

7-21-2008

A Comprehensive Portfolio Construction Under Stochastic Environment

Ahmed Elshahat

Florida International University, aelshahat@gmail.com

DOI: 10.25148/etd.FI10022509

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

Recommended Citation

Elshahat, Ahmed, "A Comprehensive Portfolio Construction Under Stochastic Environment" (2008). *FIU Electronic Theses and Dissertations*. 187.

<https://digitalcommons.fiu.edu/etd/187>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

A COMPREHENSIVE PORTFOLIO CONSTRUCTION UNDER
STOCHASTIC ENVIRONMENT

A dissertation submitted in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

BUSINESS ADMINISTRATION

by

Ahmed Elshahat

2008

© Copyright 1998 by Ahmed Elshahat

All rights reserved

DEDICATION

I dedicate this dissertation to my parents, my sisters, my brother, and my wife. Without their patience, understanding, and support, the completion of this work would not have been possible.

ACKNOWLEDGMENTS

I wish to thank the members of my committee for their support and patience. Their gentle but firm direction has been most appreciated. My major professor, Professor Ali M. Parhizgari, was particularly helpful in guiding me toward a qualitative and quantitative methodology. Professor Giri Narasimhan was a huge support in the programming and the coding of the dissertation methodology. I would like to thank Dr. Shadab Anwar for his help in programming, debugging, and detailed discussions about the application stages, Dr. Qasem Kharma for his valuable contribution in shaping the algorithm, and Mr. Ali Bakhtiar for his programming contribution. I would also like to thank Drs. Prakash and Chang and all the committee members for all the support that they gave me and for all they have taught me. I would like also to thank Dr. Welch for his continuous support.

ABSTRACT OF THE DISSERTATION
A COMPREHENSIVE PORTFOLIO CONSTRUCTION
UNDER STOCHASTIC ENVIRONMENT

by

Ahmed Elshahat

Florida International University, 2008

Miami, Florida

Professor Ali M. Parhizgari, Major Professor

Prior research has established that idiosyncratic volatility of the securities prices exhibits a positive trend. This trend and other factors have made the merits of investment diversification and portfolio construction more compelling.

A new optimization technique, a greedy algorithm, is proposed to optimize the weights of assets in a portfolio. The main benefits of using this algorithm are to: a) increase the efficiency of the portfolio optimization process, b) implement large-scale optimizations, and c) improve the resulting optimal weights. In addition, the technique utilizes a novel approach in the construction of a time-varying covariance matrix. This involves the application of a modified integrated dynamic conditional correlation GARCH (IDCC - GARCH) model to account for the dynamics of the conditional covariance matrices that are employed.

The stochastic aspects of the expected return of the securities are integrated into the technique through Monte Carlo simulations. Instead of representing the expected returns as deterministic values, they are assigned simulated values based on their historical measures. The time-series of the securities are fitted into a probability

distribution that matches the time-series characteristics using the Anderson-Darling goodness-of-fit criterion. Simulated and actual data sets are used to further generalize the results. Employing the S&P500 securities as the base, 2000 simulated data sets are created using Monte Carlo simulation. In addition, the Russell 1000 securities are used to generate 50 sample data sets.

The results indicate an increase in risk-return performance. Choosing the Value-at-Risk (VaR) as the criterion and the Crystal Ball portfolio optimizer, a commercial product currently available on the market, as the comparison for benchmarking, the new greedy technique clearly outperforms others using a sample of the S&P500 and the Russell 1000 securities. The resulting improvements in performance are consistent among five securities selection methods (maximum, minimum, random, absolute minimum, and absolute maximum) and three covariance structures (unconditional, orthogonal GARCH, and integrated dynamic conditional GARCH).

TABLE OF CONTENTS

| CHAPTER | PAGE |
|--|------|
| 1. Introduction..... | 1 |
| 2. Portfolio Construction Framework | 6 |
| 2.1. Portfolio Objective..... | 7 |
| 2.2. Securities Screening and Selection..... | 9 |
| 2.2.1. Trend Strategies. | 12 |
| 2.2.2. Strategies using exogenous factors | 13 |
| 2.2.3. Strategies using Econometric models..... | 13 |
| 2.3. Common Portfolio Constraints..... | 14 |
| 2.3.1. Linear and quadratic Constraints..... | 14 |
| 2.3.1.1. Budget constraint..... | 15 |
| 2.3.1.2. Long-only constraint..... | 15 |
| 2.3.1.3. Turnover constraint..... | 16 |
| 2.3.1.4. Bound constraint..... | 16 |
| 2.3.1.5. Risk constraint..... | 17 |
| 2.3.1.6. Benchmark exposure and tracking error constraints.... | 17 |
| 2.3.2. Combinatorial Constraints..... | 19 |
| 2.3.2.1. Minimum holding and transaction size constraints..... | 19 |
| 2.3.2.2. Cardinality constraint..... | 20 |
| 2.3.2.3. Round lot constraint..... | 20 |
| 2.4. Portfolio Optimization..... | 22 |
| 2.4.1. Why the algorithm used matters? | 24 |
| 2.4.2. The search techniques mechanism..... | 26 |
| 2.4.3. Optimality conditions..... | 27 |
| 2.5. Portfolio Revision..... | 28 |
| 2.6. Portfolio Performance Evaluation..... | 29 |
| 3. Estimating portfolio parameters: Theoretical and empirical background | 31 |
| 3.1. Financial data irregularity..... | 32 |
| 3.1.1. Parameters transition..... | 33 |
| 3.1.2. Fat tails. | 34 |
| 3.2. Data frequency..... | 35 |
| 3.3. Estimating Return..... | 37 |
| 3.3.1. Theoretical models..... | 37 |
| 3.3.2. Econometric models..... | 40 |
| 3.4. Estimating Volatility..... | 41 |
| 3.4.1. Dispersion volatility models..... | 42 |
| 3.4.2. Downside volatility models..... | 42 |
| 3.4.2.1. Traditional models..... | 43 |
| 3.4.2.2. VaR, CVaR, and CaViaR..... | 43 |

| | | |
|----------|--|-----|
| 3.4.2.3. | Implied measurement..... | 45 |
| 3.4.2.4. | Markov Chains and volatility clustering..... | 46 |
| 3.4.2.5. | ARCH/GARCH models..... | 46 |
| 3.5. | Estimating Co-volatility..... | 48 |
| 3.5.1. | Rolling window and Exponential smoothing..... | 51 |
| 3.5.2. | ARMA and ARIMA models..... | 51 |
| 3.5.3. | Stein or Shrinkage Covariance matrix..... | 52 |
| 3.5.4. | Covariance matrix discounting..... | 53 |
| 3.5.5. | Implied co-volatility..... | 53 |
| 3.5.6. | Incorporating outliers..... | 54 |
| 3.5.7. | Decomposition of co-volatility..... | 55 |
| 4. | Empirical Results..... | 56 |
| 4.1. | Data..... | 57 |
| 4.2. | Methodology..... | 60 |
| 4.2.1. | Estimating volatility and co-volatility..... | 61 |
| 4.2.1.1. | Traditional Covariance matrix..... | 61 |
| 4.2.1.2. | Orthogonal GARCH model..... | 64 |
| 4.2.1.3. | Constant Conditional Correlated GARCH model..... | 65 |
| 4.2.1.4. | Integrated Dynamic Conditionally Correlation GARCH | 66 |
| 4.2.2. | Expected Return and Stochastic Programming..... | 67 |
| 4.2.3. | Greedy Algorithm application..... | 68 |
| 4.2.3.1. | Candidate set..... | 70 |
| 4.2.3.2. | Selection function..... | 71 |
| 4.2.3.3. | Feasibility function..... | 73 |
| 4.2.3.4. | Objective function..... | 73 |
| 4.2.3.5. | Solution function..... | 74 |
| 4.3. | Results..... | 75 |
| 5. | Conclusions..... | 86 |
| 5.1. | Summary of the results..... | 86 |
| 5.2. | Limitations..... | 88 |
| 5.3. | Future work..... | 89 |
| 5.4. | Concluding remarks..... | 90 |
| | REFERENCES..... | 91 |
| | VITA..... | 103 |

LIST OF TABLES

| TABLE | PAGE |
|--|------|
| 1. Average Sharpe Ratio using return relative..... | 76 |
| 2. Average Sharpe Ratios using residual return relative..... | 77 |
| 3. Descriptive Statistics for the VaR values using traditional covariance matrix.. | 78 |
| 4. T-tests for Max-Selection Criteria Superiority over other Selection Methods.. | 82 |
| 5. Z-tests for Max-Selection Criteria Superiority over other Selection Methods.. | 82 |
| 6. OGARCH Superiority over the other two Estimation Techniques..... | 85 |

LIST OF FIGURES

| FIGURE | PAGE |
|---|------|
| 1. The Portfolio Construction Process..... | 7 |
| 2. Time-Varying Volatility using 60-day Rolling Window..... | 34 |
| 3. Time-varying correlation Coefficient between the US and EU..... | 34 |
| 4. US Time-Varying Volatility with different frequencies..... | 36 |
| 5. S&P500 Price movement over the period 1988 – 2008..... | 58 |
| 6. The Standard Error Value as the number of Simulations increase..... | 60 |
| 7. The Greedy Algorithm five stages..... | 70 |
| 8. VaR for 50 portfolios created using the traditional covariance matrix..... | 79 |
| 9. VaR for 50 portfolios created using the IDCC-GARCH covariance matrix... | 80 |
| 10. VaR for 50 portfolios created using the OGARCH covariance matrix..... | 80 |
| 11. VaR Normal and Empirical Cumulative Distributions..... | 81 |
| 12. VaR time-series of the Proposed Model vs. Crystal Ball..... | 83 |
| 13. Out-of-Sample VaR time-series of the Proposed Model vs. Crystal Ball..... | 84 |

Chapter 1

Introduction

Prior research has established that idiosyncratic volatility of the securities prices exhibits a positive trend (see, for instance, Campbell et al. (2001), Morck et al. (2000), PASTOR and Pietro (2003), and Wei and Zhang (2006)). This trend and other factors have made the merits of investment diversification and portfolio construction more compelling now than in the past. Furthermore, today's portfolio construction models are more computationally complex, and they include more parameters and constraints than Markowitz's classical Mean-Variance (MV) optimization model.

In addition to the complexity added by including constraints, the estimation process of the parameters of the models are more computationally complex than the models used in the past. For example, instead of using the unconditional covariance matrix, today's models use multivariate variants of conditional variances and covariances, like the multivariate GARCH. With all these complexities, it does not make much sense to use models that cannot optimize a portfolio of securities, especially a large one, in a reasonable time.

This dissertation proposes the use of a greedy algorithm in portfolio optimization. The main benefits of using this algorithm are to: a) increase the efficiency of the portfolio optimization process, b) implement large-scale optimizations, and c) improve the resulting optimal weights. The greedy algorithm is a very efficient algorithm; it does not reconsider any previous selections but moves to newer iterations. The efficiency of an

algorithm is a significant factor especially with today's computationally complex mathematical models.

As with any model, the quality of the output depends on the quality of the input, other things being equal. The main criticism of the MV optimization was its sensitivity to the inputs used. The main input to a portfolio optimization model is the variance-covariance matrix. In the proposed model, much attention is paid to the input estimation, so that the model is fed with reliable inputs. A conditional covariance matrix is used instead of the traditional unconditional covariance matrix to account for the non-stationary in time-series of the returns. Four different conditional estimation techniques are used to estimate the covariance matrix. These are the traditional unconditional covariance matrix, the Constant Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity (GARCH) covariance, the Integrated Dynamic Conditional Correlation GARCH covariance matrix, and the Orthogonal GARCH covariance matrix. The best results were reached using an adaptive variant of the Orthogonal GARCH.

In this dissertation the stochastic nature of the expected returns is integrated using stochastic programming and Monte Carlo simulation. Stochastic programming is a class of methods that incorporate the stochastic nature of variables into the traditional mathematical programming framework (see, Ruszczyrski and Shapiro (2006)). Instead of representing the expected returns as deterministic values, they are assigned scenarios generated in advance, based on their historical values. The historical time-series used in

our model are fitted into a probability distribution that matches the time-series characteristics using the Anderson-Darling goodness-of-fit criterion.

After the time-series are represented by a probability distribution, the Monte Carlo simulation uses the probability distribution to generate different expected returns. For example, if a time-series is fitted to a normal distribution, the mean of the distribution is the most likely value, so it has a better chance of being generated as an input, but that does not prevent other values – with lower probability – from appearing in the simulations.

To make the results more generalized, simulated data sets are used in addition to the original data set. Using the initial data set as a base, 2000 simulated data sets are created using the Monte Carlo simulation. The methodology used to create the simulated data sets is similar to the methodology used in the portfolio resampling techniques (Scherer (2002)).

The remaining parts of this dissertation are organized as follows. The second chapter provides a review of portfolio construction framework. The framework covers six stages that an investor goes through to construct an optimal portfolio. The first stage is determining the portfolio's objective. The second stage determines a procedure to screen out securities that do not meet the requirements of the investors. The third stage discusses in detail the different types of portfolio constraints that are used in practice. The fourth stage focuses on the portfolio optimization and the optimization algorithms. The fifth and the sixth stages discuss the portfolio revisions and performance evaluation issues.

The third chapter covers some aspects of the theoretical and empirical techniques to estimate inputs of the portfolio model. The chapter starts by discussing the financial data irregularities that need to be considered when estimating the parameters, mainly, data non-stationarity and the return probability distribution's fat tails. The different return estimation models along with their pros and cons are then discussed. The chapter concludes by a discussion about the volatility and the covalatility estimation models.

The fourth chapter discusses the empirical results. The chapter starts by presenting the data set used, and the simulations used to create more variants of the data set. The methodology used to estimate the model's inputs is discussed next, followed by a discussion of the application of the greedy algorithm in the portfolio construction. The chapter ends by presenting the results reached. Five different selection techniques are proposed to select the securities to be included in the optimal portfolio. All the five selection techniques are used in the same greedy algorithm framework. Consistent with some aspects of the modern portfolio theory, the technique that maximizes the weights of securities with the least correlation outperformed the other selection techniques. The results of this selection technique are compared to three different benchmarks: S&P500 index, Russell 1000 index, and Crystal Ball. Crystal Ball is a commercial software used to optimize the weights of securities portfolios.

The results of the proposed model, using the traditional unconditional covariance matrix, significantly outperform the three benchmarks used. Furthermore, the use of the conditional covariance matrices improves the results further, especially with the Orthogonal GARCH covariance matrix.

The fifth chapter provides a summary of the results reached, the limitations of the results, and the potential for future work.

Chapter 2

Portfolio Construction Framework

Harry Markowitz's seminal works on portfolio selection (Markowitz (1952) and Markowitz (1987)) have created a shift in the investment field's focus from individual security selection to portfolio construction. The focus of most investment managers is now on how an individual security will contribute to the portfolio's risk and return, rather than on estimating the expected risk and return profile of that specific security. This chapter provides a review of portfolio construction. The portfolio construction process can be summarized in the six steps shown in Figure (1).

The first and the most important step is setting the objective of the portfolio, as it sets the stage for everything to follow. The second step is to select the securities to be included in the portfolio. The third step is to set the constraints to be considered in the model. The fourth step is to select the correct algorithm in order to optimize values. The fifth step is to set the points for revising the portfolio. The final step is to evaluate the portfolio to see whether it has achieved the stated objectives.

The second step should be distinguished from the fourth step (portfolio optimization). Step two involves building expectations about the return, volatility, and co-volatility between securities, before applying a selection criterion. This selection is meant to narrow down the securities pool to a manageable number. The discretion of the investor can be integrated at this stage. Compared with step two, step four is an automated selection, based on a pre-specified procedure.

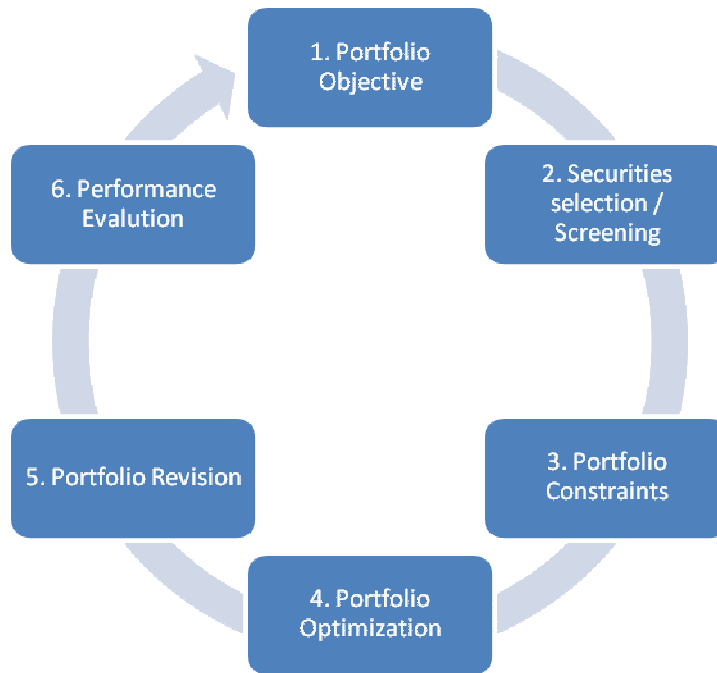


Figure 1: The Portfolio Construction Process

2.1. Portfolio Objective

The statement of the portfolio objective is the most important step in creating a portfolio as it forms the basis for everything else in the process (Barksdale and Green (1990); Fogler and Russell (1978); Ramaswami et al. (1992)). Particularly in today's financial markets, where every step is governed by sophisticated rules and regulations, the portfolio objective is heavily emphasized. The area of securities fraud and mismanagement emphasize the importance of clearly stating the portfolio's objective and its implications to the investor (see NASD Manual). For example, the retirement plans in the United States, as stated in Section 402 (b)(1) of the Employee Retirement Income Security Act of 1974 (ERISA), requires that:

“Every employee benefit plan shall provide a procedure for establishing and carrying out a funding policy and method consistent with the objectives of the plan.”

Setting the objective of the portfolio may appear to be an easy task, but a number of difficulties exist. These difficulties include: semantics, indecision, subjectivity, and multiple beneficiaries (Strong (1999)). With respect to semantics, not every investor has the same understanding of the terminologies used in the field. Common terms, like growth, income, and liquidity, may mean different things to different people. Clarifying the semantics still leaves some investors undecided, particularly due to the subjective nature of investing. The same security might be perceived as a growth security to one investor and as an income security to another.

Although the portfolio objective terminology may differ from one investor to another, still the objective can be categorized in one of four categories; preserving the principal, income, growth of income, or capital appreciation (Strong (1999)). The most conservative objective is the capital preserving. As the name indicates, this objective is ensuring that the principal is not jeopardized. Theoretically, this objective generates the lowest return. The income objective focuses on generating a smooth flow of income and, unlike principal preservation, there is no prohibition against decrease in the principal.

The growth of income objective attempts to generate an income level that will maintain the purchasing power of the investor. Generally, this is achieved by providing an increasing level of income over time. The level of increase is generally dictated by the level of inflation and risk. Sacrificing current return for higher future return is a common

practice under this objective, and therefore there is no prohibition against decrease in the principal.

The fourth objective, capital appreciation, is a long term objective. It does not require any income generation, but rather seeks to obtain capital appreciation. This objective involves different tiers based on the investment horizon of the investor, and his/her risk tolerance. An investor needs to determine how long he is willing to sacrifice liquidity (to determine the investment horizon), and how much he/she is willing to lose (to determine the risk tolerance).

2.2. Securities Screening and Selection

Security screening is a critical step in portfolio construction that is heavily emphasized in practice (Antia and Fridson (2008); Gold and Lebowitz (1999)), but not addressed seriously in the literature. The reason might be the trend towards portfolio management rather than securities analysis. A number of papers in the literature provide some evidence that selecting random stocks is as successful as selecting favorite stocks (Evans and Archer (1968)). Securities screening is a procedure aimed at reducing the securities population to a manageable size. An investor must choose one security out of the thousands listed in the NYSE, Amex, and the NASDAQ, and the many more thousands listed overseas.

The main advantage of securities screening is saving time and avoiding data overload. For example, if an investor spends only five minutes analyzing a specific stock, he/she will need forty hours to analyzing five hundred stocks, and even after such a long period of time, the investor will most likely only be more confused. Securities screening

also ensures that the securities population meets the investor's minimum requirements. Feeding an optimizer with parameters for thousands of securities might lead to many problems. First is the scalability problem, where the optimizer might not be able to handle the whole data set. Second is the efficiency problem, where the optimizer might take long time to provide optimal solutions. Third, the optimizer might select stocks that do not meet the investor's objective.

