vector of Expected Returns .....	174
Figure 5.30: Box Constraints .....	177

## Notation

<b>A</b>	bold capital letter is a matrix	<b>A'</b>	matrix transpose
<b>a</b>	bold small letter is a vector	<b>a'</b>	vector transpose
$V_{I,t}$	value of index at time $t$		
$V_{P,t}$	value of portfolio at time $t$		
$S_{i,t}$	price of asset $i$ at time $t$		
$R_{I,t}$	return of index at time $t$		
$R_{P,t}$	return of portfolio at time $t$		
$r_{i,t}$	return of asset $i$ at time $t$		
$w_i$	weight of asset $i$		
$n_i$	number of units of asset $i$		
$t = 1, \dots, T$	time periods		
$i = 1, \dots, N$	assets in the tracking portfolio		
$Pr$	probability		
$E$	expectation		
$\Sigma$	covariance matrix of assets returns		
$\bar{\mu}$	vector of expected assets returns		
<b>S</b>	squared root of covariance matrix of assets returns		
$d$	degrees of freedom		
$h$	VC dimension		

# 1 Introduction

At the present time, there is a major trend in the financial markets of focusing on the common factors rather than the individual characteristics of assets. This tendency can be seen in the growing popularity of investing in indices. In June 2011, the total value of indexed assets under management worldwide was \$5.994 trillion,<sup>1</sup> 25% higher than the previous year (\$4.781 trillion).<sup>2</sup> The reason for this growth in popularity is that the portfolios, which replicate the performance of the indices, simplify the investment process and, at the same time, act as building blocks that can be used to construct larger investment portfolios that meet the objectives of investors.

The idea of indexing, of investing in a portfolio that covers the market or a segment of it, appeared in the early 1970s, as an answer to academic publications that had argued that it is difficult or even impossible to systematically beat the comprehensive stock market averages.<sup>3</sup> Professor *Burton Malkiel* wrote, in 1973, in his book *A Random Walk Down Wall Street*:

*What we need is a no-load, minimum-management-fee mutual fund that simply buys the hundreds of stocks making up the broad stock-market averages and does no trading from security to security in an attempt to catch the winners. Whenever below-average performance on the part of any mutual fund is noticed, fund spokesmen are quick to point out, "You can't buy the averages." It's time the public could.*<sup>4</sup>

Nobel laureate *Paul Samuelson* continued along this line in 1974, in his article *Challenge to Judgment* in the *Journal of Portfolio Management*:

*What logic can demonstrate is that not everybody, nor even the average person, can do better than the comprehensive market averages. That would contradict the tautology that the whole is the sum of its parts. If you select at random a list, of say 100 stocks and buy them with weights proportional to their respective total outstanding market values, although your sample's performance will not exactly duplicate that of a comprehensive market average, it will come close to doing so – closer than if you throw a dart at only one stock, but of course not quite as close as with a sample of 200, 300, or all the stocks available in the marketplace.*<sup>5</sup>

---

<sup>1</sup> See Olsen (2011).

<sup>2</sup> See Kozlowski (2010).

<sup>3</sup> See Jensen (1968).

<sup>4</sup> See Malkiel (2007, p. 359).

<sup>5</sup> See Samuelson (1974, p. 18).

These ideas came into being on December 31, 1975,<sup>6</sup> when *John Bogle* founded *First Index Investment Trust*, the first index fund for retail investors. Recently, after more than thirty years of index investing, *John Bogle* systematized the main requirements of indexing:

*[T]he most effective means of building wealth is simply to emulate the annual returns provided by the financial markets, and reap the benefits of long-term compounding. This goal, as it turns out, can best be achieved by minimizing the costs of investing—sales commissions, advisory fees, taxes, and the like—and seeking to earn the highest possible portion of the annual return earned in each sector of the financial markets in which you invest, recognizing, and accepting, that that portion will be less than 100%. To achieve this goal, the ideal investment program includes four elements:*

*Simplicity: Matching, not beating, the markets; asset allocation that is strategic, not tactical.*

*Focus: Maximizing the productive economics (earnings and dividends, interest yields) of investing; minimizing the counterproductive emotions of investing (changing price-earnings ratios).*

*Efficiency: Economical operations; minimization of the frictional costs of fees and commissions and taxes.*

*Stewardship: Placing the interest of the client first; unyielding emphasis on human beings and eternal values—integrity, honesty, candour.<sup>7</sup>*

These four elements, simplicity, focus, efficiency, and stewardship are taken into account by index investing. Moreover, the investment portfolios for index investing are highly diversified portfolios, which allow investors to obtain the average market return at a minimum costs. This is in contrast to the costs that arise when a market participant tries to obtain an excess return over the market. In this case, the transaction costs and costs of research are much higher, because more trading and more research are needed to obtain returns above the market average. Hence, the goal of investing in the index is to invest in a diversified portfolio that gives the same return as the market while minimizing transaction costs.

However, to invest into the selected index one needs to construct the portfolio, which approximates this index. This is called index tracking. Hence, the solution of the index tracking problem is a portfolio that approximates the index. This portfolio is called the tracking portfolio. It is easy to build a tracking portfolio on paper: simply buy the stocks in the same proportions as they appear in the index and hold this portfolio. However, there are some pitfalls in practice. The computation of an index value does not take into account transaction costs, the liquidity of the assets, or taxes. These factors can reduce the gains from indexing. To overcome these problems, various approaches have been developed.

---

<sup>6</sup> See Bogle (2001, p. 98).

<sup>7</sup> See Bogle (2001, p. xvii).



These approaches have their strengths and weaknesses. By selecting the method to track the index, investor evaluates the index tracking approaches using the following criteria: applicability of the tracking method, set up costs for a tracking portfolio, structural tracking error, maintenance costs, tax efficiency, flexibility of tracking portfolio, transparency of tracking method, and counterparty risk. As we show in this thesis, the simple replication of the index by purchasing the same fixed quantities of assets as are in the index is not always a suitable and applicable method. For example, if the index contains a group of illiquid assets, then transaction costs can increase significantly. Moreover, this method requires a large amount of initial capital. It is possible to reduce costs and the initial capital by decreasing the number of assets in the tracking portfolio. Based on this example, we can state the goal of the work.

The goal of this work is to make a comparison of the approaches to index tracking, highlighting their advantages and disadvantages, and identifying any problems with their application. On the basis of this comparison, a new method will be developed that improves on the existing ones. To achieve this goal, the research is organized as follows.

Chapter 2 starts with a discussion of the rationale behind indexing. Then, we describe the index construction methodology. After that, the common approaches to index tracking, which include physical (full replication and sampling) and synthetic replication, are presented and evaluated. The evaluation of these approaches indicates that the sampling approach is more flexible than full replication and more transparent than the synthetic approach.

In Chapter 3 we focus on the sampling approach used for index tracking. In this chapter, a literature review of the existing methods of sampling is carried out. Then, we show that there are weaknesses in the traditional measures that are used to evaluate the quality of tracking. In addition, we introduce a quality measure, which is focused on the crucial investment period. After that, we identify the requirements for an effective sampling approach.

Among the requirements for a sampling approach, one can point to the absolute and the relative stability of the tracking quality of the tracking portfolio. This is because, firstly, it is necessary that the sampling method constructs tracking portfolios with similar levels of tracking quality in both the estimation and investment periods (absolute stability of tracking quality). Moreover, it is desirable that those tracking portfolios with high tracking quality in the estimation period compared to alternative tracking portfolios also have high relative tracking quality in the investment period (relative stability of tracking quality). To evaluate the absolute stability, we compute the ratio of the tracking quality in the investment period to the tracking quality in the estimation period. To assess the relative stability, we introduce the stability ratio.

The next requirement for an effective sampling method is the resistance of the tracking portfolio estimation, which means that the tracking portfolio does not change dramatically when new information arises and, if the estimation period contains some less reliable data, then the optimized sampling method should not produce a result that is overoptimized towards these data.<sup>8</sup> Resistant tracking portfolios are characterized by a low rate of turnover that reduces the transaction costs.

In Chapter 4, the traditional methods of sampling are verified empirically. First, we analyse the problem of how to model the portfolio structure: using weights or numbers of assets. We empirically show that, for a buy-and-hold strategy, there is a need to correct the asset weights in the tracking portfolio over time so as to be consistent with this strategy. Second, we analyse the absolute and relative stability of the tracking quality of tracking portfolios estimated using traditional sampling methods. In this context we empirically show that the tracking quality measure alone is not a good criterion for the selection of tracking portfolios. Third, we empirically show that the traditional sampling methods are not resistant to small changes in the data used to estimate the tracking portfolio weights. Slight modifications of the initial data in the estimation period lead to different estimates of the weights and, as a consequence, to different levels of tracking quality in the investment period.

Chapter 5 contains our contribution to the design of effective methods of sampling. We use techniques from statistical learning theory<sup>9</sup> to overcome the drawbacks of traditional approaches for solving index tracking problem. These techniques include support vector machines for solving classification problems and support vector regression for solving regression problems. We apply these techniques for solving index tracking problem. Moreover, we extend support vector techniques and introduce the structural tracking error minimization (STEM) approach, which is based on the structural risk minimization principle in statistical learning theory. The main idea of this approach is that by adding additional constraints into the optimization problem, which represent the preliminary knowledge about the weights of the tracking portfolio, it is possible to improve the robustness of the solution. Moreover, using this approach it is possible to control the quality of fitting and to avoid the problem of overfitting. Additionally, we empirically show that this approach is able to improve the absolute and the relative stability of the tracking quality and increase the resistance of the tracking portfolio estimation.

The main contributions of this work include:

---

<sup>8</sup> Adapted from Schöttle/Werner (2006, p. 129) to apply to index tracking.

<sup>9</sup> See Vapnik (1998).

- We analysed different methods for index tracking. Each of the methods of index tracking have strengths and weaknesses. However, sampling is a physical replication method that can be applied to any index. It is more flexible than full replication, and more transparent than synthetic. Moreover, it allows the index to be replicated by focusing on a trade-off between transaction costs and tracking error. These properties make it a good alternative.
- By applying sampling, the tracking quality in the investment period can be evaluated by traditional measures, such as MAD, MSE and TEV. However, we have shown, that these measures have weaknesses when used for this purpose. Therefore, we introduced the ex post tracking quality measure, which is the area between the normalized curves of the evolution of the values of the index and the tracking portfolio. Moreover, to evaluate the different methods of optimized sampling, the concepts of absolute and relative stability of tracking quality were introduced. Additionally, the importance of estimating resistant tracking portfolios was discussed.
- We showed empirically that tracking portfolios that are constructed by minimizing the traditional measures of tracking quality lack absolute and relative stability of tracking quality. We showed empirically that tracking portfolio estimation using the current sampling methods is not resistant to small changes in the dataset used for estimation. We showed that slight modifications of the data in the estimation period leads to different tracking portfolios being generated and, as a consequence, to different levels of tracking quality in the investment period.
- We introduced the STEM approach to overcome the drawbacks of the existing methods. This approach uses a capacity control on the set of models, which makes the estimation of the tracking portfolio resistant to the data used for the estimation. This helps to solve the problem of overfitting. STEM-based methods form a group of methods for index tracking, combining the data with preliminary knowledge expressed in informational structures. These methods have roots in statistical learning theory and can be based on different measures of tracking quality. We conducted an empirical comparison between the STEM-based methods and the traditional methods. The STEM-based methods extend the properties of the other methods. The empirical results show that, by using STEM-based methods, we can improve the resistance of the tracking portfolio estimation by controlling the complexity of the set of models. Moreover, the STEM-based methods can be superior to the traditional methods in terms of their tracking quality in the investment period. Additionally, we have shown that the STEM-based methods generate tracking portfolios with a better absolute stability of tracking quality than those produced by traditional methods. If we use an equal weighted (EW) structure, it is also possible to improve the relative stability.

## **2 Foundations of index tracking**

Recently, complex index-based strategies have emerged to extend the classical strategy of investing in a broad market index. This has become a stimulus for the introduction of new indices, leading in turn to new ways of investing in and tracking the indices.

Hence, at the beginning of this chapter we discuss index-based strategies. After that, we explore the indices and their construction methodologies and describe the different types of index funds. Finally, we compare and evaluate tracking approaches, and assess their strengths and weaknesses.

## 2.1 Index-based investment

Index-based investment has become an integral part of asset management. The indices most widely used for investment are broad market indices. Investment in a broad market index is the main choice used by those wishing to have a stake in the development of the whole market. For example, investing in the S&P1500 index allows one to make returns based on the greater part of the US stock market.

There are four main reasons why such investment activity is advantageous for investors:

- The first is diversification.<sup>10</sup> The market index is a highly diversified investment portfolio. The interdependence between the assets within this portfolio reduces the influence of individual assets on its overall performance.
- The second reason is that it is difficult to earn a return that exceeds the return on the market, as represented by the broad market index.<sup>11</sup>
- The third reason is that investing in the market is a zero-sum game relative to the market average, which is represented by the market index. This means that the sum of gains and losses, which are calculated relative to the market average, among all participants in the market, is equal to zero.<sup>12</sup> Hence, an excess return over the market average can only be obtained at the expense of the other market participants. As a consequence, there are market participants who obtain an above-market return, and there are those who receive a below-market return. Indexing allows for all market participants to obtain the average market return.
- The fourth reason is that the costs of investing in the index are low. This is in contrast to the costs that arise when a market participant tries to obtain an excess return over the market. In this case, the transaction costs and costs of research are much higher, because more trading and more research are needed to obtain returns above the market average.<sup>13</sup>

Hence, the goal of investing in the index is to invest in a diversified portfolio that gives the same return as the market while minimizing transaction costs.

---

<sup>10</sup> See, for example, Malkiel (2007, pp. 179-214). For a comparison of different diversification strategies, see DeMiguel/Garlappi/Uppal (2009).

<sup>11</sup> See, for example, Sharpe (1966), Treynor/Mazuy (1966), and Jensen (1968).

<sup>12</sup> This statement can be found in: Samuelson (1974, p. 18), Ellis (1975, p. 23), Wagner (2002, p. 814), Malkiel (1997, p. 3), Harris (1993), and Siegel (2008, p. 342).

<sup>13</sup> See Sharpe (1991, pp. 7-9) and Fuller/Han/Tung (2010, p. 40).

Additionally, investing in a broad market index is supported by modern portfolio theory. *Markowitz* introduces the risk/return trade-off and the concept of the mean-variance efficiency of portfolios.<sup>14</sup> Using this concept and adding a cash asset (risk-free asset), *Tobin*<sup>15</sup> proves the two funds separation theorem, which states that the investor, to maximize his utility, divides his wealth between some risky portfolio and the risk-free asset. Moreover, the proportions of the assets within the risky portfolio are always the same, and independent of the sum invested in the cash asset. *Sharpe*, *Lintner* and *Mossin* generalize the two funds separation theorem to all investors and, as a consequence, develop the Capital Asset Pricing Model (CAPM).<sup>16</sup>

This theoretical model claims that, under very strict assumptions, all investors hold in equilibrium the mean-variance efficient portfolio, which is called the market portfolio, with the same proportions of risky assets. If this is not the case, then the investment is not optimal because investment in a portfolio other than the market portfolio produces a worse return/risk ratio than combining the market portfolio with the cash asset. As a consequence, if all investors hold the same portfolio, then the market portfolio can be represented by the broad market capitalization-weighted index.

The statement that investing in the broad market capitalization-weighted index is the only correct choice is widely criticized. First, the usage of the broad market index as a proxy for the market portfolio cannot be justified, since there are other assets, for example human capital and real estate, that must be taken into account.<sup>17</sup> Second, empirical studies show that the broad market indices are not mean-variance efficient.<sup>18</sup> Third, the presence of taxes changes the preferences of investors. For example, pension funds, which are not sensitive to taxes, prefer stocks with higher dividend rates, while individuals, who are sensitive to taxes, invest in stocks that provide higher capital gains.<sup>19</sup> These criticisms show that there could be other ways of meeting investment objectives.

Studies by *Fama* and *French* demonstrate empirically that there are other important factors, besides risk, that influence the return on assets.<sup>20</sup> These factors are the size of the company and its value. The size of the company is expressed in terms of market capitalization, with

---

<sup>14</sup> See Markowitz (1952, pp. 77-91).

<sup>15</sup> See Tobin (1958, p. 84).

<sup>16</sup> The CAPM was developed by Sharpe (1964, pp. 425-442), Lintner (1965, pp. 13-37), and Mossin (1966, pp. 768-783).

<sup>17</sup> See Roll (1977, p. 130) and Goltz/Le Sourd (2010, p. 6).

<sup>18</sup> See, for example, Grinold (1992), who checks five markets: the United States, the United Kingdom, Australia, Japan, and Germany.

<sup>19</sup> See Miller (1978, p. 142).

<sup>20</sup> See Fama/French (1992, pp. 427-465) and Fama/French (1993, pp. 3-56).

companies usually categorized into small, medium, or large. The value of the company is expressed by the book-to-market ratio. Using this ratio it is possible to assign the company to one of two classes: it is called a value company if the ratio is low and a growth company if the ratio is high. Empirical results show that growth and value stocks behave differently, and that small companies' stocks may provide excess returns over those of large companies.<sup>21</sup> This fact becomes the basis for the so-called style strategies.<sup>22</sup> These strategies can also be implemented on the basis of indices. Currently, indices are developed that reflect the behaviour of the companies, grouped by size or value.

Additionally, there are empirical findings that show that geographical location or membership of a particular industrial sector can explain the returns of a particular company.<sup>23</sup> There are appropriate indices, covering industrial sectors and regions. By investing in these indices, it is possible to implement different investment strategies, for example, the sector rotation strategy,<sup>24</sup> the idea of which is to change the weights of the sectors in the portfolio depending on the business cycle, underweighting weak sectors and overweighting strong ones.

Additional diversification of the portfolio can be achieved by investing in commodity, fixed-income, and real estate markets.<sup>25</sup> These markets are typically uncorrelated with the stock market and differ in performance and risk. Investing in these markets can also be achieved by investing in appropriate indices.

A large number of index-based strategies belong to the group of semi-active strategies. This means they are index-based but have the potential to outperform the broad market index. The goal of these strategies is to earn excess returns over an index but at the same time pay attention to the quality of tracking. In other words, these strategies seek a trade-off between excess portfolio returns and tracking quality. We describe two of them: core-satellite and enhanced indexing.<sup>26</sup>

The core-satellite strategy separates the “beta” of the investment portfolio from its “alpha”. Part of the portfolio is invested in a broad market index. This part produces the “beta”. The other part is invested with the aim of outperforming the index, producing the “alpha”. The excess return of the portfolio (“alpha”) can be obtained by overweighting a certain group of

---

<sup>21</sup> See Fama/French (1992, pp. 427-465) and Malkiel (1997, pp. 2-17).

<sup>22</sup> See Sharpe (1992, pp. 7-19).

<sup>23</sup> See Strongin et al. (2002) and Cavaglia/Moroz (2002, p. 94).

<sup>24</sup> See Jhirad/Ozkul/Qian (2004, p. 317).

<sup>25</sup> See Gorton/Rouwenhorst (2005).

<sup>26</sup> See, for example, Michalik (2008, pp. 52-57), Adams/Maeda/Schoenfeld (2004, p. 492), and Sorensen/Miller/Samak (1998, pp. 18-31).

stocks, for example, by investing in a more prospective sector. The overweighting or underweighting can also apply to individual assets, for example using fundamental analysis to determine stocks with performance potential.<sup>27</sup> The core-satellite strategy can be extended by the so-called tax-optimized core, in which the core is designed to offset the capital gains from the satellite.<sup>28</sup> This strategy is widely used by professional investors.

There are also enhanced indexing strategies of both non-statistical and statistical natures.

The non-statistical enhanced indexing strategies are based on using a form of arbitrage opportunity to enhance the return of the portfolio over the return of the broad market index. There are different enhancement possibilities, for example convertible securities arbitrage: buying convertible securities and short selling the common stocks of the same issuer. The arbitrage opportunity is the difference between the current price of the stock and the price of the convertible security. Another example is futures/index arbitrage: selling the futures on the index and constructing a tracking portfolio to track this index. The arbitrage opportunity is the difference between the futures price and the value of the tracking portfolio.<sup>29</sup>

The first statistical enhanced indexing strategy is tracking an artificial index.<sup>30</sup> Using this strategy, a base index is selected and a new artificial index is constructed as the sum of the base index return plus a moderate positive constant value. Then, using optimization, an investment portfolio is selected to replicate this new artificial index. There is an expectation that this investment portfolio will produce higher returns than the base index.

The second statistical strategy is a market-neutral one, where two artificial indices are defined: the return of the base index plus a constant value and the return of the base index minus a constant value.<sup>31</sup> Then, two portfolios are constructed, replicating these two indices. The first portfolio produces a higher return, and the second a lower return, than the market. The assets in the first portfolio are bought and the assets in the second portfolio are sold, without coverage. This allows the investor to increase the portfolio value due to the growth in stock prices in the first part of the portfolio and the fall in stock prices in the second part. Hence, this strategy extends the return of the index in both directions.

---

<sup>27</sup> See Loftus (1997, pp. 35-53).

<sup>28</sup> See Quisenberry (2009, pp. 299-310).

<sup>29</sup> See Schoenfeld/Yang (2004, pp. 277-296) and Alexander/Dimitriu (2002).

<sup>30</sup> See Alexander/Dimitriu (2002, p. 2).

<sup>31</sup> See Alexander/Dimitriu (2002, p. 2).



Indexing is not only restricted to meeting investment objectives. It can also be used to perform other portfolio management-related tasks, for example, risk management, hedging, rebalancing, and cash management.<sup>32</sup>

As shown above, different investment strategies can be implemented on the basis of indices. As a consequence, the intended purpose of the indices, to track the performance of a market or one of its segments, is changed. Nowadays, there is a close relationship between indices and investment portfolio management. Indices are used as a basis for index investment strategies and as benchmarks for the evaluation of the performance of active investment strategies. Therefore, in the next section we describe the methodology used to construct indices from the perspective of a portfolio manager.

---

<sup>32</sup> See Hamid/Sandford (2007a).

## 2.2 Index construction

### 2.2.1 Equity indices

#### 2.2.1.1 Construction rules

Equity indices are constructed using predefined sets of rules. These rules are active decisions made by index providers. There are various index providers. The most popular are Standard & Poors, Stoxx, Russell, MSCI, NYSE Euronext, and Deutsche Börse. Generally, the index provider selects the following:<sup>33</sup>

- the index universe;
- the number of assets in the index;
- selection criteria for the assets;
- the weighting schema;
- rules for index readjustments;
- a reinvestment policy for periodic cash flows;
- the treatment of tax.

#### Index universe

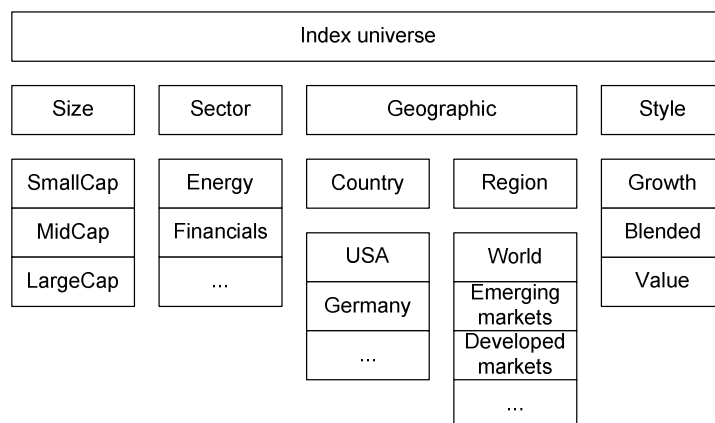


Figure 2.1: Index universe

The first step in index construction is the selection of the index universe. This is a list of assets, from which the index provider draws assets to make up the index. The assets in the list typically have one property in common (see Figure 2.1), for example, size, sector, geographic region, or style.

---

<sup>33</sup> See Fuller/Han/Tung (2010, pp. 36-38) and Maginn et al. (2007, p. 413).

By creating indices with different index universes, the index provider meets two objectives that are important to portfolio managers. The first is that the portfolio manager investing in the index has a choice between different investment strategies. The second is that the portfolio manager can evaluate investment portfolios against the index, which can be used as a benchmark.

### **Number of assets**

The second step in index construction is the selection of the number of assets to be included. The number of assets is one of the measures of index diversification. Here, there is a trade-off. The fewer assets included, the easier it is to invest in the index. However, the index also becomes less diversified and loses its ability to represent the performance of the underlying group of assets.

### **Selection criteria for assets**

The next decision to be made is how to select those assets to be used in the index from the index universe. Predefined selection criteria vary across index providers and indices.

The liquidity of the stocks is an important criterion for their selection in investable benchmarks.<sup>34</sup> If the stocks are illiquid, it will be difficult to use the index as a basis for index-based strategies. The liquidity of the stock is characterized by different factors. Market capitalization is the most important, indicating the investment capacity of the stock. The market capitalization of the company is calculated as the price of the stock multiplied by the number of stocks outstanding. For usage in indices, market capitalization is typically corrected using the free float factor, to give the percentage of market capitalization available for trading, since some stocks are held by company owners or management and are not publicly traded. The next factor influencing stock liquidity is the tradability of the stock. This is expressed in terms of traded volume, listing on exchanges, and price level. Hence, the index providers select companies for inclusion in indices based on whether they satisfy liquidity constraints, which in turn are based on the above factors.

As well as liquidity, there are other criteria for inclusion in indices. For inclusion in the S&P500, S&P MidCap 400, and S&P SmallCap 600 indices, companies should have had four consecutive quarters of positive earnings.<sup>35</sup> Companies are eligible for inclusion in the DAX, MDAX, SDAX and TecDAX indices if they are listed in the Prime Standard segment of the Deutsche Börse, which imposes special requirements on the reporting and viability of

---

<sup>34</sup> See Neubert (1997, p. 61).

<sup>35</sup> See S&P (2009, p. 6).

companies.<sup>36</sup> In contrast, the rules for inclusion in the Russell indices are less restrictive; these indices include common stocks of US companies and have no special requirements regarding the companies' viability.<sup>37</sup>

The selection criteria are important because they determine the differences between indices. For example, the performance of one index, based on the 500 companies from the S&P500 index, will be significantly different from that of another, based on the 500 companies in the Fortune 500 list.<sup>38</sup>

## Weighting schema

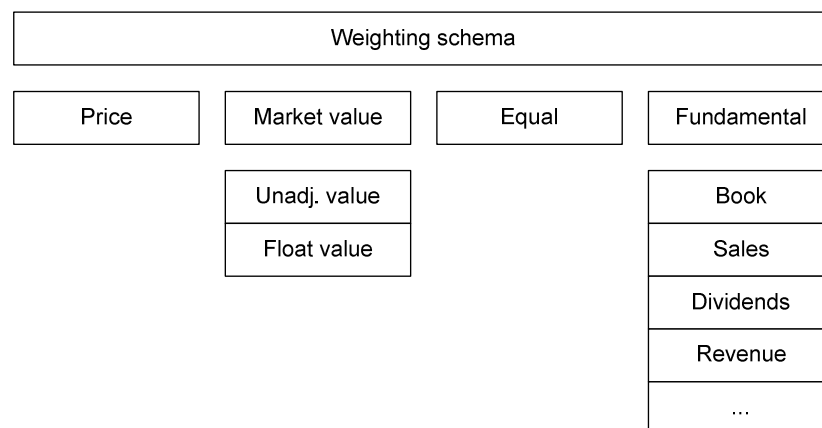


Figure 2.2: Weighting schemas

The next choice made by the index provider is the weighting of the stocks. There are various weighting schemas (see Figure 2.2):

- price weighting;
- market value (capitalization) weighting;
- equal weighting;
- fundamental weighting.

The oldest weighting schema is price weighting, which is computed as an arithmetical average of the assets' prices in the index. It was introduced on July 3, 1884, with the calculation of the Dow Jones Average index. This index is still price weighted and is now called the Dow Jones Industrial Average (DJIA). The advantage of the price-weighting schema is its simplicity. Its main disadvantage is that price does not represent the market

<sup>36</sup> See Deutsche Börse (2011, p. 20).

<sup>37</sup> See Russell (2011, pp. 3-6).

<sup>38</sup> See Arnott/Kuo (2011, pp. 36-46).

value of a company. As a consequence, a small company with a high stock price will influence the index more than a large company with a low stock price.

Capitalization weighting is a more comprehensive approach. It uses additional information about the numbers of stocks outstanding. In recent times, the capitalization-weighting schema has been extended to a free-float-weighting schema that reduces the group of stocks to those available on the market. As was mentioned earlier, the capitalization-weighted schema is supported by the CAPM. It has both advantages and disadvantages. Advantages include the fact that constant numbers of units of assets, rather than weights, are included in the index. Hence, it is possible to build a tracking portfolio that does not need to be rebalanced.<sup>39</sup> Moreover, the most liquid companies receive greater weight in the index, since market capitalization is an indicator of liquidity and investment capacity.<sup>40</sup> However, the main problem with this schema is that it causes a shift towards overpriced stocks, which should be avoided when investing.<sup>41</sup>

Another approach is the equal weighting of stocks. This schema includes all stocks in the index with the same weight. The advantage is that all stocks influence the index equally. This is equivalent to the naïve diversification of the investment portfolio.<sup>42</sup> However, there are some disadvantages in weighting assets equally. The first is that the small companies influence the index as much as the large companies.<sup>43</sup> Moreover, a liquidity problem arises for large equal-weighted indices, since investments in small companies need to be made at the same level as they are made in large companies.<sup>44</sup> Additionally, the numbers of each stock included in the index are not constant; they vary with the prices of the stocks to keep their weights constant.

Fundamental weighting of assets has appeared recently.<sup>45</sup> The assets in the index are weighted based on fundamental factors, for example, book value, revenue, sales and dividends.<sup>46</sup> This weighting schema has the same disadvantage as equal weighting: the numbers of stocks used are not constant. The advantage of this schema is that only fundamental factors are used to

---

<sup>39</sup> Details are discussed in Section 3.3.4.

<sup>40</sup> See Arnott/Hsu/Moore (2005, p. 84).

<sup>41</sup> See Miller (1978, pp. 17-28), Hsu (2006), and Arnott/Hsu/Moore (2005, p. 84).

<sup>42</sup> See DeMiguel/Garlappi/Uppal (2009).

<sup>43</sup> See Arnott/Hsu/Moore (2005, p. 84).

<sup>44</sup> See Arnott/Hsu/Moore (2005, p. 84).

<sup>45</sup> See Arnott/Hsu/Moore (2005, p. 83).

<sup>46</sup> See Arnott/Hsu/Moore (2005, p. 83).

determine the weights of the stocks and, as a consequence, overpriced stocks have less of an effect on the index value.<sup>47</sup>

There are also other, less widely-used weighting schemas. For example, the weights can be determined by optimization procedures. One can compute the minimal variance portfolio, and use the weights in this portfolio to weight the assets in the index.<sup>48</sup> Another method of calculating index weights is the maximization of the Sharpe ratio (the excess return of the portfolio divided by its standard deviation). The solutions of both optimization problems are portfolios that lie on the efficient frontier of risky assets. As a consequence, an index using such weighting schemas is mean-variance efficient, in contrast to the market capitalization-weighted indices, which are mean-variance inefficient, as shown by empirical tests.<sup>49</sup>

Sometimes, additional restrictions are placed on the weights of the companies in an index. For example, in DAX indices, the weight of an individual share should be less than 10%.<sup>50</sup> This limits the influence of individual companies on the index. However, during the computation of the index, this restriction needs to be taken into account.

## **Index readjustments**

Additions and deletions of companies are one of the reasons why index readjustments might be needed. A deletion might happen if a company filed for bankruptcy or no longer met the criteria for inclusion in the index. Analysing Table 2.1, we can conclude that additions and deletions are not rare events. The broad market indices, for example the S&P1500, and small capitalization indices, for example the S&P Small 600 and the SDAX 50, are more susceptible to changes than the other indices.

---

<sup>47</sup> See Siegel (2008, p. 354).

<sup>48</sup> See Greifeneder (2011).

<sup>49</sup> See Amenc et al. (2011).

<sup>50</sup> See Deutsche Börse (2011, p. 12).

	S&P 500	S&P Mid 400	S&P Small 600	S&P 1500	DAX 30	TecDAX 30	MDAX 50	SDAX 50
<b>2000</b>								
additions	58	87	86	231	4	-	9	-
deletions	57	87	86	230	4	-	8	-
<b>2001</b>								
additions	30	54	75	159	2	-	12	-
deletions	30	54	75	159	2	-	12	-
<b>2002</b>								
additions	24	28	47	96	2	-	11	10
deletions	23	28	47	96	2	-	12	10
<b>2003</b>								
additions	9	33	48	63	1	3	8	43
deletions	9	33	48	63	1	3	28*	42
<b>2004</b>								
additions	22	37	59	90	0	7	5	12
deletions	22	37	59	90	0	7	5	12
<b>2005</b>								
additions	20	30	54	91	1	5	5	12
deletions	20	31	54	92	1	5	5	12
<b>2006</b>								
additions	33	31	62	91	1	4	6	17
deletions	33	31	62	90	1	4	6	17
<b>2007</b>								
additions	42	61	88	121	1	5	6	10
deletions	41	60	88	120	1	5	6	10
<b>2008</b>								
additions	37	55	83	114	3	7	8	10
deletions	37	55	83	116	3	7	8	10
<b>2009</b>								
additions	19	26	40	55	4	6	7	11
deletions	19	26	40	56	4	6	7	11

Table 2.1: Index additions and deletions. Source: Standard & Poor's, Deutsche Börse.

\*MDAX has changed its composition from 70 constituents to 50.

There are different rules specifying when changes take effect. The Russell applies changes to its indices once a year.<sup>51</sup> The Deutsche Börse makes a review of the DAX indices once a quarter.<sup>52</sup> Standard & Poor's applies changes to its indices as they are needed.<sup>53</sup>

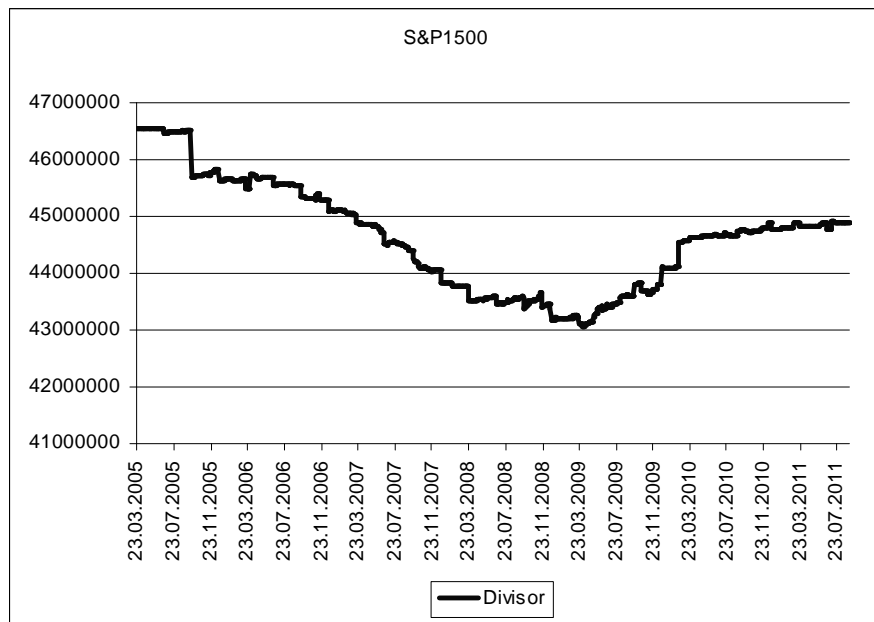


Figure 2.3: S&P1500 Index divisor. Source: Datastream

Additions and deletions are not the only source of readjustments. Corporate actions that change market capitalization also contribute to changes in the market-weighted indices. The

<sup>51</sup> See Russell (2011, p. 2).

<sup>52</sup> See Deutsche Börse (2011, p. 43).

<sup>53</sup> See S&P (2009, p. 6).

dynamic of these actions can be demonstrated using the values of the index divisor<sup>54</sup> for the S&P1500 index (see Figure 2.3). The index divisor is used to compensate for changes in the numbers of stocks of a company, so as to hold the value of the index fixed.

Corporate actions that cause changes in the index divisor include the following:<sup>55</sup>

- company added/deleted;
- change in shares outstanding;
- spin-off;
- change in float factor;
- special dividend;
- rights offering.

### Reinvestment policy for periodic cash flows and tax treatment

According to the reinvestment policy for periodic cash flows, indices are divided into price, total return (performance), and net return indices (see Figure 2.4). Price indices only reflect the movements of stock prices. Total return indices also include the dividends. Net return indices are adjusted for taxes as well as including dividends. Some indices have several forms. For example, DAX indices can be calculated as either price or total return indices. The S&P500 is a price index, but there is also the S&P500 dividend index. By combining the two, it is possible to calculate the S&P500 total return index. Russell indices are total return indices.

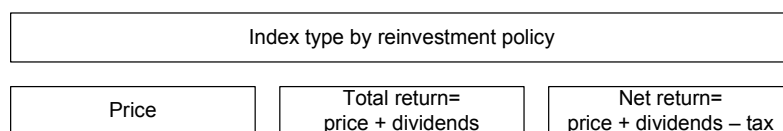


Figure 2.4: Index type by reinvestment policy

There are two possibilities for reinvesting dividends: at the stock level, where the dividends from the stock are fully invested in the same stock, and at the index level, where the dividends from one stock are reinvested in all the stocks in the index, in proportion to their weights. The reinvestment of dividends is typically done on the ex-date<sup>56</sup>, not on the pay date. Hence, the prices of the stocks reflect the dividend payments.<sup>57</sup>

<sup>54</sup> Details about the index divisor are given in Section 2.2.1.2.

<sup>55</sup> See S&P (2008, p. 9).

<sup>56</sup> The ex-date is the date when the security is traded without the announced dividends.

<sup>57</sup> See Russell (2011, p. 21).



The tax treatment varies between indices and index providers. For example, the Standard & Poor's and Russell families of indices do not include taxes. However, Stoxx assumes the reinvestment of dividends under the highest tax rate, which depends on the country, in calculating the net return indices. For example, the tax rate for Germany was 26.38% in 2011.<sup>58</sup>

### 2.2.1.2 S&P1500

To illustrate the methodology used to construct indices, we will use the S&P1500 index as an example. This index will be used throughout the thesis, since it is a broad market index, containing stocks with different levels of market capitalization and from different industries. Hence, it is useful for showing the problems that can arise when tracking indices of such types.

The S&P1500 is a price index, covering the entire US market. It combines a large cap index, the S&P500, a mid cap index, the S&P400, and a small cap index, the S&P600.<sup>59</sup> The main decisions made by the index provider for this index are shown in Table 2.2.

index universe	U.S. Companies large cap, mid cap and small cap
number of assets in index	1500
selection criteria for assets	Assets with large capitalizaion within its group, where free float factor is larger than 50%
weighting schema	float weighting
time for index readjustments	if required
reinvestment policy for periodic cash flows	No
treatment of tax	No

Table 2.2: S&P1500 Methodology

The index value  $V_{I,t}$  at time  $t$  is calculated as follows:<sup>60</sup>

$$V_{I,t} = \sum_{i=1}^N \frac{n_{i,t}}{D_t} S_{i,t},$$

---

<sup>58</sup> See Stoxx (2011a, p. 14) and Stoxx (2011b).

<sup>59</sup> See S&P (2009, p. 3).

<sup>60</sup> See S&P (2008, p. 5).

where  $n_{i,t}$  is the free float number of stock  $i$ ,  $S_{i,t}$  is the price of stock  $i$ , and  $D_t$  is the index divisor.

The weight of stock  $i$  at time  $t$  is calculated as follows:

$$w_{i,t} = \frac{\frac{n_{i,t}}{D_t} S_{i,t}}{\sum_{j=1}^N \frac{n_{j,t}}{D_t} S_{j,t}} = \frac{n_{i,t} S_{i,t}}{\sum_{j=1}^N n_{j,t} S_{j,t}}.$$

Hence, the stock weight can be calculated using the free-float-adjusted market capitalization, without knowing the value of the index divisor.

The mechanics of the index divisor can be illustrated using an example. Suppose the number of stocks of company  $i$  in the index changes from  $n_{i,t}$  to  $n'_{i,t}$  as a result of corporate actions. However, the performance of the market should stay unchanged. Hence, to maintain the value of the index at the same level, the index divisor must be adjusted so that:

$$\sum_{i=1}^N \frac{n_{i,t}}{D_t} S_{i,t} = \sum_{i=1}^N \frac{n'_{i,t}}{D'_t} S_{i,t}$$

This means that the new value of the divisor  $D'_t$  is:

$$D'_t = D_t \frac{\sum_{i=1}^N n'_{i,t} S_{i,t}}{\sum_{i=1}^N n_{i,t} S_{i,t}}$$

As a result, the divisor reflects the change in the number of stocks.

### 2.2.2 Alternative indices

There are also some alternative indices. The most popular among them are the fixed-income and commodity indices.

Fixed-income indices are composed of bonds. They can be based on government bonds, for example Barclays Capital Euro Government Bond Indices and FTSE UK Conventional Gilts, or corporate bonds, for example the Markit iBoxx Corporate Bond Index and Barclays Capital Euro Corporate Bond Index. These indices are weighted using market debt, which is similar to using market capitalization for equity indices.<sup>61</sup> However, for debt indices, this weighting

---

<sup>61</sup> See Barclays (2008).

schema can lead to the bums problem, where low-quality bonds with a lot of outstanding debt are heavily weighted in the index, but high-quality bonds with little outstanding debt are not.<sup>62</sup>

Commodity indices, which represent investment in commodities such as gold, copper or sugar, can be constructed based on commodity spot or futures prices. For example, the DJ-UBSCI, RICI, and S&P GSCI indices are all based on futures contracts. The weighting schema used for the S&P GSCI index is a world-production weighting. The schema used for DJ-UBSCI is based on liquidity and production.<sup>63</sup>

There are also some less popular indices, which track real estate, hedge funds, inflation, private equity, and so on. These indices are typically not transparent and have structures that make it more difficult to invest in them. However, it is possible to invest in them. Tracking methods have been developed that can be used to track their performance.

---

<sup>62</sup> See Goltz/Campani (2011, p. 19).

<sup>63</sup> See DJUBS (2011, p. 2).

## 2.3 Types of index funds

For a typical portfolio manager there are different ways to invest in indices. One way is to build and maintain a tracking portfolio. Another way is to invest in an index fund. In this case, the manager of the index fund has to solve the tracking problem. Hence, for every existing index product there must be a tracking of the index.

Index funds can be categorized into open-end funds (mutual funds) and exchange-traded funds (ETFs).<sup>64</sup> The fund invests in a diversified portfolio of securities and sells parts of this portfolio to investors in the form of shares. Mutual funds offer shares on the over-the-counter (OTC) market, while shares in ETFs can be purchased on the exchange. Hence, ETFs are traded intraday, while mutual funds are traded at most once a day.<sup>65</sup>

Depending on the tracked index, funds can be accumulative or non-accumulative. Accumulative funds reinvest dividends, while non-accumulative funds distribute them to the investors.

Funds are measured by their net asset value (NAV), which is the real value of the shares in the fund. The NAV is calculated as the ratio of the market value of the portfolio minus the liabilities to the number of shares outstanding:<sup>66</sup>

$$NAV = \frac{\text{market value of portfolio} - \text{liabilities}}{\text{number of shares outstanding}}.$$

The mutual fund manages the sale and purchase of shares itself (see Figure 2.5).

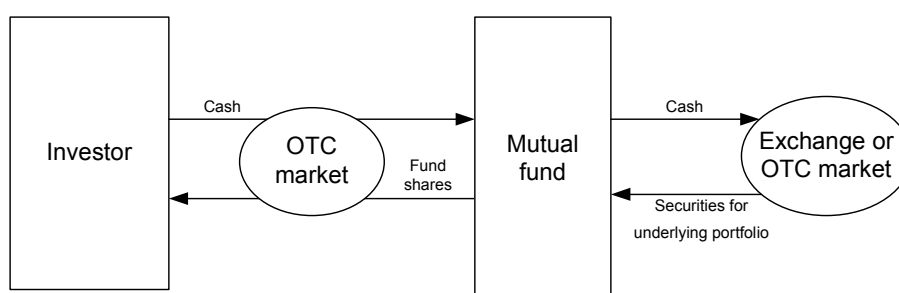


Figure 2.5: Operation of a Mutual Fund

<sup>64</sup> There are other types of fund: closed-end funds and unit trusts. For details, see Jones/Fabozzi (2002, pp. 503-530).

<sup>65</sup> See Kosev/Williams (2011, p. 52).

<sup>66</sup> See, for example, Jones/Fabozzi (2002, pp. 503-530) and Gastineau (2002, p. 531-554).

When purchasing, the investor pays the fund and receives shares. The mutual fund reinvests this cash in securities. If the investor wishes to sell his shares, the mutual fund sells a part of the portfolio and pays the investor back.

An ETF has a complex investment structure (see Figure 2.6).

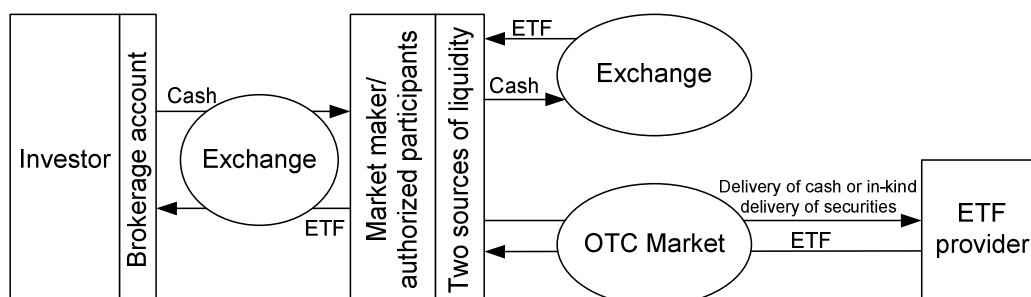


Figure 2.6: The ETF Share Creation Process<sup>67</sup>

The investor buys and sells ETF shares directly at the exchange, in the same way as he might buy and sell stocks. The market maker controls the liquidity of the ETF shares at the exchange. If there is a difference between the NAV and ETF prices, then the market maker will sell or buy ETF shares to remove this difference. The market maker has two sources of liquidity available to him so that he can accomplish this task; the first is to buy ETF shares at the exchange, and the second is the OTC market. At the OTC market, the market maker requests large blocks of ETF shares from the ETF provider in exchange for cash or securities (the latter is called in-kind delivery). This is a process that creates ETF shares. In reverse, the redemption process occurs when the market maker requests cash or securities from the ETF provider in exchange for ETF shares.

The difference in the way ETFs and mutual funds are constructed produces a difference in costs.<sup>68</sup> The ETF provider always operates with large blocks of fund shares, because it trades with market makers. In contrast, the mutual fund operates with small blocks of fund shares, because it trades directly with investors. Hence, the rebalancing needs of the mutual fund are higher than those of the ETF. This produces a larger number of transactions for the mutual fund and, as a consequence, leads to higher transaction costs and taxes on capital gains. Moreover, the accounting costs are higher for mutual funds than for ETFs because, in the first case the accounting is done at the fund level, while in the second case it is done at the shareholder level.

The ETF has the option of achieving tax optimization internally. This can be illustrated by an example. Suppose we have two units, created at two different points in time. The first unit

<sup>67</sup> Based on Tucker/Laipply (2010, p. 43).

<sup>68</sup> See Kostovetsky (2003, p. 84).

contains asset A, with price 100, and the second unit contains the same asset A, with price 150. Assume that the price of asset A at the current time is 160. Which unit is it better to redeem? For the first unit, the capital gain would be  $160-100=60$ , while for the second unit it would be  $160-150=10$ . It is thus clear that, for tax reduction, it is better to redeem the second unit. However, it is important to note that the unit in which asset A has price 100 still exists, and will be redeemed later. Hence, it is not possible to avoid the payment of tax, but it is possible to delay payment, keeping it as part of the portfolio and thus possibly producing additional profit.<sup>69</sup> The same method of tax optimization cannot be done for mutual funds, since the mutual fund has no way of controlling the inflows/outflows of investments that compel it to buy and sell shares without delay. In contrast, the ETF controls its own creation/redemption process.

However, with mutual funds, the purchasing costs are lower, because ETFs contain the bid/ask spread (deviations between the purchase price and the sell price around the NAV) and the brokerage commissions. Hence, for investors who regularly invest a small amount of capital, mutual funds can be a better choice.

The legal regulations for ETFs and mutual funds are different in Europe and the USA. In Europe, there is no difference between ETFs and mutual funds; both must conform to the UCITS<sup>70</sup> Directive. However, ETFs are not seen as mutual funds in the USA.<sup>71</sup>

ETFs have become a very popular investment vehicle. The total value of assets managed by ETFs rose from \$74.3 billion in 2000 to \$1,366.1 billion in 2011.<sup>72</sup> There is concern among regulatory authorities that the creation/redemption process of ETFs is an OTC transaction. Under stressed conditions, delays in the liquidation of ETFs can lead to significant losses for investors.<sup>73</sup>

ETFs and mutual funds use the same approaches to replicate the index. We describe these approaches in the next section.

---

<sup>69</sup> See Kostovetsky (2003, pp. 83-84).

<sup>70</sup> Undertakings for Collective Investments in Transferable Securities.

<sup>71</sup> See Ramaswamy (2011, p. 3).

<sup>72</sup> See Blackrock (2011).

<sup>73</sup> See IMF (2011, p. 71).

## 2.4 Tracking approaches

There are two groups of methods that can be used for index replication: physical replication and synthetic replication. The first group, which includes full replication and sampling, is based on direct investment in the assets. The second uses derivative contracts.

### 2.4.1 Physical replication

#### 2.4.1.1 Full replication

The full replication approach can be used to replicate indices that have a transparent, investable structure. This means that the index is computed as a sum of the products of asset prices and their quantities. Hence, to achieve the same performance as the index, the tracking portfolio can be built by purchasing the same fixed quantities of assets as are in the index.<sup>74</sup> On paper, full replication is the simplest method of constructing a tracking portfolio. However, this does not always hold in practice.

Assume that the investor's goal is to set up a new tracking portfolio using full replication. The investor faces two problems in using this method.

The first problem is related to the purchase price of the assets. There are various costs associated with the acquisition of an asset, which increase its purchase price:<sup>75</sup>

- *Opportunity costs.* These costs occur if the price of the asset changes over the time required to execute the market order. Another cause of such costs is the failure of the execution of the market order.
- *Market impact costs.* The bid/ask spread, which depends on the liquidity of the asset, means that the purchase price of the asset may differ from its price in the index. This is especially relevant for illiquid assets.
- *Transaction costs.* These include commission and custodial fees.

Hence, the purchase price of the asset is not equal to the price of the asset, which is used to calculate the index.

The second problem relates to the quantity of the asset purchased. There is a problem with rounding.<sup>76</sup> If the amount of the asset that must be purchased is determined in the index to

---

<sup>74</sup> See Rudd (1980, p. 58).

<sup>75</sup> See, for example, Hafner/Puetz/Werner (2003, p. 20).

three decimal places then, to avoid problems with rounding, the initial value of the portfolio will need to be increased. As a result, the full replication method can be very capital intensive.<sup>77</sup> Depending on the number of underlying assets in the index, a huge amount of capital may be needed to completely rebuild the index and avoid problems with rounding. Moreover, if the quantity of a certain asset in the index is very small, then it is sometimes better to exclude it from the portfolio altogether, to reduce the administrative work. The alternative solution is to increase the value of the portfolio, thereby increasing the amount of this asset that one needs to purchase. As a result, the full replication method is most suitable for large tracking portfolios in which it is possible to avoid problems with rounding and small positions.

### 2.4.1.2 Sampling

The second method of physical replication is sampling.<sup>78</sup> Here, the index is reproduced using a tracking portfolio containing a smaller number of assets, mostly components of the index. Hence, the tracking portfolio is different from the index and the reproduction is imperfect. In this method, three steps have to be performed. In the first step, the size of the tracking portfolio is chosen. In the second step, the assets included in the tracking portfolio are selected. In the third step, weights are assigned to each asset. The assets are selected and weighted using either meta-heuristic or classical optimization methods, with the aim of approximating the risk structure of the index.<sup>79</sup>

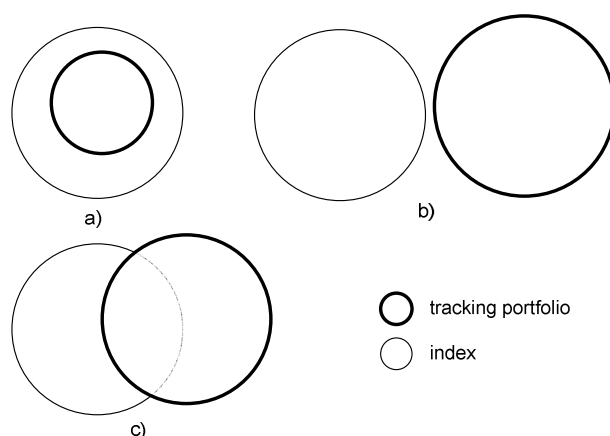


Figure 2.7: Overlapping of Assets from the Tracking Portfolio and the Index<sup>80</sup>

<sup>76</sup> See Dorfleitner (1999, pp. 493-502).

<sup>77</sup> See Wagner (2002, p. 821).

<sup>78</sup> See Montfort/Visser/van Draat (2008, p. 144).

<sup>79</sup> See Smeets/Wiesner (2009, p. 24).

<sup>80</sup> See Kusiak (2010).



Under the sampling approach, the tracking portfolio can contain various assets, which could be (see Figure 2.7)

- a) a part of the benchmark
- b) not a part of the benchmark or
- c) a combination of a) and b).

When applying sampling, it is necessary to choose the structure of the sampling universe, that is, to decide whether to use a), b), or c). This choice depends on the structure of the index. For indices with transparent, investable structures, the typical choice is a). However, variant c) can extend the flexibility of the replication at the expense of transparency. For some exotic indices, which do not have investable structures behind them, the best choice is b).

## **2.4.2 Synthetic replication**

Synthetic replication is the replication of the index using derivative products, such as futures, options, swaps or equity-linked notes.

### **2.4.2.1 Instruments**

#### **2.4.2.1.1 Futures**

The first method involves buying futures contracts on the index to be replicated.<sup>81</sup> The price of the futures depends on the spot level of the underlying index. To fulfil the goal of index tracking, the purchase price of the futures contracts should be only slightly different from the current index level. This difference can be interpreted as a cost of tracking.

The futures markets are very liquid and there is no counterparty risk. Moreover, futures contracts are leveraged, which means that no initial capital is required to open positions in futures contracts. The capital is only required later, to cover the margin between the purchase price and the current level of the index.

The spread between the futures price and the spot price is called the cost of carry. The cost of carry depends on the ability to replicate the return of the index using physical replication. For example, if the cost of physical replication is lower than the cost of carry on the market, then there is an opportunity for index arbitrage. One can replicate the index using physical

---

<sup>81</sup> For details and examples, see Bruce/Eisenberg (1992, pp. 45-47), Waring/Attwood (1998, pp. 17-30), Cano/Feldman/Smith (2009), Ronalds/Anderson (2006, pp. 30-35), and Hamid/Sandford (2007b).

replication and at the same time sell the index futures. The difference between the cost of carry and the cost of physical replication is an arbitrage profit.

The main drawback of using futures for index tracking is that they have expiration dates. On the expiration date, open positions must be closed and new positions opened. This operation is called rollover. A negative difference between the closing and new opening prices will increase the costs of tracking.

It is worth noting that futures on price indices deliver only a price return.<sup>82</sup> Hence, by using futures to track indices, the investor loses the dividends that would be obtained by physically replicating the index. Transaction costs must also be considered when selecting futures to track the index.

#### **2.4.2.1.2 Options**

The second instrument that can be used to track an index is options. In this case, one needs to buy call and sell put options on the index one wishes to track.<sup>83</sup> There are two strategies in use.

The first strategy is the synthetic long index (see Figure 2.8). It consists of buying a call and selling a put with the same strike price. The combined pay-offs from these two contracts gives the value of the index (bold black line).

The second strategy is the long combo (see Figure 2.9). This strategy involves buying a call and selling a put with different strike prices. The strike price of the call should be lower than that of the put. This strategy is very similar to the synthetic long index, with the exception that there is an insensitive area between the call strike and the put strike. If the index value is in the interval between the two strikes then the total pay-off from the two contracts is equal to zero. If the level of the index is either above the put strike or below the call strike, then the value of the tracking portfolio containing the short put and the long call is equal to the value of the index.

---

<sup>82</sup> See Cano/Feldman/Smith (2009, p. 2).

<sup>83</sup> See Hill/Mueller (2004, p. 505).

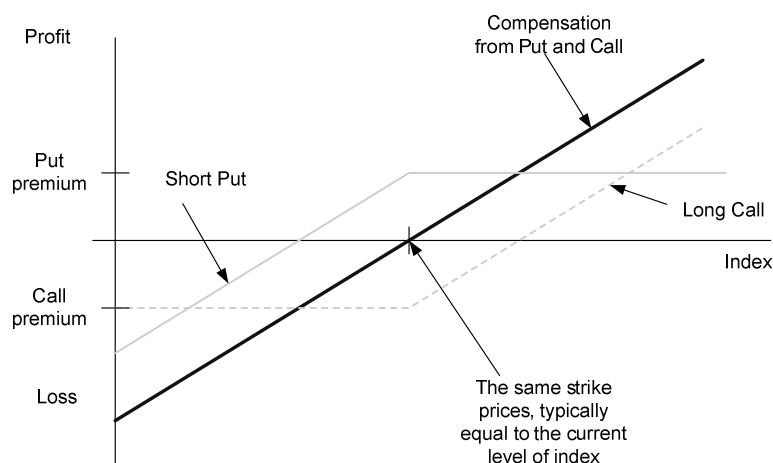


Figure 2.8: Options Strategy for Tracking an Index (Synthetic Long Index)

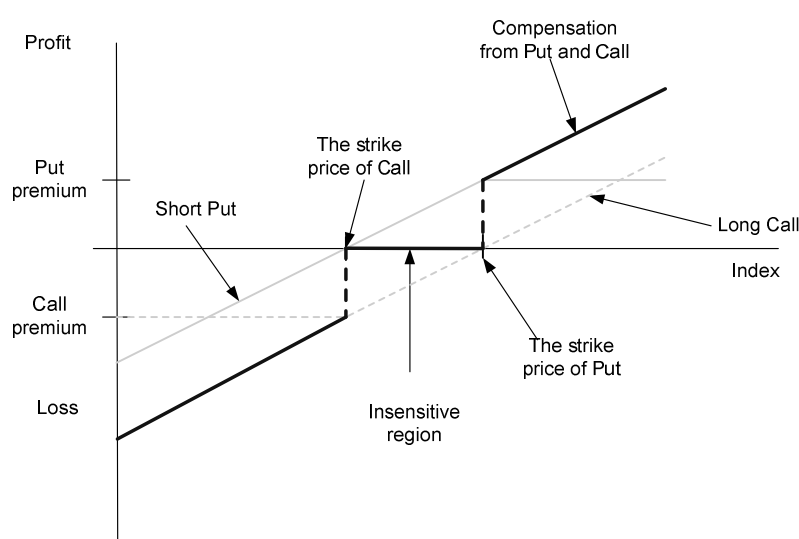


Figure 2.9: Options Strategy for Tracking an Index (Long Combo)

Options strategies are leveraged strategies. The costs of obtaining the index return include the call premium minus the put premium, rollover costs, and transaction costs. Moreover, there are no dividends when using options to track price indices, which must be taken into consideration.

### 2.4.2.1.3 Swaps

The third method is to use index swaps. A swap is an OTC instrument, requiring a counterparty. The index swap contract entails one party agreeing to pay the other party the index return in exchange for a small fee and the return on the portfolio, which is held by the other party (see Figure 2.10).<sup>84</sup>

<sup>84</sup> See Hill/Mueller (2004, pp. 522-523).

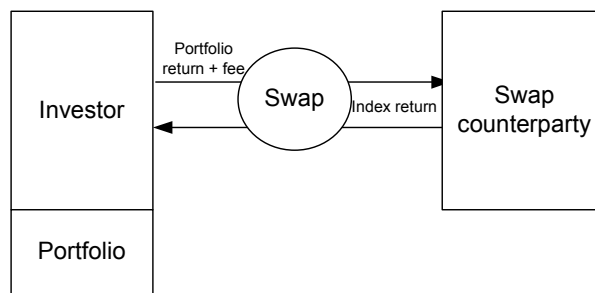


Figure 2.10: Index Swap

The swap contract is a simple way of obtaining the index return, but it is necessary to find a counterparty.

#### 2.4.2.1.4 *Equity-linked notes*

The fourth method is to invest in an equity-linked note.<sup>85</sup> This is another OTC instrument, which is very similar to a swap (see Figure 2.11).

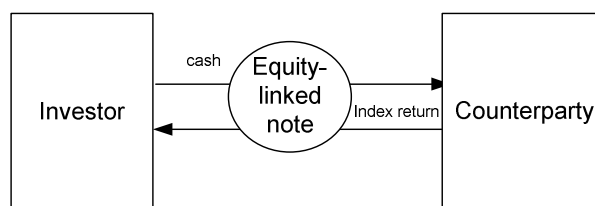


Figure 2.11: Equity-linked note

The investor invests in the equity-linked note and receives the index return. Typically, the principal is protected. The only counterparty risk is related to the index return.

### 2.4.2.2 Structuring the index fund

Index funds are legally limited in their use of synthetic replication. This has led to the creation of special structures, designed for index funds. The most commonly used structures are the unfunded swap and the funded swap.<sup>86</sup>

<sup>85</sup> See Ramaswamy (2011, p. 6).

<sup>86</sup> See Johnson/Bioy/Rose (2011, p. 7).

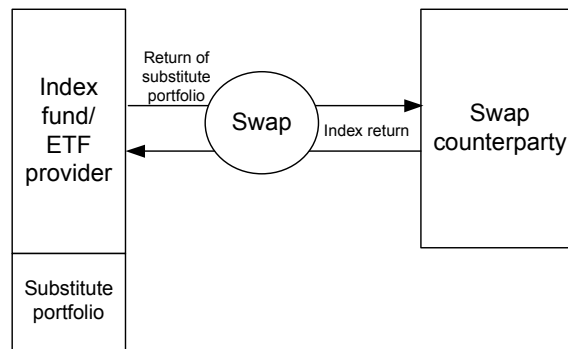


Figure 2.12: Unfunded Swap<sup>87</sup>

In the unfunded swap (see Figure 2.12), the issuer of the index fund or ETF constructs an in principle arbitrary portfolio and the counterparty agrees to pay the index return in exchange for a small fee and the return on the issuer's portfolio. The substitute portfolio is a pledge, made to the investor in case the swap counterparty goes bankrupt. Moreover, the value of the swap part is legally limited so as to reduce the counterparty risk.<sup>88</sup> For example, in the EU, this risk is limited to a maximum of 10% exposure to the swap counterparty.<sup>89</sup> The swap part in this case is not collateralized.

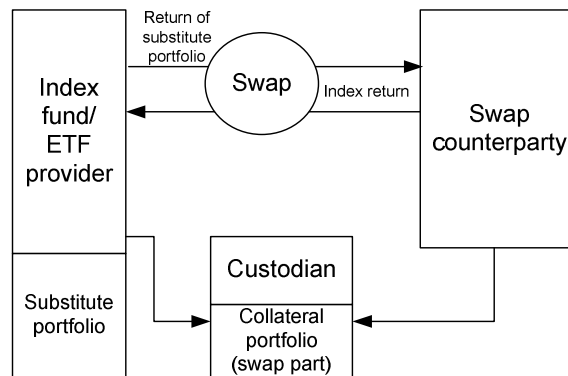


Figure 2.13: Funded Swap I

The funded swap structure is an extension of the unfunded one, in which the swap part is collateralized or even overcollateralized (see Figure 2.13). The collateral is held by a custodian. This variant is used by the ComStage ETF.<sup>90</sup>

<sup>87</sup> Based on Michalik (2008, p. 63).

<sup>88</sup> See Ramaswamy (2011, p. 7) and Kosev/Williams (2011, p. 59).

<sup>89</sup> See Article 52 of the UCITS Directive.

<sup>90</sup> See ComStage (2011, p. 14).

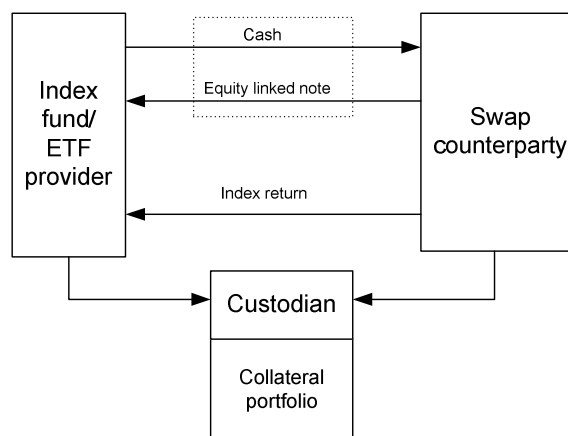


Figure 2.14: Funded Swap II

There is another version of the funded swap (see Figure 2.14), in which the substitute portfolio and the swap collateral are combined into one portfolio that is then placed with the custodian. Among practitioners, this transaction is called a swap but, as noted by *Ramaswamy*<sup>91</sup>, this is misleading because there is no actual swap transaction as there is in the first type of funded swap. This transaction type is very similar to an equity-linked note that is fully collateralized. It is used by iShares and db X-trackers.<sup>92</sup>

Group	Fund	Type	Description
I	iShares	all	Correlation between the collateral basket and the underlying index is not taken into consideration.
	Credit Suisse	all	Correlation between the substitute baskets and the underlying indices is not taken into consideration.
	Amundi	all	Correlation between the substitute baskets and the underlying indices is not taken into consideration.
	ComStage	all	Correlation between the fund holdings and the underlying indices is not taken into consideration
II	RBS	all	The funds don't target a specific degree of correlation between their holdings and reference indices
	ETF Securities	all	ETF Securities does not intend to maintain a specific degree of correlation between the substitute basket and the fund's benchmark.
III	EasyETF	all	When constituting the substitute baskets, EasyETF prioritizes the liquidity of the securities over their correlation with the reference index.
	Source	all	Source strives to achieve the highest correlation possible between the assets held by the fund and the fund's benchmark index.
	db X-trackers	fixed income and money market	The aim is to have a high correlation between the substitute basket and the relevant underlying index.
IV	db X-trackers	equity	not specified
	Lyxor	all	not specified
	UBS	all	not specified
	XACT	all	not specified

Table 2.3: Correlation of Substitute Basket / Collateral Basket with the Index<sup>93</sup>

An enhanced study was carried out by Morningstar to shed light on the substitute/collateral portfolios that are used in practice.<sup>94</sup> Morningstar interviewed twelve ETF providers. We have

<sup>91</sup> See Ramaswamy (2011, p. 6).

<sup>92</sup> See Deutsche Bank (2011, p. 7) and iShares (2010, p. 4).

<sup>93</sup> Based on research by Johnson/Bioy/Rose (2011).

<sup>94</sup> See Johnson/Bioy/Rose (2011).

organized the answers into four groups, according to the correlation of the substitute/collateral portfolio with the index (see Table 2.3).

The common tendency is for there to be a little dependence between the substitute/collateral basket and the index (groups I and II). However, some of the providers do not mention the structuring of the portfolio (group IV). Only three providers (group III) say that they consider the dependence between the portfolio and the index. Hence, in most cases, the substitute portfolio is not intended to replicate the index and has no influence on the quality of tracking. The desired tracking quality is produced by the swap contract (see Figure 2.15).

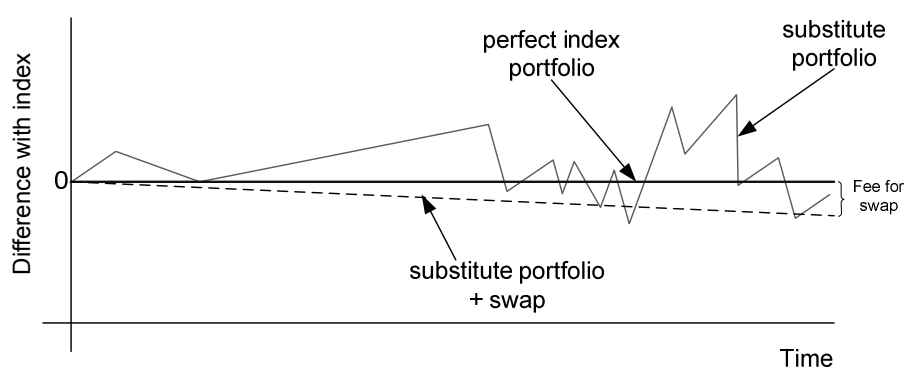


Figure 2.15: Combination of Substitute Portfolio and Swap<sup>95</sup>

The swap counterparty is typically an investment bank. There is little information about how banks hedge the risk in the swap contracts that produce the return on these indices.<sup>96</sup> Most likely, they will match the open positions. For example, ETF providers offer two types of ETFs for the same index: a simple ETF and an inversed one (see Figure 2.16).

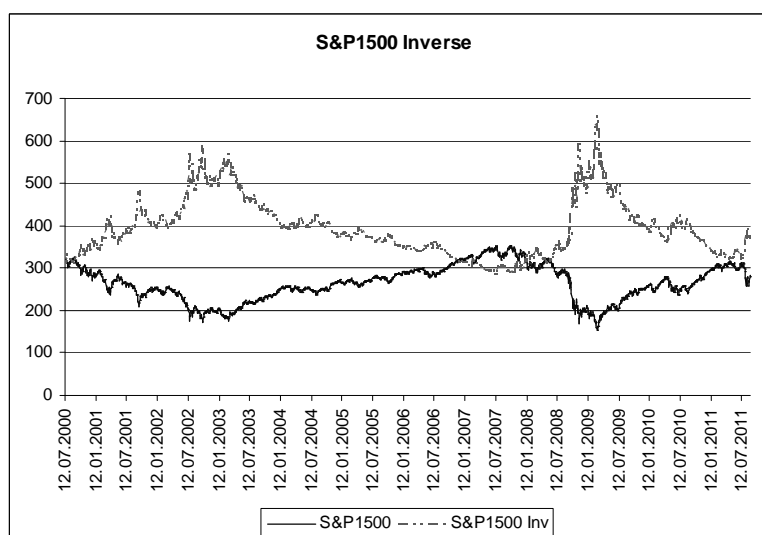


Figure 2.16: Index and Index Inverse

<sup>95</sup> Based on Michalik (2008).

<sup>96</sup> The problem of pricing index swaps and their non-transparency is discussed in Berke (2009, pp. 19-20).

As a consequence, there are two swap contracts with the swap counterparty. Matching the two types of contracts creates a market-neutral position, i.e. the position at which the market risk is neutralized, for the swap counterparty but it receives fees from both contracts.

Swap contracts can be also hedged using other derivative instruments, for example options or futures, or by using a hedging portfolio, which can be constructed using the sampling method.<sup>97</sup> An additional benefit from the maintenance of swap contracts is the optimization of the substitute/collateral portfolio. Typically, this portfolio is constructed with the following aims:<sup>98</sup>

- cost optimization;
- liquidity;
- diversification;
- tax efficiency.

The costs of a swap contract are visible to the investor through the swap fees, which depend on the ability of the counterparty to reduce the maintenance costs.

### **2.4.3 Evaluation and discussion**

In this section, we evaluate the abovementioned methods of index tracking according to the following criteria:

- applicability of the tracking method;
- set up costs;
- structural tracking error;
- maintenance costs;
- tax efficiency;
- flexibility of tracking portfolio;
- transparency of tracking method;
- counterparty risk;
- risks to stability of the financial system.

This assessment will allow us to develop a decision flow which we can use to select the most appropriate index-tracking approaches in specific cases.

---

<sup>97</sup> See Kusiak (2010, p. 32) and Ramaswamy (2011, p. 8).

<sup>98</sup> See Bootz (2009).



## **Applicability of tracking method**

The applicability of the tracking method for the selected index depends on the index structure. The full replication method can be applied to indices whose assets and their quantities are known. Additionally, it has to be possible to invest in these assets. If these conditions are not met, then the full replication method cannot be applied. Moreover, in some cases, legal regulations may limit the application of the full replication method. For example, for funds satisfying the UCITS Directive, no asset can make up more than 10% of the fund portfolio. Therefore, full replication will be restricted if the weight of any asset in the index exceeds 10%.<sup>99</sup>

In comparison to full replication, sampling is a very flexible method that can be applied to any kind of index.

The replication of indices using futures or exchange-traded options can be achieved if there are liquid markets for these types of contracts. Alternatively, it is possible to use synthetic replication if there is a counterparty for derivatives contracts (options, swaps, or equity-linked notes) on the OTC market.<sup>100</sup> However, if OTC contracts are too expensive, then the sampling approach can be preferable.

## **Set up costs**

Set up costs for the tracking portfolio are very high if the full replication method is used. Additionally, if the index contains a group of illiquid assets, then transaction costs can increase significantly. Moreover, as was discussed above, the full replication method requires a large amount of initial capital.

To overcome these problems, one can use sampling. In this way, a tracking portfolio can be built with a small amount of capital, without investing in all of the illiquid assets that are in the index.

In comparison to physical replication methods, synthetic replication methods have very low initial costs. Moreover, the instruments are leveraged and so the full amount of capital is not required to build the tracking portfolio. However, if the index return is obtained using options, the set up costs do include the difference between the premiums of the put and call options. If futures are used, the set up cost consists of the difference between the futures price and the

---

<sup>99</sup> See Günther (2002, p. 229).

