DYNAMIC PORTFOLIO OPTIMIZATION USING MEAN-SEMIVARIANCE

A Dissertation
Presented to
The Academic Faculty

By

Carl Morris

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Operations Research

Georgia Institute of Technology

December, 2017

DYNAMIC PORTFOLIO OPTIMIZATION USING MEAN-SEMIVARIANCE

Hayriye Ayhan
School of Industrial & Systems Engineering
Georgia Institute of Technology

Shijie Deng
School of Industrial & Systems Engineering
Georgia Institute of Technology

Sebastian Pokutta
School of Industrial & Systems Engineering
Georgia Institute of Technology

Chelsea C. White III
School of Industrial & Systems Engineering
Georgia Institute of Technology

Jun Xu
School of Computer Science
Georgia Institute of Technology

Date Approved: November 2, 2017

Our greatest glory is, not in never falling, but in rising every time we fall.

Oliver Goldsmith, The citizen of the world

To Si, Evelyn, and my family, with whose support this thesis is possible.

ACKNOWLEDGEMENTS

I want to thank my committee for their assistance and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

VaR	Value at risk
CVaR	Conditional value at risk
CAPM	Capital asset pricing model
LPM	Lower partial moment
KKT	Karush-Kuhn-Tucker
r	Asset returns
p	Asset excess returns
r_0	Risk free asset return
u	Portfolio weights
x	Wealth
Ω	Sample space of random variables.
n	Number of assets
h	Semivariance wealth target
η	Semivariance excess wealth target
λ	Semivariance activity indicator
U	Optimal portfolio map
Λ	Optimal portfolio semivariance indicator map
$\mathrm{E}(\xi)$	Expectation of random variable ξ
$\mathbb{P}(\Xi)$	Probability of random event Ξ
$\mathrm{V}(\xi)$	Variance of random variable ξ
$x(i)$	Element i of vector x
x'	Transpose of vector or matrix x
$[p]$	Binary indicator function of boolean predicate p
$\mathrm{ri}\,X$	Relative interior of a set X
$\mathrm{co}(x_1, x_2, \ldots, x_m)$	Convex hull of vector space elements $\{x_1, x_2, \ldots, x_m\}$
$(x)_+$	Nonnegative part of x
A^{-1}	Inverse of matrix A
$\ker A$	Kernel or null space of linear mapping A
X^∞	Recession cone of set X
$\|x\|$	Norm of vector x
$\langle x, y\rangle$	Inner product of vectors x and y
$A \succeq 0$	Matrix A is positive semidefinite
$\mathbb{S}$	Set of symmetric matrices

SUMMARY

This dissertation studies the mean-semivariance portfolio optimization problem. We describe the relationship of this kind of optimization in the context of other types of portfolio optimization. We construct a novel analysis of mean-semivariance in the context of piecewise quadratic optimization. The unique structure of mean-semivariance is leveraged to provide insight into properties of the optimal portfolio as a function of its key input parameters. This characterization allows us to introduce a new approach to solving a multi-period dynamic mean-semivariance portfolio problem. The proposed methodology provides significant improvements over naive approaches not leveraging the unique structure of the mean-semivariance value function. Finally, we develop a novel, distributionally robust piecewise quadratic formulation using semidefinite programming. We apply the robust formulation to the mean-semivariance portfolio problem to construct a distributionally robust mean-semivariance portfolio. We prove that the robust mean-semivariance portfolio is actually equivalent to the classical mean-variance portfolio.

CHAPTER 1

1.1 Introduction and Background

The goal of portfolio optimization is to allocate wealth among several assets to optimally achieve a goal. One of the most common objectives is to maximize mean return while controlling risk. The origin of this methodology is the mean-variance criterion introduced by Markowitz [2, 3]. Markowitz also proposed mean-semivariance because it only penalizes performance below a given target. While mean return is well accepted as a measure of portfolio returns, the choice of risk measure is not so clearly defined. The most common alternatives are Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), and mixed CVaR.

None of these measures clearly dominates the others in the academic literature, so there is a more or less parallel development of portfolio optimization utilizing each of these measures combined with the mean. Outside the mean-risk framework, some authors choose various utility functions to represent investor preferences. Alternatively, investors wishing to maximize portfolio growth optimize the geometric mean, or equivalently the log wealth. Each of these methods has benefits and disadvantages toward its use in portfolio optimization.

The mean-risk model, where the investor maximizes mean return while minimizing risk, remains popular today. Practitioners, like Michaud [4], continue to publish white papers advocating the mean-variance criterion. Mean-variance models, however, are frequently criticized in academic literature as being inappropriate, for example by Estrada, Hogan and Warren, Mao, Nantell and Price, Nawrocki, and Porter [5, 6, 7, 8, 9, 10]. This thesis provides a framework and theory for constructing mean-semivariance portfolios in single-period and multi-period formulations with extensions to accommodate distributional robustness.