A "security screen" determines which securities to be considered and which to be screened out. A major criterion in selecting a sufficient screen is its relevance to the investor's objective. In practice, a multistage screening process is always used, as one cannot find a screen that will cover the entire relevant requirement. For example, the use of profitability ratios is not very meaningful in isolation. It is necessary to integrate it with other ratios to obtain a larger and more relevant picture. Some screens utilize objective criterion such as financial ratios, industry averages, and index levels, while others use subjective screens implied by the investors. Some screens reflect the past, like financial ratios, while others utilize forward-looking criterion, such as expected growth rates and expected P/E ratios.

Socially responsible investing (SRI) is one of the well-known subjective screens used as a positive and creative means for allocation of resources (Bakshi (2007)). It is defined by the social investment forum as follows (Social Investment Forum: Socially Responsible Investing Basics for Individuals):

"SRI is a broad-based approach to investing that now encompasses an estimated \$2.71 trillion out of \$25.1 trillion in the U.S. investment marketplace today. SRI recognizes that corporate responsibility and societal concerns are valid parts of investment decisions. SRI

considers both the investor's financial needs and an investment's impact on society. SRI investors encourage corporations to improve their practices on environmental, social, and governance issues. You may also hear SRI-like approaches to investing referred to as mission investing, responsible investing, double or triple bottom line investing, ethical investing, sustainable investing, or green investing."

A large number of researchers have investigated the SRI area and found positive results. Some researchers found that using the SRI screens leads to an annual 8.7% abnormal return (Kempf and Osthoff (2007)). Others studied the effect of the SRI in the portfolio context to see whether the usage of SRI screens would affect the level of diversification. Evidence supports the claim that portfolios using the SRI screens do not differ significantly from those that do not use the SRI (Bello (2005)).

Academic researchers frequently use screeners such as the S&P Compustat tapes, where it is possible to set conditional statements using a list of variables, comparison operators, and desired values to assemble a population of securities that meet the pre-specified criteria. For practitioners, Bloomberg is the strongest platform in the field that allows the screening of securities from around the world. Brokerage houses such as Fidelity, Schwab, and Merrill Lynch provide screening tools to their clients. Furthermore, Yahoo (Stock Screener - Yahoo! Finance) and MSN (MSN Money - Stock Screener: Custom Stock Search), provide free powerful securities screeners that enable investors to create their own queries using several hundred statistical characteristics.

Different investors use different strategies to select securities. Some of these strategies are quantitative in nature, others are subjective. The quantitative strategies are based on models that utilize numerical data. Most of the fundamental analysis fall into

this category. The qualitative models are subjective in nature, and they include most of the technical analysis. Regardless of whether the strategy used is quantitative or qualitative, an effective strategy must be consistent, applicable in a large scale context, and programmable (Fabozzi et al. (2007)).

The securities selection strategies can be categorized into the following three categories: trend strategies, strategies that use exogenous predictors, and strategies that use econometric models.

2.2.1. Trend Strategies

Trend strategies are probably the most widely-used securities selection strategy. They include such strategies as the momentum strategy, and reversal (contrarian) strategy. The momentum strategy capitalizes on the persistence of trends over periods between 3 to 18 months (Chan et al. (1996a); Jegadeesh (1990); Parhizgari and Duong, (2008)). Researchers show that weekly and monthly stock returns tend to have negative autocorrelation. Thus, the securities with the best performance during the previous week are likely to perform weakly during the following week, and the securities with the worst performance in the previous week tend to have the best performance during the following week.

Jegadeesh and Lehmann show that a strategy that buys the winner and sells the losers will consistently generate almost 30% annual return before transaction cost (Jegadeesh (1990); Lehmann (1990)). More and more research has accumulated evidence in favor of momentum strategy profits (Avramov et al. (2007); Miffre and Rallis (2007); Scowcroft and Sefton (2005)). Reversal strategy attempts to identify a

change of direction in a trend. The reversal strategy tends to occur either over a very short period (less than three months) or on longer horizons (three to five years). Since the initial evidence of DeBondt and Thaler (DeBondt and Thaler (1985)), further findings have been reached to support the reversal strategy (Avramov et al. (2006); Platt (2006)).

Although most academic studies agree that these strategies exist, the trend driver is not agreed upon (Jegadeesh and Titman (2001); Scowcroft and Sefton (2005)). Furthermore, regardless of the trend driver, these strategies lead to frequent trading, which leads to transaction costs. Korajczyk and Sadka (2004) show that these strategies are still profitable even after considering the transaction cost.

2.2.2. Strategies using exogenous factors

The second category of securities selection models includes regression models of return (or excess return) on exogenous explanatory variable. These exogenous variables include accounting variables such as the price per earnings (P/E), price per dividends (P/D), price per book value (P/B), as well as non-accounting variables such as the market capitalization. Many studies have provided evidence for these relationships, including Gordon, who demonstrated the negative relationship between stock returns and the firms' size (Gordon (1962a)), (Basu (1977); Campbell and Shiller (2001)).

2.2.3. Strategies using Econometric models

Strategies using more sophisticated econometric and time series models are getting more acceptances for forecasting purposes (as will be detailed in the following two chapters). These models include autoregressive (AR) models (Campbell et al.

(1998a)), dynamic factor approach (Stock and Watson (2005)), and cointegration techniques (Fabozzi, Kolm and Pachamanova (2007)).

2.3. Common Portfolio Constraints

Undoubtedly Markowitz's seminal work on portfolio selection has had a major impact on the field, and the classical mean-variance (MV) framework can still serve as a starting point. In the MV optimization framework, the inclusion of a security in a portfolio is merely a function of its interaction with other assets in the portfolio, without considering other factors such as the investor's objective, transaction cost, and investment regulations. The MV optimization is often extended in several directions. In the previous section, a general coverage of securities screening and selection was provided. As stated previously, this step is used to start with a generally accepted pool of securities, and to ensure that the number of securities in this pool is manageable. In practice, Markowitz's MV optimization is often amended with various types of constraints to satisfy specific objectives. In this section, the common constraints used in practice are explored. These constraints can be grouped into two categories: linear (or quadratic) constraints, and combinatorial constraints.

2.3.1. Linear and quadratic Constraints

These are the commonly used constraints and can be easily handled by typical optimizers for solving the MV problem. These constraints include (Fabozzi, Kolm and Pachamanova (2007)): budget constraint, long-only constraints, turnover constraints, holding constraints, risk factor constraints, and benchmark exposure and tracking error constraints. The following notations are used:

W_0 Vector of portfolio's initial weights.

| | |
|---------------|---|
| W_P | Vector of optimal Weights |
| \tilde{w}_i | The amount to be traded ($W_P - W_0$) |
| x_i | An individual security |
| N | The total number of available securities |
| i | " i " as a subscript refers to a security " i " |
| p | " p " as a subscript refers to a portfolio |

2.3.1.1. Budget Constraint

The budget constraint is a linear equality constraint on the optimization. It restricts the portfolio weights sum to one. It is also known as full investment constraint. It is stated as follows:

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

2.3.1.2. Long-only constraint

The long only (no-short-selling) constraint is a linear constraint that sets the sign of all the weights to be non-negative. It reflects the avoidance of unlimited liability investment that institutional investors are required to follow:

$$w_i \geq 0$$

The budget and the long-only constraints are standard constraints used in many optimizations. However, advances in trading technology has made short-selling strategies more economically viable (Michaud (1998)). Furthermore, research has found that the inclusion of the short-selling constraint leads to a sub-optimal portfolio (Gómez and Sharma (2006)).

2.3.1.3. Turnover constraint

Frequently revising the portfolio leads to high transaction costs. Thus, it is common to have a turnover constraint in a portfolio optimization setting (Schreiner (1980)). The constraint can be set for individual assets, using the following relation (where U is the upper bound vector):

$$|t_i| \leq U_i$$

or it can be set for the whole portfolio, using the following relation:

$$\sum_{i=1}^N |t_i| \leq U_p$$

Alternatively one can use some combination of both constraints. A number of academic studies have provided evidence against the benefit of the turnover constraint, stating that including this constraint simply limit the investor's ability to exploit beneficial trades (Clarke, de Silva, and Thorley (2001)). Often the upper bound (U_i) of the turnover constraint is determined based on the average daily volume of security x_i .

2.3.1.4. Bound Constraint

Another linear constraint commonly used in the portfolio setting is the bound constraint, also known as the holding constraint. The objective of this constraint is to set a bound on the weight of specific security. Without the bound constraint, the weight of a security can take any value between 0% (which could be negative if there is no short sales constraint) and 100% (which could be more than 100% if there is no budget

constraint). The bound constraint can be set as follows (where L is the lower bound vector):

$$L_i \leq W_i \leq U_i$$

Without the bound constraint, the ultimate portfolio might not be well-diversified. The main objective of the bound constraint is to avoid large concentrations in any specific asset's class, industry, sector, or country. The bound constraint can also be used to restrict a specific industry or asset class for the portfolio. This can be set as follows:

$$L_i \leq \sum_{j \in N_i} w_j \leq U_i$$

2.3.1.5. Risk constraint

If the optimization objective is to maximize the portfolio's return, then there must be a risk constraint. Without the risk constraint the optimizer will pick up only the stock with the highest return to maximize the objective. To achieve the portfolio diversification effect, the risk factor must be properly included in the settings. Risk can be the optimization objective, part of the optimization objective, or part of the constraints. The risk constraint depends on the risk measurement used. Risk measurement is discussed in detail in a later chapter.

2.3.1.6. Benchmark exposure and tracking error constraint

Some investors – typically passive investors or index fund managers – prefer to compare their portfolio to a benchmark. Their objective is to consistently perform better than the benchmark. Even a slight over-performance as compared to the benchmark is generally acceptable. For this case, a benchmark exposure needs to be included in the setting. The easiest way is to restrict the securities weight in the portfolio to be in the

neighborhood of the weight of the securities' weights in the benchmark. Let w_b be the benchmark weights (market capitalization weights), then a common benchmark constraint could be set as follows to restrict the portfolio weights from deviating from the benchmark weight (where D is a subjective deviation level):

$$|w_p - w_b| \leq D$$

Another way to achieve the same objective is by using the tracking error of the variance (TEV) constraint (El-Hassan and Kofman (2003)). The tracking error is the variance of the difference between the portfolio's return (R_p) and the benchmark return (R_b). Note that $R_p = R \cdot w_p$, and $R_b = R \cdot w_b$. The TEV can be calculated as:

$$TEV_p = (w_p - w_b)' \cdot \Sigma \cdot (w_p - w_b)$$

where Σ is the covariance matrix of the securities returns. The TEV_p is then restricted to a subjective value determined by the investor. Note that a portfolio with the tracking error constraint as the only constraint overlooks the portfolio risk. This can lead to inefficient MV portfolios (Jorion (2003)).

The constraints discussed so far can be easily integrated into the MV quadratic programming framework. Different algorithms that can integrate these constraints are available. However, some of these algorithms only search for local optimum value (approximate MV optimizers). It is worth mentioning that the use of these constraints depends mainly on the investor's objective. An investor does not have to use all of them; it is just a matter of personal objective. In the following section combinatorial and integer constraints are presented.

2.3.2. Combinatorial Constraints

These are integer constraints that are combinatorial in nature. They are more difficult to handle in comparison to the linear and quadratic constraints mentioned above. Typically they are not preprogrammed in the commonly used MV optimizers. These constraints include (Fabozzi, Kolm and Pachamanova (2007)), minimum holding and transaction-size constraints, cardinality constraints, and round lot constraints.

2.3.2.1. Minimum holding and transaction-size constraints

One of the practical weaknesses of the MV optimization is the weights concentration. Often the MV optimization results in a few large weights and many small weights. These small weights lead to relatively high transaction costs and they contribute little to the portfolio diversification, which makes them not profitable to hold in the portfolio. To eliminate the small weights an investor can include a minimum holding constraint or a transaction size constraint. These two constraints can either be included in the portfolio optimization setting or can be used after the optimization to adjust the resulting weights. The minimum holding constraint can be set as (L_{w_i} is the minimum holding size for asset i);

$$|w_i| \geq L_{w_i} \delta_i \quad \text{where} \quad \delta_i = \begin{cases} 1, & \text{if } w_i \neq 0 \\ 0, & \text{if } w_i = 0 \end{cases}$$

The transaction size constraint can be similarly set as (where t is the amount to be traded, calculated as the difference between w_p and w_0):

$$|t_i| \geq L_{t_i} \delta_i$$

2.3.2.2. Cardinality constraint

Even after using the minimum holding constraint, and the transaction size constraint, the optimizer can result in a large number of positions or large number of trades. To ensure that this problem is overcome, an investor can use a constraint commonly used simultaneously with the minimum holding and/or transaction size constraint, known as cardinality constraint. The cardinality constraint restricts the number of securities to be included in the optimal portfolio. The cardinality constraint is commonly used by investors attempting to track a benchmark using a small number of securities. The cardinality constraint can be set as:

$$\sum_{i=1}^N \delta_i = k$$

where K is a positive integer smaller than I (investment pool).

2.3.2.3. Round lot constraint

The MV optimization assumes a perfect fractionability of the securities. Furthermore, it does not consider the round lot transaction cost savings. To consider the round lot transaction cost savings and to relax the assumption of the perfect fractionability of the market, the round lot constraint is included in the optimization model. Furthermore, if stock options are used in the portfolio setting, then using round lots becomes a necessity, because options can only be written in round lots. The easiest way to integrate the round lot constraint to the optimization problem is to integrate it into the portfolio's weights (w_i). The portfolio weights can be represented as follows (Fabozzi, Kolm and Pachamanova (2007)):

$$w_i = z_i \cdot f_i$$

$$f_i = \frac{P_0 \cdot T}{\Psi}$$

where z_i is an integer number representing the number of round lots to be considered, f_i is a fraction of the portfolio's wealth, P_0 is the initial security's price, T is the number of securities in a round lot, and the Ψ is the total portfolio wealth. Note that w_i is not just a percentage anymore – it has become a dependent variable. The inclusion of the round lot constraint makes a lot of sense, as it has a direct effect on transaction cost reduction. However, occasionally it produces a small increase in risk for a pre-specified level of return (Chiam et al. (2008)). Often the resulting optimal portfolio using the round lot constraint is different than the rounded standard MV optimal portfolio rounded to the nearest round lot (Chang et al. (2000)).

The resulting efficiency frontier from the unconstrained optimization is continuous and relatively smooth, which means that the optimization process is relatively straightforward. However, the inclusion of these constraints complicates the optimization process. The inclusion of the binary and integer variable complicates the MV optimization as well as the original quadratic program. With the combinatorial constraints, the new formulation becomes a quadratic mixed integer program which requires more sophisticated and specialized algorithms that often require significant computing time.

2.4. Portfolio Optimization

The optimization concept is fundamental to finance theory. Optimization, in a literal sense, refers to finding the optimal value for a problem given a number of constraints. In a portfolio context, portfolio optimization is a computational procedure that searches for securities optimal weights. Thus, portfolio optimization is a selection process. If the optimal weight includes a zero percent for a certain security, this security is screened out. As compared to the second section of this chapter, portfolio optimization can be perceived as a designed and/or automated selection process. Markowitz introduced the classical framework for mean-variance (MV) optimization. For an up-to-date revision of the MV optimizations see Steinbach (2001).

The mathematical formulation of the MV optimization is as follows. Let;

N = Number of securities available

W = Vector of portfolio weights on the N securities

R = Vector of expected returns of the N securities

Σ = Covariance matrix of the N securities' return

1 = Vector of ones of length N

The portfolio's mean is calculated as the weighted average of the individual securities' return, and the portfolio's risk is calculated based on the co-movements between the securities included.

$$R_p = w'_p * R$$

$$\sigma_p^2 = w'_p * \Sigma * w_p$$

In the MV setting, the objective function is either to minimize the portfolio's risk for a given level of (minimum required) return (R^*), or to maximize the portfolio's return for a given level of risk (maximum accepted risk). For the former objective (risk minimization), the optimization setting is as follows;

$$\text{Minimize: } \sigma_p^2$$

$$\text{Subject to the constraint: } R_p = R^*$$

As discussed in the previous section, more and more constraints are added to the classical MV optimization to satisfy different objectives. One of the useful alternative settings for the MV optimization is the parametric quadratic programming (Stone (1973)). In the parametric quadratic programming a parameter is added to the objective function to reflect the return. Thus, the objective function includes both risk and return. Other parametric quadratic methods include the Markowitz's critical-line algorithm (Beale (1955); Markowitz (1952); Markowitz (1987)), and the simplex algorithm (Beale (1959); Wolfe (1959)). For a more recent review of the algorithms used for solving convex optimization problems see (Boyd and Vandenberghe (2004)).

Optimization is the process of attempting to find the optimal solution to a problem that may have different possible solutions. Most of these problems involve many variables that interact based on a set of predetermined relations and constraints. The portfolio optimization techniques start with an initial portfolio with initial weights, and continue to improve them. The optimization algorithm lies at the core of the process, which is how improvements are made.

This section will answer the question of why an algorithm matters in the optimization process. Next, the algorithms commonly available will be introduced, along with the optimality conditions, and the section ends by pinpointing the major optimality pitfalls to be avoided.

2.4.1. Why does the algorithm matter?

An algorithm is defined as a sequence of computational procedures that transforms inputs to outputs (Cormen (2001)). Two main properties are used to assess the quality of any algorithm. These two properties are correctness and efficiency. An algorithm is said to be correct if for every input the correct output is produced. That does not mean that incorrect algorithms are useless. Incorrect algorithms could be useful if their error rate can be controlled. In the algorithms literature, efficiency refers to speed of processing, in other words, how long it will take for the algorithm to produce the results. These two properties are investigated in the field of “Theory of Computation and Complexity”.

Theory of computation helps to formalize the question whether a problem can be solved on a computational model using an algorithm. If a problem can be solved, the theory then addresses the efficiency. This field is thus divided into two branches – computability and complexity. Computability theory, as the name indicates, addresses the question of whether a problem is computationally solvable. Complexity theory studies how efficiently a problem can be solved and it addresses the algorithm’s scalability. In other words, the complexity theory monitors the relationship between the input size to the memory requirements and running time of the algorithm. This is also

known as the scalability problem, which can be described as the ability of the system or process to handle growing quantities of data or to be readily enlarged.

In the literature of the complexity theory, algorithms are categorized into complexity classes. Two major classes that are relevant to the portfolio optimization discussion is the NP and P complexity classes. The class NP is the class of problems that can be solved using non-deterministic polynomial-time algorithms. This class is a set of decision problems whose solutions can be verified by a deterministic Turing machine¹ in polynomial time (Turing (1936)). The traveling salesman problem is an example of a problem from the NP class. The class P is the class of problems that can be solved using deterministic polynomial-time algorithms. The problems in class P are efficiently solvable. The class NP includes a subclass of problems called NP-complete. The solutions to problems of this subclass suffer from scalability. Many problems from the area of operations research fall under this category; see (Cockshott and Michaelson (2007)). Another subclass is called NP-hard, which consists of problem that are at least as hard as the hardest problems in NP. Examples of this subclass include portfolio optimization problems.

Algorithms devised to solve the same problem often differ dramatically in their efficiency (speed). Algorithms, like computer hardware, are a technology. Particularly in today's financial markets, characterized by rapid information dissemination and increasingly large offerings of securities, the speed of an algorithm becomes critical.

¹ Turing machines are basic abstract symbol-manipulating devices that can adapt to the simulated logic of any computer.

Even with fast processors and assuming an abundance of memory space, an algorithm significantly affects the speed of processing (Cormen (2001)). It may seem surprising that many portfolio optimizers are not exact solution algorithms, but are instead approximate optimizers. Approximate optimizers have limitations; they may find suboptimal solutions or even infeasible solutions. Approximate optimizers are used to handle NP-complete problems. A significant number of algorithms used today for portfolio optimization attempt only local optima (Cornuejols and Tutuncu (2006)).

2.4.2. The search techniques mechanism

This area of optimization is highly technical and providing a full theoretical coverage of these optimizations is beyond the scope of this dissertation. However, a basic understanding of how an algorithm works is necessary to use an optimizer. For a detailed technical coverage of these algorithms, see (Nocedal and Wright (1999)).

Most optimization algorithms are of an iterative nature, where the algorithm generates a number of solutions that gets closer and closer to the optimal value. Since the optimal value is not known in advance, there must be a termination or convergence criteria that determines when the algorithm should stop searching. Among the common algorithms used in the field for linear problems are the simplex methods and the interior-point methods. For the unconstrained nonlinear models, the Newton-type algorithm is the most common. For the constrained nonlinear models, modern interior point methods and sequential quadratic programming are common. For the more sophisticated model that contains combinatorial and integer constraints, the following algorithms are common: branch and bound, cutting planes algorithms, and special purpose heuristics.