<sup>100</sup> See Hamid/Sandford (2007a, p. 5).

spot price of the index. In the case of swaps, the cost of swap replication must be paid to the counterparty.

### **Structural tracking error**

The tracking error measures the difference between the returns from the tracking portfolio and those from the index. There is a difference between the structural tracking error, which characterizes the asynchronicity in the movements of the index and the tracking portfolio, and the tracking error that comes from the maintenance costs of the tracking portfolio.<sup>101</sup> The structural tracking error reflects the ability of the tracking portfolio to track the index. The reason for this type of error is a mismatch between the two.

The full replication method is not affected by variations between the price of the holdings and the index if it is applied to indices with constant quantities of holdings. However, there will be a structural tracking error if the index provider changes the structure of the index and the index fund does not apply these changes to the tracking portfolio. To follow the changes exactly, the index fund needs to know about them as soon as they occur. Typically, the index provider will use a set of rules to regulate index readjustments.<sup>102</sup> If these rules are not defined, then the risk of incurring a structural tracking error when applying full replication increases.

There are other sources of structural tracking error in the full replication method. There is a high risk of such error if the method is applied to indices with variable amounts of holdings, for example, equal-weighted indices. Dividend reinvestment is another source, when physical replication approaches are used for performance indices, because there is a delay between the reinvestment of dividends in the index, and the receipt of the dividends and their reinvestment in the tracking portfolio.

Replication using sampling bears the largest risk of structural tracking error among the tracking methods. The tracking portfolio in this case is very sensitive to variations between the prices of the holdings and the index, because it does not contain all of the assets from the index.

The cause of structural tracking error in synthetic replication using futures and options is the expiration of the contracts. The rollover costs can be seen as structural tracking errors in this case. However, structural tracking error is eliminated when using swap contracts, since then

---

<sup>101</sup> See Budinsky (2002, p. 156).

<sup>102</sup> See Wagner (2002, p. 820).

the risk of tracking error is fully absorbed by the counterparty. Therefore, synthetic replication using a swap provides much more accurate index tracking than other methods.<sup>103</sup>

The final cause of structural tracking error is cash drag, which can appear when physical replication is used. The inflows and outflows of investments and dividends contribute to the cash position of the index fund.<sup>104</sup> The cash position in the tracking portfolio between receiving and reinvesting cash is not synchronized with the movements of the replicated index, and this has a negative impact on the tracking quality. As a consequence, the tracking portfolio outperforms the index in a falling market and underperforms it in a rising market. The recommended way of managing the cash position is to rebalance after achieving some given, critical, amount of cash. In the intermediate period it is possible to cover the cash position using futures contracts on the index.<sup>105</sup>

### **Maintenance costs**

The maintenance costs contribute to the mismatch between the performance of the index and the tracking portfolio. They include transaction costs, market impact costs, opportunity costs, and management costs.

The index providers make a periodic rebalancing of their indices. As was shown in Section 2.2.1, such changes are not rare. Hence, the tracking portfolio must also be revised in order to keep it synchronized with the underlying index, resulting in transaction costs.<sup>106</sup> The size of the necessary transaction costs depends on the frequency of revisions and can be relatively high. As a rule, the more constituents there are in the fund portfolio, the higher are the costs.

Illiquid assets are an additional source of maintenance costs in full replication. The large bid/ask spread and essential transaction costs, which are typical for illiquid assets, can affect the tracking ability, since the original index is constructed without paying attention to any costs.<sup>107</sup> Moreover, the inflows and outflows of investments and the inflows of dividends also make it necessary to rebalance fund portfolios, increasing the maintenance costs further.

Additionally, the weighting schema of the index may increase the maintenance costs. The full replication method is suitable if the index can be represented as the sum of the products of the asset prices and their quantities, where the quantities are constant. The types of indices with

---

<sup>103</sup> See Hamid/Sandford (2007a, p. 13).

<sup>104</sup> See Costandinides et al. (2010, p. 5).

<sup>105</sup> See Schoenfeld/Maeda (2004, p. 367).

<sup>106</sup> See Schioldager et al. (2004, pp. 391-393).

<sup>107</sup> See Wagner (2002, p. 818) and Günther (2002, p. 229).

this property are price-weighted, for example the DJIA, and capital-weighted, for example the DAX30 or S&P500. In this case, despite the changes in the asset prices, the weights of the assets in the index and the tracking portfolio change identically. Hence, the tracking portfolio mimics the index without any need for rebalancing. However, if the full replication method is applied to indices with non-constant quantities of assets, the maintenance costs can be large. For example, a tracking portfolio for a constant-weighted index needs to be rebalanced constantly, because the quantities of the assets will need to change every time there is an asynchronous change in their prices, so as to keep the weights constant. Hence, the full replication approach is very expensive for constant-weighted indices.

The problem of wealth erosion is inherent in full replication.<sup>108</sup> This is a hidden cost of tracking. It is caused by the fact that changes in indices are predictable and pre-announced. This leads to abnormally high prices for new assets and low prices for removed assets.<sup>109</sup> After inclusions and deletions, the prices of these assets move back towards their previous levels. As a consequence, the investor loses capital.

To reduce the maintenance costs of the tracking portfolio, especially for large indices, instead of full replication, the sampling method can be used.<sup>110</sup> The aim is to reduce costs by decreasing the number of assets in the tracking portfolio. However, additional costs are involved if there is a significant deviation from the index and the tracking portfolio has to be rebalanced.

The maintenance costs for both full replication and sampling can be reduced if the index fund participates in stock-lending activities. There are more opportunities to participate in stock lending when using full replication than when using sampling. The disadvantage of this practice is the counterparty risk.

The maintenance costs for synthetic replication result from the brokerage commissions and fees. For example, the fees for swap contracts depend on the ability of the counterparty to maintain the swap contract.

### **Tax efficiency**

There are two common taxes applied to investments: taxes on capital gains and taxes on dividends. The taxes on dividends are typically much higher than those on capital gains. The full replication method has a lower portfolio turnover in comparison to sampling, which leads

---

<sup>108</sup> See Schoenfeld/Maeda (2004, p. 377-380).

<sup>109</sup> See Bos (2000), Blume/Edelen (2004, p. 37), and Siegel (2008, p. 352).

<sup>110</sup> See Wagner (2002, p. 822).

to lower capital gains, and as a consequence, lower taxes. However, the taxation of dividends cannot be optimized under this method. In contrast, in sampling and synthetic replication it is possible to implement tax optimization strategies. For example, sampling can eliminate much of the dividend income without sacrificing total return by selecting the stocks with low dividend yield for tracking portfolio.<sup>111</sup>

### **Flexibility of tracking portfolio**

The full replication approach has an inflexible portfolio structure in comparison to other tracking approaches. When using full replication, the tracking portfolio should follow the index exactly. Hence, there are only limited possibilities for optimizing maintenance costs and taxes. Moreover, there is a problem with wealth erosion.

Sampling provides flexibility. It is possible to minimize maintenance costs and optimize taxes using this method. Moreover, the flexibility of sampling can be used to gain excess returns over the index when there are active management possibilities.<sup>112</sup>

Synthetic replication is also a very flexible approach because the return on the index is achieved using derivative products and so the structure of the substitute/collateral portfolio can be changed easily.

### **Transparency of tracking method**

The main difference between physical replication and synthetic replication is their transparency to the investor (see Figure 2.17). Full replication can be seen as a white box approach, where the investor knows exactly which assets are in the fund portfolio, because they are the same as those in the index. However, full replication loses some transparency if stock lending is used.

Sampling is less transparent than full replication, because the tracking portfolio is not the same as the index. To increase the transparency of sampling, US legislation specifies rules for the naming of funds; if a fund contains less than 80% of the assets in the index, this must be specified in its name.<sup>113</sup> Hence, the assets and their quantities cannot differ greatly from those in the index.

---

<sup>111</sup> See Booth (2001).

<sup>112</sup> See Wagner (2002, p. 821), Ezra/Warren (2010, p. 5) and Section 2.1.

<sup>113</sup> See Investment Company Act of 1940 and Ramaswamy (2011, p. 4).

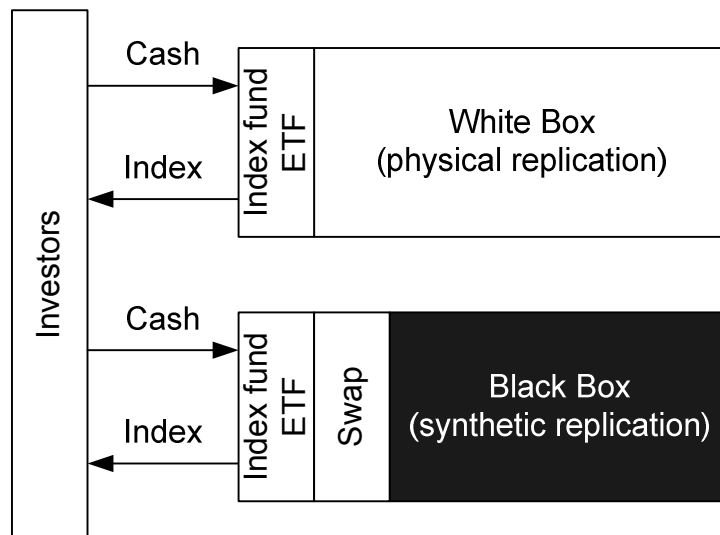


Figure 2.17: White Box and Black Box Approaches to Index Tracking

Synthetic replication is a black box approach that hides all details of the index tracking. Under this approach, the index fund decides which method is most suitable for the replication. As was shown above, the structure of the substitute portfolio may not reflect that of the index and, as a result, may not reflect the expectations of the investor either. However, this non-transparency helps to reduce the maintenance costs, by making use of synergies with the activities of the bank, which is typically the swap counterparty.

### Counterparty risk

A disadvantage of synthetic replication is the inherent counterparty risk, because the counterparty's creditworthiness is the only guarantee that the investor will receive the return on the index. In the EU, this risk is limited by regulations to a maximum of 10% exposure to the swap counterparty.<sup>114</sup> This means that the fund manager has to have at least 90% collateral to back the swap. Otherwise, the collateral has to be revised, resulting in transaction costs. In the event of the fund issuer's insolvency, only the collateral acts as a liability. All risks, including tracking error risk, and the costs of synthetic replication, are assumed by the swap provider.<sup>115</sup>

The physical replication methods can also incur counterparty risk if stock lending is used to enhance the returns of the index and reduce maintenance costs.

<sup>114</sup> See Article 52 of the UCITS Directive.

<sup>115</sup> See Hamid/Sandford (2007a, p. 13).

## **Risks to stability of the financial system**

Full replication and sampling are criticized for their non-transparency as a result of stock lending.<sup>116</sup> This activity can affect the stability of the financial system. However, there is more concern over the non-transparent structure of synthetic replication. The main criticism is the liquidity risk due to the mismatch between the assets in the fund portfolio and the index. This can lead to withdrawal from the index funds if the solvency of the fund providers is in question. This can affect the stability of the financial markets.<sup>117</sup> To mitigate this risk, the regulatory authorities require more transparency in the composition of collateral/substitute portfolios, and that the quality of the assets in fund portfolios must comply with guidelines developed by the Committee of European Securities Regulators (CESR).<sup>118</sup>

## **Which method to use**

There is a debate over whether it is better to use physical or synthetic replication. The two methods enjoy very similar levels of popularity in Europe. In 2010, 55% of ETF assets were physically replicated and 45% were synthetically replicated. In 2011, the proportions were 56% and 44%, respectively.<sup>119</sup>

As shown above, full replication, sampling and synthetic replication all have weaknesses and strengths. To select a suitable approach, we have developed the following decision flow (see Figure 2.18).

The applicability of the full replication method is limited; if the index does not have an investable structure,<sup>120</sup> then sampling or synthetic replication should be used. The maintenance costs are the central factor in method selection. If the assets in the index are not liquid, and/or the number of assets is very large, then sampling or synthetic replication should be used.

Two important properties of replication methods, transparency and flexibility, are included in the decision flow. The full replication approach is the most transparent, but it is not flexible. Synthetic replication is very flexible, but not transparent. The sampling approach is in the

---

<sup>116</sup> See FSB (2011, pp. 4-5).

<sup>117</sup> See Ramaswamy (2011), FSB (2011), and IMF (2011, p. 72).

<sup>118</sup> Guidelines on Risk Measurement and the Calculation of Global Exposure and Counterparty Risk for UCITS (2010).

<sup>119</sup> See Blackrock (2011, p.136).

<sup>120</sup> See Rudd (1986, p. 251).

middle. This approach is characterized by being both sufficiently flexible and sufficiently transparent. Therefore, hereinafter we investigate the application of the sampling approach.

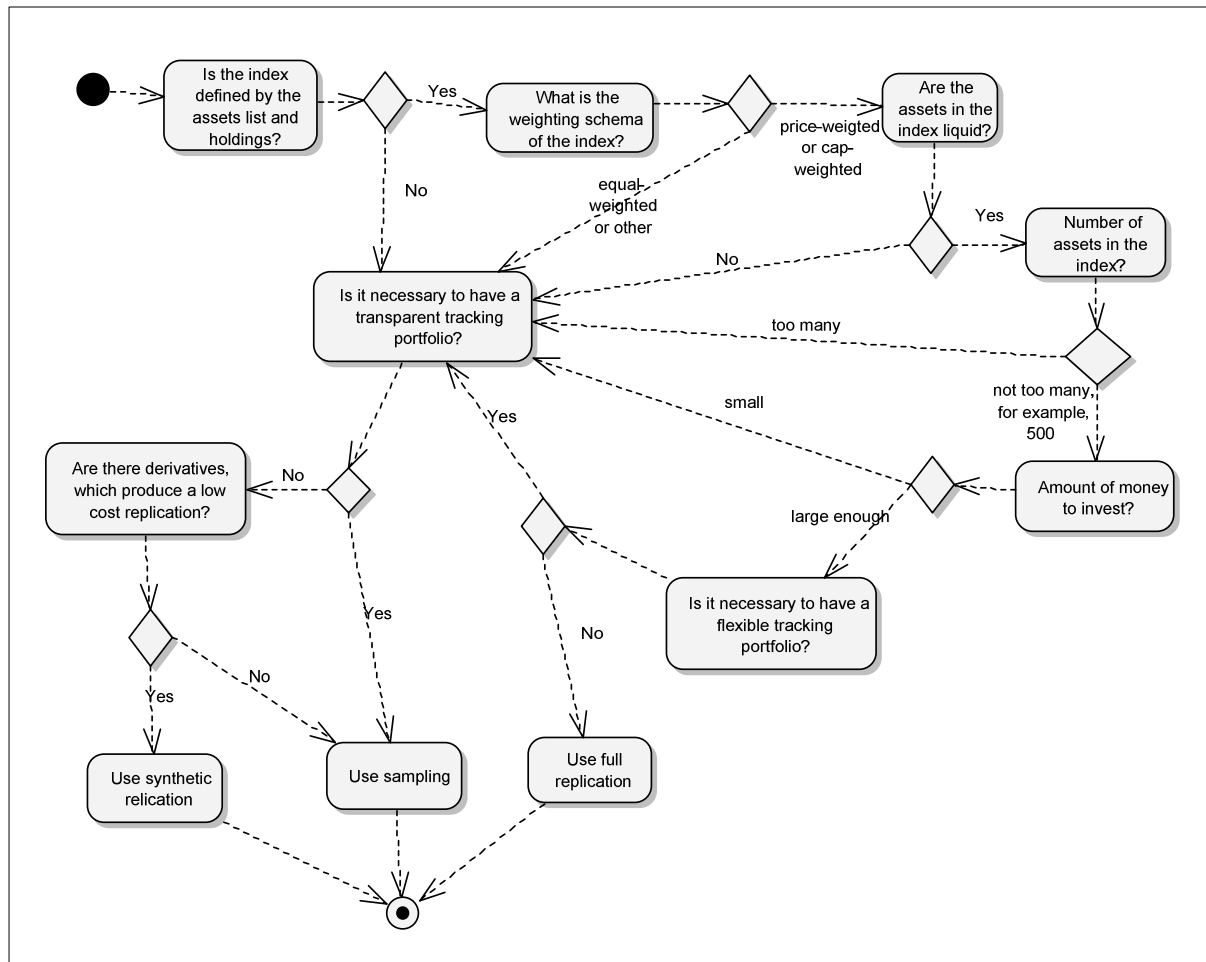


Figure 2.18: Which Method of Index Tracking should be used?



## 2.5 Summary

Index-based investments have both theoretical and empirical justifications. They simplify the stock investment process by focusing on the common factors that may affect companies, rather than the companies themselves. In addition, they allow the diversification of a portfolio, through the inclusion of the assets, for example, from different sectors, regions, and countries.

The methodology used to calculate an index is based on decisions made by the index provider, and this may have an impact on the ability to replicate index returns and on the returns of the investment. Therefore, it is necessary to know the structure of the index in order to achieve successful investment in it through replication.

Index funds make it easier to invest in an index. One can choose between ETFs and mutual funds. The main difference between them is that ETFs can be bought and sold at the exchanges and mutual funds cannot.

There are different methods of tracking indices, which can be classified into two groups: physical replication and synthetic replication. Physical replication involves investing directly in assets contained in the index. Synthetic replication uses derivative contracts, such as options, futures and swaps, to track the index.

Each of the methods of index tracking have strengths and weaknesses. However, sampling is a physical replication method that can be applied to any index. It is more flexible than full replication, and more transparent than synthetic. Moreover, it allows the index to be replicated by focusing on a trade-off between transaction costs and tracking error. These properties make it a good alternative.

### **3 Sampling**

This chapter starts with a review of the related work on sampling methods for index tracking. After that, the traditional measures of tracking quality are described, and an ex post tracking quality measure, which can be used to evaluate tracking quality during the investment period, is introduced. Then, selected methods of optimized sampling are presented, and the problem of modelling the portfolio structure is discussed. Finally, the requirements of sampling methods are formulated.

### 3.1 Related work

When applying sampling, the construction of a tracking portfolio consists of two tasks (see Figure 3.1): asset selection and asset weighting. Carrying out these two tasks provides a solution to the index tracking problem.

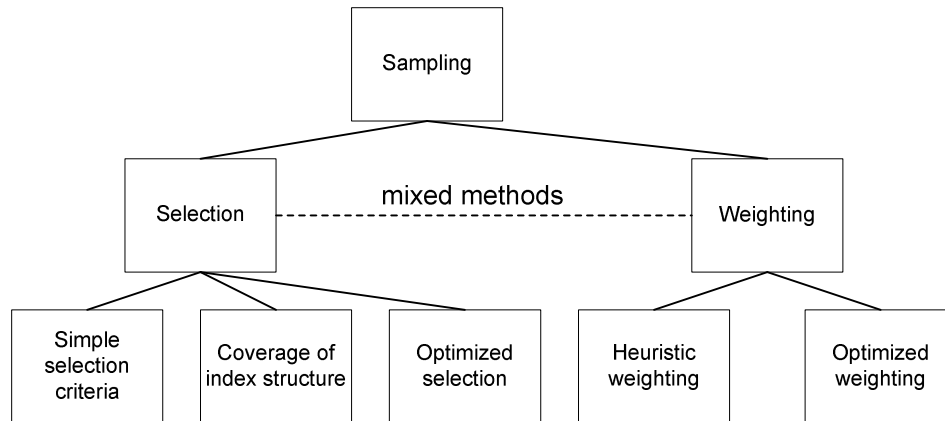


Figure 3.1: Overview of Sampling Methods

The asset selection for the tracking portfolio is based on two principles.<sup>121</sup> The first principle is that if two assets duplicate information about the index then one of them can be omitted. For example, if there are two strongly correlated assets, then one can be removed from the tracking portfolio. The second principle is that, if an asset has weak informational content, then it can be removed. For example, the variability of the price of an asset over time can be seen as information that is transmitted by that asset. As a consequence, an asset with a constant price, for example, does not contain any information about the index's variability and can thus be removed from the tracking portfolio.

There are three approaches to asset selection, based on these two principles. The first uses selection criteria. In this case, the initial asset universe is reduced by selecting only those assets that satisfy the predefined criteria. A typical criterion relates to the market capitalization of the assets. Assets with large market capitalization are selected since they are more informative than others, being heavily weighted in the market capitalization indices. The second approach aims to cover the structure of the index, by removing assets that contain duplicated information. For example, assets are selected from each industrial sector in the index, with the aim of representing its sectorial structure. The third approach is optimization based. Here, the selected assets are the solution to an optimization problem, formulated using the two principles of asset selection.

---

<sup>121</sup> These principles are based on Aivazian/Mhitarian (1998, p. 516).

The weighting approaches, meanwhile, can be divided into heuristic weighting and optimized weighting. For example, if the selected stocks are weighted according to their market capitalization, this would be a heuristic weighting. In optimized weighting, the weights are determined by solving an optimization problem.

### 3.1.1 Selection

#### 3.1.1.1 Simple selection criteria

The simplest criterion for the selection of assets is to select those assets with the largest weights in the index. In the case of capitalization-weighted indices, then, we would select those assets in the index with the largest market capitalization.<sup>122</sup> In the case of price-weighted indices, it would be those with the highest prices.<sup>123</sup> Empirical studies support this criterion.<sup>124</sup> Obviously, the assets in the index that have larger weights influence the index more than those that have smaller weights.

In the case of performance indices, assets with high dividends may affect the index behaviour more than other assets.<sup>125</sup> As a consequence, additional information about stocks that pay dividends should also be taken into consideration when selecting assets for the tracking portfolio.

Another criterion is the traded volumes of the assets. However, *Rey and Seiler* find this approach to be inferior to using the capitalization-weighted criterion.<sup>126</sup> Additionally, they combine turnover and capitalization-weighted selection criteria, using the ratio of turnover to market capitalization. However, the results of using this ratio are worse than those obtained using the market capitalization criterion alone.

Every asset makes its own contribution to the index. *Rafaely and Bennell*<sup>127</sup> suggest using each asset's relative contribution to the index as a criterion for asset selection. This relative contribution can be calculated as the product of the average weight of the asset and its average

---

<sup>122</sup> See, for example, Meade/Salkin (1989), Bamberg/Wagner (2000), Rey/Seiler (2001), Bonafede (2003), Montfort/Visser/van Draat (2008, p. 146), and Alighanbari/Mougeot (2009).

<sup>123</sup> See Alexander/Dimitriu (2004a).

<sup>124</sup> See Larsen/Resnik (1998).

<sup>125</sup> See Howaldt (2011)

<sup>126</sup> See Rey/Seiler (2001).

<sup>127</sup> See Rafaely/Bennell (2006).

price over the estimation period. The assets making the greatest contributions are selected for the tracking portfolio.

The correlation between the returns of the index and the asset returns is one of the possible statistical selection criteria that can be used.<sup>128</sup> In this case, those assets that are highly correlated with the index are selected for the tracking portfolio. However, the correlation can be seen as a short-term dependency measure, which is not appropriate if there is long-term dependency. Long-term dependency is better measured by testing the cointegration<sup>129</sup> relationship between the prices of the index and the assets. The presence of cointegration between the stocks and the index is used by *Dunis and Ho*<sup>130</sup> to select stocks for a tracking portfolio.

Another criterion, which is very similar to correlation, is proposed by *Montfort et al.*<sup>131</sup> They suggest regressing the returns of every asset on the returns of the index and selecting those assets whose beta coefficients in the regression are close to one. A value close to one would indicate that the asset and the index had similar reactions to market information.

An alternative selection method is to select assets randomly.<sup>132</sup> Using this approach, it is possible to neutralize the effect of stock selection in order to test weighting methods.

It is also possible to construct complex selection criteria. For example, several different criteria can be combined into a decision tree, using the algorithms from machine learning (i.e. ID3, C4.5, and CHAID), in the same way as is done for portfolio optimization.<sup>133</sup>

### **3.1.1.2 Coverage of index structure**

The next group of methods is aimed at selecting assets for the tracking portfolio so as to mimic the index's structure.

---

<sup>128</sup> See Montfort/Visser/van Draat (2008, p. 147).

<sup>129</sup> Two time series are cointegrated if there is a linear combination of them, which is stationary, i.e. the joint probability distribution of this linear combination does not change when shifted in time.

<sup>130</sup> See Dunis/Ho (2005).

<sup>131</sup> See Montfort/Visser/van Draat (2008, p. 147).

<sup>132</sup> See Alexander/Dimitriu (2003).

<sup>133</sup> See Sorensen/Miller/Ooi (2000).

		Market Capitalization			
		Large	Medium		Small
		Q1	Q2	Q3	Q4
Industry	0001 Oil & Gas				
	1000 Basic Materials				
	2000 Industrials				
	3000 Consumer Goods				
	4000 Health Care				
	5000 Consumer Services				
	6000 Telecommunications				
	7000 Utilities				
	8000 Financials				
	9000 Technology				

Cell with assets with medium market capitalization from health care

Index Universe

Figure 3.2: Stratified Sampling

Stratified sampling is one such method.<sup>134</sup> The index universe is divided into segments using one or a combination of characteristics. For example, one could use the market capitalization of the assets, and the assets' affiliations to industrial sectors (see Figure 3.2). As a consequence, the index universe can be represented as a cell structure. Each cell contains assets that share common characteristics. Then, representative assets from each cell are selected for the tracking portfolio.<sup>135</sup>

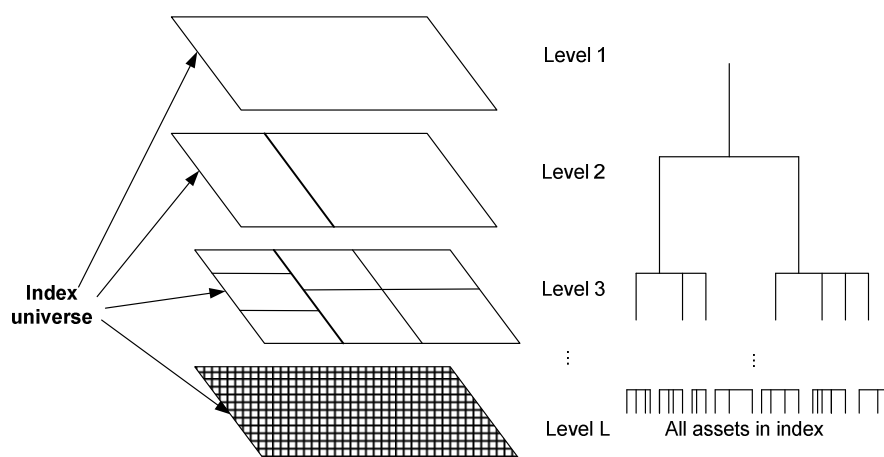


Figure 3.3: Hierarchical Clustering

The grouping of assets can also be done using quantitative methods. Some methods assign the assets into a number of groups, where the number is pre-specified. Examples include k-

<sup>134</sup> See Maginn et al. (2007, p. 425).

<sup>135</sup> An example of a stratification method is given in Section 4.3.

means, expectation-maximization (EM) clustering, and self-organizing maps.<sup>136</sup> There are also methods that build a hierarchical structure of groups (see Figure 3.3). There are two approaches that can be taken here. The first is a bottom-up approach: the assets are sequentially joined into groups using some measure of similarity. The second is a top-down approach: initially, all assets are placed in one group. Then, this group is divided into several groups, which in turn are divided further.

*Focardi and Fabozzi*<sup>137</sup> apply the methodology of hierarchical clustering of time series to the index tracking problem. They use the Euclidean distance between stock prices as a measure of similarity.<sup>138</sup> Other similarity measures could be used, for example, the correlation coefficient between stocks returns or percentage differences between the prices of stocks.<sup>139</sup> The hierarchical clustering allows one to group stocks using different levels of approximation. Then, one must select the level of approximation and a representative stock from each group, to form the tracking portfolio. For example, this could be the stock with the largest or median market capitalization within the group.

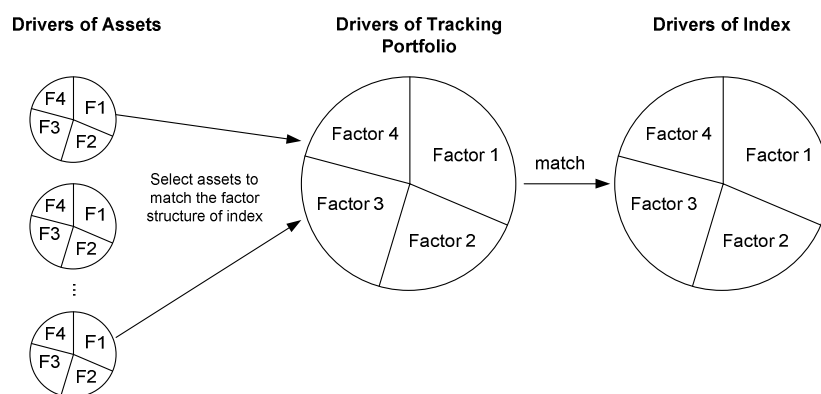


Figure 3.4: Matching the Factor Structure of the Index

Another way to select assets for a tracking portfolio is to use dimensionality reduction. One method that uses this approach is factor analysis. The idea behind it is to represent the returns on assets as a linear combination of several factors (common portion) plus an error term (unique portion).<sup>140</sup> The number of factors must be lower than the initial dimensionality of the data set. By applying factor analysis to the index tracking problem, one constructs a tracking portfolio that matches the common factors that drive the index performance (see Figure 3.4). The factors can be estimated using principal component analysis (PCA).

<sup>136</sup> Information about these methods can be found, for example, in Hastie/Tibshirani/Friedman (2009).

<sup>137</sup> See Focardi/Fabozzi (2004, pp. 417-425) and Dose/Cincotti (2005, pp. 145-151).

<sup>138</sup> See Focardi/Fabozzi (2004, p. 420).

<sup>139</sup> See Dose/Cincotti (2005, p. 148).

<sup>140</sup> See Cherkassky/Mulier (2007, p. 235).

*Corielli and Marcellino*<sup>141</sup> apply PCA to the index tracking problem. They estimate the factor model based on a matrix of stock prices. After that, they determine the factor structure of the index, calculating the correlation between the factors and the index prices. Then, those assets that are most correlated with the index factors are selected.

*Alexander and Dimitriu*<sup>142</sup> use PCA to replicate the common trend in the stock market. They apply PCA to the matrix of stock returns and select only the first component for the factor model. They then construct a portfolio that replicates this component. Their results show that it is possible to replicate the common trend in the stock market.

### 3.1.1.3 Optimized selection

Another possible way of selecting assets for a tracking portfolio is by formulating and solving an optimization problem. It is possible to formulate an optimization problem that groups assets according to their similarity, using some similarity measure, and selects only a pre-specified number of representative assets.<sup>143</sup> Such an optimization problem will typically include a cardinality constraint, to restrict the number of assets in the tracking portfolio. Hence, the problem is an integer programming problem, which is difficult to solve. One possible solution method is to use a branch-and-bound algorithm.

Instead of a similarity-between-assets measure, one can use a measure of tracking quality to select assets for the tracking portfolio.<sup>144</sup> To do so, one must calculate the tracking error for portfolios made up of all possible combinations of the assets, for a prespecified portfolio size, and then select the asset combination with the lowest tracking error. However, depending on the size of the index, the number of possible combinations can be enormous. For example, there are 155,117,520 possible combinations of tracking portfolios if we want to track the DAX 30 index using 15 of its 30 stocks. One of these combinations will have the best tracking quality but to find it, we would have to check all possible combinations. This procedure would require a huge amount of computing power and a huge amount of time. As the indices get bigger, the number of possible combinations grows exponentially.

Searching different combinations of assets is equivalent to adding in the cardinality constraint. Hence, we have a mixed integer programming problem. It is possible to solve this

---

<sup>141</sup> See Corielli/Marcellino (2006).

<sup>142</sup> See Alexander/Dimitriu (2003, p.7).

<sup>143</sup> See Cornuejols/Tütüncü (2007, p. 212).

<sup>144</sup> See Section 3.2 for details.



optimization problem using algorithms for meta-heuristic optimization: genetic algorithms<sup>145</sup>, threshold accepting<sup>146</sup>, differential evolution<sup>147</sup>, hybrid algorithms<sup>148</sup>, and simulated annealing<sup>149</sup>. These search methods are very efficient even when dealing with large search spaces. They do not guarantee that the overall best combination will be found, but with a proper setup they are able to find solutions that are close to the overall optimum.

There is also an approach that avoids solving an integer programming problem. The simplest version is to solve the optimization problem to find the minimum tracking error, using all of the assets in the index universe. As a result, the weights are computed for every asset. Then, only the assets with the largest weights are included in the tracking portfolio. The number of assets selected is equal to the size of tracking portfolio desired.<sup>150</sup> This approach can lead to poor solutions if there is a sector or region with many stocks that are highly correlated. The optimization problem assigns similar, small weights to these highly-correlated stocks, which can lead to their exclusion from the tracking portfolio and in turn to the exclusion of a whole sector or region from the portfolio.<sup>151</sup>

To solve this problem, a sequential method can be applied.<sup>152</sup> After solving the optimization problem, a small subset of stocks with the smallest weights, is removed. Then, the procedure is repeated with the remaining stocks, until the desired number of shares is reached.

The main disadvantage of optimization methods that look for stocks that are very similar to the index is the low level of diversification in the resulting tracking portfolio. The optimal tracking portfolio may contain stocks that are very similar to the index and, as a consequence, these stocks will be similar to each other, which reduces the diversification of the portfolio. There are two ways of avoiding this problem. The first is to add additional constraints, which control the diversification of the portfolio. For example, it is possible to add sector constraints, which specify that the sectorial structure of the portfolio must mimic that of the index, or a portfolio concentration constraint, which limits the weights that can be assigned to individual stocks. The second possibility is to filter out similar stocks before carrying out the optimization.

---

<sup>145</sup> See Ruiz-Torrubiano/Suarez (2009, p. 58), Rafaelly/Bennell (2006), and Beasley/Meade/Chang (2003).

<sup>146</sup> See Gilli/Kellezi (2001, p. 9).

<sup>147</sup> See Maringer (2008, pp. 12-14) and Krink/Mittnik/Paterlini (2009).

<sup>148</sup> See Fastrich/Paterlini/Winker (2010).

<sup>149</sup> See Kwiatkowski (1992) and Derigs/Nickel (2004).

<sup>150</sup> See Gaivoronski/Krylov/van der Wijst (2005, p. 125).

<sup>151</sup> See Montfort/Visser/van Draat (2008, p. 147).

<sup>152</sup> See Jansen/van Dijk (2002) and Coleman/Li/Henniger (2006).

### **3.1.2 Weighting**

#### **3.1.2.1 Heuristic weighting**

A heuristic weighting can be applied if the structure of the index is known. For example, weighting the assets using their market capitalization produces good results for capitalization-weighted indices.<sup>153</sup> Moreover, selecting companies with the largest market capitalization and weighting them by their capitalization is recommended for active portfolio management, to reduce significant deviations of the investment portfolio from the performance benchmark.<sup>154</sup> The other heuristic weighting schemas are equal weighting, fundamental weighting and price weighting.

The advantage of the heuristic weighting schemas is their simplicity. The disadvantage of applying heuristic weighting in sampling is that it does not take into account the correlation structure of the assets in the index. As a result, the weights of the assets in the tracking portfolio increase proportionally, in comparison to the weights in the index, to compensate for the use of a smaller number of assets in the tracking portfolio.

#### **3.1.2.2 Optimized weighting**

The next group of weighting methods uses optimized weighting. These methods are based on optimization procedures, which can be characterized by the following aspects:

- the type of tracking quality measure used;
- the way the portfolio structure is modelled: weights or numbers;
- which data are used for the estimation period;
- the type of model used: one-period or multi-period;
- the treatment of transaction costs and the use of rebalancing strategies;
- the use of other constraints.

#### **Type of tracking quality measure used**

The optimization methods used to determine portfolio weights are based mainly on minimizing appropriate measures of tracking quality, which are based on the tracking error. The tracking error is typically calculated as the difference (in terms of simple returns or log returns) between the tracking portfolio and the index. However, the tracking error can also be

---

<sup>153</sup> See, for example, Alighanbari/Mougeot (2009) and Larsen/Resnik (1998).

<sup>154</sup> See Strongin/Petsch/Sharenow (2000, p. 18).

calculated based on value, that is using the difference between the value of the tracking portfolio and the value of the index.

The most popular measures of tracking quality are returns-based and quadratic; they are called tracking error variance (TEV) and mean squared error (MSE).<sup>155</sup> There is also a group of linear measures: mean absolute deviation (MAD), maximal absolute deviation (MAXD), mean absolute downside deviation (MADD), and maximal absolute downside deviation (MAXDD).<sup>156</sup> Measures based on the values of the index and the tracking portfolio are less popular. They have been used by, among others, *Gaivoronski et al.*<sup>157</sup> and *Rafaely and Bennell*.<sup>158</sup> A detailed review of the various measures of tracking quality is provided in Section 3.2.

There are also approaches to finding the optimal weights for the tracking portfolio that are not based on tracking quality measures. For example, *Bamberg and Wagner*<sup>159</sup> test different regression estimators to find the optimal weights. Some estimators for the regression problem, however, coincide with the tracking quality measures. The ordinary least squares estimator (OLS) is equivalent to MSE. The least absolute value (LAV) estimator is equivalent to MAD. Two other estimators, the Huber and Tukey M-estimators, which are also tested by *Bamberg and Wagner*, can be seen as special forms of tracking quality measures. The same can be said about the conditional value-at-risk measure (CVAR), which is applied to the index tracking problem by *Rockafellar and Uryasev*.<sup>160</sup>

### **The way the portfolio structure is modelled: weights or numbers**

The open question when formulating the index tracking optimization problem is how to model the portfolio structure, that is, whether to use weights or numbers of stocks. The usage of weights or numbers is also connected to the choice over the measure of tracking quality. If the measure is returns-based, then one must use weights, because the return of the tracking portfolio is calculated as the sum of the products of the asset weights and their returns. If the measure is value-based, however, one must use numbers as the value of the tracking portfolio is the sum of the products of the numbers of assets and their prices. However, it is possible to use numbers of stocks with returns-based measures by converting the numbers into weights in the optimization problem. There is no consensus over which approach is more appropriate

---

<sup>155</sup> See, for example, Roll (1992) and Beasley/Meade/Chang (2003).

<sup>156</sup> See Rudolf/Wolter/Zimmermann (1999).

<sup>157</sup> See Gaivoronski/Krylov/van der Wijst (2005).

<sup>158</sup> See Rafaely/Bennell (2006).

<sup>159</sup> See Bamberg/Wagner (2000).

<sup>160</sup> See Rockafellar/Uryasev (2002).

(see Table 3.1). Some studies adopt both concepts. We discuss the problem of “weights versus numbers” further in Section 3.3.4.

Weights	Numbers
Alexander/Dimitriu, (2004b)	Bamberg/Wagner, (2000)
Alighanbar/Mougeot, (2009)	Beasley/Meade/Chang, (2003)
Bamberg/Wagner, (2000)	Canakgoz/Beasley, (2008)
Bonafede, (2003)	Corielli/Marcellino, (2006)
Coleman/Li/Henniger, (2006)	Gaivoronski/Krylov/van der Wijst, (2005)
Derigs/Nickel, (2004)	Gilli/Kellezi, (2001)
Dose/Cincotti, (2005)	Konno/Wijayanayake, (2001)
Dunis/Ho, (2005)	Maringer, (2008)
Fastrich/Paterlini/Winker, (2010)	Meade/Salkin, (1989)
Gaivoronski/Krylov/van der Wijst, (2005)	Montfort/Visser/van Draat, (2008)
Jansen/van Dijk, (2002)	Pliska/Suzuki, (2004)
Krink/Mitnik/Paterlini, (2009)	Karlow/Roßbach, (2011)
Kwiatkowski, (1992)	Primbs/Sung, (2008)
Meade/Salkin, (1989)	Rafaely/Bennell, (2006)
Primbs/Sung, (2008)	Rockafellar/Uryasev, (2002)
Rudd, (1980)	Stoyan/Kwon, (2009)
Rudolf/Wolter/Zimmermann, (1999)	Zenious, (2005)
Ruiz-Torrubiano/Suarez, (2009)	
Wu/Chou/Yang/Ong, (2007)	

Table 3.1: Literature Review: Weights or Numbers

### Which data are used for the estimation period

The optimization procedures use past asset prices (or asset returns, which are calculated based on prices) and index levels to estimate the tracking portfolio. These data can be characterized by their frequency and the period length. A review of the literature shows that there is no consensus on data frequency or the length of the estimation period that should be used (see Table 3.2). Obviously, if the data are of lower frequency, then the length of the estimation period should be greater.

Most studies use daily data for the estimation of the tracking portfolio. However, there is some criticism of this. The first argument against it is that daily prices contain noise.<sup>161</sup> The second is that the daily returns are autocorrelated.<sup>162</sup> Monthly data contain less noise and are not autocorrelated. However, the problem with monthly data is that there are only twelve data points in a year. As a consequence, one needs a long estimation period in order to retrieve sufficient data to carry out the estimation but there is a danger of bias if we go too far back in the past. As a result, some authors suggest using weekly data,<sup>163</sup> which contain less noise and are less autocorrelated. However, weekly data still contain less information than daily data and the estimation period needs to be longer than one year, because 52 data points would not provide a stable estimation of the weights of the tracking portfolio.

<sup>161</sup> See Beasley/Meade/Chang (2003).

<sup>162</sup> See Pope/Yadav (1994).

<sup>163</sup> See Beasley/Meade/Chang (2003).

Study	Length of the estimation period	Frequency
Alexander/Dimitriu, (2004b)	minimum 3 years (approx. 720 data points) up to 5 years	daily
Alighanbar/Mougeot, (2009)	-	daily
Bamberg/Wagner, (2000)	1 year	daily
Beasley/Meade/Chang, (2003)	2.5 years (145 data points)	weekly
Bonafede, (2003)	5 years (60 data points)	monthly
Canakgoz/Beasley, (2008)	2.5 years (145 data points)	weekly
Coleman/Li/Henniger, (2006)	2.5 years (145 data points)	weekly
Consiglio/Zenios, (2001)	starting with 6 years (72 data points)	monthly
Corielli/Marcellino, (2006)	approx. 1.5 years (350 data points)	daily
Derigs/Nickel, (2004)	-	daily
Dose/Cincotti, (2005)	40 days	daily
Dunis/Ho, (2005)	minum 2 years (approx. 480 data points) up to 4.5 years	daily
Fastrich/Paterlini/Winker, (2010)	approx. 1 year (250 data points)	daily
Gaivoronski/Krylov/van der Wijst, (2005)	1) approx. 1 year (250 data points) 2) approx. 7 months (150 data points) 3) approx. 3.5 months (75 data points)	daily
Gilli/Kellezi, (2001)	2.5 years (145 data points)	weekly
Jansen/van Dijk, (2002)	-	-
Karlow/Roßbach, (2011)	12 months (approx. 240 data points)	daily
Konno/Wijayanayake, (2001)	3 years (36 data points)	monthly
Krink/Mittnik/Paterlini, (2009)	approx. 10 months (200 data points)	daily
Kwiatkowski, (1992)	1 year (approx. 240 data points)	daily
Larsen/Resnick, (1998)	5 years (60 data points)	monthly
Maringer, (2008)	approx. 2 years (500 data points)	daily
Meade/Salkin, (1989)	2 years (approx. 100 data points)	weekly
Meade/Salkin, (1990)	2 years (approx. 100 data points)	weekly
Montfort/Visser/van Draat, (2008)	1 year (approx. 240 data points)	daily
Pliska/Suzuki, (2004)	-	-
Primbs/Sung, (2008)	approx. 15 years (3750 data points)	daily
Rafaely/Bennell, (2006)	2, 5, 10, 30 data points	daily
Rey/Seiler, (2001)	1 year (12 data points)	monthly
Rockafellar/Uryasev, (2002)	approx. 2.5 years (600 data points)	daily
Rudd, (1980)	-	monthly
Rudolf/Wolter/Zimmermann,	approx. 5 years (62 data points)	monthly
Ruiz-Torrubiano/Suarez, (2009)	2.5 years (145 data points)	weekly
Stoyan/Kwon, (2009)	11 scenarios	monthly
Takeda/Gotoh/Sugiyama, (2010)	1) 10 years monthly (120 data points) 2) 3 years weekly (150 data points)	1) monthly 2) weekly
Wu/Chou/Yang/Ong, (2007)	2 years (approx. 480 data points)	daily
Zenios, (2005)	500 scenarios	monthly
Zhang/Maringer, (2010)	approx. 1 year (250 data points)	daily

Table 3.2: Literature Review: Length of Estimation Period and Data Frequency

### Type of model used: one-period or multi-period

The models used for index tracking are usually one-period. To use these models in a multi-period setting, the optimization problem is simply solved sequentially for each future time period. However, there are some approaches that are multi-period. One of these is stochastic programming, which is based on scenarios organized into event trees.<sup>164</sup> Another is dynamic programming, sometimes called stochastic control.<sup>165</sup>

<sup>164</sup> For details, see Stoyan/Kwon (2009).

<sup>165</sup> For details, see Primbs/Sung (2008) and Pliska/Suzuki (2004).

The advantage of the multi-period models is that they can include managing and rebalancing strategies for future time periods.<sup>166</sup> Their disadvantages are their computational complexity in comparison to the one-period models and difficulty in forecasting for future periods.

### **Treatment of transaction costs and the use of rebalancing strategies**

To limit transaction costs, information about current holdings is needed. Then, the difference between the weights or numbers of asset positions in the current portfolio and the optimal portfolio is computed. This difference is then used as an input into the transaction cost function. The output of this function is then limited in the optimization problem. Typically, the function is linear.<sup>167</sup> It is possible to use non-linear functions but they increase the complexity involved in solving the optimization problem.<sup>168</sup>

The tracking portfolio should be reviewed from time to time to reflect new market information and reduce the deviation from the index. Different rebalancing strategies can be implemented. They are external to the one-period optimization problems, but are encoded into the multi-period ones. The following types of rebalancing strategies exist:<sup>169</sup>

- triggered rebalancing, when the tracking error exceeds a predefined threshold;
- calendar rebalancing, which is performed at fixed time intervals.

### **The use of other constraints**

Other types of constraints can be used in the optimization problem. For example, short-selling constraints are used to allow only long positions in the tracking portfolio, which avoids unstable portfolios.<sup>170</sup> Concentration constraints limit the size of positions that can be taken in the assets in the tracking portfolio, to increase its diversification. Constraints on the sectors that assets can belong to ensure that the tracking portfolio is exposed to the same sectors as the index.<sup>171</sup>

---

<sup>166</sup> See Stoyan/Kwon (2009, p. 248).

<sup>167</sup> See, for example, Montfort/Visser/van Draat (2008, p. 147).

<sup>168</sup> See, for example, Beasley/Meade/Chang (2003).

<sup>169</sup> See Zhang/Maringer (2010).

<sup>170</sup> See Montfort/Visser/van Draat (2008, p. 146).

<sup>171</sup> See Stoyan/Kwon (2009).

## 3.2 Measures of tracking quality

The evaluation of how well the tracking portfolio replicates the index can be done using measures of tracking quality. These measures are based on the tracking error. In this section, we describe the traditional measures of tracking quality, and highlight their strengths and weaknesses. After that, we introduce the ex post tracking quality measure.<sup>172</sup>

### 3.2.1 Traditional measures of tracking quality

Measures of tracking quality can be defined based on simple returns (also called arithmetic or discrete returns) or log returns (also called geometric or continuous returns). Simple returns are calculated as follows:

$$R_{I,t} = \frac{V_{I,t} - V_{I,t-1}}{V_{I,t-1}}, \quad R_{P,t} = \frac{V_{P,t} - V_{P,t-1}}{V_{P,t-1}},$$
$$r_{i,t} = \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}},$$

where  $R_P$  is the return of the tracking portfolio,  $R_I$  is the return of the index,  $r_i$  is the return of asset  $i$ ,  $V_I$  is the value of the index,  $V_P$  is the value of the portfolio, and  $S_i$  is the price of asset  $i$ .

Log returns are computed as

$$\tilde{R}_{I,t} = \ln \frac{V_{I,t}}{V_{I,t-1}}, \quad \tilde{R}_{P,t} = \ln \frac{V_{P,t}}{V_{P,t-1}},$$
$$\tilde{r}_{i,t} = \ln \frac{S_{i,t}}{S_{i,t-1}}.$$

The connection between log returns and simple returns can be expressed as follows:

$$\tilde{R}_t = \ln(1 + R_t)$$
$$R_t = \exp(\tilde{R}_t) - 1.$$

---

<sup>172</sup> This section is adapted from the working paper by Roßbach/Karlow (2011).

Hence, the difference between log returns and simple returns is equal to<sup>173</sup>

$$\tilde{R}_t = \ln(1 + R_t) \approx R_t - \frac{R_t^2}{2}.$$

As a consequence, the lower is the value of  $R_t$ , the lower is the difference between the simple and log returns.

The log returns can be better approximated by a probability distribution, because they are in the interval  $\tilde{R}_t \in (-\infty, +\infty)$ . In contrast, simple returns are in the interval  $R_t \in (-1, +\infty)$ . Moreover, the sum of log returns in consecutive periods gives the total return for the whole period. The advantage of simple returns is that it is possible to aggregate them across assets. This means that the return of the portfolio can be calculated as a weighted sum of the simple returns of its constituents. This type of aggregation is important for index tracking. Therefore, we use simple returns in this study.

The majority of measures of tracking quality are based on the tracking error (TE). The TE measures the difference between the returns of the tracking portfolio and the returns of the index. It can be formulated as

$$TE_t = R_{I,t} - R_{P,t} = R_{I,t} - \sum_{i=1}^N w_i r_{i,t}, \quad (3.1)$$

where  $w_i$  is the weight of asset  $i$ , and  $N$  is the number of assets. The sum of the weights in the tracking portfolio is equal to one:

$$\sum_{i=1}^N w_i = 1.$$

Different measures of tracking quality use the TE in different ways. The MAD measures the mean absolute difference between the returns over a given time interval:

$$MAD = \frac{1}{T} \sum_{t=1}^T |R_{I,t} - R_{P,t}| = \frac{1}{T} \sum_{t=1}^T |TE_t|,$$

where there are  $t = 1, \dots, T$  time periods.

---

<sup>173</sup> See Gouriéroux/Jasiak (2001). Here the logarithm of returns are represented as two members of Maclaurin series with an assumption that  $-1 < R_t \leq 1$ .



The MSE measures the mean of the squared differences between the returns over a given time interval:

$$MSE = \frac{1}{T} \sum_{t=1}^T (R_{I,t} - R_{P,t})^2 = \frac{1}{T} \sum_{t=1}^T (TE_t)^2.$$

Sometimes, the equivalent root mean square error (RMSE) measure is used:

$$RMSE = \sqrt{MSE}. \quad (3.2)$$

The advantage of this measure is that it is of the same scale as the TE.

Figure 3.5 illustrates the difference between MAD and MSE. It can be seen that MSE assigns relatively higher measures to larger differences in returns compared to smaller ones than MAD does.

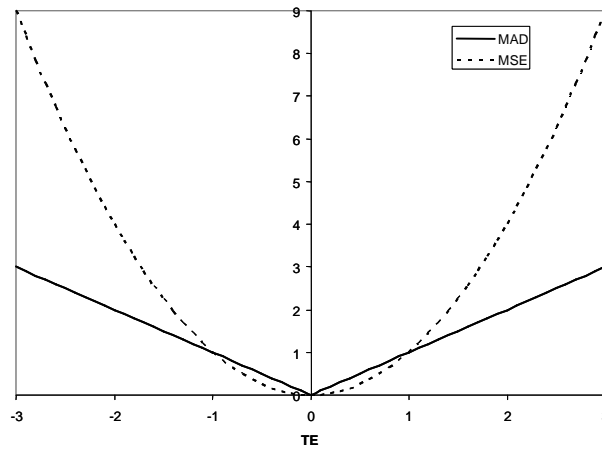


Figure 3.5: Difference between MAD and MSE<sup>174</sup>

A variant of MSE is the tracking error variance (TEV), which is defined as the variance of the differences in returns:

$$TEV = \frac{1}{T} \sum_{t=1}^T (R_{I,t} - R_{P,t} - (\bar{R}_I - \bar{R}_P))^2 = \frac{1}{T} \sum_{t=1}^T (TE_t^*)^2,$$

where  $\bar{R}_P$  is the mean of the returns of the tracking portfolio,  $\bar{R}_I$  is the mean of the returns of the index, and  $TE_t^*$  is the TE, adjusted for bias. Sometimes, the standard deviation of the TE is used:

$$TESTD = \sqrt{TEV}.$$

---

<sup>174</sup> See Roßbach/Karlow (2011 p. 8).

While the TEV is used very often, it does have the disadvantage of including a shift.<sup>175</sup> This means that, if the return on the tracking portfolio minus the return on the index is constant, then the TE is a constant and its variance is zero, although there is a deviation from the index. This is illustrated in Figure 3.6. Portfolio 1 has an obviously worse tracking quality than Portfolio 2 but the TEV of Portfolio 1 is equal to zero while the TEV of Portfolio 2 is greater than zero.

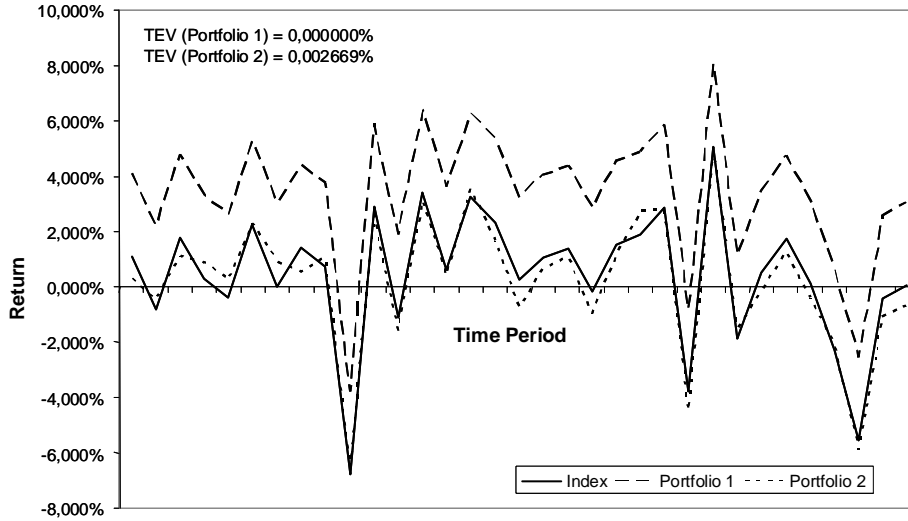


Figure 3.6: Weakness of the TEV<sup>176</sup>

It is possible to show that MSE and TEV are related:

$$\begin{aligned} TEV &= Var[R_I - R_P] = E[(R_I - R_P - (\bar{R}_I - \bar{R}_P))^2] = \\ &E[(R_I - R_P)^2] - (E[R_{I,t} - R_{P,t}])^2 = MSE - (\bar{R}_I - \bar{R}_P)^2. \end{aligned}$$

TEV is not suitable to be used alone to evaluate the quality of a tracking portfolio if  $(\bar{R}_I - \bar{R}_P)^2 \neq 0$ . The expectation of the TE should also be reported. Therefore, MSE is more appropriate, since it is the sum of the variance of the TE and its squared expected value.

A further weakness of the traditional measures of tracking quality (MAD, MSE, and TEV) is that they do not take into account negative serial autocorrelation in the TE.<sup>177</sup> This leads to an upward bias if more frequent data are used, so that the TE measure resulting from daily data will be larger than that resulting from weekly or monthly data. The negative serial

<sup>175</sup> See Beasley/Meade/Chang (2003, p. 629).

<sup>176</sup> See Roßbach/Karlow (2011 p. 8).

<sup>177</sup> See Pope/Yadav (1994, p. 30). The negative serial autocorrelation is the negative correlation of TE with itself.

autocorrelation of the TE follows from the empirical fact that the asset returns are serially autocorrelated.<sup>178</sup>

A good measure of tracking quality must evaluate the ex post tracking quality from the viewpoint of an investor. For an investor, a tracking portfolio is of high quality if it performs similarly to the index. Similarity in this context means that the tracking portfolio and the index do not really deviate from each other in terms of growth in value. The more similar a tracking portfolio is to the index, the higher quality it is in the eyes of the investor. The period over which the quality of tracking is validated we call the investment period.

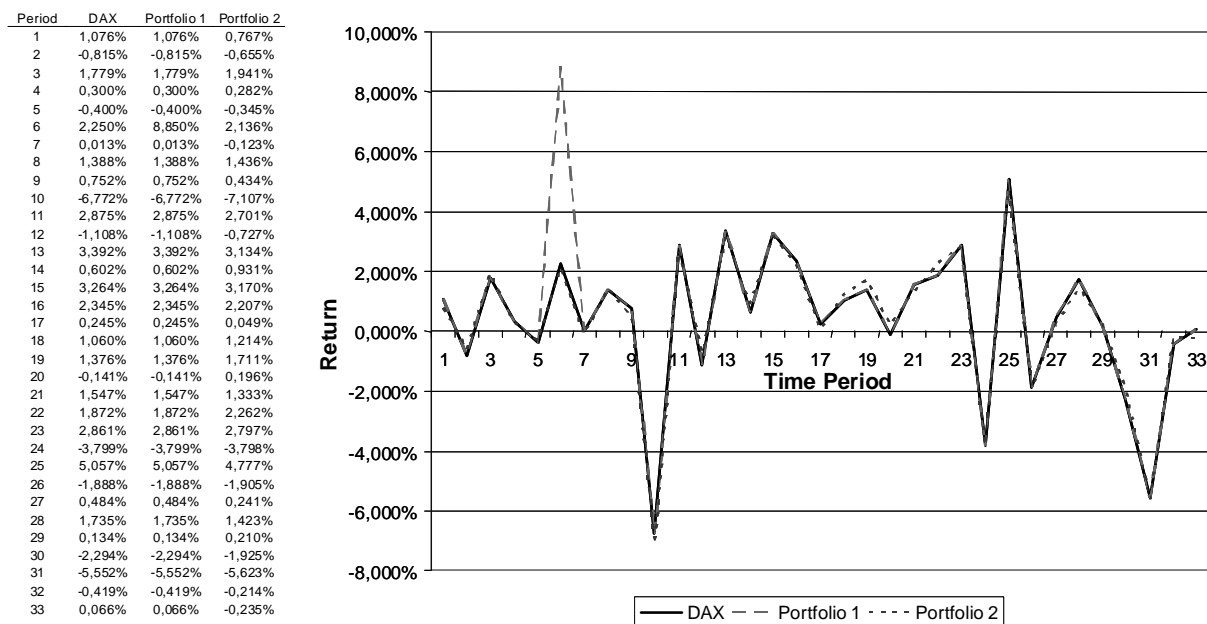


Figure 3.7: Returns of the DAX and Two Tracking Portfolios with Equal MADs<sup>179</sup>

Unfortunately, neither the MAD nor the MSE reflect this definition of quality regarding deviations in growth in value. This can be demonstrated easily. Figure 3.7 shows the returns of the DAX and two artificially constructed tracking portfolios. Tracking Portfolio 1 has exactly the same returns as the index, except in period 6 where the return has a surplus of 6.6 percent. On the other hand, Tracking Portfolio 2 has small deviations from the index in every period. The MADs of both tracking portfolios are equal. This means that, according to MAD, both tracking portfolios have identical tracking quality.

<sup>178</sup> See Lo/MacKinlay (1988).

<sup>179</sup> See Roßbach/Karlow (2011 p. 11).

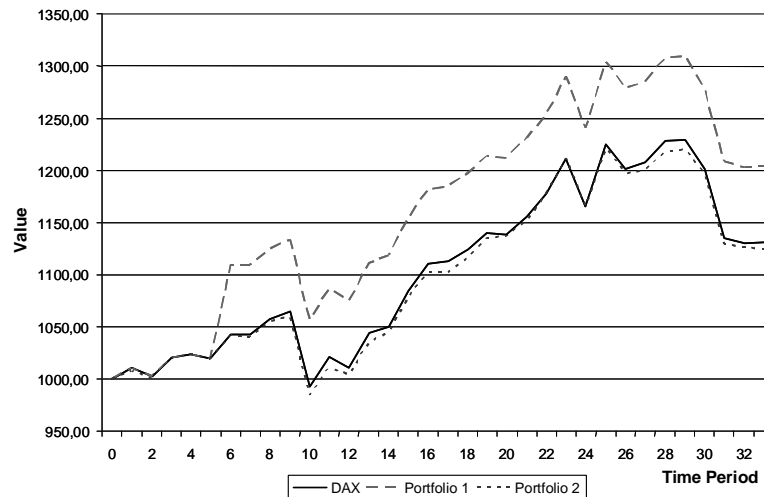


Figure 3.8: Values of a DAX and Two Tracking Portfolios with Equal MADs<sup>180</sup>

Assuming one invests 1,000 Euros in the index and in each tracking portfolio in period 0, the growth of the three portfolios would be as illustrated in Figure 3.8. It is obvious that Tracking Portfolio 1 is much worse at tracking the index than Tracking Portfolio 2. Thus, the MAD is not a feasible way of measuring the ex post tracking quality.

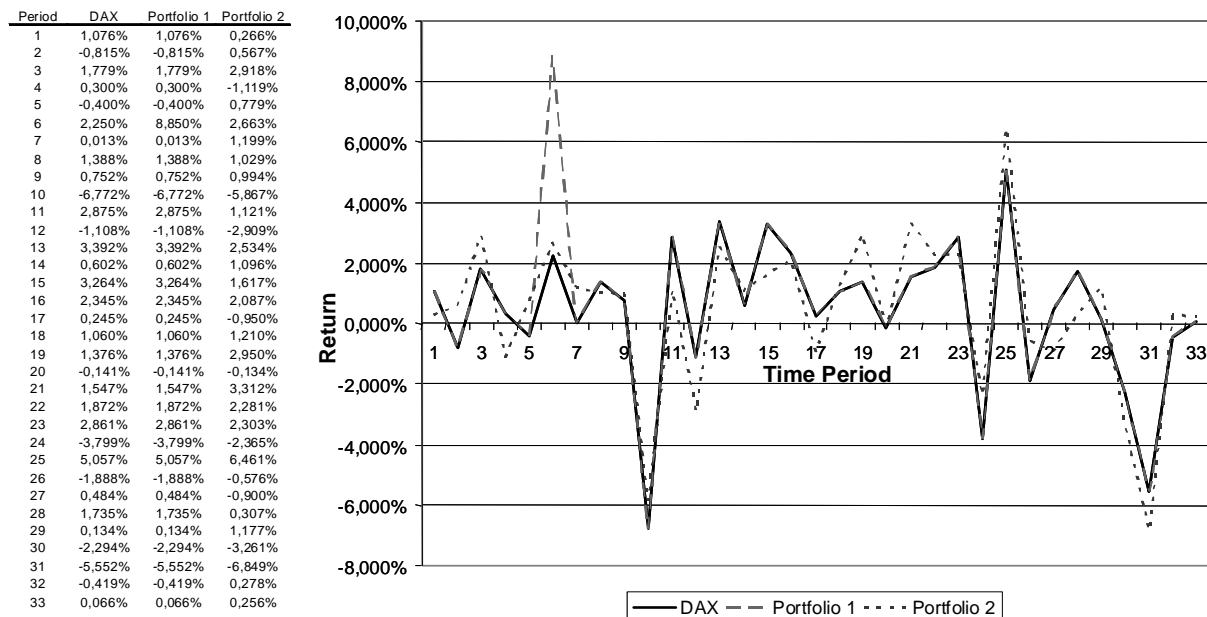


Figure 3.9: Returns of the DAX and Two Tracking Portfolios with Equal MSEs<sup>181</sup>

Repeating this experiment with the MSE measure shows similar results. The index and Tracking Portfolio 1 are the same as in the previous example but a different Tracking Portfolio 2, this time with the same MSE as Tracking Portfolio 1, is generated (see Figure 3.9).

<sup>180</sup> See Roßbach/Karlow (2011 p. 12).

<sup>181</sup> See Roßbach/Karlow (2011 p. 12).

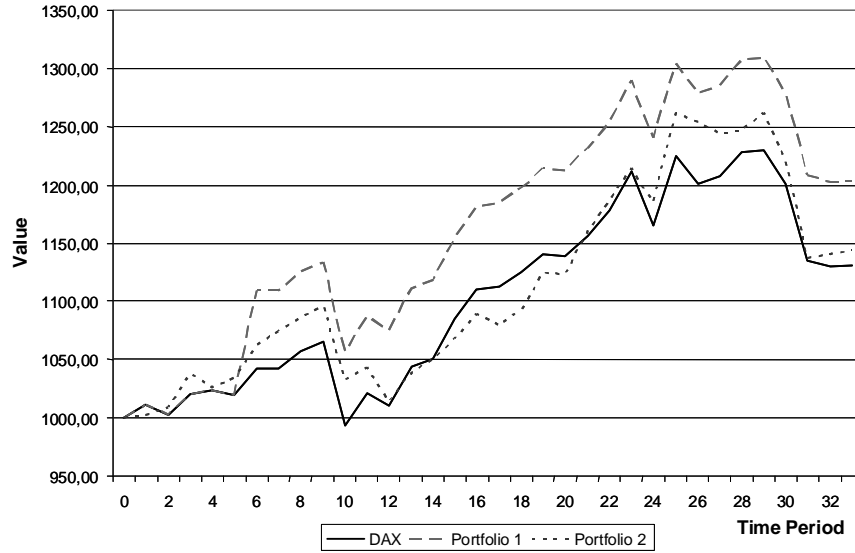


Figure 3.10: Values of the DAX and Two Tracking Portfolios with Equal MSEs<sup>182</sup>

Figure 3.10 shows that Tracking Portfolio 1 is again worse than Tracking Portfolio 2 at tracking the index. Thus, the MSE is not a feasible approach either, in terms of measuring the ex post tracking quality.

This raises the question of which measure should be used for tracking quality in the investment period. In the next section, we introduce the ex post tracking quality measure, which is able to evaluate the closeness of a tracking portfolio to an index, in terms of value.

### 3.2.2 Ex post tracking quality

An ex post tracking measure that calculates the deviations between the values of the tracking portfolio and the index seems to be most appropriate from the investors' point of view. Simply taking the difference between the terminal value of the tracking portfolio and that of the index portfolio at the end of the time interval could be subject to too much chance. A useful measure of the ex post tracking quality should measure the mean absolute differences across all periods of the time interval.

An ex post tracking quality (ETQ) measure satisfying this requirement can be calculated using the formula<sup>183</sup>

$$ETQ = \frac{1}{T} \sum_{t=1}^T \left| \frac{V_{I,t}}{V_{I,0}} - \frac{V_{P,t}}{V_{P,0}} \right|, \quad (3.3)$$

<sup>182</sup> See Roßbach/Karlow (2011 p. 13).

<sup>183</sup> A similar measure is used by Gaivoronski/Krylov/van der Wijst (2005, p. 120). However, their measure cannot be interpreted as the area between the value curves of the index and the tracking portfolio.

where  $V_P$  is the value of the tracking portfolio, and  $V_I$  is the value of the index.

Graphically, this measure can be interpreted as the area between two curves — one for the index and one for the tracking portfolio — that are the result of investing one monetary unit in the index and in the tracking portfolio, respectively, at the beginning of the time period (see Figure 3.11).

This measure expresses the average difference between the value of the tracking portfolio and the value of the index relative to their initial values at time  $t = 0$ . Its application clearly shows that Tracking Portfolio 1 is worse in both of the cases illustrated in Figure 3.8 and Figure 3.10. Tracking Portfolio 1 has an ETQ of 6.149%, while Tracking Portfolio 2 in Figure 3.8 has an ETQ of 0.468% and Tracking Portfolio 2 in Figure 3.10 has an ETQ of 1.884%. Thus, the measure meets the requirements stated above.

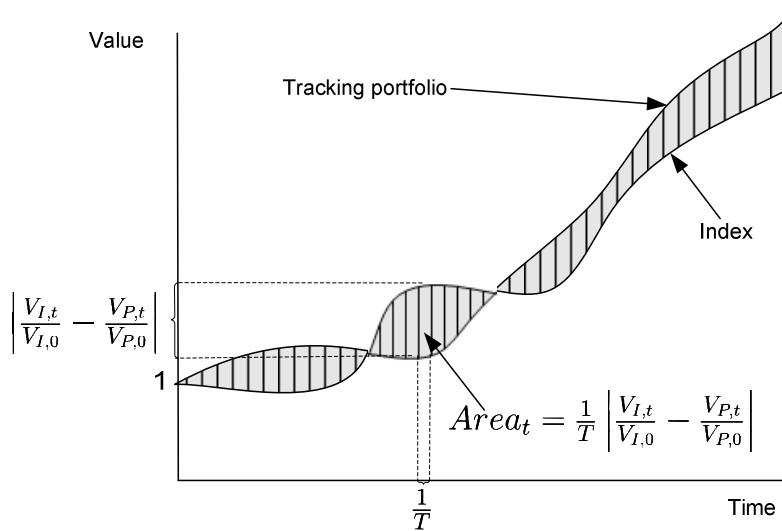


Figure 3.11: Ex Post Tracking Quality Shown as the Area between Two Curves

Moreover, this measure is not affected by the negative autocorrelation of the returns if more frequent data are used since it depends on the values of the tracking portfolio and the index. Additionally, the compounding of these values is taken into consideration. This can be seen from the equation for compounded returns:

$$R_t^c = \frac{V_t}{V_0} - 1 = \prod_{k=1}^t (1 + R_k) - 1, \quad (3.4)$$

where  $R_t^c$  is the return compounded up to period  $t$ , and  $R_k$  is the simple return for period  $k$ .

Using this equation, the ETQ (3.3) can be rewritten as the sum of the absolute differences between the compounded returns:

$$ETQ = \frac{1}{T} \sum_{t=1}^T |R_{I,t}^c - R_{P,t}^c|.$$

Since ETQ is based on the compounded returns it is a good measure for evaluating the long-term quality of a tracking portfolio. The use of the ETQ demonstrates the difference between the values of the estimated tracking portfolios and that of the index. In contrast, the traditional measures are better for evaluating short-term tracking quality, since they are computed based on the TE over a short time period and the sequence of the returns by their computation is not taken into consideration.

### 3.3 Selected methods of optimized sampling

Tracking quality measures are not only used to indicate the performance of a tracking portfolio. They can also be used as a basis for constructing tracking portfolios. In this section, we describe methods of optimized sampling that are based on minimizing tracking quality measures.

#### 3.3.1 Minimizing traditional measures

The general optimization problem for optimized sampling to find the weights for the tracking portfolio is based on the minimization of the measure of tracking quality (e.g., MAD), which in turn depends on the TE:

$$\underset{\mathbf{w}}{\text{minimize}} \quad \text{measure}(\text{TE}). \quad (3.5)$$

The weights of the tracking portfolio should sum to one:

$$\sum_{i=1}^N w_i = 1.$$

Any of the traditional measures described in Section 3.2 can be used as the measure of tracking quality in the optimization problem (3.5).

The MAD and RMSE measures from Section 3.3.1 can be extended using the generalized traditional measure (GTM):<sup>184</sup>

$$GTM = \left( \sum_{t=1}^T \Delta_t |TE_t|^p \right)^{1/p},$$

where  $\Delta_t$  is a weighting factor and  $p > 0$  is a power. Depending on the choice of factor  $p$  it is possible to construct different measures of tracking quality. For example, this measure is equal to MAD if  $p = 1$  and to RMSE if  $p = 2$ .

Basically, the weighting factor  $\Delta_t$  is equal to  $1/T$  for all  $t$ . However, it is possible to assign larger weights to more important data points and smaller weights to less important data points. For example, the older data points are less important, therefore smaller weighting

---

<sup>184</sup> See Beasley/Meade/Chang (2003, p. 629).



factors  $\Delta_t$  can be assigned to them than to the more recent data points. It is possible to use an exponential weighting function to achieve this, such as the exponentially-weighted moving average:<sup>185</sup>

$$\Delta_t = (1 - \alpha)\alpha^{t-1}, t = 1, \dots, T, 0 < \alpha \leq 1.$$

Additionally, different constraints can be added to the optimization problem (3.5). Typical constraints, which prohibit the short selling of assets, are

$$w_i \geq 0, i = 1, \dots, N.$$

Depending on the measure of tracking quality (for example MAD, MSE, or TEV) that is plugged into the optimization problem, tracking portfolios with different risk/return properties can be constructed<sup>186</sup> but what would be an appropriate measure to be minimized? It is possible to rewrite the optimization problem as a regression problem. The tracking error  $TE_t^*$  is an error term, usually denoted by  $e_t$ . Thus, we can write

$$R_{I,t} = R_{P,t} + b + TE_t^* = R_{P,t} + b + e_t = \sum_{i=1}^N w_i r_{i,t} + b + e_t,$$

where  $b$  is a constant difference between the tracking portfolio and the index, usually equal to zero. We can adjust the TE for bias by adapting the TE given in equation (3.1):

$$TE_t^* = R_{I,t} - R_{P,t} - b = TE_t - b \quad (3.6)$$

If term  $b \neq 0$ , then we are interested in minimizing the variability of the returns of the tracking portfolio around the returns of the index. The value of parameter  $b$  is the value, which is determined by the optimization procedure in the estimation period. The usage of the term  $b$  allows for the fact that we are not taking into account the constant deviations between the index returns and the tracking portfolio returns during the optimization. This allows us to concentrate on the variability of the tracking portfolio around the index. The TE in form (3.6) can be used in the optimization instead of the non-adjusted  $TE_t$  (equation (3.1)), but for performance measurement the usage of the term  $b$  has the disadvantage of causing a shift, as was demonstrated for TEV, where  $b = \bar{R}_I - \bar{R}_P$ , in Section 3.2.1.