Our contribution spans four areas:

1. Properties of the single-period and robust mean-semivariance portfolio are analyzed.

We also contrast the single-period and multi-period portfolios.

2. Algorithmic improvements are provided for the efficient computation of the single-period efficient frontier and the multi-period portfolio.

3. Methodology is developed for deriving deterministic counterparts for piecewise quadratic optimization over moment-based ambiguity sets.

4. Analytical solutions are derived for the robust mean-semivariance portfolio problem over moment-based ambiguity sets.

Single-period mean-semivariance problems are the most well studied and have been studied by Cumova and Nawrocki, Hogan and Warren, Markowitz, and, Markowitz, Todd, Xu, and Yamane [3, 11, 12, 13]. We develop a method based on a nonsmooth version of Newton's method by Sun [14]. Multi-period models are not well studied, and their research is limited, with more results for continuous-time by Jin, Yan, and Zhou, Jin, Markowitz, and Zhou, and Lari-Lavassani and Li [15, 16, 17] than discrete-time, which has be studied by Yan, Miao, and Li [18]. Continuous-time portfolios are generally restricted to those that are continuous with respect to time, narrowing the solution space of portfolio functions. Discrete-time portfolio functions are especially difficult to construct because there is no notion of continuity to restrict the space of optimal portfolio functions. We propose the development of a tractable, exact algorithm for multi-period mean-semivariance portfolio problems in discrete-time, which is a significant contribution of this thesis. Such an algorithm does not exist in the literature, and it would permit the calculation of both optimal portfolios for practitioners, as well as benchmark portfolios for researchers developing mean-semivariance heuristics.

Sensitivity analysis of mean-semvariance portfolios in the literature is limited to observations about the efficient frontier's convexity. This thesis therefore pursues several results in this area. Parametric quadratic programming has been successfully applied to the mean-variance portfolio problem by Hirschberger, Qi, and Steuer, Markowitz, Niedermayer and Niedermayer, Steuer, Qi, and Hirschberger [3, 13, 19, 20, 21]. Recent research by Patrinos and Sarimveis [22] facilitate development of parametric piecewise quadratic programming algorithms to calculate mean-semivariance efficient frontiers. Furthermore, these same methods apply to calculating the multi-period value functions. We present parametric formulations of the value functions for the multi-period problem. We use this formulation to provide an algorithm for computing multi-period portfolios via a nonsmooth Newton's method that is more efficient than a naive multi-period stochastic programming formulation. This representation also forms a basis for future work to construct more efficient algorithms for the multi-period mean-semivariance portfolio problem.

Finally, DeMiguel and Nogales, Garlappi, Uppal, and Wang, Ghaoui, Oks, and Oustry, Goldfarb and Iyengar, Harvey, Liechty, and Liechty [23, 24, 25, 26, 27] all conclude that mean-variance portfolios suffer from poor out of sample performance. A priori there is no reason to believe mean-semivariance portfolios are any less sensitive, and our preliminary results indicate the existence of such performance issues. The academic literature indicates one of three approaches are typically used to overcome these performance issues: resampling, robust optimization, and Bayesian methods. We provide a novel robust formulation for the mean-semivariance portfolio problem to address these concerns. We prove that the robust formulation is equivalent to the mean-variance portfolio under reasonable conditions. This equivalence demonstrates a novel property of the mean-variance portfolio.

1.2 Importance of the Mean-Semivariance Criterion

The mean-variance criterion for efficiency, originally introduced by Markowitz [3], continues to be used today for several reasons. First, it is simple to use, intuitively appealing, and Levy and Markowitz, and Markowitz [28, 29] both found it generally correlates well with other utility functions. Beginning with Sharpe's [30] paper establishing the basis for the capital asset pricing model (CAPM), the mean-variance model became justified as consistent with economic theory of market equilibria. A measure called beta, derived from CAPM theory, captures an asset's relative risk contribution, and its continued usage demonstrates the importance of mean-variance to investors.

Despite its strong history and established place in investment theory, even Markowitz [3] was doubtful of variance as a suitable measure of risk, so he suggested semivariance as an alternative. Estrada, Hogan and Warren, Mao, Nantell and Price, Nawrocki, and Porter [5, 6, 7, 8, 9, 10] all have raised similar concerns about variance and have often suggested semivariance as an alternative because it is conceptually very similar to variance but in a sense more rational. Mean-semivariance also has a strong utility function interpretation via stochastic dominance. Porter [10] showed that mean-semivariance efficiency is a sufficient condition for a return distribution to be non-dominated in the sense of second order stochastic dominance. Given a family of return distributions, second order stochastic dominance means that the cumulative distribution function is not dominated by any other cumulative distribution function in the family. Following a reasoning similar to the development of CAPM, Estrada, Hogan and Warren, and Nantell and Price [5, 6, 8] develop asset pricing models using semivariance. These asset pricing models lead to measures generally called downside-beta. Together these research results suggest that mean-semivariance provides a rational model without increasing computational complexity unreasonably.