2.4.3. Optimality conditions

The first order optimality condition for a smooth function is that the derivative of the function is equal to zero. This condition is easy to comprehend in two-dimensional space. Let \mathbf{x}^* be a local minimum of the optimization function, and $f'(x) > 0$ for all $x > x^*$. The unconstrained optimization problem can be set as:

$$\min f(x)$$

where f is an N-dimensional function. The necessary optimality condition is given the following gradient condition:

$$\nabla f(\mathbf{x}^*) = \left(\frac{\partial}{\partial x_1} f(\mathbf{x}^*), \dots, \frac{\partial}{\partial x_N} f(\mathbf{x}^*) \right) = 0$$

If an equality constraint exists, the same optimality condition can be used, by converting the constraint problem into an unconstrained problem. This can be achieved by including the equality constraint in the objective function using the Lagrangian multiplier method. The only difference will be partially deriving the objective function once with respect to the \mathbf{x} and once with respect to the λ (Lagrangian multiplier). If an inequality constraint exists, then the optimality condition is given by Karush-Kuhn-Tucker (Karush (1939); Kjeldsen (2000); Kuhn and Tucker (1951)). The Karush-Kuhn-Tucker (KKT) is a necessary optimality condition for nonlinear problems (given the regularity conditions). The KKT is a generalization of the Lagrangian multiplier method.

2.5. Portfolio Revision

Portfolio revision is an integral part of managing a portfolio. Although there is a large body of academic research that supports passive investment, portfolio revision remains a critical factor. The statement of investment policy and investment objective often requires portfolio revision to maintain a certain level of income, risk, or return. Portfolio revision (rebalancing) includes different strategies. Some strategies attempt to maintain the same percentages of the asset classes involved, such as the constant mix strategy (Wilkins, Heck and Cochran (2006)).

Although the constant mix strategy attempts to maintain the status quo, it sells appreciating assets and buys depreciating assets. Whether the market is increasing or decreasing, the constant mix strategy literally does that. Some portfolio managers are forced to use this strategy due to investment policy restrictions. Others hope to make profits from trend reversals. However, given that portfolios generally get revised once every three to twelve months (a period known for momentum trends), it is not often the case that profits are made from trend reversals (Grinblatt et al. (1995); Jegadeesh and Titman (2001)).

Another frequently used portfolio revision strategy is the Constant Proportion Portfolio Insurance (CPPI) (Kingston (1989)). The CPPI only maintains a certain level (floor value) invested in a low-risk asset class (often fixed income security) that is perceived as the portfolio insurance. Any excess value over the floor can be invested elsewhere using a multiplier that determines the aggressiveness of the investor. This strategy gives the portfolio manager more freedom to buy more of the appreciating asset classes and sell the depreciating classes. The downside of this strategy is its tendency to

concentrate on specific asset classes which may affect the overall portfolio's risk. Other portfolio revision strategies focus on other parameters such as the portfolio's beta, its relation to an index, the portfolio's standard deviation, etc.

Revising a portfolio comes at a cost. Each time a revision is made, trading fees are incurred, commissions are paid, and probable transfer taxes are imposed (only in some states). Indirect consequences of portfolio revision also include increasing the noise in the market place. Institutional investors revising their holding to coincide with the investment objective, these revisions in turn increase price volatility and create noise in the market place. Another indirect cost of portfolio revision is the management time and potential income tax implication (Feldstein and Slemrod (1980)).

2.6. Portfolio Performance Evaluation

Portfolio evaluation is one of the critical stages in portfolio management that is often not addressed intensively. The objective of this stage is to see whether a portfolio achieved its objective. To properly evaluate the portfolio performance, one needs a reliable measure for both return and risk. Although fund managers are actually appraised on the realized return, with little risk consideration (Strong (1988)), from an academic point of view, risk should be an integral part of the performance appraisal.

The most commonly used performance measurement is the Sharpe ratio (Sharpe (1966); Sharpe (1998)), the excess return of the portfolio divided by the portfolio standard deviation as a risk-adjusted return measurement for the portfolio performance. Treynor used a more generic measurement that applies to both individual securities as well as to portfolios by using the beta in the denominator (Treynor (1965)). Jensen proposed another measure for portfolio performance (Jensen (1967)). Jensen stated that

running a regression of the excess market return on a portfolio's excess return should lead to a zero intercept (alpha). A positive intercept reflects outperforming the market. Many researchers have documented weaknesses in the Jensen measurement.

In practice, many fund managers are still being assessed based on the portfolio's return figures as compared to a certain benchmark. In case of frequent cash deposits and withdrawal, time-weighted or money-weighted rates of return are appropriate performance appraisals (Fama (1972)). The Value-at-Risk (VaR) approach is gaining more acceptance as a performance evaluation variable (Alexander and Baptista (2003)). The VaR and its conditional forms are discussed in details in a later chapter. Many researchers have proposed improvements to the Sharpe ratio including the Sharpe ratio with expert betas (Bilbao et al. (2007)), Dowd's adjusted Sharpe ratio (Dowd (2000)), and others (Israelsen (2005); Lo (2002); Nielsen and Vassalou (2008)).

Chapter 3

Estimating Portfolio Parameters:

Theoretical and Empirical Background

The investment value of a portfolio depends on the estimated parameters, before it depends on the optimization process. A perfect optimization process will lead to a sub-optimal portfolio when it optimizes inaccurate parameters. Traditionally, investors use historical data to calculate statistical parameters needed to calculate the MV portfolio. The investor's objective is to maximize the expected return of his or her portfolio given a certain level of risk. The expected return and the estimated risk are measured by the mean and standard deviation of the historical return available, respectively. In many cases, this approach leads to sub-optimal portfolios. Using the mean and the standard deviation as measures for return and risk is proven to lead to unstable and highly sensitive portfolios (Jobson and Korkie (1980); Jobson and Korkie (1981a); Jobson and Korkie (1981b)).

No single estimation technique can be perceived as the best. Different estimation techniques attempt to establish a balance among different dimensions. For instance, a balance must be maintained between the need to capture an accurate time-varying volatility and the imprecision that results from using only recent data. For this trade-off, the investment horizon plays a critical role. The longer the time horizon, the stronger is the need to use a time-varying volatility technique. Another trade-off has to do with the data frequency. The investor may want to extract as much information as possible by using higher frequencies, but this might contaminate the resulting estimates, due to mean-

reverting noise over short periods, different spacing between observation, stale prices, and other high frequency data irregularities.

This chapter presents the different estimation techniques along with their extensions and improvements. Before we start the estimation techniques discussion, an overview of the financial data irregularities and the choice of the sampling frequencies is illustrated.

3.1. Financial data irregularity

Relevant to any estimation process is the distribution assumption. Financial models tend to assume normality as the probability distribution that fits the data, but the financial data are proven not to be always normally distributed. Academic research shows evidence that the financial data tend to be skewed, with excess kurtosis (fat tails), which makes the higher moments non-negligible. Thus, relying on the mean and variance to fully describe the securities' return distribution is not enough. One of the reasons that the financial data is not normally distributed is the regime shifting, which is typically modeled by Markov switching techniques.

This regime shifting makes the financial data transitional in nature, where the volatility and correlation of return tend to change over time and asset returns tend to be fat-tailed. These observed facts negate the normality assumption. Financial data time-series, especially data with higher frequencies, are notoriously auto-correlated and heteroscedastic. Thus, an estimation model must account for these irregularities as well.

3.1.1. Parameters Transition

It is now widely accepted that volatility varies over time. Different models have been used to model the time-varying volatility. Among these models are simple filters like the rolling window standard deviation used by Officer (1973), the exponential smoothing, the more sophisticated models like the univariate ARCH by Engle (1982), and several generalizations to the multivariate settings. Survey papers covering the vast literature in this area exist (see Bollerslev et al. (1992); Campbell *et al.* (1998b); Ghysels et al. (1996); and Hentschel (1995)).

As an example to show volatility transition, the daily standard deviation of the Morgan Stanley Capital International (MSCI) indices for the United States and for the European Union's aggregate stock market are depicted in Figure 2. The excess return over the risk free rate of return is used for both time-series. For the MSCI-EU, the LIBOR is used as the risk-free return, and, for the MSCI-US, the return on three-month Treasury bill is used. We depicted the standard deviation from January 1989 through September 2007. Over the observed period, the EU volatility changes between 0.48-2.9 percent, and the US volatility changes between 0.55-3.6 percent.

Similar to the volatility, correlation is also time-varying. Research has documented the sensitivity of the financial models to changes in the correlation coefficients (Ingersoll (1987); Rebonato (1999)). Thus, it is critical to consider the returns-correlation transition. Many attempts to model the time-varying correlations exist (Alexander (2001); Bollerslev, Engle and Wooldridge (1988); Christodoulakis and Satchell (2002); Christodoulakis (2007); Engle and Manganelli (2004)). Figure 3 compares two correlation estimates of the co-movements between the MSCI-US, and

MSCI-EU. During the 20-year period, the correlation coefficient ranged from -26.8% to $+67.7\%$.

Figure 2: Time-Varying Volatility using 60-day Rolling Window

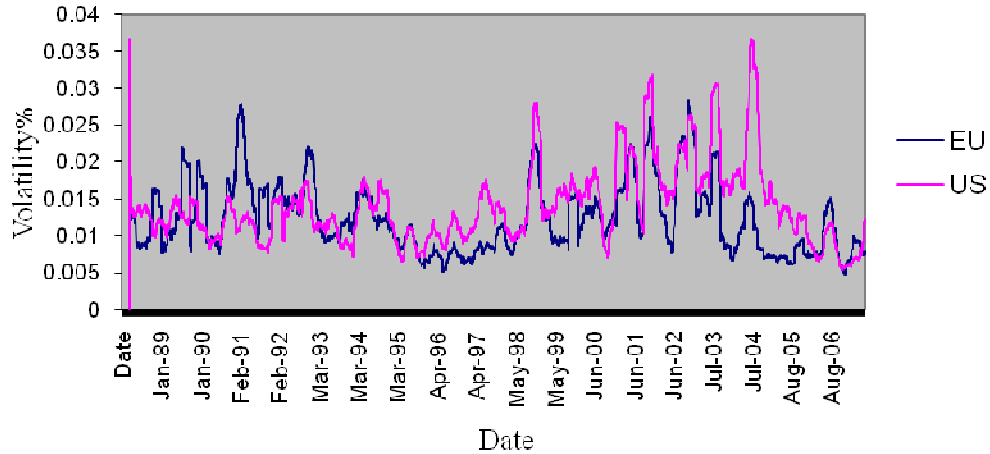
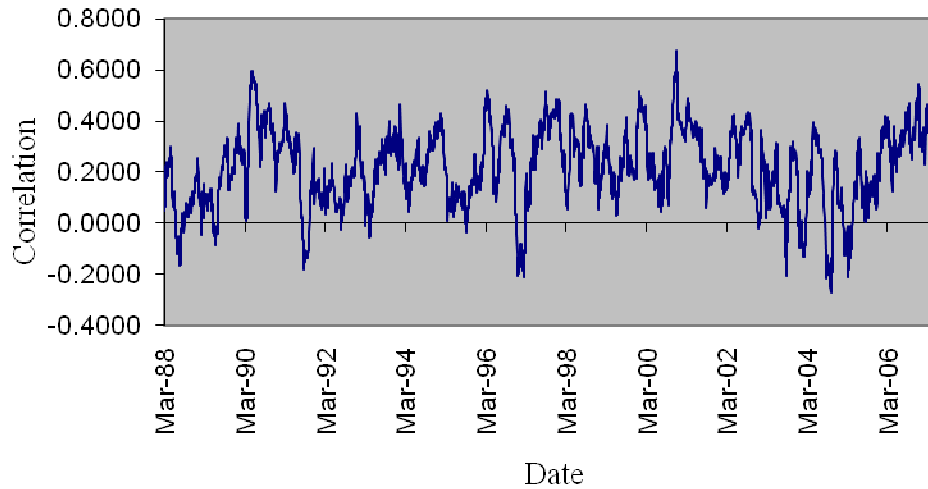


Figure 3: Time-varying correlation Coefficient between the US and EU



3.1.2. Fat tails

Research has documented that financial data possess excess kurtosis, or what is known as the fat tails, see (Kang and Yoon (2007); Kirchler and Huber (2007)). Normal distribution has a kurtosis of three (excess kurtosis over zero), which means that normal distribution cannot fit the financial data, as it understates the probability of extreme

values (tails). This means that assuming normality would understate the true risk measurement. The fact that financial data has fat tails does not come as a surprise. Normal distributions assume independence between error terms; however, this independence does not exist in the financial data.

It is clear that the assumption about the data distribution is critical to the risk measurement and the portfolio construction. Ignoring this fact will affect the accuracy of the estimated parameters. Bidarkota and Dupoyet (2007) report that, by taking the fat tails into consideration, the implied mean risk-free rate is 20% lower, the equity premium is 80% higher, and the term premium is 20% higher, as compared to a Gaussian process. Their results make the model implications closer to the empirically found observations. They also document an increase in the volatility of both risk-free rate and of the equity premium, as implied by their model. Doganoglu et al. (2007) propose a portfolio selection approach using multivariate stable distribution that considers the financial data fat tails as well as the conditionally varying volatility.

3.2. Data frequency

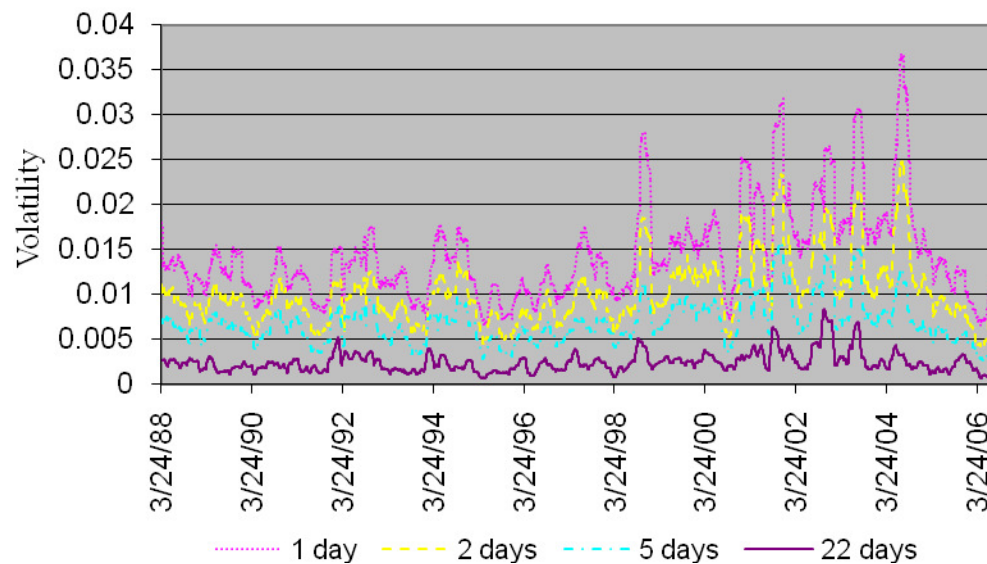
In addition to the data distribution, the data frequency is not of any less importance. Data provide information, so more data frequency, within the same time interval, results in more information. Merton shows that a long history of data is needed to estimate an accurate expected return (Merton (1980)). However, as the data frequency increases, the data quality decreases, where the data becomes contaminated by more noise. Price changes may reflect bid/ask spread, non-synchronous trading due to different time zones, and unequally spaced observation. Furthermore, high frequency data, if not

handled carefully, can capture effects not accounted for, like volatility due to economic data release, time of the day, or trading hours of another exchange.

There are many data-clustering algorithms that model the data and fit them to a distribution. However, their performance depends to a large extent on the model's parameters. Saltenis (2006) proposed a data-clustering algorithm that does not rely on the model's parameters, where the parameters are taken as a natural adaptation of distances between observations.

The optimal data frequency needed for estimating volatility and co-volatility is different from the optimal data frequency for estimating expected return. Volatility and co-volatility estimates can be improved by increasing the sampling frequency. On the other hand, increasing the sampling frequency increases the noise level in estimating the expected return (Garman and Klass (1980); Parkinson (1980)).

Figure 4: US Time-Varying Volatility with different frequencies



In this dissertation, we use monthly data, which are expected to reflect information about the primary return and volatility trends. Monthly data are also less contaminated by the factors mentioned in this section. For example, unlike intra-daily and daily data, the Monthly data are equally spaced, not affected by holidays, and stale prices. To show how different data frequency behaves differently, Figure 4 uses daily, two days, weekly (five days), and monthly (22 days) data frequencies over the period of 1988-2006. It is intuitive that lower frequency averages the trend.

3.3. Estimating return

Since the creation of stock markets, forecasting a security's return has drawn most of the attention of researchers. Many models have been proposed, but still no one model has proved to be completely satisfactory. Some of the models focus on the evolution of returns (time-series analysis), while others focus on the relationship between the behavior of a securities' return at a given point in time (cross-sectional analysis). As in the case of any forecasting model, return estimation models are not exact ones. Return estimation models are only stochastic models that are expected to have forecasting power under certain assumptions.

The return estimation models can be categorized into two groups: theoretical and empirical (econometric). In the following section, both models are introduced.

3.3.1. Theoretical models

These models are forms of the economic general equilibrium theory (GET). The GET is based on a number of assumptions. These assumptions include the assumption of supply and demand equilibrium and that agents maximize their utility function (rational

agents). The most famous theoretical model is the Capital Asset Pricing Model (CAPM) which was developed by Tobin, Treynor, Sharpe, and Lintner independently (Lintner (1965); Sharpe (1964); Tobin (1958); Treynor (1961)). The CAPM is the simplest form of the GET. The CAPM states that, under the pre-defined assumptions, the expected return of a security is a positive linear function of its beta, where beta reflects a security's systematic risk. The higher the beta, the higher the expected return.

During the 1980s and 1990s, evidence was gathered against the CAPM. A number of academic researchers did not find the beta to be a good measurement of the systematic risk and a number of anomalies were accumulating (Fama and French (1992b); Fama and French (1993)). In the absence of other reliable models (Fama and French (1996); Fama and French (1996)), improvements were made to the original CAPM. These improvements include the Conditional CAPM (C-CAPM), in which the CAPM is conditioned on information set at a lagged period (Jagannathan and Wang (1996)). The main problem with the conditional CAPM is the difficulty of identifying the information set (Hansson and Hordahl (1998); Lewellen and Nagel (2006)). Other modifications to the original model include: the zero-beta CAPM (Shanken (1985)), the Consumption-Oriented CAPM (Breedon, Gibbons and Litzenberger (1989)). For a detailed discussion of the CAPM and its different variants see Fama and French (2004).

Stephen Ross proposed an alternative to the CAPM, based on arbitrage arguments, called the arbitrage pricing theory (APT). The APT stated that, unlike the CAPM, a security's expected return is affected by different factors, not just the market risk (as suggested by the CAPM). The APT postulates a positive linear relationship between a security's expected return and some factors (Roll and Ross (1980); Ross

(1976)). The APT proved to have some disadvantages. The main disadvantage is that the factors are not defined by the model (Dhrymes, Friend and Gultekin (1984); Shanken (1982)).

Other techniques from the traditional fundamental analysis that can be used for estimation purposes include Gordon's Dividend Discount Models (DDM) and its improved variant, the Residual Income Model (RIM). The DDM calculates the expected price of a security by discounting its future cash flows (Gordon (1962b)). The DDM was found to be very sensitive to errors in the inputs, which is highly based on assumptions about the security's growth. An improved variant of the DDM is the RIM, which was proved to be much less sensitive to errors in the input than the DDM. As the name indicates, the RIM uses residual income (abnormal return) rather than the absolute return. The RIM states that the value of a stock is equal to its book value per share, plus the present value of expected future residual income per share (Claus and Thomas (2001); Philips (2003)).

Sharpe and Fisher, in two independent papers, introduced another model to accurately estimate the expected return, without using the historical prices; see, (Fisher (1975); Sharpe (1974)). The model was referred to as Reverse Optimization or Implied Expected Returns. Their model computed the expected return on a security as the product of the market price of risk, the covariance matrix, and the market capitalization weights. The Black-Litterman portfolio model, and its predecessor, the Treynor-Black model, used this model to estimate the expected return (see Black and Litterman (1991); Treynor and Black (1973)).

3.3.2. Econometric models

These models are based on an econometric hypothesis, which has the status of an economic theory. The most famous econometric models used for estimating securities return are the Random Walk (RW) models, and the multifactor models. The simple RW model is based on the market efficiency assumption, where the securities return can be viewed as fluctuations around their mean return (Kac (1947)). It is also assumed that the returns are independently and identically distributed (iid).