---

<sup>185</sup> See Jorion (2007, p. 230).

<sup>186</sup> See, for example, Rudolf/Wolter/Zimmermann (1999, p. 102), or Rey/Seiler (2001).

The probabilistic characteristics of the error term have an essential influence on the estimation of the parameters  $w_i$ . Thus, it can be shown that, if the error terms are IID<sup>187</sup> and follow a normal distribution, the minimization of the MSE gives the best linear unbiased estimator.<sup>188</sup> However, if the error terms are IID and follow a Laplace distribution, then it can be shown that the minimization of the MAD gives the most likely estimator.

The first problem that arises from the regression formulation is that the regression problem for the index tracking problem violates the assumptions (zero expectation of TE, independence of TEs, constant variance of TEs, and independence of the estimated parameters and TEs) used for regression problems, since the explained and explanatory variables are correlated returns and the weights of the tracking portfolio are not time invariant.<sup>189</sup> The second problem is that the constraints used in the formulation of the index tracking problem can add bias to the estimator.<sup>190</sup> To fix these problems, it is preferable to use robust methods to estimate the weights of the tracking portfolio. In the following sections, we describe the minimax and conditional value-at-risk approaches.

### 3.3.2 Minimax approach

The minimax approach is the worst-case approach, where the aim is to minimize the MAXD between the index and the tracking portfolio.<sup>191</sup> In this case, the measure of tracking quality used in the optimization problem (3.5) is formulated as follows:

$$MAXD = \max_{t=1, \dots, T} |TE_t^*|$$

To solve the resulting optimization problem, it must be reformulated as

$$\begin{aligned} & \underset{\mathbf{w}, \theta}{\text{minimize}} && \theta \\ & \text{subject to} && TE_t^* \leq \theta \\ & && -TE_t^* \leq \theta \\ & && t = 1, \dots, T \\ & && \sum_{i=1}^N w_i = 1 \\ & && w_i \geq 0, i = 1, \dots, N. \end{aligned}$$

---

<sup>187</sup> Independent and identically distributed.

<sup>188</sup> See Draper/Smith (1998, p. 567).

<sup>189</sup> See Bamberg/Wagner (2000).

<sup>190</sup> See Hansen (2012, p. 151)

<sup>191</sup> See Rudolf/Wolter/Zimmermann (1999, p. 88).

Solving this optimization problem, we find the smallest value  $\theta^*$ , such that, for all data points  $t = 1, \dots, T$ , the difference between the return of the index and that of the tracking portfolio is less than or equal to this value:

$$-\theta^* \leq TE_t^* \leq \theta^*.$$

The minimax approach can be extended to the conditional value-at-risk approach by allowing a certain number of absolute deviations to fall outside the interval  $[-\theta^*, \theta^*]$ .

### 3.3.3 Minimization of conditional value-at-risk

The conditional value-at-risk (CVAR) approach to index tracking can be formulated using the following measure of tracking quality:

$$CVAR_\alpha = \varepsilon + \frac{1}{T(1-\alpha)} \sum_{t=1}^T \max(0, |TE_t^*| - \varepsilon),$$

where parameter  $(1-\alpha) \in [0, 1]$  is pre-selected and specifies the probability that the deviations from the index are larger than parameter  $\varepsilon$ .<sup>192</sup> This parameter  $\varepsilon$  is not fixed and is determined by solving the optimization problem with fixed level  $\alpha$ . Hence, the value of parameter  $\varepsilon$  is equal to the value-at-risk (VAR) with confidence level  $\alpha$ . The CVAR measure can be interpreted as the mean of all absolute deviations of the tracking portfolio from the index that are larger than  $\varepsilon$ . Any deviations that are smaller than  $\varepsilon$  are ignored. The MAXD and CVAR measures are illustrated in Figure 3.12.

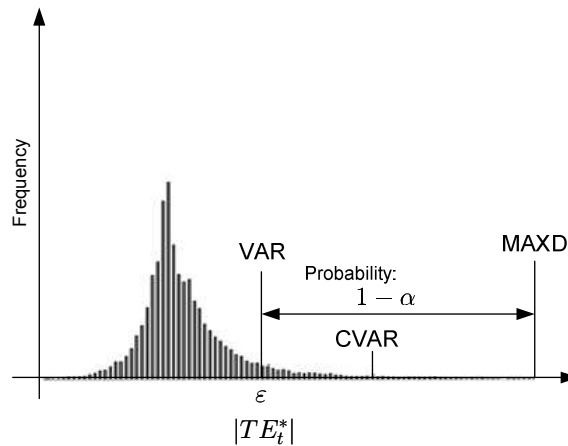


Figure 3.12: MAXD and CVAR<sup>193</sup>

<sup>192</sup> See Rockafellar/Uryasev (2002).

<sup>193</sup> This figure is adapted from Uryasev (2000).

By introducing error slopes  $\xi_t, \xi_t^*$  such that

$$\begin{aligned}\xi_t &= \max(0, TE_t^* - \varepsilon) \\ \xi_t^* &= \max(0, -TE_t^* - \varepsilon) \\ t &= 1, \dots, T,\end{aligned}$$

we can formulate the optimization problem (3.5) for the CVAR measure as

$$\begin{aligned}\underset{\mathbf{w}, \varepsilon, \xi, \xi^*}{\text{minimize}} \quad & \varepsilon + \frac{1}{T(1-\alpha)} \sum_{t=1}^T (\xi_t + \xi_t^*) \\ \text{subject to} \quad & TE_t^* \leq \varepsilon + \xi_t \\ & -TE_t^* \leq \varepsilon + \xi_t^* \\ & \xi_t \geq 0, \xi_t^* \geq 0 \\ & t = 1, \dots, T \\ & \sum_{i=1}^N w_i = 1 \\ & w_i \geq 0, i = 1, \dots, N.\end{aligned} \tag{3.7}$$

Hence, any absolute differences between the returns of the tracking portfolio and the index that are larger than parameter  $\varepsilon$  are penalized by the constant  $\frac{1}{T(1-\alpha)}$ .

### 3.3.4 Portfolio structure: weights versus numbers

The evolution of the value of the tracking portfolio through time depends on the assets that make up the portfolio. There are two ways to describe the positions of these assets within the portfolio. The first way is to use the number of units of each asset. The second is to use the weight of each asset.

Studies of the index tracking problem model the portfolio structure in different ways, using fixed weights or fixed numbers (see Section 3.1.2.2). We can illustrate the difference between these two methods using the following example:

Suppose we take a portfolio containing 41 MOLX (Molex Incorporated) stocks, which corresponds to 1.21% of the portfolio value at the beginning of the investment period. Over time, the number of stocks required to hold this weight constant must vary due to the relative price changes of the portfolio components (see left side of Figure 3.13). This follows from the formula

$$n_{i,t} = \frac{w_i V_{P,t}}{S_{i,t}}.$$

Hence, transactions must be made to keep the weights constant. Alternatively, we can fix the number of MOLX stocks instead. Now, the relative weight of this stock will vary over time (see right side of Figure 3.13).

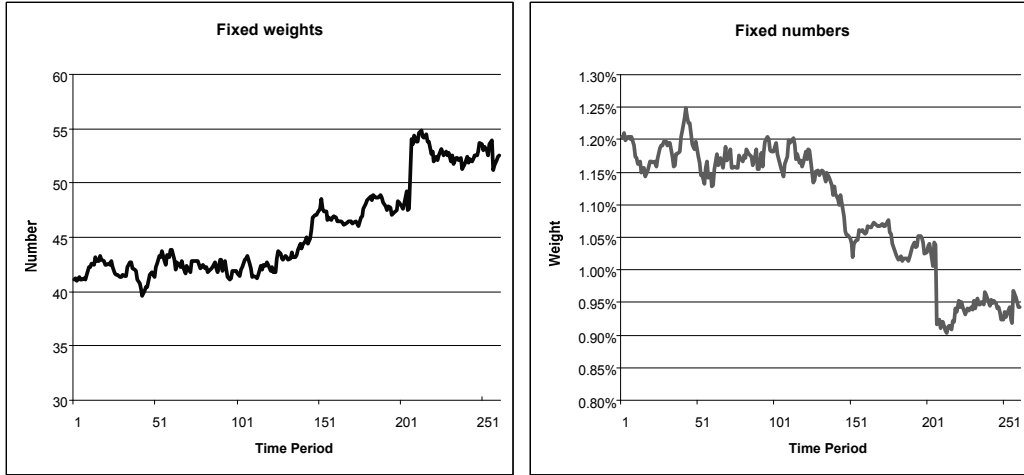


Figure 3.13: Weights versus Numbers

The weight at a fixed point in time is given by

$$w_{i,t} = \frac{n_i S_{i,t}}{V_{P,t}}.$$

Hence, we can use either fixed weights or fixed numbers of stocks to describe the evolution of the value of an asset within the portfolio over time but these two approaches are not equivalent.

A buy-and-hold strategy is based on using fixed numbers of stocks. The implementation of this strategy assumes that the assets in the tracking portfolio are bought at some point in time and then held without making any changes. From the example above, we can see that a portfolio made up of fixed weights of assets is not consistent with a buy-and-hold strategy. In contrast, a portfolio based on fixed numbers of assets is a natural choice for tracking the index.

To make the sampling consistent with a buy-and-hold strategy, different methods can be used. The return on the portfolio at time  $t$  can either be calculated based on time-dependent variable weights or on fixed numbers of stocks:

$$R_{P,t} = \sum_{i=1}^N w_{i,t} r_{i,t} = \frac{\sum_{i=1}^N n_i S_{i,t}}{\sum_{i=1}^N n_i S_{i,t-1}} - 1. \quad (3.8)$$

If we use weights for the tracking portfolio, we should remove their dependence from time. If we use numbers, we should remove the variables in the denominator to make the function of returns linear. The different strategies that can be used to transform the returns of the tracking portfolio, shown in equation (3.8), so that they can be used in the optimization problem are as follows:

### Simple approximation

To avoid computational complexity, the variable weights are usually approximated by fixed weights:<sup>194</sup>

$$R_{P,t} = \sum_{i=1}^N w_{i,t} r_{i,t} = [w_{i,t} \approx w_i] \approx \sum_{i=1}^N w_i r_{i,t}.$$

This is the simplest method.

### Correction at the beginning of the period

Now we assume that the relationships between the value of the portfolio at the start of the estimation period and its values at any other points in time are approximately equal to the corresponding relations for the index itself. Then, the variable weights can be transformed into fixed weights using the following correction factor:<sup>195</sup>

$$\begin{aligned} w_{i,t} = \frac{n_i S_{i,t}}{V_{P,t}} &= \left[ n_i = \frac{w_{i,0} V_{P,0}}{S_{i,0}} \right] = \\ w_{i,0} \frac{V_{P,0}}{V_{P,t}} \frac{S_{i,t}}{S_{i,0}} &= \left[ \frac{V_{P,0}}{V_{P,t}} \approx \frac{V_{I,0}}{V_{I,t}} \right] \approx w_{i,0} \frac{V_{I,0}}{V_{I,t}} \frac{S_{i,t}}{S_{i,0}}. \end{aligned} \quad (3.9)$$

*Bamberg and Wagner*<sup>196</sup> correct the weights at the beginning of the period, using returns, as follows:

$$w_{i,t} \approx w_{i,0} \prod_{k=1}^t \frac{1 + r_{i,k}}{1 + R_{I,k}}. \quad (3.10)$$

---

<sup>194</sup> See, for example, Meade/Salkin (1989, p. 878 (ii)).

<sup>195</sup> See Montfort/Visser/van Draat (2008).

<sup>196</sup> See Bamberg/Wagner (2000).

It is possible to show that (3.9) is equivalent to (3.10). Reformulating equation (3.9) gives:

$$w_{i,t} \approx w_{i,0} \frac{V_{I,0}}{V_{I,t}} \frac{S_{i,t}}{S_{i,0}} = w_{i,0} \frac{1 + \left[ \frac{S_{i,t}}{S_{i,0}} - 1 \right]}{1 + \left[ \frac{V_{I,t}}{V_{I,0}} - 1 \right]}. \quad (3.11)$$

Substituting the expressions in squared brackets in (3.11) with the equation for compounded returns (3.4) completes the proof.

### Correction at the end of the period

In the same way, we can make adjustments to the weights at the end of the period:

$$\begin{aligned} w_{i,t} &= \frac{n_i S_{i,t}}{V_{P,t}} = \left[ n_i = \frac{w_{i,T} V_{P,T}}{S_{i,T}} \right] = \\ &= w_{i,T} \frac{V_{P,T}}{V_{P,t}} \frac{S_{i,t}}{S_{i,T}} = \left[ \frac{V_{P,T}}{V_{P,t}} \approx \frac{V_{I,T}}{V_{I,t}} \right] \approx w_{i,T} \frac{V_{I,T}}{V_{I,t}} \frac{S_{i,t}}{S_{i,T}}. \end{aligned}$$

### The use of compounded returns

It is also possible to use a measure of tracking quality that is based on compounded returns. In this case,

$$TE_t^c = R_{I,t}^c - R_{P,t}^c$$

We can define different measures of tracking quality in the same way as we did for simple returns. For example, we can define the MAD measure for compounded returns as follows:

$$MAD^c = \frac{1}{T} \sum_{t=1}^T |R_{I,t}^c - R_{P,t}^c| = \frac{1}{T} \sum_{t=1}^T |TE_t^c|.$$

Substituting equation (3.4) into this instead of the compounded returns gives

$$MAD^c = \frac{1}{T} \sum_{t=1}^T \left| \frac{V_{I,t}}{V_{I,0}} - \frac{V_{P,t}}{V_{P,0}} \right|. \quad (3.12)$$

Note that this measure is the same as the ex post tracking quality measure (see equation (3.3)).

To use this measure for the optimization, we must transform it. First, we rewrite the relation between the portfolio values as

$$\frac{V_{P,t}}{V_{P,0}} = \sum_{i=1}^N \frac{S_{i,t} n_i}{V_{P,0}} = \sum_{i=1}^N \frac{S_{i,t} n_i}{V_{P,0}} \left[ \frac{S_{i,0} V_{I,0}}{S_{i,0} V_{I,0}} \right] = \sum_{i=1}^N \left[ \frac{n_i S_{i,0}}{V_{I,0}} \right] \frac{V_{I,0}}{V_{P,0}} \frac{S_{i,t}}{S_{i,0}} = \sum_{i=1}^N w_{i,0} \frac{V_{I,0}}{V_{P,0}} \frac{S_{i,t}}{S_{i,0}}. \quad (3.13)$$

Substituting (3.13) into (3.12) we obtain

$$MAD^c = \sum_{t=1}^T \left| \frac{V_{I,t}}{V_{I,0}} - \sum_{i=1}^N w_{i,0} \frac{V_{I,0}}{V_{P,0}} \frac{S_{i,t}}{S_{i,0}} \right|.$$

Then we can use this measure in the optimization problem (3.5). The  $MAD^c$  measure, based on compounded returns, takes into account the sequence of the returns, which may help us to find tracking portfolios with high tracking quality.

#### **Assumption that the portfolio value is equal to the index value in the previous period**

For the next approximation, we assume that

$$V_{P,t-1} \approx V_{I,t-1}. \quad (3.14)$$

The returns of the portfolio can then be written as

$$R_{P,t} = \frac{V_{P,t} - V_{P,t-1}}{V_{P,t-1}} \approx \frac{V_{P,t} - V_{I,t-1}}{V_{I,t-1}},$$

where the value of the portfolio in the previous period is completely replaced by the index value. We need an additional constraint here:

$$\sum_{i=1}^N n_i S_{i,0} = V_{I,0}, \quad (3.15)$$

where the portfolio value and the index value are equal at the beginning of the estimation period.

#### **Assumption that the portfolio return is related to the index value in the previous period**

Here we use the same assumption (3.14) as above but now the returns of the tracking portfolio are approximated as follows:<sup>197</sup>

---

<sup>197</sup> See Meade/Salkin (1989, p. 878 (i)).



$$R_{P,t} = \frac{V_{P,t} - V_{P,t-1}}{V_{P,t-1}} \approx \frac{V_{P,t} - V_{P,t-1}}{V_{I,t-1}},$$

where the value of the portfolio is replaced by the value of the index in the denominator. As before, it is necessary to use constraint (3.15).

We systematize the approximation methods described above in Table 3.3.

Simple approximation	$R_{P,t} = \sum_{i=1}^N w_i r_{i,t}$
Correction at the beginning of the period	$R_{P,t} = \sum_{i=1}^N w_{i,0} \frac{V_{I,0}}{V_{I,t}} \frac{S_{i,t}}{S_{i,0}} r_{i,t}$
Correction at the end of the period	$R_{P,t} = \sum_{i=1}^N w_{i,T} \frac{V_{I,T}}{V_{I,t}} \frac{S_{i,t}}{S_{i,T}} r_{i,t}$
The use of compounding returns	$R_{P,t}^c = \sum_{i=1}^N w_{i,0} \frac{V_{I,0}}{V_{P,0}} \frac{S_{i,t}}{S_{i,0}} - 1$
Assumption that the portfolio value is equal to the index value in the previous period	$R_{P,t} = \frac{\sum_{i=1}^N n_i S_{i,t} - V_{I,t-1}}{V_{I,t-1}}$
Assumption that the portfolio return is related to the index value in the previous period	$R_{P,t} = \frac{\sum_{i=1}^N n_i S_{i,t} - \sum_{i=1}^N n_i S_{i,t-1}}{V_{I,t-1}}$

Table 3.3: Approximations of Tracking Portfolio Returns

We have seen that there are different ways to approximate the tracking portfolio returns for use in the optimization problem, but which is the best? We test these approaches empirically in Section 4.1.

### 3.4 Requirements and performance assessment

In this section, we define the requirements of the sampling process and ways of assessing the performance of different methods.

#### 3.4.1 The absolute stability of the tracking quality

The index tracking problem is solved by a sampling procedure using past data. We call the period used to estimate the portfolio weights the estimation period. To validate the quality of tracking, it is necessary to invest in the portfolio and check how well it tracks the index over a new period. We call this the investment period.

The quality of a tracking portfolio, selected using one of the above methods of optimized sampling, can be measured by any of the measures described in Section 3.2. We expect the absolute value of this measure to be at roughly the same level in the estimation period as in the investment period. The ratio and the difference between the absolute values of a measure of tracking quality in the two periods are alternative indicators of the absolute stability (AS) of the tracking quality:

$$AS_{ratio} = \frac{err_{inv}}{err_{est}} \quad (3.16)$$

$$AS_{diff} = err_{inv} - err_{est}, \quad (3.17)$$

where  $err_{inv}$  is the value of the tracking quality measure in the estimation period and  $err_{est}$  is its value in the investment period. If the sampling method produces a tracking portfolio with absolutely stable tracking quality, then  $AS_{ratio}$  will be equal to one and  $AS_{diff}$  to zero. Equations (3.16) and (3.17) can be rewritten as:

$$err_{inv} = err_{est} AS_{ratio}$$

$$err_{inv} = err_{est} + AS_{diff}.$$

Hence, the value of the tracking quality measure in the investment period depends on the value of the tracking quality measure in the estimation period, corrected by the  $AS$  indicator. It is possible to quantify the  $AS_{ratio}$  and  $AS_{diff}$  indicators using the data in the estimation period. This will be shown in Section 5.1.4.3. As a consequence, one can estimate the value of a tracking quality measure in the investment period.

### 3.4.2 The relative stability of the tracking quality

The selection of tracking portfolios from a set of candidate tracking portfolios assumes that those portfolios with high tracking quality during the estimation period will also have high tracking quality during the investment period. However, two types of error can occur (see Figure 3.14).

		Investment	
		High	Low
Estimation	High	Correct prediction	Error Type I
	Low	Error Type II	Correct prediction

Figure 3.14: Prediction Errors <sup>198</sup>

The first type is when the predicted quality of the tracking portfolio using the estimation period is high but the portfolio actually has low tracking quality in the investment period. Investing in this tracking portfolio will cause the investor to incur a direct loss. The second type of error involves opportunity loss. The tracking portfolio in this case is predicted to have low tracking quality according to the estimation period, but in fact will have high tracking quality in the investment period.

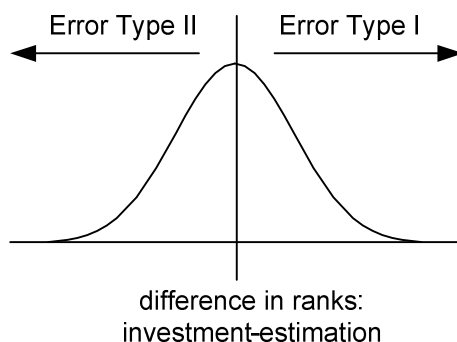


Figure 3.15: Prediction Errors Based on Relative Quality in the Investment and Estimation Periods

To identify the two types of error we can sort the tracking portfolios from highest to lowest tracking quality according to the estimation period and assign ranks to them, and then do the same for the investment period. A positive difference between a tracking portfolio's rank in the investment period and its rank in the estimation period is an indication of a type I error. A negative difference indicates a type II error. We can draw a histogram of the distribution of

<sup>198</sup> Adapted from Sobhart/Keenan/Stein (2000, p.6) and applied to the index tracking problem.

differences in ranks for different tracking portfolios (see Figure 3.15). Type II errors are on the left and type I errors on the right.

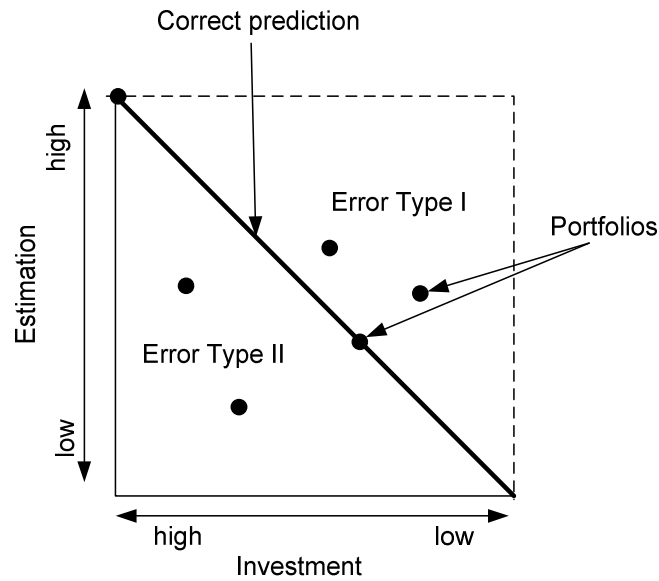


Figure 3.16: Illustration of the Two Types of Error

Another example illustration of this effect is given in Figure 3.16. As before, the tracking portfolios are sorted by tracking quality measure from high to low quality, separately, for the estimation and investment periods. If the rank of the tracking portfolio in the estimation period is the same as its rank in the investment period, then the portfolio will lie on the diagonal. If the tracking quality of the portfolio is lower in the investment period, then its quality has been overestimated and it will lie above the diagonal. Portfolios whose quality has been underestimated lie below the diagonal.

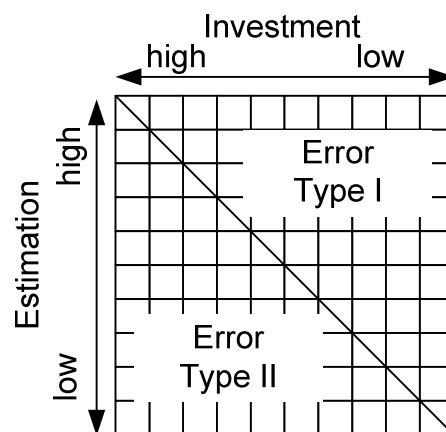


Figure 3.17: The Two Types of Error Depicted in a Dispersion Table

To quantify the two types of error we can also use quantiles. In this case, we rank candidate tracking portfolios according to their tracking quality in the estimation and investment periods and then separate the portfolios into quantiles according to their ranks, for each time period.

The resulting table indicates poor stability if a high percentage of the tracking portfolios are in different quantiles according to the estimation and investment periods (see Figure 3.17).

Formally, the proportion of portfolios in each quantile can be defined as

$$a_{i,j} = \frac{M}{P} \sum_{p=1}^P \delta_{i,j}^p,$$

where  $M$  is the total number of quantiles,  $P$  is the total number of portfolios, and

$$\delta_{i,j}^p = \begin{cases} 1 & \text{if portfolio } p \text{ is in } i \text{ quantile in the estimation} \\ & \text{period and in } j \text{ quantile in the investment period,} \\ 0 & \text{otherwise.} \end{cases}$$

As a result, the table illustrated in Figure 3.17 can be represented by the matrix  $\mathbf{A}$ :

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & \dots & a_{1,M} \\ \dots & \dots & \dots \\ a_{M,1} & \dots & a_{M,M} \end{pmatrix}.$$

The sum of the values in each row and each column is equal to one by definition:

$$\sum_{i=1}^M a_{i,j} = \sum_{j=1}^M a_{i,j} = 1.$$

If none of the tracking portfolios change quantiles, then all of the elements on the diagonal of matrix  $\mathbf{A}$  will be equal to one:

$$a_{i,i} = 1, \quad i = 1, \dots, M.$$

If this is not the case, we can calculate a stability measure as follows:

$$RS = \frac{1}{M} \sum_{i=1}^M a_{i,i}.$$

Unfortunately, this measure does not take into account the values of the other off-diagonal elements. To do so, we need another measure of relative stability. The pictures of dispersion tables below show three possible cases (see Figure 3.18): the best case, a random case, and the worst case.

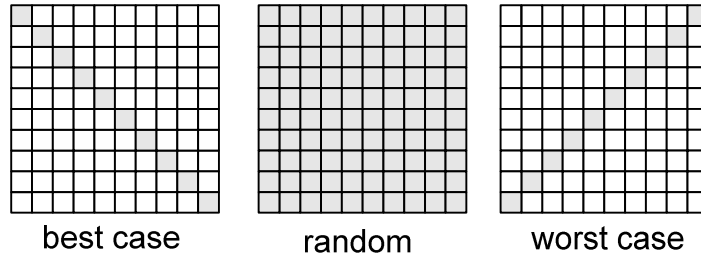


Figure 3.18: Three Cases of Relative Stability

In the best case, all portfolios remain in the same quantile in both periods. In the random case the proportion of portfolios in each cell of the dispersion table is equal. In the worst case, all portfolios move into the opposite quantile when we look at the investment period rather than the estimation period.

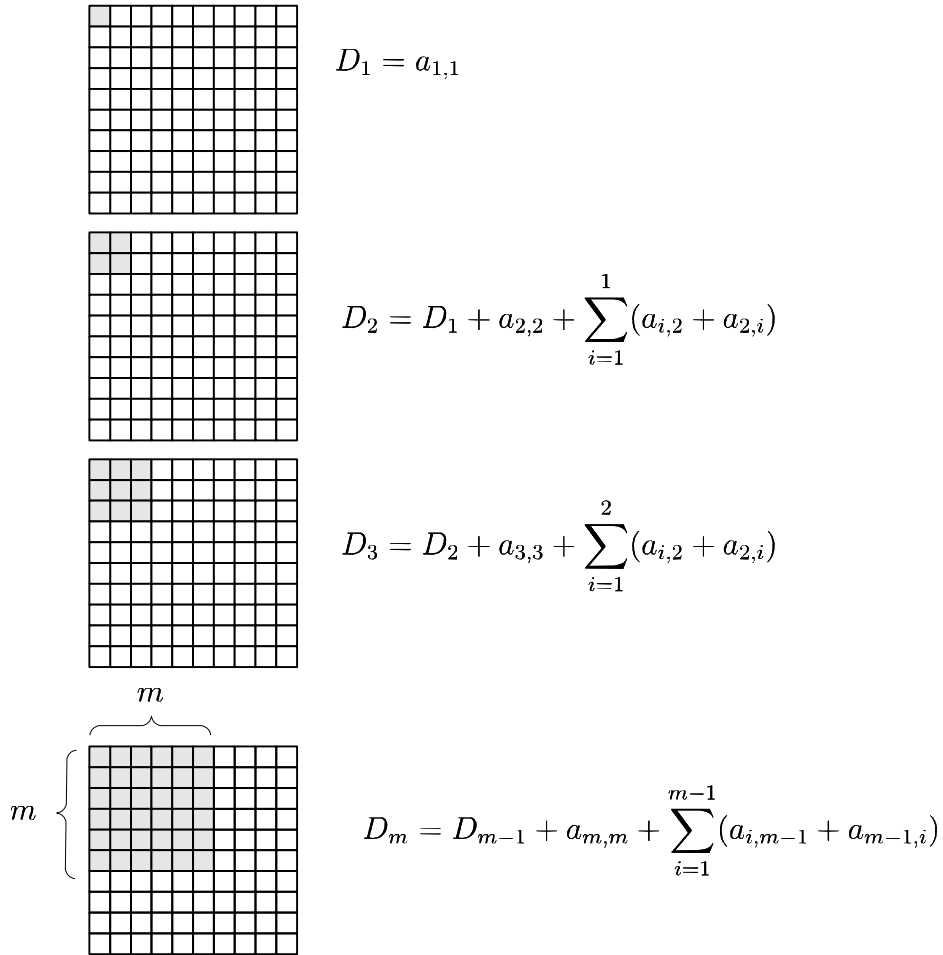


Figure 3.19: Type I and II Errors Together

To distinguish between these three cases we compute the indicator of relative stability (RS), which quantifies the dispersion effect as a single number. To compute RS, we first calculate the elements  $D_m$  in the dispersion table as follows:

$$D_1 = a_{1,1}$$

$$D_m = D_{m-1} + a_{m,m} + \sum_{i=1}^{m-1} (a_{i,m-1} + a_{m-1,i}).$$

These elements are illustrated in Figure 3.19. Computing them that way we quantify the concentration of the portfolios from high to low tracking quality in both (estimation and investment) periods. Then, RS can be computed as

$$RS = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \frac{D_m}{M} \right). \quad (3.18)$$

This indicator shows the average difference between the best-case dispersion table and the current dispersion table.

In the best case, all elements on the diagonal are equal to one, therefore

$$D_m = m.$$

Substituting  $D_m$  into equation (3.18), we have

$$RS^{best} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \frac{m}{M} \right) = 0.$$

In the random case, all elements in matrix  $\mathbf{A}$  are equal to  $\frac{1}{M}$ , therefore

$$D_m = \frac{m^2}{M}.$$

Substituting these values for  $D_m$  into equation (3.18), we obtain (see Appendix 1):

$$RS^{rand} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \left( \frac{m}{M} \right)^2 \right) = \frac{1}{6} - \frac{1}{6M^2}. \quad (3.19)$$

From equation (3.19) we have that, if the number of quantiles goes to infinity,  $M \rightarrow \infty$ , then

$$RS^{rand} \rightarrow \frac{1}{6}.$$

In the worst case, all elements on the reverse diagonal are equal to one. We have two cases here, if  $M$  is an odd number, then

$$D_m = \max[0, 2(m - \frac{M}{2})].$$

If  $M$  is an even number, then

$$D_m = \max[0, 2(m - 1 - \frac{M}{2})].$$

Hence, in both cases (see Appendix 1)

$$RS^{worst} = \frac{1}{4}.$$

We define the stability ratio (SR) relative to the random case for large  $M$ :

$$SR = \frac{RS}{RS^{random}} = 6RS. \quad (3.20)$$

Then, the optimization procedure gives no gain compared to the random case when the ratio approaches one:

$$SR^{random} = 6 * \frac{1}{6} = 1.$$

If the ratio is greater than one, then the optimization procedure will rate the portfolios in reverse order. For the worst case,

$$SR^{worst} = 6 * \frac{1}{4} = 1.5.$$

The ratio for the best case is equal to

$$SR^{best} = 6 * 0 = 0.$$

The stability ratio also has a graphical interpretation (see Figure 3.20).



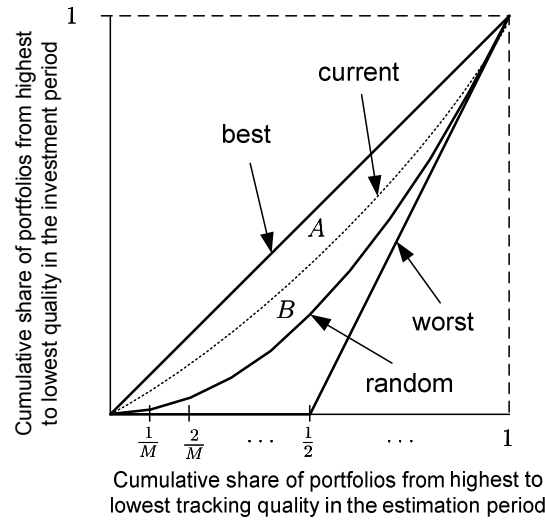


Figure 3.20: Stability Ratio<sup>199</sup>

The two axes show the cumulative shares of portfolios from the highest to the lowest quality in the estimation and investment periods respectively. We can display the best, worst and random cases of the dispersion table as curves. A point on a curve can be interpreted as “from  $n\%$  of portfolios of the highest quality in the estimation period, only  $m\%$  are of the highest quality in the investment period”. The current dispersion table is displayed as a dotted line. We compare it with the best case. The smaller is the area of the shape  $A$ , the more stable is the tracking quality of the portfolios produced by the sampling. We can define the SR as the ratio of the area of shape  $A$  to the area of shape  $B$  (the area between the random case and the best case).

$$SR = \frac{RS}{RS_{random}} = \frac{A}{B}.$$

This definition of the SR coincides with the ratio (3.20). The equivalence follows from equation (3.18), which defines the areas of  $A$  and  $B$ .

The next two indicators that can be used to define the relative stability are the Spearman and Kendall rank correlation coefficients. To calculate these coefficients, the portfolios are sorted from the highest to the lowest quality in both the estimation and investment periods. Then, ranks are assigned for both periods. The Spearman correlation coefficients are calculated as:

---

<sup>199</sup> For the worst case, the cumulative share of the portfolios from the highest to the lowest tracking quality in the investment period (axis  $y$ ) is equal to zero for the first 50% of portfolios in the estimation period (axis  $x$ ). Then, the cumulative share of portfolios goes up to 100%. This can be explained using an example. In the estimation period, six portfolios are ranked according to their quality: [1 2 3 4 5 6]. In the investment period, for the worst case, they are sorted as [6 5 4 3 2 1]. As a result, only the first 50% of portfolios in the estimation period have a better relative quality in that period in comparison to the investment period.

$$\rho = \frac{\sum_{p=1}^P RE_p RI_p - \frac{1}{P} \sum_{p=1}^K RE_p \sum_{i=1}^K RI_p}{\sqrt{\left( \sum_{p=1}^P RE_p^2 - \frac{1}{P} \left[ \sum_{p=1}^P RE_p \right]^2 \right) \left( \sum_{p=1}^P RI_p^2 - \frac{1}{P} \left[ \sum_{p=1}^P RI_p \right]^2 \right)}},$$

where  $RE_p$  is the rank of portfolio  $p$  in the estimation period and  $RI_p$  is the rank of the same portfolio  $p$  in the investment period.<sup>200</sup>

The Kendall rank correlation coefficient is computed as:<sup>201</sup>

$$\tau = \frac{2(P_c - P_d)}{P(P - 1)},$$

where  $P_c$  is the number of concordant portfolios,  $P_d$  is the number of discordant portfolios, and  $P$  is the total number of portfolios. Two portfolios are concordant if  $(RE_1 < RE_2$  and  $RI_1 < RI_2)$  or  $(RE_1 > RE_2$  and  $RI_1 > RI_2)$ . Two portfolios are discordant if  $(RE_1 < RE_2$  and  $RI_1 > RI_2)$  or  $(RE_1 > RE_2$  and  $RI_1 < RI_2)$ .

These two coefficients both fall in the interval  $[-1, 1]$ . If the coefficient is equal to zero, then there is no dependence between the tracking quality in the estimation period and that in the investment period. If the coefficient is larger than zero, then there is positive dependence. The closer the coefficient is to unity, the greater the relative stability of the tracking quality.

### 3.4.3 Resistance of the tracking portfolio estimation

Other than the absolute and relative stability of tracking quality, a good method of optimized sampling should have the following properties:<sup>202</sup>

- The tracking portfolio should not change dramatically when new information arises.
- If the estimation period contains any less reliable data, then the method of optimized sampling should not be overoptimized towards these data points.

Both of these properties depend on the resistance of the tracking portfolio estimation to variations in the inputs to the optimization problem.

If the sampling procedure constructs tracking portfolios that are very different as a result of

---

<sup>200</sup> See Rinne (1995, p. 93).

<sup>201</sup> See Rinne (1995, p. 96).

<sup>202</sup> Adapted from Schöttle/Werner (2006, p. 129) to apply to the case of index tracking.

small changes in data in the estimation period, then the procedure is very sensitive to the inputs, and one can say that the estimation of the tracking portfolio weights is uncertain. Using a sampling method that produces uncertain estimations of portfolio weights will lead to high transaction costs that should be avoided when tracking an index. Moreover, investors expect tracking portfolios to be characterized by a low rate of turnover.

If the estimation period contains outliers, then the estimation of the weights of the tracking portfolio could be strongly biased towards this group of assets, which again is not desirable when solving an index tracking problem. Since it is unlikely that these outlying assets will exhibit the same critical behaviour in the investment period, this could lead to a deterioration in tracking quality.

### 3.5 Summary

Sampling methods can be implemented since assets can contain duplicated information about the index and some assets may have weak informational content. This review of different approaches to sampling has shown that some solve the asset selection and weighting problems separately while others use a mixed approach.

The tracking quality in the investment period can be evaluated by traditional measures, such as MAD, MSE and TEV. However, these measures have weaknesses when used for this purpose. Therefore, we introduced the ex post tracking quality measure, which is the area between the normalized curves of the evolution of the values of the index and the tracking portfolio.

In this chapter, we selected certain methods of optimized sampling and further analysed their empirical properties. The typical form is based on the minimization of a traditional measure of tracking quality. However, it is possible to use more robust variants, namely the minimax and CVAR approaches.

The question of how to model the tracking portfolio structure so as to be consistent with a buy-and-hold strategy was also raised. Different variants of sampling can be implemented to deal with this issue.

To evaluate the different methods of optimized sampling, the concepts of absolute and relative stability of tracking quality were introduced. Moreover, the importance of estimating resistant tracking portfolios was discussed.

## **4 Empirical analysis**

In this chapter, first, we verify which approximations of portfolio returns are appropriate for index tracking, second, we empirically analyse the absolute and relative stability of the tracking quality of different methods of optimized sampling and, third, we test the resistance of the tracking portfolio estimation to slight changes of the data in the estimation period.

## 4.1 Portfolio structure: weights versus numbers

### 4.1.1 Research method

We empirically analyse six possible ways of modelling the returns of the tracking portfolio, which were defined in Section 3.3.4, Table 3.3. For this analysis, we use the S&P1500 price index, which has 1,500 underlying US stocks. To avoid the possibility of our conclusions being caused by any one specific capital market phase, we select five time periods between 2004 and 2010. Each time period covers a different capital market phase (see Figure 4.1).<sup>203</sup>

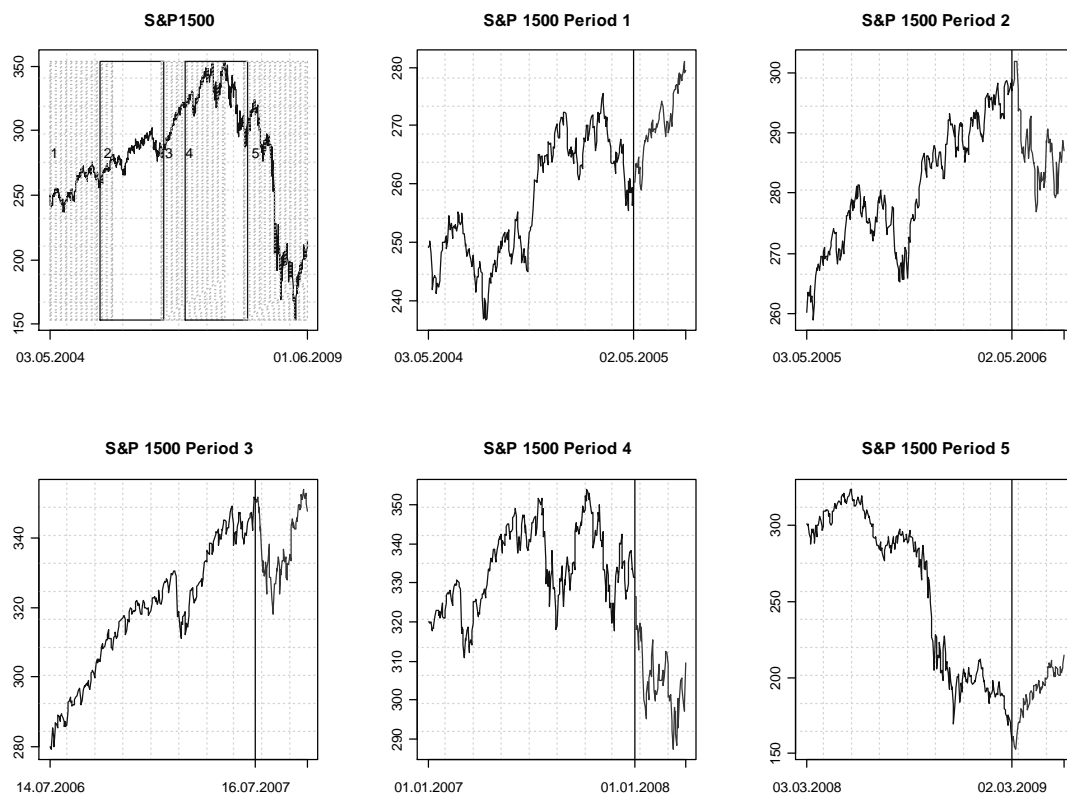


Figure 4.1: S&P1500

We divide each selected time period into two parts: one estimation part with a length of 12 months and one investment part with a length of 3 months. The data from the estimation part is used to estimate the tracking portfolios and the data from the investment part is used to evaluate these portfolios. The values of the tracking portfolios in the investment part are computed using fixed numbers of assets, which are determined using the estimation part. We use daily asset prices and daily index levels for our calculations. Moreover, we select portfolio sizes from 30 to 240, in increments of 30, resulting in sizes of 30, 60, 90, ..., 240.

---

<sup>203</sup> This research method is based on Roßbach/Karlow (2011).

For every time period and every portfolio size, tracking portfolios are optimized using the MSE as the objective. The MSE is computed using six different methods of computing the portfolio returns, taken from Table 3.3:

- MSE with simple approximation of portfolio returns (SA) (fixed weights);
- MSE with corrected portfolio returns at the beginning of the estimation period (CB) (fixed numbers);
- MSE with corrected portfolio returns at the end of the estimation period (CE) (fixed numbers);
- MSE using compounded portfolio and index returns (CR) (fixed numbers);
- MSE with corrected portfolio returns, using the assumption that the portfolio value is equal to the index value at the previous point in time (CPI1) (fixed numbers);
- MSE with corrected portfolio returns, using the assumption that the portfolio value is equal to the index value at the previous point in time, only in the denominator (CPI2) (fixed numbers).

Hence, there are six different optimization problems used to choose the amount of each asset that is included in each portfolio, one for each approximation method. First, however, we need to select which combination of assets to include in each portfolio.

To investigate how the approximation of portfolio returns influences the tracking quality, we would need to estimate all the portfolios (i.e. all possible combinations of assets), of each given fixed size, for each method of computing the portfolio returns, and then determine the central tendency in the measure of tracking quality in the investment period. However, the estimation of all possible portfolios is an impractical task since the number of possible asset combinations that can be drawn from 1,500 assets is huge (see Table 4.1).

Portfolio size	Combinations
30	5.39881E+62
60	1.3363E+108
90	3.1113E+146
120	1.5175E+180
150	2.0329E+210
180	3.3969E+237
210	1.9164E+262
240	7.3759E+284

Table 4.1: The Number of Possible Combinations of Assets for the Tracking Portfolios

As a consequence, to accomplish our task, we have to make the assumption that the measure of tracking quality is a random variable, and follows some probability distribution. To identify the nature of this variable, we select 5,000 combinations of assets for each time period, each portfolio size. For each portfolio, the stocks are randomly selected from the 1,500 stocks of the S&P1500 index. Then, using those selected stocks, we perform the optimization, based on our approximation method and estimation period, to choose the weight

of each stock. Hence, using this selection method, we neutralize the effect of stock selection and increase the representativeness of the sample of tracking portfolios. In total,  $(5,000 \text{ (combinations)} * 8 \text{ (sizes)} * 6 \text{ (approximation methods)} * 5 \text{ (periods)}) = 1.2 \text{ million tracking portfolios}$  are generated.<sup>204</sup> These portfolios are used to analyse the influence of the approximation method on the tracking quality.

We then test whether the measure of tracking quality for one approximation method is larger than the measure of tracking quality for another method. We use a non-parametric paired Wilcoxon test<sup>205</sup> with significance level  $\alpha = 0.05$  to do this. We select a paired test because the approximation methods are applied to the same combination of assets; therefore it is possible to determine which method produces the tracking portfolio with the best tracking quality for a given subset of stocks. For every period, and every portfolio size, we calculate the table shown in Figure 4.2.

Is H0 rejected with significance level 0.05?

	SA	CB	CE	CR	CPI1	CPI2
SA	-	Y or N	Y or N	Y or N	Y or N	Y or N
CB	Y or N	-	Y or N	Y or N	Y or N	Y or N
CE	Y or N	Y or N	-	Y or N	Y or N	Y or N
CR	Y or N	Y or N	Y or N	-	Y or N	Y or N
CPI1	Y or N	Y or N	Y or N	Y or N	-	Y or N
CPI2	Y or N	Y or N	Y or N	Y or N	Y or N	-

Figure 4.2: Comparison of Methods for the Approximation of Portfolio Returns using Wilcoxon Test

The null hypothesis is formulated as follows: the probability that the tracking quality measure for the method in the row is larger than or equal to the measure for the method in the column is larger than 0.5. If the null hypothesis is rejected then we can say that the measure of tracking quality of the method in the row has a smaller value than that of the method in the column. Hence, the best method should have all “Y” values in its row.

## 4.1.2 Empirical results

### Results for the estimation period

We start with the analysis of the results for the estimation period. We report in Table 4.2 the medians and inter-quartile ranges (IQRs) of the tracking qualities, which are expressed as

<sup>204</sup> We use the MOSEK solver to formulate and solve the quadratic optimization problems.

<sup>205</sup> See Rinne (2003, p. 552). We use Octave to perform this test. The Wilcoxon test is based on the construction of the rank ordering of the absolute differences of pairs of values.



percentage RMSE (see equation (3.2)), for the tracking portfolios created using each of the six objectives.

		SA		CB		CE		CR		CPI1		CPI2	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	0.2761	0.0417	0.2769	0.0440	0.2857	0.0457	0.3530	0.0654	0.3529	0.0653	0.2752	0.0428
	60	0.2000	0.0285	0.1999	0.0302	0.2079	0.0302	0.2555	0.0399	0.2553	0.0400	0.1979	0.0291
	90	0.1633	0.0222	0.1621	0.0232	0.1699	0.0240	0.2081	0.0309	0.2080	0.0308	0.1605	0.0222
	120	0.1401	0.0194	0.1386	0.0207	0.1460	0.0205	0.1775	0.0278	0.1774	0.0277	0.1367	0.0200
	150	0.1244	0.0176	0.1221	0.0184	0.1301	0.0184	0.1566	0.0241	0.1565	0.0240	0.1203	0.0180
	180	0.1115	0.0163	0.1088	0.0169	0.1172	0.0168	0.1393	0.0225	0.1392	0.0225	0.1069	0.0166
	210	0.1011	0.0145	0.0980	0.0156	0.1068	0.0153	0.1254	0.0198	0.1252	0.0198	0.0960	0.0151
	240	0.0923	0.0139	0.0886	0.0148	0.0983	0.0143	0.1131	0.0193	0.1129	0.0192	0.0866	0.0144
Period 2	30	0.2691	0.0428	0.2664	0.0451	0.2896	0.0508	0.3415	0.0686	0.3416	0.0682	0.2652	0.0446
	60	0.1943	0.0283	0.1894	0.0297	0.2109	0.0329	0.2441	0.0412	0.2441	0.0410	0.1883	0.0293
	90	0.1590	0.0218	0.1531	0.0226	0.1734	0.0251	0.1982	0.0315	0.1981	0.0316	0.1520	0.0223
	120	0.1366	0.0181	0.1291	0.0186	0.1495	0.0211	0.1677	0.0250	0.1677	0.0250	0.1280	0.0184
	150	0.1211	0.0163	0.1128	0.0170	0.1336	0.0191	0.1468	0.0236	0.1468	0.0235	0.1117	0.0170
	180	0.1089	0.0150	0.0999	0.0154	0.1211	0.0184	0.1297	0.0211	0.1298	0.0212	0.0986	0.0153
	210	0.0994	0.0137	0.0895	0.0140	0.1114	0.0171	0.1157	0.0198	0.1157	0.0198	0.0881	0.0140
	240	0.0914	0.0128	0.0806	0.0129	0.1029	0.0158	0.1039	0.0179	0.1039	0.0179	0.0790	0.0130
Period 3	30	0.2467	0.0399	0.2464	0.0411	0.2608	0.0469	0.3254	0.0629	0.3247	0.0620	0.2455	0.0405
	60	0.1755	0.0248	0.1746	0.0252	0.1854	0.0281	0.2294	0.0379	0.2287	0.0377	0.1737	0.0251
	90	0.1419	0.0192	0.1407	0.0194	0.1495	0.0214	0.1834	0.0288	0.1829	0.0287	0.1397	0.0194
	120	0.1211	0.0160	0.1195	0.0162	0.1277	0.0179	0.1556	0.0236	0.1552	0.0235	0.1185	0.0161
	150	0.1058	0.0145	0.1039	0.0148	0.1120	0.0160	0.1350	0.0210	0.1346	0.0209	0.1028	0.0148
	180	0.0945	0.0131	0.0922	0.0133	0.1004	0.0142	0.1194	0.0191	0.1190	0.0190	0.0910	0.0132
	210	0.0851	0.0116	0.0829	0.0117	0.0904	0.0130	0.1068	0.0167	0.1064	0.0168	0.0816	0.0118
	240	0.0776	0.0116	0.0751	0.0118	0.0826	0.0126	0.0963	0.0165	0.0961	0.0165	0.0737	0.0119
Period 4	30	0.2988	0.0440	0.2935	0.0449	0.3206	0.0514	0.3832	0.0695	0.3837	0.0693	0.2920	0.0443
	60	0.2133	0.0276	0.2062	0.0277	0.2317	0.0325	0.2693	0.0417	0.2693	0.0418	0.2048	0.0273
	90	0.1744	0.0227	0.1655	0.0229	0.1924	0.0267	0.2151	0.0325	0.2151	0.0323	0.1638	0.0225
	120	0.1497	0.0196	0.1397	0.0197	0.1671	0.0227	0.1810	0.0284	0.1810	0.0283	0.1379	0.0196
	150	0.1332	0.0174	0.1217	0.0177	0.1510	0.0203	0.1564	0.0245	0.1563	0.0244	0.1196	0.0177
	180	0.1202	0.0154	0.1078	0.0158	0.1384	0.0188	0.1379	0.0217	0.1378	0.0217	0.1055	0.0157
	210	0.1098	0.0142	0.0964	0.0141	0.1286	0.0174	0.1225	0.0199	0.1225	0.0198	0.0938	0.0142
	240	0.1016	0.0139	0.0873	0.0134	0.1210	0.0171	0.1102	0.0189	0.1102	0.0188	0.0843	0.0138
Period 5	30	0.5973	0.0912	0.5707	0.0989	0.6467	0.1057	0.7321	0.1404	0.7272	0.1413	0.5644	0.0990
	60	0.4553	0.0624	0.4097	0.0635	0.5148	0.0801	0.5310	0.0917	0.5244	0.0908	0.4011	0.0635
	90	0.3928	0.0537	0.3341	0.0515	0.4650	0.0794	0.4310	0.0767	0.4231	0.0743	0.3227	0.0519
	120	0.3556	0.0496	0.2878	0.0443	0.4402	0.0781	0.3664	0.0635	0.3592	0.0622	0.2745	0.0448
	150	0.3306	0.0485	0.2542	0.0381	0.4242	0.0728	0.3191	0.0563	0.3113	0.0545	0.2385	0.0388
	180	0.3123	0.0448	0.2303	0.0360	0.4173	0.0738	0.2829	0.0540	0.2755	0.0520	0.2122	0.0361
	210	0.2988	0.0448	0.2113	0.0325	0.4102	0.0701	0.2532	0.0478	0.2460	0.0464	0.1902	0.0337
	240	0.2891	0.0424	0.1956	0.0306	0.4099	0.0664	0.2287	0.0451	0.2211	0.0437	0.1722	0.0314

Table 4.2: Portfolio Structure: Results for the Estimation Period (RMSE)

The medians and IQRs are reported for the 5,000 random tracking portfolios of each size, for each estimation period and for every method of computing portfolio returns used in the objective in the optimization problem. For example, 0.2761, in the first row and column of the table, is the median tracking quality, expressed as RMSE, of the 5,000 tracking portfolios of size 30 using estimation period 1, generated by minimizing MSE using a simple approximation of portfolio returns (SA).

A comparison of the median tracking qualities for the six methods of approximation indicates that, in the estimation period, the CPI2 method produces portfolios that track the index much better than all the other methods (demonstrated by lower values of RMSE). This holds for all sizes of tracking portfolio and all periods.

The results for the CR and CPI1 methods are very similar to each other. This can be explained by the fact that both of these methods estimate the numbers of assets to be placed in the tracking portfolios at the beginning of the estimation period. However, this similarity is reduced in the turbulent period 5.

We also compute percentage ETQ measures (see equation (3.3)) for the tracking portfolios in the estimation period, to illustrate the closeness of the index values to the values of the tracking portfolios (see Table 4.3).

		SA		CB		CE		CR		CPI1		CPI2	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	3.2124	2.5891	1.6243	1.2940	4.4522	2.8433	0.3947	0.1102	0.3951	0.1105	1.6573	1.3260
	60	2.4873	1.8498	1.1149	0.8525	3.5630	2.0193	0.2259	0.0515	0.2260	0.0516	1.1289	0.8884
	90	2.1176	1.5862	0.9428	0.7343	3.0752	1.6753	0.1619	0.0337	0.1618	0.0338	0.9316	0.7014
	120	1.8775	1.3375	0.8441	0.6687	2.7401	1.4194	0.1267	0.0266	0.1267	0.0266	0.7965	0.5937
	150	1.7237	1.1690	0.7885	0.6492	2.4943	1.2528	0.1052	0.0215	0.1052	0.0216	0.7189	0.5328
	180	1.5386	1.0240	0.7767	0.6623	2.2848	1.0779	0.0892	0.0180	0.0892	0.0179	0.6480	0.4895
	210	1.4347	0.9337	0.7599	0.6658	2.1274	1.0316	0.0771	0.0151	0.0771	0.0150	0.5960	0.4232
Period 2	240	1.3320	0.8897	0.7791	0.7226	1.9581	0.9881	0.0676	0.0137	0.0677	0.0137	0.5521	0.4139
	30	2.7970	2.7168	1.7658	1.3130	4.9742	3.7607	0.4003	0.1175	0.4005	0.1175	1.7026	1.2642
	60	2.3247	2.0620	1.3070	1.0427	4.2538	2.5348	0.2270	0.0522	0.2272	0.0523	1.1987	0.8956
	90	2.1488	1.6753	1.1018	0.9017	3.7738	2.0145	0.1635	0.0361	0.1636	0.0363	0.9557	0.7023
	120	2.0437	1.4225	0.9945	0.8116	3.4718	1.6622	0.1274	0.0268	0.1275	0.0269	0.8111	0.5724
	150	1.9544	1.2540	0.9103	0.7634	3.2559	1.4693	0.1050	0.0224	0.1050	0.0224	0.7056	0.5124
	180	1.8590	1.1349	0.8692	0.7293	3.0240	1.3913	0.0886	0.0181	0.0887	0.0181	0.6323	0.4645
Period 3	210	1.7899	1.0188	0.8488	0.7047	2.8740	1.2130	0.0760	0.0160	0.0761	0.0160	0.5734	0.4076
	240	1.6976	0.9193	0.8667	0.6794	2.7045	1.1189	0.0661	0.0140	0.0662	0.0140	0.5161	0.3725
	30	1.9103	1.5606	2.3743	2.0267	2.7512	2.5302	0.3874	0.1030	0.3880	0.1038	2.1562	1.7793
	60	1.3124	1.0068	1.9883	1.6002	2.1961	1.7226	0.2187	0.0495	0.2190	0.0496	1.5665	1.3405
	90	1.0753	0.8231	1.7106	1.2830	1.9005	1.4397	0.1564	0.0334	0.1566	0.0335	1.2197	1.0071
	120	0.9410	0.7353	1.5834	1.1047	1.7333	1.2284	0.1231	0.0252	0.1232	0.0252	1.0146	0.8001
	150	0.8601	0.6975	1.4694	0.9933	1.6072	1.0806	0.1010	0.0206	0.1010	0.0206	0.8479	0.6487
Period 4	180	0.7854	0.6565	1.4062	0.8647	1.4864	1.0341	0.0856	0.0175	0.0858	0.0176	0.7280	0.5303
	210	0.7387	0.6127	1.3668	0.8108	1.3759	0.8718	0.0739	0.0145	0.0740	0.0146	0.6439	0.4764
	240	0.7113	0.6142	1.3399	0.7201	1.3034	0.8522	0.0647	0.0139	0.0649	0.0140	0.5681	0.4002
	30	2.3180	2.0094	2.2272	1.6627	3.9788	3.3494	0.4139	0.1085	0.4137	0.1079	2.0563	1.5030
	60	2.2338	1.8648	1.5085	1.0612	3.9513	2.4023	0.2323	0.0493	0.2323	0.0494	1.3273	0.8869
	90	2.2527	1.5950	1.2180	0.8658	3.9039	1.9090	0.1646	0.0327	0.1646	0.0326	1.0291	0.6708
	120	2.2057	1.4039	1.0539	0.7485	3.7243	1.6209	0.1288	0.0253	0.1289	0.0252	0.8562	0.5635
Period 5	150	2.2025	1.2455	0.9543	0.6685	3.6399	1.5144	0.1052	0.0199	0.1052	0.0200	0.7430	0.4762
	180	2.1463	1.1098	0.8940	0.6245	3.5183	1.2838	0.0893	0.0170	0.0893	0.0170	0.6594	0.4270
	210	2.0763	0.9993	0.8632	0.5752	3.4097	1.2236	0.0763	0.0147	0.0764	0.0147	0.5925	0.4140
	240	1.9965	0.9207	0.8592	0.5788	3.2717	1.1613	0.0669	0.0134	0.0670	0.0134	0.5420	0.3687
	30	5.4668	3.0461	2.4632	1.4530	6.3602	3.2416	0.5882	0.1640	0.5963	0.1675	2.3556	1.4567
	60	4.9146	2.2445	1.8980	0.9546	5.6583	2.3752	0.3355	0.0731	0.3407	0.0739	1.6895	0.9837
	90	4.7173	1.8753	1.6771	0.7864	5.4094	2.0280	0.2443	0.0499	0.2482	0.0503	1.3904	0.8144
Period 5	120	4.4537	1.6425	1.5644	0.6605	5.1441	1.7698	0.1943	0.0390	0.1975	0.0394	1.2105	0.7185
	150	4.3318	1.5044	1.4714	0.5950	5.0459	1.6495	0.1624	0.0326	0.1651	0.0331	1.0781	0.6606
	180	4.1899	1.3829	1.4175	0.5452	4.9337	1.4806	0.1389	0.0292	0.1413	0.0296	0.9697	0.6274
	210	4.0217	1.2705	1.3985	0.4971	4.7914	1.3662	0.1213	0.0255	0.1235	0.0261	0.8896	0.5913
	240	3.9087	1.1423	1.3685	0.4557	4.6958	1.3296	0.1072	0.0233	0.1094	0.0239	0.8296	0.5709

Table 4.3: Portfolio Structure: Results for the Estimation Period (ETQ)

The CR and CPI1 methods both generate tracking portfolios that approximate the index values very closely in the estimation period. The quality of tracking of these two methods is approximately ten times that of the other methods. The high tracking quality of the CR and CPI1 methods is due to the fact that these two methods use tracking quality measures that are based on compounded returns and therefore contain information about the sequence of changes in the asset prices. The other methods do not contain such information.

Comparing tracking quality expressed as ETQ and MSE we can see that CR and CPI1 methods approximate index values better than other methods, which is shown by ETQ measure, but these two methods are not the best one in respect to MSE measure. Hence, the both measures are not equivalent.

### Results for the investment period

Next, we analyse the tracking quality in the investment period. We report the medians and IQRs of the tracking qualities using the ETQ measure (see Table 4.4).

		SA		CB		CE		CR		CPI1		CPI2	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	1.3782	1.3426	1.3646	1.3476	1.4223	1.3805	1.6544	1.6022	1.5997	1.5132	1.3406	1.2793
	60	1.0944	1.0227	1.0700	0.9989	1.1302	1.0522	1.3154	1.2958	1.3229	1.2552	1.0621	0.9797
	90	0.8929	0.8397	0.8953	0.8341	0.9362	0.8615	1.1113	1.1022	1.1207	1.0657	0.8811	0.8168
	120	0.8071	0.7372	0.7929	0.7367	0.8315	0.7492	1.0030	0.9296	1.0359	0.9351	0.8014	0.7244
	150	0.7563	0.6824	0.7458	0.6747	0.7650	0.6988	0.9208	0.8709	0.9660	0.8752	0.7440	0.6773
	180	0.7001	0.6235	0.6925	0.6331	0.7053	0.6253	0.8704	0.7954	0.9042	0.8033	0.6995	0.6176
	210	0.6739	0.6076	0.6645	0.6084	0.6792	0.6067	0.8343	0.7639	0.8714	0.7815	0.6815	0.6072
	240	0.6341	0.5738	0.6217	0.5619	0.6380	0.5634	0.7559	0.6966	0.8022	0.7163	0.6443	0.5806
Period 2	30	1.0541	0.9420	1.2067	1.0948	0.9976	0.8583	1.3041	1.1356	1.1653	1.0133	1.0227	0.9166
	60	0.8174	0.7684	0.9993	0.9055	0.7464	0.6447	0.9985	0.8763	0.8671	0.7305	0.7647	0.6722
	90	0.7525	0.7000	0.9405	0.8291	0.6652	0.5836	0.8886	0.7790	0.7353	0.6507	0.6749	0.5955
	120	0.6792	0.5981	0.9004	0.7281	0.5768	0.4796	0.8328	0.7012	0.6519	0.5355	0.5837	0.4840
	150	0.6441	0.5702	0.8541	0.7292	0.5550	0.4611	0.8087	0.6901	0.6174	0.5078	0.5540	0.4670
	180	0.6181	0.5522	0.8291	0.6692	0.5262	0.4449	0.7847	0.6605	0.5729	0.4688	0.5206	0.4294
	210	0.5961	0.5284	0.8128	0.6361	0.5002	0.4367	0.7734	0.6288	0.5453	0.4494	0.4831	0.4099
	240	0.5806	0.5358	0.7934	0.6506	0.4893	0.4417	0.7640	0.6236	0.5075	0.4282	0.4683	0.4013
Period 3	30	1.2883	1.1008	1.3118	1.0985	1.2845	1.1659	1.5118	1.3383	1.5068	1.3827	1.2740	1.0990
	60	1.0237	0.8557	1.0447	0.8404	0.9989	0.9144	1.1794	1.0093	1.1595	1.0605	1.0016	0.8826
	90	0.8839	0.7240	0.8999	0.7092	0.8531	0.7684	0.9815	0.8398	0.9736	0.9179	0.8508	0.7537
	120	0.7902	0.6563	0.8032	0.6369	0.7470	0.6877	0.8790	0.7289	0.8546	0.7730	0.7526	0.6791
	150	0.7315	0.5881	0.7448	0.5849	0.6934	0.6177	0.7845	0.6502	0.7461	0.6900	0.6916	0.6119
	180	0.6891	0.5603	0.7028	0.5453	0.6504	0.5746	0.7466	0.5891	0.7127	0.6426	0.6567	0.5880
	210	0.6481	0.5271	0.6627	0.5109	0.6088	0.5348	0.6851	0.5370	0.6545	0.5741	0.6119	0.5457
	240	0.6280	0.5102	0.6437	0.5048	0.5867	0.5201	0.6608	0.5087	0.6129	0.5539	0.5921	0.5434
Period 4	30	1.7506	1.4804	2.0742	1.7641	1.5814	1.3334	2.1336	1.8293	1.8822	1.5834	1.6601	1.4155
	60	1.4463	1.2425	1.8160	1.4882	1.2801	1.1107	1.8217	1.5307	1.4332	1.2001	1.3241	1.1518
	90	1.3349	1.1042	1.7488	1.3690	1.1663	0.9988	1.7426	1.4241	1.2858	1.1139	1.1866	1.0255
	120	1.2346	1.0858	1.6700	1.3489	1.0672	0.9590	1.6418	1.3462	1.1473	0.9912	1.0706	0.9522
	150	1.1747	1.0228	1.6339	1.2435	1.0267	0.9189	1.6065	1.3160	1.0690	0.9261	1.0062	0.9082
	180	1.1185	0.9806	1.5769	1.1951	0.9821	0.8736	1.5449	1.2234	0.9907	0.8681	0.9538	0.8373
	210	1.1007	0.9381	1.5828	1.1442	0.9647	0.8369	1.5187	1.1900	0.9539	0.8189	0.9174	0.8067
	240	1.0765	0.8999	1.5759	1.1560	0.9401	0.8232	1.4930	1.1450	0.8867	0.7838	0.8964	0.7905
Period 5	30	3.5333	3.2762	9.3875	7.1526	2.3045	1.6691	8.3556	6.2522	2.7219	2.0645	2.7628	2.5480
	60	3.2455	2.7013	9.7465	5.7560	1.7816	1.2467	8.2066	5.1100	1.9728	1.3780	2.0546	1.6475
	90	3.1839	2.4244	10.0252	4.9701	1.5786	1.0542	8.2935	4.6605	1.6920	1.1758	1.7008	1.3060
	120	3.2345	2.3416	10.2796	4.5495	1.4622	0.9723	8.5899	4.4619	1.5170	1.0169	1.5172	1.1424
	150	3.1820	2.2009	10.3894	4.3119	1.3922	0.8929	8.7466	4.2847	1.4159	0.9358	1.3795	1.0208
	180	3.2465	2.0659	10.4970	4.0254	1.3171	0.8644	8.8387	4.0016	1.3027	0.8415	1.2797	0.8975
	210	3.3195	2.0436	10.6699	3.9703	1.2805	0.8054	9.0102	4.0095	1.2241	0.7859	1.2104	0.8504
	240	3.4309	1.9098	10.7663	3.7232	1.2354	0.7925	9.2573	3.9068	1.1374	0.7361	1.1450	0.7758

Table 4.4: Portfolio Structure: Results for the Investment Period (ETQ)

The results indicate that the CR and CPI1 procedures are not similar to one another, as they were in the estimation period. The reason for this is that the weights of the tracking portfolio under the CR procedure are transformed into numbers using the stock prices and index value at the beginning of the investment period. The numbers for the tracking portfolio under the

CPI1 procedure, however, remain unchanged. As a result, the tracking qualities for the CR procedure are worse than those for the CPI1 procedure in most cases.

To analyse the influence of the approximation method on the tracking quality, we select the best methods, according to the median ETQs, from Table 4.4. The results are shown in Table 4.5.

	Period 1	Period 2	Period 3	Period 4	Period 5
<b>30</b>	CPI2	CE	CPI2	CE	CE
<b>60</b>	CPI2	CE	CE	CE	CE
<b>90</b>	CPI2	CE	CPI2	CE	CE
<b>120</b>	CB	CE	CE	CE	CE
<b>150</b>	CPI2	CPI2	CPI2	CPI2	CE
<b>180</b>	CB	CPI2	CE	CPI2	CPI2
<b>210</b>	CB	CPI2	CE	CPI2	CPI2
<b>240</b>	CB	CPI2	CE	CPI1	CPI1

Table 4.5: Best Methods by Median (ETQ)

Two of the methods produce tracking portfolios with better quality, on average, than the other methods in most of the cases. These methods are CE and CPI2, which model the portfolio structure using fixed numbers. The procedure that is most used in literature, SA, which models the portfolio structure using fixed weights, gives worse results than these methods. Hence, the usage of fixed numbers of assets can improve the tracking quality in the investment period in comparison to fixed weights. We test this hypothesis again later in this section, using the Wilcoxon test.

We illustrate the distribution of the tracking qualities, measured by ETQ (see equation (3.3)), for the portfolios of size 240, for the investment part of period 1, by computing a histogram for each approximation method (see Figure 4.3).

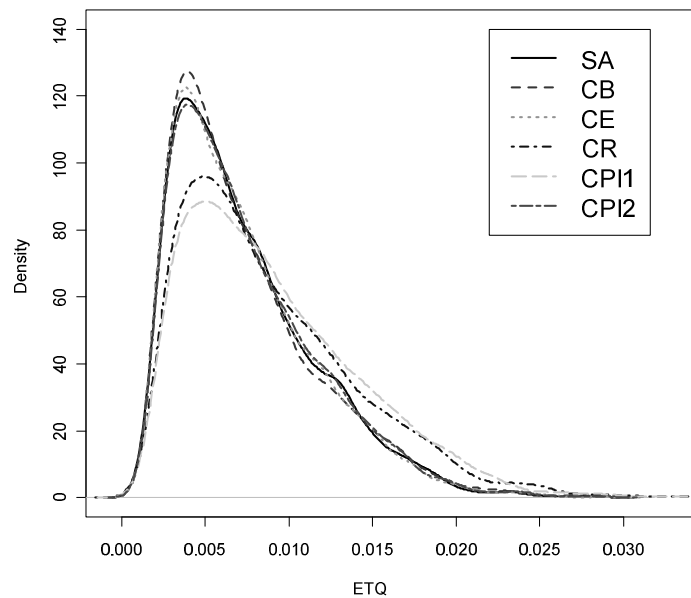


Figure 4.3: Histograms for each of the Approximation Procedures (Period 1, Size 240, ETQ)

The distribution of the ETQ measure is not symmetric; it has a positive skew. Therefore, the use of the median as the measure of the central tendency is more appropriate than the mean. The best method in this particular case is CB. The histogram for this method has the highest peak, and is shifted most to the left, indicating better tracking quality.

We also compute the RMSE tracking quality measure in the investment period (see Table 4.6).

		SA		CB		CE		CR		CPI1		CPI2	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	0.2975	0.0621	0.3028	0.0659	0.3011	0.0658	0.3655	0.0871	0.3592	0.0838	0.2972	0.0641
	60	0.2340	0.0435	0.2388	0.0446	0.2360	0.0452	0.2808	0.0583	0.2778	0.0574	0.2339	0.0444
	90	0.2043	0.0370	0.2084	0.0386	0.2057	0.0385	0.2406	0.0475	0.2383	0.0466	0.2046	0.0378
	120	0.1864	0.0333	0.1900	0.0343	0.1867	0.0336	0.2168	0.0418	0.2148	0.0415	0.1865	0.0337
	150	0.1741	0.0300	0.1781	0.0315	0.1741	0.0314	0.1986	0.0379	0.1965	0.0368	0.1740	0.0311
	180	0.1636	0.0291	0.1677	0.0296	0.1633	0.0291	0.1855	0.0351	0.1839	0.0351	0.1641	0.0294
	210	0.1570	0.0279	0.1604	0.0286	0.1561	0.0286	0.1749	0.0324	0.1736	0.0325	0.1569	0.0284
	240	0.1505	0.0258	0.1543	0.0268	0.1490	0.0259	0.1671	0.0299	0.1652	0.0310	0.1502	0.0259
Period 2	30	0.3066	0.0576	0.3204	0.0585	0.3175	0.0703	0.3783	0.0857	0.3745	0.0860	0.3078	0.0596
	60	0.2393	0.0418	0.2543	0.0454	0.2440	0.0478	0.2922	0.0569	0.2844	0.0577	0.2389	0.0427
	90	0.2096	0.0336	0.2250	0.0370	0.2109	0.0377	0.2521	0.0464	0.2425	0.0455	0.2083	0.0353
	120	0.1917	0.0300	0.2074	0.0336	0.1910	0.0322	0.2294	0.0408	0.2180	0.0404	0.1900	0.0309
	150	0.1784	0.0287	0.1950	0.0315	0.1769	0.0295	0.2129	0.0365	0.1998	0.0357	0.1767	0.0294
	180	0.1694	0.0272	0.1863	0.0297	0.1671	0.0273	0.2013	0.0341	0.1868	0.0338	0.1671	0.0272
	210	0.1611	0.0263	0.1790	0.0290	0.1582	0.0262	0.1909	0.0336	0.1760	0.0313	0.1584	0.0261
	240	0.1552	0.0247	0.1731	0.0283	0.1517	0.0250	0.1822	0.0314	0.1666	0.0296	0.1520	0.0254
Period 3	30	0.4169	0.0987	0.4297	0.1058	0.4092	0.0980	0.4952	0.1320	0.4947	0.1322	0.4109	0.0960
	60	0.3159	0.0651	0.3261	0.0699	0.3115	0.0643	0.3651	0.0850	0.3648	0.0827	0.3111	0.0632
	90	0.2713	0.0528	0.2803	0.0554	0.2648	0.0539	0.3067	0.0658	0.3051	0.0656	0.2652	0.0532
	120	0.2441	0.0465	0.2525	0.0483	0.2392	0.0469	0.2721	0.0561	0.2707	0.0557	0.2399	0.0460
	150	0.2259	0.0418	0.2337	0.0432	0.2207	0.0430	0.2470	0.0486	0.2456	0.0499	0.2214	0.0420
	180	0.2122	0.0385	0.2201	0.0397	0.2077	0.0388	0.2297	0.0440	0.2273	0.0445	0.2087	0.0384
	210	0.2012	0.0364	0.2085	0.0378	0.1960	0.0370	0.2156	0.0407	0.2127	0.0422	0.1974	0.0383
	240	0.1925	0.0348	0.1999	0.0356	0.1870	0.0353	0.2052	0.0379	0.2004	0.0384	0.1879	0.0356
Period 4	30	0.4904	0.0957	0.5407	0.1174	0.4830	0.0928	0.6024	0.1378	0.5725	0.1267	0.4851	0.0947
	60	0.3872	0.0682	0.4361	0.0874	0.3753	0.0673	0.4704	0.0916	0.4341	0.0864	0.3779	0.0669
	90	0.3404	0.0610	0.3923	0.0755	0.3282	0.0582	0.4159	0.0797	0.3731	0.0722	0.3302	0.0594
	120	0.3128	0.0545	0.3672	0.0683	0.2995	0.0538	0.3797	0.0729	0.3355	0.0622	0.3001	0.0538
	150	0.2936	0.0516	0.3508	0.0642	0.2804	0.0495	0.3586	0.0665	0.3078	0.0546	0.2797	0.0492
	180	0.2789	0.0477	0.3381	0.0600	0.2662	0.0462	0.3410	0.0620	0.2860	0.0525	0.2646	0.0458
	210	0.2682	0.0455	0.3293	0.0568	0.2543	0.0448	0.3289	0.0602	0.2706	0.0483	0.2517	0.0445
	240	0.2583	0.0457	0.3225	0.0583	0.2440	0.0437	0.3181	0.0586	0.2566	0.0463	0.2406	0.0436
Period 5	30	0.7392	0.1865	1.1269	0.4358	0.6373	0.1309	1.1876	0.4089	0.7607	0.1741	0.6622	0.1580
	60	0.6240	0.1367	1.0576	0.3309	0.5087	0.0925	1.0263	0.3057	0.5795	0.1147	0.5217	0.1073
	90	0.5764	0.1236	1.0372	0.2872	0.4482	0.0818	0.9647	0.2753	0.4983	0.0967	0.4547	0.0927
	120	0.5485	0.1153	1.0320	0.2730	0.4090	0.0729	0.9481	0.2701	0.4486	0.0845	0.4118	0.0808
	150	0.5295	0.1088	1.0228	0.2535	0.3842	0.0691	0.9347	0.2596	0.4125	0.0778	0.3834	0.0743
	180	0.5111	0.1060	1.0150	0.2435	0.3626	0.0640	0.9289	0.2532	0.3841	0.0752	0.3595	0.0699
	210	0.5027	0.1043	1.0123	0.2425	0.3461	0.0602	0.9294	0.2421	0.3622	0.0657	0.3401	0.0657
	240	0.4952	0.1009	1.0211	0.2256	0.3321	0.0586	0.9270	0.2418	0.3447	0.0638	0.3241	0.0614

Table 4.6: Results for the Investment Period (RMSE)

The best methods according to the medians are reported in Table 4.7.

	Period 1	Period2	Period 3	Period 4	Period 5
30	CPI2	SA	CE	CE	CE
60	CPI2	CPI2	CPI2	CE	CE
90	SA	CPI2	CE	CE	CE
120	SA	CPI2	CE	CE	CE
150	CPI2	CPI2	CE	CPI2	CPI2
180	CE	CPI2	CE	CPI2	CPI2
210	CE	CE	CE	CPI2	CPI2
240	CE	CE	CE	CPI2	CPI2

Table 4.7: Best Methods by Median (RMSE)

Here, again, the CE and CPI2 methods are best in most cases. Hence, modelling the portfolio structure using numbers of assets also improves the tracking quality when it is measured using RMSE.

### Stability of the medians

As was shown above (see Table 4.1), the total number of possible portfolios of a given size is very large. We must then question the validity of using just 5,000 randomly-selected portfolios, sampled from the entire set of possible portfolios. Figure 4.4 illustrates that the median stabilizes as the number of sampled portfolios increases. To construct this figure, the medians are computed, first using the tracking quality (expressed in ETQ) for the first two portfolios, then for the first three portfolios and so on up to 5,000. The figure below shows the medians for the investment period, for portfolios of size 240 in period 5.

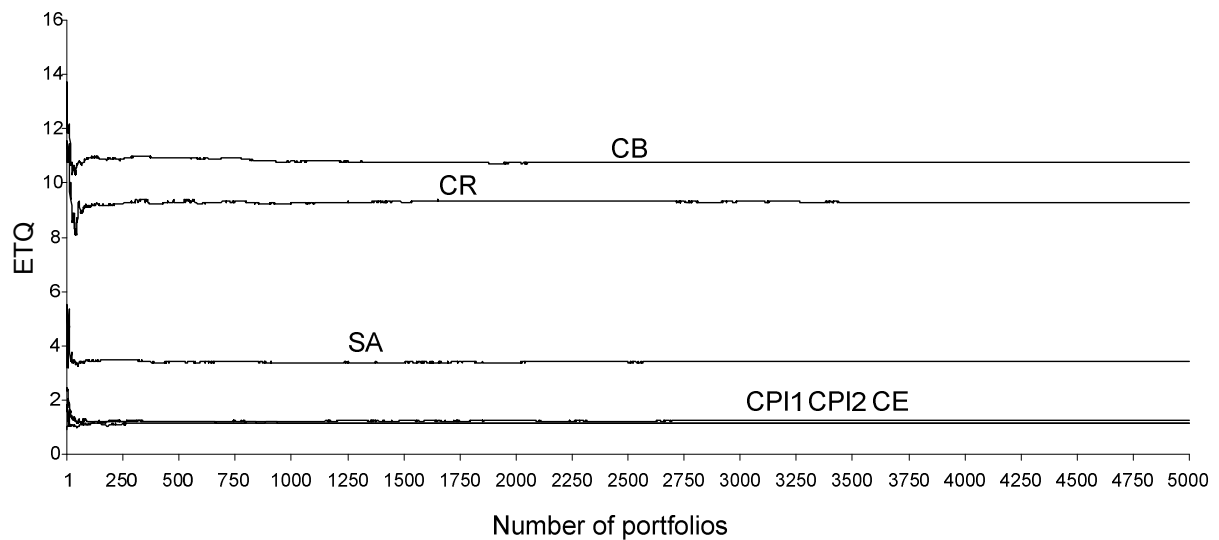


Figure 4.4: Stability of the Median (ETQ)

The estimated value of the median is relatively stable after 1,000 portfolios have been included. We can see that the methods CPI1, CPI2 and CE are very similar. This coincides with the results shown in Table 4.4, where CPI1 is the best method for period 5, portfolio size 240.

## Wilcoxon test in the investment period

To test the significance of our findings, we use a Wilcoxon test with significance level  $\alpha = 0.05$ . We construct the summary table as depicted earlier in Figure 4.2. For example, for period 1 and tracking portfolio size 240, using the ETQ measure, the resulting table is as shown below in Table 4.8.

	SA	CB	CE	CR	CPI1	CPI2
SA	-	N	N	Y	Y	Y
CB	Y	-	N	Y	Y	Y
CE	Y	N	-	Y	Y	Y
CR	N	N	N	-	Y	N
CPI1	N	N	N	N	-	N
CPI2	N	N	N	Y	Y	-

Table 4.8: Wilcoxon Test Summary Table for Period 1, Tracking Portfolio Size 240

Here, the best methods are CB and CE, because the rows for these methods contain more “Y” values than the rows for the other methods. We construct such tables for the other periods and tracking portfolio sizes. Then, we select the best methods for each portfolio size and period (see Table 4.9).

	Period 1	Period 2	Period 3	Period 4	Period 5
30	CPI2	CE	CPI2	CE	CE
60	CPI2	CE	CE, CPI2	CE	CE
90	CPI2	CE	CE, CPI2	CE	CE
120	CB, CPI2	CE	CE, CPI2	CE, CPI2	CE
150	CB, CPI2	CE, CPI2	CE	CPI2	CE
180	CB	CPI2	CE	CPI2	CE, CPI1, CPI2
210	CB	CPI2	CE	CPI2	CPI1, CPI2
240	CB, CE	CPI2	CE	CPI2	CPI1, CPI2

Table 4.9: Best Methods, with Significance Level 0.05 (ETQ)

Based on the results in Table 4.9 and Table 4.5, we can conclude that the CPI2 and CE methods of approximating returns are superior to the other methods in most cases.

We make the same calculation for the RMSE measure (see Table 4.10).

	Period 1	Period 2	Period 3	Period 4	Period 5
30	SA	SA	CE	CE	CE
60	SA, CPI2	SA, CPI2	CE	CE	CE
90	SA, CPI2	CPI2	CE	CE	CE
120	SA, CPI2	CPI2	CE	CE	CE
150	CE	CPI2	CE	CPI2	CE, CPI2
180	CE	CPI2, CE	CE	CPI2	CPI2
210	CE	CE	CE	CPI2	CPI2
240	CE	CE	CE	CPI2	CPI2

Table 4.10: Best Methods, with Significance Level 0.05 (RMSE)

Again, the CPI2 and CE methods of approximating returns produce better tracking quality, on average, than the other methods in most of the cases.

### 4.1.3 Conclusion

The first conclusion is that modelling the tracking portfolio structure using fixed numbers of assets is superior to using fixed weights. Among the methods tested, the best two are both based on fixed numbers. The first of these is CPI2, which is proposed by *Meade and Salkin*.<sup>206</sup> The second is CE, which is proposed by *Montfort et al.*<sup>207</sup> The significance of this finding is supported by the Wilcoxon test.

The second conclusion we can draw is that the closeness of the values of the tracking portfolio to the values of the index in the estimation period is not a guarantee of the same degree of closeness in the investment period. This is supported by the fact that the optimization of the portfolios using the CR and CPI1 methods produces tracking portfolios that are very close to the index in the estimation period in terms of the ETQ measure but this level of tracking quality does not remain in the investment period.

---

<sup>206</sup> See Meade/Salkin (1989, p. 878 (i)).

<sup>207</sup> See Montfort/Visser/van Draat (2008).



## 4.2 The stability of the minimization of traditional measures

### 4.2.1 Research method

It is necessary to analyse the sampling methods empirically to check their ability to produce tracking portfolios with good tracking quality in the investment period. Therefore, in this section we focus on analysing the stability of the tracking qualities of the portfolios. Here, we estimate tracking quality by minimizing three different traditional measures (MAD, MSE, and TEV), which were defined in Section 3.2.1. We use these measures as objectives in the optimization problem used to solve the index tracking problem. The optimization problem was defined in Section 3.3.1. This means that we are no longer searching for the approximation of returns that produces the tracking portfolio with the highest tracking quality. Instead, we want to compare the results for the estimation periods with those for the investment periods, in order to prove the stability of these three different measures of tracking quality. Furthermore, the assets selected for the tracking portfolios are not important in this context.

We use the data for the S&P1500 index and the same combinations of assets that we used in Section 4.1.1.<sup>208</sup> For each time period (as defined earlier) and each portfolio size, 5,000 tracking portfolios are estimated using three different measures of tracking quality (MAD, MSE and TEV) as the objectives. In total,  $(5,000 \text{ (combinations)} * 8 \text{ (sizes)} * 3 \text{ (objectives)} * 5 \text{ (periods)}) = 600,000$  tracking portfolios are computed. These portfolios are used for three separate analyses.

In the first piece of analysis, we examine the absolute values of the three traditional measures of tracking quality. Thus, we want to analyse whether there are significant differences in quality between the estimation period and the investment period. A significant difference would mean that the tracking quality in the estimation period was not a good indicator of the tracking quality in the investment period. To achieve this, we compare the value of the measure used as the objective for the estimation period with the value of the same measure in the investment period.

In the second piece of analysis, we evaluate the stability of the relative tracking quality of each of the three measures used as objectives. For this, we use each of the 5,000 tracking portfolios for every combination of portfolio size, time period and objective. For each one, we calculate the ratio of relative stability, which was defined in Section 3.4.2. Additionally, to identify whether there is dependence among the relative positions of the tracking portfolios

---

<sup>208</sup> The results for the HDAX index can be found in Roßbach/Karlow (2011).

regarding their qualities in the estimation and investment periods, the Spearman and Kendall rank correlation coefficients are computed.

The third piece of analysis is a comparison of the ETQs of the measures used as objectives. We analyse whether it is possible to predict the ETQ measure in the investment period from the values of the traditional measures in the estimation period.

## 4.2.2 Empirical results

### Absolute stability of traditional measures

	Size	MAD				MSE				TEV			
		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	0.2092	0.0315	0.2292	0.0466	0.2761	0.0417	0.2975	0.0621	0.2763	0.0419	0.2956	0.0609
	60	0.1511	0.0209	0.1831	0.0329	0.2000	0.0285	0.2340	0.0435	0.2003	0.0286	0.2325	0.0430
	90	0.1230	0.0164	0.1612	0.0291	0.1633	0.0222	0.2043	0.0370	0.1637	0.0222	0.2033	0.0362
	120	0.1055	0.0143	0.1472	0.0257	0.1401	0.0194	0.1864	0.0333	0.1405	0.0195	0.1854	0.0322
	150	0.0937	0.0125	0.1379	0.0241	0.1244	0.0176	0.1741	0.0300	0.1248	0.0179	0.1732	0.0298
	180	0.0841	0.0118	0.1300	0.0227	0.1115	0.0163	0.1636	0.0291	0.1120	0.0165	0.1630	0.0285
	210	0.0766	0.0106	0.1246	0.0218	0.1011	0.0145	0.1570	0.0279	0.1017	0.0146	0.1562	0.0271
Period 2	240	0.0703	0.0102	0.1195	0.0198	0.0923	0.0139	0.1505	0.0258	0.0932	0.0141	0.1497	0.0249
	30	0.2058	0.0324	0.2391	0.0450	0.2691	0.0428	0.3066	0.0576	0.2692	0.0428	0.3068	0.0594
	60	0.1490	0.0216	0.1889	0.0336	0.1943	0.0283	0.2393	0.0418	0.1948	0.0283	0.2392	0.0429
	90	0.1227	0.0168	0.1668	0.0272	0.1590	0.0218	0.2096	0.0336	0.1596	0.0218	0.2093	0.0340
	120	0.1056	0.0143	0.1526	0.0252	0.1366	0.0181	0.1917	0.0300	0.1372	0.0181	0.1910	0.0304
	150	0.0942	0.0127	0.1427	0.0234	0.1211	0.0163	0.1784	0.0287	0.1219	0.0165	0.1776	0.0293
	180	0.0853	0.0120	0.1354	0.0220	0.1089	0.0150	0.1694	0.0272	0.1098	0.0151	0.1684	0.0277
Period 3	210	0.0785	0.0112	0.1288	0.0213	0.0994	0.0137	0.1611	0.0263	0.1006	0.0137	0.1598	0.0266
	240	0.0724	0.0104	0.1243	0.0202	0.0914	0.0128	0.1552	0.0247	0.0926	0.0128	0.1536	0.0251
	30	0.1874	0.0304	0.3167	0.0717	0.2467	0.0399	0.4169	0.0987	0.2466	0.0401	0.4191	0.1009
	60	0.1334	0.0191	0.2433	0.0484	0.1755	0.0248	0.3159	0.0651	0.1755	0.0249	0.3172	0.0658
	90	0.1077	0.0146	0.2107	0.0394	0.1419	0.0192	0.2713	0.0528	0.1419	0.0192	0.2721	0.0537
	120	0.0918	0.0122	0.1920	0.0358	0.1211	0.0160	0.2441	0.0465	0.1211	0.0161	0.2450	0.0470
	150	0.0803	0.0112	0.1779	0.0327	0.1058	0.0145	0.2259	0.0418	0.1059	0.0145	0.2266	0.0423
Period 4	180	0.0721	0.0097	0.1674	0.0294	0.0945	0.0131	0.2122	0.0385	0.0946	0.0130	0.2126	0.0393
	210	0.0650	0.0090	0.1581	0.0289	0.0851	0.0116	0.2012	0.0364	0.0854	0.0116	0.2011	0.0369
	240	0.0595	0.0087	0.1517	0.0269	0.0776	0.0116	0.1925	0.0348	0.0778	0.0116	0.1924	0.0355
	30	0.2247	0.0336	0.3739	0.0706	0.2988	0.0440	0.4904	0.0957	0.2990	0.0439	0.4910	0.0968
	60	0.1611	0.0216	0.2972	0.0522	0.2133	0.0276	0.3872	0.0682	0.2138	0.0276	0.3865	0.0696
	90	0.1320	0.0179	0.2627	0.0454	0.1744	0.0227	0.3404	0.0610	0.1749	0.0229	0.3382	0.0610
	120	0.1136	0.0151	0.2418	0.0410	0.1497	0.0196	0.3128	0.0545	0.1502	0.0195	0.3100	0.0553
Period 5	150	0.1014	0.0136	0.2262	0.0389	0.1332	0.0174	0.2936	0.0516	0.1337	0.0175	0.2901	0.0517
	180	0.0919	0.0121	0.2155	0.0359	0.1202	0.0154	0.2789	0.0477	0.1207	0.0155	0.2747	0.0480
	210	0.0844	0.0112	0.2056	0.0344	0.1098	0.0142	0.2682	0.0455	0.1105	0.0143	0.2634	0.0457
	240	0.0782	0.0107	0.1983	0.0339	0.1016	0.0139	0.2583	0.0457	0.1023	0.0139	0.2531	0.0458
	30	0.4376	0.0652	0.5726	0.1464	0.5973	0.0912	0.7392	0.1865	0.5982	0.0913	0.7284	0.1825
	60	0.3362	0.0441	0.4885	0.1120	0.4553	0.0624	0.6240	0.1367	0.4566	0.0626	0.6058	0.1322
	90	0.2906	0.0389	0.4516	0.0994	0.3928	0.0537	0.5764	0.1236	0.3941	0.0541	0.5527	0.1179
Period 5	120	0.2635	0.0358	0.4299	0.0943	0.3556	0.0496	0.5485	0.1153	0.3569	0.0499	0.5197	0.1105
	150	0.2448	0.0337	0.4186	0.0890	0.3306	0.0485	0.5295	0.1088	0.3315	0.0486	0.4960	0.1058
	180	0.2310	0.0323	0.4046	0.0859	0.3123	0.0448	0.5111	0.1060	0.3131	0.0450	0.4744	0.1015
	210	0.2197	0.0310	0.3977	0.0874	0.2988	0.0448	0.5027	0.1043	0.2996	0.0447	0.4637	0.0987
	240	0.2117	0.0294	0.3917	0.0833	0.2891	0.0424	0.4952	0.1009	0.2897	0.0431	0.4516	0.0955

Table 4.11: Comparison of Levels of Tracking Quality over Different Time Periods

The first piece of analysis is the comparison between the absolute tracking quality in the estimation period and that in the investment period. This means that we want to examine the difference in the level of tracking quality. To do this, we report the medians and IQRs of the

absolute values of the tracking qualities in the estimation and investment periods, in Table 4.11.

The medians and IQRs are reported for the 5,000 random tracking portfolios of each size, for each period and for each measure of tracking quality used as an objective in the optimization problem. For example, 0.2092, in the first row and column, is the median tracking quality, expressed as MAD, of the 5,000 tracking portfolios of size 30 based on period 1. These tracking portfolios were generated by minimizing MAD. The value 0.2761 in the same row represents the median tracking quality, expressed in  $\text{RMSE}^{209}$ , of the 5,000 tracking portfolios generated by minimizing MSE over the 5,000 portfolios of size 30 based on period 1. For the optimization problem that used TEV as the objective, we report the medians and IQRs expressed in STDTE (standard deviation of tracking error).<sup>210</sup> We report the square roots of the MSEs and TEVs because MAD is a linear measure and MSE and TEV are quadratic measures. This means that the measures are of the same scale and can be compared directly.

Table 4.11 indicates that the tracking quality improves as we increase the size of the tracking portfolio. This holds for the estimation period and for the investment period, for all market phases, and for all measures of tracking quality.

To compare the tracking qualities in the estimation period with those in the investment period, we compute the ratio of the value of the measure of tracking quality in the investment period to its value in the estimation period, for every tracking portfolio. Then, the medians, minimal and maximal values of the tracking qualities for every sample of 5,000 portfolios are computed, and reported in Table 4.12. A value below one indicates that the tracking quality is underestimated in the estimation period compared to the investment period, while a value greater than one represents overestimation of tracking quality in the estimation period.

Table 4.12 suggests that the difference between the tracking quality in the investment period and that in the estimation period increases as the size of the tracking portfolios increases. The ratios for all measures are highest in periods 3 and 4. However, the ratios in the turbulent period 5 are low. The quality of tracking is overestimated in the estimation period, on average, for all periods. However, the minimal values of the ratios, which are lower than unity, indicate that there are some portfolios for which the tracking quality is underestimated. The maximal values of the ratios suggest meanwhile that there are cases for which the tracking quality is highly overestimated in the estimation period. In the worst cases, the quality in the investment period is around four times worse than in the estimation period. For example, this occurs in periods 3 and 4.

---

<sup>209</sup> The square root of MSE.

<sup>210</sup> The square root of TEV.

	Size	MAD			MSE			TEV		
		Med	Min	Max	Med	Min	Max	Med	Min	Max
Period 1	30	1.09	0.70	1.65	1.07	0.71	1.76	1.06	0.70	1.75
	60	1.21	0.78	1.84	1.17	0.72	2.07	1.16	0.71	2.07
	90	1.31	0.79	1.92	1.25	0.75	1.94	1.24	0.72	1.91
	120	1.40	0.85	2.18	1.33	0.88	2.09	1.31	0.88	2.05
	150	1.47	0.91	2.12	1.41	0.82	2.26	1.39	0.80	2.19
	180	1.55	1.04	2.41	1.48	0.92	2.29	1.46	0.88	2.21
	210	1.63	0.96	3.01	1.55	0.99	2.76	1.54	1.00	2.73
	240	1.70	1.07	2.69	1.64	1.06	2.68	1.61	1.01	2.61
Period 2	30	1.16	0.74	1.90	1.14	0.72	1.90	1.14	0.74	1.93
	60	1.27	0.78	1.96	1.23	0.79	1.86	1.23	0.75	1.88
	90	1.36	0.84	2.05	1.32	0.85	1.96	1.31	0.83	1.92
	120	1.44	0.90	2.28	1.40	0.84	2.14	1.39	0.83	2.09
	150	1.51	0.97	2.26	1.47	0.99	2.25	1.46	0.99	2.23
	180	1.59	0.99	2.61	1.55	1.03	2.66	1.53	0.99	2.63
	210	1.64	0.98	2.41	1.63	1.06	2.41	1.60	1.03	2.32
	240	1.72	1.07	2.61	1.70	1.07	2.73	1.66	1.07	2.66
Period 3	30	1.68	0.99	2.79	1.68	1.04	3.52	1.69	1.03	3.54
	60	1.83	1.14	2.94	1.80	1.14	3.15	1.81	1.12	3.18
	90	1.96	1.23	3.10	1.91	1.18	3.44	1.92	1.16	3.46
	120	2.10	1.27	3.26	2.02	1.27	3.50	2.03	1.27	3.57
	150	2.21	1.41	3.43	2.13	1.34	3.57	2.14	1.29	3.62
	180	2.33	1.55	3.56	2.25	1.46	3.58	2.25	1.45	3.59
	210	2.44	1.53	3.83	2.36	1.49	4.05	2.36	1.48	4.04
	240	2.56	1.60	4.00	2.48	1.48	4.12	2.47	1.47	4.09
Period 4	30	1.67	1.10	2.82	1.64	1.04	3.31	1.64	1.02	3.41
	60	1.84	1.19	2.68	1.81	1.19	2.74	1.80	1.17	2.79
	90	1.99	1.27	3.10	1.95	1.29	3.41	1.93	1.29	3.33
	120	2.12	1.42	3.43	2.08	1.39	3.38	2.06	1.37	3.28
	150	2.23	1.37	3.36	2.20	1.43	3.50	2.17	1.42	3.51
	180	2.34	1.40	3.77	2.31	1.48	4.00	2.27	1.46	3.89
	210	2.44	1.56	3.78	2.42	1.29	3.93	2.37	1.31	3.86
	240	2.53	1.61	3.90	2.54	1.64	4.25	2.47	1.50	4.07
Period 5	30	1.31	0.79	2.55	1.24	0.76	2.25	1.22	0.76	2.24
	60	1.45	0.86	2.44	1.37	0.78	2.42	1.33	0.76	2.33
	90	1.56	0.82	2.61	1.47	0.81	2.46	1.41	0.78	2.37
	120	1.64	0.95	2.80	1.54	0.91	2.49	1.46	0.88	2.36
	150	1.71	0.88	2.92	1.60	0.89	2.63	1.50	0.86	2.48
	180	1.75	0.94	3.00	1.64	0.93	2.68	1.52	0.85	2.56
	210	1.81	0.99	3.03	1.68	0.98	2.92	1.54	0.90	2.76
	240	1.86	1.01	3.04	1.72	0.91	2.99	1.57	0.85	2.67

Table 4.12: Comparison of Ratios: Investment/Estimation

The results of the first piece of analysis suggest that it is possible to improve the quality of tracking in the estimation period and in the investment period by increasing the size of the tracking portfolio. However, at the same time, the discrepancy between the tracking quality in the estimation period and that in the investment period also increases. Hence, the absolute stability of the tracking quality deteriorates. This indicates an increasing overfitting effect, meaning that the tracking quality enhancement in the estimation period resulting from a greater portfolio size does not apply proportionally to the investment period. In general, the tracking portfolios that are estimated by minimizing the traditional measures of tracking quality lack absolute stability in their tracking quality.

## Relative stability of traditional measures

The second piece of analysis is the comparison of the relative positions of the tracking portfolios in the two time periods, as ranked by their quality. If a tracking portfolio is one of the best in the estimation period, is it also one of the best in the investment period, and vice versa?

MAD			Investment period									
			1	2	3	4	5	6	7	8	9	10
Estimation period	Period 1	1	34.60%	19.60%	13.00%	9.60%	8.60%	6.20%	4.60%	1.60%	1.80%	0.40%
		2	21.20%	15.60%	13.80%	13.40%	9.80%	8.60%	7.60%	6.20%	2.60%	1.20%
		3	10.60%	12.80%	13.20%	13.20%	13.40%	10.80%	8.40%	7.80%	6.20%	3.60%
		4	10.20%	13.00%	14.40%	12.00%	10.80%	9.20%	10.80%	7.00%	8.80%	3.80%
		5	8.60%	10.20%	10.80%	10.60%	12.00%	11.00%	12.40%	10.20%	9.60%	4.60%
		6	6.80%	8.20%	11.60%	12.40%	9.40%	14.00%	9.60%	10.60%	9.80%	7.60%
		7	3.40%	7.20%	9.80%	9.20%	12.40%	12.00%	11.80%	11.60%	13.20%	9.40%
		8	3.40%	6.80%	5.80%	8.00%	10.00%	12.60%	10.60%	14.20%	14.00%	14.60%
		9	0.20%	3.60%	3.80%	7.80%	8.40%	9.60%	13.00%	15.20%	18.80%	19.60%
		10	1.00%	3.00%	3.80%	3.80%	5.20%	6.00%	11.20%	15.60%	15.20%	35.20%
	Period 2	1	27.20%	15.60%	13.60%	10.80%	10.20%	4.40%	7.00%	6.40%	2.80%	2.00%
		2	13.80%	18.00%	13.80%	13.40%	9.40%	8.00%	6.20%	7.40%	4.40%	5.60%
		3	14.20%	14.60%	10.40%	11.80%	8.80%	10.20%	10.80%	8.20%	6.80%	4.20%
		4	12.40%	12.20%	12.80%	8.40%	12.00%	10.00%	10.00%	8.40%	9.40%	4.40%
		5	9.40%	10.00%	9.80%	10.20%	9.80%	13.60%	10.20%	9.40%	10.80%	6.80%
		6	8.40%	9.60%	9.60%	10.40%	10.00%	13.20%	8.80%	9.20%	11.60%	9.20%
		7	7.40%	7.00%	9.80%	12.20%	9.40%	12.20%	10.00%	13.40%	10.00%	8.60%
		8	3.60%	5.60%	9.00%	9.60%	10.80%	10.20%	12.60%	10.80%	14.60%	13.20%
		9	2.60%	4.60%	6.80%	7.40%	11.00%	10.40%	13.80%	12.80%	14.80%	15.80%
		10	1.00%	2.80%	4.40%	5.80%	8.60%	7.80%	10.60%	14.00%	14.80%	30.20%
	Period 3	1	28.40%	17.80%	13.60%	11.80%	9.60%	6.60%	4.80%	4.60%	1.80%	1.00%
		2	18.00%	18.20%	12.80%	12.40%	10.80%	7.00%	7.60%	6.00%	5.00%	2.20%
		3	14.20%	10.00%	14.00%	10.60%	12.40%	10.40%	8.60%	8.60%	5.80%	5.40%
		4	10.20%	14.00%	13.60%	11.80%	8.60%	12.40%	9.20%	8.60%	5.80%	5.80%
		5	8.00%	9.20%	11.00%	12.40%	13.00%	11.00%	9.60%	9.80%	10.40%	5.60%
		6	8.00%	9.80%	7.40%	10.00%	8.20%	12.00%	12.80%	11.00%	11.20%	9.60%
		7	5.00%	8.00%	10.00%	8.20%	12.00%	11.20%	12.40%	12.00%	10.40%	10.80%
		8	3.80%	5.80%	7.40%	9.20%	11.00%	11.80%	11.20%	12.20%	14.80%	12.80%
		9	3.40%	4.20%	4.60%	7.40%	7.80%	8.60%	10.60%	14.40%	17.60%	21.40%
		10	1.00%	3.00%	5.60%	6.20%	6.60%	9.00%	13.20%	12.80%	17.20%	25.40%
	Period 4	1	28.60%	16.80%	13.40%	11.60%	8.40%	8.00%	4.60%	4.40%	3.00%	1.20%
		2	15.60%	18.60%	12.60%	8.80%	10.00%	11.20%	7.20%	7.60%	3.60%	4.80%
		3	13.00%	14.80%	14.60%	9.80%	11.40%	10.40%	10.60%	7.20%	4.60%	3.60%
		4	13.80%	10.60%	10.60%	11.60%	13.00%	10.00%	10.20%	8.00%	7.60%	4.60%
		5	10.60%	8.80%	10.20%	10.40%	10.60%	14.00%	12.40%	10.20%	8.00%	4.80%
		6	7.60%	7.40%	10.20%	14.00%	11.40%	6.20%	10.00%	11.20%	11.60%	10.40%
		7	3.80%	8.60%	11.60%	10.00%	11.00%	9.00%	10.80%	13.40%	11.80%	10.00%
		8	3.80%	7.00%	8.20%	9.20%	7.40%	12.40%	11.20%	11.80%	14.80%	14.20%
		9	2.40%	4.80%	5.40%	10.40%	9.40%	11.40%	11.60%	12.40%	18.00%	14.20%
		10	0.80%	2.60%	3.20%	4.20%	7.40%	7.40%	11.40%	13.80%	17.00%	32.20%
	Period 5	1	22.00%	16.60%	14.60%	12.40%	9.80%	7.00%	6.20%	5.80%	3.40%	2.20%
		2	17.40%	14.60%	11.80%	13.60%	9.00%	10.60%	6.40%	6.80%	5.40%	4.40%
		3	15.00%	11.60%	15.00%	9.80%	11.20%	9.60%	8.40%	9.00%	6.60%	3.80%
		4	10.00%	12.20%	12.00%	9.80%	10.20%	10.20%	12.40%	8.80%	7.60%	6.80%
		5	8.40%	12.00%	8.00%	12.40%	11.80%	10.80%	9.40%	11.00%	8.20%	8.00%
		6	6.40%	11.80%	8.80%	9.40%	11.00%	9.60%	11.00%	11.40%	11.00%	9.60%
		7	6.80%	5.80%	11.20%	10.80%	7.80%	9.60%	12.60%	10.40%	11.80%	13.20%
		8	5.80%	5.80%	7.40%	8.60%	9.20%	13.00%	11.40%	12.40%	13.60%	12.80%
		9	5.20%	6.00%	6.80%	5.60%	11.00%	11.40%	12.40%	11.60%	13.80%	16.20%
		10	3.00%	3.60%	4.40%	7.60%	9.00%	8.20%	9.80%	12.80%	18.60%	23.00%

Table 4.13: Stability of Relative Tracking Quality (MAD, TP size: 120)

Here, for each measure, tracking portfolio size and time period, we rank the 5,000 tracking portfolios according to their tracking quality in the estimation period and again according

their tracking quality in the investment period. Then, we allocate the tracking portfolios into deciles according to their ranks, separately, for each period. Finally, we build a matrix according to these deciles (see Table 4.13).

Table 4.13 shows the matrices for the five periods using the measure MAD and the tracking portfolio size 120. On the main diagonal of the matrices one can find, for each decile of 500 portfolios based on the estimation period, the percentages of those 500 portfolios that are allocated to the same decile in the investment period. The values on this diagonal indicate the stability of the relative tracking quality. This means, for example, looking at the row and column labelled 1, for period 1, that 34.60% of the tracking portfolios that are in the best 10 percent in the estimation period are also in the best 10 percent in the investment period. The other values in the rows of the matrices show the percentage of the tracking portfolios in a given decile in the estimation period that change decile in the investment period. In row 1, column 4 for period 1, the 9.6% indicates that 48 out of the 500 best tracking portfolios (top decile) in the estimation period moved to the fourth decile in the investment period.

Because of the number of matrices that would be produced if we used all of the our tracking portfolios and measures, we select the tracking portfolio size of 120 and the MAD measure for presentation. The values in the matrices show that majority of the tracking portfolios do not remain in the same decile and instead have a broad distribution across deciles in the investment period.

The stability ratio from Section 3.4.2 can be used to characterize the matrices in Table 4.13 in terms of a single number. Hence, to analyse the relative stability of the other measures of tracking quality and the other sizes of tracking portfolios, we calculate the stability ratio, and the Kendall and Spearman rank correlation coefficients (see Table 4.14).

The smaller is the stability ratio, the better is the predictive power of the measure of tracking quality. If the stability ratio is near to 100%, this indicates an almost entirely random case, where the relative positions of the tracking portfolios in the estimation period are not related to their relative positions in the investment period at all. The values of the Spearman and Kendall rank correlation coefficients are in the interval  $[-100\%, 100\%]$ . Here, a value of zero corresponds to the random case. We test whether the values of the Kendall and Spearman correlation coefficients are different from zero.

	TP Size	Stability Ratio			Kendall Rank Correlation			Spearman Rank Correlation		
		MAD	MSE	TEV	MAD	MSE	TEV	MAD	MSE	TEV
Period 1	30	58.03%	58.74%	59.94%	44.28%*	43.39%*	42.18%*	62.21%*	60.96%*	59.50%*
	60	64.79%	64.59%	65.75%	37.74%*	37.59%*	36.40%*	54.00%*	53.89%*	52.35%*
	90	67.00%	67.60%	68.74%	35.56%*	34.62%*	33.33%*	51.14%*	49.95%*	48.21%*
	120	67.55%	67.97%	69.12%	34.78%*	34.08%*	33.04%*	50.11%*	49.10%*	47.74%*
	150	69.73%	69.74%	70.81%	32.84%*	32.39%*	31.21%*	47.80%*	46.94%*	45.35%*
	180	69.02%	68.88%	70.57%	33.03%*	33.01%*	31.40%*	47.80%*	47.73%*	45.54%*
	210	70.64%	70.20%	71.07%	31.74%*	32.05%*	31.09%*	46.18%*	46.68%*	45.34%*
Period 2	240	72.19%	71.87%	73.18%	30.16%*	30.42%*	28.86%*	44.02%*	44.29%*	42.06%*
	30	66.83%	66.51%	65.93%	35.32%*	35.92%*	36.44%*	50.67%*	51.57%*	52.27%*
	60	71.34%	70.93%	70.28%	31.12%*	31.30%*	31.92%*	45.24%*	45.40%*	46.26%*
	90	74.42%	73.26%	72.46%	27.84%*	29.20%*	30.11%*	40.79%*	42.79%*	44.04%*
	120	75.75%	75.06%	73.89%	26.32%*	27.10%*	28.18%*	38.57%*	39.67%*	41.17%*
	150	76.41%	74.29%	73.52%	25.54%*	27.92%*	28.75%*	37.55%*	40.84%*	42.01%*
	180	78.66%	75.82%	75.01%	23.70%*	26.51%*	27.38%*	35.02%*	38.96%*	40.21%*
Period 3	210	76.01%	73.88%	72.56%	26.08%*	28.22%*	29.65%*	38.36%*	41.26%*	43.22%*
	240	78.76%	76.79%	75.25%	23.26%*	25.84%*	27.37%*	34.36%*	38.11%*	40.29%*
	30	67.94%	70.14%	71.17%	34.45%*	32.34%*	31.23%*	49.71%*	46.96%*	45.46%*
	60	70.25%	71.43%	72.31%	31.86%*	30.82%*	30.04%*	46.22%*	44.86%*	43.81%*
	90	70.80%	72.59%	73.26%	31.29%*	29.67%*	28.91%*	45.39%*	43.18%*	42.12%*
	120	73.02%	74.79%	75.57%	29.15%*	27.37%*	26.69%*	42.66%*	40.20%*	39.26%*
	150	71.26%	74.62%	75.33%	31.05%*	28.01%*	27.23%*	45.23%*	41.12%*	40.04%*
Period 4	180	72.90%	75.09%	75.71%	28.95%*	27.09%*	26.52%*	42.18%*	39.69%*	38.90%*
	210	73.87%	75.54%	75.84%	28.20%*	26.48%*	26.11%*	41.15%*	38.69%*	38.16%*
	240	75.66%	77.06%	77.29%	26.74%*	25.48%*	25.23%*	39.24%*	37.52%*	37.15%*
	30	65.57%	66.49%	66.47%	36.61%*	35.71%*	35.65%*	52.56%*	51.26%*	51.19%*
	60	70.03%	70.57%	70.37%	32.11%*	31.43%*	31.56%*	46.56%*	45.61%*	45.78%*
	90	71.94%	71.90%	71.69%	30.08%*	29.60%*	29.91%*	43.78%*	43.02%*	43.48%*
	120	73.43%	73.34%	72.91%	28.85%*	28.71%*	29.11%*	42.14%*	41.89%*	42.45%*
Period 5	150	74.26%	74.88%	74.47%	28.28%*	27.37%*	27.75%*	41.39%*	39.95%*	40.52%*
	180	76.43%	76.72%	75.95%	25.57%*	25.28%*	26.13%*	37.47%*	37.09%*	38.31%*
	210	78.54%	78.79%	78.29%	23.48%*	23.10%*	23.66%*	34.57%*	34.03%*	34.82%*
	240	77.39%	78.45%	77.83%	24.89%*	23.49%*	24.09%*	36.63%*	34.55%*	35.39%*
	30	72.74%	71.23%	70.52%	29.43%*	31.23%*	31.89%*	42.89%*	45.39%*	46.25%*
	60	79.11%	76.56%	75.71%	23.24%*	25.43%*	26.28%*	34.36%*	37.31%*	38.51%*
	90	76.69%	73.52%	72.56%	25.68%*	29.15%*	30.07%*	37.76%*	42.65%*	43.90%*
Period 5	120	78.56%	75.57%	74.87%	23.41%*	26.71%*	27.35%*	34.59%*	39.14%*	40.00%*
	150	78.11%	74.77%	74.21%	23.62%*	27.25%*	27.85%*	34.85%*	39.99%*	40.84%*
	180	78.89%	76.87%	76.31%	22.83%*	25.45%*	26.07%*	33.72%*	37.51%*	38.42%*
	210	77.26%	75.06%	75.09%	24.92%*	27.25%*	27.49%*	36.63%*	39.94%*	40.30%*
	240	77.56%	76.98%	76.28%	24.69%*	25.11%*	25.69%*	36.39%*	36.90%*	37.72%*

\* Significantly different from zero with level 0.05

Table 4.14: Stability Ratio and Correlation Coefficients for Traditional Measures

The stability ratios in Table 4.14 indicate that it is possible to predict the tracking quality for the investment period from the quality in the estimation period, since the values are lower than 100%. However, all of the values are larger than 50%. The best prediction of the relative positions of the tracking portfolios can be made in the relatively quiet period 1. The results also suggest that the tracking portfolio size influences predictability. Predictability increases as the size of the tracking portfolios decreases. This can be explained by the fact that the tracking quality is improved by increasing the size of the tracking portfolio, and it thus becomes difficult to distinguish the portfolios from each other if we only use the measure of tracking quality in the estimation period. These findings are confirmed by the Kendall and Spearman correlation coefficients, in the middle and right hand panels of the table, which are significantly different from zero.

We also analyse the relative stability by selecting the best 10% of the 5,000 tracking portfolios and comparing their tracking qualities to the entire sample of 5,000. First, the tracking qualities are reported in Table 4.15.

Size		MAD				MSE				TEV			
		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
Period 1	30	0.1757	0.0093	0.1969	0.0347	0.2314	0.0124	0.2527	0.0452	0.2315	0.0126	0.2528	0.0443
	60	0.1277	0.0060	0.1609	0.0253	0.1688	0.0087	0.2067	0.0326	0.1689	0.0088	0.2057	0.0316
	90	0.1046	0.0058	0.1418	0.0217	0.1384	0.0076	0.1805	0.0305	0.1387	0.0075	0.1803	0.0304
	120	0.0889	0.0049	0.1301	0.0213	0.1176	0.0070	0.1636	0.0264	0.1178	0.0072	0.1639	0.0255
	150	0.0790	0.0045	0.1247	0.0167	0.1041	0.0057	0.1557	0.0246	0.1045	0.0058	0.1558	0.0242
	180	0.0707	0.0039	0.1160	0.0177	0.0930	0.0054	0.1472	0.0234	0.0935	0.0053	0.1472	0.0230
	210	0.0645	0.0036	0.1113	0.0178	0.0843	0.0052	0.1404	0.0219	0.0848	0.0053	0.1403	0.0217
	240	0.0592	0.0033	0.1085	0.0163	0.0771	0.0044	0.1357	0.0187	0.0777	0.0043	0.1360	0.0179
Period 2	30	0.1710	0.0096	0.2148	0.0355	0.2241	0.0121	0.2739	0.0442	0.2245	0.0120	0.2733	0.0447
	60	0.1250	0.0065	0.1703	0.0280	0.1639	0.0083	0.2144	0.0350	0.1645	0.0083	0.2135	0.0356
	90	0.1037	0.0055	0.1545	0.0228	0.1350	0.0075	0.1921	0.0290	0.1357	0.0074	0.1918	0.0274
	120	0.0897	0.0046	0.1402	0.0227	0.1165	0.0061	0.1761	0.0290	0.1171	0.0061	0.1752	0.0288
	150	0.0795	0.0045	0.1321	0.0184	0.1024	0.0056	0.1644	0.0228	0.1030	0.0057	0.1621	0.0236
	180	0.0713	0.0039	0.1267	0.0204	0.0916	0.0053	0.1567	0.0242	0.0924	0.0053	0.1541	0.0258
	210	0.0654	0.0036	0.1192	0.0184	0.0836	0.0045	0.1474	0.0221	0.0845	0.0046	0.1449	0.0222
	240	0.0604	0.0032	0.1150	0.0182	0.0768	0.0041	0.1427	0.0214	0.0779	0.0042	0.1403	0.0207
Period 3	30	0.1558	0.0092	0.2725	0.0546	0.2056	0.0112	0.3615	0.0733	0.2055	0.0112	0.3649	0.0764
	60	0.1123	0.0055	0.2174	0.0359	0.1480	0.0074	0.2794	0.0514	0.1479	0.0077	0.2817	0.0515
	90	0.0908	0.0046	0.1883	0.0326	0.1201	0.0064	0.2431	0.0468	0.1198	0.0065	0.2439	0.0473
	120	0.0776	0.0043	0.1726	0.0283	0.1026	0.0057	0.2227	0.0383	0.1025	0.0056	0.2242	0.0384
	150	0.0684	0.0035	0.1602	0.0238	0.0902	0.0043	0.2054	0.0294	0.0903	0.0043	0.2058	0.0310
	180	0.0606	0.0036	0.1495	0.0249	0.0792	0.0044	0.1917	0.0331	0.0795	0.0044	0.1925	0.0345
	210	0.0550	0.0030	0.1435	0.0228	0.0717	0.0043	0.1804	0.0316	0.0720	0.0042	0.1816	0.0327
	240	0.0498	0.0029	0.1374	0.0210	0.0646	0.0037	0.1752	0.0292	0.0650	0.0036	0.1745	0.0308
Period 4	30	0.1892	0.0089	0.3326	0.0544	0.2520	0.0124	0.4345	0.0721	0.2523	0.0123	0.4332	0.0732
	60	0.1370	0.0076	0.2654	0.0433	0.1818	0.0103	0.3432	0.0590	0.1821	0.0101	0.3424	0.0590
	90	0.1125	0.0053	0.2380	0.0360	0.1488	0.0073	0.3080	0.0531	0.1493	0.0073	0.3071	0.0513
	120	0.0969	0.0054	0.2206	0.0370	0.1277	0.0067	0.2848	0.0468	0.1281	0.0069	0.2812	0.0476
	150	0.0859	0.0043	0.2052	0.0361	0.1130	0.0060	0.2661	0.0507	0.1134	0.0058	0.2632	0.0481
	180	0.0779	0.0042	0.1971	0.0308	0.1020	0.0051	0.2565	0.0416	0.1027	0.0053	0.2517	0.0425
	210	0.0714	0.0040	0.1907	0.0301	0.0937	0.0046	0.2468	0.0451	0.0942	0.0048	0.2415	0.0439
	240	0.0659	0.0036	0.1815	0.0302	0.0862	0.0046	0.2380	0.0390	0.0868	0.0048	0.2332	0.0420
Period 5	30	0.3663	0.0206	0.4924	0.1077	0.4969	0.0315	0.6411	0.1328	0.4973	0.0320	0.6293	0.1268
	60	0.2844	0.0149	0.4411	0.0860	0.3835	0.0227	0.5580	0.1163	0.3841	0.0220	0.5386	0.1031
	90	0.2459	0.0139	0.4050	0.0771	0.3305	0.0194	0.5108	0.0998	0.3316	0.0188	0.4895	0.0967
	120	0.2224	0.0119	0.3918	0.0744	0.2977	0.0162	0.4863	0.0932	0.2982	0.0163	0.4574	0.0925
	150	0.2063	0.0099	0.3823	0.0825	0.2781	0.0151	0.4747	0.0964	0.2792	0.0150	0.4414	0.0887
	180	0.1955	0.0102	0.3708	0.0709	0.2630	0.0144	0.4633	0.0891	0.2630	0.0146	0.4290	0.0771
	210	0.1850	0.0090	0.3648	0.0715	0.2500	0.0148	0.4559	0.0820	0.2507	0.0141	0.4184	0.0777
	240	0.1788	0.0098	0.3600	0.0659	0.2422	0.0129	0.4564	0.0830	0.2424	0.0130	0.4138	0.0794

Table 4.15: Levels of Tracking Quality over Different Time Periods (Best 10%)

The results in Table 4.15 repeat the point for the 5,000 tracking portfolios shown in Table 4.11: the tracking quality improves as the size of the tracking portfolio increases.

To see whether there is an improvement in tracking quality when we select the best 10% of the tracking portfolios rather than take the average across all 5,000 tracking portfolios, we compute the ratios of the median tracking quality for all 5,000 portfolios to the median for the best 10% of portfolios, in the estimation period and in the investment period. These ratios are reported in Table 4.16.



	Size	MAD		MSE		TEV	
		Est	Inv	Est	Inv	Est	Inv
Period 1	30	1.19	1.16	1.19	1.18	1.19	1.17
	60	1.18	1.14	1.19	1.13	1.19	1.13
	90	1.18	1.14	1.18	1.13	1.18	1.13
	120	1.19	1.13	1.19	1.14	1.19	1.13
	150	1.19	1.11	1.19	1.12	1.19	1.11
	180	1.19	1.12	1.20	1.11	1.20	1.11
	210	1.19	1.12	1.20	1.12	1.20	1.11
	240	1.19	1.10	1.20	1.11	1.20	1.10
Period 2	30	1.20	1.11	1.20	1.12	1.20	1.12
	60	1.19	1.11	1.19	1.12	1.18	1.12
	90	1.18	1.08	1.18	1.09	1.18	1.09
	120	1.18	1.09	1.17	1.09	1.17	1.09
	150	1.19	1.08	1.18	1.09	1.18	1.10
	180	1.20	1.07	1.19	1.08	1.19	1.09
	210	1.20	1.08	1.19	1.09	1.19	1.10
	240	1.20	1.08	1.19	1.09	1.19	1.10
Period 3	30	1.20	1.16	1.20	1.15	1.20	1.15
	60	1.19	1.12	1.19	1.13	1.19	1.13
	90	1.19	1.12	1.18	1.12	1.18	1.12
	120	1.18	1.11	1.18	1.10	1.18	1.09
	150	1.17	1.11	1.17	1.10	1.17	1.10
	180	1.19	1.12	1.19	1.11	1.19	1.10
	210	1.18	1.10	1.19	1.12	1.19	1.11
	240	1.20	1.10	1.20	1.10	1.20	1.10
Period 4	30	1.19	1.12	1.19	1.13	1.19	1.13
	60	1.18	1.12	1.17	1.13	1.17	1.13
	90	1.17	1.10	1.17	1.11	1.17	1.10
	120	1.17	1.10	1.17	1.10	1.17	1.10
	150	1.18	1.10	1.18	1.10	1.18	1.10
	180	1.18	1.09	1.18	1.09	1.18	1.09
	210	1.18	1.08	1.17	1.09	1.17	1.09
	240	1.19	1.09	1.18	1.09	1.18	1.09
Period 5	30	1.19	1.16	1.20	1.15	1.20	1.16
	60	1.18	1.11	1.19	1.12	1.19	1.12
	90	1.18	1.12	1.19	1.13	1.19	1.13
	120	1.18	1.10	1.19	1.13	1.20	1.14
	150	1.19	1.09	1.19	1.12	1.19	1.12
	180	1.18	1.09	1.19	1.10	1.19	1.11
	210	1.19	1.09	1.20	1.10	1.20	1.11
	240	1.18	1.09	1.19	1.09	1.20	1.09

Table 4.16: Ratios of Tracking Qualities (5000/Best 10%)

In general, the average absolute tracking quality of the best 10% of the portfolios is better than the average across all 5,000 portfolios. This is indicated by the values of the ratios, which are all greater than unity. In the estimation period, the discrepancy between all 5,000 and the best 10% is larger than in the investment period. Moreover, this improvement in the investment period over the estimation period is larger for smaller portfolios.

Hence, the tracking quality can be improved on average in the investment period by selecting the portfolios with better tracking quality in the estimation period. This holds for all periods, all measures, and all portfolio sizes.

We check our findings by plotting the tracking portfolios of size 120 (see Figure 4.5), to demonstrate the distribution of tracking quality in the estimation (x-axis) and investment (y-axis) periods.

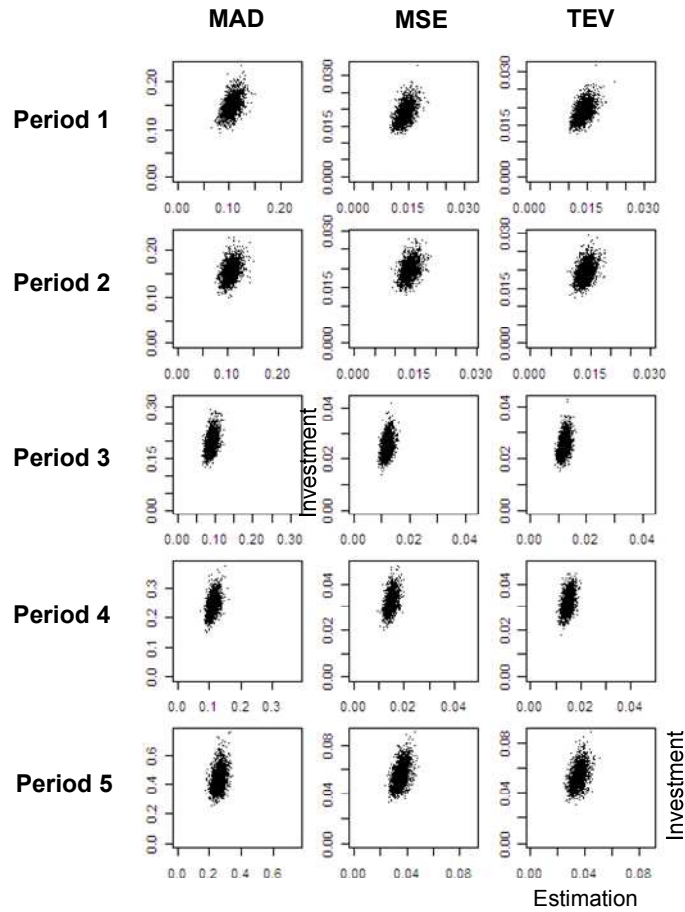


Figure 4.5: Scatter Plots for Traditional Measures, for Tracking Portfolios of Size 120

The ellipsoidal shape of the distribution, which is elongated along the axis for the investment period, confirms our previous findings, that the quality of tracking in the estimation period is better than in the investment period for all cases, and that the tracking quality in the investment period is poorly predicted by the tracking quality in the estimation period.

### Prediction of ETQ

The third piece of analysis is a comparison of the ETQs (see equation (3.3)) of the three traditional measures, to evaluate whether there are any significant differences in their predictive value. We report the medians and IQRs of the ETQs for the investment period, in Table 4.17.

		MAD		MSE		TEV	
Size		Med	IQR	Med	IQR	Med	IQR
Period 1	30	1.3797	1.3202	1.3782	1.3426	1.3851	1.3431
	60	1.1016	1.0164	1.0944	1.0227	1.1044	1.0276
	90	0.9474	0.8862	0.8929	0.8397	0.9015	0.8485
	120	0.8492	0.7722	0.8071	0.7372	0.8191	0.7458
	150	0.7928	0.7177	0.7563	0.6824	0.7692	0.6997
	180	0.7611	0.6596	0.7001	0.6235	0.7127	0.6441
	210	0.7044	0.6472	0.6739	0.6076	0.6921	0.6178
	240	0.6866	0.6002	0.6341	0.5738	0.6494	0.5863
Period 2	30	1.0554	0.9282	1.0541	0.9420	1.0474	0.9431
	60	0.8284	0.7424	0.8174	0.7684	0.8063	0.7471
	90	0.7534	0.6832	0.7525	0.7000	0.7290	0.6740
	120	0.7024	0.6064	0.6792	0.5981	0.6511	0.5746
	150	0.6520	0.5748	0.6441	0.5702	0.6214	0.5481
	180	0.6326	0.5553	0.6181	0.5522	0.5847	0.5185
	210	0.5946	0.5311	0.5961	0.5284	0.5574	0.4994
	240	0.5910	0.5426	0.5806	0.5358	0.5360	0.5011
Period 3	30	1.3248	1.1294	1.2883	1.1008	1.2903	1.1069
	60	1.0699	0.8903	1.0237	0.8557	1.0280	0.8609
	90	0.9182	0.7629	0.8839	0.7240	0.8873	0.7224
	120	0.8162	0.6662	0.7902	0.6563	0.7930	0.6532
	150	0.7574	0.6314	0.7315	0.5881	0.7316	0.5908
	180	0.7196	0.6003	0.6891	0.5603	0.6897	0.5646
	210	0.6751	0.5366	0.6481	0.5271	0.6490	0.5325
	240	0.6516	0.5411	0.6280	0.5102	0.6305	0.5182
Period 4	30	1.7823	1.5273	1.7506	1.4804	1.7509	1.4819
	60	1.4739	1.2938	1.4463	1.2425	1.4240	1.2322
	90	1.3978	1.1817	1.3349	1.1042	1.2996	1.0827
	120	1.2857	1.1210	1.2346	1.0858	1.1971	1.0535
	150	1.2218	1.0783	1.1747	1.0228	1.1294	0.9944
	180	1.1765	1.0004	1.1185	0.9806	1.0734	0.9465
	210	1.1535	1.0103	1.1007	0.9381	1.0324	0.9076
	240	1.1297	0.9469	1.0765	0.8999	1.0160	0.8621
Period 5	30	3.5612	3.3031	3.5333	3.2762	3.3395	3.1193
	60	3.3402	2.8623	3.2455	2.7013	2.9327	2.5469
	90	3.2712	2.5019	3.1839	2.4244	2.7118	2.2661
	120	3.2459	2.4145	3.2345	2.3416	2.6463	2.1756
	150	3.2457	2.2873	3.1820	2.2009	2.5110	2.0555
	180	3.3062	2.2051	3.2465	2.0659	2.4764	1.9719
	210	3.3873	2.1883	3.3195	2.0436	2.4831	1.8958
	240	3.4720	2.1088	3.4309	1.9098	2.5228	1.7932

Table 4.17: Comparison of Levels of Tracking Quality in the Investment Period (ETQ)

The results suggest that the ETQ is improved and the IQRs become smaller if the number of assets in the tracking portfolios is increased. ETQ is greatly dependent on the market phase. The comparison of values shows that it is nearly the same for periods 1, 2 and 3. In period 4 it is worse, and in period 5 it is much worse. In other words, the tracking portfolios track the index much less closely in period 5 than in the other periods.

To analyse whether it is possible to predict the ETQ from the values of the traditional measures of tracking quality in the estimation period, we compute the stability ratios and the Kendall and Spearman rank correlation coefficients (see Table 4.18).

	TP Size	Stability Ratio			Kendall Rank Correlation			Spearman Rank Correlation		
		MAD	MSE	TEV	MAD	MSE	TEV	MAD	MSE	TEV
Period 1	30	78.10%	77.66%	77.70%	24.20%*	24.64%*	24.55%*	35.79%*	36.36%*	36.22%*
	60	82.39%	81.82%	81.73%	19.68%*	20.49%*	20.44%*	29.18%*	30.37%*	30.29%*
	90	81.45%	82.55%	82.38%	20.01%*	19.08%*	19.18%*	29.51%*	28.22%*	28.35%*
	120	83.56%	83.41%	83.47%	18.33%*	18.12%*	18.03%*	27.22%*	26.89%*	26.75%*
	150	84.41%	84.59%	84.75%	17.21%*	17.00%*	16.91%*	25.60%*	25.28%*	25.15%*
	180	84.92%	84.78%	84.60%	17.49%*	17.01%*	17.14%*	26.09%*	25.27%*	25.44%*
	210	86.56%	86.50%	86.59%	15.71%*	15.29%*	15.15%*	23.48%*	22.81%*	22.60%*
Period 2	240	85.61%	85.34%	85.59%	16.01%*	15.72%*	15.45%*	23.79%*	23.35%*	22.94%*
	30	100.74%	101.19%	100.86%	-0.45%	-0.95%	-0.55%	-0.68%	-1.43%	-0.83%
	60	101.02%	101.18%	100.93%	-0.16%	-0.63%	-0.30%	-0.28%	-0.95%	-0.46%
	90	102.77%	102.35%	101.86%	-2.63%*	-2.22%*	-1.68%	-3.96%*	-3.31%*	-2.49%
	120	100.94%	101.48%	100.78%	-0.77%	-0.99%	-0.35%	-1.17%	-1.50%	-0.53%
	150	102.56%	101.95%	101.67%	-2.12%*	-1.53%	-1.05%	-3.19%*	-2.29%	-1.57%
	180	101.40%	100.77%	100.36%	-0.99%	-0.50%	-0.13%	-1.50%	-0.75%	-0.20%
Period 3	210	99.78%	99.87%	99.60%	0.47%	0.27%	0.50%	0.70%	0.42%	0.76%
	240	99.13%	98.30%	98.25%	1.27%	2.18%*	2.25%*	1.93%	3.27%*	3.40%*
	30	93.73%	93.47%	93.66%	7.44%*	7.24%*	7.05%*	11.15%*	10.86%*	10.57%*
	60	93.53%	93.30%	93.49%	7.75%*	8.09%*	7.97%*	11.62%*	12.12%*	11.93%*
	90	93.87%	93.15%	93.20%	7.25%*	7.90%*	7.85%*	10.84%*	11.80%*	11.73%*
	120	91.26%	91.68%	91.80%	10.08%*	9.29%*	9.16%*	15.08%*	13.88%*	13.70%*
	150	91.29%	90.44%	90.59%	9.92%*	10.42%*	10.32%*	14.80%*	15.55%*	15.41%*
Period 4	180	92.33%	92.89%	93.09%	8.19%*	7.90%*	7.79%*	12.23%*	11.81%*	11.64%*
	210	93.35%	93.41%	93.30%	7.06%*	7.44%*	7.52%*	10.53%*	11.13%*	11.24%*
	240	92.61%	92.70%	92.47%	8.52%*	8.64%*	8.80%*	12.75%*	12.96%*	13.21%*
	30	95.52%	94.57%	94.45%	5.53%*	6.08%*	6.24%*	8.27%*	9.09%*	9.33%*
	60	97.07%	97.47%	97.27%	3.57%*	3.60%*	3.77%*	5.38%*	5.40%*	5.66%*
	90	97.57%	96.13%	95.92%	3.30%*	4.50%*	4.69%*	4.95%*	6.76%*	7.02%*
	120	96.34%	96.05%	96.14%	5.07%*	5.05%*	4.98%*	7.63%*	7.56%*	7.48%*
Period 5	150	95.80%	96.49%	96.35%	4.75%*	4.54%*	4.75%*	7.11%*	6.82%*	7.13%*
	180	98.19%	99.26%	99.05%	2.00%*	1.15%	1.43%	2.99%*	1.70%	2.12%
	210	97.86%	98.13%	98.05%	2.93%*	2.77%*	2.99%*	4.40%*	4.17%*	4.49%*
	240	97.04%	97.40%	96.91%	3.79%*	3.46%*	3.86%*	5.67%*	5.20%*	5.81%*
	30	88.40%	86.91%	86.58%	13.40%*	14.85%*	15.23%*	19.97%*	22.07%*	22.63%*
	60	90.75%	89.91%	89.13%	10.89%*	11.70%*	12.55%*	16.36%*	17.51%*	18.75%*
	90	89.73%	88.27%	87.08%	11.50%*	13.26%*	14.58%*	17.14%*	19.78%*	21.73%*
Period 5	120	91.51%	89.71%	88.65%	9.91%*	11.88%*	13.08%*	14.82%*	17.80%*	19.55%*
	150	91.76%	89.98%	88.85%	9.92%*	11.83%*	13.06%*	14.90%*	17.71%*	19.51%*
	180	92.84%	91.24%	90.09%	8.48%*	10.10%*	11.56%*	12.70%*	15.09%*	17.29%*
	210	89.93%	88.73%	87.48%	11.30%*	12.95%*	14.20%*	16.94%*	19.34%*	21.21%*
	240	90.47%	89.08%	87.78%	10.96%*	11.88%*	13.10%*	16.39%*	17.79%*	19.57%*

\* Significantly different from zero with level 0.05

Table 4.18: Stability Ratios, and Kendall and Spearman Rank Correlation Coefficients (ETQ)

The stability ratios for the ETQ are very high for all of the periods. This is an indication that the values of the traditional measures in the estimation period are weak predictors of the ETQ in the investment period. The worst situation is observed in period 2. The stability ratio is close to 100% and the Kendall and Spearman rank correlation coefficients are not significantly different from zero. Hence, it is not possible to predict the ETQ in the investment period based on the tracking quality (as measured by the traditional measures) in the estimation period. The best situation is observed in period 1, for which the relative stability indicators do show that the ETQ is predictable.

To analyse the predictability of the ETQ further, we select the best 10% of the portfolios, according to the values of the traditional measures in the estimation period. The results are reported in Table 4.19.

		MAD		MSE		TEV	
Size		Med	IQR	Med	IQR	Med	IQR
Period 1	30	0.9097	0.8627	0.9201	0.8874	0.9140	0.8929
	60	0.7733	0.6835	0.7452	0.6681	0.7552	0.7073
	90	0.6781	0.6765	0.6406	0.5944	0.6389	0.5789
	120	0.6183	0.5370	0.5928	0.5354	0.5979	0.5459
	150	0.6221	0.5041	0.5601	0.4598	0.5639	0.4588
	180	0.5409	0.4599	0.5050	0.4861	0.5000	0.4768
	210	0.5323	0.4552	0.5197	0.4073	0.5272	0.4209
	240	0.4981	0.4259	0.4991	0.4153	0.5054	0.4413
Period 2	30	1.0932	1.0809	1.1117	1.0814	1.1057	1.0849
	60	0.8085	0.6851	0.8170	0.7761	0.8029	0.7890
	90	0.7784	0.7328	0.7889	0.7031	0.7558	0.6978
	120	0.7412	0.6237	0.6884	0.6200	0.6584	0.6254
	150	0.6912	0.5752	0.6351	0.5527	0.6225	0.5325
	180	0.6206	0.5076	0.6234	0.5430	0.5906	0.5337
	210	0.5927	0.5155	0.5840	0.5576	0.5522	0.5165
	240	0.5408	0.4968	0.5159	0.5135	0.4994	0.4772
Period 3	30	1.1749	0.9931	1.1721	0.9561	1.1892	0.9312
	60	0.9358	0.6950	0.9010	0.7014	0.8903	0.6716
	90	0.8225	0.6461	0.7642	0.6463	0.7635	0.6324
	120	0.7129	0.4943	0.6766	0.4879	0.6717	0.4634
	150	0.6579	0.5523	0.6441	0.4768	0.6422	0.4820
	180	0.6410	0.5323	0.6071	0.5264	0.6222	0.5392
	210	0.6030	0.4446	0.5607	0.4247	0.5607	0.4228
	240	0.5778	0.4063	0.5404	0.3873	0.5438	0.4165
Period 4	30	1.6374	1.5971	1.6275	1.4787	1.6032	1.4755
	60	1.4476	1.1870	1.3684	1.1987	1.3513	1.1486
	90	1.3767	1.0123	1.2970	1.0684	1.2570	1.0582
	120	1.1933	0.9735	1.1679	1.0136	1.1604	0.9650
	150	1.1347	1.0033	1.1293	0.9415	1.0810	0.9213
	180	1.2308	1.0325	1.1110	0.9805	1.0417	0.9610
	210	1.1011	0.9656	1.0715	0.8481	0.9873	0.8433
	240	1.0566	0.8417	1.0390	0.8273	0.9796	0.7755
Period 5	30	2.7488	2.2807	2.7800	2.3151	2.5924	2.2022
	60	2.7724	2.3527	2.6564	2.3185	2.3474	2.1074
	90	2.7313	2.3079	2.6112	2.2551	2.1319	1.9412
	120	2.7984	1.9497	2.7433	1.9523	2.1192	1.7741
	150	2.9410	2.1137	2.8017	1.9692	2.0209	1.6536
	180	2.9964	1.8413	2.8710	1.7325	2.0986	1.5519
	210	3.1258	1.8908	2.9337	1.6833	2.0701	1.4777
	240	3.1384	1.8489	3.1955	1.7534	2.2543	1.6185

Table 4.19: Comparison of Levels of Tracking Quality (ETQ, Best 10%)

To see how the selection of the best 10% of the portfolios from the original 5,000 influences the ETQ, we calculate the ratios of the medians in Table 4.17 to the medians in Table 4.19. These ratios are reported in Table 4.20.

Size		MAD	MSE	TEV
Period 1	30	1.52	1.50	1.52
	60	1.42	1.47	1.46
	90	1.40	1.39	1.41
	120	1.37	1.36	1.37
	150	1.27	1.35	1.36
	180	1.41	1.39	1.43
	210	1.32	1.30	1.31
	240	1.38	1.27	1.28
Period 2	30	0.97	0.95	0.95
	60	1.02	1.00	1.00
	90	0.97	0.95	0.96
	120	0.95	0.99	0.99
	150	0.94	1.01	1.00
	180	1.02	0.99	0.99
	210	1.00	1.02	1.01
	240	1.09	1.13	1.07
Period 3	30	1.13	1.10	1.08
	60	1.14	1.14	1.15
	90	1.12	1.16	1.16
	120	1.14	1.17	1.18
	150	1.15	1.14	1.14
	180	1.12	1.13	1.11
	210	1.12	1.16	1.16
	240	1.13	1.16	1.16
Period 4	30	1.09	1.08	1.09
	60	1.02	1.06	1.05
	90	1.02	1.03	1.03
	120	1.08	1.06	1.03
	150	1.08	1.04	1.04
	180	0.96	1.01	1.03
	210	1.05	1.03	1.05
	240	1.07	1.04	1.04
Period 5	30	1.30	1.27	1.29
	60	1.20	1.22	1.25
	90	1.20	1.22	1.27
	120	1.16	1.18	1.25
	150	1.10	1.14	1.24
	180	1.10	1.13	1.18
	210	1.08	1.13	1.20
	240	1.11	1.07	1.12

Table 4.20: Ratios of Tracking Qualities (ETQ, 5,000/Best 10%)

The selection of the best 10% of the tracking portfolios based on the estimation period (see Table 4.20) generally improves the tracking quality: the values of the ratios are large than unity. Period 2 is an exception. Here, in some cases, the median values of the tracking quality measures over the top 10% are larger than those across all 5,000 tracking portfolios.

To verify our findings graphically, we draw scatter plots for the tracking portfolios of size 120, shown in Figure 4.6.

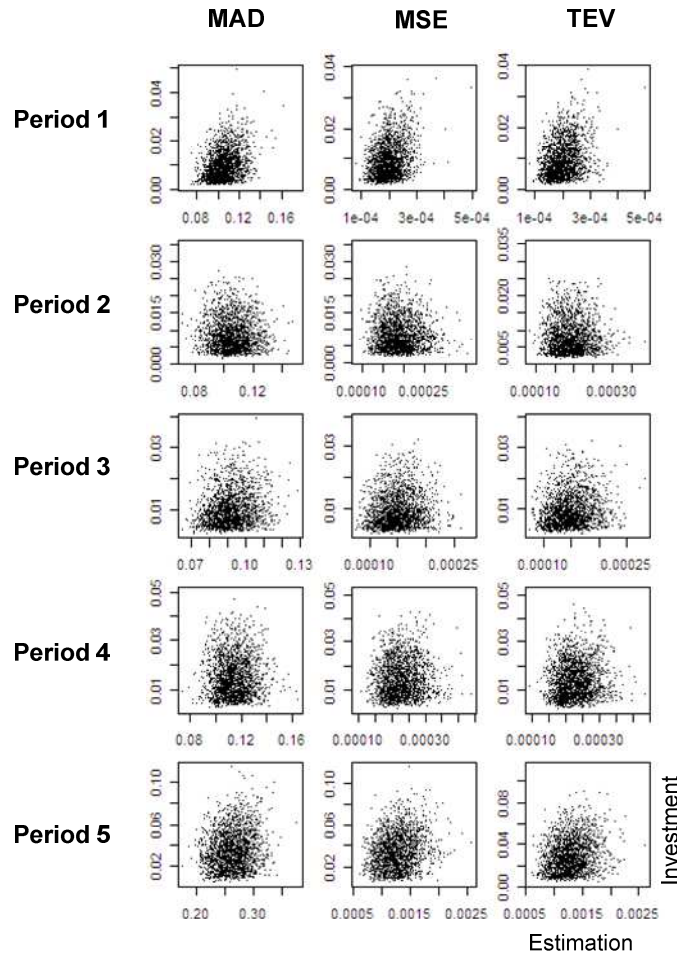


Figure 4.6: Comparing Traditional Measures with ETQs

This figure shows that there is little dependence between the ETQ measure (y-axis) in the investment period and the traditional measures of tracking quality (x-axis) in the estimation period.

### 4.2.3 Conclusion

The purpose of Section 4.2 was to analyse the stability of the performance of traditional measures of tracking quality when they are used as the objective in sampling. In the empirical part, we produced several findings.

Our analyses show that the use of traditional measures in sampling not only lead to weaknesses in absolute stability but also to poor relative stability on average. Independent of the stock market phase, the relative tracking quality in the estimation period compared to alternative tracking portfolios is not a good predictor of the relative tracking quality in the investment period. Thus, those tracking portfolios with the best tracking quality in the estimation period might not necessarily be among the best in the investment period and vice versa.

Furthermore, adding extra stocks to the tracking portfolio, in general, deteriorates the relative tracking quality for the S&P1500 index. As a consequence, the commonly used sampling procedure of searching for the tracking portfolio with the minimal value of a given tracking quality measure solely based on past data does not lead to the desired result, regardless of which measure is used in the optimization process.

All three measures show very poor results when used as predictors for the ETQ measure we defined earlier. On average, there is just a marginal coherence in the results. The selection of the best tracking portfolios in the estimation period may improve the tracking quality, but not always; sometimes there are counterintuitive results, when the ETQ decreases with increasing portfolio size, or stays the same for all portfolio sizes.

This leads to the following question for further research: can an alternative measure be found with a better ability to predict future tracking quality and therefore a better ability to avoid overfitting.