Different RW models were established by making incremental improvements. The models include: the arithmetic RW, the geometric RW, and the multivariate RW. The Arithmetic RW (ARW) standardizes the positive and negative deviation from the mean by using a Bernoulli variable. The ARW differentiates between two different noise terms (deviation). Unlike the RW model, the ARW states that there is no mean reverse process. The Geometric RW (GRW) corrects the drawbacks of the ARW where it avoids negative prices by using an absorbing barrier, and it is more consistent with empirical results (Guerre and Jouneau (1998)). Another advantage of the GRW is that it does not assume linearity and it approximates a lognormal model. Both the ARW and the GRW are univariate models, which make them too simple to be applicable empirically, as they implicitly assume that the return time-series are independent. The multivariate RW (MRW), as the name indicates, improves the ARW and the GRW by using multivariate time-series (Harvey, Ruiz and Shephard (1994a)).

Factor models are models where the expected return is estimated using different factors. The APT was one of the first models to provide theoretical support for this type

of models. Factor models can be categorized into the following three categories (Connor (1995)): statistical factor models, macroeconomic factor models, and fundamental factor models. The statistical factor model uses historical and cross-sectional data on a security's return to explain the return's behavior using factors that are linear return combinations and uncorrelated. These types of models use the Principle Component Analysis (PCA) to determine the factors. The problem with the statistical factor model is that the factors do not have any economic meaning.

The Macroeconomic Factor Model (MFM) uses the historical securities return along with macroeconomic variables (raw descriptors) to determine those variables with explanatory power (Burmeister, Roll and Ross (2003)). The fundamental factor models use company and industry variables and market data as independent variables that are expected to have relevant predictive powers (Fama and French (1992a); Fama and French (1993)).

3.4. Estimating volatility

It has been perceived that good return expectations will overpower weak volatility estimation. However, extensive empirical evidence has proved the importance of the accuracy of the volatility estimation technique. Unlike prices and returns, volatility and co-volatility are not directly observable in the market and must instead be estimated. Since the 1990s, the volatility estimation techniques have experienced considerable innovation and improvements. Among the key improvements in the volatility estimation are volatility decomposition and the attention paid to the risk of the tails. Volatility measures can be categorized in two groups: dispersion and downside.

3.4.1. Dispersion volatility models

Dispersion models measure the security's deviation from the mean. These models consider both the positive deviation and the negative deviation as a risk, which reflects an unfavorable sign. However, this is not always the case. For an investor with a long position, an increasing price (positive deviation) over the mean reflects good news, and for an investor with a short position, a decreasing price (negative deviation) over the mean reflects good news. The dispersion volatility measurements include the standard deviation and the variance, the mean-absolute deviation, and the mean-absolute moments. The MV optimization use of a dispersion volatility measurement is the most common critique levied against it, as it penalizes the positive deviation.

3.4.2. Downside volatility models

These models assume that the investor is taking a long position, and considers only the downside, negative deviation, as a volatility measure (Bawa (1975); Nawrocki (1999)). Theoretically, these models are more appealing, but they have some practical limitations. Typically, these models are computationally more complicated to use in a portfolio setting, cannot be aggregated into portfolio downside risk, leads to higher risk measurement, and are accompanied by a higher estimation error (Grootveld and Hallerbach (1999)). In this section, six categories of downside volatility measurements are presented. These are traditional models, VaR models, implied volatility models, Markov Chains and volatility clustering, ARCH/GARCH models, and stochastic models.

3.4.2.1. Traditional models

Roy's paper on safety first was one of the first attempts to consider only the downside risk (Roy (1952)). Harry Markowitz proposed using the semi-variance, since the variance penalizes over-performance (Jin, Markowitz and Yu Zhou (2006); Markowitz (1991)). Bawa (1976) provided a more general measure for the downside risk called lower partial moment risk measure (Bawa (1976)).

3.4.2.2. VaR, CVaR, and CaViaR

JP Morgan developed the Value-at-Risk (VaR) measurement, which is now perceived to be the most well-known downside risk measurement (RiskMetrics). The VaR measures the predicted maximum loss at a pre-specified probability level, over a certain time horizon. Since the year 2000, many academic researchers have investigated the VaR as a measure of downside risk (Alexander and Baptista (2002); Mittnik, Rachev and Schwartz (2002); Pflug and Gaivoronski (2005); Yamai and Yoshida (2005)). Before defining the VaR, two dimensions need to be decided upon. The first dimension is the time interval used in computing the VaR. This time interval reflects the investment horizon, or the time period before revising the portfolio. The second is the confidence interval (α), which is the probability that the VaR is not expected to exceed a certain maximum loss. VaR can be defined mathematically in several ways. The following notations are used in standard references:

$$VaR_{\alpha} = \text{Min} \{X | P(L \geq X) \leq 1 - \alpha\}$$

where

VaR_{α} is the portfolio's (security's) Value-at-risk with α confidence interval.

X is an arbitrary small value of the portfolio (security)

P is the probability function

L is the possible portfolio loss

α is the confidence interval in percentage.

The above equation states that the VaR for a confidence level α is the value X , such that the probability that the possible loss L exceeds X is at most some small number $(1 - \alpha)$. Even with the VaR, a number of drawbacks exist. These drawbacks can be summarized as follows (Rau-Bredow and Str (2004)): subadditivity, non-smooth and non-convex function, and ignoring the magnitude of losses beyond the VaR value. First, the subadditivity problem refers to the fact that the VaR computed for a portfolio may be higher than the weighted average of the VaR computed for the individual securities in the optimal portfolio.

Thus, due to the subadditivity, the VaR penalizes the portfolio construction and it negates the portfolio diversification concept. Second, the VaR leads to a non-smooth and non-convex function, which will hinder the optimizer from finding the optimality condition. The VaR leads to multiple stationary points that complicate the optimization process (Grootveld and Hallerbach (2004)). Third, the VaR only computes the probable losses until the VaR value; it does not consider the magnitude of the losses beyond the VaR value.

All these drawbacks lead to the conditional variant of the VaR, Conditional Value-at-Risk (CVaR). The CVaR measures the expected losses in the tail of the distribution of possible portfolio losses, beyond the portfolio VaR (Artzner et al. (1999)).

Thus, the CVaR measures the market risk conditional on the current information. The CVaR can be mathematically defined using the following formula:

$$CVaR_{\alpha} = E(L|L \geq VaR_{\alpha})$$

A concept that is defined very closely to the CVaR is the expected shortfall (Acerbi and Tasche (2002)). The expected shortfall measurement result has been shown to be equivalent to the CVaR, even though it is defined in a different way. Engle and Manganelli (2004) used a different methodology to compute a different variant of the CVaR. They called their measure conditional autoregressive Value at Risk (CaViaR). Using an autoregressive process, their model specifies the evolution of the CVaR over time.

3.4.2.3. Implied Volatility

Implied Volatility (IV) measures the volatility as implied by the stock option market. Instead of using the historical data to estimate the volatility, this technique uses the actual options' prices (using an option pricing model) to determine the asset's implied volatility. This model has the obvious advantage of using a forward looking estimate, in which it relies on the expectations of the investors, rather than relying of the historical prices.

The IV model suffers from the following shortcomings, which affect its applicability and accuracy. First, option contracts are not available for every single security in which an investor might be interested. Second, the estimation could be perceived as a snapshot or a volatility measure at one point in time. Third, the usage of an option pricing model to calculate the prices adds another risk, i.e., the modeling risk.

Thus, this technique could be used as a diagnostic tool, rather than a stand-alone estimation model. The evidence available in the literature about whether the estimated IV is superior to the realized historical volatility is mixed (Duque and Paxson (1999); Fung and Hsieh (1991)).

3.4.2.4. Markov Chains and volatility clustering

The Markov Chains is perceived as a clustering technique that forms groups that can be distinguished from other groups. The Markov Chains is a discrete-time stochastic process that drops the concept of regimes independence over time. The Markov Chains can be perceived as a chain with no memory. The use of the Markov Chains in portfolio construction became more evident by the seminal work of Jacquier et al. (1994) who fully carried out Bayesian inference through a Markov chain Monte Carlo scheme. The transition probability has the benefit of producing volatility clustering. However, the main shortcoming of using the Markov Chain is the increased number of parameters that must be estimated due to the inclusion of the transition probability. For more detailed applications of the volatility clustering measurements, see Duda et al. (2000) and Focardi (2004).

3.4.2.5. ARCH/GARCH models

Regression models assume that the observations used are independent in nature. Volatility levels tend to be persistent, and thus volatility of different periods tends to be related. The Autoregressive Conditionally Heteroscedastic (ARCH) process accounts for the dependency that some variables possess (Engle (1982)). Bollerslev introduced a

generalization for the ARCH process (GARCH), (see Bollerslev (1986)). The GARCH model generally assumes dependency of the variance over time. Under the GARCH (p, q) process, the conditional variance h_t^2 evolves according to the following equation (where α_0, α_i and γ_j are the parameters to be estimated):

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j}^2$$

The GARCH models allow for long memory processes; they use the past squared residuals to estimate the current variance. If the value of α_i and γ_i are positive, the GARCH model would imply that the current variance will be above its long-run mean during high volatility periods. The GARCH accounts for the fat-tailed distribution and it captures the volatility clustering. A major advantage of the GARCH model is its consistency in estimating volatilities over longer time horizons (Chang and Yang 2005). The GARCH model has been comprehensively tested. See for instance, Bollerslev, Chou and Kroner (1992)' Bollerslev, Engle and Wooldridge (1988), and Ding and Engle (2001).

However, the large scale application of the GARCH model in portfolio construction is not explored well. The main reason is the huge correlation matrix that the GARCH model would produce, which makes it hard for optimization. Further, the GARCH models are very dependent on the data frequency, and thus, the results will be different by the use of different frequencies. The first attempt to create a multivariate GARCH model was made by Bollerslev, Engle and Wooldridge (1988). The model was

referred to as VECH. The major disadvantage was that the conditional variance-covariance cannot often be positive semi-definite (Lien and Luo (1994)). Among the seminal empirical work in the multivariate analysis of variance are the works of Aguilar and West (2000); Chib, Nardari and Shephard (2006); Harvey, Ruiz and Shephard (1994b); Lopes and Migon (2002); and Pitt and Shephard (1999).

Engle and Kroner proposed a multivariate GARCH model that will ensure that the covariance matrix is positive semi-definite (Engle and Kroner (1995)). The model is known as BEKK (standing for Baba, Engle, Kraft, Kroner). The VECH and the BEKK models had to be re-estimated each time period for the whole data set. Engle (2002) proposed a new class of multivariate GARCH models called the Dynamic Conditional Correlation (DCC) models (Engle (2002)). His model has the flexibility of the univariate GARCH but not the complication of the multivariate GARCH. This new setting solves the main problem that the other GARCH models have in the portfolio construction context. In the DCC-GARCH, the number of parameters to be estimated is independent of the number of series to be correlated. Thus, it makes the estimation of very large matrices applicable. The applicability of this model was analyzed by (Engle and Sheppard (2001)). The results showed that the DCC-GARCH often provided the most accurate estimation results.

3.5. Estimating Co-Volatility

The estimation of the co-movement or the co-volatility lies at the heart of the portfolio construction, as well as of the risk diversification. The estimation of the co-volatility can be claimed as the most important parameter in the whole process. The most

commonly used measure for the co-volatility is the sample covariance matrix estimator, calculated using historical prices. The covariance matrix represents a summary of the estimated volatility (matrix diagonal) and co-volatility (off-diagonal) of the assets used in a portfolio. In the N securities case, the sample covariance matrix (Σ) can be defined, using matrix form, as:

$$\Sigma = \frac{1}{N-1} \mathbf{X} \mathbf{X}'$$

where \mathbf{X} is the N by T matrix of the time series of the deviation of the N securities over their historical mean (\bar{R}_i):

$$\mathbf{X} = \begin{bmatrix} R_{11} & \dots & R_{1T} \\ \vdots & \ddots & \vdots \\ R_{N1} & \dots & R_{NT} \end{bmatrix} - \begin{bmatrix} \bar{R}_1 & \dots & \bar{R}_1 \\ \vdots & \ddots & \vdots \\ \bar{R}_N & \dots & \bar{R}_N \end{bmatrix}$$

Markets and economic conditions change from one period to another, and firms react differently towards these changing conditions. Thus, assuming that the past is a good estimate for the future leads to biased estimators. A large number of academic researchers showed the non-stationarity of different estimators, and thus, showing that extrapolating these estimators to the future provides very poor forecasting power (Fama and French (2002)). One of the first techniques used to account for the non-stationarity problem was to use the excess return, instead of using the return figures. The advantage of using the excess return is the subtraction of a return figure, which is being affected by the same market and economic conditions.

Using the excess return time series does not solve the problem. The sample mean, standard deviation, and covariance matrix are still poor estimates for a number of reasons. First, the sample estimators provide good estimators only for normally distributed time series, or generally for distributions with no excess kurtosis. Financial time series are known to be heavy-tailed. Thus, using the sample estimators to forecast the future is expected to be biased (Ibragimov (2005)). Second, estimating the covariance matrix for N securities requires N time periods of observed return. Thus, the estimation process could be bounded by data availability. Third, the covariance matrix is highly sensitive to the estimation error, which might affect the value of the resulting optimal portfolio.

Fourth, the covariance matrix needs to be inverted to reach the optimality condition. A non-singular, and thus non-invertible, covariance matrix will prohibit an exact optimizer from reaching its optimality condition (Michaud (1998)). Many commercial optimizers are non-exact optimizers that are insensitive to whether the covariance matrix is well-defined or not. The optimal portfolios produced by these optimizers are inaccurate and should be avoided. The ill-conditioned covariance matrix is the most important reason behind the instability of an optimizer and should be handled with care.

Several improvements were suggested to improve the forecasting power of the covariance matrix. The first improvements attempted to capture the changing nature of the markets and the economic conditions by using weighted data. Given that the most recent data is more relevant. Thus, more weight is assigned to it, and less weight is

assigned to the older data (Litterman and Winkelmann (1998); Pafka, Potters and Kondor (2004)). However, for the weighted covariance matrix to have a viable forecasting power, the distribution of the return time series must be independent and identical, which is not always the case. In the general case, when the return is not independent and identically distributed, the sample covariance matrix will be highly sensitive to the estimation error, thus the results of the MV optimization will be inaccurate.

The remaining part of this section discusses different attempts to improve the forecasting power of the covariance matrix. These attempts are the rolling windows, ARMA and ARIMA models, covariance matrix shrinkage, discounting, using implied co-volatility, incorporating outliers, and decomposition of co-volatility.

3.5.1. Rolling window and Exponential smoothing

Officer (1973) was among the first to account for the volatility and co-volatility transitions by using monthly moving series for the period 1897-1969 to assess the volatility transition. The calculation of the moving average progresses from a simple moving average, to a weighted moving average, to an exponential moving averages (Brown's exponential smoothing). See Brown and Meyer (1961) and GARDNER Jr (1985).

3.5.2. ARMA and ARIMA models

The Autoregressive Moving Average (ARMA) models and the Autoregressive Integrated Moving Average (ARIMA) models are generalizations of the moving averages (Campbell, Lo and MacKinlay (1996)). The ARIMA models have three stages:

identification, estimation, and forecasting. In the identification stage, the sample autocorrelation and partial autocorrelation that can be examined to establish a pattern for the model is identified. The estimation stage serves two purposes: it estimates the model's parameters, and it serves as a check point for adequacy of the model. The forecasting stage, as the name indicates, provides the forecast.

3.5.3. Stein or Shrinkage covariance matrix

In the mid 1950s, Charles Stein provided evidence that there are uniformly better methods for estimating parameters of a time series than the sample parameters (Stein (1955)). The Stein (Shrinkage) estimators can be perceived as an example of the Bayesian estimation procedures. A number of Stein estimators were developed for the MV optimization. These attempts include Frost and Savarino (1986), James and Stein (1961), Ledoit and Wolf (2004), Ledoit (1994), and Stein (1955). Ledoit and Olivier (2004) used a covariance matrix that is composed from a constant correlation matrix and a CAPM. Their results indicate that using the Shrinkage covariance matrix provides more superior results than the sample covariance matrix.

Theoretically, the Shrinkage estimators are superior to the sample estimators, especially when the estimation error is high. But from an application point of view, Disatnik and Benninga, using historical data of NYSE stocks, showed that the optimal portfolios produced using the Shrinkage estimators are not significantly better than those produced using the traditional sample covariance matrix, see Disatnik and Benninga (2007).

3.5.4. Covariance matrix discounting

A major problem with volatility and co-volatility is the non-stationarity. Both volatility and co-volatility vary over time. One of the early models that attempted to deal with time varying co-volatility was introduced by Quintana and West (1987). They named their model the Covariance Matrix Discounting (CMD). They estimated the covariance matrix as a weighted average of past covariance matrices. They used weights that decay exponentially. The decay rate is estimated using maximum likelihood estimation. Like the Stein estimators, the covariance matrix discounting is theoretically better than the sample covariance matrix, but does not add significant predictive power (Li (1997)). Thus, it is not expected to have a superior investment value. A more viable improvement could be achieved by using dynamic (Bayesian) factors that reflects changes in the market conditions (Aguilar and West (2000)).

3.5.5. Implied co-volatility

As discussed before, implied parameters are forward-looking parameters implied by the stock option market. One way to calculate a forward looking covariance matrix is the use of implied volatility. The covariance matrix can be calculated as the matrix product of the static correlation matrix and the implied volatility diagonal matrix (Fabozzi, Kolm and Pachamanova (2007)). Formally, the implied co-volatility (Σ) can be defined as:

$$\Sigma = A C A'$$

where \mathbf{C} is the static correlation matrix calculated using the securities historical returns, and \mathbf{A} is the diagonal matrix of the implied volatility of the securities used. Using this methodology gives the advantage of both using the simple historical correlation matrix and adding a forward-looking component using the implied volatility. However, the question about whether the implied volatility adds any investment value remains unanswered.

3.5.6. Incorporating outliers

Another way to account for the time-varying covariance matrix is a simple methodology introduced by Chow et al. (1999). The contribution of this model is to account for outliers or co-volatility during high volatility periods. They estimated two covariance matrices during two different time periods. One time period is characterized by low volatility and the other by high volatility. The full sample covariance matrix is then calculated as a weighted average of the two matrices. The full sample covariance matrix can be formally defined as:

$$\mathbf{\Sigma} = p \mathbf{\Sigma}_{\text{low}} + (1 - p) \mathbf{\Sigma}_{\text{high}}$$

An improvement to this methodology was developed by Kritzman, Lowry and VAN Royen (2001). Their contribution was the use of a two-stage Markov Chain regime switching model to determine the weights. Thus, instead of having only one full covariance matrix, the result will be a time-varying covariance matrix. The time-varying full sample covariance matrix can be formally defined as:

$$\Sigma_t = p_t \Sigma_{\text{low}} + (1 - p_t) \Sigma_{\text{high}}$$

3.5.7. Decomposition of co-volatility

Several studies have researched the decomposition of variance into stationary and transitory components, but not many have attempted to decompose the co-volatility, especially under a multivariate setting. The apparent reason is the notorious complication of the multivariate co-volatility models, see, Campbell *et al.* (2001). Returns can be perceived as being driven by both stationary and transitory components. During stable periods, the stationary component is the main driver, and during the unstable periods the asset returns drift away from the stationary components, which are the transitory components.

The return distribution is assumed to be normal with constant mean and variance. Consequently, the return volatility over the long run would not have a positive trend. However, in the transitory periods, returns are assumed to be produced by distribution with a constant daily variance, but whose expected return is mean-reverting. Thus, volatility over longer periods would depend on the parameters of the mean reverting process.

Chapter 4

Empirical Results

The merits of investment diversification and portfolio construction are more evident now than in the past. Using daily data for the period of 1962-1997, Campbell et al. (2001) constructed monthly realized volatility measures and showed that, unlike market volatility and industrial volatility, firm-level volatility increased over time for the period of their study. Thus, the number of securities needed to reach optimality increased due to the lower correlation coefficient. Morck, Yeung and Yu (2000) reached similar results using emerging market data. Their research explained the existence of positive volatility trend over time by the continuous increase in the number of publicly traded firms, the serial correlation non-stationarity in the daily data, and/or the increased trading volume.

Wei and Zhang (2006) investigated the same fact and found similar results. Using daily data for the period of 1976-2000, they showed that the increase in the corporate earnings' volatility and the general decrease in the corporate earnings as measured by ROE have led to the increase in securities' return volatility. Their relation was more evident for the newly listed securities, which could be due to the reason documented by PASTOR and PIETRO (2003) who analyzed the firm specific characteristics in a rational asset pricing model structure. They found that profitability induces an age effect that, in turn, increases confinement and thus reduces volatility. New firms lack this history of profitability; therefore they suffer from higher idiosyncratic volatility. There is now more

evidence that the diversifiable risk is increasing over time. Thus, regardless of what drives the positive trend of the firm-specific-volatility, its existence increases the importance of portfolio construction.

In this chapter optimal portfolios are created using a multistage optimization. First, the Monte Carlo simulation algorithm is used to generate multiple scenarios, and then an iterative variant of the greedy algorithm is used to create optimal portfolios. This chapter is organized as follows. The first section provides a description of the data set used. The second section describes in detail the multistage optimization methodology. The third section presents the results. The conclusion is discussed separately in the following chapter.

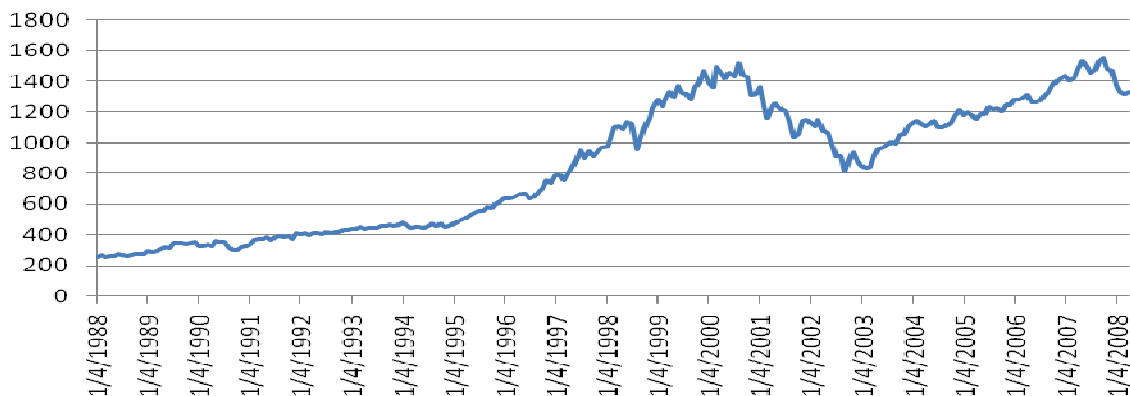
4.1. Data

USA securities markets include approximately 10,000 different securities of 10,000 corporations. A significant large percentage of these securities are not actively traded, so, their prices are generally stale. Attempting to feed a portfolio optimizer with this large number of securities will either decrease the efficiency of the optimizer or lead to computational failure. An approach to reduce the number of securities is the use of “securities screeners” which was discussed in chapter two. A security screener specifies selection criteria and eliminates securities that do not meet these criteria. Some securities screeners work better than others; thus, using a security screener may affect our results. If superior results are reached using a specific optimization technique in combination with a good securities screener, the superior results could be linked to either the securities

screeners or the optimization technique. To avoid this problem, no security screeners were used in this dissertation.

The initial attempt to reduce the total number of securities used in this dissertation was made by considering only the large securities as measured by their market capitalization. The 3000 securities prices of the corporations included in the Russell 3000 are used. These 3000 securities represent 98 percent of the investable US securities. After making a number of runs, it was found that the prices of a significant number of these 3000 securities are stale. Consequently, the securities' pool is narrowed down to the securities included in the Russell 1000, and then narrowed down further to the securities included in the S&P500. Fifty securities were repeatedly selected randomly from the S&P500 securities. The monthly adjusted closing prices for these 50 securities were acquired from Bloomberg. The closing prices were adjusted for both dividends and security splits. The data set used ranged from January 2001 until July 2008, a period characterized by its high volatility (see figure 5). The adjusted closing prices of the S&P500 index are used as a benchmark for the same period.

Figure 5: S&P500 Price movement over the period 1988 - 2008



Securities prices are not normally distributed and are not I(0). Thus, the initial step in data preparation was to convert the closing prices (p_t) to returns:

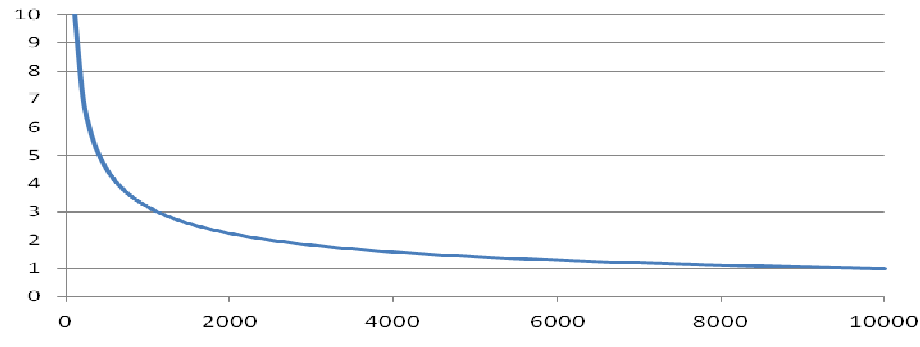
$$r_t = \ln(p_t) - \ln(p_{t-1})$$

The resulting returns, especially for the period chosen, included many negative returns. Negative returns cannot be fitted to a number of probability distribution, including the lognormal distribution as will be explained in the methodology section. To avoid the problems associated with the negative returns, return relatives are calculated by adding one to the natural log returns. The natural log return of the S&P500 index is then subtracted from each return relative to calculate the residual return relative. The benefits of the use of the residual return measurement have been documented in the literature (Michaud (1998)). One of the main benefits is to partially account for non-stationarity. The residual return subtracts a benchmark return figure – both being affected by the same market and economic conditions – from the return figures used. A different variation of the MV optimization was created using the residual returns, known as benchmark optimization (Michaud (1998); Roll (1992)).

To improve the accuracy of the results, simulated data sets are used in addition to of the original data set. Using the residual-return relative time series of the 50 securities, as a base, 2000 simulated data sets were created using the Monte Carlo simulation. The methodology used to create the simulated data sets is similar to the methodology used in the portfolio resampling techniques (Scherer (2002)). The reason for choosing 2000 simulations is to reduce the estimated standard error (see figure 6). After 2000

simulations, the incremental benefits are not justified. The first 2000 simulations approximately reduce the standard error from ten to two, but the following 8000 simulations only reduce the standard error from two to one.

Figure 6: The Standard Error Value as the Number of Simulations Increase



4.2. Methodology

In this section the methodology used to construct the optimal portfolios is discussed. Using the simulated data, we start by estimating the parameters of the 50 securities: volatility, co-volatility, and expected return. Four different estimation techniques are applied to estimate the volatility and the co-volatility. Initially, the traditional – unconditional – variance and covariance estimation using the historical returns is used. The traditional variance-covariances matrix is calculated to enable the comparison with the traditional mean-variance portfolio optimization. The conditional variance and covariance is then calculated using three different types of GARCH models: CCC-GARCH, IDCC-GARCH, and the O-GARCH.

The expected return for the 50 securities is then calculated by fitting the return historical time series to a probability distribution, using the Anderson-Darling goodness-of-fit methodology (Anderson and Darling (1952)). Kolmogorov-Smirnov and Chi-

square were also calculated, but not reported as they gave very close results to the Anderson and Darling methodology (Massey (1951)). After calculating the securities parameters, portfolios are created using the greedy algorithm.

4.2.1. Estimating volatility and co-volatility

Evidence in the academic literature indicates that the MV efficiency is highly error-prone and unstable. The most important limitation is its high sensitivity to changes in the covariance matrix. A minor change in the covariance matrix will lead to a significant change in the portfolio's optimal weights. In other words, the MV efficiency cannot afford any minor estimation error, especially when it comes to the covariance matrix. In this sub-section, four different estimation techniques are used to estimate the covariance matrix.

4.2.1.1. Traditional covariance matrix

As discussed in chapter three, the covariance matrix is a summary of the estimated volatility (matrix diagonal) and co-volatility (off-diagonal) of the assets used in a portfolio. In the N securities case, the sample covariance matrix (Σ) can be defined using matrix form as:

$$\Sigma = \frac{1}{N - 1} \mathbf{X}\mathbf{X}'$$

where \mathbf{X} is the N by T matrix of the time-series of the deviation of the N securities over their historical mean (\bar{R}_i).

$$\mathbf{X} = \begin{bmatrix} R_{11} & \cdots & R_{1T} \\ \vdots & \ddots & \vdots \\ R_{N1} & \cdots & R_{NT} \end{bmatrix} - \begin{bmatrix} \bar{R}_1 & \cdots & \bar{R}_1 \\ \vdots & \ddots & \vdots \\ \bar{R}_N & \cdots & \bar{R}_N \end{bmatrix}$$

In this dissertation N is 50 securities, T is the monthly observations, and \mathbf{R} is the time-series of the residual return relatives. The first criticism on the tradition covariance matrix is its stationarity. Markets and economic conditions change over time, and firms react differently towards these changes. Thus, assuming that the past is a good estimate for the future leads to biased estimators.

The use of the residual return improves the estimation process, but the traditional covariance matrix still produces poor estimates (see chapter three for more details and a literature review.) To fully account for the parameters' non-stationarity and to improve the estimation results, three conditional variants of volatility and co-volatility are used. These models are: the Orthogonal (o) GARCH, the Constant Conditional Correlation (CCC) GARCH model, and the Integrated Dynamic Conditional Correlation (IDCC) GARCH. The three models are discussed in the following three sub-sections. Following the standard ARCH/GARCH notations, the notations are defined as follows:

(\mathbf{Y}_t) A vector of stochastic process of dimension N x 1,

(\mathbf{I}_{t-1}) The information available at time t-1,

$\mu_t(\theta)$ The conditional mean vector,

$\overset{\pm}{H}_t^2(\theta)$ The N x N positive definite matrix,

ε_t

An $N \times 1$ random vector with mean of zero and standard deviation of one

$$y_t = \mu_t(\theta) + \varepsilon_t$$

$$\varepsilon_t = H_t^{\frac{1}{2}}(\theta) z_t$$

$$\text{Var}(y_t | I_{t-1}) = \text{Var}_{t-1}(y_t) = \text{Var}_{t-1}(\varepsilon_t)$$

$$= H_t^{\frac{1}{2}} \text{Var}_{t-1}(z_t) \left(H_t^{\frac{1}{2}} \right)'$$

$$= H_t$$

H_t is the conditional covariance matrix of y_t . Both the H_t and the y_t are dependent on the unknown parameter vector θ . The specification of H_t differs from one GARCH model to the other.

The multivariate GARCH models could roughly be categorized into the following classes (see Bauwens, Laurent and Rombouts (2006)): (1) direct generalization of the univariate GARCH, which includes the VEC, MGARCH, and the BEKK, (2) linear combinations of the univariate GARCH, which includes the OGARCH, and (3) nonlinear combinations of univariate GARCH, which includes the constant and dynamic conditional correlation GARCH, general dynamic covariance, and copula-GARCH. In this dissertation, only the GARCH (1, 1) is used, so the p and the q are always 1. Only the

second and third categories of the multivariate GARCH are used for estimating the covariance matrix. More specifically, the Orthogonal GARCH, Constant Conditional Correlation GARCH, and the Integrated Dynamic Conditional Correlation GARCH are used. The first category of the multivariate GARCH – VEC and BEKK – is not used due to the fact that the resulting conditional covariance matrix is not positive semi-definite.

4.2.1.2. Orthogonal GARCH model

The OGARCH is a linear combination of univariate models. These models could be ARCH, GARCH, EGARCH, etc. The data are generated by an orthogonal transformation of N univariate GARCH. The matrix generated is the orthogonal matrix of eigenvectors of the unconditional covariance matrix population. The original OGARCH was introduced by Alexander and Chibumba (1997). In the initial model, the N x N conditional covariance matrix is generated by $m \leq N$ univariate GARCH model. The OGARCH (1,1,m) is formally defined as follows:

$$V^{-1/2} \varepsilon_t = u_t = \Lambda_m f_t$$

where $V = \text{diag}(v_1, v_2, \dots, v_N)$, $u_t = \varepsilon_{it} / \sqrt{h_{it}}$, and v_i is the population variance of ε_{it} , and Λ_m is a matrix of dimension N x m. The parameters of the model are V, Λ_m and the parameters of the GARCH. The number of parameters is N(N+5)/2, if m = N. “m” is chosen by principal component analysis applied to the standardized residuals \hat{u}_t .

4.2.1.3. Constant Conditional Correlation GARCH model

The multivariate GARCH models started with the VEC model proposed by Bollerslev, Engle and Wooldridge (1988) but the model suffered from the fact that the resulting covariance matrix is not positive definite or semi-definite, thus affecting the optimality condition of any optimization. Furthermore, its covariance matrix needs to be re-calculated for every time period. Bollerslev (1990) proposed the Constant Conditional Correlation (CCC) GARCH model. The CCC-GARCH can be used in a multivariate setting to produce a time varying conditional variance and covariance and a constant correlation matrix. Bollerslev (1990) stated that the maximum likelihood estimate of the correlation matrix is equal to the sample correlation matrix, and as the sample correlation matrix is positive semi-definite, the positive semi-definiteness of the conditional covariance matrix can be ensured. The CCC-GARCH can be formally defined as follows;

$$H_t = D_t R D_t = (\rho_{ij} \sqrt{h_{iit} h_{jjt}})$$

$$D_t = \text{diag}(h_{11t}^{-\frac{1}{2}} \dots h_{NNt}^{-\frac{1}{2}})$$

where h_{iit} is the conditional variance calculated using a GARCH model, and R is the constant symmetric positive definite conditional correlation matrix. The CCC-GARCH model contains $N(N+5)/2$ parameters, and the unconditional covariances are difficult to calculate due to the model nonlinearity. The CCC-GARCH is computationally intensive and the assumption that the conditional correlations are constant is unrealistic for many empirical applications.

4.2.1.4. Integrated Dynamic Conditionally Correlated GARCH

Engle and Kroner proposed an improved model, BEKK, which ensures the positive semi-definiteness of the covariance matrix (Engle and Kroner (1995)). The problem with the BEKK is that it is computationally intensive. Several authors propose a generalization to the CCC-GARCH by making the constant conditional correlation time-dependent; their model is called the Dynamic Conditional Correlation (DCC) GARCH (Christodoulakis and Satchell (2002); Engle (2002); Tse and Tsui (2002)). Unlike the BEKK, the resulting conditional covariance matrix is guaranteed to be positive semi-definite. The models proposed by Engle (2002) and Tse and Tsui (2002) are multivariate models, whereas the model proposed by Christodoulakis and Satchell (2002) is only bivariate. The DCC-GARCH can be defined as follows:

$$H_t = D_t R_t D_t$$

Unlike the CCC-GARCH, the conditional correlation matrix is not constant any more. The conditional correlation matrix is time-varying in the DCC-GARCH, as follows:

$$R_t = \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2})$$

where the $N \times N$ symmetric positive definite matrix $Q_t = (q_{ij,t})$ is given by:

$$Q_t = (1 - \alpha - \beta) Q + \alpha u_{t-1} u'_{t-1} + \beta Q_{t-1}$$

where $u_t = \varepsilon_{it}/\sqrt{h_{it}}$ and \bar{Q} is the $N \times N$ unconditional variance matrix of u_t , and α and β are non-negative scalar parameters that satisfy $\alpha + \beta < 1$. If $\alpha + \beta < 1$, then Q_t is mean reverting. The Integrated DCC-GARCH is a special case of the DCC when the $\alpha + \beta = 1$.

4.2.2. Expected return and stochastic programming

While portfolio optimization algorithms and sophisticated software allow investors to handle large number of securities, the optimal solution reached can still be very sensitive to small fluctuations in the input. Furthermore, regardless of the accuracy, most estimation processes produce results that are deterministic. Thus, to add to the viability of a model, the stochastic nature of the input must be accounted for. One of the ways to account for the stochastic nature of the input is the use of sensitivity analysis. However, the sensitivity analysis deals with uncertainty after the optimal solution is reached. It shows the input ranges over which the current optimal result stays optimal. Furthermore, it can only show changes in one input at a time.

In this dissertation the stochastic nature of the expected return is integrated using stochastic programming and Monte Carlo simulation. Stochastic programming is a class of methods that incorporate the stochastic nature of variables into traditional mathematical programming framework (Ruszczyriski and Shapiro (2006)). Instead of representing the expected returns as deterministic values, they are assigned scenarios generated in advance, based on their historical values. The time-series of the return of the

securities used in our model are fitted into a probability distribution that matches the time-series characteristics using the Anderson-Darling goodness-of-fit criterion.

After each of the time-series is represented by a probability distribution, the Monte Carlo simulation uses the probability distribution to generate different expected returns. For example, if a time-series was fitted to a normal distribution, the mean of the distribution is the most likely value, so it has a better chance of being generated as an input, but that does not prevent other values – with lower probability – from appearing in the simulations.

The main contribution of this dissertation is the use of the greedy algorithm in portfolio optimization. Finding an accurate expected return measure is highly desirable, but is not the major objective of this dissertation. Therefore, large numbers of possible inputs are chosen to feed the proposed optimizer. In the following section, the greedy algorithm proposed in this dissertation is discussed.

4.2.3. Greedy algorithm application

Optimization algorithms are of an iterative nature, i.e., the optimization algorithms go through a sequence of steps with a set of choices at each step that take it closer and closer to the optimal solution. The greedy algorithm is a powerful algorithm that is applicable to a wide range of problems. It follows a problem-solving metaheuristic, in which it makes the best choice at the moment without considering the future consequences. In each iteration the greedy algorithm chooses the optimal solution for that stage, so that with more iterations the global optimal solution is reached. In other

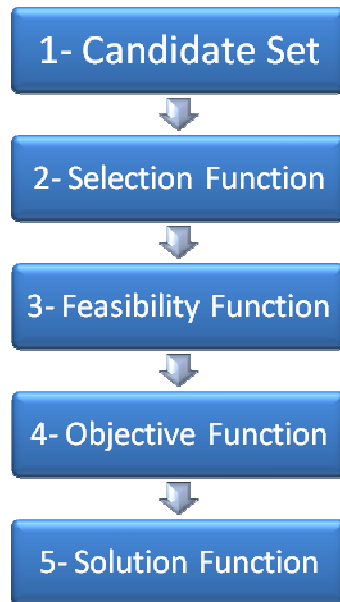
words, the greedy algorithm reaches the optimal solution by making a sequence of choices. In every iteration the choice that seems best at the moment is chosen.

This dissertation proposes the use of the greedy algorithm to optimize the portfolio weights. The greedy algorithm is expected to add value to the portfolio construction for a number of reasons. First, the greedy algorithm is very efficient; it does not reconsider any previous selections but moves to a newer iterations. The efficiency of an algorithm is a significant factor especially with today's computationally complex mathematical models. Second, the greedy algorithm selection process is similar to many investors' selection process, not only due to the rational – greedy – nature of investors, but also due to the fact that investors seek the investment that looks better for them in the short term. Third, the greedy algorithm offers the ability to work with a large number of variables, which addresses the scalability problems.

A greedy algorithm generally consists of five stages (see figure 7). First, it creates a candidate set. Second, a selection procedure is created to choose the optimal value from that candidate set. Third, after an optimal value is selected, the optimal value goes through a screening process to see whether it is feasible or not. Generally this stage is performed using a set of constraints. Fourth, the objective function determines the quality of the current set of choices. Fifth, the solution function determines when the algorithm should be stopped.

The portfolio optimization – using greedy algorithm – proposed in this dissertation follows the same five steps of the general greedy algorithm. The following five sub-sections explain these five steps.

Figure 7: The Greedy Algorithm five stages



4.2.3.1. Candidate set

The candidate set is the input – from which a solution will be chosen – that will be fed to the optimizer. In this dissertation the candidate set is a set of securities for which the covariance matrix is provided. The covariance matrix is a summary of the volatility and the co-volatility of the securities' time-series. As with any model, the quality of the output depends on the quality of the input, other things being equal. The main critique to the MV optimization was its sensitivity to the inputs used. The main input to a portfolio optimization model is the variance-covariance matrix and the expected return. In the model proposed, much attention is paid to the input estimation, so that the model is fed with reliable inputs.

Four different covariance matrices are used. They correspond to the four estimation techniques used. The first covariance matrix is the traditional covariance

matrix, created using the residual return relative time-series. Even though using the residual return improves the traditional covariance matrix, it is still not a robust matrix. The main reason for using the traditional covariance matrix is to compare the results to other models that are built with the traditional covariance matrix, like MV optimization.

The remaining three candidate sets use the GARCH model. The Constant Conditional Correlation (CCC) GARCH model, the Orthogonal GARCH, and the Integrated Dynamic Conditional Correlation (IDCC) GARCH are used. The CCC-GARCH uses a conditional covariance matrix along with the constant correlation. Its major problem was the very long processing time needed. The OGARCH is used to ensure that the covariance matrix is positive semi-definite. All 50 securities are used to force the principle factor component analysis to maintain the identity of each security. The IDCC-GARCH has a number of very appealing qualities. First, it is computationally less intensive than the other multivariate GARCH models. Second, it uses a conditional correlation matrix rather than constant correlation. Third, it guarantees that the conditional covariance matrix is positive semi-definite.