## 4.3 Resistance of the tracking portfolio estimation

### 4.3.1 Research method

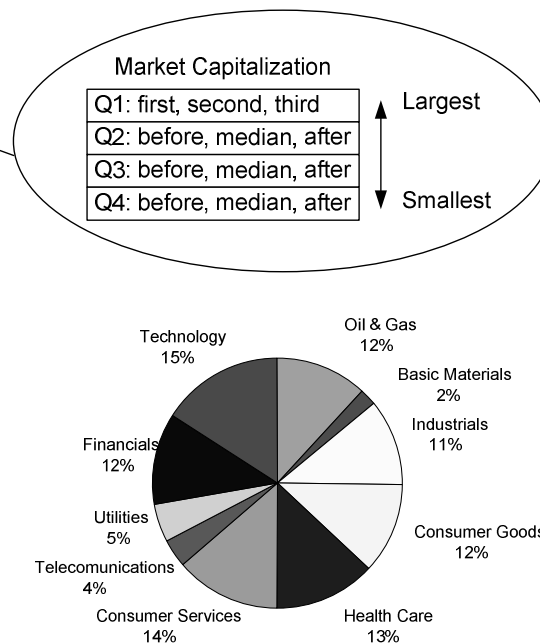
In this section, we analyse the resistance of the tracking portfolio estimation (using optimized sampling) to the use of different estimation sets. The empirical results are applied to the S&P1500 index.

We select five time periods, the same ones as in Section 4.1.1. For each time period, we construct a tracking portfolio using stratification, which is based on the procedure in *Schoenfeld and Maeda*<sup>211</sup> (see Figure 4.7). This procedure consists of three steps. In the first step, the index universe is divided into ten industrial sectors. In the second step, the stocks are sorted within each sector into descending order of market capitalization, and assigned to four quartiles. In the third step, twelve stocks from each sector are selected. The first, second, and third highest stocks in terms of market capitalization are selected from the first quartile. The median stock, the stock before the median and the stock after the median are selected from the remaining three quartiles. Hence, the tracking portfolio includes 120 stocks, covers all sectors, and contains small, medium and large capitalization stocks.

#### S&P1500\*

Industry	Companies
0001 Oil & Gas	78
1000 Basic Materials	68
2000 Industrials	296
3000 Consumer Goods	151
4000 Health Care	149
5000 Consumer Services	217
6000 Telecommunications	17
7000 Utilities	71
8000 Financials	258
9000 Technology	195

\*02.03.2009



**Tracking portfolio:** 10 Sectors \* 12 Stocks = 120 Stocks

Figure 4.7: Stock Selection for the Tracking Portfolio

<sup>211</sup> See Schoenfeld/Maeda (2004, p. 372).

To analyse the resistance of the tracking portfolio estimation to changes in the estimation datasets, for each time period we generate 1,000 random samples using 75% of the data from the estimation period (see Figure 4.8). Then, we estimate the tracking portfolios weights using these samples and evaluate the estimated tracking portfolios using the data in the investment period.

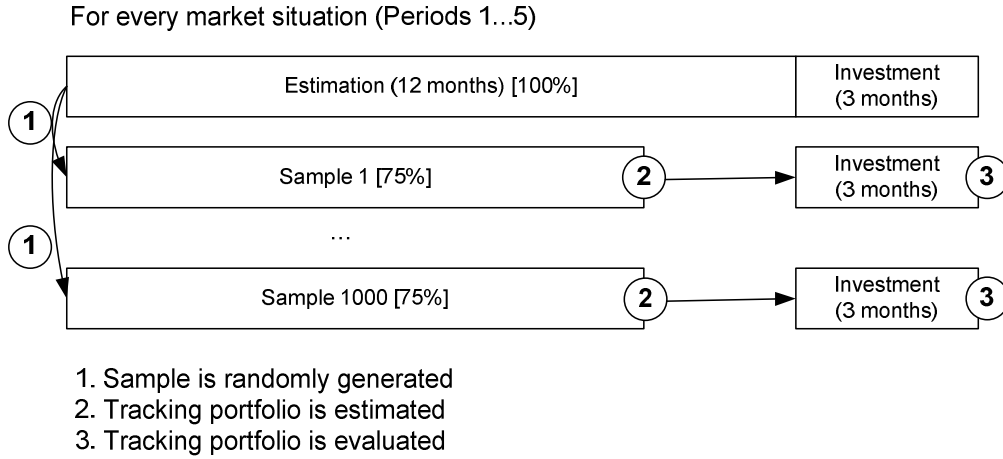


Figure 4.8: Design of the Experiment

The estimation of the weights of the tracking portfolios is done using the optimization procedure from Section 3.3. We test the resistance of the tracking portfolio estimation using the minimization of the traditional measures of tracking quality (MAD, MSE, and TEV). Additionally, we analyse the estimation by minimizing the robust measures MAXD and CVAR, which were described in Section 3.3.2 and Section 3.3.3 respectively.

For the CVAR method, it is necessary to pre-select the value of the parameter  $\alpha$ . We select two values for this parameter:  $\alpha = 0.95$  and  $\alpha = 0.5$ . The parameter  $\alpha = 0.95$  controls 5% of the largest deviations from the index.<sup>212</sup> In other words, using this parameter in the objective, the method pays attention only to the largest deviations from the index. The parameter  $\alpha = 0.5$  controls the largest 50% of deviations and provides a trade-off between approximation and robustness.<sup>213</sup> We call the method with  $\alpha = 0.95$  CVAR and that with  $\alpha = 0.5$  CVAR2.

Two versions of each measure are used. In the first version, the measure of tracking quality is based on the tracking error without the term  $b$ ; in the second version, term  $b$  is minimized together with the objective of the optimization problem.<sup>214</sup>

<sup>212</sup> See Rockafellar/Uryasev (2002, p. 1463). For their empirical results, they use  $\alpha = 0.1$ .

<sup>213</sup> See Mattera/Haykin (1998, p. 226).

<sup>214</sup> See Section 3.3.1 and equation (3.6) for details.

### 4.3.2 Empirical results

The empirical results for all measures are reported in Table 4.21.

	Method	Estimation, RMSE%				Investment, ETQ%			
		Med	IQR	Min	Max	Med	IQR	Min	Max
Period 1	TEV	0.0673	0.0040	0.0573	0.0757	0.3538	0.0925	0.1983	0.7320
	MSE	0.0672	0.0041	0.0573	0.0757	0.3528	0.0956	0.1984	0.7232
	MSE (b)	0.0672	0.0041	0.0573	0.0757	0.3521	0.0952	0.1983	0.7323
	MAD	0.0784	0.0057	0.0650	0.0899	0.4238	0.1872	0.1912	1.0624
	MAD (b)	0.0784	0.0057	0.0650	0.0899	0.4238	0.1884	0.1912	1.0624
	MAXD	0.0805	0.0049	0.0671	0.0896	0.4708	0.2305	0.2296	1.3496
	MAXD (b)	0.0804	0.0049	0.0673	0.0893	0.4710	0.2270	0.2268	1.2121
	CVAR	0.0805	0.0049	0.0671	0.0896	0.4705	0.2298	0.2294	1.3496
	CVAR (b)	0.0804	0.0049	0.0673	0.0893	0.4710	0.2270	0.2268	1.2128
	CVAR2	0.0722	0.0049	0.0612	0.0821	0.3948	0.1232	0.1856	0.9298
	CVAR2 (b)	0.0723	0.0048	0.0612	0.0816	0.3901	0.1224	0.1946	0.9441
Period 2	TEV	0.0738	0.0041	0.0620	0.0831	0.2873	0.1641	0.1363	0.8132
	MSE	0.0737	0.0041	0.0620	0.0829	0.2907	0.1689	0.1366	0.8364
	MSE (b)	0.0736	0.0041	0.0620	0.0827	0.2963	0.1735	0.1363	0.8772
	MAD	0.0850	0.0053	0.0704	0.0982	0.3995	0.2626	0.1514	1.1180
	MAD (b)	0.0850	0.0054	0.0698	0.1005	0.3931	0.2579	0.1527	1.1180
	MAXD	0.0878	0.0055	0.0751	0.0987	0.2704	0.1640	0.1337	1.0092
	MAXD (b)	0.0881	0.0057	0.0758	0.0984	0.2725	0.1648	0.1343	0.9763
	CVAR	0.0878	0.0055	0.0751	0.0988	0.2702	0.1648	0.1325	1.0091
	CVAR (b)	0.0881	0.0057	0.0758	0.0984	0.2725	0.1669	0.1343	0.9661
	CVAR2	0.0785	0.0046	0.0665	0.0875	0.3326	0.2066	0.1384	0.9836
	CVAR2 (b)	0.0787	0.0046	0.0665	0.0884	0.3285	0.2060	0.1396	1.0345
Period 3	TEV	0.0673	0.0038	0.0577	0.0758	0.4958	0.2673	0.1757	1.4797
	MSE	0.0671	0.0039	0.0577	0.0754	0.5207	0.2726	0.2039	1.5694
	MSE (b)	0.0671	0.0039	0.0576	0.0754	0.5156	0.2739	0.2041	1.5694
	MAD	0.0780	0.0049	0.0644	0.0900	0.5405	0.3146	0.2356	1.8451
	MAD (b)	0.0780	0.0050	0.0645	0.0900	0.5270	0.3067	0.2336	1.8451
	MAXD	0.0804	0.0049	0.0685	0.0907	0.4830	0.2244	0.2288	1.3859
	MAXD (b)	0.0805	0.0046	0.0693	0.0895	0.4611	0.2182	0.2169	1.7659
	CVAR	0.0804	0.0049	0.0685	0.0907	0.4833	0.2248	0.2288	1.3855
	CVAR (b)	0.0805	0.0046	0.0693	0.0895	0.4611	0.2189	0.2169	1.7659
	CVAR2	0.0720	0.0044	0.0610	0.0809	0.5764	0.3555	0.2180	1.5121
	CVAR2 (b)	0.0723	0.0046	0.0611	0.0816	0.5332	0.3354	0.1722	1.4747
Period 4	TEV	0.0772	0.0045	0.0680	0.0864	0.7446	0.3343	0.2455	1.9167
	MSE	0.0771	0.0046	0.0679	0.0862	0.7466	0.3311	0.2420	1.9198
	MSE (b)	0.0771	0.0045	0.0680	0.0862	0.7464	0.3339	0.2459	1.9201
	MAD	0.0879	0.0052	0.0744	0.1004	0.8426	0.4481	0.2610	2.0326
	MAD (b)	0.0880	0.0052	0.0743	0.1004	0.8499	0.4450	0.2610	2.0333
	MaxD	0.0932	0.0058	0.0781	0.1043	0.7332	0.4586	0.1995	2.1408
	MaxD (b)	0.0933	0.0059	0.0776	0.1054	0.7308	0.4393	0.1902	2.1263
	CVaR	0.0932	0.0058	0.0781	0.1041	0.7341	0.4586	0.1995	2.1410
	CVaR (b)	0.0933	0.0059	0.0776	0.1054	0.7284	0.4390	0.1902	2.1262
	CVaR2	0.0818	0.0050	0.0716	0.0917	0.8574	0.3686	0.2443	1.8401
	CVaR2 (b)	0.0821	0.0049	0.0716	0.0913	0.8584	0.3833	0.2410	1.8917
Period 5	TEV	0.1842	0.0145	0.1498	0.2126	4.2698	0.9283	2.3453	6.6341
	MSE	0.1819	0.0142	0.1495	0.2084	4.8509	1.0159	2.7509	7.2736
	MSE (b)	0.1818	0.0143	0.1495	0.2084	4.8423	1.0167	2.8589	7.2940
	MAD	0.2125	0.0175	0.1647	0.2448	4.3208	1.2320	1.9102	7.4818
	MAD (b)	0.2131	0.0183	0.1642	0.2450	4.2491	1.2307	1.9371	7.4817
	MaxD	0.2238	0.0226	0.1766	0.2537	5.7520	1.4763	2.6897	8.8691
	MaxD (b)	0.2217	0.0201	0.1761	0.2550	5.0440	1.3028	2.0924	8.3031
	CVaR	0.2228	0.0221	0.1766	0.2537	5.7432	1.4379	2.6897	8.8691
	CVaR (b)	0.2213	0.0201	0.1761	0.2550	5.0223	1.3170	2.0926	8.3030
	CVaR2	0.1981	0.0171	0.1572	0.2267	4.4370	1.0661	2.2879	7.4242
	CVaR2 (b)	0.2005	0.0169	0.1576	0.2278	3.7256	1.0166	1.6624	6.4132

Table 4.21: Resistance of the Tracking Portfolio Estimation

We report the RMSE for the estimation period and the ETQ for the investment period. To illustrate the results, we select the MSE (b) method for period 2. Using this method, the

median portfolio of the 1,000 portfolios has a tracking quality in the estimation period equal to 0.0736. The best and worst tracking portfolios have tracking qualities equal to 0.0620 and 0.0827 respectively. The IQR, the difference between the third quartile and the first quartile, which covers 500 portfolios (50% of 1,000), for this method, is equal to 0.0041. These numbers demonstrate that, despite the fact that the data points used for the estimation were randomly selected from the same time period, the quality of the tracking varies, even in the estimation period.

In the investment period, the results in Table 4.21 indicate that the difference between the qualities of the best and worst portfolios is large. This behaviour is independent of the measure used in the optimization problem. Figure 4.9 illustrates graphically the development of the tracking portfolios in the investment period, for the MSE (b) method for period 2, in comparison to the index.

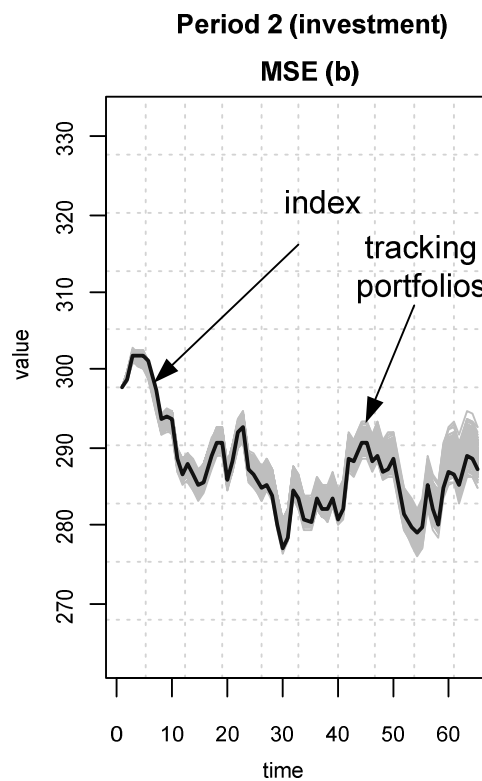


Figure 4.9: Example of the Sensitivity of the Tracking Portfolios

In this case, the minimal value of the ETQ is equal to 0.1363 and the maximal value is equal to 0.8772. Hence, the tracking portfolio estimation is not resistant to a modification of the data used in the estimation period. Similar analyses of the results for other measures and periods indicate that the same behaviour occurs.

### **4.3.3 Conclusion**

In this section, we have shown that even small modifications of the data in the estimation period can lead to significant differences in the tracking portfolio estimation and, as a consequence, to very different tracking qualities in the investment period. Therefore, there is a need for an approach that can handle this problem.

## 4.4 Summary

In this chapter, we have carried out an empirical analysis of optimized sampling methods.

First, we analysed the problem of modelling portfolio structures using weights or numbers, and showed empirically that, for a buy-and-hold strategy, it is necessary to correct the portfolio weights in order to be consistent with holding fixed numbers of assets. This improves the tracking quality.

Then, we showed empirically that tracking portfolios that are constructed by minimizing the traditional measures of tracking quality lack absolute and relative stability of tracking quality.

Finally, we showed empirically that tracking portfolio estimation using the current sampling methods is not resistant to small changes in the dataset used for estimation. We showed that slight modifications of the data in the estimation period leads to different tracking portfolios being generated and, as a consequence, to different levels of tracking quality in the investment period.

In this chapter, we have empirically identified some problems with the traditional methods of optimized sampling. In the next chapter, we develop a new approach that overcomes the shortcomings of the existing methods.

## **5 A method for robust index tracking**

This chapter focuses on the development of new robust index tracking approaches. First, we make an overview of statistical learning theory. Then, we describe support vector techniques, which are based on this theory, for solving classification and regression problems. Then, using support vector techniques and principles of statistical learning theory, we develop the methods for solving index tracking problem and compare them with the existing approaches. Finally, we analyse the methods empirically.

## 5.1 Statistical learning theory

### 5.1.1 Overview of statistical learning theory<sup>215</sup>

Statistical learning theory (SLT) was developed by *Vapnik and Chervonenskis* to solve the problem of the estimation of the functional dependence (model) between variables from a limited number of available data points.<sup>216</sup> The aim of SLT is to build learning algorithms for estimating models with high generalizability from a given set of data. Due to the problem of overfitting (see Section 5.1.4.1 for details), the main goal is to maximize the prediction quality and not to fit the data that is used for the model estimation as well as possible. This is equivalent to minimizing the risk that the model will have weak predictive ability.

One method SLT provides for avoiding overfitting is to constrain the set of admissible models.<sup>217</sup> The aim is to filter out those models that may lead to overfitting. This can be done by incorporating additional information into the estimation process. Hence, generalizability is achieved by incorporating both data and knowledge.

The size of the set of admissible models is called the capacity (see Section 5.1.4.1 for details).<sup>218</sup> This capacity determines the model's complexity. High capacity equates to high model complexity and vice versa. The key finding of SLT is that the generalizability, and therefore the prediction quality, of an estimated model depends on its complexity.<sup>219</sup> The capacity of the model set can be measured direct from the data (see Section 5.1.4.4 for details).

In SLT, one approach used to improve the generalizability of a model is called structural risk minimization (SRM) (see Section 5.1.4.2 for details).<sup>220</sup> The idea is to choose a sequence of sets of models of increasing capacity and to optimize within each model set. Finally, from this sequence, the model with the lowest prediction risk is chosen. The prediction risk (in SLT this is named the test error) can be measured using various methods (see Section 5.1.4.2 and Section 5.1.4.3 for details).

---

<sup>215</sup> This section is based on Rossbach/Karlow (2012).

<sup>216</sup> See Vapnik (1998).

<sup>217</sup> See Bousquet/Boucheron/Lugosi (2004).

<sup>218</sup> See Cherkassky/Mulier (2007).

<sup>219</sup> See Vapnik (1998, 2006).

<sup>220</sup> See Vapnik (1998, 2006).



SLT introduces the concept of consistency of learning, which we describe below in Section 5.1.2. It is proved, that the SRM learning principle is consistent.<sup>221</sup> Hence, the SRM principle is theoretically justified.

There are also learning algorithms, which implements SRM principle of SLT. They are support vector machines (SVMs) (see Section 5.1.3.1) and support vector regression (SVR) (see Section 5.1.3.2).

In the following sections we will concretize the building blocks of SLT. We start with the formulation of a general setting of learning problem. After that, we describe the concept of learning consistency. Then, we review support vector techniques for classification and regression problems (SVM and SVR). Then, we discuss in details bias-variance trade-off, model complexity, and the problem of overfitting. Model selection methods are then described. After that, we discuss existing approaches for quantifying and measuring the capacity of a set of models, which we then use for model selection. Then, we introduce an extended version of SVR. Finally, we show that a regression problem that is based on support vectors can be transformed into a classification problem.

## 5.1.2 General setting of a learning problem and learning consistency

There are two common types of learning problems. The first is the classification problem, where the goal is to classify the data points into two or more classes. The second is a regression problem, where the goal is to estimate a functional dependence between one or many input variables and one output variable. The general setting of a learning problem can be formulated as follows.<sup>222</sup> Assume that a set of independent and identically distributed data pairs,

$$\mathbf{z}_l := (y_l, \mathbf{x}_l), l = 1, \dots, L$$

is given, where  $y_l$  is the output scalar and  $\mathbf{x}_l$  is the input vector. These data pairs  $\mathbf{z}_l \in Z$  are characterized by an unknown probability measure  $P(\mathbf{z})$ , which is defined on the space  $Z$ . One needs to estimate the parameters  $\mathbf{w} \in W$  of the functional dependence  $f(\mathbf{x}, \mathbf{w})$ . To do so, it is necessary to define the loss function  $Q(\mathbf{z}_l, \mathbf{w})$ , which describes the discrepancy between  $y_l$  and  $f(\mathbf{x}_l, \mathbf{w})$ . This can be done in different ways. For example, one can use quadratic loss:

$$Q(\mathbf{z}_l, \mathbf{w}) = (y_l - f(\mathbf{x}_l, \mathbf{w}))^2$$

---

<sup>221</sup> See Vapnik (1998, p. 229).

<sup>222</sup> See Vapnik (1999).

or absolute loss:

$$Q(\mathbf{z}_l, \mathbf{w}) = |y_l - f(\mathbf{x}_l, \mathbf{w})|.$$

Hence, the goal of learning is to estimate parameters  $\mathbf{w}$  that give a minimum value for the error functional:

$$err(\mathbf{w}) = \int Q(\mathbf{z}, \mathbf{w}) dP(\mathbf{z}), \mathbf{w} \in W. \quad (5.1)$$

In SLT this functional is called the risk functional.

However, the direct estimation of parameters  $\mathbf{w}$  by minimizing (5.1) is impossible, since the probability measure  $P(\mathbf{z})$  is unknown. Therefore, one approximates the risk functional (5.1) via its empirical counterpart:

$$err_{emp}(\mathbf{w}) = \frac{1}{L} \sum_{l=1}^L Q(z_l, \mathbf{w}). \quad (5.2)$$

There are two conditions, under which the minimization of the approximation (5.2) of the actual risk (5.1) in order to estimate the parameters  $\mathbf{w}$  is valid. On the one hand, the actual error  $err(\mathbf{w}_L)$  for the estimated parameters  $\mathbf{w}_L$  should converge in probability to the minimal actual error  $\inf_{\mathbf{w}} err(\mathbf{w})$ . On the other hand, the empirical error  $err_{emp}(\mathbf{w}_L)$  should also converge to the minimal actual error as the size of the dataset used for the estimation increases. Formally, this can be written as follows:<sup>223</sup>

$$\begin{aligned} err(\mathbf{w}_L) &\xrightarrow[L \rightarrow \infty]{P} \inf_{\mathbf{w}} err(\mathbf{w}) \\ err_{emp}(\mathbf{w}_L) &\xrightarrow[L \rightarrow \infty]{P} \inf_{\mathbf{w}} err(\mathbf{w}) \end{aligned} \quad (5.3)$$

If there is such a convergence in probability with as the size of the data  $L$  increases, one says that the learning process is consistent. This is illustrated in Figure 5.1, where  $\inf_{\mathbf{w}} err(\mathbf{w})$  is the smallest possible actual error.

---

<sup>223</sup> See Vapnik (2000, p.36).

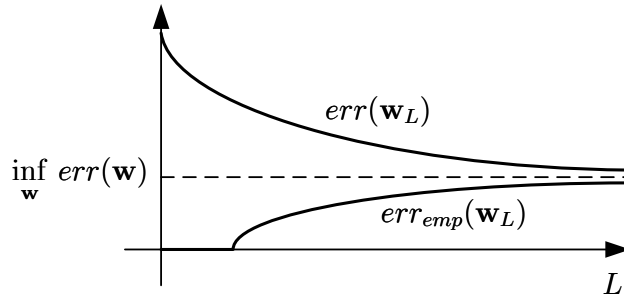


Figure 5.1: Consistency of the learning process<sup>224</sup>

One can test whether the learning is consistent, i.e. whether the convergence conditions (5.3) are satisfied, by using the key theorem of SLT, which was proved by *Vapnik and Chervonenkis*.

**Key Theorem of SLT (Vapnik and Chervonenkis).**<sup>225</sup> *Let  $Q(\mathbf{z}, \mathbf{w}), \mathbf{w} \in W$  be a set of functions that has bounded loss for the probability measure  $P(\mathbf{z})$ :*

$$A \leq \int Q(\mathbf{z}, \mathbf{w}) dP(\mathbf{z}) \leq B, \forall \mathbf{w} \in W.$$

*Then, for the learning from the data to be consistent, it is necessary and sufficient that the empirical risk  $err_{emp}(\mathbf{w})$  converges uniformly to the actual risk  $err(\mathbf{w})$  over the set  $Q(\mathbf{z}, \mathbf{w}), \mathbf{w} \in W$  as follows:*

$$\lim_{L \rightarrow \infty} Pr \left[ \sup_{\mathbf{w} \in W} (err(\mathbf{w}) - err_{emp}(\mathbf{w})) > \varepsilon \right] = 0, \forall \varepsilon.$$

As a result, consistency of learning is equivalent to the uniform convergence of the empirical error to the actual error. As will be shown further in Section 5.1.4, the speed of uniform convergence can be characterized by some function  $\Phi(L/h)$ , which depends on the size of the data  $L$  and some complexity measure  $h$ , as follows:

$$err(\mathbf{w}_L) - err_{emp}(\mathbf{w}_L) \leq \Phi(L/h). \quad (5.4)$$

As a result, the lower the value of the function  $\Phi(L/h)$  is, the smaller is the difference between the empirical and actual errors. In Section 5.1.4, we discuss the properties of learning algorithms that are important for consistency and for increasing the speed of the uniform convergence.

<sup>224</sup> See Vapnik (2000, p.36).

<sup>225</sup> See Vapnik/Chervonenkis (1989).

Moreover, inequality (5.4) can be rewritten as

$$err(\mathbf{w}_L) \leq err_{emp}(\mathbf{w}_L) + \Phi(L/h).$$

Hence, the actual risk for a given data sample of size  $L$  is bounded by the empirical error  $err_{emp}(\mathbf{w}_L)$  and some confidence interval  $\Phi(L/h)$ . To minimize the value of the guaranteed actual error  $err(\mathbf{w}_L)$ , it is necessary to minimize concurrently the empirical error  $err_{emp}(\mathbf{w}_L)$  and the confidence interval  $\Phi(L/h)$  that characterizes the speed of the uniform convergence.

To do this, Vapnik and his colleagues designed new learning algorithms SVM and SVR to solve the classification and regression problems. We discuss these algorithms further in Section 5.1.3.

### 5.1.3 Support vector (SV) techniques

#### 5.1.3.1 Support vector machines (SVMs)

##### 5.1.3.1.1 Separable case

A SVM is a supervised learning algorithm, used for the binary classification of data points.<sup>226</sup> The algorithm computes an optimal function, the so-called classification rule, which assigns class labels (scalars) to given data points (vectors). The classification rule is estimated by using data points  $\mathbf{x}_l$  with known class labels  $y_l$ . The class label takes the value -1 for the first class and +1 for the second class:

$$\begin{aligned} &(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_L, y_L) \\ &\mathbf{x}_l = (x_{1,l}, \dots, x_{N,l}) \\ &y_l \in \{-1, 1\}. \end{aligned}$$

The simplest classification rule is based on a linear function, where the weights vector  $\mathbf{w} = (w_1, \dots, w_N)$  and the constant  $b$  are estimated by a learning algorithm. The classification rule can be written as<sup>227</sup>

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^N w_i x_i + b\right), \quad (5.5)$$

where the sign function is defined as

---

<sup>226</sup> See Boser/Guyon/Vapnik (1992) and Vapnik (1995).

<sup>227</sup> See Vapnik (1995)

$$\text{sign}(a) = \begin{cases} 1 & \text{if } a > 0, \\ 0 & \text{if } a = 0, \\ -1 & \text{if } a < 0. \end{cases}$$

Thus, the classification rule assigns the class label -1 to the data point  $\mathbf{x}_l$  if  $\sum_{i=1}^N w_i x_{i,l} + b < 0$  and the class label +1 to the data point  $\mathbf{x}_l$  if  $\sum_{i=1}^N w_i x_{i,l} + b > 0$ . If  $\sum_{i=1}^N w_i x_{i,l} + b = 0$  then the membership of the data point  $\mathbf{x}_l$  is not defined. Hence, the classification rule (5.5) can be interpreted as a hyperplane, which passes between data points, separating them into two classes  $\{-1, 1\}$  in a Hilbert space (a space with any number of dimensions) (see Figure 5.2). However, more than one hyperplane can pass between the given data points. Therefore, it is necessary to compute the hyperplane that separates the data points optimally. *Boser et al.*<sup>228</sup> suggest selecting a hyperplane that separates two classes with a maximal margin. This unique<sup>229</sup> hyperplane is called the optimal separating hyperplane and can be found by solving the relevant optimization problem.

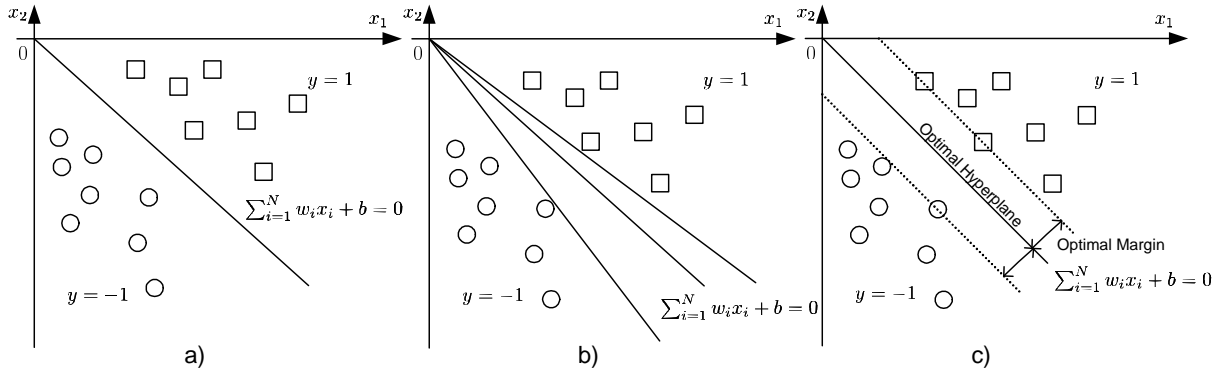


Figure 5.2: a) Separating Hyperplane b) Separating Hyperplane is not Unique c) Optimal Separating Hyperplane

Assuming that two classes are separated by the hyperplane  $\sum_{i=1}^N w_i x_i + b = 0$  correctly, the distance  $d_l$  between data vector  $\mathbf{x}_l$  and this hyperplane is equal to<sup>230</sup>

$$d_l = \frac{|\sum_{i=1}^N w_i x_{i,l} + b|}{\|\mathbf{w}\|_2} \geq 0, \quad (5.6)$$

where the 2-norm is defined as

$$\|\mathbf{w}\|_2 = \sqrt{\sum_{i=1}^N w_i^2}.$$

<sup>228</sup> See Boser/Guyon/Vapnik (1992).

<sup>229</sup> See Vapnik (1998, p. 402, Theorem 10.1).

<sup>230</sup> See Cherkassky/Mulier (2007, p. 419).

Multiplying the numerator in (5.6) by the class label  $y_l \in \{-1, 1\}$  allows us to omit the modulus:

$$d_l = \frac{y_l(\sum_{i=1}^N w_i x_{i,l} + b)}{\|\mathbf{w}\|_2} \geq 0. \quad (5.7)$$

Assuming that the distance  $d_l$  between vector  $\mathbf{x}_l$  and the separating hyperplane is at least  $\gamma$  for all  $l = 1, \dots, L$ , we can rewrite equation (5.7) as follows:

$$d_l = \frac{y_l(\sum_{i=1}^N w_i x_{i,l} + b)}{\|\mathbf{w}\|_2} \geq \gamma.$$

Thus, the optimal separating hyperplane can be found by fixing the scaling factor  $\|\mathbf{w}\|_2 = 1$  and maximizing the margin  $\gamma$

$$\begin{aligned} & \underset{\mathbf{w}, b}{\text{maximize}} && \gamma \\ & \text{subject to} && y_l \left( \sum_{i=1}^N w_i x_{i,l} + b \right) \geq \gamma \\ & && \|\mathbf{w}\|_2 = 1 \\ & && l = 1, \dots, L. \end{aligned} \quad (5.8)$$

To solve the optimization problem (5.8), there is a need to reformulate it as a quadratic programming problem in the following way:

Instead of fixing  $\|\mathbf{w}\|_2 = 1$ , we fix

$$\gamma \|\mathbf{w}\|_2 = 1.$$

It follows that the margin can be found as

$$\gamma = \frac{1}{\|\mathbf{w}\|_2}.$$

Therefore, the minimization of  $\|\mathbf{w}\|_2$  is equivalent to the maximization of the margin  $\gamma$ , since the smaller is the 2-norm of the weights  $\|\mathbf{w}\|_2$ , the larger is the margin  $\gamma$ . As a result, the optimization problem (5.8) can be reformulated as

$$\begin{aligned}
& \underset{\mathbf{w}, b}{\text{minimize}} && \frac{1}{2} \|\mathbf{w}\|_2^2 \\
& \text{subject to} && y_l \left( \sum_{i=1}^N w_i x_{i,l} + b \right) \geq 1 \\
& && l = 1, \dots, L.
\end{aligned} \tag{5.9}$$

Solving this optimization problem, we can find an optimal separating hyperplane that separates the data points into two classes with a maximal margin.

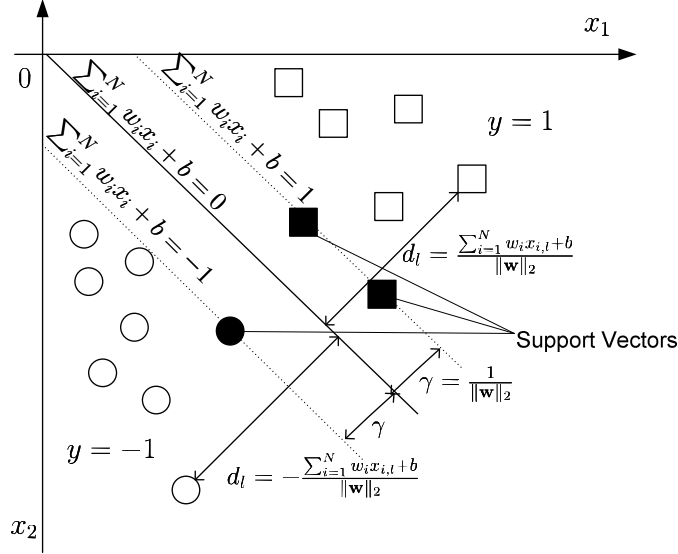


Figure 5.3: Optimal Separating Hyperplane and Support Vectors

The vectors,  $\mathbf{x}_l$ , that lie on the hyperplanes  $\sum_{i=1}^N w_i x_i + b = 1$  and  $\sum_{i=1}^N w_i x_i + b = -1$  are called support vectors (see Figure 5.3). The support vectors determine the position of the separating hyperplane.

#### 5.1.3.1.2 Non-separable case

In some cases, it is impossible to find the optimal separating hyperplane by means of the optimization problem described above since it is not possible to separate data points into classes without errors using a linear function. For that reason, there is another SVM formulation that allows errors  $\xi_l$  (see Figure 5.4). The optimization problem (5.9) is accordingly modified as follows, allowing for errors in separation:<sup>231</sup>

<sup>231</sup> See Cortes/Vapnik (1995, p. 280).

$$\begin{aligned}
& \underset{\mathbf{w}, b, \xi}{\text{minimize}} && \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{l=1}^L \xi_l \\
& \text{subject to} && y_l \left( \sum_{i=1}^N w_i x_{i,l} + b \right) \geq 1 - \xi_l \\
& && \xi_l \geq 0 \\
& && l = 1, \dots, L.
\end{aligned} \tag{5.10}$$

The new parameter  $C$  is introduced to penalize errors. Depending on the value of the parameter  $C$ , the influence of the errors on the position of the separating hyperplane changes. A smaller  $C$  gives a smaller penalization and, as a consequence, a smaller impact from errors on the solution of the SVM. For  $C = 0$  we have a unique solution  $\mathbf{w}' = (0, \dots, 0)$ . If the two classes are linearly separable and  $C$  is large enough (errors are heavily penalized), then the solutions of (5.9) and (5.10) will coincide.

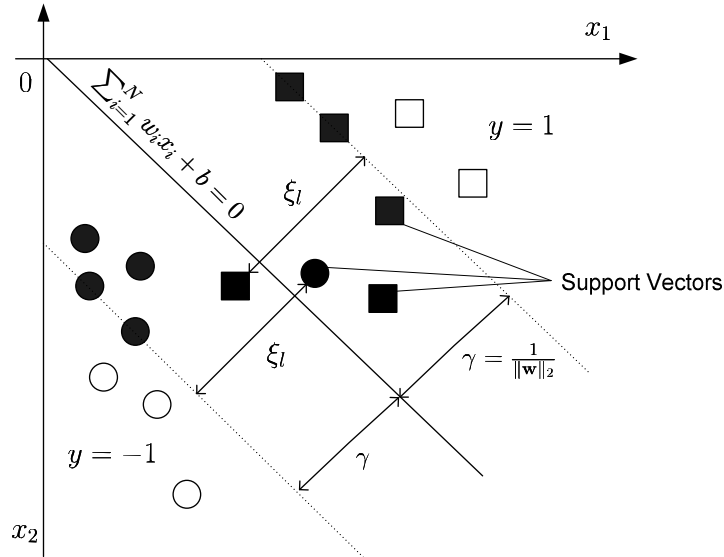


Figure 5.4: SVM for Non-Separable Case

The solution of the optimization problem (5.10) is determined by a saddle point of the Lagrange functional<sup>232</sup>

$$L = \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{l=1}^L \xi_l - \sum_{l=1}^L \alpha_l [y_l (\sum_{i=1}^N w_i x_{i,l} + b) - 1 + \xi_l] - \sum_{l=1}^L \beta_l \xi_l, \tag{5.11}$$

where  $\alpha_l \geq 0, \beta_l \geq 0$  are Lagrange multipliers.

The saddle point of the Lagrange functional fulfils the Karush-Kuhn-Taker (KKT) conditions

<sup>232</sup> See Vapnik/Chapelle (2000, p. 2016).



$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial w_i} = w_i - \sum_{l=1}^L \alpha_l y_l x_{i,l} = 0, \quad (5.12)$$

$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial \xi_l} = C - \alpha_l - \beta_l = 0, \quad (5.13)$$

$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial b} = \sum_{l=1}^L \alpha_l y_l = 0,$$

together with the complementary conditions

$$\alpha_l [y_l (\sum_{i=1}^N w_i x_{i,l} + b) - 1 + \xi_l] = 0, \quad (5.14)$$

$$\beta_l \xi_l = 0.$$

From the complementary KKT condition (5.14), the conclusion can be drawn that the data points  $\mathbf{x}_l$ , where the conditions of  $\alpha_l \neq 0$  and accordingly  $y_l (\sum_{i=1}^N w_i x_{i,l} + b) - 1 + \xi_l = 0$  are fulfilled, are support vectors.

From (5.12) it follows that the parameters ( $w_i$ ) of the model can be expressed as a weighted sum of the data points

$$w_i = \sum_{l=1}^L \alpha_l y_l x_{i,l}. \quad (5.15)$$

This feature means that, the greater is the importance of the data point, which is also a support vector, the greater is its influence on the parameters of the model.

From (5.13),  $C$  can be expressed as

$$C = \alpha_l + \beta_l \quad (5.16)$$

and, using the fact that the Lagrange multipliers ( $\alpha_l$  and  $\beta_l$ ) are positive, it follows that every weight ( $\alpha_l$ ) for a given data point in equation (5.15), is less than  $C$ :

$$0 \leq \alpha_l \leq C.$$

In other words, the parameter  $C$  limits the weights ( $\alpha_l$ ) to a given range, thereby limiting the impact of support vectors on the model parameters.

Substituting (5.15) and (5.16) into (5.11) gives a formulation of the dual optimization problem:

$$\begin{aligned}
 & \underset{\alpha}{\text{maximize}} && \sum_{l=1}^L \alpha_l - \frac{1}{2} \sum_{l=1}^L \sum_{k=1}^L \alpha_l \alpha_k y_l y_k \sum_{i=1}^N x_{i,l} x_{i,k} \\
 & \text{subject to} && 0 \leq \alpha_l \leq C \\
 & && \sum_{l=1}^L \alpha_l y_l = 0 \\
 & && l = 1, \dots, L, i = 1, \dots, N.
 \end{aligned} \tag{5.17}$$

Optimization problem (5.17) can also be used for the non-linear case.

### 5.1.3.1.3 Kernel SVM

In the previous sections, SVM, which separates data points into two classes using linear functions, was introduced. In practical problems, the linear separability of classes does not always exist. For such cases, SVM constructs a non-linear separation, by transforming the input data space into the feature space using some function  $\phi(\mathbf{x})$ , and constructing a separating hyperplane in the feature space (see Figure 5.5). Hence, the linear separation in the feature space gives non-linear separation in the input space.<sup>233</sup>

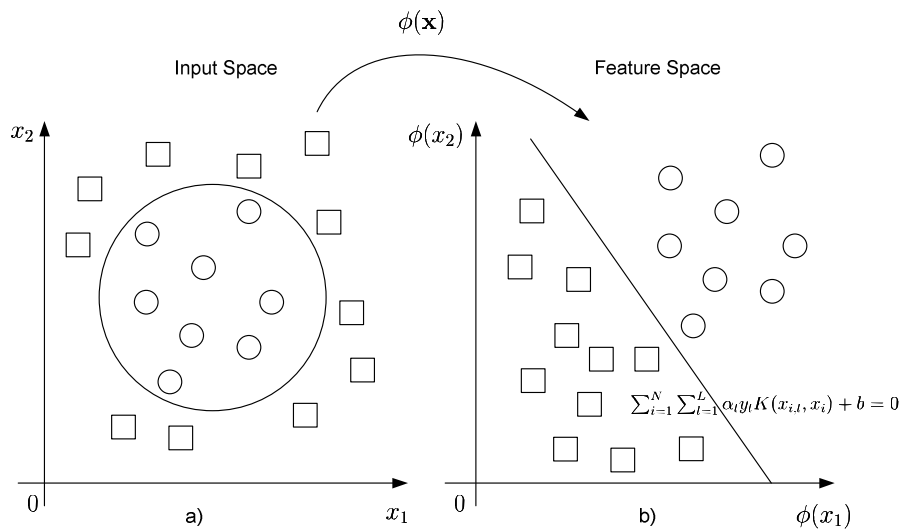


Figure 5.5: a) Non-linear Separating in Input Space b) Linear Separating in Feature Space

<sup>233</sup> See Cherkassky/Mulier (2007, p. 427).

Solving problem (5.17) with modified data allows us to find the separating hyperplane in the feature space. To do this, however, it is necessary to apply the transformation given by function  $\phi(\mathbf{x})$  to all data points, which is computationally intensive.

To solve this problem, we can apply the transformation to the inner product

$$(\mathbf{x}_l \cdot \mathbf{x}_k) = \sum_{i=1}^N x_{i,l} x_{i,k},$$

finding kernel function  $K(\mathbf{x}_l, \mathbf{x}_k)$ , such that

$$(\phi(\mathbf{x}_l) \cdot \phi(\mathbf{x}_k)) = \sum_{i=1}^N \phi(x_{i,l}) \phi(x_{i,k}) = K(\mathbf{x}_l, \mathbf{x}_k).$$

Thus, the optimal separating hyperplane for the feature space can be found by modifying the dual problem (5.17) such that<sup>234</sup>

$$\begin{aligned} & \underset{\alpha}{\text{maximize}} && \sum_{l=1}^L \alpha_l - \frac{1}{2} \sum_{l=1}^L \sum_{k=1}^L \alpha_l \alpha_k y_l y_k K(\mathbf{x}_l, \mathbf{x}_k) \\ & \text{subject to} && 0 \leq \alpha_l \leq C \\ & && \sum_{l=1}^L \alpha_l y_l = 0 \\ & && l = 1, \dots, L, i = 1, \dots, N. \end{aligned}$$

This modification is known as the “kernel trick”.<sup>235</sup>

The classification rule can also be represented using the kernel function:

$$f(\mathbf{x}) = \text{sign}\left(\sum_{l=1}^L \alpha_l y_l K(\mathbf{x}_l, \mathbf{x}) + b\right).$$

As a consequence, if the kernel function is specified, it is not necessary to know the form of function  $\phi$  to find the optimal separating hyperplane in the feature space.

For the function to be a kernel, and to define the inner product in the feature space, it must be a symmetric function that satisfies the Mercer condition.<sup>236</sup>

---

<sup>234</sup> See Schölkopf/Smola (2002, p. 15).

<sup>235</sup> See Schölkopf/Smola (2002, p. 15).

<sup>236</sup> See Vapnik (2000, p. 140).

**Theorem (Mercer).** *To guarantee that the symmetric function  $K(u, v)$  from  $L_2$  has an expansion*

$$K(u, v) = \sum_{k=1}^{\infty} c_k \phi_k(u) \phi_k(v)$$

*with positive coefficients  $c_k$  (i.e.,  $K(u, v)$  describes an inner product on some feature space), it is necessary and sufficient that the condition*

$$\int \int K(u, v) g(u) g(v) du dv > 0$$

*is valid for all  $g \neq 0$  for which*

$$\int g^2(U) du < \infty.$$

There are various functions that satisfy the Mercer condition, and which can therefore be used as kernels, for example<sup>237</sup>

- Inner product

$$K(\mathbf{x}_l, \mathbf{x}_k) = (\mathbf{x}_l \cdot \mathbf{x}_k)$$

- Polynomial kernel

$$K(\mathbf{x}_l, \mathbf{x}_k) = [(\mathbf{x}_l \cdot \mathbf{x}_k) + 1]^d$$

- Radial basis function

$$K(\mathbf{x}_l, \mathbf{x}_k) = e^{[-c(\mathbf{x}_l - \mathbf{x}_k)^2]}.$$

As a consequence, SVM is a wide class of learning algorithms, which can be built by plugging in different kernel functions. Moreover, the use of SVM can be extended, from classification to regression problems.

---

<sup>237</sup> See Vapnik (2000, p. 143).

### 5.1.3.2 Support vector regression (SVR)

Support vector regression (SVR) is a special type of SVM.<sup>238</sup> The goal of SVR is to estimate the weights vector  $\mathbf{w}$  and bias  $b$  for linear functional dependence between input data points (vectors)  $\mathbf{x}_l$  with known scalar outputs  $y_l$

$$y_l = \sum_{i=1}^N w_i x_{i,l} + b + u_l, \quad l = 1, \dots, L,$$

where  $u_l$  is the error term. The error term describes the deviations between the outputs and the functional approximation:

$$u_l = y_l - \sum_{i=1}^N w_i x_{i,l} - b.$$

SVR is based on a so-called  $\varepsilon$ -insensitive loss function (see Figure 5.6), which does not penalize acceptable deviations defined by a parameter  $\varepsilon$ . Formally, this loss function can be written as

$$|u_l|_\varepsilon = \max(0, |u_l| - \varepsilon). \quad (5.18)$$

In other words, deviations that are larger than parameter  $-\varepsilon$  and smaller than parameter  $\varepsilon$ , have a zero loss. One says that these deviations are in the  $\varepsilon$ -tube.<sup>239</sup>

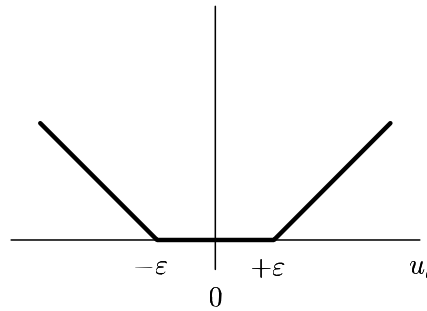


Figure 5.6:  $\varepsilon$ -Insensitive Loss Function

Deviations outside of the  $\varepsilon$ -tube have a loss larger than zero. One can introduce error slopes  $\xi_l, \xi_l^* \geq 0$  for these deviations as follows:

---

<sup>238</sup> See Vapnik (2000).

<sup>239</sup> See Schölkopf/Smola (2002, p. 79).

$$\xi_l = \begin{cases} u_l - \varepsilon, & \text{if } u_l - \varepsilon \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

$$\xi_l^* = \begin{cases} -(u_l - \varepsilon) & \text{if } u_l - \varepsilon \leq 0 \\ 0, & \text{otherwise.} \end{cases}$$

The optimization problem for SVR, which minimizes the  $\varepsilon$ -insensitive loss function (5.18), can be formulated as

$$\begin{aligned} & \underset{\mathbf{w}, b, \xi, \xi^*}{\text{minimize}} && \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{l=1}^L (\xi_l + \xi_l^*) \\ & \text{subject to} && y_l - \sum_{i=1}^N w_i x_{i,l} - b \leq \varepsilon + \xi_l \\ & && \sum_{i=1}^N w_i x_{i,l} + b - y_l \leq \varepsilon + \xi_l^* \\ & && \xi_l, \xi_l^* \geq 0 \\ & && l = 1, \dots, L, \end{aligned} \tag{5.19}$$

where the insensitive region defined by parameter  $\varepsilon$  and the trade-off parameter  $C$  are pre-selected.

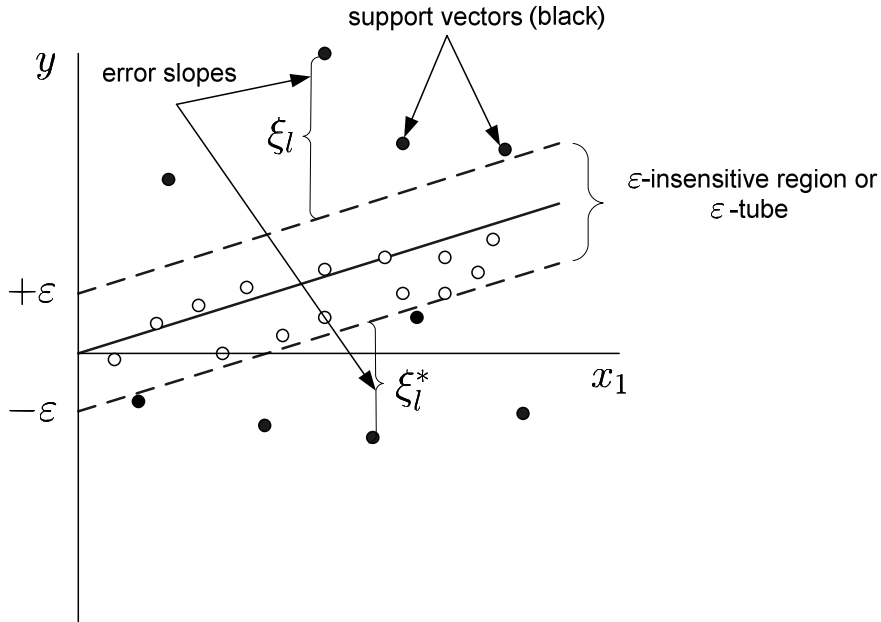


Figure 5.7: Support Vector Regression

The minimization of the error slopes  $\xi_l, \xi_l^*$  leads to a good approximation of the  $y$  by the function  $f(\mathbf{x}) = \sum_{i=1}^N w_i x_i + b$  (see Figure 5.7). Varying the size of the insensitive region can help to achieve a good generalization, that is good performance on data sets other than the given one. Reducing the size of the region increases the approximation error and may lead to

overfitting. This means that the estimated functional dependence will be specific to the given data points.<sup>240</sup> In contrast, increasing the size of the region reduces the approximation error and may lead to underfitting, where the estimated functional dependence does not capture the structure of the given data set. Thus, an optimal trade-off between approximation and generalization, expressed by parameters  $C$  and  $\varepsilon$ , must be found.

The insensitive region can be characterized by data points that are outside or lie on the boundary of the  $\varepsilon$ -insensitive region. These points are called support vectors. In the standard formulation of the SVR margin,  $\varepsilon$  is fixed. *Schölkopf et al.*<sup>241</sup> introduce another formulation of the SVR, a so-called  $\nu$ -SVR, where the margin defined by parameter  $\varepsilon$  is not fixed and is selected automatically based on input data. The influence of the data on the parameter  $\varepsilon$  is specified by the parameter  $\nu \in [0, 1]$ . One can show that  $\nu$  is “an upper bound on the fraction of margin errors and is a lower bound on the fraction of support vectors”.<sup>242</sup> For example, if  $\nu = 0.5$ , then less than 50% of data points are margin errors (where  $\xi_l > 0$  or  $\xi_l^* > 0$ ) and more than 50% of data points are support vectors (where  $\xi_l \geq 0$  or  $\xi_l^* \geq 0$ ).

$$\begin{aligned} \underset{\mathbf{w}, b, \varepsilon, \xi, \xi^*}{\text{minimize}} \quad & \frac{1}{2} \|\mathbf{w}\|_2^2 + C[\nu\varepsilon + \frac{1}{L} \sum_{l=1}^L (\xi_l + \xi_l^*)] \\ \text{subject to} \quad & y_l - \sum_{i=1}^N w_i x_{i,l} - b \leq \varepsilon + \xi_l \\ & \sum_{i=1}^N w_i x_{i,l} + b - y_l \leq \varepsilon + \xi_l^* \\ & \xi_l, \xi_l^* \geq 0 \\ & l = 1, \dots, L. \end{aligned}$$

The solutions of  $\nu$ -SVR and  $\varepsilon$ -SVR coincide if, in  $\varepsilon$ -SVR, the parameter  $\varepsilon$  is equal to  $\varepsilon^*$ , which is an optimal margin detected by  $\nu$ -SVR.<sup>243</sup>

SVR has the following advantages over the classical regression formulation. The first advantage is that SVR is based on the  $\varepsilon$ -insensitive loss function. As a consequence, SVR is robust to outliers. The second advantage is that SVR is designed for predictive learning. This means that its goal is to predict or imitate the output from the inputs and not to explain the dependence between the output and the inputs, as is the goal of a classical regression formulation.

---

<sup>240</sup> We describe overfitting in detail in Section 5.1.4.

<sup>241</sup> See Schölkopf et al. (2000).

<sup>242</sup> See Schölkopf et al. (2000, p. 1226).

<sup>243</sup> See Schölkopf et al. (2000, p. 1215, proposition 2).

### 5.1.4 Model selection

SVR and SVM contain parameters that should be pre-selected before optimization. These parameters control the behaviour of the algorithms and are important for achieving good prediction quality. In this section, we discuss some of the methods that can assist with the selection of parameters for both algorithms.

#### 5.1.4.1 Bias-variance trade-off, model complexity, and overfitting

When solving classification or regression problems, a single model should be selected from a set of available models, for example from a set of linear functions, and its prediction quality should be evaluated. For this purpose, a given dataset  $\mathbf{A}$ , which consists of pairs  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_L, y_L), \dots, (\mathbf{x}_{L'}, y_{L'})$ , is divided randomly into two subsets: a training set (sample),  $\mathbf{A}_{train}$ , containing the first  $L$  data points, and a test set (out of sample),  $\mathbf{A}_{test}$ , containing the remaining  $L' - L$  data points (see Figure 5.8). The training set is used for the estimation and validation of different models, in order to choose the best one. The test set is used to make a further assessment of the prediction quality of the selected model.<sup>244</sup>

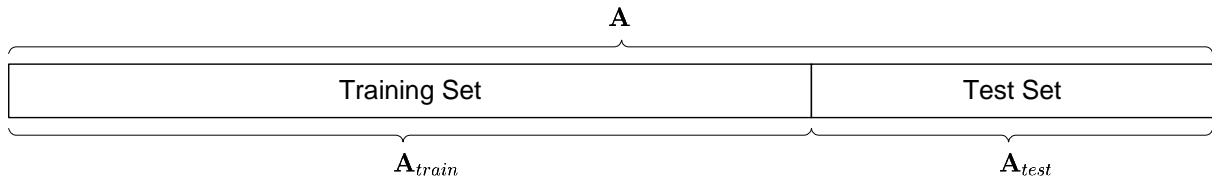


Figure 5.8: Training Set and Test Set

To select the best model from a set of available models, one needs to measure their quality using the data in the training set, and choose the best one. The training set quality of any model  $f(\mathbf{x})$  can be measured by its empirical error (training error), which can be calculated as a mean-squared difference between responses  $y_l$  and their estimations by the model  $f(\mathbf{x}_l)$  on the training set  $\mathbf{A}_{train}$ .<sup>245</sup>

$$err_{emp} = \frac{1}{L} \sum_{l=1}^L (y_l - f(\mathbf{x}_l))^2. \quad (5.20)$$

The training error can also be measured as the mean absolute deviation:

<sup>244</sup> See Hastie/Tibshirani/Friedman (2009, p. 224).

<sup>245</sup> See Cherkassky/Mulier (2007, p. 63).



$$err_{emp} = \frac{1}{L} \sum_{l=1}^L |y_l - f(\mathbf{x}_l)|. \quad (5.21)$$

To evaluate the generalizability or prediction quality of the model  $f(\mathbf{x})$ , the test error, sometimes called the generalization error, is calculated over an independent test set,  $\mathbf{A}_{test}$ .<sup>246</sup>

$$err_{test} = \frac{1}{L' - L} \sum_{l=L+1}^{L'} (y_l - f(\mathbf{x}_l))^2.$$

The empirical error  $err_{emp}$  can be used as a prediction of the test error  $err_{test}$ . In this case, there is an assumption that the best model on the training set will remain the best one on the test set. Unfortunately, this is not generally true. The cause of this behaviour can be illustrated using a bias-variance decomposition of the expected test error.

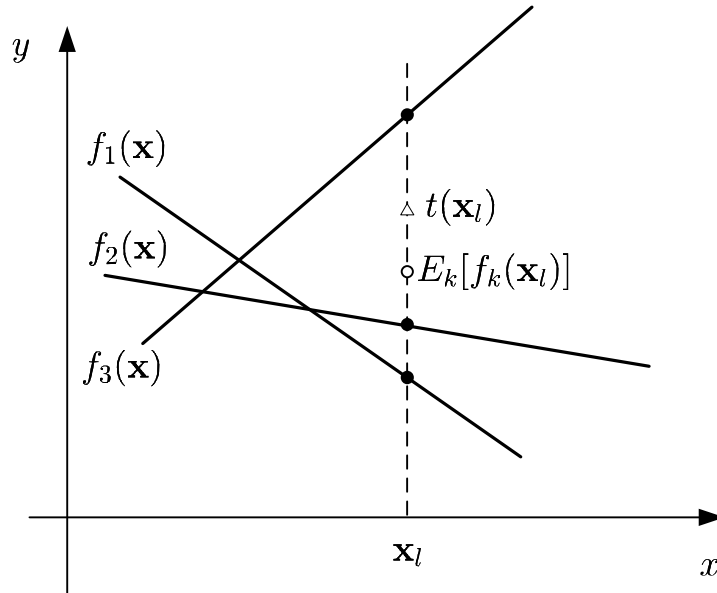


Figure 5.9: True Value and Estimated Values

The model  $f(\mathbf{x})$  in the form  $y = f(\mathbf{x}) + u$ , where  $u$  is an error term with  $E[u] = 0$  and  $Var[u] = \sigma_u^2$ , can be selected using different training sets. The true model  $y = t(\mathbf{x})$  is unknown; only its realizations  $(y_l, \mathbf{x}_l)$ ,  $l = 1, \dots, L$ , corrupted by noise, are observable. If  $K$  training sets are given, then  $K$  models  $f_k(\mathbf{x})$  are selected by the learning algorithm over these training sets (see Figure 5.9). Selecting any data point  $\mathbf{x}_l$ , and computing  $K$  different estimated values  $f_k(\mathbf{x}_l)$  of the true value  $t(\mathbf{x}_l)$  at this point, then the average test error can be decomposed as the sum of three parts:<sup>247</sup>

<sup>246</sup> See Hastie/Tibshirani/Friedman (2009, p. 220).

<sup>247</sup> See Hastie/Tibshirani/Friedman (2009, p. 223).

$$\begin{aligned}
E_k[err_{test}(\mathbf{x}_l)] &= E_k[(y_l - f_k(\mathbf{x}_l))^2] \\
&= E_k[(y_l - E_k[f_k(\mathbf{x}_l)] + E_k[f_k(\mathbf{x}_l)] - f_k(\mathbf{x}_l))^2] \\
&= \sigma_u^2 + (t(\mathbf{x}_l) - E_k[f_k(\mathbf{x}_l)])^2 + E_k[(f_k(\mathbf{x}_l) - E_k[f_k(\mathbf{x}_l)])^2] \\
&= \text{Irreducible Error} + \text{Bias}^2 + \text{Variance}.
\end{aligned} \tag{5.22}$$

Irreducible error is the variance of the realizations  $y_l$  around the true value  $t(\mathbf{x}_l)$ , which cannot be reduced by the learning algorithm. Squared bias is the squared difference between the true value  $t(\mathbf{x}_l)$  and the average of the estimates  $f_k(\mathbf{x}_l)$ . The variance is the variance of the estimates  $f_k(\mathbf{x}_l)$  around their mean  $E_k[f_k(\mathbf{x}_l)]$ . This variance measures the sensitivity of the learning algorithm to the training sets that are used to estimate the model.

To reduce the expectation of the test error in equation (5.22), it is necessary to reduce the bias and the variance simultaneously. The variance of the estimates  $f_k(\mathbf{x}_l)$  around their mean  $E_k[f_k(\mathbf{x}_l)]$  depends on the capacity of the model set, from which the learning algorithm selects estimates  $f_k(\mathbf{x}_l)$ .

Capacity is a measure of content.<sup>248</sup> As it is defined by Sewell:<sup>249</sup>

*Capacity is a measure of complexity and measures the expressive power, richness or flexibility of a set of functions by assessing how wiggly its members can be.*

There are different measures of capacity of the model set. The most known is VC dimension, which we describe in Section 5.1.4.2.

The capacity concept can be illustrated using the following example. There are two model sets: a set of linear functions  $y = bx + c$  and a set of quadratic functions  $y = ax^2 + bx + c$ . The first set is a subset of the second set, since the set of linear functions is a subset of the quadratic functions. One can say that the set of quadratic functions is richer than the set of linear functions.<sup>250</sup> In other words, the capacity of the set of quadratic functions is higher since there are more possibilities to approximate the true function  $t(\mathbf{x}_l)$  by selecting a quadratic function rather than a linear one.

In this context, one can distinguish between complex and simple models. If the model is selected from the set of low capacity, then it is called a simple model, otherwise it is called complex.

---

<sup>248</sup> See Merriam-Webster dictionary.

<sup>249</sup> See Sewell (2008).

<sup>250</sup> See Vapnik (2000).

By increasing the set of available models, one can ensure that the algorithm is capable of finding a model that better fits the data points in the training set. Hence, estimation bias decreases but variance increases. By decreasing the set of available models, one reduces the amount of freedom available to fit the model to the data points. As a consequence, estimation bias increases and variance decreases. In general, if the set of available models for an algorithm is rich, the bias is small but the variance is large, while if the richness of the set of models is lower, the bias will be greater, but the variance smaller. To keep the average test error small, one must minimize the sum of the squared bias and the variance, to find a bias-variance trade-off.<sup>251</sup>

It is possible to reduce the variance of the estimation by increasing the number of data points used to select the model from the set of the models. As a rule, to select a model from a rich set of models, the number of data points needs to be large. If the model set is rich but the number of data points is low, then it is possible to select a model that fits the data points well but whose performance on the test set is significantly worse. This problem is called overfitting.<sup>252</sup>

Summarizing, to ensure the generalizability of the model  $f(\mathbf{x})$ , the important factors are as follows: the capacity of the model set from which the learning algorithm selects the model and the number of data points used to select the model.

To analyse the problem of overfitting, one must consider the difference between the test error and the empirical error. To do so, one can express the test error in terms of the empirical error, in additive form with a confidence interval of  $[err_{test} - err_{emp}]$ , or in multiplicative form with a confidence interval of  $\left[\frac{err_{test}}{err_{emp}}\right]$ .<sup>253</sup>

$$err_{test} = err_{emp} + [err_{test} - err_{emp}] \quad (5.23)$$

$$err_{test} = err_{emp} \left[ \frac{err_{test}}{err_{emp}} \right]. \quad (5.24)$$

We show in the next sections that the confidence interval in either form can be characterized by a function that depends on the capacity of the set of models and on the size of the training set.<sup>254</sup> This function is an increasing function of capacity, and a decreasing function of the size of the training set. In other words, the confidence interval is larger for a rich set of models and a small size of training set, and smaller for a poor set of models and a large

---

<sup>251</sup> See, for example, Hastie/Tibshirani/Friedman (2009).

<sup>252</sup> See, for example, Hastie/Tibshirani/Friedman (2009, p. 221).

<sup>253</sup> See Cherkassky/Mulier (2007, p. 76) and Bousquet/Boucheron/Lugosi (2004, p. 181).

<sup>254</sup> For details see Section 5.1.4.2 and Section 5.1.4.3.

training set. As a consequence, the goal of model selection is to select the model that minimizes the sum of the empirical error and the confidence interval.

Model selection for SVM and SVR is performed by selecting parameters for these algorithms. The parameters limit the capacity of a set of models.<sup>255</sup> By varying these parameters and estimating the test error using (5.23) or (5.24) and the average test error using (5.22), it is possible to select a model of optimal complexity that generalizes the data well.

There are two groups of model selection methods. The first group uses analytic criteria to estimate the discrepancy between the empirical and the test error, based on some measures of model complexity. The second group consists of data-driven methods, which use resampling to estimate the average test error.<sup>256</sup>

#### 5.1.4.2 Analytic method of model selection for classification problems

SVM is based on statistical learning theory, which focuses on the generalizability of learning algorithms.<sup>257</sup> The main statement of the theory is that simple models generalize well.<sup>258</sup> The measurement of model complexity, and a further formalization of this concept which introduced the Vapnik-Chervonenkis (VC) dimension, have become the foundation for constructing SVMs.

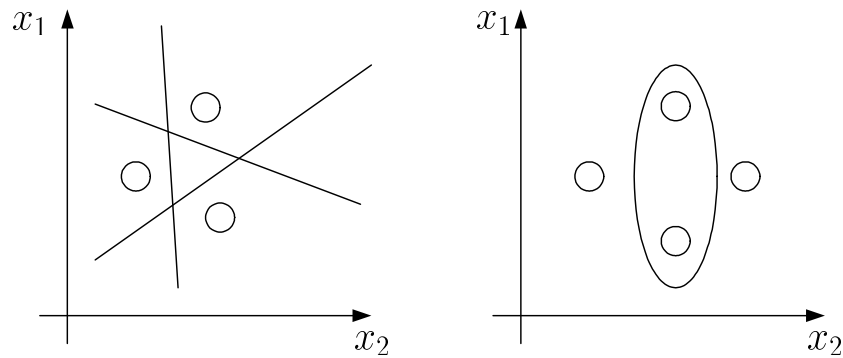


Figure 5.10: VC Dimension for Indicator Functions

The VC dimension  $h$  measures the model complexity, that is the capacity, of the set of models that can be selected by a learning machine.<sup>259</sup> For example, the VC dimension  $h$ , for a set of models that is restricted to a class of indicator functions  $\mathbb{R}^2 \rightarrow \{-1, 1\}$  in two dimensional

<sup>255</sup> For details see Section 5.1.4.2 and Section 5.1.4.3.

<sup>256</sup> See Cherkassky/Mulier (2007, p. 129).

<sup>257</sup> See Vapnik (1998).

<sup>258</sup> See Vapnik (1995, p. 129).

<sup>259</sup> See Vapnik (1995)

space, is given by  $h = 3$ . This is because an indicator function can be represented as a line on the plane that separates the given data points and, as a result, it is possible to split up at least any three points into two classes in any way (see Figure 5.10 left). However, it is not possible to do the same for four points using a line (see Figure 5.10 right). Therefore, the VC dimension of the set of lines on the plane is  $h = 3$ , which is the maximal number of points that can be divided up by this set of lines. It is possible to generalize this concept to  $N$  dimensions. The  $N$ -dimensional hyperplane has a VC dimension of  $h = N + 1$ .

The hyperplane that separates the classes with a margin can have a lower VC dimension than a simple line. We can illustrate this using an example. It is possible to separate three points in any way using a line (see Figure 5.11 left). Moreover, it is also possible to separate points 1 and 3 from point 2 by a margin hyperplane. However, it is not possible to separate points 1 and 2 from point 3 as well as points 3 and 2 from point 1 using a hyperplane with the same margin (see Figure 5.11 right).

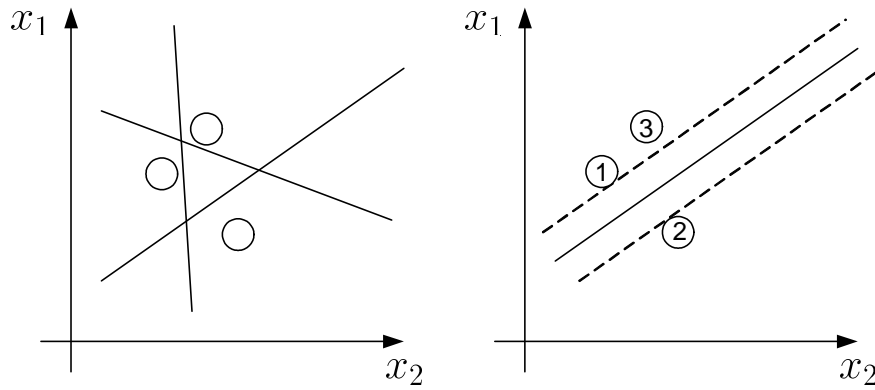


Figure 5.11: Margin Hyperplane with a Lower VC Dimension

Hence, the margin hyperplane can be simpler than a line. The larger is the margin, the lower is the VC dimension. SVM selects the hyperplane with the maximal margin, which is the simplest possible model. *Vapnik*<sup>260</sup> proves that the VC dimension for a margin hyperplane is bounded by

$$h \leq \min\left(\frac{R^2}{\gamma^2}, N\right) + 1 \quad (5.25)$$

where  $R$  is the radius of the minimal sphere, which includes all training points, and  $\gamma$  is the margin. If the margin goes to zero ( $\gamma \rightarrow 0$ ), this formula gives the VC dimension for a line ( $h = N + 1$ ).

---

<sup>260</sup> See Vapnik (2000, p. 132, Theorem 5.1).

Moreover, the discrepancy between the test error and the empirical error  $[err_{test} - err_{emp}]$  in equation (5.23) can be quantified by the function  $\Phi(L/h)$ :

$$err_{test} - err_{emp} \leq \Phi(L/h), \quad (5.26)$$

where  $\Phi(L/h)$  is a decreasing function that depends on the relation between the number of examples in the training set  $L$  and the VC dimension  $h$ . Statistical learning theory introduces an explicit form for  $\Phi(L/h)$ <sup>261</sup>

$$\Phi(\tau) = \begin{cases} 1 & \text{if } \tau \leq 0.5, \\ a^{\frac{\ln 2\tau+1}{\tau-k}} \left( 1 + \sqrt{1 + \frac{b(\tau-k)}{\ln 2\tau+1}} \right) & \text{if } \tau > 0.5 \end{cases} \quad (5.27)$$

where  $\tau = L/h$  and  $a = 0.16, b = 1.2, k = 0.15$ . The values of the constants  $a, b, k$  were found empirically by Vapnik *et al.*<sup>262</sup> The estimation of the VC dimension  $h$  is given by (5.25).

If the relation  $L/h$  is smaller than 0.5, then the function  $\Phi(L/h)$  takes the largest value, 1. In this case, there is not enough data to estimate the model of complexity  $h$ .

One can rewrite (5.26) as

$$err_{test} \leq err_{emp} + \Phi(L/h). \quad (5.28)$$

The sum of the empirical error and the confidence interval  $\Phi(L/h)$  is an upper bound for the test error, which can then be used to select the model. Two strategies for finding the optimal model can be derived from this inequality.

The first strategy, called the empirical risk minimization principle (ERM), pays no attention to model complexity but simply selects the model that minimizes the empirical error. However, this method has a disadvantage in that the term  $\Phi(L/h)$  can be quite large and, as a consequence, the prediction of the test error will not be accurate.

The second strategy, called the structural risk minimization principle (SRM), solves the problem in the following way. Firstly, the structure, which consists of nested sets of models with different capacities, as measured by the VC dimension  $h$ , is defined ( $S_1 \subset S_2 \subset \dots \subset S_n$ ). This structure should be defined in advance, without the use of

---

<sup>261</sup> See Vapnik/Levin/Cun (1994, p. 7).

<sup>262</sup> See Vapnik/Levin/Cun (1994, p. 7).

training data. Secondly, the models with the least empirical risk in each of these sets are selected. Thirdly, from the models chosen in step 2, the model from the set with VC dimension  $h^*$  is selected that has the lowest upper bound (5.28) (see Figure 5.12).<sup>263</sup>

For example, the SVM implements the SRM. Parameter  $C$  controls the complexity of the set of models available to the learning algorithm, since it puts a limit on the lowest possible margin between two classes. Therefore, one can introduce the structure through a sequence of parameters  $C_1 < C_2 < \dots < C_n$  with corresponding complexities  $h_1 \leq h_2 \leq \dots \leq h_n$ . By solving the SVM for different parameters  $C$ , one can choose the model with the minimal bound on the test error.