4.2.3.2.Selection function

The steps in the greedy algorithm are as follows. First, two securities are selected using a selection function to be discussed later. Then these two securities are merged into a portfolio with weights determined using the objective function which will be described in this section. Using the optimal weights the two time-series of the two securities are merged into a single time-series. Thus, the total number of securities left to be considered is reduced by one. Then, the process repeats till only one time-series –reflecting the

optimal portfolio's time series – remains. All the optimal weights reached from the different iterations are saved in a log file and then are used to determine the optimal weights to be invested on each security. In other words, the greedy algorithm's iterations repeatedly shrink the number of time-series, and thus the covariance matrix, until only one time series remains, at which point the process stops.

The selection function determines the criteria for selecting the best candidates to merge from the candidate set. In this dissertation, five different selection functions are used, corresponding to five different models. Given the four candidate sets and the five selection functions, 20 different variants are used. The first selection function was designed to select the two securities with the lowest covariance, the least two correlated securities. If the data set contains both positively and negatively correlated securities, then this procedure selects, for merging, a pair of securities with the most negative correlation. This selection function's objective is to bundle or merge the assets with the least correlation to maximize the diversification benefits. However, it was found that the securities that got bundled earlier in the process are more susceptible to be diluted from further and further bundling, and thus end up with a small percentage in the final portfolio.

The second selection function, as opposed to the first, selects the two securities with the highest correlation. This is to preserve the securities with low correlation from being diluted and to bundle the securities that behave similarly. When the securities with the high correlation get bundled repeatedly, they have a better chance of being eliminated. The third selection function selects the two securities with the lowest absolute

correlation. The securities with the least absolute correlation are perceived not to be correlated. Their security price movement can be said to be independent, so early in the process the third selection function bundles the securities that seem to be unrelated. The fourth selection function selects the securities with the highest absolute correlation. The fifth selection function selects the securities randomly. The fifth function is used to see whether the selection process matters at all.

4.2.3.3. Feasibility function

As the name indicates, the feasibility function is used to determine if the optimal weight is a feasible weight or not. The feasibility function is a set of constraints that the optimal weight needs to satisfy before moving to the next stage. A list of possible constraints is described in details in chapter two of this dissertation. In the model proposed, only two constraints are applied; budget constraint and long-only constraint. The budget constraint is used to ensure that the total weights are not more than 100 percent of the investment money available. The long-only constraint prevents the weights from having any negative values, which reflects short positions.

4.2.3.4. Objective function

The objective function is not the function to be maximized or minimized; rather, it is the procedure used to determine the optimal weight for each iteration. As explained in the previous section, the portfolio selection problem is reduced to two-security portfolio. Thus, only the optimal weight between two securities that will minimize the portfolio's risk is needed. It is well known that the risk as measured by variance of a two-security-

portfolio (σ_p^2) can be calculated using the following equation, where σ_{12} is the unconditional covariance between the two securities and w_i is the weight to be optimized:

$$\sigma_p^2 = w_1^2 \cdot \sigma_1^2 + w_2^2 \cdot \sigma_2^2 + 2 \cdot w_1 \cdot w_2 \cdot \sigma_{12}$$

Taking the first derivative of this equation, equating the result to zero, and solving for the desired weight, leads to the optimal weight that will minimize this two-security portfolio's risk. Differentiating the two-security portfolio's standard deviation with respect to w_1 gives the following equation:

$$w_1 = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2 \cdot \sigma_{12}}$$

When the GARCH model is used, the conditional covariance $h_{12,t}$ is used instead of the unconditional covariance σ_{12} .

4.2.3.5. Solution function

The solution function determines when the solution is reached. As mentioned earlier, the greedy algorithm has an iterative nature: a sequence of solutions is generated that gets closer and closer to the optimal weight. However, since the optimal weight is not known in advance and the optimization should not continue running indefinitely, the process must be terminated using some criteria. Different criteria can be used. One of the common criteria is to stop when no progress has been made; other criteria are based on a time interval. The greedy algorithm proposed in this dissertation is an exact optimizer, which means that, using the same inputs, each time the optimizer is used, it will give the

exact same results. The optimization will be terminated when the number of time-series in the initial data set used reaches one time-series.

4.3. Results

Before the optimization is started, the data sets used are fitted to probability distributions using the Anderson-Darling goodness-of-fit criteria to integrate the stochastic nature of the input. The probability distributions were restricted to the normal and the lognormal distributions. Given that different securities are affected to varying degrees by the same market and economic factors, the correlation between the time-series was taken into consideration when the simulations were generated. Spearman rank correlation coefficient, a non-parametric estimator of the correlation coefficient calculated from the ranks of the observations, is used instead of the Pearson product-moment correlation coefficient because the Pearson correlation does not generalize to all the distributions.

Using the possible expected values based on the fitted probability distribution, Monte Carlo simulation generates 2000 different scenarios of the expected return for each security. These expected returns are fed to the optimizer to locate the optimal weights. After calculating the expected return scenarios, the covariance matrix is estimated. Using the 50 time-series, a 50 by 50 covariance matrix is estimated, and then the selection function picks two securities and merges them using the weights determined by the objective function. Now the total number of time-series remaining is 49 time-series. The remaining 49 time-series are then used to estimate a new 49 by 49 covariance matrix. The process continues until only one time-series remains.

As explained in the methodology section, the covariance matrices are estimated using different techniques: traditional unconditional, OGRARCH, CCC-GARCH, and IDCC-GARCH covariance matrices. Each estimating technique is used with five different selection criteria: absolute maximum, absolute minimum, maximum, minimum, and random. Therefore, 20 variants of the model are created. Each model generates a set of optimal weights. These weights are used to calculate the portfolio's risk and return, which are then used to calculate the Sharpe ratio and the VaR for the portfolio.

The first attempt of the proposed model used 50 securities randomly selected from the S&P500. Table 1 shows the Sharpe's ratio for the different models. The CCC-GARCH model was found to be very inefficient in the portfolio setting proposed. Each run took approximately 27 hours and 18 minutes to generate the results. All the other variants of the model used much less time, ranging from a couple of seconds to 14 minutes. It does not make sense to use a model that cannot produce the optimal values in a reasonable time. Thus, the results of the CCC-GARCH were not investigated further.

Table 1
Average Sharpe Ratios using return relative

| Covariance matrix | Traditional | OGARCH | IDCC-GARCH | CCC-GARCH |
|--------------------------|--------------------|---------------|-------------------|------------------|
| Abs. Maximum | 0.52 | 0.79* | 0.66 | --- |
| Abs. Minimum | 0.12 | 0.29 | 0.40 | --- |
| Maximum | 0.27 | 0.27 | 0.52 | 0.82* |
| Min | 0.27 | 0.76 | 0.78* | 0.16 |
| Random | 0.62* | 0.72 | 0.73 | --- |

*The largest average Sharpe ratios within each estimation method.

A commercially used stochastic portfolio optimizer is used as a benchmark to compare the results. The commercial optimizer is allowed 15 minutes to run the optimization. Using the same input, the average Sharpe ratio calculated for the benchmark is 0.56. This value is better than most of the values calculated using the traditional covariance except with the random selection. Using the OGARCH and the IDCC-GARCH models, six out of the ten trials performed better than the benchmark.

Table 2
Average Sharpe Ratios using residual return relative

| Covariance matrix | Traditional | OGARCH | IDCC-GARCH |
|--------------------------|--------------------|---------------|-------------------|
| Abs. Maximum | 0.65 | 0.91 | 0.58* |
| Abs. Minimum | 0.25 | 0.56 | 0.40 |
| Maximum | 0.58 | 0.56 | 0.57 |
| Min | 0.20 | 0.88 | 0.37 |
| Random | 0.67* | 0.94* | 0.52 |

*The largest average Sharpe ratios within each estimation method.

When the residual return relative is used instead of the return relative time-series, the commercial optimizer's Sharpe ratio became 0.52. Table 2 shows that three out of the five selection criteria performed better using the traditional and the IDCC-GARCH model. Using the OGARCH model, all the selection criteria performed better. Tables 1 and 2 summarize the initial tests performed to assess the performance of the proposed model.

To generalize the results, a larger data set – 850 securities included in the Russell 1000 – was used to test the model. 150 securities were not used due to data availability. 50 portfolios were created from the 850 securities pool. Each portfolio was composed of 50 securities selected randomly from the pool. The 50 portfolios were fed – one at a time – to the optimizer to determine the optimal weight. The optimal weights were then used

to calculate the portfolio's risk and return, which were then used to calculate the portfolio's VaR. Table 3 shows the descriptive statistics of the VaR calculated for the 50 portfolios using the different selection criteria and the different covariance estimation techniques.

Table 3a

Descriptive statistics for the VaR values using the traditional covariance matrix

| | Tmax | Tmin | Tran | Tabsmax | Tabsmin |
|-----------------------|-------------|-------------|-------------|----------------|----------------|
| Mean | \$61.24 | \$321.40 | \$68.02 | \$58.53 | \$360.46 |
| Variance | 319.30 | 6496.09 | 871.29 | 364.12 | 7121.20 |
| Std. Dev. | \$17.87 | \$80.60 | \$29.52 | \$19.08 | \$84.39 |
| Skewness | 1.5193 | 0.1551 | 3.1758 | 1.3416 | -0.4287 |
| Kurtosis | 6.5468 | 2.6424 | 14.9286 | 3.8185 | 2.5896 |
| Mean Abs. Dev. | \$12.93 | \$66.12 | \$17.73 | \$14.66 | \$69.25 |

Table 3b

Descriptive statistics for the VaR values using the OGARCH covariance matrix

| | Omax | Omin | Oran | Oabsmax | Oabsmin |
|-----------------------|-------------|-------------|-------------|----------------|----------------|
| Mean | \$50.78 | \$210.84 | \$58.77 | \$47.24 | \$180.39 |
| Variance | 216.22 | 14682.38 | 126.85 | 108.19 | 21340.75 |
| Std. Dev. | \$14.70 | \$121.17 | \$11.26 | \$10.40 | \$146.08 |
| Skewness | 2.6263 | -0.0563 | 1.9077 | 1.7287 | 0.4466 |
| Kurtosis | 12.4636 | 1.9983 | 7.3597 | 6.3184 | 1.6849 |
| Mean Abs. Dev. | \$10.11 | \$99.55 | \$7.68 | \$7.66 | \$131.72 |

Table 3c

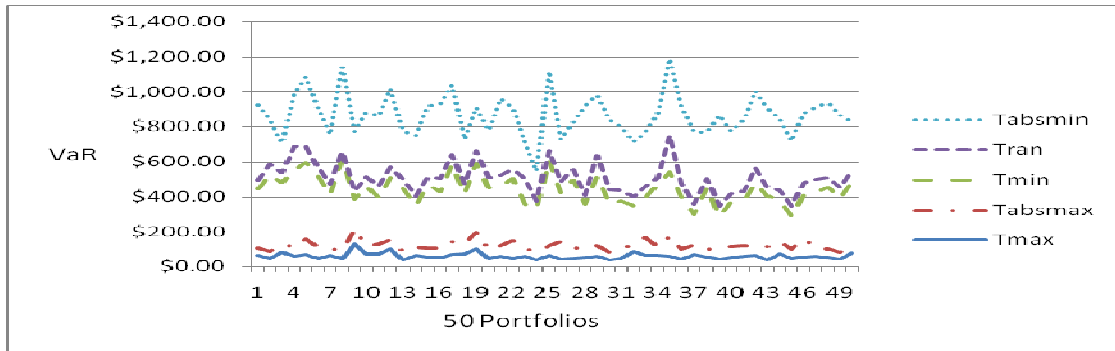
Descriptive statistics for the VaR values using the IDCC-GARCH covariance matrix

| | Imax | Imin | Iran | Iabsmax | Iabsmin |
|-----------------------|-------------|-------------|-------------|----------------|----------------|
| Mean | \$57.91 | \$219.34 | \$68.69 | \$59.96 | \$232.26 |
| Variance | 164.51 | 21522.10 | 127.56 | 325.08 | 7409.96 |
| Std. Dev. | \$12.83 | \$146.70 | \$11.29 | \$18.03 | \$86.08 |
| Skewness | 1.5161 | 0.1345 | 0.7224 | 1.2450 | 0.6211 |
| Kurtosis | 6.2388 | 1.5222 | 4.3332 | 4.3079 | 2.8720 |
| Mean Abs. Dev. | \$8.30 | \$128.72 | \$8.72 | \$13.70 | \$67.94 |

Figure 8 shows the VaR of the 50 optimized portfolios using the traditional covariance matrix assuming an initial investment of \$1000. The best results are reached when the selection function selects the max value in the covariance matrix. The max selection function merges the securities with the maximum correlation – securities that behave similarly. When these securities merge again their weight in the optimal portfolio gets smaller and smaller. Ultimately the securities with low correlation will have a higher weight in the optimal portfolio. The results reached are consistent with the literature, where a portfolio composed of securities with low correlation is expected to have low risk.

Figure 8

Assuming an initial investment of \$1000, the 1%-VaR for 50 portfolios created using the five different selection criteria are calculated, using the traditional covariance matrix.



The use of the conditional covariance matrix improved the results reached. By comparing figure 8, which uses the traditional covariance matrix, and figure 9, which uses the IDCC-GARCH covariance matrix, the improvement can be seen clearly. Further improvements are reached by using the OGARCH, as can be seen in figure 10.

Figure 9

Assuming a \$1000 initial investment, the 1%-VaR for 50 portfolios created using the five different selection criteria are calculated, using the IDCC-GARCH covariance matrix.

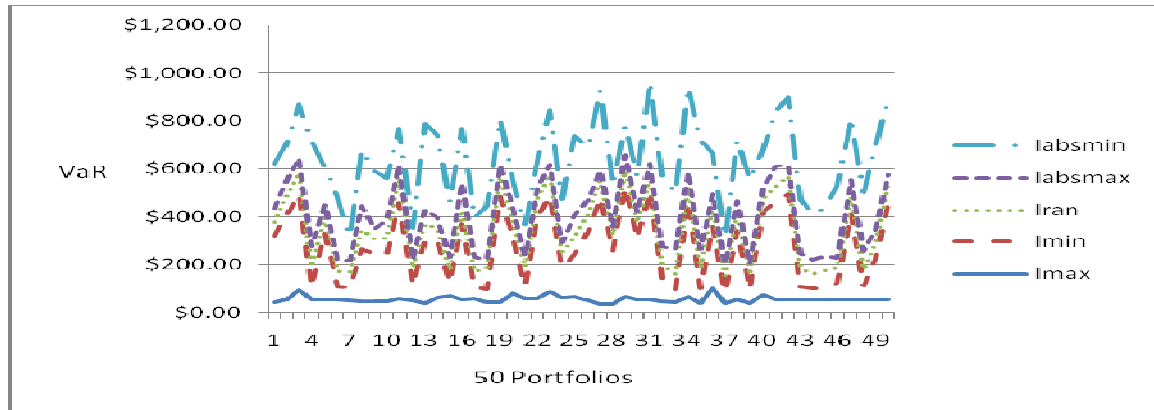
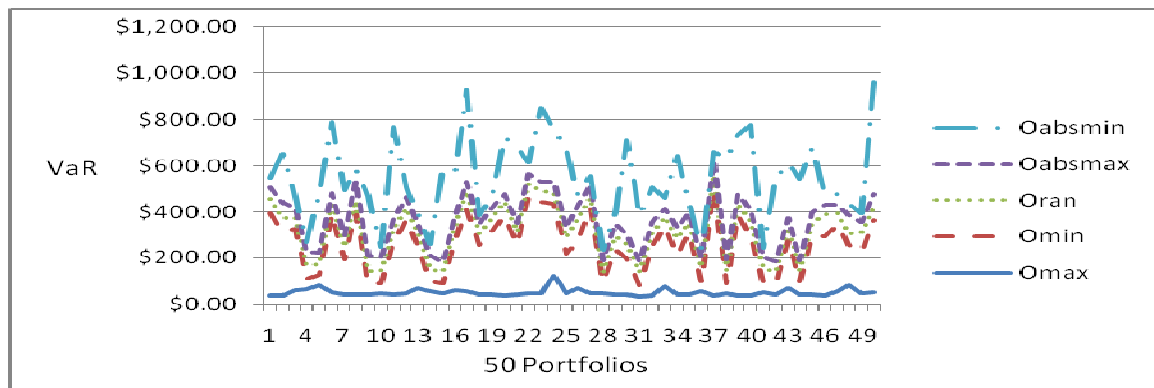


Figure 10

Assuming an initial investment of \$1000, the 1%-VaR for 50 portfolios created using the five different selection criteria are calculated, using the OGARCH covariance matrix.



The Lilliefors test is used to measure the VaR values' normality or departure from normality. The following three graphs show the cumulative distribution functions (cdfs) of the VaR values calculated using the different selection criteria and different covariance estimation techniques. The Lilliefors test measures the maximum vertical distance between the normal cdf and the empirical cdf, and compares this maximum to tabulated values. No significant departure from normality was found (see Figure 11).

Figure 11a
VaR Normal and Empirical Cumulative Distributions (Using traditional Covariance)

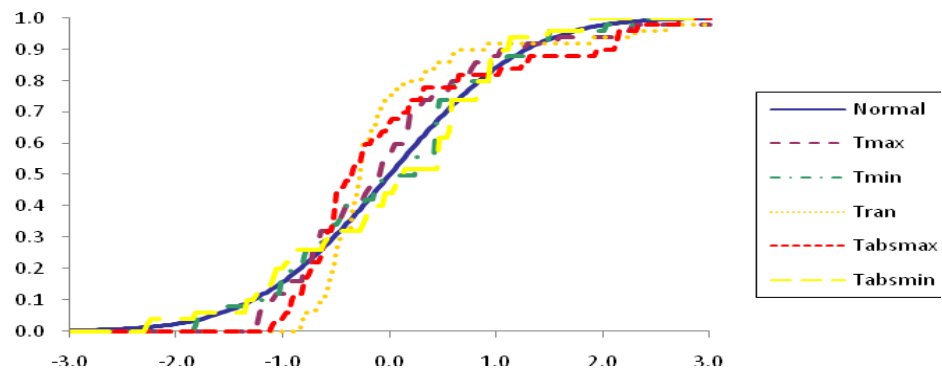


Figure 11b
VaR Normal and Empirical Cumulative Distributions (Using OGARCH Covariance)

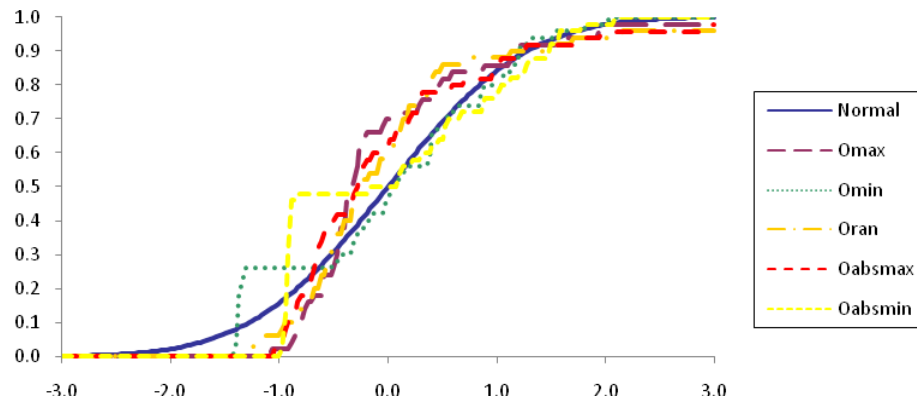
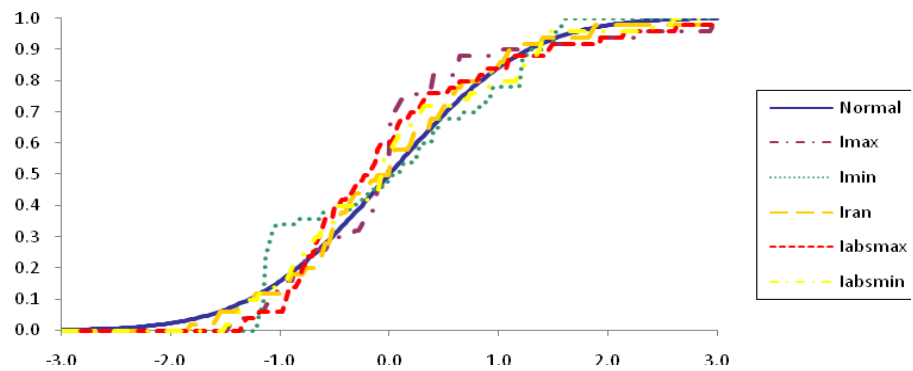


Figure 11c
VaR Normal and Empirical Cumulative Distributions (Using IDCC-GARCH Covariance)



To validate the results shown in the previous three figures, the t-values and the z-values are calculated. The alternative hypothesis states that the mean of the max-selection is lower than the mean of the other selection methods. The t-values (see table 4-4) and the z-values (see table 4-5) show that the max-selection significantly outperforms the min, ran, and abs-min selection techniques in risk reduction.