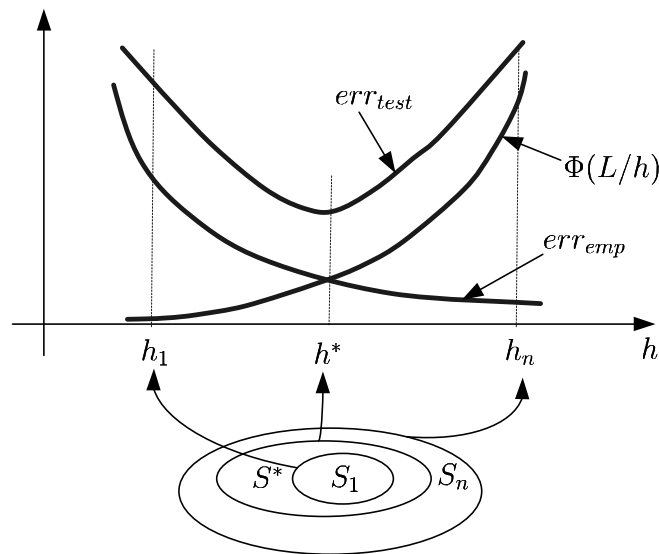


Figure 5.12: SRM and Bounds on Test Error<sup>264</sup>

Moreover, *Lugosi and Zeger* have proved the following theorem:

**Theorem (Lugosi, Zeger).**<sup>265</sup> *For any distribution function the SRM method provides convergence to the best possible solution with probability one (i.e. the learning using SRM method is consistent).*<sup>266</sup>

### 5.1.4.3 Analytic methods of model selection for regression problems

There are different methods of estimating the discrepancy between the empirical error and the test error in the forms (5.23) and (5.24) for regressions. These criteria, analogous to the

<sup>263</sup> See Vapnik (2000, p. 96).

<sup>264</sup> See Vapnik (2000, p. 96).

<sup>265</sup> See Vapnik (1998, p. 229).

<sup>266</sup> In Section 5.1.2 we defined the concept of the learning consistency.

classification case, are based on the size of the training set and on some measure of the capacity of the set of admissible models for the learning algorithm. Two capacity measures can be used here: the number of degrees of freedom  $d$ , which represents the independent pieces of information given by the data,<sup>267</sup> and the VC dimension  $h$ .<sup>268</sup>

The criteria used most often are the Akaike information criterion (AIC) and the Bayesian information criterion (BIC),<sup>269</sup> which can be written as

$$AIC(d) = err_{emp}(d) + 2\frac{d}{L}\hat{\sigma}_u^2 \quad (5.29)$$

$$BIC(d) = err_{emp}(d) + (\ln L)\frac{d}{L}\hat{\sigma}_u^2, \quad (5.30)$$

where  $\hat{\sigma}_u^2$  is an estimation of the variance of the noise, and  $err_{emp}(d)$  is an empirical error, taking the form of (5.20) or (5.21).

If empirical error in the form of (5.20) is used, then the estimation of the noise in (5.29) and (5.30) can be written as

$$\hat{\sigma}_u^2 = \frac{L}{L-d} \frac{1}{L} \sum_{l=1}^L (y_l - f(y_l))^2 = \frac{L}{L-d} err_{emp}(d). \quad (5.31)$$

Substituting expression (5.31) into (5.29) and (5.30) respectively, gives the final prediction error (FPE) and the Schwartz criterion (SC):<sup>270</sup>

$$FPE = err_{emp}(d)(1+p)(1-p)^{-1} \quad (5.32)$$

$$SC = err_{emp}(d) (1+p(1-p)^{-1} \ln L) \quad (5.33)$$

where  $p = d/L$ .

There are other criteria, for example, generalized cross validation (GCV) and Shibata's model selector (SMS), that can also be used to estimate test error:<sup>271</sup>

---

<sup>267</sup> See Glenn/Littler (1984, p. 46).

<sup>268</sup> Practical methods for estimating the degrees of freedom and the VC dimension are discussed in Section 5.1.4.4.

<sup>269</sup> See Hastie/Tibshirani/Friedman (2009, pp. 231 and 233).

<sup>270</sup> See Cherkassky/Mulier (2007, p. 131) or Vapnik (1998, p. 524).

<sup>271</sup> See Cherkassky/Mulier (2007, p. 77) or Vapnik (1998, p. 524).



$$GCV = \frac{err_{emp}(d)}{(1-p)^2} \quad (5.34)$$

$$SMS = err_{emp}(d)(1+2p). \quad (5.35)$$

There is also an estimation of test error based on statistical learning theory. The practical bound on the test error for unbounded indicator functions (Vapnik method (VM)), for the regression problem where  $h$  is the VC dimension, has the form<sup>272</sup>

$$VM = err_{emp}(q) \left( 1 - \sqrt{q - q \ln q + \frac{\ln L}{2L}} \right)^{-1}_+, \quad (5.36)$$

where  $q = h/L$  and

$$(a)_+ = \max(0, a).$$

The criteria FPE, SC, GCV, SMS, and VM, which estimate the test error in the multiplicative form (5.24), are used to select parameters  $C$  and  $\varepsilon$  for SVR, producing the minimal estimated test error.

The growth of the correction factors for the empirical error  $err_{emp}$  for the above criteria is displayed in Figure 5.13.

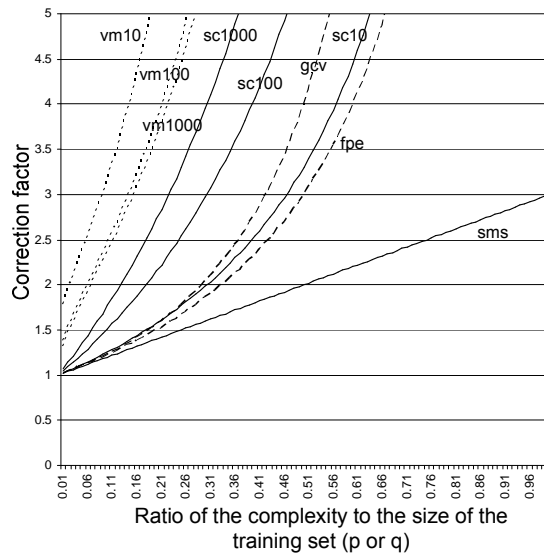


Figure 5.13: Correction Factors for Complexity<sup>273</sup>

<sup>272</sup> See Cherkassky/Mulier (2007, p. 129).

<sup>273</sup> Based on Cherkassky/Mulier (2007, p. 130).

For example for the SMS method the correction factor is equal to  $(1 + 2p)$  and it is displayed as a line. The correction factors for the VM and SC methods depend on the size of the training set. Therefore, we show three variants of these methods for training set sizes 10, 100, and 1000. Figure 5.13 indicates that the VM is more conservative than the other methods. In contrast, the SMS method is the most optimistic.

#### 5.1.4.4 Measuring model complexity

Bias-variance trade-off and statistical learning theory establish the fact that the generalizability of the learning algorithm depends on the empirical error and the capacity of the set of models. Imposing restrictions on the parameters of the required model, one can reduce the capacity of the available set of models.<sup>274</sup>

Figure 5.14 illustrates the reduction of the capacity of a set of linear models. The set of linear models without constraints on the parameters is the richest set. One can reduce this set by adding constraints on the parameters. In the extreme case, with the smallest possible parameter  $C$ , the set is limited to one variant. In this case, the algorithm will return the same model regardless of the data, so the variation will be zero but the bias will be large.

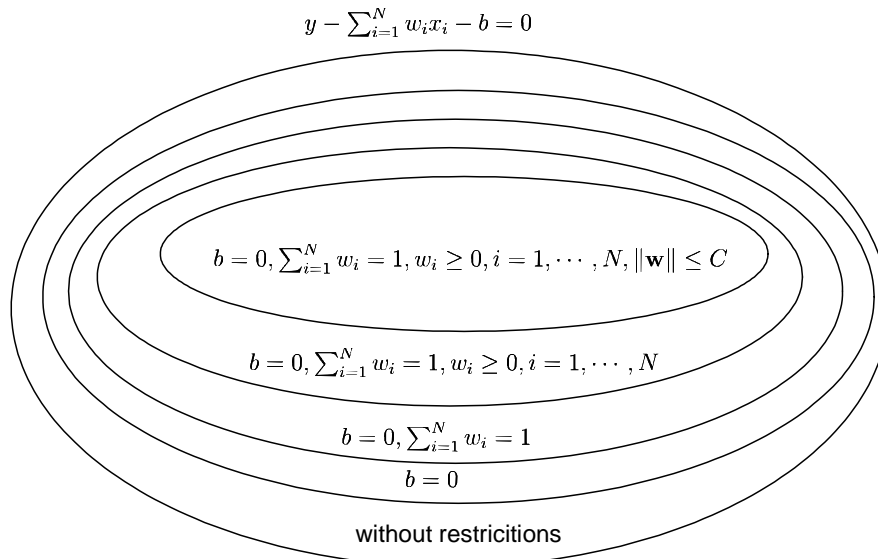


Figure 5.14: Reducing Complexity for the Hyperplane

In many cases, the capacity of the set of models cannot be measured theoretically. Therefore, it is necessary to use approaches that estimate the model complexity using the data. In the following, we will discuss such approaches: one approach to estimate the effective degrees of freedom and one to estimate the effective VC dimension.

<sup>274</sup> See Vapnik (2006, p. 452).

#### 5.1.4.4.1 Measuring the effective degrees of freedom

The number of degrees of freedom (DoF) can be obtained by analysing the covariance between the predicted value, which is measured based on different training sets, and the observed value. Based on the covariance, one can introduce the effective DoF as<sup>275</sup>

$$d = \sum_{l=1}^L \frac{Cov(\hat{y}_l, y_l)}{\sigma_u^2}. \quad (5.37)$$

Unfortunately, the covariance is not observable but it can be estimated using bootstrap methods.<sup>276</sup>

Another algorithm for estimating DoF is proposed by Ye.<sup>277</sup> It estimates DoF using  $M$  training sets, which are derived from the original training set using perturbations of size  $\tau$ :

- Repeat  $m = 1, \dots, M$
- Generate perturbations  $\Delta_m = (\delta_{m1}, \dots, \delta_{mL})$  from the density  $\frac{1}{\tau^L} \prod_{l=1}^L \phi(\delta_{ml}/\tau)$
- Estimate  $\hat{\mathbf{y}}_m$  using the modified data set  $\mathbf{y} + \Delta_m$
- Calculate  $\hat{h}_l$  as the regression slope from  $\hat{y}_m = \alpha + \hat{h}_m \delta_{ml}$
- Generalized degrees of freedom are equal to  $d = \sum_{l=1}^L \hat{h}_l$

This algorithm is based on definition (5.37) for DoF.<sup>278</sup> The resulting DoF  $d$  is then used in conjunction with the criteria from Section 5.1.4.3 to select the model.

#### 5.1.4.4.2 Measuring the effective VC dimension

The VC dimension can be measured by estimating the maximum discrepancy between the empirical error and the test error. For this purpose, one can construct an experiment that assesses the VC dimension using the data (see Figure 5.15):<sup>279</sup>

- Repeat  $k = 1, \dots, K$

---

<sup>275</sup> See Hastie/Tibshirani/Friedman (2009, p. 233).

<sup>276</sup> See Efron (2004, p. 620)

<sup>277</sup> See Ye (1998, p. 122).

<sup>278</sup> See Ye (1998, p. 122, Definition 1).

<sup>279</sup> See Vapnik/Levin/Cun (1994, p. 7).

- Select a random training set with size  $s_k$  from the original training set
- Assign uniformly random labels  $\{-1, 1\}$  to the training data points
- Divide the training sample into two equal parts
- Flip the labels to their reverse values for the second part
- Combine these sets and estimate the model
- Flip the labels for the second part back to what they were originally
- Calculate the errors for the first part and the second part
- Compute the discrepancy:

$$d_k = |err_2 - err_1|. \quad (5.38)$$

It is necessary to minimize  $err_1$  and maximize  $err_2$  to determine the maximal value of the discrepancy  $d_k$  using the data in the training set. One can achieve this goal by using a training set made up of two parts: the first part with the original labels and the second part with the inverted labels. The first part of the training set is used to minimize  $err_1$ , and the second part to maximize  $err_2$ .

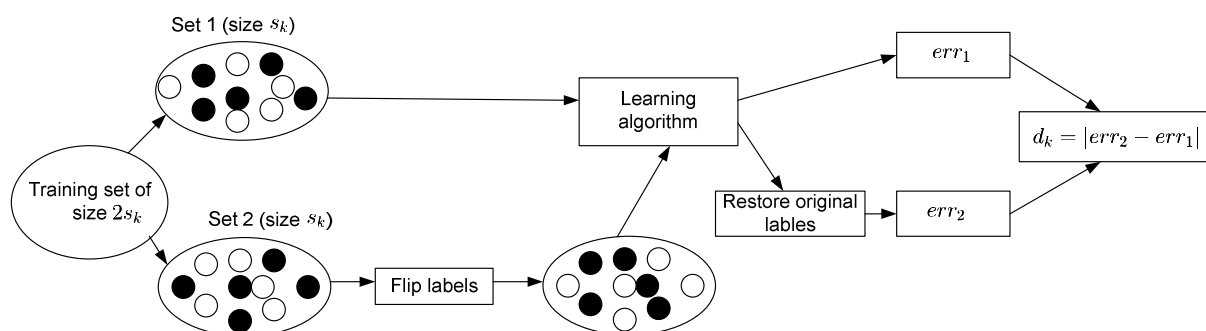


Figure 5.15: Measuring the VC Dimension<sup>280</sup>

By definition, from statistical learning theory, the values  $d_k$  are approximately equal to

$$d_k \approx \Phi(s_k/h),$$

where the function  $\Phi$  represents the discrepancy between the training error and the test error and has the analytical form (5.27). This makes it possible to determine the VC dimension  $h$  by solving the optimization problem

$$\underset{h}{\text{minimize}} \quad \sum_{k=1}^K (\Phi(s_k/h) - d_k)^2. \quad (5.39)$$

<sup>280</sup> Based on Shao/Cherkassky/Li (2007, p. 28).

The sizes  $s_k$  of the training sets should be selected in advance. *Shao et al.*<sup>281</sup> have shown that the estimation of the VC dimension is more stable if the experiments are repeated for values  $s_k/h$  that are greater than 5.

The estimated VC dimension used for the model selection in this research will be calculated using bounds on the test error (5.27) for the classification problem and (5.36) for the regression problem.

#### 5.1.4.5 Resampling methods of model selection

Resampling methods of model selection aim to estimate the average test error. The simplest resampling method is a holdout.<sup>282</sup> The data set is randomly divided into three parts: a training set, a validation set, and a test set (see Figure 5.16). For every given set of models,<sup>283</sup> the learning algorithm selects the model with the minimal empirical error, using the training set. In the next step, out of these selected models, the one with the minimal validation error on the validation set is chosen. The quality of this model is then estimated using the test set.

Training Set	Validation Set	Test Set
--------------	----------------	----------

Figure 5.16: Hold Out

Cross-validation is a more complicated resampling method.<sup>284</sup> Suppose that the sets of models to be used in the learning algorithm are given. The purpose of cross-validation is to choose, from these sets, the one that gives the lowest estimate of expected test error. Initially, the training set  $A$  is divided into  $K$  subsets (folds),  $\mathbf{A}_k$  (see Figure 5.17).

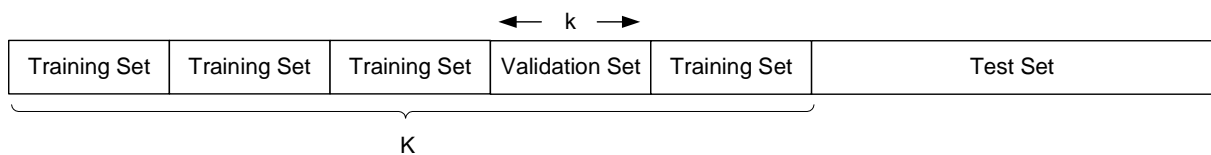


Figure 5.17: K-Fold Cross-Validation

Next, one fixes a set of models through an appropriate parameterization of the learning algorithm, selects model  $f_k$  from this set using the training data  $\mathbf{A} \setminus \mathbf{A}_k$ , and, finally, validates every model  $f_k$  on the set  $\mathbf{A}_k$ , to compute the total validation error:

<sup>281</sup> Shao/Cherkassky/Li (2000, p. 21).

<sup>282</sup> See Kohavi (1995, Section 2.1).

<sup>283</sup> For example, for SVM the model sets are determined by the values of parameter  $C$ .

<sup>284</sup> See Kohavi (1995, Section 2.2).

$$CV = \frac{1}{L} \sum_{k=1}^K \sum_{l \in A_k} (y_l - f_k(x_{i,l}))^2.$$

This step is repeated for all given sets of models. Next, the set of models with the lowest total validation error is selected. The learning algorithm then estimates a model using the chosen model set, for the whole set  $\mathbf{A}$ . The quality of this model is determined using the test set.

Typical numbers of folds used are  $K = 5$  and  $K = 10$ . The special case, when every fold contains one example ( $K = L$ ), is called leave-one-out (LOO) cross-validation.

#### 5.1.4.6 Model selection based on the number of support vectors

Another method of model selection is based on the number of support vectors:<sup>285</sup>

$$err_{test} \leq \frac{\#SV}{L}.$$

The test error in this case is bounded by the ratio of the number of support vectors to the number of data points. Model selection using this bound is implemented using  $\nu$ -SVR.

#### 5.1.4.7 Model selection process

Having listed the model selection methods, it is now possible to define the model selection process (see Figure 5.18).

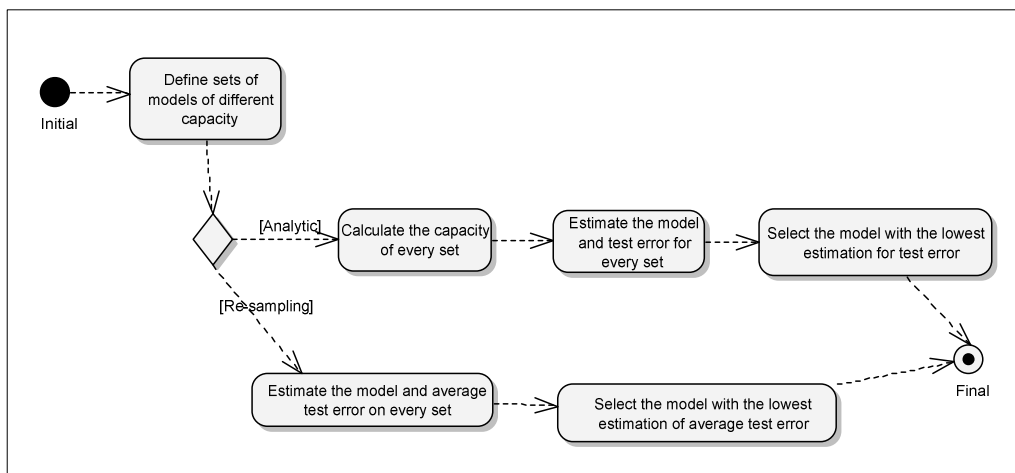


Figure 5.18: Model Selection Process

<sup>285</sup> See Cortes/Vapnik (1995, p. 275).

Initially, sets of models with different capacities are defined. One can proceed further with analytic or resampling methods of model selection. For analytic methods, it is necessary to compute the capacity of every set, which is measured by the VC dimension or the DoF. The capacity of the set is also used to estimate the test error. Alternatively, one can estimate the test error directly via resampling methods, using a part of the data set that is not used for the model estimation. The estimation of the test error, whether obtained through analytic or resampling methods, is then used to select the best model.

### 5.1.5 Extended version of SVR

To control the capacity of the model set directly in the optimization problem for SVR, it is necessary to rewrite the SVR given in (5.19) as

$$\begin{aligned}
& \underset{\mathbf{w}, b, \xi, \xi^*}{\text{minimize}} && \sum_{l=1}^L (\xi_l + \xi_l^*) \\
& \text{subject to} && y_l - \sum_{i=1}^N w_i x_{i,l} - b \leq \varepsilon + \xi_l \\
& && \sum_{i=1}^N w_i x_{i,l} + b - y_l \leq \varepsilon + \xi_l^* \\
& && \xi_l, \xi_l^* \geq 0 \\
& && \frac{1}{2} \|\mathbf{w}\|_2^2 \leq C' \\
& && l = 1, \dots, L,
\end{aligned} \tag{5.40}$$

The difference between SVR (5.19) and SVR (5.40) is the treatment of the 2-norm. The 2-norm in SVR (5.19) appears in the objective of the optimization problem, but the 2-norm in SVR (5.40) is a part of the constraint. Although it is possible for the fixed  $C'$  to find the value of the penalty parameter  $C$  in SVR (5.19) such that the solutions of both formulations coincide,<sup>286</sup> SVR (5.40) allows us to quantify the model capacity directly using the penalty parameter  $C'$ .<sup>287</sup>

Moreover, it is possible to extend the 2-norm constraint in SVR (5.40) to<sup>288</sup>

$$\frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C',$$

---

<sup>286</sup> See Vapnik (2000, p. 186).

<sup>287</sup> See Vapnik (2006, p. 435).

<sup>288</sup> This constraint has the same form as when it is used for the regularization of linear programs in Tikhonov/Arsenin (1979, p. 269).

where  $\mathbf{S}$  is a positive semi-definite matrix and  $\mathbf{w}_0$  is a fixed vector. Using this type of constraint it is possible to include different types of preliminary information about vector  $\mathbf{w}$ .<sup>289</sup>

### 5.1.6 The transformation of SVR into SVM

The model selection for SVR includes the selection of two meta-parameters that influence the generalizability of the algorithm: the trade-off parameter  $C$  and the size of the insensitive region  $\varepsilon$ . There are various ways of accomplishing this task:

1. The parameters are selected based on analytic bounds (see Section 5.1.4.3).
2. The parameters are selected using resampling methods (see Section 5.1.4.5).
3. The user selects the parameters manually, based on experience.
4. The parameters are selected to achieve a fixed percentage of support vectors out of the total number of data points. This can be done using  $\nu$ -SVR. For example, if  $\nu = 0.5$ , then 50% of the data points are support vectors and 50% (the rest) are inside the insensitive region.<sup>290</sup>
5. The parameters are selected using analytic criteria based on the training data: parameter  $C$  is equal to  $C := \max(|\bar{y} + 3\sigma_y|, |\bar{y} - 3\sigma_y|)$ , where  $\bar{y}$  and  $\sigma_y$  are the mean and variance of the scalar outputs  $y_l$  in the training set, parameter  $\varepsilon$  is equal to  $\varepsilon := 3\sigma\sqrt{\frac{\ln L}{L}}$ , where  $\sigma$  is the variance of the noise, and  $L$  is the size of the training set.<sup>291</sup>

We use analytic and resampling methods for our model selection because they have a strong theoretical background (see Section 5.1.4). To simplify their application, one can transform SVR with two parameters  $(C, \varepsilon)$  into SVM with one parameter  $(C^c)$ . This transformation reduces the computational burden of the model selection.

The relation between SVR and SVM has been studied previously by *Pontil et al.* and *Ancona*.<sup>292</sup> Their work shows that SVR can be seen as an extension of SVM and that it is possible to use it for the classification task. In addition, *Bi and Benett*<sup>293</sup> show that the regression problem can be converted to a classification problem and, based on this fact, they

---

<sup>289</sup> See, for example, Section 5.2.3.

<sup>290</sup> See Schölkopf et al. (2000). Value  $\nu = 0.5$  is recommended by Mattera/Haykin (1998, p. 226).

<sup>291</sup> See Cherkassky/Ma (2004).

<sup>292</sup> See Pontil/Rifkin/Evgeniou (1999) and Ancona (1999).

<sup>293</sup> See Bi/Benett (2003).



introduce new algorithms H-SVR and RH-SVR. We extend this idea and show that SVR can be transformed into SVM.

The main idea behind the transformation of SVR into SVM is to modify the initial training set in such a way that, instead of the regression line, one must look for the separating hyperplane. For example, suppose we have the data point  $(\mathbf{x}_l, y_l)$ . We can modify this data point by adding and subtracting a fixed value  $\theta > 0$  and, as a result, we have two data points  $(\mathbf{x}_l, y_l + \theta)$  and  $(\mathbf{x}_l, y_l - \theta)$ . We can say that the first pair belongs to class 1, and the second pair to class -1. If we apply this transformation to all data points (see Figure 5.19), then the regression problem is converted into the classification problem.<sup>294</sup>

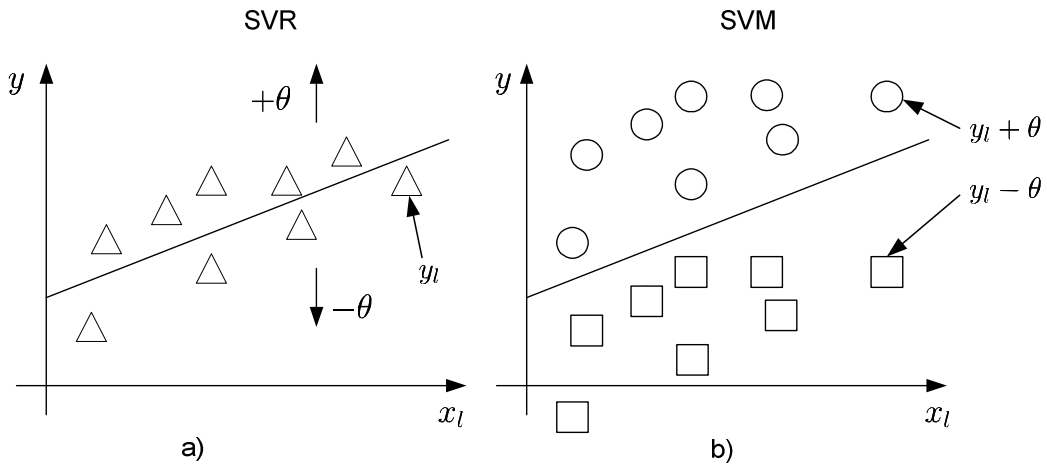


Figure 5.19: Conversion of SVR into SVM: a) Original Data and Regression Line b) Shifted Data and Separating Plane<sup>295</sup>

To formulate an optimization problem to find the separating hyperplane, one can fix  $\theta > 0$  such that, for any  $l = 1, \dots, L$ , all errors  $u_l = y_l - \sum_{i=1}^N w_i x_{i,l} - b$  are in the interval  $[-\theta, \theta]$ :

$$-\theta \leq u_l \leq \theta, \quad (5.41)$$

Rewriting (5.41) as

$$\begin{aligned} u_l + \theta &\geq 0 \\ u_l - \theta &\leq 0, \end{aligned} \quad (5.42)$$

one can define two classes: the first class  $\{+1\}$  has  $u_l + \theta \geq 0$  and the second class  $\{-1\}$  has  $u_l - \theta \leq 0$ .

<sup>294</sup> See Bi/Bennett (2003, p. 2).

<sup>295</sup> Modified from Bi/Bennett (2003, p. 3).

The following proposition shows the equivalence of SVR and SVM, by fixing parameter  $\theta$  and imposing the constraints  $\sum_{i=1}^N w_i = 1, w_i \geq 0, i = 1, \dots, N$ :

**Proposition 5.1:** For any fixed  $\theta > \varepsilon > 0$ , the solutions of the classification problem

$$\begin{aligned}
& \underset{\mathbf{w}^c, b^c, \xi^c, \xi^{*c}}{\text{minimize}} && \sum_{l=1}^L (\xi_l^c + \xi_l^{*c}) \\
& \text{subject to} && \sum_{i=1}^N w_i^c (y_l - x_{i,l} - \theta) - b^c \leq -1 + \xi_l^c \\
& && - \sum_{i=1}^N w_i^c (y_l - x_{i,l} + \theta) + b^c \leq -1 + \xi_l^{*c} \\
& && \sum_{i=1}^N w_i^c \geq \frac{1}{\theta} \\
& && w_i^c \geq 0, i = 1, \dots, N \\
& && \xi_l^c, \xi_l^{*c} \geq 0 \\
& && \frac{1}{2} \|\mathbf{S}(\mathbf{w}^c - \mathbf{w}_0 \sum_{i=1}^N w_i^c)\|_2^2 \leq C^c \\
& && l = 1, \dots, L.
\end{aligned} \tag{5.43}$$

and the regression problem

$$\begin{aligned}
& \underset{\mathbf{w}, b, \xi, \xi^*}{\text{minimize}} && \sum_{l=1}^L (\xi_l + \xi_l^*) \\
& \text{subject to} && y_l - \sum_{i=1}^N w_i x_{i,l} - b \leq \varepsilon + \xi_l \\
& && \sum_{i=1}^N w_i x_{i,l} + b - y_l \leq \varepsilon + \xi_l^* \\
& && \sum_{i=1}^N w_i = 1 \\
& && w_i \geq 0, i = 1, \dots, N \\
& && \xi_l, \xi_l^* \geq 0 \\
& && \frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C \\
& && l = 1, \dots, L.
\end{aligned}$$

coincide, where

$$w_i = \frac{w_i^c}{\sum_{i=1}^N w_i^c}$$

$$b = \frac{b^c}{\sum_{i=1}^N w_i^c}$$

$$\varepsilon = \theta - \frac{1}{\sum_{i=1}^N w_i^c} \geq 0$$

$$\xi_l = \frac{\xi_l^c}{\sum_{i=1}^N w_i^c}$$

$$\xi_l^* = \frac{\xi_l^{*c}}{\sum_{i=1}^N w_i^c}$$

$$C = \frac{C^c}{(\sum_{i=1}^N w_i^c)^2}.$$

**Proof:** see Appendix 2.

The above proposition describes how to convert a regression problem with two parameters  $C, \varepsilon$  into a classification problem with one parameter  $C^c$ . However, to make such a transformation, it is necessary to choose the parameter  $\theta$ . Parameter  $\theta$  is selected in advance and specifies the required precision of the regression problem. On the one hand, if the value of  $\theta$  is very small, we may see a situation in which one of the inequalities (5.42) is not satisfied, meaning that there is no way to define the classes. On the other hand, as the value of  $\theta$  increases, the model accuracy deteriorates.

One possible way to choose the parameter  $\theta$  is to find the smallest possible value for which the inequalities (5.42) are still satisfied. This can be done by solving the following optimization problem:

$$\begin{aligned} & \underset{\mathbf{w}, \theta}{\text{minimize}} && \theta \\ & \text{subject to} && y_l - \sum_{i=1}^N w_i x_{i,l} - b \leq \theta \\ & && y_l - \sum_{i=1}^N w_i x_{i,l} - b \geq -\theta \\ & && \sum_{i=1}^N w_i = 1 \\ & && l = 1, \dots, L \end{aligned}$$

Then, we would use the solution  $\theta = \theta^*$  for the conversion.

The geometrical interpretation of the SVM (5.43) in the regression data space is illustrated in Figure 5.20. The margin  $\gamma$  is an interval that is cut off on the axis  $y$  by two parallel hyperplanes  $y - \sum_{i=1}^N w_i x_i - b = \theta^*$  and  $y - \sum_{i=1}^N w_i x_i - b = \varepsilon$ .<sup>296</sup> The larger the margin, the more probable it is that errors  $u_l$  remain in the interval  $[-\theta^*, \theta^*]$ . SVM (5.43) automatically adjusts margin  $\gamma$  for a fixed capacity of the set of models, which is specified by the 2-norm constraint and the constant  $C^c$ .

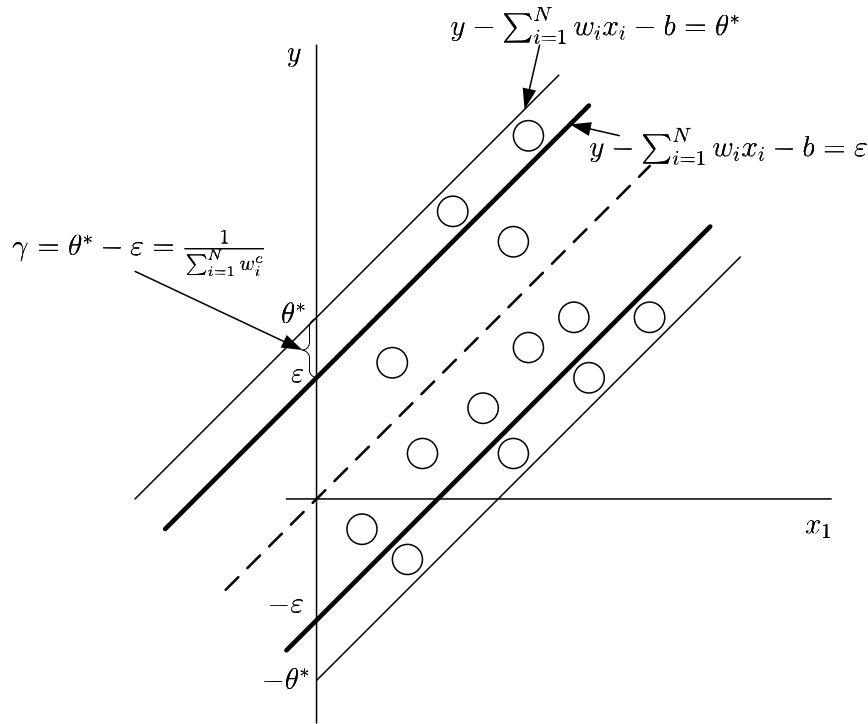


Figure 5.20: Precision and Margin for SVR

SVM (5.43) implements the SRM principle.<sup>297</sup> The nested structure is defined by the sequence of parameters  $C_1^c < C_2^c < \dots < C_M^c$ . The smaller is parameter  $C^c$ , the smaller is the capacity of the set of models. Using the model selection methods from Section 5.1.4, one can find an appropriate bias-variance trade-off and select the optimal value of parameter  $C^c$ .

<sup>296</sup> See Shawe-Taylor/Christianini (1999, p. 10).

<sup>297</sup> See Section 5.1.4.2.

## 5.2 Tracking portfolio selection using SV techniques

In this section we introduce the selection of tracking portfolios using SV techniques. First, we formulate SVR for index tracking, and then we show different ways of selecting the size of the insensitive region and the capacity of the set of tracking portfolio models for the SVR. After that, we develop the structural tracking error minimization approach. Finally, we describe further possible modifications of SVR for the index tracking problem.

### 5.2.1 SVR for index tracking

In Section 3.4 we formulated the requirements for the solution of the index tracking problem. By further analysing the traditional methods of index tracking empirically in Section 4.2 and Section 4.3, we found that those methods do not meet the defined requirements. In particular, first, the portfolios constructed have weak absolute and relative stability of tracking quality and, second, the tracking portfolio estimation is not resistant to slight changes in the data used for the estimation, which leads to the problem of overfitting. To overcome these problems, we formulate SVR to solve the index tracking problem.

Unlike other methods (e.g. MSE and MAD), SVR is based on the  $\varepsilon$ -insensitive loss function (see Figure 5.21). This function penalizes any deviations between the returns of the index and those of the tracking portfolios that are smaller than  $-\varepsilon$  or larger than  $\varepsilon$ .

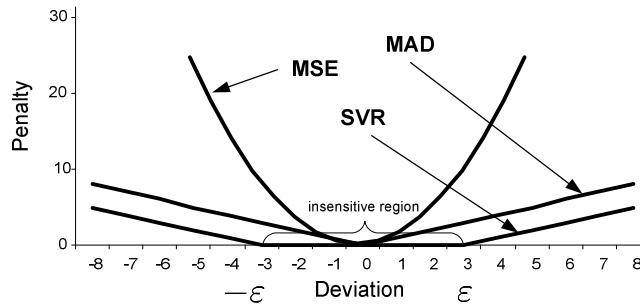


Figure 5.21: Comparison of Different Measures of Tracking Quality

The optimization problem, using the  $\varepsilon$ -insensitive loss function as an objective, for the index tracking problem has the following form:

$$\begin{aligned}
& \underset{\mathbf{w}, b, \xi, \xi^*}{\text{minimize}} && \sum_{t=1}^T (\xi_t + \xi_t^*) \\
& \text{subject to} && R_{I,t} - \sum_{i=1}^N w_i r_{i,t} - b \leq \varepsilon + \xi_t \\
& && \sum_{i=1}^N w_i r_{i,t} + b - R_{I,t} \leq \varepsilon + \xi_t^* \\
& && \xi_t, \xi_t^* \geq 0, t = 1, \dots, T,
\end{aligned} \tag{5.44}$$

where  $R_{I,t}$  are the returns on the index,  $r_{i,t}$  are the returns on asset  $i$ ,  $w_i$  is the weight of asset  $i$  in the tracking portfolio,  $b$  is the shift between the index and the tracking portfolio,<sup>298</sup>  $\varepsilon$  is the pre-selected size of the insensitive region, and  $\xi_t, \xi_t^*$  are the error slopes. Allowing optimization procedure to determine shift  $b$ , we concentrate on the minimizing the variability of the tracking portfolio around the index. Additionally, there are two constraints on the portfolio weights, namely, the budget constraint and a prohibition on short selling:

$$\begin{aligned}
& \sum_{i=1}^N w_i = 1 \\
& w_i \geq 0, i = 1, \dots, N
\end{aligned} \tag{5.45}$$

To control the capacity of the set of admissible portfolios that is used by the optimization procedure, a constraint is also placed on the 2-norm:<sup>299</sup>

$$\frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C, \tag{5.46}$$

where the parameter  $C$  is a pre-selected constant, and the matrix  $\mathbf{S}$  and vector  $\mathbf{w}_0$  define the preliminary structure of the tracking portfolio. We discuss possible structures in Section 5.2.3.

This formulation has two important properties. The first is that SVR is robust to outliers, since it is based on the  $\varepsilon$ -insensitive loss function. The second is that complexity is controlled for using the penalty parameter  $C$ , which makes the solution of the SVR less sensitive to the data used for estimation, improving the robustness of the solution.

Before applying SVR to solving the index tracking problem, it is necessary to specify the parameters  $\varepsilon$  and  $C$ , and the preliminary structure of the tracking portfolio, expressed by the matrix  $\mathbf{S}$  and the vector  $\mathbf{w}_0$ . Next, we discuss the different ways of parameterizing SVR.

---

<sup>298</sup> See Section 3.3.1 for details.

<sup>299</sup> See Section 5.1.3 for details.

## 5.2.2 Selecting the size of the insensitive region

There are two methods of selecting the size of the insensitive region. The first is based on the minimax approach, the second on  $\nu$ -SVR.

### 5.2.2.1 Minimax approach

The minimax approach (see Section 3.3.2) involves the minimization of the maximal deviation between the index and the tracking portfolio:

$$\begin{aligned} & \min_{\mathbf{w}} \max_{t=1, \dots, T} |R_{I,t} - R_{P,t} - b| \\ & \text{subject to} \quad \sum_{i=1}^N w_i = 1, \end{aligned}$$

where  $R_{P,t} = \sum_{i=1}^N w_i r_{i,t}$ .

This optimization problem can be rewritten in the following equivalent form, where the inequality constraints determine the maximal deviation between the index and the tracking portfolio:

$$\begin{aligned} & \underset{\mathbf{w}, b, \theta}{\text{minimize}} \quad \theta \\ & \text{subject to} \quad R_{I,t} - \sum_{i=1}^N w_i r_{i,t} - b \leq \theta \\ & \quad \quad \quad R_{I,t} - \sum_{i=1}^N w_i r_{i,t} - b \geq -\theta \\ & \quad \quad \quad \sum_{i=1}^N w_i = 1 \\ & \quad \quad \quad t = 1, \dots, T. \end{aligned} \tag{5.47}$$

The solution  $\theta^*$  of this problem is the minimal value of the largest deviation between the index and the tracking portfolio. As a consequence, for all  $t = 1, \dots, T$ , the inequalities

$$-\theta^* \leq R_{I,t} - R_{P,t} - b \leq \theta^*$$

are satisfied.

Using the optimal value  $\theta^*$ , it is possible, by relaxing the minimax approach, to formulate the SVR (see Figure 5.22).

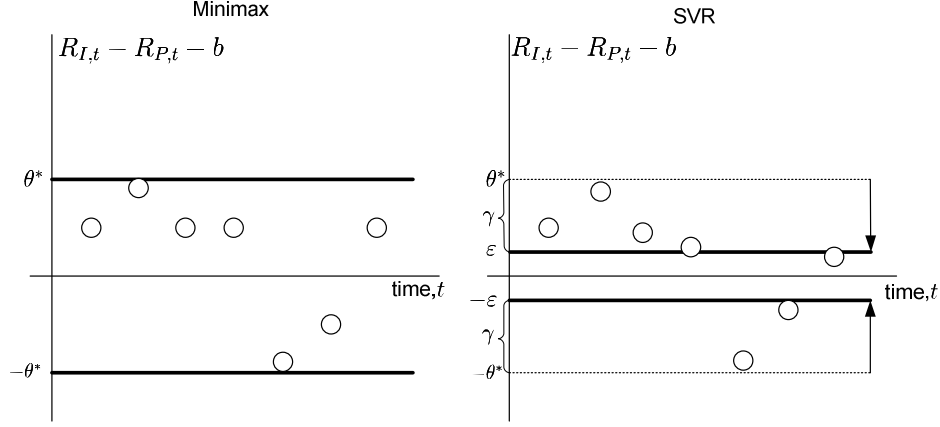


Figure 5.22: SVR as a Relaxation of the Minimax Approach

To do this, we introduce a margin  $\gamma$ , which is less than or equal to  $\theta^*$  (see right side of Figure 5.22). The optimal margin  $\gamma$  can be found by using Proposition 5.1 from Section 5.1.6 to transform the SVR into a SVM. Hence, to solve the index tracking problem we need to solve the following classification problem in the form of a SVM:

$$\begin{aligned}
 & \underset{\mathbf{w}^c, b^c, \xi^c, \xi^{*c}}{\text{minimize}} && \sum_{t=1}^T (\xi_t^c + \xi_t^{*c}) \\
 & \text{subject to} && \sum_{i=1}^N w_i^c (R_{I,t} - r_{i,t} + \theta) - b^c \geq 1 - \xi_t^c \\
 & && \sum_{i=1}^N w_i^c (R_{I,t} - r_{i,t} - \theta) - b^c \leq -1 + \xi_t^{*c} \\
 & && \sum_{i=1}^N w_i^c \geq \frac{1}{\theta^*} \\
 & && w_i^c \geq 0, i = 1, \dots, N \\
 & && \xi_t^c, \xi_t^{*c} \geq 0 \\
 & && \frac{1}{2} \|\mathbf{S}(\mathbf{w}^c - \mathbf{w}_0 \sum_{i=1}^N w_i^c)\|_2^2 \leq C^c \\
 & && t = 1, \dots, T.
 \end{aligned} \tag{5.48}$$

The margin  $\gamma$  is automatically determined by solving this optimization problem, and is given by  $\gamma = \frac{1}{\sum_{i=1}^N w_i^c}$ . The weights for the tracking portfolio are computed using the following expression:

$$w_i = \frac{w_i^c}{\sum_{i=1}^N w_i^c}.$$

The size of the insensitive region is given by



$$\varepsilon = \theta^* - \gamma = \theta^* - \frac{1}{\sum_{i=1}^N w_i^c}.$$

### 5.2.2.2 $\nu$ -SVR

An alternative method of determining the size of the insensitive region is to use  $\nu$ -SVR (see Section 5.1.3.2). The index tracking problem in this case is formulated as:

$$\begin{aligned} \underset{\mathbf{w}, \xi, \xi^*, \varepsilon, b}{\text{minimize}} \quad & \nu\varepsilon + \frac{1}{T} \sum_{t=1}^T (\xi_t + \xi_t^*) \\ \text{subject to} \quad & R_{I,t} - \sum_{i=1}^N w_i r_{i,t} - b \leq \varepsilon + \xi_t \\ & \sum_{i=1}^N w_i r_{i,t} + b - R_{I,t} \leq \varepsilon + \xi_t^* \\ & \sum_{i=1}^N w_i = 1 \\ & \frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C \\ & w_i \geq 0 \\ & \xi_t, \xi_t^* \geq 0. \end{aligned} \tag{5.49}$$

The parameter  $\nu$  is selected from the interval  $[0, 1]$ . This parameter specifies the percentage of support vectors that are data points that are allowed to be outside the insensitive region.

### 5.2.3 Preliminary informational structures for the tracking portfolio

To solve the optimization problem for the index tracking, and to choose the size of the insensitive region, it is first necessary to specify the preliminary structure of the tracking portfolio, which is determined by the matrix  $\mathbf{S}$  and the vector  $\mathbf{w}_0$  according to constraint (5.46). There are various ways of doing this.

The constraint on the 2-norm defines an ellipsoid with its centre at the point  $\mathbf{w}_0$  in the space of portfolio weights (see left side of Figure 5.23).

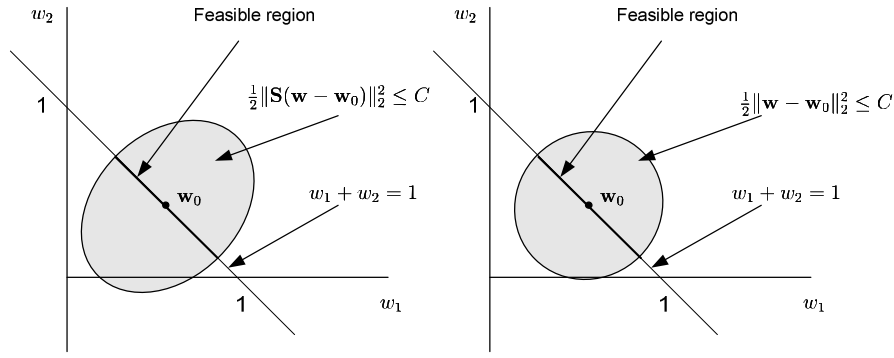


Figure 5.23: Ellipsoidal and Spherical Constraints in 2-Dimensional Space

The constant  $C$  sets the size of the ellipsoid. The feasible region for the optimization problem, given the constraints on the weights (5.45), is an interval. The ellipsoidal constraint can take the special form of a spherical constraint when the matrix  $\mathbf{S} = \mathbf{I}$  (i.e. an identity matrix, see right side of Figure 5.23).

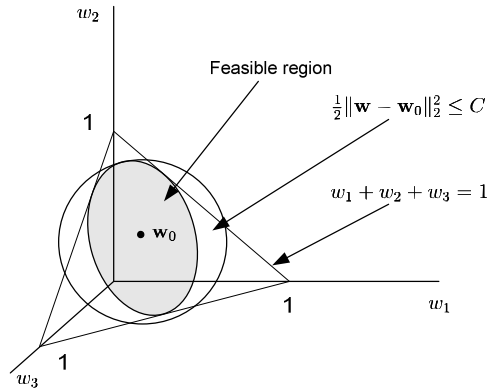


Figure 5.24: Spherical Constraint in 3-Dimensional Space

The feasible region in 3-dimensional space is an ellipse when there is an ellipsoidal constraint and a sphere when there is a spherical constraint (see Figure 5.24). The ellipse or the sphere lies on the plane defined by the constraint  $\sum_{i=1}^N w_i = 1$ .

In spaces with more dimensions, the feasible region is a hypersphere under the spherical constraint and an ellipsoid under the ellipsoidal constraint. The budget constraint is represented by a hyperplane.

Using the constraint (5.46), we can formulate the different sources of information that determine the preliminary structure of the tracking portfolio. These sources are

- an equal weighted portfolio
- the current tracking portfolio holdings
- the fundamentals of the company
- the market capitalization
- the sector structure

- the covariance structure
- mixed structures.

### Equal weighted portfolio

If there is no additional information, a useful assumption is that all assets should contribute equally to the portfolio. Hence, it is possible to define a vector  $\mathbf{w}_0$ , where all components are equal to  $\frac{1}{N}$ . This case corresponds to a naïve diversification of the portfolio. Hence, by controlling the constant  $C$ , it is possible to control the distance of the tracking portfolio from the equal weighted portfolio. Matrix  $\mathbf{S}$  is an identity matrix in this case.

### Current tracking portfolio holdings

The information about the current portfolio holdings, again included in the optimization problem using vector  $\mathbf{w}_0$ , can help to reduce the transaction costs, since the goal is to minimize the difference between the current portfolio and the new portfolio. Hence, by controlling constant  $C$  it is possible to keep the transaction costs within a particular budget. Matrix  $\mathbf{S}$  is again an identity matrix.

### Fundamentals of the company

Incorporating the fundamentals of the company, for example, its book value, revenue, sales or dividends, reduces the sensitivity of the solution to the prices of the assets. This information can be encoded into the vector  $\mathbf{w}_0$  in the same way as it is done in the fundamental indices, by weighting the components according to their fundamental values:

$$w_{0,i} = \frac{F_{i,T}}{\sum_{j=1}^N F_{j,T}},$$

where  $F_{i,T}$  is the value of the fundamental factor for company  $i$  at the end of the estimation period. Once again, matrix  $\mathbf{S}$  is an identity matrix.

### Market capitalization

The market capitalization is another source of information that can be plugged into the problem using vector  $\mathbf{w}_0$ . The components of the vector in this case are calculated using the formula

$$w_{0,i} = \frac{f_i S_{i,T}}{\sum_{j=1}^N f_j S_{j,T}}, \quad (5.50)$$

where  $f_i$  is the free-float adjusted number of stocks (in company)  $i$  outstanding, and  $S_{i,T}$  is the price of stock  $i$  at the end of the estimation period. The free-float adjustment is an adjustment of the number of stocks outstanding, to include only those available on the market. Matrix  $\mathbf{S}$  is an identity matrix again.

### Sector structure

The sector structure of the tracking portfolio is specified by introducing new variables  $z_j, j = 1, \dots, M$ , defined as follows:

$$z_j = \sum_{i=1}^N \delta_{i,j} w_i, j = 1, \dots, M, \quad (5.51)$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if company } i \text{ belongs to sector } j, \\ 0 & \text{otherwise.} \end{cases}$$

The index structure is defined as

$$z_{0,j} = \sum_{i=1}^N \delta_{i,j} w_{0,i}, j = 1, \dots, M,$$

where  $w_{0,i}$  is the weight of company  $i$  in the index. This weight could, for example, be based on market capitalization or fundamentals.

Using vectors  $\mathbf{z}$  and  $\mathbf{z}_0$ , we can define the constraint as

$$\frac{1}{2} \|\mathbf{z} - \mathbf{z}_0\|_2^2 \leq C. \quad (5.52)$$

Hence, by adding constraints (5.51) and (5.52), we can include the sector structure in the optimization problem.

### Covariance structure

It is possible to define the matrix  $\mathbf{S}$  using the covariance matrix of the asset returns. This matrix can be decomposed using Cholesky decomposition as follows:<sup>300</sup>

---

<sup>300</sup> See Fabozzi et al. (2007, p. 315).

$$\Sigma = \mathbf{S}'\mathbf{S}. \quad (5.53)$$

We can extend all of the above cases by defining the matrix  $\mathbf{S}$  in this way. As a consequence, we can reduce or increase the influence of particular weights on the solution of the optimization problem.

### Mixed structures

It is also possible to combine the above-described structures and formulate an index tracking problem that incorporates different types of information.

#### 5.2.4 Structural tracking error minimization (STEM)

The STEM approach that we define in this section is used to select the parameter  $C$ , which controls the capacity of the set of admissible portfolios used in the optimization procedure.

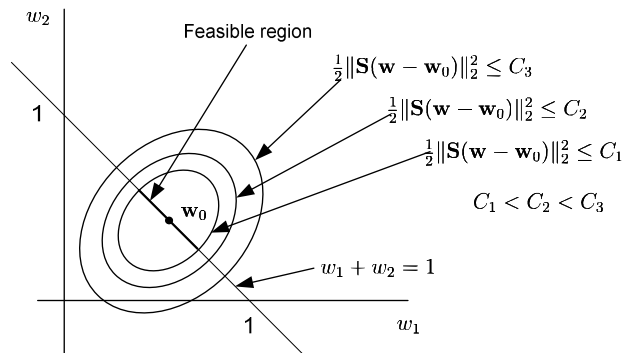


Figure 5.25: The Nested Structure for STEM

The constraint in the general form (5.46) specifies a trade-off between the informational structures and the data used for the estimation. The more sure we are about the preliminary portfolio weights  $\mathbf{w}_0$ , the lower our constant  $C$  should be. Using this constraint, it is possible to formulate the STEM approach. Defining a sequence of constants  $C$  as

$$C_1 < C_2 < C_3 < \cdots < C_l,$$

we determine the nested informational structure for the optimization problem (see Figure 5.25).

The smaller is the constant  $C$ , the smaller is the number of possible combinations of portfolio weights. In other words, the capacity of the set of admissible portfolios becomes lower as we decrease the constant  $C$ . In the extreme case, with  $C = 0$ , the set of admissible portfolios

reduces to just one, with  $\mathbf{w} = \mathbf{w}_0$ . The capacity of the set of admissible portfolios can be measured using the effective degrees of freedom (DoF) or by the effective VC dimension (see Section 5.1.4.4).

STEM approach is based on the SRM principle, which was introduced by *Vapnik* in statistical learning theory (see Section 5.1 and Section 5.1.4.2 for details). Using SRM principle one minimizes both the empirical error and the confidence interval, which depends on the capacity of the set of admissible portfolios. As it is proved by *Lugosi and Zeger* and described in Section 5.1.4.2, the solution of SRM method converges to the best possible solution with probability one. Hence, the learning using SRM approach is consistent.<sup>301</sup>

The optimal constant  $C$ , which determines the optimal complexity for modelling the tracking portfolio, can be selected using analytic methods of model selection (see Section 5.1.4.3), based on the DoF and the VC dimension, or using resampling methods (see Section 5.1.4.5). The model selection process is described in Section 5.1.4.7.

## 5.2.5 Further modifications of the SVR

The SVR for the index tracking problem (equations (5.44), (5.45) and (5.46)) can be extended by using monotonic convex functions  $F(\xi_t)$ ,  $F^*(\xi_t^*)$  to penalize outliers:<sup>302</sup>

$$\begin{aligned}
& \underset{\mathbf{w}, b, \xi, \xi^*}{\text{minimize}} && \sum_{t=1}^T (F(\xi_t) + F^*(\xi_t^*)) \\
& \text{subject to} && R_{I,t} - \sum_{i=1}^N w_i r_{i,t} - b \leq \varepsilon + \xi_t \\
& && \sum_{i=1}^N w_i r_{i,t} + b - R_{I,t} \leq \varepsilon + \xi_t^* \\
& && \sum_{i=1}^N w_i = 1 \\
& && \frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C \\
& && w_i \geq 0 \\
& && \xi_t, \xi_t^* \geq 0.
\end{aligned} \tag{5.54}$$

Before solving the index tracking problem (5.54), we need to select the functions  $F(\xi_t)$ ,  $F^*(\xi_t^*)$ , the size of the insensitive region  $\varepsilon$ , the informational structure  $\mathbf{S}$ ,  $\mathbf{w}_0$ , and the

---

<sup>301</sup> In Section 5.1 we define the concept of the learning consistency.

<sup>302</sup> See Cortes/Vapnik (1995, p. 281).

trade-off parameter  $C$ . Then, using the general form (5.54), it is possible to implement different variants of optimized sampling.

For example, SVR is equivalent to MAD if the informational constraint is omitted and

$$\begin{aligned} F(\xi_t) &= \xi_t \\ F(\xi_t^*) &= \xi_t^* \\ \varepsilon &= 0 \\ b &= 0. \end{aligned}$$

It is equivalent to MSE if the informational constraint is omitted and

$$\begin{aligned} F(\xi_t) &= \xi_t^2 \\ F(\xi_t^*) &= \xi_t^{2*} \\ \varepsilon &= 0 \\ b &= 0. \end{aligned}$$

Moreover, the traditional methods can be extended further, into informational strategies, by using the informational structures from Section 5.2.3.

The traditional methods can be extended further by introducing the insensitive region, which is defined by the parameter  $\varepsilon > 0$ . For example, the minimization of MAD can be extended by an  $\varepsilon$ -insensitive loss function (see Figure 5.6), and the minimization of MSE can be extended by an  $\varepsilon$ -insensitive loss function with quadratic loss (see Figure 5.26).

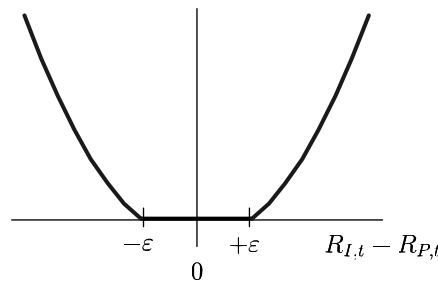


Figure 5.26:  $\varepsilon$ -Insensitive Loss Function with a Quadratic Loss

Moreover, SVR can be used for the asymmetric case<sup>303</sup>, for example for penalizing only downside deviations of the tracking portfolio from the index (see Figure 5.27).

---

<sup>303</sup> See Smola (1998, p. 18).

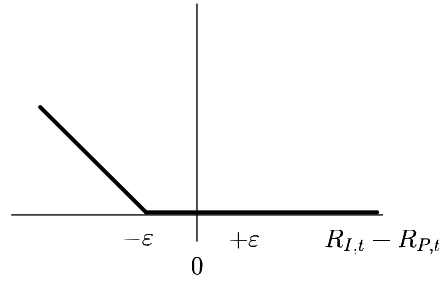


Figure 5.27:  $\varepsilon$ -Insensitive Loss Function with a Downside Loss

In this case, SVR is equivalent to MADD (mean absolute downside deviation)<sup>304</sup> if the informational constraint is omitted and

$$\begin{aligned} F(\xi_t) &= 0 \\ F(\xi_t^*) &= \xi_t^* \\ \varepsilon &= 0 \\ b &= 0. \end{aligned}$$

Summarizing, the general form of SVR can be used to formulate different modifications of optimized sampling.

---

<sup>304</sup> See Rudolf/Wolter/Zimmermann (1999, p. 88).



## 5.3 Comparison of STEM-based methods with other approaches

In this section we compare STEM-based methods of index tracking with other approaches to portfolio optimization and index tracking. This allows us to identify links of STEM with other methods.

### 5.3.1 Mean-variance portfolio optimization

The idea of solving the problem of portfolio optimization by adding a constraint on the 2-norm  $\|\mathbf{w}\|_2$  of portfolio weights is implicitly implemented in the Markowitz mean-variance framework.<sup>305</sup> The optimization problem can be written as

$$\begin{aligned} & \underset{\mathbf{w}}{\text{maximize}} && \bar{\boldsymbol{\mu}}' \mathbf{w} \\ & \text{subject to} && \mathbf{1}' \mathbf{w} = 1 \\ & && \frac{1}{2} \mathbf{w}' \boldsymbol{\Sigma} \mathbf{w} \leq C, \end{aligned} \tag{5.55}$$

where  $\bar{\boldsymbol{\mu}}$  is the vector of expected returns and  $\boldsymbol{\Sigma}$  is the covariance matrix of returns.

Using the Cholesky decomposition, the constraint on the variance of the portfolio returns can be written as follows (see equation (5.53)):

$$\frac{1}{2} \mathbf{w}' \boldsymbol{\Sigma} \mathbf{w} = \frac{1}{2} \|\mathbf{S} \mathbf{w}\|_2^2 \leq C$$

Moreover, *Markowitz and Todd*<sup>306</sup> use the mean-variance framework to solve the index tracking problem. The constraint on the portfolio variance is replaced by the TEV:

$$TEV = \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)' \boldsymbol{\Sigma} (\mathbf{w} - \mathbf{w}_0) = \frac{1}{2} \|\mathbf{S} (\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C, \tag{5.56}$$

where the vector  $\mathbf{w}_0$  represents the weights in the benchmark. A constraint of this form limits the variance of the tracking portfolio around the index.

STEM-based methods of index tracking use the same form of constraint as (5.56) to include the required information in the optimization problem. However, it is possible to include different sources of information from Section 5.2.3.

---

<sup>305</sup> See Markowitz (1952).

<sup>306</sup> See Markowitz/Todd (1987, p. 50).

Moreover, STEM-based methods can be used not only for index tracking but also for portfolio optimization, using the same objective as in the optimization problem (5.55) and combining it with the different sources of information described in Section 5.2.3.

### 5.3.2 Generalized framework for portfolio optimization

The mean-variance approach is extended further by *DeMiguel et al.*<sup>307</sup> They develop a generalized framework for portfolio optimization, with the focus on the performance in the investment period. They solve the following minimal-variance portfolio optimization problem:

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \mathbf{w}'\Sigma\mathbf{w} \\ & \text{subject to} && \mathbf{1}'\mathbf{w} = 1. \end{aligned} \tag{5.57}$$

This problem is solved by placing additional constraints on the portfolio weights.<sup>308</sup>

The first constraint is a 1-norm constraint (see Figure 5.28):

$$\|\mathbf{w}\|_1 \leq C,$$

where  $\|\mathbf{w}\|_1 = \sum_{i=1}^N |w_i|$ .

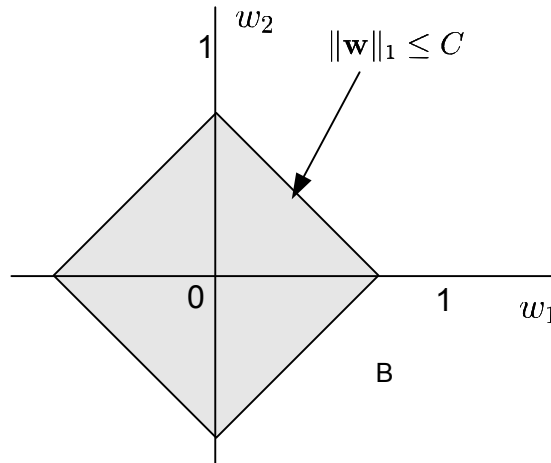


Figure 5.28: 1-Norm Constraint on the Portfolio Weights

The second constraint is a 2-norm constraint:

<sup>307</sup> See DeMiguel et al. (2009, p. 809).

<sup>308</sup> See DeMiguel et al. (2009, pp. 802-804).

$$\|\mathbf{w}\|_2^2 \leq C,$$

where  $\|\mathbf{w}\|_2^2 = \sum_{i=1}^N w_i^2$ .

The third constraint is an A-norm constraint:

$$\mathbf{w}'\mathbf{A}\mathbf{w} \leq C,$$

where  $\mathbf{A}$  is a matrix.

The generalized framework assumes that there is no additional information about the weights of the portfolio. It can be shown that, if  $\mathbf{1}'\mathbf{w} = 1$ , then<sup>309</sup>

$$\|\mathbf{w} - \frac{1}{N}\mathbf{1}\|_2^2 = \|\mathbf{w}\|_2^2 - \frac{1}{N}.$$

As a result, by limiting the 2-norm of the portfolio weights, one can impose a limit on the distance from the equal weighted portfolio.

The parameter  $C$  is chosen using cross-validation.<sup>310</sup> The empirical results in *DeMiguel et al. (2009)* demonstrate that, in the investment period, the norm-constrained portfolios have a higher Sharpe ratio than the other portfolios examined.

This approach can also be applied to solving the index tracking problem if, instead of minimizing the variance of the tracking portfolio in (5.57), we minimize one of the measures of tracking quality described in Section 3.3.

The comparison of this approach with STEM-based methods of index tracking indicates that the informational constraints for the STEM-based methods have a more general form, which includes this case as a special case when there is no information available about the weights. Moreover, there is a theoretical justification for the STEM-based methods provided by statistical learning theory in that, by constraining the norms of the portfolio weights, one controls the capacity of the set of admissible portfolios, which leads to robust tracking portfolio estimation.

---

<sup>309</sup> See DeMiguel et al. (2009, p. 803).

<sup>310</sup> See DeMiguel et al. (2009, p. 805).

### 5.3.3 Portfolio optimization with a chance constraint

The next approach extends the mean-variance approach and handles the estimation error in the vector of expected returns  $\boldsymbol{\mu}$  with covariance matrix  $\boldsymbol{\Sigma}$  (see Figure 5.29). This vector is itself a random vector. This means that estimating it using different datasets gives different results. One assumes that this vector is normally distributed with mean  $\bar{\boldsymbol{\mu}}$ . One assumes further that the random vector of estimation errors is also normally distributed with zero mean and covariance matrix  $\tau\boldsymbol{\Sigma}$ , which is proportional to  $\boldsymbol{\Sigma}$ , where the scalar  $\tau$  satisfies  $0 \leq \tau \leq 1$ . Then, the vector  $\boldsymbol{\mu}$  follows a multivariate normal distribution with mean  $\bar{\boldsymbol{\mu}}$  and covariance matrix  $\tau\boldsymbol{\Sigma}$ .<sup>311</sup>

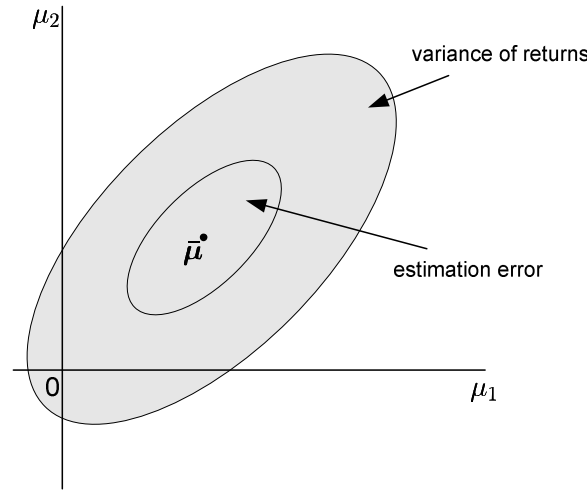


Figure 5.29: Estimation Error for the Vector of Expected Returns

Then, it is possible to formulate the portfolio optimization problem with a constraint on the expected returns that takes into account an estimation error. This constraint can be formulated as<sup>312</sup>

$$Pr(\boldsymbol{\mu}'\mathbf{w} \leq \alpha) \leq \beta, \quad (5.58)$$

where  $\beta$  is the probability that the portfolio returns are smaller than the level  $\alpha$ . Portfolio optimization with a chance constraint can then be formulated as

$$\begin{aligned} & \underset{\mathbf{w}}{\text{maximize}} && \bar{\boldsymbol{\mu}}'\mathbf{w} \\ & \text{subject to} && \mathbf{1}'\mathbf{w} = 1 \\ & && Pr(\boldsymbol{\mu}'\mathbf{w} \leq \alpha) \leq \beta. \end{aligned} \quad (5.59)$$

<sup>311</sup> See Black/Litterman (1992).

<sup>312</sup> See Lobo et al. (1998, p. 214).

The chance constraint (5.58) can be reformulated as a second-order cone programming (SOCP) constraint:<sup>313</sup>

$$\bar{\boldsymbol{\mu}}'\mathbf{w} + \Phi^{-1}(\beta)\|\sqrt{\tau}\mathbf{S}\mathbf{w}\|_2 \geq \alpha,$$

where  $\Phi(x)$  is the cumulative distribution function of the standard normal distribution and  $\mathbf{S}$  is the square root of the covariance matrix  $\boldsymbol{\Sigma}$ , which can be computed using the Cholesky decomposition. The constraint is convex if  $\beta \leq 0.5$ . In this case,  $\Phi^{-1}(\beta) \leq 0$ .

As a result, the SOCP constraint can be rewritten as

$$\|\sqrt{\tau}\mathbf{S}\mathbf{w}\|_2 \leq \frac{\alpha - \bar{\boldsymbol{\mu}}'\mathbf{w}}{\Phi^{-1}(\beta)}.$$

This inequality is feasible only if  $\alpha - \bar{\boldsymbol{\mu}}'\mathbf{w} \leq 0$ , since  $\Phi^{-1}(\beta) \leq 0$ . In this case, the right hand side of the inequality is greater than or equal to zero. Squaring the two parts of the inequality gives

$$\begin{aligned} \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} &\leq \frac{1}{\tau} \left( \frac{\alpha - \bar{\boldsymbol{\mu}}'\mathbf{w}}{\Phi^{-1}(\beta)} \right)^2 \\ \bar{\boldsymbol{\mu}}'\mathbf{w} &\geq \alpha \end{aligned}$$

Analysing the reformulated SOCP constraint, we can see that we have limited the variance of the portfolio. When combined with the objective of maximizing  $\bar{\boldsymbol{\mu}}'\mathbf{w}$ , the robust optimization problem and the Markowitz problem are equivalent.<sup>314</sup>

The optimization problem (5.59) can be used to solve the index tracking problem if the chance constraint is rewritten as

$$Pr(\boldsymbol{\mu}'(\mathbf{w} - \mathbf{w}_0) \leq \alpha) \leq \beta, \quad (5.60)$$

where the vector  $\mathbf{w}_0$  represents the weights in the benchmark. This constraint means that the probability that the difference between the expected returns of the tracking portfolio and those of the benchmark is lower than  $\alpha$  is less than  $\beta$ .

The constraint (5.60) can be written as the SOCP constraint

---

<sup>313</sup> See Lobo et al. (1998, p. 215).

<sup>314</sup> The equivalence to the Markowitz approach has also been observed by Ben-Tal/Nemirovski (1999).

$$\|\sqrt{\tau}\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2 \leq \frac{\alpha - \bar{\boldsymbol{\mu}}'(\mathbf{w} - \mathbf{w}_0)}{\Phi^{-1}(\beta)}.$$

In this form the constraint on the 2-norm of the portfolio weights limits the estimation error. As a consequence, STEM-based methods of solving the index tracking problem by limiting the norm of the portfolio weights also reduce the estimation error.

### 5.3.4 Conditional value-at-risk

CVAR, which was described in Section 3.3.3, is very similar to  $\nu$ -SVR, which was described in Section 5.2.2.2. Comparing the problem for CVAR (3.7) with that for  $\nu$ -SVR (5.49), one can see that both would have the same objective function if that in the CVAR problem was multiplied by  $1 - \alpha > 0$  and the expression  $(1 - \alpha)$  was then replaced with  $\nu$ .<sup>315</sup>

The only difference is that when  $\nu$ -SVR is used for index tracking it has an informational constraint, which allows the user to control the capacity of the set of admissible portfolios. Hence,  $\nu$ -SVR can be seen as a robust form of CVAR. An analysis of CVAR with the chance constraint from Section 5.3.3 included can be found in *Zhu et al.*<sup>316</sup>

### 5.3.5 Box constraints

Information can be incorporated into the optimization problem, not only using ellipsoidal constraints as in the STEM-based methods, but also using box constraints.

These constraints take the form

$$w_i^d \leq w_i \leq w_i^u, i = 1, \dots, N. \quad (5.61)$$

They ensure that the stock weights remain between a lower and an upper level. For example, if an investor holds a portfolio with weights  $\mathbf{w}_0$ , and does not want to move too far away from this portfolio, then the upper and lower limits can be defined as

$$w_i^d = w_{0,i} - c, w_i^u = w_{0,i} + c, i = 1, \dots, N,$$

where  $c > 0$  is the maximal permitted distance from the original weights.

---

<sup>315</sup> The same result can be found in Takeda/Gotoh/Sugiama (2011).

<sup>316</sup> See Zhu/Coleman/Li (2009, p. 64).

The most commonly used constraints ensure the prohibition of short selling:

$$w_i \geq 0, i = 1, \dots, N.$$

These constraints are a special case of the constraints from (5.61). The 1-norm constraint (see Figure 5.28) also belongs to the box constraints category. Moreover, it is possible to formulate robust portfolio optimization using box constraints.<sup>317</sup>

We now illustrate the constraints (5.61) geometrically. First, consider a portfolio consisting of two stocks. The feasible region for the optimization problem is a line in two-dimensional space (see left side of Figure 5.30). For a portfolio consisting of three stocks, the feasible region is a polygon (see right side of Figure 5.30). In a space with more dimensions, the region would be a polytope.

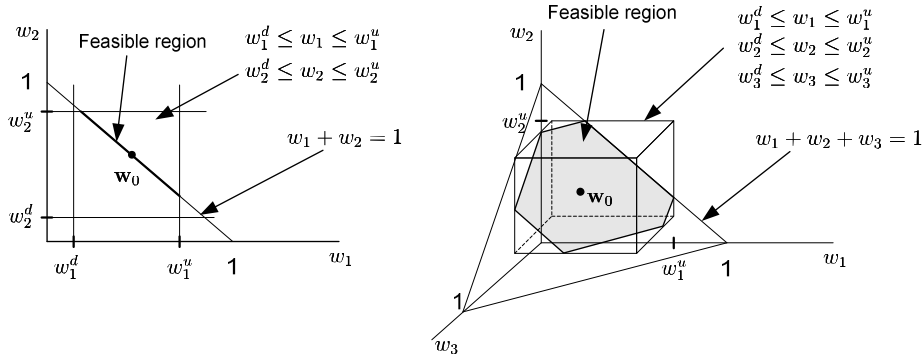


Figure 5.30: Box Constraints

It is well known that adding box constraints reduces the estimation error and improves the robustness of the resulting portfolios.<sup>318</sup> However, there is no explanation for this empirical fact. We have developed the STEM-based methods, in which we also add constraints on the weights of the assets in the tracking portfolio. In this case, adding the constraints can be justified using statistical learning theory. Namely, adding the constraints reduces the capacity of the set of admissible tracking portfolios; as a result, it is possible to find a bias-variance trade-off and, as a consequence of that, reduce the test error.

<sup>317</sup> See Cornuejols/Tütüncü (2007, p. 313).

<sup>318</sup> See Jagannathan/Ma (2003), Chopra (1993), Eichhorn/Gupta/Stubbs (1998), Frost/Savarino (1988), and Grauer/Shen (2000).

## 5.4 Empirical analysis of STEM

In this section, the empirical analysis of the STEM approach is carried out. First, we test the resistance of the tracking portfolio estimation, achieved using STEM-based methods, to different training sets. Second, we demonstrate the application of the methods of selecting the model of optimal complexity. Third, we analyse the absolute and relative stability of the tracking quality of tracking portfolios estimated using STEM-based methods.

### 5.4.1 Resistance of tracking portfolio estimation using STEM

#### 5.4.1.1 Research method

This section complements the results from Section 4.3, in which we tested the resistance of tracking portfolio estimation to different training sets, using traditional methods of sampling. Here, we analyse eight different methods of sampling using STEM, as displayed in Table 5.1.