Table 4

T-tests for Max-Selection Criteria Superiority over other Selection Methods

The table shows t-values for a one-tail hypothesis test. The alternative hypothesis is that the mean of the max-selection is lower than the mean of other selection methods listed.

| Covariance | Traditional | Orthogonal | IDCC |
|-------------------|--------------------|-------------------|-------------|
| Min | 22.28* | 9.27* | 7.75* |
| Ran | 1.389* | 3.05* | 4.46* |
| Abs-Max | -0.73 | -1.39 | 0.65 |
| Abs-Min | 24.528* | 6.24* | 14.17* |

*Null hypothesis is rejected at a 1% Significant.

Now that the results show that the max selection is outperforming the other selection methods, the VaR of the max selection, using the traditional covariance matrix, is compared to the VaR of the Russell 1000 index, S&P500 index, and to the VaR values of 50 portfolios constructed using Crystal Ball. The Crystal Ball (CB) is one of the well known commercial portfolio construction package, currently available in the market.

Table 5

Z-tests for Max-Selection Criteria Superiority over other Selection Methods

The table shows the z-values for a one-tail hypothesis test. The alternative hypothesis is the mean of the max-selection is lower than the mean of other selection methods listed.

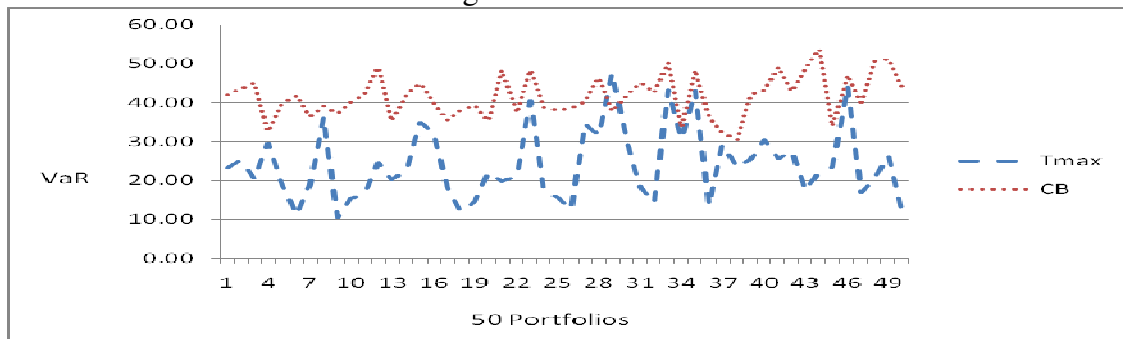
| Covariance | Traditional | OGARCH | IDCC |
|-------------------|--------------------|---------------|-------------|
| Min | 6.15* | 5.44* | 5.21* |
| Ran | 1.03 | 3.7* | 4.21* |
| Abs-Max | -1.31 | -1.22 | 0.27 |
| Abs-Min | 6.15* | 3.93* | 6.15* |

*Null hypothesis is rejected at a 1% Significant.

Figure 12 shows that the model proposed consistently outperform the Crystal Ball. The mean of the VaR values produced by the proposed model were found to be significantly lower than the VaR values produced by the CB at a 1% level. The t-value was 12.73 and the z-value was 6.15. The 50 portfolios used in this experiment were created using the traditional matrix to make a fair comparison with the CB which uses the traditional covariance for the most part.

Figure 12
VaR time-series of the Proposed Model vs. Crystal Ball

Assuming an initial investment of \$1000, VaR for 50 portfolios constructed using the max selection compared to the VaR for 50 portfolios constructed using Crystal Ball. Covariance matrix is estimated using the traditional covariance matrix.



The model proposed outperformed the Russell 1000 and the S&P500 indices, the VaR for the Russell 1000 index calculated for the same time period is 66.1, and the VaR for the S&P500 is 66.2. The VaR for all the 50 portfolios constructed using the model proposed are significantly lower than both values.

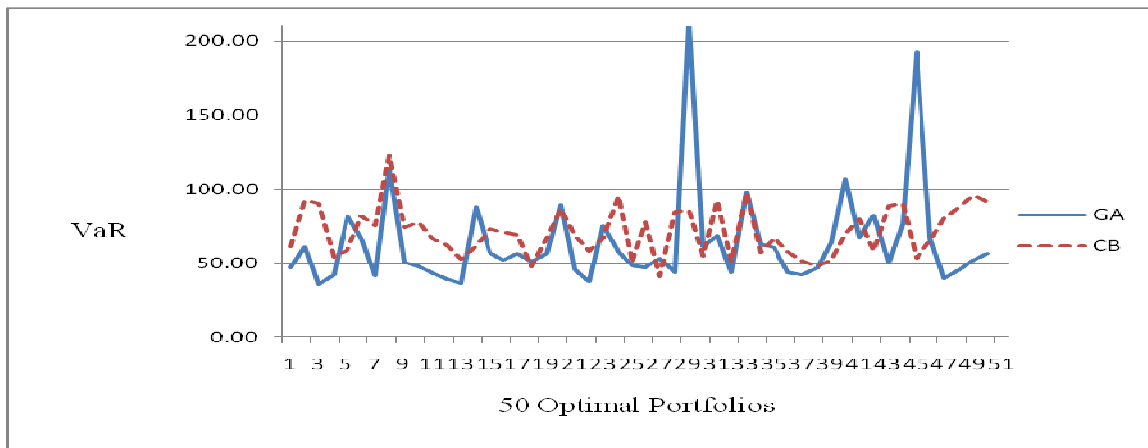
The question that needs to be answered now is whether the proposed model performs well when applied to a new data set – out-of-sample data. To answer this question the Russell 1000 data set, 850 securities for that period from January 2001 to July 2008, was divided into two parts. The first part, from January 2001 to April 2005,

was used for estimating the weights of the assets to be included in the optimal portfolios. The second part, from May 2005 to July 2008, was used to test the performance of these optimal portfolios over time. Figure 13 shows the VaR values for the proposed model as compared to the VaR values for the Crystal Ball. The mean of the VaR values produced by the proposed model are significantly lower than the VaR values produced by the CB at a 1% level with a z-value of 3.12. Using the t-test the mean of the proposed model's VaR is significantly lower at a 10% level with a t-value of 1.42.

Figure 13

Out-of-Sample VaR time-series of the Proposed Model vs. Crystal Ball

Assuming an initial investment of \$1000, VaR for 50 portfolios constructed using the max selection compared to the VaR for 50 portfolios constructed using Crystal Ball. Covariance matrix is estimated using the traditional covariance matrix.



Now that the model proposed outperforms the CB, the effect of the covariance estimation techniques within the proposed model can be addressed. Table 6 compares the max-selection model using the three different covariance estimation techniques. The max-selection using the OGARCH is found to significantly outperform the IDCC-GARCH and the traditional covariance, which shows that the usage of a conditional covariance matrix added further value to the portfolio construction.

Table 6
OGARCH Superiority over the other two Estimation Techniques
The table shows the t-values and the z-values for a one-tail hypothesis test.

| | Omax-Imax | Omax-Tmax | Imax-Tmax |
|---------------|------------------|------------------|------------------|
| T-test | 2.59* | 3.20* | 1.07 |
| Z-text | 3.06* | 3.35* | 0.62 |

*Null hypothesis is rejected at a 1% Significant.

Chapter 5

Conclusions

Prior research has established that idiosyncratic volatility of securities prices exhibits a positive trend. This trend and other factors have made the merits of investment diversification and portfolio construction more appealing. This dissertation proposes the use of a greedy algorithm in portfolio optimization to: a) increase the efficiency of the portfolio optimization process, b) implement large-scale optimizations, and c) improve the resulting optimal weights. This chapter starts with a summary of the results followed by their limitations. The potential future work is then discussed, followed by some concluding remarks.

4.3. Summary of the results

In this dissertation, a new methodology for optimizing the weights of equity portfolios using a greedy algorithm was proposed and implemented. Twenty variants were created using the same greedy algorithm; these variants included five different selection methods and four different estimation techniques. The selection methods selected the securities to be merged based on five criteria (on the pair-wise covariance values): the minimum, the maximum, the absolute minimum, the absolute maximum, and on a random basis. The estimation techniques were used to estimate the variance covariance matrices. These included: the traditional unconditional covariance matrix, the CCC-GARCH covariance matrix, the IDCC-GARCH covariance matrix, and the OGARCH covariance matrix. The first five variants used the traditional covariance to

estimate the variance and the covariance. The second five variants used the IDCC-GARCH. The third five variants used the OGARCH, and the last five variants used the CCC-GARCH estimation technique.

The empirical part of this dissertation was divided into two major stages. The first stage had the objective of determining whether the proposed twenty variants of the model were efficient and stable. The second stage compared the results to three different benchmarks to assess the performance. Two different sets of data were used in this process. In the first set of experiments, fifty securities were randomly selected from the S&P500 securities. These fifty securities were used to create simulated data sets which were then fed into the proposed optimizer, as well as into the benchmark optimizer to determine the optimal weights. A second set of experiments were performed with synthetically created data sets as explained below.

After the algorithms computed their best solutions, Sharpe ratio and VaR were calculated as criteria to compare the model's performance. Both criteria gave similar results, thus, only the average Sharpe ratios were reported. The objective of this run was to assess the efficiency and the stability of the proposed models. The five variants of the model using the CCC-GARCH were found to be inefficient, and thus were not investigated further. All the other fifteen variants, using the three other estimation techniques, were found to be efficient and stable, so they were investigated further.

In the second set of experiments, a different data set was used. Fifty portfolios were created, each composed of fifty different securities selected randomly from the Russell 1000 securities. Securities were drawn using a uniform distribution to make the

selection of securities equally likely. No duplicate securities were allowed within each portfolio. These fifty portfolios were fed one at a time into the fifteen variants of the model as well as into the benchmark to determine the optimal weights. The optimal weights were then used to calculate the VaR, which was used as a criterion to assess performance. For each variant of the greedy algorithm, fifty VaR values were calculated, one for each optimal portfolio. The VaR calculated for the three variants using the max selection dominated all the other four variants in their corresponding estimation technique. Thus, only these three variants were investigated further.

The simple VaR of the Russell 1000 and the S&P500 indices, and the VaR values of 50 portfolios constructed using Crystal Ball were calculated as benchmarks for comparison. Given that the Crystal Ball uses the traditional covariance matrix in its calculations, only the variant using the traditional covariance matrix was used for comparison. The proposed model significantly outperforms these three benchmarks at a 1% significance level. The use of the conditional covariance matrix, especially the OGARCH, showed further improvement over the variant of the model using the traditional covariance.

4.4. Limitations

Although the results of the proposed model have shown significant improvement over the benchmarks used, more tests are needed to generalize the conclusions. The first limitation is the data period: only one time period, January 2001 – July 2008, was used, and to generalize the conclusions, different time periods need to be investigated. The second limitation is the data frequency, only monthly data were used. Different data

frequencies have their own pros and cons, but for the sake of validating our model, different frequencies need to be tested. The third limitation is the length of the data period: only an eight-year period was used in this dissertation, and the effect of the data period length was not tested. In addition to these three limitations, different lead-lag structure of the GARCH models could be investigated: only the GARCH (1, 1) was used. Experiments of other lead-lag structures could be useful.

4.5. Future work

The first stage in the future work is to validate the model using different data periods, different data frequencies, different data length, and different lead-lag structures. After these tests, different potential improvements need to be considered. These potential improvements include the following:

- a) Using the time-varying covariance matrix to construct a time-varying portfolio. The time-varying portfolio can then be used to determine the portfolio turnover rate.
- b) Integrating and testing more constraints. In the current variant, only the long-only and the full-investment constraints were included. Other constraints, as discussed in section 2.3, are still to be integrated and tested.
- c) Using the CVaR instead of the VaR to account for the three drawbacks of the VaR, as discussed in section 3.4.2.2.

- d) Using a more advanced model for the return expectations, like the implied return calculations proposed by Black-Litterman or the Wavelet for multiple data filtering and trend extrapolation.
- e) Using parallel computing to improve the speed of processing by using parallel for-loops and parallel-numeric algorithms.
- f) Integrating a data acquisition module to the model to import and manipulate the data needed to run the model.

4.6. Concluding remarks

Using fifty different data sets, fifty portfolios were constructed using the proposed model. The VaR of these portfolios significantly outperformed the VaR of the S&P500 and the Russell 1000 indices, and the VaR of the portfolios constructed using Crystal Ball at a 1% significance level.

The proposed model's superior performance, particularly the max variant, is consistent with modern portfolio literature, where the securities with lower correlation coefficients add more diversification benefits. Even though more testing is necessary to validate and generalize the results, the results show that the portfolios constructed using the model proposed are superior to the three benchmarks that were computed.

Even though this dissertation has some limitations that should be addressed, the model proposed provides an improved platform that could be perceived as a starting point for a new approach for portfolio construction, an approach characterized by its simplicity, efficiency, and effectiveness.

REFERENCES

- Acerbi, C., D. Tasche, 2002, On the Coherence of Expected Shortfall, *Journal of Banking and Finance* 26, 1487-1503.
- Aguilar, O., and M. West, 2000, Bayesian Dynamic Factor Models and Portfolio Allocation, *Journal of Business & Economic Statistics* 18, 338.
- Alexander, C., and A. Chibumba, 1997, Multivariate Orthogonal Factor GARCH, University of Sussex, Mimeo.
- Alexander, C., 2001, Orthogonal GARCH, *Mastering Risk* 2, 21–38.
- Alexander, G. J., and A. M. Baptista, 2002, Economic Implications of Using a Mean-VaR Model for Portfolio Selection: A Comparison with Mean-Variance Analysis, *Journal of Economic Dynamics and Control* 26, 1159-1193.
- . 2003, Portfolio Performance Evaluation Using Value at Risk, *The Journal of Portfolio Management* 29, 93–102.
- Anderson, T. W., and D. A. Darling, 1952, Asymptotic Theory of Certain Goodness of Fit Criteria Based on Stochastic Processes, *Ann.Math.Statist* 23, .
- Antia, M. J., and M. S. Fridson, 2008, The Standard & Poor's Guide to Selecting Stocks (a Review), *Book Reviews* 3, 1-1.
- Artzner, P., et al, 1999, Coherent Measures of Risk, *Mathematical Finance* 9, 203-228.
- Avramov, D., T. Chordia, and A. Goyal, 2006, Liquidity and Autocorrelations in Individual Stock Returns, *The Journal of Finance* 61, 2365-2394.
- Avramov, D., T. Chordia, G. Jostova, and A. Philipov, 2007, Momentum and Credit Rating, *The Journal of Finance* 62, 2503-2520.
- Bakshi, R., 2007, Transforming Markets in the 21st Century: Socially Responsible Investing as a Tool, *Futures* 39, 523.
- Barksdale, E., and W. Green, 1990, Performance is Useless in Selecting Managers, *Pensions and Investments* 16.
- Basu, S., 1977, Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis, *Journal of Finance* 32, 663–682.
- Bauwens, L., S. Laurent, and J. V. K. Rombouts, 2006, Multivariate GARCH Models: A Survey, *Journal of Applied Econometrics* 21, 79-109.

- Bawa, V. S., 1976, Admissible Portfolios for all Individuals, *Journal of Finance* 23, 1169-1183.
- . 1975, Optimal Rules for Ordering Uncertain Prospects, *Journal of Financial Economics* 2, 95-121.
- Beale, E. M. L., 1955, On Minizing A Convex Function Subject to Linear Inequalities, *Journal of the Royal Statistical Society. Series B (Methodological)* 17, 173-184.
- . 1959, On Quadratic Programming, *Naval Research Logistics Quarterly* 6, 227-243.
- Bello, Zakri Y., 2005, Socially Responsible Investing and Portfolio Diversification, *The Journal of Financial Research* 28, 41.
- Bidarkota, P. V., and B. V. Dupoyet, 2007, The Impact of Fat Tails on Equilibrium Rates of Return and Term Premia, *Journal of Economic Dynamics and Control* 31, 887-905.
- Bilbao, A., et al, 2007, On Constructing Expert Betas for Single-Index Model, *European Journal of Operational Research* 183, 827-847.
- Black, F., and R. Litterman, 1991, Asset Allocation: Combining Investor Views with Market Equilibrium, *Journal of Fixed Income* 1, 7-18.
- Bollerslev, T., R. Y. Chou, and K. F. Kroner, 1992, ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics* 52, 5-59.
- Bollerslev, T., R. F. Engle, and J. M. Wooldridge, 1988, A Capital Asset Pricing Model with Time-Varying Covariances, *The Journal of Political Economy* 96, 116.
- Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307-327.
- . 1990, Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model, *Review of Economics and Statistics* 72, 498-505.
- Boyd, S., and L. Vandenberghe, 2004, *Convex Optimization, Semidefinite Programming and Recent Applications*, COLT, Banff, Canada, July.
- Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger, 1989, Empirical Tests of the Consumption-Oriented CAPM, *Journal of Finance* 44, 231-262.
- Brown, R. G., and R. F. Meyer, 1961, The Fundamental Theorem of Exponential Smoothing, *Operations research* 9, 673-687.

- Burmeister, E., R. Roll, and S. A. Ross, 2003, Using Macroeconomic Factors to Control Portfolio Risk, BIRR Portfolio Analysis, Inc., www.birr.com/documen2.htm.
- Campbell, J. Y., Lettau M, Malkiel B G, Xu Y, 2001, Have Individual Stocks Become More Volatile? an Empirical Exploration of Idiosyncratic Risk, *The Journal of Finance* 56, 1-43.
- Campbell, J. Y., A. Lo, and C. MacKinlay, 1996, *The Econometrics of Financial Markets*, 1997, (Princeton University Press, Princeton, NJ).
- Campbell, J. Y., Lo AW, MacKinlay AC, 1998, *The Econometrics Of Financial Markets*, *Macroeconomic Dynamics* 2, 559-562.
- Campbell, J. Y., and R. J. Shiller, 2001, Valuation Ratios and the Long-Run Stock Market Outlook: An Update, NBER Working Paper.
- Chan, L. K. C., N. Jegadeesh, and J. Lakonishok, 1996a, Momentum Strategies, *Journal of Finance* 51, 1681-1713.
- Chang, T. J., N Meade, JE Beasley, YM Sharaiha, 2000, Heuristics for Cardinality Constrained Portfolio Optimisation, *Computers and Operations Research* 27, 1271-1302.
- Chiam, S. C., K. C. Tan, and A. Al Mamum, 2008, Evolutionary Multi-Objective Portfolio Optimization in Practical Context, *International Journal of Automation and Computing* 5, 67-80.
- Chib, S., F. Nardari, and N. Shephard, 2006, Analysis of High Dimensional Multivariate Stochastic Volatility Models, *Journal of Econometrics* 134, 341-371.
- Chow, G., et al, 1999, Optimal Portfolios in Good Times and Bad, *Financial Analysts Journal* 55, 65-74.
- Christodoulakis, G. A., and S. E. Satchell, 2002, Correlated ARCH (CorrARCH): Modelling the Time-Varying Conditional Correlation between Financial Asset Returns, *European Journal of Operational Research* 139, 351-370.
- Christodoulakis, G. A., 2007, Common Volatility and Correlation Clustering in Asset Returns, *European Journal of Operational Research* 182, 1263-1284.
- Clarke, R., H. de Silva, and S. Thorley, 2001, *Portfolio Constraints and the Fundamental Law of Active Management*, CFA publication .
- Claus, J., and J. Thomas, 2001, Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets, *The Journal of Finance* 56, 1629-1666.

- Cockshott, P., and G. Michaelson, 2007, Are there New Models of Computation? Reply to Wegner and Eberbach, *The Computer Journal* 50, 232.
- Connor, G., 1995, The Three Types of Factor Models: A Comparison of their Explanatory Power, *Financial Analysts Journal* 51, 42-46.
- Cormen, T. H., 2001. *Introduction to Algorithms* (MIT Press).
- Cornuejols, G., and R. Tutuncu, 2006. *Optimization Methods in Finance* (Cambridge University Press).
- DeBondt, W. F. M., and R. Thaler, 1985, Does the Stock Market Overreact, *Journal of Finance* 40, 793-805.
- Dhrymes, P. J., I. Friend, and N. B. Gultekin, 1984, A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory, *Journal of Finance* 39, 323–346.
- Ding, Z., and R. F. Engle, 2001, Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing, *Academia Economic Papers* 29, 157–184.
- Disatnik, D. J., and S. Benninga, 2007, Shrinking the Covariance Matrix, *Journal Of Portfolio Management* 33, 55.
- Doganoglu, T., C. Hartz, and S. Mittnik, 2007, Portfolio Optimization when Risk Factors are Conditionally Varying and Heavy Tailed, *Computational Economics* 29, 333-354.
- Dowd, K., 2000, Adjusting for Risk: An Improved Sharpe Ratio, *International Review of Economics and Finance* 9, 209-222.
- Duda, R. O., P. E. Hart, and D. G. Stork, 2000. *Pattern Classification* (Wiley-Interscience).
- Duque, J., and D. A. Paxson, 1999, Empirical Evidence on Volatility Estimators, Working Paper, University of Manchester.
- El-Hassan, N., and P. Kofman, 2003, Tracking Error and Active Portfolio Management. *Australian Journal of Management* 28, 183-208.
- Engle, R. F., and K. F. Kroner, 1995, Multivariate Simultaneous Generalized ARCH, *Econometric Theory* 11, 122-150.
- Engle, R. F., and S. Manganelli, 2004, CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics* 22, 367-382.

- Engle, R. F., and K. Sheppard, 2001, Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, NBER Working Paper.
- Engle, R. F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007.
- . 2002, Dynamic Conditional Correlation: A Simple Class of Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* 20, 339-350.
- Evans, J. L., and S. H. Archer, 1968, Diversification and the Reduction of Dispersion: An Empirical Analysis, *Journal of Finance* 23, 761-767.
- Fabozzi, F., P. Kolm, and D. Pachamanova, 2007, Robust Portfolio Optimization and Management.
- Fama, E. F., and K. French, 1993, Common Risk Factors in the Returns on Bonds and Stocks, *Journal of Financial Economics* 33, 3-53.
- Fama, E. F., and K. R. French, 1996, The CAPM is Wanted, Dead or Alive, *Journal of Finance* 51, 1947-1958.
- . 2004, The CAPM: Theory and Evidence, *Journal of Economic Perspectives* 18, 25-46.
- . 1992a, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427-465.
- . 2002, The Equity Risk Premium, *Journal of Finance* 57, 637-659.
- . 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance* 51, 55-84.
- Fama, E. F., 1972, Components of Investment Performance, *Journal of Finance* 27, 551-567.
- Feldstein, M. S., and J. Slemrod, 1980, Personal Taxation, Portfolio Choice, and the Effect of the Corporation Income Tax, *The Journal of Political Economy* 88, 854.
- Fisher, L., 1975, Using Modern Portfolio Theory to Maintain an Efficiently Diversified Portfolio, *Financial Analysts Journal* 31, 73-85.
- Focardi, S. M., 2004, Clustering Economic and Financial Time Series: Exploring the Existence of Stable Correlation Conditions, *Finance letters* 2, 1-9.

- Fogler, Peter O. D., and H Russell, 1978, Pension Portfolio Objective: Connecting the Loop, *Financial Management* (pre-1986) 7, 56.
- Frost, P. A., and J. E. Savarino, 1986, An Empirical Bayes Approach to Efficient Portfolio Selection, *Journal of Financial and Quantitative Analysis* 21, 293-305.
- Fung, W. K. H., and D. A. Hsieh, 1991, Empirical Analysis of Implied Volatility: Stocks, Bonds and Currencies, Manuscript, Duke University .
- Gardner Jr, E. S., 1985, Exponential Smoothing: The State of the Art, *Journal of Forecasting* 4, 1-28.
- Garman, M. B., and M. J. Klass, 1980, On the Estimation of Security Price Volatilities from Historical Data, *Journal of Business* 53, 67.
- Ghysels, E., A. C. Harvey, and E. Renault, 1996, Stochastic Volatility, *Handbook of Statistics* 14, 119-191.
- Gold, S. C., and P. Lebowitz, 1999, Computerized Stock Screening Rules for Portfolio Selection, *Financial Services Review* 8, 61-70.
- Gómez, J. P., and T. Sharma, 2006, Portfolio Delegation under Short-Selling Constraints, *Economic Theory* 28, 173-196.
- Gordon, M. J., 1962a. *The Investment, Financing, and Valuation of the Corporation* (Homewood, Ill.: RDIrwin).
- Grinblatt, M., S. Titman, and R. Wermers, 1995, Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior, *American Economic Review* 85, 1088-1105.
- Grootveld, H., and W. G. Hallerbach, 2004, Upgrading Value-at-Risk from Diagnostic Metric to Decision Variable: A Wise Thing to do? *Risk Measures for the 21*, 33-50.
- Grootveld, H., and W. Hallerbach, 1999, Variance Vs Downside Risk: Is there really that Much Difference? *European Journal of Operational Research* 114, 304-319.
- Guerre, E., and F. Jouneau, 1998, Geometric Versus Arithmetic Random Walk the Case of Trended Variables, *Journal of Statistical Planning and Inference* 68, 203-220.
- Hansson, B., and P. Hordahl, 1998, Testing the Conditional CAPM using Multivariate GARCH-M, *Applied Financial Economics* 8, 377-388.
- Harvey, A. C., E. Ruiz, and N. Shephard, 1994a, Multivariate Stochastic Variance Models, *Review of Economic Studies* 61, 247-264.

- Hentschel, L., 1995, All in the Family Nesting Symmetric and Asymmetric GARCH Models, *Journal of Financial Economics* 39, 71-104.
- Ibragimov, R., 2005, On Efficiency of Linear Estimators under Heavy-Tailedness, Harvard Institute of Economic Research, Harvard University 2085.
- Ingersoll, J. E., 1987. *Theory of Financial Decision Making* (Rowman & Littlefield Publishers).
- Israelson, C. L., 2005, A Refinement to the Sharpe Ratio and Information Ratio, *Journal of Asset Management* 5, 423-427.
- Jacquier, E., N. G. Polson, and P. E. Rossi, 1994, Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics* 12, 371-389.
- Jagannathan, R., and Z. Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, *Journal of Finance* 51, 3-53.
- James, W., and C. Stein, 1961, Estimation with Quadratic Loss, *Proc Fourth Berkeley Symposium Math Statist Prob* 1, 361-379.
- Jegadeesh, N., and S. Titman, 2001, Profitability of Momentum Strategies: An Evaluation of Alternative Explanations, *The Journal of Finance* 56, 699-720.
- Jegadeesh, N., 1990, Evidence of Predictable Behavior of Security Returns, *Journal of Finance* 45, 881-898.
- Jensen, M. C., 1967, The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance* 23, 389-416.
- Jin, H., H. Markowitz, and X. Yu Zhou, 2006, A Note on Semivariance, *An International Journal of Mathematics, Statistics and Financial Economics* 16, 53-61.
- Jobson, J. D., and B. Korkie, 1980, Estimation for Markowitz Efficient Portfolios, *Journal of the American Statistical Association* 75, 544-554.
- Jobson, J. D., and B. M. Korkie, 1981a, Performance Hypothesis Testing with the Sharpe and Treynor Measures, *Journal of Finance* 36, 889-908.
- Jobson, J. D., and B. Korkie, 1981b, Putting Markowitz Theory to Work, *Journal of Portfolio Management* 7, 70-74.
- Jorion, P., 2003, Portfolio Optimization with Tracking-Error Constraints, *Financial Analysts Journal* 59, 70-82.
- Kac, M., 1947, Random Walk and the Theory of Brownian Motion, *Amer.Math.Monthly* 54, 369-391.

- Kang, S. H., and S. M. Yoon, 2007, Value-at-Risk Analysis of the Long Memory Volatility Process: The Case of Individual Stock Returns, *The Business Review*.
- Karush, W., 1939, Minima of Functions of several Variables with Inequalities as Side Constraints, Department of Mathematics, University of Chicago.
- Kempf, Alexander, and Peer Osthoff, 2007, The Effect of Socially Responsible Investing on Portfolio Performance, *European Financial Management* 13, 908.
- Kingston, G., 1989, Theoretical Foundations of Constant Proportion Portfolio Insurance, *Economics Letters* 29, 345-347.
- Kirchler, M., and J. Huber, 2007, Fat Tails and Volatility Clustering in Experimental Asset Markets, *Journal of Economic Dynamics and Control* 31, 1844-1874.
- Kjeldsen, T. H., 2000, A Contextualized Historical Analysis of the Kuhn–Tucker Theorem in Nonlinear Programming: The Impact of World War II, *Historia Mathematica* 27, 331-361.
- Korajczyk, R. A., and R. Sadka, 2004, Are Momentum Profits Robust to Trading Costs? *The Journal of Finance* 59, 1039-1082.
- Kritzman, M., K. Lowry, and A. S. VAN Royen, 2001, Risk, Regimes, and Overconfidence, *Journal of Derivatives* 8, 32-42.
- Kuhn, H. W., and A. W. Tucker, 1951, Nonlinear Programming. J. Neyman, Ed, *Proc. Second Berkeley Sympos. on Math. Statist. Probab* 481–492.
- Ledoit, O., 1994, Portfolio Selection: Improved Covariance Matrix Estimation, Working paper, Sloan School of Management.
- Ledoit, Olivier, and Michael Wolf, 2004, Honey, I Shrunk the Sample Covariance Matrix, *Journal of Portfolio Management* 30, 110.
- Lehmann, B., 1990, Fads, Martingales, and Market Efficiency, *Quarterly Journal of Economics* 105, 1-28.
- Lewellen, J., and S. Nagel, 2006, The Conditional CAPM does Not Explain Asset-Pricing Anomalies, *Journal of Financial Economics* 82, 289-314.
- Li, Frank S., 1997, Time Deformation Modeling: Theory and Application.
- Lien, D., and X. Luo, 1994, Multiperiod Hedging in the Presence of Conditional Heteroscedasticity, *Journal of Futures Markets* 14, 927-955.
- Lintner, J., 1965, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics* 47, 13-37.

- Litterman, R., and K. Winkelmann, 1998, Estimating Covariance Matrices, Risk Management Series, Goldman Sachs 2.
- Lo, A. W., 2002, The Statistics of Sharpe Ratios, *Financial Analysts Journal* 58, 36–52.
- Lopes, H. F., and H. Migon, 2002, Comovements and Contagion in Emergent Markets: Stock Indexes Volatilities, *Case Studies in Bayesian Statistics* 6, 285–300.
- Markowitz, H. M., 1991, Foundations of Portfolio Selection, *Journal of Finance* 46, 469–477.
- Markowitz, H., 1987, *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. New York, B. Blackwell.
- . 1952, Portfolio Selection, *The Journal of Finance* 7, 77–91.
- Massey, F. J., 1951, The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association* 46, 68–78.
- Merton, R. C., 1980, On Estimating the Expected Return on the Market, *Journal of Financial Economics* 8, 323–361.
- Michaud, R. O., 1998. *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation* (Oxford University Press).
- Miffre, J., and G. Rallis, 2007, Momentum Strategies in Commodity Futures Markets, *Journal of Banking and Finance* 31, 1863–1886.
- Mittnik, S., S. Rachev, and E. Schwartz, 2002, Value-at-Risk and Asset Allocation with Stable Return Distributions, *Allgemeines Statistisches Archiv* 86, 53–67.
- Morck, R., B. Yeung, and W. Yu, 2000, The Information Content of Stock Markets: Why do Emerging Markets have Synchronous Stock Price Movements? *Journal of Financial Economics* 58, 215–260.
- Nawrocki, D., 1999, A Brief History of Downside Risk Measures, *Journal of Investing* 8, 9–25.
- Nielsen, L. T., and M. Vassalou, 2008, Sharpe Ratios and Alphas in Continuous Time (Digest Summary), *Journal of Financial and Quantitative Analysis* 39, 103–114.
- Nocedal, J., and S. J. Wright, 1999. *Numerical Optimization* Springer.
- Officer, R. R., 1973, The Variability of the Market Factor of the New York Stock Exchange, *Journal of Business* 46, 434.

- Pafka, S., M. Potters, and I. Kondor, 2004, Exponential Weighting and Random-Matrix-Theory-Based Filtering of Financial Covariance Matrices for Portfolio Optimization, Arxiv preprint cond-mat/0402573 .
- Parkinson, M., 1980, The Extreme Value Method for Estimating the Variance of the Rate of Return, *Journal of Business* 53, 61.
- PAstor, L., and V. Pietro, 2003, Stock Valuation and Learning about Profitability, *The Journal of Finance* 58, 1749-1790.
- Pflug, G., and A. A. Gaivoronski, 2005, Value-at-Risk in Portfolio Optimization: Properties and Computational Approach, *Journal of Risk* 7, 1-31.
- Philips, Thomas K., 2003, Estimating Expected Returns, *Journal of Investing* 12, 49-73.
- Pitt, M. K., and N. Shephard, 1999, Time Varying Covariances: A Factor Stochastic Volatility Approach (with Discussion), *Bayesian Statistics* 6, 547–570.
- Platt, H. D., 2006, Revisiting the Reversal of Large Stock-Price Declines, *The Journal of Alternative Investments*, 45-67.
- Quintana, J. M., and M. West, 1987, An Analysis of International Exchange Rates using Multivariate DLM's. *Statistician* 36, 275-282.
- Ramaswami, S. N., K. Srivastava R., and McInish T. H., 1992, An Exploratory Study of Portfolio Objectives and Asset Holdings, *Journal of Economic Behavior & Organization* 19, 285-301.
- Rau-Bredow, H., and L. W. Str, 2004, Value at Risk, Expected Shortfall, and Marginal Risk Contribution, *Risk Measures for the 21st Century*, Wiley Finance, Chichester 61–68.
- Rebonato, R., 1999. *Volatility and Correlation in the Pricing of Equity, FX and Interest-Rate Options* (Wiley Chichester).
- RiskMetrics, T. M., Technical Document, JP Morgan/Reuters, New York, December 1996.
- Roll, R., and S. A. Ross, 1980, An Empirical Investigation of the Arbitrage Pricing Theory, *Journal of Finance* 35, 1073-1103.
- Roll, R., 1992, A Mean/Variance Analysis of Tracking Error, *Journal of Portfolio Management* 18, 13-22.
- Ross, S. A., 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory* 13, 341-360.

- Roy, A. D., 1952, Safety First and the Holding of Assets, *Econometrica* 20, 431–449.
- Ruszczyski, A., and A. Shapiro, 2006, Optimization of Risk Measures, Probabilistic and Randomized Methods for Design Under Uncertainty.
- Saltenis, V., 2006, Data Clustering Based on Maximization of Outlier Factor, *Journal of Global Optimization* 35, 625-635.
- Scherer, B., 2002, Portfolio Resampling: Review and Critique, *Financial Analysts Journal* 58, 98–109.
- Schreiner, J., 1980, Portfolio Revision, A Turnover Constrained Approach, *Financial Management* 9, 67-75.
- Scowcroft, A., and J. Sefton, 2005, Understanding Momentum, *Financial Analysts Journal* 61, 64-82.
- Shanken, J., 1982, The Arbitrage Pricing Theory: Is it Testable? *Journal of Finance* 37, 1129–40.
- . 1985, Multivariate Tests of the Zero-Beta CAPM, *Journal of Financial Economics* 14, 327–348.
- Sharpe, W. F., 1964, Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk, *Journal of Finance* 19, 425-442.
- . 1974, Imputing Expected Security Returns from Portfolio Composition, *Journal of Financial and Quantitative Analysis* 9, 463-472.
- . 1966, Mutual Fund Performance, *Journal of Business* 39, 119.
- . 1998, The Sharpe Ratio, *Streetwise: The Best of the Journal of Portfolio Management*.
- Social Investment Forum: Socially Responsible Investing Basics for Individuals, 2008.
- Stein, C., 1955, Inadmissibility of the Usual Estimator of the Mean of a Multivariate Normal Distribution, *Third Berkeley Symposium Math Statist Prob* 1, 361-379.
- Steinbach, M. C., 2001, Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis, *SIAM Review* 43, 31-85.
- Stock, J. H., and M. W. Watson, 2005, Implications of Dynamic Factor Models for VAR Analysis, NBER Working Paper.
- Stone, B. K., 1973, A Linear Programming Formulation of the General Portfolio Selection Problem, *Journal of Financial and Quantitative Analysis* 8, 621-636.

- Strong, R. A., 1999. Portfolio Construction, Management, and Protection (South-Western Educational Publishing).
- Strong, Robert A., 1988, A Behavioral Investigation of Three Paradigms in Finance, Northeast Journal of Business & Economics 14, 1.
- Tobin, J., 1958, Liquidity Preference as Behavior Towards Risk, Review of Economic Studies 25, 65-86.
- Treynor, J. L., and F. Black, 1973, How to use Security Analysis to Improve Portfolio Selection, Journal of Business 46, 66.
- Treynor, J. L., 1965, How to Rate Management of Investment Funds, Harvard business review 43, 63-75.
- . 1961, Market Value, Time, and Risk, Unpublished Manuscript dated August 8, 1961.
- Tse, Y. K., and A. Tsui, 2002, A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations, Journal of Business & Economic Statistics 20, 351-362.
- Turing, A. M., 1936, On Computable Numbers, with an Application to the Entscheidungsproblem, Proceedings of the London Mathematical Society 42, 230-265.
- Wei, S. X., and C. Zhang, 2006, Why did Individual Stocks Become More Volatile?, The Journal of Business 79, 259-292.
- Wilkins, K. A., J. L. Heck, and S. J. Cochran, 2006, The Effects of Mean Reversion on Alternative Investment Strategies, Managerial Finance 32, 14-38.
- Wolfe, P., 1959, The Simplex Method for Quadratic Programming, Econometrica 27, 382-398.
- Yamai, Y., and T. Yoshida, 2005, Value-at-Risk Versus Expected Shortfall: A Practical Perspective, Journal of Banking and Finance 29, 997-1015.

VITA
AHMED ELSHAHAT

| | |
|-----------------|---|
| January 5, 1978 | Born, Alexandria, Egypt |
| 1995 – 1999 | B.A. Accounting Alexandria University Alexandria, Egypt |
| 2000-2003 | MBA – Finance Arab Academy for Science and Tech, Alexandria, Egypt |
| 2004 | Certified Financial Manager (CFM) Institute of Management Accounting, New Jersey, NJ. |
| 2004 – 2008 | Doctoral candidate, Finance Florida International University, Miami, FL Adjunct Faculty, School of Accounting, Florida International University, Miami, FL Graduate Teaching Assistant, Finance dept, Florida International University, Miami, FL |
| 2005 | Most dedicated Teacher award, Beta Alpha Psi - FIU Miami, FL |

PUBLICATIONS

“Forecasting and determining Effect of FFR Announcements on Euro/Dollar Exchange using High Frequency Data” Co-Authored with A. Parhizgari, SWFA.

“Assessing the performance of the privatized Egyptian construction state-owned-enterprises” Master’s Thesis, Published in the proceedings of the AAST, 2003.

“Risk Assessment of the banking stocks after the FBAs announcements” Working paper.

PRESENTATIONS

| | |
|---|----------|
| 4 th International Conference on Banking and Finance, Malaysia | Jan 2008 |
| National Sun Yat-sen University, Kaohsiung, Taiwan. | Dec 2007 |
| Financial Management Association (FMA) conference, Orlando, FL | Oct 2007 |
| Arab Academy for Science and Tech (AAST), Alexandria | Jan 2007 |
| Financial Management Association (FMA) conference, SLC, Utah | Oct 2006 |
| South Western Finance Association (SWFA) conference, OK city | Mar 2006 |
| American finance Association (AFA) annual Conference, Boston | Jan 2006 |

PROFESSIONAL MEMBERSHIPS

Institute of Management Accounting (IMA)
Certified Financial Analyst (CFA) institute.
American Finance Association (AFA)
Financial Management Association (FMA)
Social Science Research Network (SSRN)

CONSULTING AND REVIEWS

Principles of Financial Management, April 2005, 1st edition, by David Kidwell and Robert Parrino, John Wiley & Sons. Paper Reviewer for the Latin American conference.

Ad-hoc referee for the *Journal of Business Research*

REFERENCES

| | | |
|--|--|---|
| Name: ALI PARHIZGARI | Name: WILLIAM WELCH | MANUEL DIEGUEZ |
| Title: PROF/ FINANCE | Title: CHAIR, FINANCE | Title: ASSO DIRECTOR |
| Office: RB 204B | Office: RB 208A | Office: RB 237A |
| Phone: 305-348-3326 | Phone: 305-348-4202 | Phone: 305-348-3263 |
| E-mail: parhiz@fiu.edu | E-mail: welchw@fiu.edu | Email: dieguezm@fiu.edu |