I	II		III
SVM_EW	CVAR_EW	CVAR2_EW	MSE_EW
SVM_CW	CVAR_CW	CVAR2_CW	MSE_CW

Table 5.1: STEM-Based Methods of Optimized Sampling

The first group of methods is that developed in Section 5.2.2.1, namely, optimization problem (5.48). It is a relaxation of a minimax approach and is based on SVM. Two variants of this method are considered, one using an equal weighted portfolio (EW) and the other a capital weighted portfolio (CW) as the structure. In the first variant, the weights vector  $\mathbf{w}_0$  is the vector  $(1/N, \dots, 1/N)'$ . In the second variant,  $\mathbf{w}_0$  consists of capital weights, which are calculated using (5.50). For both structures, the matrix  $\mathbf{S}$  is the identity matrix. In applying this method, initially the MAXD problem (5.47) is solved to determine the parameter  $\theta^*$ , which is then used to convert SVR to SVM (5.48) using Proposition 5.1.

The second group of methods is  $\nu$ -SVR, which is described in Section 5.2.2.2. For this method it is necessary to pre-select the value of the parameter  $\nu$ . We select two values for this parameter:  $\nu = 0.05$  and  $\nu = 0.5$ . The parameter  $\nu = 0.05$  indicates that we control the largest 5% of deviations from the index.<sup>319</sup> Using this parameter in the objective means that the method only pays attention to these largest deviations of the tracking portfolio returns from the index returns. Using the parameter  $\nu = 0.5$  ensures that we control half of the deviations (the largest ones), which provides a robust trade-off between approximation and

---

<sup>319</sup> See Rockafellar/Uryasev (2002, p. 1463). To obtain their empirical results, they use  $\nu = 0.1$ .



robustness.<sup>320</sup> We call the method using  $\nu = 0.05$  CVAR and that where  $\nu = 0.5$  CVAR2. As above, we use two structures for each method: EW and CW.

The third group of methods is MSE. This is the most popular optimization objective in sampling. We combine the minimization of MSE with the EW and CW structures, in the form of constraint (5.46), to analyse whether the STEM approach is able to improve the resistance of the tracking portfolio estimation to modifications of the training set.

All of the above methods use a tracking error with constant term  $b$ , where  $b$  is the constant difference between the returns of the index and those of the tracking portfolio (see Section 3.3.1 for details). Allowing optimization procedure to determine shift  $b$ , we concentrate on the minimizing the variability of the tracking portfolio around the index.

For all the experiments in this section, we select 120 assets for the tracking portfolio for each period, using the stratified sampling procedure described in Section 4.3. To test the resistance of the tracking portfolio estimation, we use the same experiment design as in Section 4.3.1, Figure 4.8. We generate 1,000 training sets for each time period, and test the resistance of the estimations as we vary parameter  $C$  in the 2-norm constraint (5.46), which determines the structure for STEM. The values of parameter  $C$  are taken from the following, exponentially growing sequence:<sup>321</sup>

$$C \in [2^{-20}, 2^{-19}, \dots, 2^{-1}]. \quad (5.62)$$

As a result, we can analyse the influence of parameter  $C$  on the tracking portfolio estimation.

#### 5.4.1.2 Empirical results

First we analyse whether the STEM approach is able to improve the resistance of the tracking portfolio estimation. The tracking portfolios are computed using the eight methods described above (see Table 5.1), for 1,000 modifications of the estimation set (see Figure 4.8 for details), for every value of the penalty parameter  $C$ , and for five time periods. In the estimation period, the tracking quality is measured by percentage RMSE. In the investment period, the tracking quality is measured by percentage ETQ. The results for the MSE method using the CW structure, for Period 5, are presented in Table 5.2.

---

<sup>320</sup> See Mattera/Haykin (1998, p. 226).

<sup>321</sup> Hsu/Chang/Lin (2009, p. 5) recommend the use of the exponentially growing sequence.

	C	Estimation, RMSE%				Investment, ETQ%			
		MSE_CW				MSE_CW			
		Med	IQR	Min	Max	Med	IQR	Min	Max
Period 5	$2^{-1}$	0.1842	0.0145	0.1498	0.2126	4.2697	0.9289	2.3461	6.6340
	$2^{-2}$	0.1842	0.0145	0.1498	0.2126	4.2699	0.9291	2.3462	6.6340
	$2^{-3}$	0.1842	0.0145	0.1498	0.2126	4.2700	0.9292	2.3462	6.6340
	$2^{-4}$	0.1842	0.0145	0.1498	0.2126	4.2697	0.9285	2.3458	6.6342
	$2^{-5}$	0.1842	0.0145	0.1498	0.2126	4.2697	0.9285	2.3458	6.6342
	$2^{-6}$	0.1843	0.0145	0.1498	0.2126	4.2695	0.9211	2.3458	6.3571
	$2^{-7}$	0.1914	0.0165	0.1549	0.2197	3.8664	0.7449	2.2684	5.5076
	$2^{-8}$	0.2156	0.0179	0.1728	0.2424	3.1815	0.5926	1.7116	4.5397
	$2^{-9}$	0.2523	0.0187	0.2049	0.2783	2.2208	0.4742	0.9404	3.1856
	$2^{-10}$	0.2975	0.0182	0.2463	0.3267	1.0135	0.3344	0.2487	1.7481
	$2^{-11}$	0.3481	0.0184	0.2939	0.3791	0.3889	0.1315	0.2546	0.9053
	$2^{-12}$	0.3983	0.0188	0.3368	0.4300	1.4635	0.1356	1.1615	1.7895
	$2^{-13}$	0.4416	0.0199	0.3734	0.4776	2.3682	0.0930	2.1636	2.5647
	$2^{-14}$	0.4768	0.0205	0.4033	0.5159	3.0420	0.0631	2.9008	3.1789
	$2^{-15}$	0.5038	0.0215	0.4266	0.5453	3.5330	0.0440	3.4368	3.6316
	$2^{-16}$	0.5241	0.0220	0.4443	0.5674	3.8826	0.0308	3.8155	3.9535
	$2^{-17}$	0.5389	0.0226	0.4574	0.5836	4.1296	0.0214	4.0825	4.1799
	$2^{-18}$	0.5498	0.0229	0.4669	0.5953	4.3080	0.0148	4.2741	4.3432
	$2^{-19}$	0.5577	0.0232	0.4738	0.6037	4.4350	0.0104	4.4102	4.4601
	$2^{-20}$	0.5634	0.0234	0.4788	0.6097	4.5240	0.0072	4.5062	4.5421

Table 5.2: Reduction in the Variability of the Tracking Portfolio Estimation in the Investment Period, with Decreasing  $C$  (Period 5)

In the estimation period, by reducing the value of parameter  $C$ , we reduce the quality of the approximation of the index by the tracking portfolios. This is indicated by the median tracking quality (Med). Moreover, the variability of the tracking quality, which is expressed by the IQR, increases as we decrease the value of parameter  $C$ .

In the investment period, we see a different picture. The variability of the tracking quality decreases as parameter  $C$  decreases and, at the same time, the median tracking quality initially improves and only later starts to decline. Hence, it is possible to improve the resistance of the tracking portfolio estimation by reducing the value of parameter  $C$ , which limits the capacity of the set of models. Additionally, it is possible to find the optimal value of parameter  $C$ , so as to produce the best possible tracking quality in the investment period. Similar results are obtained for the other periods and other methods.

Further comparisons of the results in the estimation and investment periods show that the smallest value of the measure of tracking quality in the estimation period does not correspond to the best tracking quality in the investment period. Hence, it is not enough to analyse the tracking quality in the estimation period in order to select parameter  $C$  optimally. In Section 5.4.2 we use model selection methods for this purpose.

To analyse the potential of the STEM approach in comparison to the traditional methods, we report the minimal values of the median tracking qualities in the investment period by varying parameter  $C$ , for all methods, in Table 5.3. The results for the traditional methods that do not use STEM (CVAR, CVAR2, and MSE) are shown for reference.

	Structural Tracking Error Minimization								Traditional		
	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW	CVAR	CVAR2	MSE
1	0.4102	0.3341	0.4711	0.3324	0.3901	0.3246	0.3511	0.3178	0.4710	0.3901	0.3521
2	0.1798	0.2747	0.2036	0.2725	0.1631	0.3290	0.1668	0.2870	0.2725	0.3285	0.2963
3	0.4283	0.4559	0.4418	0.3435	0.4328	0.4211	0.4365	0.4265	0.4611	0.5332	0.5156
4	0.8154	0.8056	0.7252	0.7308	0.8584	0.8533	0.7444	0.7442	0.7284	0.8584	0.7464
5	3.9843	0.3817	4.6715	0.6854	3.7256	0.4648	4.2696	0.3889	5.0223	3.7256	4.8423

Table 5.3: Minimal Values of the Median ETQs for the Five Time Periods

The results in Table 5.3 suggest that the STEM-based methods are able to produce better tracking quality than the traditional methods (for each period, the lowest median tracking quality among the eight STEM methods is marked with grey). This is especially noticeable when comparing the tracking qualities for the turbulent Period 5, where the problem of overfitting is particularly acute. The median tracking quality of the SVM\_CW method in period 5 is approximately ten times better than the median tracking quality of the traditional methods (CVAR, CVAR2, and MSE).

### 5.4.1.3 Conclusion

We apply the STEM approach to one combination of assets for each period, selected using stratification techniques. Our analysis supports the hypothesis that, by controlling the capacity of the set of models, it is possible to improve the tracking quality in the investment period, and the resistance of the tracking portfolio estimation to modifications of the training set.

## 5.4.2 Selecting the model of optimal complexity

### 5.4.2.1 Research method

In Section 5.4.1, we show that, with an appropriate choice of parameter  $C$  in the 2-norm constraint on the portfolio weights (5.46), it is possible to improve the tracking quality in the investment period and reduce the dependence of the tracking portfolio estimation on the training set. The open question then is how to select this penalty parameter  $C$  optimally. The answer to this question involves the model selection process that was described in Section 5.1.4.7.

We use the model selection process to find parameter  $C$  from the range<sup>322</sup>

$$C \in [0, 2^{-16}, 2^{-15}, \dots, 2^{-6}, 2^{-5}, \infty].$$

<sup>322</sup> For the SVM method we use the constant  $C^c = \frac{C}{(\theta^*)^2}$ , since  $C^c = C(\sum_{i=1}^N w_i^c)^2 \geq \frac{C}{(\theta^*)^2}$  (see SVM (5.48)).

This range is a reduced version of the initial range (5.62). These values were selected based on the experiments presented in Section 5.4.1.

Two types of model selection methods can be used in the model selection process (for selecting penalty parameter  $C$ ): analytic and resampling. Below, we illustrate their usage and features.

### Analytic model selection methods

In Section 5.1.4.3 we described five analytic methods of model selection:

- final prediction error (FPE) (equation (5.32));
- general cross validation (GCV) (equation (5.33));
- Schwarz criterion (SC) (equation (5.34));
- Shibata's model selector (SMS) (equation (5.35));
- Vapnik measure (VM) (equation (5.36)).

Analytic methods use a measure of the complexity of the set of models to estimate the test error. There are two measures of complexity: degrees of freedom (DoF)  $d$  and the VC dimension  $h$ . These two complexity measures can be estimated using the data in the estimation period (see Section 5.1.4.4):

- The DoF  $d$  is estimated using the algorithm described in Section 5.1.4.4.1. We set the number of perturbations  $M$  of the index returns in the training set to be equal to 30. The perturbations are generated from the normal distribution with mean zero and standard deviation equal to  $\tau = 0.6\sigma$ , where  $\sigma$  is an estimation of the standard deviation of the index returns, based on the training set. The DoF  $d$  is then estimated using these perturbations.
- The VC dimension  $h$  is estimated using the procedure described in Section 5.1.4.4.2. For this procedure we randomly select  $K = 50$  data samples, which are used for the estimation of VC dimension  $h$ . These 50 data samples have sizes equal to  $s_k = T, T/2, T/3, T/4, T/5$ , where  $T$  is the size of the whole training set. For every size we generate 10 different data sets. Using these samples we estimate 50 variables  $d_k$  (equation (5.38)), which characterize the flexibility of the model set. These values are fitted by the function  $\Phi(s_k/h)$  (see equation (5.39)) for every value of VC dimension  $h = 1, \dots, N + 1$ . It is enough to fit the function  $\Phi(s_k/h)$  for these values of  $h$ , since the maximal VC dimension for linear models is equal to  $N + 1$ . As a result, the VC dimension is equal to  $h^*$ , which delivers the smallest fitting error.

The DoF  $d$  is then used in the FPE, GCV, SC and SMS model selection methods, and the VC dimension  $h$  is used in the VM method.

To demonstrate the use of analytic methods of model selection, we select the SVM\_CW method (see Table 5.4). We use the data from period 1 for the S&P1500 index (see Section 4.1.1 for details). We estimate the weights for one combination of assets, which are selected by the stratification procedure (see Section 4.3.1 for details), for different values of constant  $C$ .

	C	Emp. err., RMSE%	$d$	$h$	T	Correction factor					Estimation of test error, RMSE%					$\ w - w_0\ _2$	$\varepsilon$
						FPE	GCV	SC	SMS	VM	FPE	GCV	SC	SMS	VM		
<b>CW</b>	<b>0</b>	0.2068	1	1	261	1.00	1.00	1.01	1.00	1.11	0.2068	0.2068	0.2089	0.2068	0.2296	0.00025	0.00000
<b>SVM_CW</b>	$2^{-16}$	0.1914	2	2	261	1.01	1.01	1.02	1.01	1.14	0.1933	0.1933	0.1952	0.1933	0.2182	0.00552	0.00000
	$2^{-15}$	0.1854	2	1	261	1.01	1.01	1.02	1.01	1.11	0.1873	0.1873	0.1891	0.1873	0.2058	0.00781	0.00000
	$2^{-14}$	0.1772	3	4	261	1.01	1.01	1.03	1.01	1.20	0.1790	0.1790	0.1826	0.1790	0.2127	0.01105	0.00000
	$2^{-13}$	0.1672	4	6	261	1.02	1.02	1.04	1.02	1.24	0.1705	0.1705	0.1739	0.1705	0.2073	0.01563	0.00000
	$2^{-12}$	0.1548	6	3	261	1.02	1.02	1.06	1.02	1.17	0.1579	0.1579	0.1641	0.1579	0.1811	0.02210	0.00000
	$2^{-11}$	0.1395	10	7	261	1.04	1.04	1.11	1.04	1.26	0.1450	0.1450	0.1548	0.1450	0.1757	0.03125	0.00000
	$2^{-10}$	0.1223	14	9	261	1.06	1.06	1.15	1.05	1.29	0.1296	0.1296	0.1406	0.1284	0.1577	0.04419	0.00000
	$2^{-9}$	0.1073	20	10	261	1.08	1.08	1.21	1.07	1.31	0.1159	0.1159	0.1298	0.1148	0.1405	0.06040	0.00004
	$2^{-8}$	0.0977	26	21	261	1.11	1.11	1.27	1.10	1.48	0.1085	0.1085	0.1241	0.1075	0.1447	0.07705	0.00016
	$2^{-7}$	0.0914	35	24	261	1.14	1.15	1.36	1.13	1.52	0.1042	0.1051	0.1243	0.1032	0.1389	0.08950	0.00036
	$2^{-6}$	0.0866	42	35	261	1.18	1.19	1.44	1.15	1.68	0.1022	0.1031	0.1248	0.0996	0.1455	0.10172	0.00053
	$2^{-5}$	0.0864	42	36	261	1.18	1.19	1.44	1.15	1.69	0.1020	0.1029	0.1245	0.0994	0.1461	0.10869	0.00071
<b>MAXD</b>	$\infty$	0.0943	46	35	261	1.19	1.21	1.48	1.16	1.68	0.1123	0.1141	0.1396	0.1094	0.1585	0.13607	0.00124

Table 5.4: Analytic Model Selection Methods for the SVM\_CW Method (Period 1)

In the first column in Table 5.4 we report the penalty parameter  $C$ , which is used in the optimization problem (5.48) for the SVM\_CW method. Using the returns in the training set we calculate the empirical error for each parameter  $C$ . As expected, the empirical error reduces as we increase the parameter  $C$ . It is possible to select parameter  $C$  so that the empirical error is the smallest but this can cause overfitting, since the values of the estimated capacity measures of the set of admissible portfolios (DoF  $d$  and VC dimension  $h$ ) will be maximized. Therefore, it is necessary to adjust the empirical error using the correction factor, which is based on the capacity of the set of admissible portfolios and on the size of the training sample  $T$ .

The different analytic methods compute this correction factor differently (see equations (5.32) to (5.36)). We report the correction factors for each method and each constant  $C$  in the next panel of columns in Table 5.4. Then we compute the estimated test error as a product of the empirical error and the correction factor. For example, for method FPE and parameter  $C = 2^{-5}$ , the empirical error is equal to 0.0864 and the correction factor is equal to 1.18. The estimated test error is therefore equal to  $0.0864 \times 1.18 = 0.1020$ . Hence, the resulting estimation of the test error is based on two factors, the empirical error and the correction factor, which depends on the size of the training set and the capacity of the set of admissible portfolios. Using this estimation of the test error, it is possible to select the optimal value of the parameter  $C$ . For example, the FPE, GCV, and SMS methods indicate that a value of  $C = 2^{-5}$

should be chosen, since the estimation of the test error is the smallest in this case. Meanwhile, the SC method indicates  $C = 2^{-8}$ , and the VM method  $C = 2^{-7}$ .

Additionally, the results in Table 5.4 demonstrate the operation of the SVM\_CW method. If parameter  $C = 0$ , the SVM\_CW method is equivalent to the CW method; if  $C = \infty$ , it is equivalent to the MAXD method. Using a value of  $C$  between these two extremes provides a trade-off between the CW and MAXD portfolios. This can also be seen by analysing the distance  $\|\mathbf{w} - \mathbf{w}_0\|_2$  from the capitalization weighted portfolio. The smaller is the parameter  $C$ , the smaller is the distance between the estimated tracking portfolio weights and the capitalization weights.

Moreover, the SVM\_CW method automatically adjusts the insensitive region, expressed by the parameter  $\varepsilon$ , in the  $\varepsilon$ -insensitive loss function. By increasing penalty parameter  $C$  the insensitive region increases to compensate for the higher penalization of errors.

### **Resampling model selection methods**

The alternative group of model selection methods is the resampling methods, of which there are two kinds: hold out (HO) and cross-validation (CV) (see Section 5.1.4.5).

We use different parameterizations for the hold out method. In this method, the estimation period of length twelve months is divided into two parts. The first part is used for model estimation, the second to compute the expected test error. Recent asset returns and index returns fall into the second part. We divide the initial estimation period into two parts in different ways as follows: 8+4 months (HO8.4 method), 9+3 months (HO9.3 method), 10+2 months (HO10.2 method), 11+1 months (HO11.1 method). As a consequence, we end up with four variants of the hold out method.

For the cross-validation method, we have to select the number of folds, and we choose to set this equal to four (CV.4FOLD method). As a consequence, the size of each fold is three months.

To illustrate the use of resampling methods, we apply them to selecting parameter  $C$  for the SVM\_CW method as before. The results are displayed in Table 5.5.

	C	Empirical error, RMSE%	Estimation of expected test error, RMSE%				
			HO8.4	HO9.3	HO10.2	HO11.1	CV.4FOLD
<b>CW</b>	<b>0</b>	0.2068	0.1893	0.2099	0.1761	0.2283	0.2024
<b>SVM_CW</b>	$2^{-16}$	0.1914	0.1810	0.2047	0.1712	0.2228	0.1891
	$2^{-15}$	0.1854	0.1777	0.2024	0.1689	0.2208	0.1839
	$2^{-14}$	0.1772	0.1730	0.1992	0.1652	0.2176	0.1766
	$2^{-13}$	0.1672	0.1654	0.1945	0.1587	0.2130	0.1695
	$2^{-12}$	0.1548	0.1577	0.1897	0.1536	0.2099	0.1571
	$2^{-11}$	0.1395	0.1469	0.1848	0.1488	0.2073	0.1464
	$2^{-10}$	0.1223	0.1355	0.1805	0.1431	0.2054	0.1333
	$2^{-9}$	0.1073	0.1269	0.1786	0.1414	0.2041	0.1232
	$2^{-8}$	0.0977	0.1231	0.1802	0.1475	0.2113	0.1176
	$2^{-7}$	0.0914	0.1211	0.1775	0.1469	0.2150	0.1154
	$2^{-6}$	0.0866	0.1268	0.1776	0.1392	0.2110	0.1198
	$2^{-5}$	0.0864	0.1350	0.1892	0.1494	0.2287	0.1174
<b>MAXD</b>	$\infty$	0.0943	0.1185	0.1741	0.1444	0.2318	0.1236

Table 5.5: Resampling Methods for SVM\_CW Strategy (Period 1)

Parameter  $C$  is chosen so as to minimize the estimated expected test error. Hence, the HO8.4 and HO9.3 methods result in  $C = \infty$ , HO10.2 gives  $C = 2^{-6}$ , HO11.1 gives  $C = 2^{-9}$ , and using CV.4FOLD produces  $C = 2^{-7}$ . It can be observed here that the value of parameter  $C$  with the lowest empirical error is not selected by any of the resampling methods. This indicates that the model complexity, which depends on the penalty parameter  $C$ , contributes towards achieving a better generalizability of the data.

Next we analyse whether STEM-based methods combined with model selection methods (analytic and resampling methods) are able to produce better tracking quality than traditional methods. We make no general conclusion here, but simply investigate the application of the model selection methods. For this, we use the S&P 1500 index with the same time periods as we used earlier (see Section 4.1.1). For each period, we select 120 assets for the tracking portfolios using a stratification procedure (see Section 4.3.1). Then, we choose the parameter  $C$  for each of the methods in Table 5.1, using analytic and resampling model selection methods. This results in five combinations of the assets in the tracking portfolios, one per period. After this, we compare the results of the STEM-based methods with the results of the traditional methods.

#### 5.4.2.2 Empirical results

To compare the two sets of methods, we start by reporting the tracking qualities for the traditional methods in Table 5.6. There are two heuristic weighting methods, equal weighted (EW) and capitalization weighted (CW) and five optimization methods, with objectives MAXD, CVAR, CVAR2, MSE and MSE (b) (see Section 3.3). Each method is used to select the weights for the one combination of assets per period, which is constructed using the stratification method. As a result, there is one tracking portfolio for each period and each method. For each period, the lowest tracking quality out of all the methods is marked as grey. These results serve as benchmarks for the STEM-based methods.

		Estimation							Investment						
		EW	CW	MAXD	CVAR	CVAR2	MSE	MSE (b)	EW	CW	MAXD	CVAR	CVAR2	MSE	MSE (b)
1	RMSE%	0.3220	0.2068	0.0950	0.0950	0.0860	0.0813	0.2476	0.2835	0.1726	0.1348	0.1348	0.1508	0.1398	0.2326
	ETQ%								3.1419	2.4461	0.6020	0.6020	0.3984	0.3030	2.4631
2	RMSE%	0.2968	0.1764	0.1045	0.1052	0.0938	0.0885	0.2531	0.2383	0.2679	0.1527	0.1501	0.1595	0.1566	0.1880
	ETQ%								0.7813	1.3198	0.2483	0.2623	0.2596	0.2547	0.4409
3	RMSE%	0.2996	0.2003	0.0888	0.0885	0.0832	0.0758	0.0763	0.3599	0.2198	0.2288	0.2105	0.1814	0.1986	0.1933
	ETQ%								0.7893	1.1566	0.4480	0.3881	0.4681	0.4875	0.4583
4	RMSE%	0.2814	0.2267	0.1150	0.1165	0.1078	0.1013	0.2503	0.3557	0.3415	0.1977	0.1908	0.2131	0.2054	0.3283
	ETQ%								1.8568	1.5543	0.6677	0.5576	0.8099	0.6858	1.7405
5	RMSE%	0.5350	0.6488	0.2870	0.2972	0.2860	0.2672	0.2698	0.9064	0.4865	0.3914	0.3389	0.3178	0.3283	0.3009
	ETQ%								9.4570	4.7383	6.7292	4.2175	3.2698	5.1009	4.0100

Table 5.6: Tracking Quality in the Estimation and Investment Periods (Traditional Methods)

Next, we generate the tracking portfolios using STEM-based methods. The results for the estimation period are reported in Table 5.7. We report the parameter  $C$  selected by each model selection method. Where no value of  $C$  is reported,  $C = \infty$ .

		SVM_EW		SVM_CW		CVAR_EW		CVAR_CW		CVAR2_EW		CVAR2_CW		MSE_EW		MSE_CW	
		RMSE	C	RMSE	C	RMSE	C	RMSE	C	RMSE	C	RMSE	C	RMSE	C	RMSE	C
1	FPE	0.0847	0.0313	0.0864	0.0313	0.0946	0.0078	0.0921	0.0078	0.0862	0.0078	0.0849	0.0078	0.0817	0.0078	0.0817	0.0078
	GCV	0.0847	0.0313	0.0864	0.0313	0.0946	0.0078	0.0921	0.0078	0.0862	0.0078	0.0849	0.0078	0.0817	0.0078	0.0817	0.0078
	SC	0.0847	0.0313	0.0977	0.0039	0.0946	0.0078	0.0921	0.0078	0.0862	0.0078	0.0849	0.0078	0.0817	0.0078	0.0876	0.0039
	SMS	0.0847	0.0313	0.0864	0.0313	0.0946	0.0078	0.0921	0.0078	0.0862	0.0078	0.0849	0.0078	0.0817	0.0078	0.0817	0.0078
	VM	0.0847	0.0313	0.0914	0.0078	0.0946	0.0078	0.0921	0.0078	0.0860	0.0156	0.0849	0.0078	0.0817	0.0156	0.0817	0.0078
	HO8.4	0.0848	0.0156	0.0943		0.0946	0.0078	0.0950	0.0313	0.0960	0.0039	0.0907	0.0039	0.0929	0.0039	0.0876	0.0039
	HO9.3	0.0889	0.0078	0.0943		0.0946	0.0078	0.0921	0.0078	0.0960	0.0039	0.1046	0.0020	0.0929	0.0039	0.1015	0.0020
	HO10.2	0.0980	0.0039	0.0866	0.0156	0.1076	0.0039	0.0979	0.0039	0.0960	0.0039	0.1046	0.0020	0.0929	0.0039	0.1015	0.0020
	HO11.1	0.0980	0.0039	0.1073	0.0020	0.1076	0.0039	0.0979	0.0039	0.0960	0.0039	0.1046	0.0020	0.0929	0.0039	0.1015	0.0020
	CV.4FO	0.0848	0.0156	0.0914	0.0078	0.0946	0.0078	0.0979	0.0039	0.0960	0.0039	0.0907	0.0039	0.0929	0.0039	0.0876	0.0039
2	FPE	0.0961	0.0156	0.0987	0.0039	0.1055	0.0078	0.1050	0.0039	0.0937	0.0078	0.0936	0.0039	0.0896	0.0078	0.0899	0.0039
	GCV	0.0961	0.0156	0.0987	0.0039	0.1055	0.0078	0.1050	0.0039	0.0937	0.0078	0.0936	0.0039	0.0896	0.0078	0.0899	0.0039
	SC	0.0961	0.0156	0.1022	0.0020	0.1055	0.0078	0.1050	0.0039	0.0937	0.0078	0.0989	0.0020	0.0896	0.0078	0.0955	0.0020
	SMS	0.0961	0.0156	0.0957	0.0078	0.1055	0.0078	0.1050	0.0039	0.0937	0.0078	0.0936	0.0039	0.0896	0.0078	0.0899	0.0039
	VM	0.0961	0.0156	0.0987	0.0039	0.1052	0.0313	0.1050	0.0039	0.0937	0.0078	0.0936	0.0039	0.0887	0.0156	0.0899	0.0039
	HO8.4	0.1043		0.1043		0.1055	0.0078	0.1039	0.0078	0.0938	0.0156	0.0936	0.0039	0.2531		0.0899	0.0039
	HO9.3	0.1043		0.1043		0.1052	0.0156	0.1052	0.0313	0.0938		0.0989	0.0020	0.0887	0.0313	0.0899	0.0039
	HO10.2	0.1043		0.1043		0.1052	0.0313	0.1039	0.0078	0.0937	0.0078	0.0936	0.0039	0.0887	0.0313	0.0887	0.0156
	HO11.1	0.1043		0.0957	0.0078	0.1052		0.1144	0.0020	0.1070	0.0039	0.0989	0.0020	0.0887	0.0156	0.0887	0.0313
	CV.4FO	0.0966	0.0313	0.0987	0.0039	0.1193	0.0039	0.1144	0.0020	0.1070	0.0039	0.0989	0.0020	0.0896	0.0078	0.0955	0.0020
3	FPE	0.0841	0.0156	0.0821	0.0078	0.0885	0.0156	0.0888	0.0078	0.0825	0.0078	0.0837	0.0039	0.0763		0.0763	0.0078
	GCV	0.0841	0.0156	0.0821	0.0078	0.0885	0.0156	0.0888	0.0078	0.0825	0.0078	0.0837	0.0039	0.0763		0.0763	0.0078
	SC	0.0841	0.0156	0.0821	0.0078	0.0912	0.0078	0.0888	0.0078	0.0825	0.0078	0.0837	0.0039	0.0763		0.0812	0.0039
	SMS	0.0841	0.0156	0.0821	0.0078	0.0885	0.0156	0.0888	0.0078	0.0825	0.0078	0.0837	0.0039	0.0763		0.0763	0.0078
	VM	0.0896	0.0078	0.0821	0.0078	0.0885	0.0313	0.0885	0.0156	0.0832		0.0837	0.0039	0.0763		0.0763	
	HO8.4	0.0843	0.0313	0.0821	0.0078	0.0885		0.1044	0.0020	0.0833	0.0156	0.0837	0.0039	0.0767	0.0078	0.0951	0.0020
	HO9.3	0.0885		0.0883	0.0039	0.0885	0.0156	0.1044	0.0020	0.0832		0.0968	0.0020	0.0767	0.0078	0.0951	0.0020
	HO10.2	0.0885		0.0999	0.0020	0.0885		0.1196	0.0010	0.0825	0.0078	0.0968	0.0020	0.0767	0.0078	0.1121	0.0010
	HO11.1	0.1589	0.0010	0.0883	0.0039	0.1328	0.0020	0.0938	0.0039	0.1585	0.0010	0.0968	0.0020	0.1546	0.0010	0.0951	0.0020
	CV.4FO	0.0896	0.0078	0.0821	0.0078	0.0912	0.0078	0.0938	0.0039	0.0825	0.0078	0.0837	0.0039	0.0767	0.0078	0.0763	0.0078
4	FPE	0.1057	0.0156	0.1073	0.0313	0.1165	0.0156	0.1160	0.0078	0.1057	0.0078	0.1068	0.0078	0.1018	0.0078	0.1013	0.0078
	GCV	0.1057	0.0156	0.1105	0.0156	0.1165	0.0156	0.1160	0.0078	0.1057	0.0078	0.1068	0.0078	0.1018	0.0078	0.1013	0.0078
	SC	0.1057	0.0156	0.1181	0.0039	0.1200	0.0078	0.1160	0.0078	0.1057	0.0078	0.1068	0.0078	0.1018	0.0078	0.1013	0.0078
	SMS	0.1057	0.0156	0.1073	0.0313	0.1165	0.0156	0.1160	0.0078	0.1057	0.0078	0.1068	0.0078	0.1018	0.0078	0.1013	0.0078
	VM	0.1057	0.0156	0.1135	0.0078	0.1165	0.0156	0.1160	0.0078	0.1057	0.0078	0.1068	0.0078	0.1018	0.0078	0.1013	0.0078
	HO8.4	0.1075	0.0313	0.1105	0.0156	0.1200	0.0078	0.1249	0.0039	0.1078	0.0156	0.1138	0.0039	0.1018	0.0078	0.1075	0.0039
	HO9.3	0.1109	0.0078	0.1135	0.0078	0.1200	0.0078	0.1249	0.0039	0.1078	0.0156	0.1138	0.0039	0.1013	0.0156	0.1075	0.0039
	HO10.2	0.1075	0.0313	0.1135	0.0078	0.1351	0.0039	0.1249	0.0039	0.1078		0.1138	0.0039	0.1018	0.0078	0.1075	0.0039
	HO11.1	0.1227	0.0039	0.1135	0.0078	0.1165	0.0313	0.1165	0.0156	0.1196	0.0039	0.1138	0.0039	0.1166	0.0039	0.1075	0.0039
	CV.4FO	0.1075	0.0313	0.1181	0.0039	0.1351	0.0039	0.1249	0.0039	0.1057	0.0078	0.1078	0.0313	0.1013	0.0156	0.1075	0.0039
5	FPE	0.2855		0.2855		0.2972	0.0156	0.2973	0.0156	0.2833	0.0078	0.2859	0.0156	0.2724	0.0078	0.2699	0.0156
	GCV	0.2855		0.2855		0.2972	0.0156	0.2973	0.0156	0.2833	0.0078	0.2859	0.0156	0.2724	0.0078	0.2699	0.0156
	SC	0.2855		0.2855		0.2972	0.0156	0.3064	0.0078	0.2833	0.0078	0.2990	0.0078	0.2724	0.0078	0.2841	0.0078
	SMS	0.2855		0.2855		0.2972	0.0156	0.2973	0.0156	0.2833	0.0078	0.2859	0.0156	0.2724	0.0078	0.2699	0.0156
	VM	0.2855		0.2855		0.2972		0.2972		0.2860	0.0156	0.2859	0.0156	0.2698		0.2698	
	HO8.4	0.3024	0.0078	0.3578	0.0039	0.3255	0.0039	0.3575	0.0020	0.3026	0.0039	0.3264	0.0039	0.2891	0.0039	0.3528	0.0020
	HO9.3	0.2855		0.4247	0.0010	0.3129	0.0078	0.2973	0.0156	0.2833	0.0078	0.4038	0.0010	0.2724	0.0078	0.4019	0.0010
	HO10.2	0.2855		0.3865	0.0020	0.3129	0.0078	0.3064	0.0078	0.2860		0.4038	0.0010	0.2698		0.3138	0.0039
	HO11.1	0.2855		0.4247	0.0010	0.3129	0.0078	0.2973	0.0156	0.2833	0.0078	0.3577	0.0020	0.2724	0.0078	0.3528	0.0020
	CV.4FO	0.2901	0.0313	0.3321	0.0078	0.2972	0.0156	0.3064	0.0078	0.3026	0.0039	0.2990	0.0078	0.2724	0.0078	0.3138	0.0039

Table 5.7: Tracking Qualities in the Estimation Period (STEM)



A comparison of the tracking qualities of the STEM-based methods (see Table 5.7) with those of the traditional methods (see Table 5.6) indicates that the traditional methods have better tracking quality in the estimation period in most cases.

Next, we report the results for the investment period, in Table 5.8.

		SVM_EW		SVM_CW		CVAR_EW		CVAR_CW		CVAR2_EW		CVAR2_CW		MSE_EW		MSE_CW	
		RMSE	ETQ	RMSE	ETQ	RMSE	ETQ	RMSE	ETQ	RMSE	ETQ	RMSE	ETQ	RMSE	ETQ	RMSE	ETQ
1	FPE	0.1418	0.4561	0.1416	0.3843	0.1340	0.6371	0.1405	0.7005	0.1506	0.4049	0.1477	0.3773	0.1401	0.3050	0.1398	0.3038
	GCV	0.1418	0.4561	0.1416	0.3843	0.1340	0.6371	0.1405	0.7005	0.1506	0.4049	0.1477	0.3773	0.1401	0.3050	0.1398	0.3038
	SC	0.1418	0.4561	0.1257	0.3928	0.1340	0.6371	0.1405	0.7005	0.1506	0.4049	0.1477	0.3773	0.1401	0.3050	0.1272	0.3008
	SMS	0.1418	0.4561	0.1416	0.3843	0.1340	0.6371	0.1405	0.7005	0.1506	0.4049	0.1477	0.3773	0.1401	0.3050	0.1398	0.3038
	VM	0.1418	0.4561	0.1256	0.3247	0.1340	0.6371	0.1405	0.7005	0.1508	0.3984	0.1477	0.3773	0.1402	0.3047	0.1398	0.3038
	HO8.4	0.1381	0.4116	0.1361	0.6017	0.1340	0.6371	0.1348	0.6020	0.1329	0.8293	0.1310	0.3052	0.1348	0.6979	0.1272	0.3008
	HO9.3	0.1383	0.7055	0.1361	0.6017	0.1340	0.6371	0.1405	0.7005	0.1329	0.8293	0.1205	0.3987	0.1348	0.6979	0.1249	0.4702
	HO10.2	0.1421	0.9309	0.1408	0.3566	0.1399	0.9005	0.1423	0.3380	0.1329	0.8293	0.1205	0.3987	0.1348	0.6979	0.1249	0.4702
	HO11.1	0.1421	0.9309	0.1246	0.5535	0.1399	0.9005	0.1423	0.3380	0.1329	0.8293	0.1205	0.3987	0.1348	0.6979	0.1249	0.4702
	CV.4FO	0.1381	0.4116	0.1256	0.3247	0.1340	0.6371	0.1423	0.3380	0.1329	0.8293	0.1310	0.3052	0.1348	0.6979	0.1272	0.3008
2	FPE	0.1537	0.3506	0.1825	0.4575	0.1474	0.3164	0.1586	0.2201	0.1555	0.2336	0.1698	0.4307	0.1525	0.2340	0.1639	0.3253
	GCV	0.1537	0.3506	0.1825	0.4575	0.1474	0.3164	0.1586	0.2201	0.1555	0.2336	0.1698	0.4307	0.1525	0.2340	0.1639	0.3253
	SC	0.1537	0.3506	0.1965	0.5963	0.1474	0.3164	0.1586	0.2201	0.1555	0.2336	0.1786	0.3720	0.1525	0.2340	0.1813	0.4631
	SMS	0.1537	0.3506	0.1723	0.3654	0.1474	0.3164	0.1586	0.2201	0.1555	0.2336	0.1698	0.4307	0.1525	0.2340	0.1639	0.3253
	VM	0.1537	0.3506	0.1825	0.4575	0.1501	0.2609	0.1586	0.2201	0.1555	0.2336	0.1698	0.4307	0.1564	0.2485	0.1639	0.3253
	HO8.4	0.1494	0.2279	0.1494	0.2279	0.1474	0.3164	0.1501	0.2791	0.1595	0.2596	0.1698	0.4307	0.1880	0.4409	0.1639	0.3253
	HO9.3	0.1494	0.2279	0.1494	0.2279	0.1501	0.2610	0.1501	0.2610	0.1595	0.2596	0.1786	0.3720	0.1564	0.2485	0.1639	0.3253
	HO10.2	0.1494	0.2279	0.1494	0.2279	0.1501	0.2609	0.1501	0.2791	0.1555	0.2336	0.1698	0.4307	0.1564	0.2485	0.1564	0.2485
	HO11.1	0.1494	0.2279	0.1723	0.3654	0.1501	0.2623	0.1869	0.4692	0.1422	0.3268	0.1786	0.3720	0.1564	0.2485	0.1565	0.2485
	CV.4FO	0.1499	0.1776	0.1825	0.4575	0.1313	0.1446	0.1869	0.4692	0.1422	0.3268	0.1786	0.3720	0.1525	0.2340	0.1813	0.4631
3	FPE	0.1936	0.5631	0.1710	0.5332	0.2105	0.3881	0.2085	0.3867	0.1840	0.5113	0.1674	0.3824	0.1933	0.4583	0.1930	0.4571
	GCV	0.1936	0.5631	0.1710	0.5332	0.2105	0.3881	0.2085	0.3867	0.1840	0.5113	0.1674	0.3824	0.1933	0.4583	0.1930	0.4571
	SC	0.1936	0.5631	0.1710	0.5332	0.2163	0.3791	0.2085	0.3867	0.1840	0.5113	0.1674	0.3824	0.1933	0.4583	0.1675	0.3987
	SMS	0.1936	0.5631	0.1710	0.5332	0.2105	0.3881	0.2085	0.3867	0.1840	0.5113	0.1674	0.3824	0.1933	0.4583	0.1930	0.4571
	VM	0.1981	0.3976	0.1710	0.5332	0.2105	0.3881	0.2105	0.3881	0.1814	0.4681	0.1674	0.3824	0.1933	0.4583	0.1933	0.4583
	HO8.4	0.1916	0.3665	0.1710	0.5332	0.2105	0.3881	0.1764	0.3341	0.1814	0.4687	0.1674	0.3824	0.1953	0.4517	0.1603	0.3858
	HO9.3	0.2083	0.3744	0.1638	0.4601	0.2105	0.3881	0.1764	0.3341	0.1814	0.4681	0.1587	0.3280	0.1953	0.4517	0.1603	0.3858
	HO10.2	0.2083	0.3744	0.1650	0.5477	0.2105	0.3881	0.1817	0.4462	0.1840	0.5113	0.1587	0.3280	0.1953	0.4517	0.1675	0.4477
	HO11.1	0.2582	0.7394	0.1638	0.4601	0.2173	0.5347	0.1732	0.3074	0.2455	0.6963	0.1587	0.3280	0.2409	0.6513	0.1603	0.3858
	CV.4FO	0.1981	0.3976	0.1710	0.5332	0.2163	0.3791	0.1732	0.3074	0.1840	0.5113	0.1674	0.3824	0.1953	0.4517	0.1930	0.4571
4	FPE	0.2226	1.0943	0.2057	1.0688	0.1908	0.5576	0.1978	0.6275	0.2196	1.0079	0.2139	0.9264	0.1998	0.7639	0.2050	0.6863
	GCV	0.2226	1.0943	0.1994	0.9003	0.1908	0.5576	0.1978	0.6275	0.2196	1.0079	0.2139	0.9264	0.1998	0.7639	0.2050	0.6863
	SC	0.2226	1.0943	0.2082	1.0474	0.1801	0.5032	0.1978	0.6275	0.2196	1.0079	0.2139	0.9264	0.1998	0.7639	0.2050	0.6863
	SMS	0.2226	1.0943	0.2057	1.0688	0.1908	0.5576	0.1978	0.6275	0.2196	1.0079	0.2139	0.9264	0.1998	0.7639	0.2050	0.6863
	VM	0.2226	1.0943	0.2032	0.9191	0.1908	0.5576	0.1978	0.6275	0.2196	1.0079	0.2139	0.9264	0.1998	0.7639	0.2050	0.6863
	HO8.4	0.2210	1.1938	0.1994	0.9003	0.1801	0.5032	0.1940	0.8590	0.2131	0.8099	0.1979	0.8325	0.1998	0.7639	0.1986	0.9647
	HO9.3	0.2010	0.9499	0.2032	0.9191	0.1801	0.5032	0.1940	0.8590	0.2131	0.8099	0.1979	0.8325	0.2053	0.6844	0.1986	0.9647
	HO10.2	0.2210	1.1938	0.2032	0.9191	0.1820	0.7416	0.1940	0.8590	0.2131	0.8099	0.1979	0.8325	0.1998	0.7639	0.1986	0.9647
	HO11.1	0.1940	1.0589	0.2032	0.9191	0.1908	0.5576	0.1908	0.5576	0.2053	1.1514	0.1979	0.8325	0.1916	1.0124	0.1986	0.9647
	CV.4FO	0.2210	1.1938	0.2082	1.0474	0.1820	0.7416	0.1940	0.8590	0.2196	1.0079	0.2131	0.8099	0.2053	0.6844	0.1986	0.9647
5	FPE	0.3468	5.0284	0.3468	5.0284	0.3389	4.2174	0.3434	4.0733	0.3231	3.4199	0.3178	3.2708	0.3076	4.1048	0.3010	4.0105
	GCV	0.3468	5.0284	0.3468	5.0284	0.3389	4.2174	0.3434	4.0733	0.3231	3.4199	0.3178	3.2708	0.3076	4.1048	0.3010	4.0105
	SC	0.3468	5.0284	0.3468	5.0284	0.3389	4.2174	0.3453	3.7511	0.3231	3.4199	0.3102	2.7978	0.3076	4.1048	0.2936	3.6519
	SMS	0.3468	5.0284	0.3468	5.0284	0.3389	4.2174	0.3434	4.0733	0.3231	3.4199	0.3178	3.2708	0.3076	4.1048	0.3010	4.0105
	VM	0.3468	5.0284	0.3468	5.0284	0.3389	4.2175	0.3389	4.2175	0.3178	3.2698	0.3178	3.2708	0.3009	4.0100	0.3009	4.0100
	HO8.4	0.3567	4.8435	0.2525	1.4912	0.3832	4.7448	0.3831	3.2521	0.3564	4.8535	0.2709	1.9209	0.3515	4.8469	0.2788	2.2464
	HO9.3	0.3468	5.0284	0.2620	0.6049	0.2810	3.2820	0.3434	4.0733	0.3231	3.4199	0.2710	0.7450	0.3076	4.1048	0.2691	1.0223
	HO10.2	0.3468	5.0284	0.2652	0.9328	0.2810	3.2820	0.3453	3.7511	0.3178	3.2698	0.2710	0.7450	0.3009	4.0100	0.2883	1.5338
	HO11.1	0.3468	5.0284	0.2620	0.6049	0.2810	3.2820	0.3434	4.0733	0.3231	3.4199	0.2612	1.4173	0.3076	4.1048	0.2788	2.2464
	CV.4FO	0.3248	4.1165	0.2687	1.9125	0.3389	4.2174	0.3453	3.7511	0.3564	4.8535	0.3102	2.7978	0.3076	4.1048	0.2883	3.1538

Table 5.8: Tracking Qualities in the Investment Period (STEM)

To compare the results in Table 5.6 with the results in Table 5.8, we report the median, minimal and maximal ETQs across the five time periods for each of the traditional methods (see Table 5.9, top panel). For example, for the MSE (b) method, the ETQs in the five time periods were [2.4631, 0.4409, 0.4583, 1.7405, 4.0100]. In the table below, we report the median (1.7405), the minimum (0.4409), and the maximum (4.0100). We report the corresponding values for each of the STEM-based methods in the remaining panels of Table 5.9.

	Med	Min	Max
<b>EW</b>	1.8568	0.7813	9.4570
<b>CW</b>	1.5543	1.1566	4.7383
<b>MAXD</b>	0.6020	0.2483	6.7292
<b>CVAR</b>	0.5576	0.2623	4.2175
<b>CVAR2</b>	0.4681	0.2596	3.2698
<b>MSE (b)</b>	1.7405	0.4409	4.0100

	SVM_EW, ETQ%			SVM_CW, ETQ%				CVAR_EW, ETQ%			CVAR_CW, ETQ%		
	Med	Min	Max	Med	Min	Max		Med	Min	Max	Med	Min	Max
<b>FPE</b>	0.5631	0.3506	5.0284	0.5332	0.3843	5.0284	<b>FPE</b>	0.5576	0.3164	4.2174	0.6275	0.2201	4.0733
<b>GCV</b>	0.5631	0.3506	5.0284	0.5332	0.3843	5.0284	<b>GCV</b>	0.5576	0.3164	4.2174	0.6275	0.2201	4.0733
<b>SC</b>	0.5631	0.3506	5.0284	0.5963	0.3928	5.0284	<b>SC</b>	0.5032	0.3164	4.2174	0.6275	0.2201	3.7511
<b>SMS</b>	0.5631	0.3506	5.0284	0.5332	0.3654	5.0284	<b>SMS</b>	0.5576	0.3164	4.2174	0.6275	0.2201	4.0733
<b>VM</b>	0.4561	0.3506	5.0284	0.5332	0.3247	5.0284	<b>VM</b>	0.5576	0.2609	4.2175	0.6275	0.2201	4.2175
<b>HO8.4</b>	0.4116	0.2279	4.8435	0.6017	0.2279	1.4912	<b>HO8.4</b>	0.5032	0.3164	4.7448	0.6020	0.2791	3.2521
<b>HO9.3</b>	0.7055	0.2279	5.0284	0.6017	0.2279	0.9191	<b>HO9.3</b>	0.5032	0.2610	3.2820	0.7005	0.2610	4.0733
<b>HO10.2</b>	0.9309	0.2279	5.0284	0.5477	0.2279	0.9328	<b>HO10.2</b>	0.7416	0.2609	3.2820	0.4462	0.2791	3.7511
<b>HO11.1</b>	0.9309	0.2279	5.0284	0.5535	0.3654	0.9191	<b>HO11.1</b>	0.5576	0.2623	3.2820	0.4692	0.3074	4.0733
<b>CV.4FO</b>	0.4116	0.1776	4.1165	0.5332	0.3247	1.9125	<b>CV.4FO</b>	0.6371	0.1446	4.2174	0.4692	0.3074	3.7511

	CVAR2_EW, ETQ%			CVAR2_CW, ETQ%				MSE_EW, ETQ%			MSE_CW, ETQ%		
	Med	Min	Max	Med	Min	Max		Med	Min	Max	Med	Min	Max
<b>FPE</b>	0.5113	0.2336	3.4199	0.4307	0.3773	3.2708	<b>FPE</b>	0.4583	0.2340	4.1048	0.4571	0.3038	4.0105
<b>GCV</b>	0.5113	0.2336	3.4199	0.4307	0.3773	3.2708	<b>GCV</b>	0.4583	0.2340	4.1048	0.4571	0.3038	4.0105
<b>SC</b>	0.5113	0.2336	3.4199	0.3824	0.3720	2.7978	<b>SC</b>	0.4583	0.2340	4.1048	0.4631	0.3008	3.6519
<b>SMS</b>	0.5113	0.2336	3.4199	0.4307	0.3773	3.2708	<b>SMS</b>	0.4583	0.2340	4.1048	0.4571	0.3038	4.0105
<b>VM</b>	0.4681	0.2336	3.2698	0.4307	0.3773	3.2708	<b>VM</b>	0.4583	0.2485	4.0100	0.4583	0.3038	4.0100
<b>HO8.4</b>	0.8099	0.2596	4.8535	0.4307	0.3052	1.9209	<b>HO8.4</b>	0.6979	0.4409	4.8469	0.3858	0.3008	2.2464
<b>HO9.3</b>	0.8099	0.2596	3.4199	0.3987	0.3280	0.8325	<b>HO9.3</b>	0.6844	0.2485	4.1048	0.4702	0.3253	1.0223
<b>HO10.2</b>	0.8099	0.2336	3.2698	0.4307	0.3280	0.8325	<b>HO10.2</b>	0.6979	0.2485	4.0100	0.4702	0.2485	3.1538
<b>HO11.1</b>	0.8293	0.3268	3.4199	0.3987	0.3280	1.4173	<b>HO11.1</b>	0.6979	0.2485	4.1048	0.4702	0.2485	2.2464
<b>CV.4FO</b>	0.8293	0.3268	4.8535	0.3824	0.3052	2.7978	<b>CV.4FO</b>	0.6844	0.2340	4.1048	0.4631	0.3008	3.1538

Table 5.9: Aggregated Statistics for Traditional and STEM-Based Methods in the Investment Period (ETQ)

The results in Table 5.9 indicate that, by using STEM-based methods, it is possible to improve the tracking quality in the investment period in comparison to traditional methods (the minimal Med, Min and Max values among all methods are highlighted).

### 5.4.2.3 Conclusion

The results in this section indicate that STEM-based methods, in conjunction with the model selection methods, are able to improve the tracking quality in the investment period in comparison to traditional methods. However, our analysis has been limited to just one combination of assets per period. In the next section, we verify the results by testing the absolute and relative stability of tracking quality, for portfolios constructed using STEM-based methods.

### 5.4.3 Stability of methods using the STEM approach

#### 5.4.3.1 Research method

The goal of this section is to analyse the absolute and relative stability of the tracking quality of portfolios generated using STEM-based methods. We use the same methodology as we used for the traditional methods in Section 4.2. We analyse five periods representing different market phases. The dataset contains daily prices for the S&P 1500 index and its constituents. 5,000 portfolios are randomly generated for each period and each portfolio size (we use the same asset combinations as in Section 4.2).

We analyse the eight methods from Table 5.1. Portfolio sizes [60, 120, 240] are selected. To reduce the computational time, we choose the HO9.3 model selection method to determine the value of the penalty parameter  $C$ , since it showed good results in the previous section. The range of possible values for the constant  $C$  is reduced to the values

$$C = [2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, \infty].$$

We conduct the same three analyses as in Section 4.2. First, we analyse the absolute stability of the tracking quality. Second, we examine the relative stability of the tracking quality. Third, we inspect the ability to predict the ex post tracking quality using the measures of tracking quality in the estimation period.

Additionally, we compare the tracking qualities of the portfolios generated using STEM-based methods with those of the portfolios generated using traditional methods.

#### 5.4.3.2 Empirical results

##### Absolute stability of tracking quality

In the first piece of analysis, we examine the absolute stability of the tracking quality of the portfolios constructed using STEM-based methods of optimized sampling. The medians and IQRs, expressed in RMSE, for the estimation period and the investment period are reported in Table 5.10.

		SVM_EW, RMSE%				SVM_CW, RMSE%				CVAR_EW, RMSE%				CVAR_CW, RMSE%			
Size		Estimation		Investment		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	0.2251	0.0358	0.2477	0.0456	0.2261	0.0357	0.2451	0.0515	0.2433	0.0414	0.2656	0.0527	0.2371	0.0390	0.2556	0.0516
	120	0.1616	0.0284	0.1951	0.0339	0.1599	0.0267	0.1890	0.0368	0.1809	0.0318	0.2118	0.0386	0.1752	0.0313	0.1994	0.0386
	240	0.1110	0.0208	0.1557	0.0256	0.1116	0.0238	0.1471	0.0267	0.1247	0.0227	0.1662	0.0276	0.1236	0.0281	0.1535	0.0278
2	60	0.2193	0.0350	0.2415	0.0472	0.2203	0.0355	0.2531	0.0485	0.2387	0.0418	0.2588	0.0513	0.2316	0.0398	0.2615	0.0508
	120	0.1571	0.0256	0.1885	0.0310	0.1569	0.0252	0.1967	0.0334	0.1773	0.0287	0.1997	0.0332	0.1711	0.0305	0.2044	0.0349
	240	0.1096	0.0174	0.1497	0.0233	0.1115	0.0218	0.1518	0.0250	0.1214	0.0182	0.1519	0.0257	0.1229	0.0256	0.1553	0.0270
3	60	0.1930	0.0303	0.3138	0.0629	0.1989	0.0312	0.3149	0.0665	0.2082	0.0355	0.3302	0.0665	0.2075	0.0350	0.3263	0.0671
	120	0.1343	0.0205	0.2437	0.0467	0.1383	0.0231	0.2394	0.0461	0.1513	0.0253	0.2534	0.0464	0.1511	0.0275	0.2470	0.0460
	240	0.0896	0.0144	0.1916	0.0343	0.0944	0.0203	0.1805	0.0331	0.1004	0.0163	0.1927	0.0343	0.1053	0.0231	0.1832	0.0333
4	60	0.2376	0.0353	0.3957	0.0702	0.2422	0.0362	0.4067	0.0788	0.2569	0.0392	0.4158	0.0739	0.2523	0.0393	0.4168	0.0819
	120	0.1702	0.0273	0.3177	0.0539	0.1719	0.0270	0.3189	0.0601	0.1875	0.0288	0.3357	0.0566	0.1858	0.0316	0.3303	0.0636
	240	0.1175	0.0175	0.2562	0.0447	0.1214	0.0223	0.2490	0.0468	0.1278	0.0190	0.2658	0.0462	0.1333	0.0261	0.2535	0.0462
5	60	0.4903	0.0694	0.6398	0.1468	0.5080	0.0872	0.6140	0.1474	0.5069	0.0725	0.6892	0.1726	0.5225	0.0901	0.6504	0.1598
	120	0.3820	0.0522	0.5376	0.1197	0.3998	0.0749	0.5093	0.1206	0.4002	0.0567	0.5824	0.1343	0.4180	0.0747	0.5414	0.1325
	240	0.3083	0.0422	0.4635	0.0988	0.3260	0.0560	0.4283	0.0994	0.3197	0.0404	0.4947	0.1020	0.3333	0.0562	0.4575	0.1012

		CVAR2_EW, RMSE%				CVAR2_CW, RMSE%				MSE_EW, RMSE%				MSE_CW, RMSE%			
Size		Estimation		Investment		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	0.2210	0.0394	0.2445	0.0457	0.2124	0.0344	0.2358	0.0450	0.2158	0.0394	0.2432	0.0455	0.2072	0.0346	0.2348	0.0457
	120	0.1628	0.0273	0.1951	0.0336	0.1541	0.0279	0.1857	0.0343	0.1580	0.0281	0.1935	0.0334	0.1489	0.0272	0.1846	0.0342
	240	0.1115	0.0173	0.1547	0.0238	0.1102	0.0240	0.1452	0.0255	0.1054	0.0170	0.1531	0.0238	0.1052	0.0255	0.1440	0.0258
2	60	0.2157	0.0393	0.2425	0.0468	0.2072	0.0352	0.2453	0.0462	0.2110	0.0400	0.2399	0.0456	0.2021	0.0347	0.2424	0.0456
	120	0.1593	0.0247	0.1881	0.0309	0.1523	0.0273	0.1950	0.0337	0.1552	0.0241	0.1859	0.0306	0.1469	0.0279	0.1929	0.0335
	240	0.1092	0.0150	0.1466	0.0233	0.1112	0.0216	0.1504	0.0248	0.1040	0.0146	0.1443	0.0224	0.1059	0.0223	0.1490	0.0247
3	60	0.1868	0.0317	0.3163	0.0634	0.1867	0.0328	0.3136	0.0650	0.1823	0.0317	0.3138	0.0640	0.1823	0.0325	0.3106	0.0650
	120	0.1336	0.0217	0.2440	0.0456	0.1344	0.0252	0.2377	0.0442	0.1292	0.0226	0.2414	0.0456	0.1291	0.0251	0.2359	0.0447
	240	0.0890	0.0136	0.1885	0.0340	0.0952	0.0212	0.1767	0.0329	0.0848	0.0139	0.1860	0.0340	0.0903	0.0212	0.1751	0.0317
4	60	0.2358	0.0373	0.3933	0.0691	0.2281	0.0367	0.3935	0.0759	0.2316	0.0370	0.3907	0.0694	0.2223	0.0363	0.3902	0.0741
	120	0.1699	0.0257	0.3156	0.0530	0.1682	0.0299	0.3152	0.0612	0.1652	0.0253	0.3139	0.0531	0.1628	0.0295	0.3132	0.0615
	240	0.1168	0.0170	0.2544	0.0435	0.1222	0.0236	0.2456	0.0458	0.1118	0.0161	0.2534	0.0443	0.1182	0.0235	0.2441	0.0461
5	60	0.4795	0.0680	0.6166	0.1414	0.4903	0.0817	0.5988	0.1369	0.4666	0.0672	0.6186	0.1409	0.4802	0.0836	0.6028	0.1365
	120	0.3825	0.0571	0.5303	0.1160	0.3949	0.0712	0.5050	0.1159	0.3705	0.0557	0.5352	0.1195	0.3856	0.0719	0.5088	0.1167
	240	0.3118	0.0414	0.4617	0.0969	0.3247	0.0532	0.4268	0.0965	0.3022	0.0406	0.4649	0.0974	0.3175	0.0533	0.4285	0.0941

Table 5.10: Comparison of Levels of Tracking Quality over Different Time Periods

This shows the same picture as Table 4.11 did for traditional methods. The tracking quality improves as we increase the size of the tracking portfolio. This holds for the estimation period and for the investment period, for all market phases, and for all STEM-based methods.

To analyse the absolute stability of the tracking quality, we compute the ratio of the tracking quality in the investment period to the tracking quality in the estimation period for every portfolio. The median, minimal, and maximum values of these ratios are reported in Table 5.11.

		SVM_EW			SVM_CW			CVAR_EW			CVAR_CW			CVAR2_EW			CVAR2_CW			MSE_EW			MSE_CW		
Size		Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max	Med	Min	Max
1	60	1.10	0.64	1.90	1.08	0.53	2.05	1.09	0.46	1.81	1.08	0.47	2.15	1.11	0.70	1.95	1.11	0.62	2.03	1.12	0.71	1.96	1.13	0.67	2.05
	120	1.20	0.64	1.89	1.18	0.67	1.98	1.17	0.57	1.94	1.14	0.56	1.99	1.20	0.78	2.04	1.20	0.56	2.18	1.22	0.76	2.06	1.24	0.58	2.20
	240	1.39	0.84	2.36	1.32	0.72	2.31	1.33	0.74	2.31	1.24	0.58	2.16	1.38	0.83	2.30	1.31	0.78	2.32	1.45	0.78	2.58	1.36	0.77	2.49
2	60	1.10	0.69	1.73	1.15	0.63	1.81	1.09	0.61	1.81	1.13	0.64	1.90	1.13	0.67	1.82	1.18	0.72	1.87	1.14	0.66	1.85	1.20	0.72	1.85
	120	1.19	0.77	1.98	1.25	0.77	1.99	1.12	0.69	1.98	1.19	0.68	1.98	1.17	0.79	2.01	1.27	0.75	2.03	1.19	0.80	2.08	1.30	0.77	2.00
	240	1.36	0.84	2.19	1.35	0.76	2.32	1.25	0.72	2.23	1.26	0.64	2.23	1.33	0.84	2.39	1.35	0.73	2.39	1.38	0.90	2.54	1.40	0.82	2.54
3	60	1.63	0.91	2.85	1.58	0.84	2.82	1.58	0.97	2.83	1.57	0.86	3.16	1.69	0.99	3.06	1.67	0.85	3.00	1.71	0.96	3.15	1.70	0.90	3.15
	120	1.82	1.10	3.08	1.73	0.93	3.13	1.67	0.90	3.02	1.64	0.88	3.38	1.82	1.12	3.17	1.77	0.90	3.37	1.86	0.99	3.30	1.83	0.95	3.33
	240	2.14	1.21	3.70	1.93	0.97	3.38	1.91	1.20	3.38	1.73	0.92	3.38	2.11	1.26	3.64	1.84	0.97	3.60	2.18	0.97	3.77	1.92	0.97	3.75
4	60	1.66	1.04	2.80	1.68	0.85	3.12	1.61	1.01	3.14	1.65	0.85	2.83	1.66	0.95	2.74	1.72	0.94	3.24	1.68	1.01	2.77	1.74	0.92	3.12
	120	1.85	1.19	3.02	1.85	1.00	3.14	1.77	1.15	3.79	1.77	0.95	3.79	1.84	1.14	3.19	1.87	0.98	3.16	1.89	1.16	3.27	1.91	0.96	3.27
	240	2.16	1.28	3.98	2.04	0.91	3.71	2.05	1.27	3.63	1.90	1.02	3.40	2.13	1.25	3.88	2.00	0.83	3.93	2.22	1.38	4.04	2.05	0.86	4.04
5	60	1.30	0.74	2.60	1.21	0.50	2.66	1.36	0.72	3.15	1.24	0.61	2.97	1.29	0.74	2.26	1.22	0.53	2.35	1.33	0.77	2.35	1.26	0.59	2.91
	120	1.41	0.84	2.32	1.27	0.57	2.30	1.46	0.80	2.46	1.30	0.59	2.57	1.39	0.83	2.41	1.28	0.59	2.33	1.45	0.87	2.40	1.32	0.62	2.25
	240	1.51	0.84	2.54	1.31	0.66	2.54	1.56	0.80	2.61	1.37	0.66	2.69	1.49	0.87	2.42	1.31	0.70	2.55	1.54	0.93	2.63	1.35	0.78	2.62

Table 5.11: Comparison of Ratios of Investment Period to Estimation Period Tracking Quality

When we compare these ratios with the ratios for the traditional measures (see Table 4.12), we can see that the ratios are lower for the STEM-based methods. Thus, the absolute stability of tracking quality is improved by using preliminary structures in the optimization problem.

### Relative stability of tracking quality

To examine the relative stability of the tracking quality, we select the 10% of portfolios that perform best in the estimation period. We select them according to the method of index tracking in use. The CVAR\_EW (CW), CVAR2\_EW (CW) and SVM\_EW (CW) methods are connected with the MAD measure. Therefore, we use this measure to select the best portfolios. For the MSE\_EW (CW) methods, we use the MSE measure. The results are reported in Table 5.12.

Size		SVM_EW, RMSE%				SVM_CW, RMSE%				CVAR_EW, RMSE%				CVAR_CW, RMSE%			
		Estimation		Investment		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	0.1863	0.0126	0.2130	0.0319	0.1882	0.0134	0.2115	0.0341	0.1983	0.0127	0.2276	0.0351	0.1965	0.0111	0.2225	0.0368
	120	0.1325	0.0108	0.1705	0.0264	0.1327	0.0103	0.1651	0.0291	0.1453	0.0107	0.1826	0.0305	0.1423	0.0097	0.1785	0.0303
	240	0.0908	0.0090	0.1398	0.0211	0.0916	0.0101	0.1344	0.0190	0.0986	0.0076	0.1481	0.0247	0.0964	0.0065	0.1435	0.0237
2	60	0.1818	0.0111	0.2111	0.0322	0.1813	0.0119	0.2221	0.0382	0.1946	0.0121	0.2313	0.0405	0.1895	0.0102	0.2336	0.0407
	120	0.1318	0.0092	0.1702	0.0251	0.1307	0.0092	0.1791	0.0281	0.1456	0.0101	0.1886	0.0309	0.1411	0.0094	0.1854	0.0275
	240	0.0921	0.0087	0.1389	0.0192	0.0903	0.0087	0.1403	0.0217	0.0999	0.0068	0.1437	0.0244	0.0969	0.0064	0.1463	0.0253
3	60	0.1598	0.0107	0.2779	0.0515	0.1646	0.0126	0.2791	0.0546	0.1717	0.0099	0.2993	0.0588	0.1718	0.0090	0.2965	0.0623
	120	0.1126	0.0084	0.2181	0.0382	0.1143	0.0096	0.2170	0.0389	0.1244	0.0079	0.2375	0.0441	0.1230	0.0075	0.2301	0.0439
	240	0.0750	0.0067	0.1759	0.0307	0.0762	0.0084	0.1665	0.0276	0.0814	0.0054	0.1805	0.0364	0.0816	0.0059	0.1795	0.0337
4	60	0.1992	0.0124	0.3522	0.0547	0.2025	0.0133	0.3587	0.0732	0.2133	0.0123	0.3762	0.0581	0.2112	0.0116	0.3732	0.0675
	120	0.1418	0.0098	0.2834	0.0470	0.1433	0.0111	0.2886	0.0584	0.1550	0.0099	0.3060	0.0541	0.1537	0.0093	0.2989	0.0542
	240	0.0993	0.0089	0.2346	0.0401	0.1002	0.0087	0.2317	0.0409	0.1061	0.0070	0.2429	0.0427	0.1074	0.0066	0.2428	0.0443
5	60	0.4128	0.0261	0.5699	0.1198	0.4192	0.0330	0.5572	0.1078	0.4231	0.0245	0.6117	0.1464	0.4300	0.0287	0.5939	0.1507
	120	0.3216	0.0225	0.4825	0.0959	0.3260	0.0231	0.4712	0.0994	0.3355	0.0229	0.5194	0.1135	0.3427	0.0229	0.5213	0.1061
	240	0.2608	0.0187	0.4201	0.0761	0.2680	0.0173	0.4064	0.0797	0.2683	0.0160	0.4449	0.0893	0.2736	0.0180	0.4331	0.0958

Size		CVAR2_EW, RMSE%				CVAR2_CW, RMSE%				MSE_EW, RMSE%				MSE_CW, RMSE%			
		Estimation		Investment		Estimation		Investment		Estimation		Investment		Estimation		Investment	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	0.1794	0.0115	0.2117	0.0332	0.1762	0.0105	0.2089	0.0333	0.1748	0.0110	0.2098	0.0315	0.1713	0.0096	0.2063	0.0328
	120	0.1312	0.0099	0.1708	0.0281	0.1262	0.0087	0.1663	0.0279	0.1257	0.0089	0.1696	0.0287	0.1212	0.0079	0.1646	0.0276
	240	0.0908	0.0073	0.1410	0.0205	0.0870	0.0064	0.1381	0.0237	0.0851	0.0060	0.1394	0.0222	0.0815	0.0055	0.1366	0.0216
2	60	0.1747	0.0116	0.2195	0.0378	0.1718	0.0100	0.2166	0.0347	0.1700	0.0102	0.2164	0.0391	0.1670	0.0089	0.2148	0.0359
	120	0.1305	0.0083	0.1792	0.0292	0.1260	0.0079	0.1798	0.0293	0.1262	0.0072	0.1779	0.0307	0.1214	0.0067	0.1770	0.0277
	240	0.0917	0.0061	0.1393	0.0243	0.0886	0.0064	0.1431	0.0223	0.0869	0.0058	0.1370	0.0230	0.0833	0.0059	0.1413	0.0224
3	60	0.1543	0.0087	0.2851	0.0536	0.1543	0.0091	0.2808	0.0550	0.1500	0.0082	0.2826	0.0539	0.1501	0.0082	0.2807	0.0542
	120	0.1104	0.0071	0.2282	0.0408	0.1096	0.0068	0.2233	0.0398	0.1060	0.0065	0.2265	0.0394	0.1053	0.0059	0.2218	0.0381
	240	0.0734	0.0051	0.1790	0.0330	0.0745	0.0057	0.1755	0.0340	0.0687	0.0045	0.1761	0.0327	0.0693	0.0049	0.1728	0.0325
4	60	0.1952	0.0130	0.3568	0.0654	0.1913	0.0127	0.3492	0.0612	0.1912	0.0115	0.3486	0.0561	0.1870	0.0105	0.3456	0.0647
	120	0.1406	0.0101	0.2890	0.0506	0.1385	0.0082	0.2862	0.0511	0.1356	0.0083	0.2871	0.0463	0.1334	0.0082	0.2836	0.0523
	240	0.0982	0.0069	0.2380	0.0411	0.0984	0.0065	0.2332	0.0410	0.0939	0.0061	0.2334	0.0394	0.0949	0.0062	0.2319	0.0406
5	60	0.4005	0.0276	0.5404	0.1053	0.4041	0.0272	0.5414	0.1122	0.3884	0.0222	0.5467	0.1100	0.3961	0.0227	0.5467	0.1164
	120	0.3162	0.0232	0.4698	0.0978	0.3244	0.0215	0.4719	0.0982	0.3064	0.0183	0.4658	0.0972	0.3140	0.0194	0.4695	0.0931
	240	0.2619	0.0174	0.4196	0.0788	0.2684	0.0174	0.4057	0.0834	0.2533	0.0126	0.4192	0.0785	0.2597	0.0140	0.4007	0.0801

Table 5.12: Levels of Tracking Quality over the Five Time Periods (Best 10% of 5,000)

To analyse whether there is an improvement, we compute the ratio of the tracking quality of all 5,000 portfolios to the tracking quality of the best 10%. The results are shown in Table 5.13.

Size		SVM_EW		SVM_CW		CVAR_EW		CVAR_CW		CVAR2_EW		CVAR2_CW		MSE_EW		MSE_CW	
		Est	Inv	Est	Inv	Est	Inv	Est	Inv	Est	Inv	Est	Inv	Est	Inv	Est	Inv
1	60	1.21	1.16	1.20	1.16	1.23	1.17	1.17	1.21	1.23	3.43	1.15	1.38	1.21	3.28	1.13	1.35
	120	1.22	1.14	1.21	1.14	1.25	1.16	1.16	1.23	1.24	2.75	1.14	1.20	1.22	3.19	1.12	1.23
	240	1.22	1.11	1.22	1.09	1.27	1.12	1.12	1.28	1.23	2.35	1.10	1.16	1.27	3.73	1.05	1.08
2	60	1.21	1.14	1.22	1.14	1.23	1.12	1.12	1.22	1.23	3.39	1.10	1.24	1.21	3.52	1.13	1.33
	120	1.19	1.11	1.20	1.10	1.22	1.06	1.06	1.21	1.22	2.98	1.05	1.06	1.21	3.44	1.08	1.15
	240	1.19	1.08	1.23	1.08	1.22	1.06	1.06	1.27	1.19	2.45	1.05	0.96	1.25	3.37	1.05	1.11
3	60	1.21	1.13	1.21	1.13	1.21	1.10	1.10	1.21	1.21	3.65	1.11	1.18	1.21	3.60	1.12	1.18
	120	1.19	1.12	1.21	1.10	1.22	1.07	1.07	1.23	1.21	3.06	1.07	1.12	1.23	3.70	1.06	1.11
	240	1.19	1.09	1.24	1.08	1.23	1.07	1.07	1.29	1.21	2.65	1.05	1.03	1.28	3.75	1.01	0.97
4	60	1.19	1.12	1.20	1.13	1.20	1.11	1.11	1.19	1.21	2.87	1.10	1.06	1.19	2.89	1.13	1.24
	120	1.20	1.12	1.20	1.11	1.21	1.10	1.10	1.21	1.21	2.54	1.09	1.05	1.21	3.63	1.10	1.20
	240	1.18	1.09	1.21	1.07	1.21	1.09	1.09	1.24	1.19	2.46	1.07	1.06	1.24	3.61	1.05	1.12
5	60	1.19	1.12	1.21	1.10	1.20	1.13	1.13	1.21	1.20	2.46	1.14	1.34	1.21	3.01	1.11	1.22
	120	1.19	1.11	1.23	1.08	1.19	1.12	1.12	1.22	1.21	2.46	1.13	1.19	1.22	3.31	1.07	1.18
	240	1.18	1.10	1.22	1.05	1.19	1.11	1.11	1.22	1.19	2.38	1.10	1.23	1.21	3.05	1.05	1.16

Table 5.13: Ratios of Tracking Qualities (5,000/Best 10%)

The results indicate that the difference between all 5,000 portfolios and the best 10% is not large, with the ratios near to unity in most cases. However, the tracking quality is more than doubled when we select the best 10% for the CVAR2\_EW and MSE\_EW methods.

To examine the relative stability further, we calculate the stability ratio, which was introduced in Section 3.4.2. The results are reported in Table 5.14.

Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	57.61%	61.80%	56.83%	67.34%	57.46%	67.47%	59.55%
	120	60.01%	67.35%	56.73%	73.23%	62.02%	75.45%	61.53%
	240	68.32%	77.63%	63.58%	86.85%	69.07%	92.59%	68.57%
2	60	60.62%	68.19%	62.40%	65.92%	69.84%	67.00%	70.18%
	120	67.28%	71.11%	74.63%	72.89%	81.04%	73.77%	78.34%
	240	74.70%	79.30%	77.38%	83.80%	80.52%	88.26%	78.15%
3	60	68.75%	71.59%	76.13%	79.74%	79.86%	80.34%	79.70%
	120	70.68%	74.53%	82.08%	85.99%	84.95%	89.88%	83.30%
	240	76.79%	82.18%	81.68%	93.96%	84.99%	99.08%	81.60%
4	60	66.43%	71.50%	69.83%	74.04%	69.34%	72.11%	71.22%
	120	67.29%	72.72%	67.99%	77.67%	69.75%	78.41%	69.69%
	240	73.88%	82.53%	68.78%	88.88%	70.49%	89.28%	70.69%
5	60	76.92%	84.77%	78.63%	89.77%	74.51%	85.05%	72.99%
	120	76.41%	91.68%	76.44%	94.31%	73.06%	92.13%	68.78%
	240	77.54%	94.80%	77.36%	92.68%	77.49%	94.54%	71.56%

Table 5.14: Stability Ratios for the STEM-Based Methods

The relative stability of tracking quality is better for the methods that use the EW structures than for those that use the CW structures. Comparing the stability ratios with those for the traditional methods, shown in Table 4.14, we can also observe an improvement in the relative stability of tracking quality in most cases when the EW structures are used.

To verify the results, the Spearman and Kendall rank correlation coefficients are calculated (see Table 5.15 and Table 5.16).

Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	44.79%*	40.55%*	45.42%*	35.11%*	45.30%*	35.26%*	42.76%*
	120	42.20%*	35.24%*	45.22%*	28.80%*	40.34%*	26.89%*	40.11%*
	240	33.70%*	24.89%*	38.57%*	15.63%*	32.77%*	9.93%*	33.10%*
2	60	41.54%*	33.83%*	39.43%*	36.24%*	32.00%*	34.97%*	32.01%*
	120	34.66%*	30.64%*	27.86%*	29.30%*	20.75%*	28.07%*	23.17%*
	240	27.42%*	22.85%*	24.11%*	18.43%*	20.91%*	13.75%*	23.42%*
3	60	33.74%*	30.84%*	26.54%*	22.49%*	22.79%*	21.99%*	22.99%*
	120	32.00%*	27.73%*	20.18%*	15.91%*	17.29%*	12.57%*	18.99%*
	240	25.29%*	19.74%*	20.13%*	7.52%*	17.28%*	1.63%	20.19%*
4	60	36.25%*	30.64%*	32.49%*	27.62%*	33.20%*	29.87%*	31.12%*
	120	34.96%*	29.36%*	34.02%*	23.92%*	32.47%*	23.80%*	32.20%*
	240	28.08%*	18.72%*	33.36%*	12.67%*	31.29%*	12.29%*	31.35%*
5	60	25.17%*	16.72%*	23.19%*	11.54%*	27.68%*	16.31%*	28.84%*
	120	25.90%*	9.53%*	25.75%*	6.74%*	29.16%*	9.25%*	33.75%*
	240	24.28%*	6.91%*	24.65%*	8.59%*	24.64%*	6.67%*	30.18%*

\* Significantly different from zero with level 0.05

Table 5.15: Kendall Rank Correlation Coefficients for the STEM-Based Methods

Size		SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	62.95*	57.55*	63.68*	50.54*	63.68*	50.78*	60.28*	51.55*
	120	59.52*	50.66*	63.18*	42.05*	57.05*	39.37*	56.33*	40.15*
	240	48.46*	36.59*	54.93*	23.31*	47.13*	14.89*	47.29*	23.00*
2	60	58.70*	48.76*	55.98*	51.96*	46.08*	50.25*	46.13*	52.78*
	120	49.91*	44.49*	40.64*	42.64*	30.27*	40.96*	33.32*	46.50*
	240	40.07*	33.69*	35.12*	27.39*	30.56*	20.50*	34.20*	24.20*
3	60	48.93*	44.83*	39.03*	33.23*	33.78*	32.52*	34.09*	30.72*
	120	46.51*	40.54*	29.94*	23.67*	25.76*	18.83*	28.11*	19.44*
	240	37.15*	29.14*	29.82*	11.29%	25.76*	2.36%	29.53*	5.70*
4	60	52.12*	44.42*	47.09*	40.33*	47.97*	43.40*	45.15*	44.36*
	120	50.34*	42.89*	49.06*	35.10*	46.99*	35.05*	46.32*	37.76*
	240	40.81*	27.64%	47.81*	18.94%	44.91*	18.36%	44.74*	19.61*
5	60	36.96*	24.82%	34.12*	17.26%	40.55*	24.15*	42.07*	25.65*
	120	38.13*	14.08*	37.85*	10.05*	42.74*	13.82*	48.61*	18.90%
	240	35.76*	10.31*	36.33*	12.87*	36.35*	10.01*	43.40*	19.96%

\* Significantly different from zero with level 0.05

Table 5.16: Spearman Correlation Coefficients for the STEM-Based Methods

The results confirm the previous findings. Moreover, the tests indicate that all coefficients are significantly larger than zero. Hence, the tracking quality expressed in RMSE in the investment period is significantly positively dependent on the tracking quality expressed in RMSE in the estimation period. However, the correlation coefficients are lower than 50% in most of the cases. This shows that the tracking quality measure is not a good indicator of relative stability to use when selecting tracking portfolios.

To give an overview, we report the superior methods, according to relative stability, in Table 5.17.

		Stability ratio	Kendall Rank Correlation	Spearman Rank Correlation
1	60	CVAR_EW	CVAR_EW	CVAR_EW, CVAR2_EW
	120	CVAR_EW	CVAR_EW	CVAR_EW
	240	CVAR_EW	CVAR_EW	CVAR_EW
2	60	SVM_EW	SVM_EW	SVM_EW
	120	SVM_EW	SVM_EW	SVM_EW
	240	SVM_EW	SVM_EW	TEV
3	60	SVM_EW	SVM_EW	SVM_EW
	120	SVM_EW	SVM_EW	SVM_EW
	240	MAD	MAD	MAD
4	60	SVM_EW	SVM_EW	SVM_EW
	120	SVM_EW	SVM_EW	SVM_EW
	240	CVAR_EW	CVAR_EW	CVAR_EW
5	60	MSE_EW	MSE_EW	MSE_EW
	120	MSE_EW	MSE_EW	MSE_EW
	240	MSE_EW	MSE_EW	MSE_EW

Table 5.17: Superior Methods, According to Relative Stability

It can be observed that the use of EW structures in the optimization problem helps to improve the relative stability of tracking quality.

## Prediction of ETQ

In the next piece of analysis, we look at the ability to predict the ETQs for STEM-based methods in the investment period from the values of measures of tracking quality in the



estimation period. The tracking quality of the portfolios, generated by the STEM-based methods, is reported in Table 5.18. These results are calculated for the same tracking portfolios as were reported in Table 5.10, but here, instead of the RMSE measure, we report ETQs.

Size		SVM_EW		SVM_CW		CVAR_EW		CVAR_CW		CVAR2_EW		CVAR2_CW		MSE_EW		MSE_CW	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	1.6910	1.2949	1.1723	1.1368	1.8087	1.5040	1.2625	1.2425	1.5404	1.3220	1.0812	1.0581	1.5372	1.3405	1.1203	1.1001
	120	1.2521	1.0074	0.8300	0.7935	1.4692	1.1633	0.9042	0.8858	1.3032	1.0193	0.7940	0.7797	1.3142	1.0208	0.8146	0.7789
	240	0.9393	0.7581	0.6445	0.5905	1.1356	0.8356	0.6983	0.6716	1.0014	0.7119	0.6091	0.5623	0.9969	0.7110	0.6176	0.5902
2	60	0.7286	0.6179	0.8192	0.7427	0.7809	0.6794	0.8436	0.7712	0.7474	0.6485	0.8117	0.7245	0.7366	0.6549	0.8183	0.7333
	120	0.5530	0.4555	0.6446	0.5530	0.5667	0.4714	0.6691	0.5547	0.5412	0.4388	0.6455	0.5338	0.5341	0.4406	0.6495	0.5501
	240	0.4648	0.3967	0.5092	0.4308	0.4325	0.3618	0.5006	0.4195	0.4318	0.3565	0.5011	0.4219	0.4243	0.3629	0.4967	0.4195
3	60	1.0212	0.8505	1.0317	0.8728	1.0643	0.8986	1.0577	0.9020	1.0239	0.8763	1.0133	0.8564	1.0066	0.8557	1.0020	0.8445
	120	0.8066	0.6698	0.7736	0.6576	0.8242	0.6920	0.7955	0.6730	0.8071	0.6770	0.7741	0.6343	0.7988	0.6698	0.7576	0.6291
	240	0.6536	0.5682	0.5925	0.5095	0.6704	0.6019	0.6023	0.5003	0.6575	0.5701	0.5746	0.5025	0.6569	0.5801	0.5729	0.4930
4	60	1.3987	1.1607	1.4773	1.2700	1.4527	1.1586	1.5121	1.2817	1.4198	1.1602	1.4703	1.2895	1.4032	1.1230	1.4458	1.2499
	120	1.1760	0.9397	1.1875	1.0413	1.2161	0.9356	1.2108	1.0427	1.1734	0.9375	1.1850	1.0666	1.1643	0.9160	1.1868	1.0534
	240	1.0139	0.8108	0.9756	0.8518	1.0094	0.7865	0.9449	0.8420	1.0012	0.7912	0.9573	0.8693	1.0078	0.7787	0.9515	0.8682
5	60	3.5418	2.9085	2.8386	2.6262	3.8496	3.3945	3.0978	2.9900	3.2011	2.7746	2.8068	2.5331	3.2186	2.7755	2.9025	2.5719
	120	3.0826	2.3280	2.4610	2.2494	3.4141	2.7624	2.7340	2.4939	2.9820	2.3618	2.4798	2.2333	3.0567	2.4547	2.5410	2.2526
	240	2.7970	1.8581	2.4160	1.9374	3.0537	2.0232	2.7081	2.0829	2.8163	1.9289	2.3912	1.9344	2.8607	1.9063	2.4580	1.9176

Table 5.18: Ex Post Tracking Quality for the STEM-Based Methods

To verify whether it is possible to improve the average tracking quality in the investment period by selecting the best portfolios in the estimation period, we select the best 10% of the portfolios, according to their tracking quality in the estimation period. The medians and IQRs for this subset are reported in Table 5.19.

Size		SVM_EW		SVM_CW		CVAR_EW		CVAR_CW		CVAR2_EW		CVAR2_CW		MSE_EW		MSE_CW	
		Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR	Med	IQR
1	60	1.0659	0.9464	0.8629	0.7799	1.0722	0.9588	0.8823	0.7981	0.8522	0.7576	0.7534	0.7111	0.8466	0.7631	0.7443	0.6912
	120	0.8852	0.7744	0.5974	0.5237	0.7925	0.7731	0.6743	0.6735	0.6799	0.6137	0.6326	0.5540	0.6785	0.6110	0.5955	0.5266
	240	0.6337	0.5943	0.5048	0.4498	0.6066	0.5628	0.5344	0.5075	0.6383	0.5864	0.5012	0.4035	0.5947	0.5150	0.5000	0.4391
2	60	0.6930	0.5744	0.7792	0.7728	0.8101	0.7344	0.8485	0.7870	0.8143	0.8122	0.7725	0.7277	0.7917	0.8044	0.7974	0.7920
	120	0.5625	0.5173	0.6518	0.5824	0.6436	0.6272	0.6392	0.5747	0.6461	0.6308	0.6490	0.6343	0.6611	0.6396	0.6799	0.6463
	240	0.4644	0.4322	0.4923	0.4445	0.4772	0.4897	0.4999	0.4375	0.4792	0.4126	0.5289	0.4926	0.4707	0.4383	0.5043	0.4551
3	60	0.9231	0.7281	0.9280	0.7406	0.9599	0.8089	0.9620	0.8358	0.8842	0.7018	0.9335	0.7306	0.8840	0.6945	0.9235	0.7406
	120	0.6754	0.5381	0.6537	0.5280	0.7330	0.5583	0.7038	0.5355	0.7011	0.5344	0.7047	0.4905	0.6912	0.5411	0.6651	0.5057
	240	0.5720	0.4338	0.5317	0.4086	0.6161	0.4909	0.5686	0.4567	0.5858	0.4534	0.5292	0.4208	0.5727	0.4512	0.5203	0.3875
4	60	1.3437	1.0972	1.3960	1.0078	1.4048	1.2001	1.3721	1.2102	1.4364	1.2809	1.3603	1.1975	1.4107	1.1457	1.3263	1.1814
	120	1.0749	0.9317	1.1283	0.9742	1.1501	0.9674	1.1156	0.9559	1.1679	1.0641	1.0605	0.9193	1.1587	1.0305	1.0780	0.9543
	240	0.9950	0.7401	0.9345	0.7100	0.9354	0.8291	0.9219	0.7854	0.9460	0.8318	0.9127	0.7238	0.9904	0.7609	0.8685	0.7220
5	60	3.0154	2.4567	2.6328	2.5182	3.1336	2.9749	2.9677	2.7737	2.4316	2.1299	2.3814	2.1053	2.3869	2.1125	2.4982	2.1078
	120	2.6657	1.9347	2.4541	1.9584	2.6546	2.2566	2.8370	2.2737	2.2190	1.8928	2.4248	1.8645	2.1249	1.8264	2.3190	1.7230
	240	2.4021	1.5145	2.3767	1.6718	2.5392	1.6988	2.7366	1.8348	2.2966	1.6988	2.3227	1.7187	2.3017	1.5774	2.3180	1.7485

Table 5.19: Comparison of Ex Post Tracking Quality (Best 10%)

To compare the tracking quality of all 5,000 tracking portfolios (see Table 5.18) with the top 10%, we present the ratios of one to the other in Table 5.20.

	Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	1.59	1.37	1.36	1.46	1.69	1.57	1.43	1.56
	120	1.41	1.30	1.39	1.52	1.85	1.50	1.34	1.32
	240	1.48	1.28	1.28	1.31	1.87	1.48	1.31	1.32
2	60	1.05	1.08	1.05	0.96	0.96	0.93	0.99	0.98
	120	0.98	0.88	0.99	0.95	0.88	0.75	1.05	0.97
	240	1.00	0.92	1.03	0.97	0.91	0.74	1.00	0.96
3	60	1.11	1.17	1.11	1.18	1.11	1.11	1.10	1.08
	120	1.19	1.24	1.18	1.25	1.12	1.24	1.13	1.26
	240	1.14	1.31	1.11	1.25	1.09	1.23	1.06	1.10
4	60	1.04	1.06	1.06	1.26	1.03	0.97	1.10	1.06
	120	1.09	1.01	1.05	1.07	1.06	0.97	1.09	1.09
	240	1.02	1.10	1.04	1.20	1.08	0.95	1.02	1.07
5	60	1.17	1.18	1.08	1.04	1.23	1.14	1.04	1.08
	120	1.16	1.20	1.00	1.15	1.29	1.22	0.96	1.10
	240	1.16	1.23	1.02	1.16	1.20	1.19	0.99	1.14

Table 5.20: Ratios of Ex Post Tracking Quality (5,000 / Best 10%)

The results in Table 5.20 indicate that the tracking quality is only slightly improved in the top 10%. For period 2, there is even a decrease in the tracking quality when we select the 500 portfolios that perform best in the estimation period.

Next, we examine further whether it is possible to predict the ETQ from the tracking quality in the estimation period. To do this, we calculate the stability ratios (see Table 5.21).

	Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	76.85%	82.29%	70.83%	82.43%	67.12%	82.13%	66.53%	79.56%
	120	78.61%	84.63%	69.00%	85.06%	68.94%	85.44%	66.82%	83.31%
	240	84.28%	89.09%	71.96%	89.33%	75.82%	92.40%	73.63%	87.46%
2	60	96.95%	99.40%	98.62%	99.62%	103.82%	99.39%	103.37%	98.51%
	120	99.69%	98.21%	103.04%	97.94%	105.23%	100.35%	103.34%	98.65%
	240	99.24%	98.73%	102.01%	101.13%	102.03%	102.81%	102.01%	101.96%
3	60	93.22%	94.59%	94.97%	95.48%	94.49%	96.88%	94.76%	96.17%
	120	93.18%	93.39%	95.64%	95.96%	93.80%	96.13%	93.02%	95.43%
	240	95.01%	94.06%	94.00%	98.29%	94.26%	98.47%	92.87%	97.13%
4	60	95.63%	97.61%	96.17%	96.10%	98.73%	95.43%	97.97%	95.21%
	120	94.35%	96.93%	96.73%	95.96%	97.84%	96.47%	97.34%	95.76%
	240	98.65%	100.10%	96.98%	99.76%	98.13%	100.77%	97.42%	98.43%
5	60	92.14%	99.05%	89.92%	99.63%	86.96%	96.91%	85.04%	95.91%
	120	90.72%	103.46%	87.39%	102.75%	83.74%	102.67%	78.81%	99.35%
	240	90.72%	104.16%	88.15%	102.58%	86.49%	104.29%	80.69%	101.02%

Table 5.21: Stability Ratios for ETQ

Table 5.21 indicates that the predictability of the ETQ from the measures of tracking quality is low. In some of the cases, the stability ratio is close to 100%, indicating a random result. To verify these results, we compute the Kendall and Spearman rank correlation coefficients (see Table 5.22 and Table 5.23).

Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	25.50%*	19.70%*	31.49%*	19.47%*	35.51%*	19.55%*	22.31%*
	120	22.74%*	16.83%*	33.17%*	16.16%*	33.51%*	16.17%*	18.37%*
	240	16.90%*	12.30%*	29.98%*	12.05%*	26.15%*	8.80%*	13.72%*
2	60	3.74%*	1.28%	1.53%	1.38%	-3.84%*	1.37%	-3.52%*
	120	0.50%	2.12%*	-2.96%*	2.38%*	-5.74%*	-0.50%	-4.30%*
	240	0.85%	1.72%	-2.51%*	-0.85%	-3.03%*	-2.86%*	-3.12%*
3	60	8.32%*	6.07%*	5.91%*	4.90%*	6.50%*	3.90%*	6.31%*
	120	7.96%*	7.20%*	5.35%*	4.85%*	6.61%*	4.65%*	7.82%*
	240	6.14%*	6.92%*	6.90%*	1.80%	7.14%*	1.81%*	8.51%*
4	60	5.51%*	3.49%*	4.56%*	5.00%*	1.59%	5.34%*	2.96%*
	120	6.37%*	3.64%*	4.13%*	4.58%*	3.19%*	4.66%*	3.88%*
	240	2.47%*	0.61%	4.07%*	0.64%	2.97%*	0.22%	4.02%*
5	60	9.16%*	1.63%	11.35%*	0.72%	15.08%*	3.81%*	16.49%*
	120	11.09%*	-3.39%*	14.02%*	-2.74%*	18.12%*	-2.47%*	23.37%*
	240	10.47%*	-4.38%*	13.32%*	-2.99%*	14.82%*	-4.42%*	20.51%*

\* Significantly different from zero with level 0.05

Table 5.22: Kendall Rank Correlation Coefficients for ETQ

Size	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
1	60	37.50%*	29.23%*	45.98%*	28.89%*	51.59%*	28.89%*	52.21%*
	120	33.33%*	24.93%*	48.08%*	24.01%*	48.39%*	23.96%*	50.68%*
	240	24.82%*	18.33%*	43.58%*	17.95%*	38.21%*	13.12%*	40.89%*
2	60	5.57%*	1.92%	2.25%	2.07%	-5.84%*	2.04%	-5.36%*
	120	0.70%	3.15%*	-4.51%*	3.58%*	-8.73%*	-0.79%	-6.53%*
	240	1.27%	2.58%*	-3.82%*	-1.34%	-4.53%*	-4.33%*	-4.68%*
3	60	12.48%*	9.07%*	8.87%*	7.33%*	9.73%*	5.87%*	9.45%*
	120	11.90%*	10.75%*	8.03%*	7.26%*	9.86%*	6.92%*	11.67%*
	240	9.21%*	10.36%*	10.36%*	2.69%	10.67%*	2.70%	12.77%*
4	60	8.29%*	5.25%*	6.81%*	7.53%*	2.36%	8.01%*	4.41%*
	120	9.54%*	5.48%*	6.19%*	6.87%*	4.73%*	7.00%*	5.77%*
	240	3.72%*	0.88%	6.07%*	0.94%	4.47%*	0.25%	5.97%*
5	60	13.71%*	2.46%	16.95%*	1.11%	22.55%*	5.69%*	24.55%*
	120	16.64%*	-5.11%*	20.88%*	-4.18%*	27.09%*	-3.82%*	34.48%*
	240	15.67%*	-6.71%*	19.90%*	-4.61%*	22.16%*	-6.77%*	30.05%*

\* Significantly different from zero with level 0.05

Table 5.23: Spearman Rank Correlation Coefficients for ETQ

The coefficients indicate the low dependence between the tracking quality in the estimation period and the ETQ in the investment period. Moreover, there are cases where the correlation coefficients are not significantly different from zero or even take negative values.

The analysis of the results demonstrates the poor ability of the tracking quality measures in the estimation period to predict the ETQ in the investment period. Predicting ETQ based on the tracking quality in the estimation period also depends on how close the portfolios are to each other in the estimation period. If they are very close to each other, then it is difficult to distinguish which will produce higher tracking quality in the investment period.

### Comparison of the methods in the investment period

Next, we compare the traditional methods of index tracking with the STEM-based methods. In Table 5.24, we report the superior methods based on comparing the median ETQs for the traditional methods (reported in Table 4.17 in Section 4.2) with the median ETQs for the STEM-based methods given in Table 5.18.

	Period 1	Period 2	Period 3	Period 4	Period 5
60	CVAR2_CW	SVM_EW	MSE_CW	SVM_EW	CVAR2_CW
120	CVAR2_CW	MSE_EW	MSE_CW	MSE_EW	SVM_CW
240	CVAR2_CW	MSE_EW	MSE_CW	CVAR_CW	CVAR2_CW

Table 5.24: Best Methods Based on a Comparison of Median ETQs

The results indicate that it is possible to improve tracking quality by using structures.

To check the significance of our findings, we compare the methods using a paired Wilcoxon test. We test the null hypothesis that the median ETQ for one method is greater than or equal to the median ETQ for another method. Then we build a table, showing the results of this hypothesis for each pair of methods. If the hypothesis is rejected, then a “Y” is placed in the intersection of the column and row for this pair, otherwise a “N” is put in the cell. Table 5.25 shows an example of such a table for period 1 and a tracking portfolio (TP) of size 60.

Period 1, TP Size 60

	MAD	MSE	TEV	SVM_EW	SVM_CW	CVAR_EW	CVAR_CW	CVAR2_EW	CVAR2_CW	MSE_EW	MSE_CW
MAD	-	N	N	Y	Y	Y	Y	Y	N	Y	N
MSE	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y
TEV	Y	N	-	Y	Y	Y	Y	Y	N	Y	Y
SVM_EW	N	N	N	-	N	Y	N	N	N	N	N
SVM_CW	N	N	N	Y	-	Y	Y	Y	N	Y	N
CVAR_EW	N	N	N	N	N	-	N	N	N	N	N
CVAR_CW	N	N	N	Y	N	Y	-	Y	N	Y	N
CVAR2_EW	N	N	N	Y	N	Y	N	-	N	N	N
CVAR2_CW	Y	N	N	Y	Y	Y	Y	Y	-	Y	Y
MSE_EW	N	N	N	Y	N	Y	N	Y	N	-	N
MSE_CW	N	N	N	Y	Y	Y	Y	Y	N	Y	-

Table 5.25: Comparison Table for the Significance Tests

The results in this table can be read as follows: is the median for the method in the row significantly lower than the median for the method in the column? The results indicate that the best method is MSE, because every cell in that row contains a “Y”, meaning that the median ETQ for the MSE is lower than the medians of all the other methods.

In the same way, we compare the median ETQs for the other time periods and portfolio sizes. The best methods according to these comparison are presented in Table 5.26.

	Period 1	Period 2	Period 3	Period 4	Period 5
60	MSE	SVM_EW	MSE_EW, MSE_CW	SVM_EW, MSE_EW	CVAR2_CW
120	MSE	CVAR2_EW, MSE_EW	MSE_CW	MSE_EW	SVM_CW, CVAR2_CW
240	CVAR2_CW, MSE_CW	MSE_EW	MSE_CW	CVAR_CW, MSE_CW	SVM_CW, CVAR2_CW

Table 5.26: Best Methods as Indicated by Significance Tests Comparing ETQs

The results indicate that the STEM-based methods are significantly better than the traditional methods in most cases. It is interesting to note that the objectives that are robust to outliers (CVAR and SVM) appear to be the best objectives to use in turbulent market phases (as

shown by our period 5). In other situations, it is sufficient to use the standard MSE objective with either the EW or CW structure.

Additionally, we test the tracking quality expressed in RMSE. The results of comparing the median RMSEs and the significance tests are reported in Table 5.27 and Table 5.28 respectively.

	Period 1	Period 2	Period 3	Period 4	Period 5
<b>60</b>	MSE	TEV	MSE_CW	TEV	CVAR2_CW
<b>120</b>	MSE_CW	MSE_EW	MSE_CW	TEV	CVAR2_CW
<b>240</b>	MSE_CW	MSE_EW	MSE_CW	MSE_CW	CVAR2_CW

Table 5.27: Best Methods According to a Comparison of Median RMSEs.

	Period 1	Period 2	Period 3	Period 4	Period 5
<b>60</b>	MSE_CW	TEV	MSE_CW	TEV	CVAR2_CW
<b>120</b>	MSE_CW	MSE_EW	MSE_CW	TEV	CVAR2_CW
<b>240</b>	MSE_CW	MSE_EW	MSE_CW	MSE_CW	CVAR2_CW

Table 5.28: Best Methods as Indicated by Significance Tests Comparing RMSEs

The results once again indicate that the STEM-based methods are superior to the traditional methods in most of the cases.

### 5.4.3.3 Conclusion

Our four analyses of STEM-based methods indicate the following:

- In comparison to traditional methods, the STEM-based methods reduce the discrepancy in tracking quality between the estimation and investment periods, and as a consequence improve the absolute stability of tracking quality.
- The relative stability of tracking quality can be improved by using STEM-based methods with EW structures.
- The predictability of ETQ is not generally improved by including additional information in the optimization problem.
- Using STEM-based methods rather than traditional methods, it is possible to improve the average tracking quality. This is confirmed by the paired Wilcoxon test.

## 5.5 Summary

In this chapter, first, the SVM and SVR machine learning algorithms were described. These algorithms are constructed so as to focus on model performance over a test set and to solve the problem of overfitting to achieve a better generalization of the data. The problem of overfitting occurs when it is necessary to select one model from a complex set of models using only a small number of data points. The complexity or capacity of the set of models is measured using the VC dimension or DoF. Hence, it is possible to avoid overfitting by controlling the capacity of the set of models. The SRM principle allows us to determine the optimal capacity of the set of models, by using analytic and resampling model selection methods. Hence, the model selection methods help us to find appropriate parameterizations for the SVM and SVR methods. However, in the SVR algorithm, we have to select one further parameter: the size of the insensitive region. To do this, we can use the  $\nu$ -SVR formulation. Another method is to transform the SVR into the SVM, which does not require this parameter.

Next, we applied the SV techniques to an index tracking problem. This method has two important properties. First, it is robust to outliers, since it is based on the  $\varepsilon$ -insensitive loss function. Second, there is a capacity control on the set of models, which makes the estimation of the tracking portfolio resistant to the data used for the estimation.

To find the optimal capacity of the set of models, the STEM approach was introduced. STEM-based methods form a group of methods for index tracking, combining the data with preliminary knowledge expressed in informational structures. These methods have roots in statistical learning theory and can be based on different measures of tracking quality.

Then, we compared the STEM-based methods with other approaches to portfolio optimization and index tracking. The STEM-based methods extend the properties of the other methods, by using flexible informational structures in the space of portfolio weights.

Finally, we conducted an empirical comparison between the STEM-based methods and the traditional methods. The empirical results show that, by using STEM-based methods, we can improve the resistance of the tracking portfolio estimation by controlling the complexity of the set of models. The model selection methods can be used to determine the complexity of the model set, and this can make the STEM-based methods superior to the traditional methods in terms of their tracking quality in the investment period.

Additionally, we have shown that the STEM-based methods generate tracking portfolios with a better absolute stability of tracking quality than those produced by traditional methods. If we use an EW structure, it is also possible to improve the relative stability.

Our comparison of ex post tracking quality in the investment period indicated that, by using STEM-based methods rather than traditional, it is possible to improve the quality of tracking. Moreover, using objectives that are less dependent on outliers, for example CVAR or SVM, it is possible to lower the influence of extreme data points on the construction of tracking portfolios.

## 6 Concluding remarks

There are different methods of tracking indices, which can be classified into two groups: physical replication and synthetic replication. Each of the methods of index tracking have strengths and weaknesses as we have shown in Chapter 2. However, sampling is a physical replication method that can be applied to any index. It is more flexible than full replication, and more transparent than synthetic. These properties make it a good alternative for specific application.

But in Chapter 4 we have shown that traditional methods of sampling, which we described in Chapter 3, have weaknesses. They construct tracking portfolios that lack absolute and relative stability of tracking quality, and the estimation of tracking portfolio weights is not resistant to slight modifications of the data used for estimation. To overcome these weaknesses, we have developed in Chapter 5 a new approach to sampling, called structural tracking error minimization (STEM). This approach is based on statistical learning theory. The application of this theory has shown that the following factors should be considered when using optimization techniques to solve index tracking problem:

- the size of the dataset that is used for the tracking portfolio estimation;
- the number of assets in the tracking portfolio, whose weights we need to estimate;
- the capacity of the set of admissible portfolios that is used for the tracking portfolio estimation.

If there is a lack of data, or the data are not reliable, which is often the case in real-life problems, then it is possible to improve the tracking portfolio estimation by adding additional information about the portfolio weights. By adding this information, we reduce the set of admissible portfolios and, as a result, the tracking portfolio estimation is not so strongly affected by the data.

We empirically tested the STEM-based methods and compared them with the traditional methods. These methods may improve the absolute and relative stability of tracking quality, solve the overfitting effect, and increase the resistance of the tracking portfolio estimation.

The abovementioned properties of the STEM-based methods make them highly relevant for practical use, where the stability of the solution and its prediction quality are of high importance. Moreover, we have demonstrated by solving the index tracking problem that statistical learning theory may also be applicable for solving other optimization problems in finance, especially when the solution of the optimization problem is not stable and the problem of overfitting is acute. The improvement of the solution is due to the direct use of



preliminary information about the solution in the optimization problem. This procedure has a strong theoretical justification in statistical learning theory.

Future research is needed to further identify and empirically verify the informational structures that can be used to solve index tracking problems. Additionally, future research should investigate the possibility of using the STEM methodology to solve general portfolio optimization problems.

## Appendix 1

### Formula (Section 3.4.2)

$$RS^{rand} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \left( \frac{m}{M} \right)^2 \right) = \frac{1}{6} - \frac{1}{6M^2}$$

is valid.

### Proof

Reformulating the expression for  $RS^{rand}$  one obtains

$$RS^{rand} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \left( \frac{m}{M} \right)^2 \right) = \frac{1}{M^2} \sum_{m=1}^M m - \frac{1}{M^3} \sum_{m=1}^M m^2. \quad (6.1)$$

The formula for the sum of an arithmetic progression is as follows:

$$\sum_{m=1}^M m = \frac{M(M+1)}{2}. \quad (6.2)$$

The formula for the sum of the sequence of the squared natural numbers is as follows:

$$\sum_{m=1}^M m^2 = \frac{M(M+1)(2M+1)}{6}. \quad (6.3)$$

Substituting (6.2) and (6.3) into (6.1) gives

$$RS^{rand} = \frac{(M+1)}{2M} - \frac{(M+1)(2M+1)}{6M^2} = \frac{1}{6} - \frac{1}{6M^2}.$$

**QED**

### Formula (Section 3.4.2)

$$RS = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \frac{D_m}{M} \right) = \frac{1}{4}$$

is valid for the first case:

$$D_m = \max[0, 2(m - \frac{M}{2})]$$

and for the second case:

$$D_m = \max[0, 2(m - 1 - \frac{M}{2})].$$

### Proof

For the first case

$$RS^{worst} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \frac{\max[0, 2m - M]}{M} \right) = \frac{1}{M^2} \sum_{m=1}^M m - \frac{1}{M^2} \sum_{m=\frac{M}{2}+1}^M (2m - M) \quad (6.4)$$

The second term can be reformulated as:

$$\begin{aligned} \sum_{m=\frac{M}{2}+1}^M (2m - M) &= 2 \sum_{m=\frac{M}{2}+1}^M m - \frac{M^2}{2} = \\ &= 2 \frac{M}{2} \left( \frac{\frac{M}{2} + 1 + M}{2} \right) - \frac{M^2}{2} = \\ &= \frac{M(\frac{M}{2} + 1 + M)}{2} - \frac{M^2}{2} = \frac{M}{2} + \frac{M^2}{4}. \end{aligned} \quad (6.5)$$

Substituting (6.2) and (6.5) into (6.4) gives

$$RS^{worst} = \frac{(M+1)}{2M} - \frac{1}{2M} - \frac{1}{4} = \frac{1}{4}.$$

For the second case,

$$RS^{worst} = \frac{1}{M} \sum_{m=1}^M \left( \frac{m}{M} - \frac{\max[0, 2m - 2 - M]}{M} \right) = \frac{1}{M^2} \sum_{m=1}^M m - \frac{1}{M^2} \sum_{m=\frac{M}{2}}^M (2m - 2 - M) \quad (6.6)$$

$$\begin{aligned} \sum_{m=\frac{M}{2}}^M (2m - 2 - M) &= 2 \sum_{m=\frac{M}{2}}^M m - \sum_{m=\frac{M}{2}}^M M = \\ &= 2 \left( \frac{M}{2} + 1 \right) \frac{\frac{M}{2} + M}{2} - \left( \frac{M}{2} + 1 \right) M = \frac{M}{2} + \frac{M^2}{4} \end{aligned} \quad (6.7)$$

Substituting (6.2) and (6.7) into (6.6) gives

$$RS^{worst} = \frac{(M+1)}{2M} - \frac{1}{2M} - \frac{1}{4} = \frac{1}{4}.$$

**QED**

## Appendix 2

### Proof of Proposition 5.1 (Section 5.1.6):

#### Inequalities

Subtracting the parameter  $\theta > \varepsilon > 0$  from both sides in the inequalities for the regression problem and using  $\sum_{i=1}^N w_i = 1$  it is possible to write

$$\begin{aligned} \sum_{i=1}^N w_i (y_l - x_{i,l} - \theta) - b &\leq -(\theta - \varepsilon) + \xi_l \\ - \sum_{i=1}^N w_i (y_l - x_{i,l} + \theta) + b &\leq -(\theta - \varepsilon) + \xi_l^*. \end{aligned} \quad (7.1)$$

It is always possible to find weights  $w_i^c$ , such that

$$w_i := \frac{w_i^c}{\sum_{i=1}^N w_i^c}. \quad (7.2)$$

Substituting (7.2) into (7.1) and multiplying both sides by  $\sum_{i=1}^N w_i^c > 0$  gives

$$\begin{aligned} \sum_{i=1}^N w_i^c (y_l - x_{i,l} - \theta) - b \sum_{i=1}^N w_i^c &\leq -(\theta - \varepsilon) \sum_{i=1}^N w_i^c + \xi_l \sum_{i=1}^N w_i^c \\ - \sum_{i=1}^N w_i^c (y_l - x_{i,l} + \theta) + b \sum_{i=1}^N w_i^c &\leq -(\theta - \varepsilon) \sum_{i=1}^N w_i^c + \xi_l^* \sum_{i=1}^N w_i^c. \end{aligned} \quad (7.3)$$

It is always possible to rescale  $w_i^c$  such that

$$(\theta - \varepsilon) \sum_{i=1}^N w_i^c = 1. \quad (7.4)$$

Reformulating this expression one obtains

$$\varepsilon = \theta - \frac{1}{\sum_{i=1}^N w_i^c} \geq 0. \quad (7.5)$$

This constraint should be added to the optimization problem to limit  $\varepsilon$ .

Introducing the following new variables

$$\begin{aligned}
\xi_l^c &= \xi_l \sum_{i=1}^N w_i^c \\
\xi_l^{c*} &= \xi_l^* \sum_{i=1}^N w_i^c \\
b^c &= b \sum_{i=1}^N w_i^c,
\end{aligned} \tag{7.6}$$

and using the rescaling function (7.4) and the variables (7.6), one can rewrite equation (7.3) as

$$\begin{aligned}
&\sum_{i=1}^N w_i^c (y_l - x_{i,l} - \theta) - b^c \leq -1 + \xi_l^c \\
&-\sum_{i=1}^N w_i^c (y_l - x_{i,l} + \theta) + b^c \leq -1 + \xi_l^{c*}.
\end{aligned}$$

The constraint  $\sum_{i=1}^N w_i = 1$  can be omitted for  $w_i^c$

$$\sum_{i=1}^N w_i = \sum_{i=1}^N \frac{w_i^c}{\sum_{i=1}^N w_i^c} = 1.$$

### Norm constraint

The norm constraint in the form

$$\frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 \leq C$$

can be rewritten, using (7.2) as

$$\frac{1}{2} \|\mathbf{S}(\mathbf{w} - \mathbf{w}_0)\|_2^2 = \frac{1}{2} \|\mathbf{S}(\frac{\mathbf{w}^c}{\sum_{i=1}^N w_i^c} - \mathbf{w}_0)\|_2^2 \leq C.$$

Multiplying both sides of the inequality by  $(\sum_{i=1}^N w_i^c)^2 \neq 0$  gives

$$\frac{1}{2} \|\mathbf{S}(\mathbf{w}^c - \mathbf{w}_0 \sum_{i=1}^N w_i^c)\|_2^2 \leq C (\sum_{i=1}^N w_i^c)^2.$$

We now introduce a new constant,

$$C^c = C \left( \sum_{i=1}^N w_i^c \right)^2.$$

Using (7.5), we get

$$C = \frac{C^c}{\left( \sum_{i=1}^N w_i^c \right)^2} = C^c (\theta - \varepsilon)^2 \leq C^c \theta^2.$$

The norm can be rewritten as

$$\frac{1}{2} \left\| \mathbf{S}(\mathbf{w}^c - \mathbf{w}_0) \sum_{i=1}^N w_i^c \right\|_2^2 \leq C^c. \quad (7.7)$$

### Objective

The objective function has the form

$$\sum_{l=1}^L (\xi_l + \xi_l^*) = \frac{\sum_{l=1}^L (\xi_l^c + \xi_l^{c*})}{\sum_{i=1}^N w_i^c}.$$

The denominator is fixed because of (7.7).

Therefore, it is possible to minimize

$$\sum_{l=1}^L (\xi_l^c + \xi_l^{c*}).$$

**QED**

## References

- Adams, M., Maeda, K. and Schoenfeld, S., 2004.** Index-Based Separately Managed Accounts Delivering on the Performance Promise. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley, pp. 479-496.
- Aivazian, S.A., Mhitarian, V.S., 1998.** *Prikladnaja statistika i osnovy econometriki [in Russian]*. Moscow: Uniti.
- Alexander, C., Dimitriu, A., 2002.** The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies. *Discussion Papers in Finance ISMA Center No. 2002-08*.
- Alexander, C., Dimitriu, A., 2003.** Sources of over-performance in equity markets: mean reversion, common trends and herding. *Discussion Papers in Finance ISMA Center No. 2003-08*.
- Alexander, C., Dimitriu, A., 2004a.** A Comparison of Cointegration and Tracking Error Models for Mutual Funds and Hedge Funds. *Discussion Papers in Finance ISMA Centre No 2004-04*.
- Alexander, C., Dimitriu, A., 2004b.** Equity indexing: Optimize your passive investments. *Quantitative Finance*, 4, (3), pp. C30-C33.
- Alighanbari, M., Mougeot, N., 2009.** Benchmark Tracking Strategies: On Optimal Tracking. *Deutsche Bank*.
- Amenc, N. et al., 2011.** Improved Beta A Comparison of Index-Weighting Schemes. *An EDHEC-Risk Institute Publication*.
- Ancona, N., 1999.** Classification properties of support vector machines for regression. *Technical Report RI-IESI/CNR - Nr. 02/99. Consiglio Nazionale delle Ricerche*.
- Arnott, R., Kuo, L., 2011.** Selection Bias How an Index Chooses Stocks Really Does Make a Difference. *Journal of Indexes*, September/October, pp. 36-45.
- Arnott, R.D., Hsu, J., Moore, P., 2005.** Fundamental indexation. *Financial Analysts Journal*, 60, (2), pp. 83-99.
- Bamberg, G., Wagner, N., 2000.** Equity index replication with standard and robust regression estimators. *OR Spectrum*, 22, (4), pp. 525-543.
- Barclays, 2008.** The benchmark in fixed income: Barclays Capital Indices. *Barclays*.
- Beasley, JE, Meade, N., Chang, T.J., 2003.** An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research*, 148, (3), pp. 621-643.
- Ben-Tal, A., Nemirovski, A., 1999.** Robust solutions of uncertain linear programs. *Operations Research Letters*, 25, (1), pp. 1-14.
- Berke, L., 2009.** The optimal implementation ETFs, futures and SWAPS. *TABB Group Report*.
- Bi, J., Bennett, K.P., 2003.** A geometric approach to support vector regression. *Neurocomputing*, 55, (1-2), pp. 79-108.
- Black, F., Litterman, R., 1992.** Global portfolio optimization. *Financial Analysts Journal*, Sep/Oct , pp. 28-43.
- Blackrock, 2011.** ETF Landscape. Industry review. End H1 2011. *Blackrock*.
- Blume, M.E., Edelen, R.M., 2004.** S&P 500 indexers, tracking error, and liquidity. *The Journal of Portfolio Management*, 30, (3), pp. 37-46.
- Bogle, J.C., 2001.** *John Bogle on investing: the first 50 years*. McGraw-Hill.



- Bonafede, J., 2003.** The Wilshire 5000 Total Market Index: The Logistics Behind Managing Broad-Based Indexes. *Journal of Indexes*, 3rd Quarter. Available at: <http://www.indexuniverse.com/publications/journalofindexes/joi-articles/1593.html>. [Accessed 29.02.2011].
- Booth, D.G., 2001.** Index and Enhanced Index Funds. *Dimensional Fund Advisors*. Available at: <http://www.seiler-associates.com/downloads/DFA2/IndexEnhancedIndexFunds.pdf> [Accessed 29.02.2011].
- Bootz, R., 2009.** Konstruktionsmöglichkeiten von ETFs. [PowerPoint Slides], Lyxor, ETFTOUR 2009 Bern, Basel, St. Gallen, Zürich, Available at: [http://www.etftour.ch/files/Downloads/RogerBootz\\_Lyxor.pdf](http://www.etftour.ch/files/Downloads/RogerBootz_Lyxor.pdf) [Accessed 29.02.2011].
- Bos, R., 2000.** Event study: quantifying the effect of being added to an S&P index. *Standard & Poor's Report*.
- Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992.** A training algorithm for optimal margin classifiers. In: D. Haussler (Ed.), *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144-152.
- Bousquet, O., Boucheron, S., Lugosi, G., 2004.** Introduction to statistical learning theory. In: Bousquet, O., U. von Luxburg and G. Ratsch, eds, *Advanced Lectures on Machine Learning*. Lecture Notes in Artificial Intelligence 3176, Springer, Heidelberg, Germany, pp. 169-207.
- Bruce, B.R., Eisenberg, A., 1992.** Global Synthetic Index Funds. *The Journal of Investing*, 1, (2), pp. 45-47.
- Budinsky, R., 2002.** *Aktienindexfonds: Produktidee und Umsetzung*. Bankakademie-Verlag.
- CESR, 2010.** Guidelines on Risk Measurement and the Calculation of Global Exposure and Counterparty Risk for UCITS. .
- Canakgoz, N.A., Beasley, J.E., 2008.** Mixed-integer programming approaches for index tracking and enhanced indexation. *European Journal of Operational Research*, 196, (1), pp. 384-399.
- Cano, G., Feldman, B., Smith, J., 2009.** ETFs, swaps and futures: Trade at Index Close (TIC) and the coevolution of financial markets. *ETFs and Indexing*, FALL 2009, Vol. 2009, No. 1, pp. 50-58.
- Cavaglia, S., Moroz, V., 2002.** Cross-industry, Cross-country allocation. *Financial Analysts Journal*, 58, 6, pp. 78-97.
- Cherkassky, V., Ma, Y., 2004.** Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, 17, (1), pp. 113-126.
- Cherkassky, V., Mulier, F., 2007.** *Learning from data: Concepts, theory, and methods*. Wiley-IEEE Press.
- Chopra, V.K., 1993.** Improving optimization. *The Journal of Investing*, 2, (3), pp. 51-59.
- Coleman, T.F., Li, Y., Henniger, J., 2006.** Minimizing tracking error while restricting the number of assets. *Journal of Risk*, 8, (4), pp. 33.
- ComStage, 2011.** ComStage ETFs: Geniale Ideen sind oft sehr einfach – ComStage ETFs. *Commerzbank*. Available at: [http://www.comstage.de/SiteContent/4/1/2/192/16/broschuere\\_retail.pdf](http://www.comstage.de/SiteContent/4/1/2/192/16/broschuere_retail.pdf) [Accessed 29.02.2011].
- Consiglio, A., Zenios, S.A., 2001.** Integrated simulation and optimization models for tracking international fixed income indices. *Mathematical Programming*, 89, (2), pp. 311-339.
- Corielli, F., Marcellino, M., 2006.** Factor based index tracking. *Journal of Banking & Finance*, 30, (8), pp. 2215-2233.
- Cornuejols, G., Tütüncü, R., 2007.** *Optimization methods in finance*. Cambridge University Press.
- Cortes, C., Vapnik, V., 1995.** Support-vector networks. *Machine learning*, 20, (3), pp. 273-297.

- Costandinides, C. et al., 2010.** Global Equity Index & ETF Research. Do ETF trackers track? European equity ETF tracking error under the microscope. *Deutsche Bank Report*. Available at: [http://www.etf.db.com/de/pdf/en/research/researchspecial\\_2010\\_09\\_06\\_do\\_etf\\_trackers\\_track.pdf](http://www.etf.db.com/de/pdf/en/research/researchspecial_2010_09_06_do_etf_trackers_track.pdf) [Accessed 28.02.2012].
- DJUBS, 2011.** Commodity index handbook. *Dow Jones*. Available at: <http://www.djindexes.com/commodity/> [Accessed 25.02.2012].
- DeMiguel, V et al., 2009.** A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55, (5), pp. 798-812.
- DeMiguel, V., Garlappi, L., Uppal, R., 2009.** Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *Review of Financial Studies*, 22, (5), pp. 1915-1953.
- Derigs, U., Nickel, N.H., 2004.** On a local-search heuristic for a class of tracking error minimization problems in portfolio management. *Annals of Operations Research*, 131, (1), pp. 45-77.
- DeutscheBank, 2011.** Präzise Marktabbildung Ein Leitfaden zu den Replikationsmethoden von ETFs.. *DeutscheBank*. Available at: [http://www.etf.db.com/CH/DE/binaer\\_view.asp?BinaerNr=91](http://www.etf.db.com/CH/DE/binaer_view.asp?BinaerNr=91) [Accessed 29.02.2011].
- DeutscheBörse, 2011.** Guide to the equity indices of Deutsche Börse. *Deutsche Börse*. Available at: [http://dax-indices.com/EN/MediaLibrary/Document/Equity\\_L\\_6\\_16\\_e.pdf](http://dax-indices.com/EN/MediaLibrary/Document/Equity_L_6_16_e.pdf) [Accessed 29.02.2011].
- Dorfleitner, G., 1999.** Eine Anmerkung zur exakten Nachbildung von Aktienindizes mittels einer Multiplikator-Rundungsmethode. *OR Spectrum*, 21, (4), pp. 493-502.
- Dose, C., Cincotti, S., 2005.** Clustering of financial time series with application to index and enhanced index tracking portfolio. *Physica A: Statistical Mechanics and its Applications*, 355, (1), pp. 145-151.
- Draper, N.R., Smith, H., 1998.** *Applied regression analysis*. Wiley-Interscience.
- Dunis, C.L., Ho, R., 2005.** Cointegration portfolios of European equities for index tracking and market neutral strategies. *Journal of Asset Management*, 6, (1), pp. 33-52.
- Efron, B., 2004.** The estimation of prediction error: covariance penalties and cross validation. *Journal of the American Statistical Association*, 99, (467), pp. 619-632.
- Eichhorn, D., Gupta, F., Stubbs, E., 1998.** Using constraints to improve the robustness of asset allocation. *The Journal of Portfolio Management*, 24, (3), pp. 41-48.
- Ellis, C.D., 1975.** The loser's game. *Financial Analysts Journal*, Jul/Aug, 31, (4), pp. 19-26.
- Ezra, D., Warren, G., 2010.** When should investors consider an alternative to passive investing? *Journal of Portfolio Management*, 36, (4), pp. 5-6.
- FSB, 2011.** Potential financial stability issues arising from recent trends in ETFs. *Financial Stability Board*. Available at: [http://www.financialstabilityboard.org/publications/r\\_110412b.pdf](http://www.financialstabilityboard.org/publications/r_110412b.pdf) [Accessed 27.06.2011].
- Fabozzi, F.J. et al., 2007.** *Robust portfolio optimization and management*. Wiley.
- Fama, E.F., French, K.R., 1992.** The cross-section of expected stock returns. *Journal of Finance*, 47, pp. 427-465.
- Fama, E.F., French, K.R., 1993.** Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, (1), pp. 3-56.
- Fastrich, B., Paterlini, S., Winker, P., 2010.** Cardinality versus q-Norm Constraints for Index Tracking. Available at SSRN: <http://ssrn.com/abstract=1679690> or <http://dx.doi.org/10.2139/ssrn.1679690>. [Accessed 27.06.2011]
- Focardi, S.M., Fabozzi, F.J., 2004.** A methodology for index tracking based on time-series clustering. *Quantitative Finance*, 4, (4), pp. 417-425.

- Frost, P. A., and Savarino, J. E., 1988.** For better performance constrain portfolio weights. *Journal of Portfolio Management*, 5, pp. 29–34.
- Fuller, R.J., Han, B., Tung, Y., 2010.** Thinking about Indices and Passive versus Active Management. *The Journal of Portfolio Management*, 36, (4), pp. 35-47.
- Gaivoronski, A.A., Krylov, S., Van Der Wijst, N., 2005.** Optimal portfolio selection and dynamic benchmark tracking. *European Journal of Operational Research*, 163, (1), pp. 115-131.
- Gastineau, G.L., 2002.** Exchange traded funds and their competitors. In: Fabozzi, F. J., ed. 2002. *The Handbook of Financial Instruments*. John Wiley & Sons Ltd, pp. 531-554.
- Gilli, M., Kellezi, E., 2001.** Threshold Accepting for Index Tracking. *Availbale at:* <http://www.unige.ch/ses/dsec/static/gilli/portfolio/Yale-2001-IT.pdf>. [Accessed 27.05.2011]. Published as: Gilli, M., Kellezi, E., 2002. The Threshold Accepting Heuristic for Index Tracking. In: Pardalos, P., Tsitsiringos, V.K., Financial Engineering, E-Commerce, and Supply Chain, Kluwer, Boston, pp. 1-18.
- Glenn, J.A., Littler, G.H., 1984.** *A Dictionary of Mathematics*. Harper & Row.
- Goltz, F., Campani, C.H., 2011.** A Review of Corporate Bond Indices Construction Principles, Return Heterogeneity, and Fluctuations in Risk Exposures. *An EDHEC- Risk Institute Publication*.
- Goltz, F., Le Sourd, V., 2010.** Does finance theory make the case for capitalisation-weighted indexing? *An EDHEC- Risk Institute Publication*, 19.
- Gorton, G., Rouwenhorst, K.G., 2005.** Facts and fantasies about commodity futures. *The Rodney L. White Center for Financial Research*.
- Gourieroux, C., Jasiak, J., 2001.** *Financial econometrics: problems, models and methods*. Princeton.
- Grauer, R.R., Shen, F.C., 2000.** Do constraints improve portfolio performance?. *Journal of Banking & Finance*, 24, (8), pp. 1253-1274.
- Greifeneder, B., 2011.** All Indices are not created equal picking the right index. [PowerPoint Slides] 2nd annual Inside ETFs Europe Conference, May 5-6, 2011, Amsterdam, Netherlands. Available at: [http://www.indexuniverse.eu/docs/presentation2011/7\\_AllIndicesAreNotEqual.pdf](http://www.indexuniverse.eu/docs/presentation2011/7_AllIndicesAreNotEqual.pdf) [Accessed 29.02.2011].
- Grinold, R.C., 1992.** Are Benchmark Portfolios Efficient? *The Journal of Portfolio Management*, 19, (1), pp. 34-40.
- Günther, S., 2002.** Praktische Bedeutung und professioneller Einsatz von Benchmarkportfolios. *Handbuch Portfoliomanagement*, 2, pp. 225-252.
- Hafner, R., Pütz, A., Werner, R., 2003.** Index tracking under transaction costs rebalancing passive portfolios. *The Euromoney, Global Portfolio Trading Handbook*.
- Hamid, N., Sandford, Y., 2007a.** Portfolio Trading and Index Strategy: Expanding the European Market with Innovative ETFs. *DeutscheBank*.
- Hamid, N., Sandford, Y., 2007b.** Portfolio Trading and Index Strategy: ETFs v Futures v Swaps. *DeutscheBank*.
- Hansen, B.E., 2012.** Econometrics. *University of Wisconsin*. Available at: <http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf> [Accessed 23.06.2012].
- Harris, L., 1993.** The winners and losers of the zero sum game with Innovative ETFs Price Efficiency and Market Liquidity. *Institute for Quantitative Research in Finance, Spring 93, Wesley Chapel, Florida*.
- Hastie, T., Tibshirani, R. and Friedman, J., 2009.** *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition. Springer.

- Hill, J., Mueller, B., 2004.** Choosing among Index Vehicles How Does an Institutional Investor Select an Investment Product. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley, pp. 503-535.
- Howaldt, C., 2011.** *Analyse von Index-Tracking-Eigenschaften von Portfolios und Aktien*. Master's thesis. Frankfurt School of Finance & Management.
- Hsu, C.W., Chang, C.C., Lin, C.J., 2009.** A practical guide to support vector classification. *National Taiwan University*.
- Hsu, J., 2006.** Cap weighted portfolios are sub-optimal portfolios. *Journal of Investment Management*, 4, (3), pp. 44-53.
- IMF, 2011.** Global financial stability report, April 2011. *International Monetary Fund*. Available at: <http://www.imf.org/external/pubs/ft/gfsr/2011/01/index.htm> [Accessed 23.06.2012].
- Investment company act, 1940.**
- iShares, 2010.** Overview of ETF Structures How ETFs Track their Benchmark. *iShares*. Available at: <http://www.ishares.com> [Accessed 23.06.2012].
- Jagannathan, R., Ma, T., 2003.** Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58, (4), pp. 1651-1684.
- Jansen, R., Van Dijk, R., 2002.** Optimal benchmark tracking with small portfolios-Using the diversity measure to deal with constraints on number of stocks in a portfolio. *Journal of Portfolio Management*, 28, (2).
- Jensen, M.C., 1968.** The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23, (2), pp. 389-416.
- Jhirad, Y., Ozkul, O. and Qian, D., 2004.** Exchange-Traded Funds A Flexible and Efficient Investment Tool. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley, pp. 297-324.
- Johnson, B., Bioy, H., Rose, G., 2011.** Synthetic ETFs Under the Microscope. *Morningstar*.
- Jones, F.J., Fabozzi, F.J., 2002.** Investment companies. In: Fabozzi, F.J., ed. 2002. *The Handbook of financial instruments*. Wiley, pp. 503-530.
- Jorion, P., 2007.** *Value at Risk The New Benchmark for Managing Financial Risk*. McGraw-Hill.
- Karlow, D., Roßbach, P., 2011.** A Method for Robust Index Tracking, in: Bo Hu, Karl Morasch, Stefan Pickl, Markus Siegle (Eds.): *Operations Research Proceedings 2010*, Heidelberg: Springer, pp. 9-14.
- Kohavi, R., 1995.** A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *International joint Conference on artificial intelligence*, pp. 1137-1145.
- Konno, H., Wijayanayake, A., 2001.** Minimal cost index tracking under nonlinear transaction costs and minimal transaction unit constraints. *International Journal of Theoretical and Applied Finance*, 4, (6), pp. 939-958.
- Kosev, M., Williams, T., 2011.** Exchange traded funds. *RBA Bulletin*, pp. 51-60.
- Kostovetsky, L., 2003.** Index mutual funds and exchange-traded funds. *The Journal of Portfolio Management*, 29, (4), pp. 80-92.
- Kozlowski, R., 2010.** Blackrock takes the lead in indexing. *Pensions&Investments*.
- Krink, T., Mittnik, S., Paterlini, S., 2009.** Differential evolution and combinatorial search for constrained index-tracking. *Annals of Operations Research*, 172, (1), pp. 153-176.
- Kusiak, S., 2010.** Career Portfolios and Exotic Index Replication. [PowerPoint Slides] 23rd Annual Northfield Research Conference.

- Kwiatkowski, J.W., 1992.** Algorithms for Index Tracking. *IMA Journal of Management Mathematics*, 4, (3), pp. 279.
- Larsen Jr, G.A., Resnick, B.G., 1998.** Empirical Insights on Indexing. *The Journal of Portfolio Management*, 25, (1), pp. 51-60.
- Lintner, J., 1965.** The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47, (1), pp. 13-37.
- Lo, A.W., MacKinlay, A.C., 1988.** Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1, (1), pp. 41-66.
- Lobo, M.S. et al, 1998.** Applications of second-order cone programming. *Linear Algebra and Its Applications*, 284, (1-3), pp. 193-228.
- Loftus, J.S., 1997.** Enhanced Equity Indexing. In: Fabozzi, F., ed. 1997. *Professional perspectives on indexing*. Fabozzi Associates, pp. 35-53.
- Maginn, J.L. et al, 2007.** *Managing Investment Portfolios: A Dynamic Process*. Wiley.
- Malkiel, B.G., 1997.** Indexing: a continuing winning strategy. In: Fabozzi, F., ed. 1997. *Professional perspectives on indexing*. Fabozzi Associates, pp. 3-17.
- Malkiel, B.G., 2007.** *A random walk down Wall Street: the time-tested strategy for successful investing*. WW Norton & Co Inc.
- Maringer, D., 2008.** Constrained index tracking under loss aversion using differential evolution. *Natural Computing in Computational Finance*, pp. 7-24.
- Markowitz, H., 1952.** Portfolio Selection. *The Journal of Finance*, 7, (1), pp. 77-91.
- Markowitz, H., Todd, G.P., 1987.** *Mean-variance analysis in portfolio choice and capital markets*. Wiley.
- Mattera, D., Haykin, S., 1999.** Support vector machines for dynamic reconstruction of a chaotic system. In: Schölkopf, B., Burges, C.J.C., and Smola, A.J., Eds., *Advances in kernel methods*, MIT Press, pp. 211-241.
- Meade, N., Salkin, G.R., 1989.** Index funds-construction and performance measurement. *Journal of the Operational Research Society*, 40, pp. 871-879.
- Meade, N., Salkin, G.R., 1990.** Developing and maintaining an equity index fund. *Journal of the Operational Research Society*, 41, (7), pp. 599-607.
- Michalik, T., 2008.** Aktives Management mit passiven Produkten. [PowerPoint Slides] Deutsche Bank.
- Miller Jr, E.M., 1978.** How to Win at the Loser's Game. *The Journal of Portfolio Management*, 5, (1), pp. 17-28.
- Montfort, K., Visser, E., van Draat, L.F., 2008.** Index tracking by means of optimized sampling. *The Journal of Portfolio Management*, 34, (2), pp. 143-152.
- Mossin, J., 1966.** Equilibrium in a capital asset market. *Econometrica*, 34, (4), pp. 768-783.
- Neubert, A. S., 1997.** Knowing your indexes: Background, definitions, and rules. In: Fabozzi, F.J., ed. 1997. *Professional Perspectives on Indexing*. Fabozzi Associates, pp. 54-66.
- Olsen, K., 2011.** Indexed assets surge 25%. *Pensions&Investments*. Available at: <http://www.pionline.com/article/20110919/PRINTSUB/309199986> [Accessed 23.06.2012].
- Pliska, S.R., Suzuki, K., 2004.** Optimal tracking for asset allocation with fixed and proportional transaction costs. *Quantitative Finance*, 4, (2), pp. 233-243.



- Pontil, M., Rifkin, R., Evgeniou, T., 1999.** From regression to classification in support vector machines. *Massachusetts Institute of Technology*. Available at: <ftp://publications.ai.mit.edu/ai-publications/pdf/AIM-1649.pdf> [Accessed 13.08.2012].
- Pope, P.F., Yadav, P.K., 1994.** Discovering errors in tracking error. *The Journal of Portfolio Management*, 20, (2), pp. 27-32.
- Primbs, J.A., Sung, C.H., 2008.** A stochastic receding horizon control approach to constrained index tracking. *Asia-Pacific Financial Markets*, 15, (1), pp. 3-24.
- Quisenberry, C.H., 2009.** Core-Satellite Strategies for the High-Net-Worth Investor. In: Horan, S.M., ed. 2009. *Private Wealth: wealth management in practice*. Wiley, pp. 299-310.
- Rafaely, B., Bennell, J.A., 2006.** Optimisation of FTSE 100 tracker funds: A comparison of genetic algorithms and quadratic programming. *Managerial Finance*, 32, (6), pp. 477-492.
- Ramaswamy, S., 2011.** Market structures and systemic risks of Exchange-Traded Funds. *BIS Working Papers*.
- Rey, D.M., Seiler, D., 2001.** Indexation and tracking errors. *WWZ, Department of Finance, Working Paper*, (2/01).
- Rinne, H., 1995.** *Taschenbuch der Statistik*. Harri Deutsch Verlag.
- Rinne, H., 2003.** *Taschenbuch der Statistik*. Harri Deutsch Verlag.
- Rockafellar, R.T., Uryasev, S., 2002.** Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26, (7), pp. 1443-1471.
- Roll, R., 1977.** A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4, (2), pp. 129-176.
- Roll, R., 1992.** A mean/variance analysis of tracking error. *The Journal of Portfolio Management*, 18, (4), pp. 13-22.
- Ronalds, N., Anderson, C., 2006.** The Synthetic EAFE Index CMEs new EAEF contract may save us all a lot of trouble. *Journal of Indexes*, November/December, pp. 30-35.
- Roßbach, P., Karlow, D., 2011.** The stability of traditional measures of index tracking quality. *Frankfurt School - Working Paper Series 164*, Frankfurt am Main: Frankfurt School of Finance & Management.
- Roßbach, P., Karlow, D., 2012.** Optimizing Tracking Robustness by Structural Minimization Technique. Manuscript submitted for publication.
- Rudd, A., 1980.** Optimal selection of passive portfolios. *Financial Management*, 9, (1), pp. 57-66.
- Rudd, A., 1986.** Portfolio management another look at passive management. *Journal of Accounting, Auditing and Finance*, 7, pp. 242-252.
- Rudolf, M., Wolter, H.J., Zimmermann, H., 1999.** A linear model for tracking error minimization. *Journal of Banking & Finance*, 23, (1), pp. 85-103.
- Ruiz-Torrubiano, R., Suarez, A., 2009.** A hybrid optimization approach to index tracking. *Annals of Operations Research*, 166, (1), pp. 57-71.
- Russell, 2011.** Equity Indices Construction and Methodology. *Russel*. Available at: <http://www.russell.com/indexes/documents/methodology.pdf> [Accessed 13.08.2012].
- S&P, 2008.** Index Mathematics Methodology. *Standard & Poors*. Available at: [https://www.sp-indexdata.com/idpfiles/indexalert/prc/active/whitepapers/Index\\_Math\\_Web.pdf](https://www.sp-indexdata.com/idpfiles/indexalert/prc/active/whitepapers/Index_Math_Web.pdf) [Accessed 13.08.2012].
- S&P, 2009.** S&P US Indices Methodology. *Standard & Poors*. Available at: [https://www.sp-indexdata.com/idpfiles/indexalert/prc/active/whitepapers/SP\\_US\\_Indices\\_Methodology\\_Web.pdf](https://www.sp-indexdata.com/idpfiles/indexalert/prc/active/whitepapers/SP_US_Indices_Methodology_Web.pdf) [Accessed 13.08.2012].

- Samuelson, P.A., 1974.** Challenge to judgment. *Journal of Portfolio Management*, 1, (1), pp. 17-19.
- Schioldager A. et al, , 2004.** The unique challenges of US Equity index management. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley.
- Schoenfeld, S.A., Maeda, , 2004.** Fundamental index portfolio management techniques. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley, pp. 365-388.
- Schoenfeld, S.A., Yang, K., 2004.** Enhanced Indexing Adding Index Alpha in a Disciplined Risk-Controlled Manner. In: Schoenfeld, S., ed. 2004. *Active index investing: Maximizing Portfolio Performance and Minimizing Risk through Global Index Strategies*. Wiley, pp. 277-296.
- Schottle, K., Werner, R., 2006.** Towards reliable efficient frontiers. *Journal of Asset Management*, 7, (2), pp. 128-141.
- Schölkopf, B., Smola, A.J., 2002.** *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press.
- Schölkopf, B. et al., 2000.** New support vector algorithms. *Neural Computation*, 12, (5), pp. 1207-1245.
- Sewell, M., 2008.** VC Dimension. Available at: <http://www.svms.org/vc-dimension/vc-dimension.pdf> [Accessed 29.07.2012].
- Shao, X., Cherkassky, V., Li, W., 2000.** Measuring the VC-dimension using optimized experimental design. *Neural Computation*, 12, (8), pp. 1969-1986.
- Sharpe, W.F., 1964.** Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19, (3), pp. 425-442.
- Sharpe, W.F., 1966.** Mutual fund performance. *The Journal of Business*, 39, (1), pp. 119-138.
- Sharpe, W.F., 1991.** The arithmetic of active management. *Financial Analysts Journal*, 47, (1), pp. 7-9.
- Sharpe, W.F., 1992.** Asset allocation. *The Journal of Portfolio Management*, 18, (2), pp. 7-19.
- Shawe-Taylor, J., Cristianini, N., 1999.** Further results on the margin distribution. In: *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, pp. 278-285.
- Siegel, J.J., 2008.** *Stocks for the long run: The definitive guide to financial market returns and long-term investment strategies*. McGraw-Hill Companies.
- Smeets, H.D., Wiesner, A., 2009.** Exchange Traded Funds: Simply buying the index? *Die Bank*, 11/2009, pp. 22-27.
- Smola, A.J., , 1998.** *Learning with Kernels*. Ph. D. Thesis. Technischen Universität Berlin.
- Sobehart, J.R., Keenan, S.C., Stein, R.M., 2000.** Benchmarking quantitative default risk models: a validation methodology. *Moodys Investors Service*.
- Sorensen, E.H., Miller, K.L., Ooi, C.K., 2000.** The decision tree approach to stock selection. *The Journal of Portfolio Management*, 27, (1), pp. 42-52.
- Sorensen, E.H., Miller, K.L., Samak, V., 1998.** Allocating between active and passive management. *Financial Analysts Journal*, 54, (5), pp. 18-31.
- Stoxx, 2011a.** Stoxx equity index calculation guide. *Stoxx*. Available at: [http://www.stoxx.com/download/indices/rulebooks/stoxx\\_calculation\\_guide.pdf](http://www.stoxx.com/download/indices/rulebooks/stoxx_calculation_guide.pdf) [Accessed 29.07.2012].

- Stoxx, 2011b.** Withholding Taxes. *Stoxx*. Available at: <http://www.stoxx.com/indices/taxes.html> [Accessed 29.07.2012].
- Stoyan, S.J., Kwon, R.H., 2009.** A two-stage stochastic mixed-integer programming approach to the index tracking problem. *Optimization and Engineering*, 11, (2), pp. 247-275.
- Strongin S. et al., 2002.** A Stockpickers Reality Part III Global Portfolio Analysis. *Goldman Sachs*.
- Strongin, S., Petsch, M., Sharenow, G., 2000.** Beating benchmarks. *The Journal of Portfolio Management*, 26, (4), pp. 11-27.
- Takeda, A., Gotoh, J., Sugiyama, M., 2011.** Support Vector Regression as Conditional Value at Risk Minimization with Application to Financial Time Series. In: *2010 IEEE International Workshop on Machine Learning for Signal Processing (August 2010)*.
- Tikhonov, A.N., Arsenin, V.J., 1979.** *Metody reshenija nekorrektnykh zadach [In Russian]*. Nauka.
- Tobin, J., 1958.** Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25, (2), pp. 65-86.
- Treynor, J., Mazuy, K., 1966.** Can mutual funds outguess the market. *Harvard Business Review*, 44, (4), pp. 131-136.
- Tucker, M., Laipply, S., 2010.** Understanding bond etf premiums and discounts. *Journal of Indexes*, September/October, pp. 40-47.
- UCITS, 2009.** UCITS DIRECTIVE 2009/65/EU. .
- Uryasev, S., 2000.** Conditional value-at-risk: Optimization algorithms and applications. In: *Computational Intelligence for Financial Engineering, 2000.(CIFER) Proceedings of the IEEE/IAFE/INFORMS 2000 Conference*, pp. 49-57.
- Vapnik, V., Chapelle, O., 2000.** Bounds on error expectation for support vector machines. *Neural Computation*, 12, (9), pp. 2013-2036.
- Vapnik, V., Levin, E., Cun, Y.L., 1994.** Measuring the VC-dimension of a learning machine. *Neural Computation*, 6, (5), pp. 851-876.
- Vapnik, V.N., 1995.** *The nature of statistical learning theory*. Springer Verlag.
- Vapnik, V.N., 1998.** *Statistical Learning Theory*. Wiley-Interscience.
- Vapnik, V.N., 1999.** An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10, (5), p.988-999.
- Vapnik, V.N., 2000.** *The nature of statistical learning theory*. Second Edition. Springer Verlag.
- Vapnik, V.N., 2006.** *Estimation of dependences based on empirical data*. Springer-Verlag New York Inc.
- Vapnik, V.N., Chervonenkis, A. Ja., 1989.** The necessary and sufficient conditions for consistency of the method of empirical risk minimization, *Yearbook of the Academy of Sciences of the USSR on Recognition, Classification, and Forecasting*, vol. 2, pp. 217–249, Nauka Moscow (in Russian). English translation: *Pattern Recogn. and Image Analysis*, vol. 1, no. 3, pp. 284–305, 1991.
- Wagner, N., 2002.** Passives Management: Methoden zum Tracking von Marktindizes. In: Kleeberg, J.M., ed. 2002. *Handbuch Portfolio Management*. Bad Soden, pp. 813-839.
- Waring, B., Attwood, F., 1998.** Futures versus Physicals Implementing International Portfolio. *Journal of Investing*, Winter, 8, (4), pp. 17-30.
- Wu, L.C. et al., 2007.** Enhanced index investing based on goal programming. *The Journal of Portfolio Management*, 33, (3), pp. 49-56.



**Ye, J., 1998.** On measuring and correcting the effects of data mining and model selection. *Journal of the American Statistical Association*, 93, (441), pp. 120-131.

**Zenious, , 2005.** Optimization Models for Structuring Index Funds. In: Wallace, S.W., Ziemba, W.T., eds. 2005. *Application of stochastic programming*. MPS-SIAM, pp. 471-502.

**Zhang, J., Maringer, D., 2010.** Index Mutual Fund Replication. In: Brabazon, A., O'Neil, M., Maringer, D., eds., *Natural Computing in Computational Finance*, Springer, pp. 109-130.

**Zhu, L., Coleman, T.F., Li, Y., 2009.** Min-max robust and CVaR robust mean-variance portfolios. *Journal of Risk*, 11, (3), pp. 55-85.

## **About the Author**

Denis Karlow received in 2001 a diploma in Applied Mathematics from Belarusian State University, Minsk, Belarus. Before joining Frankfurt School of Finance and Management he spent nine years in variety of positions in software development industry. Since 2012 he works as a software architect in the one asset management company. His research interests are business analytics and optimization.

### **Eidesstattliche Versicherung**

Ich versichere hiermit an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Für alle Inhalte die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Quellen gleich welcher Art entnommen sind, habe ich die fremde Urheberschaft kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form weder im Inland noch im Ausland als Prüfungsarbeit eingereicht worden.

### **Declaration of authorship**

I hereby certify that unless otherwise indicated in the text or references, or acknowledged, this thesis is entirely the product of my own scholarly work. Appropriate credit is given where I have used language, ideas, expressions, or writings of another from public or non-public sources. This thesis has not been submitted, either in whole or part, for a degree at this or any other university or institution.

Frankfurt am Main, 10.09.2012