Semivariance is also compared to two other families of risk measures: lower-partial-moments (LPM) and CVaR. Semivariance is a special case of LPM. Cumova and Nawrocki [31] observe that solution algorithms developed for semivariance can often apply to LPM portfolio problems. Based on this observation, we hypothesize that our results for mean-semivariance portfolios have the potential to be generalized to the mean-LPM portfolio problem. Hence, we do not see mean-semivariance research as competing with LPM models. We see our research as a contribution to LPM research.

Comparisons between CVaR and mean-semivariance are not common in literature, but several differences are immediately apparent. First, CVaR has the beneficial property of being a coherent risk measure. Furthermore, it admits a linear programming representation

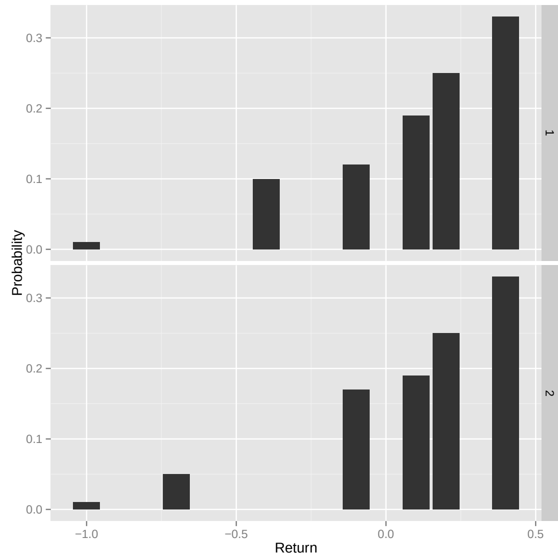

Figure 1.1: Two distributions are compared that have the same mean return. Scenario 1 has a smaller semivariance than scenario 2, but both distributions have the same CVaR. The key observation is that the low probability large negative return in Scenario 2 increases the semivariance significantly compared to the higher probability but less negative returns.

that makes optimizing it tractable. It provides a bound on VaR. When interpreted in terms of utility maximization, CVaR can be interpreted as a piecewise linear function with steeper slopes below a given target, i.e. loss averse. In this way it is very similar to Mean-semivariance which is also loss-averse, but mean-semivariance is increasingly loss averse for larger losses because it is quadratic below the target. According to this interpretation, mean-semivariance will prefer a high probability of a small loss to a small probability of a large loss even when CVaR would be indifferent to the two outcomes. See Figure 1.1 for an illustration of such a distribution.

1.3 Contribution

We developed and implemented a Newton's Method based algorithm for the single-period mean-semivariance portfolio problem. Our algorithm is comparable to existing parametric and active set approaches developed by Cumova and Nawrocki, Hogan and Warren, Markowitz, and Markowitz et al. [3, 11, 12, 13]. Our algorithm's key advantage is that it inherits the rapid convergence of Newton's method when we have a suitable choice of portfolio to initialize the algorithm. The approach results in an optimal solution with an analytical structure that permits analysis of the solution properties.

We present several sensitivity analysis results that are not discussed elsewhere in the literature. The results span three areas: critical regions, solution mapping, and transaction costs. The characterization of the critical regions and solution mapping provide insight into the behavior of the portfolio as the investor varies model parameters. The transaction costs analysis imparts a deeper understanding of which factors impact the decision to re-balance portfolios.

For the first time in the literature, we present parametric formulas for the solution mapping, efficient frontier, and value functions for the single-period mean-semivariance problem using methods similar to those of Hirschberger et al., Markowitz, Markowitz et al., Niedermayer and Neidermayer, Patrinos and Sarimveis, and Steuer et al. [3, 13, 19, 20, 21, 22]. The formulas facilitate analyses of optimal mean-semivariance portfolios and associated value functions. This algorithmic improvement provides an efficient means of computing and storing portfolio results for practitioners who need to calculate efficient frontiers or run numerous scenarios.

We provide a multi-period mean-semivariance portfolio algorithm based on the work of Lau and Womersley [32]. Such an algorithm is the first of its kind in the mean-semivariance portfolio literature. Our algorithm is the first exact algorithm for the multi-period mean-semivariance portfolio problem in discrete-time. Jin et al. [16] showed that no such algorithm could exist in the continuous-time case. Following the approach of Niedermayer and Niedermayer [20], we utilize parametric solutions for piecewise quadratic functions to make the algorithm tractable compared to a naive implementation using multi-period stochastic programming. Our algorithm improves the run time allowing larger instances to be solved in reasonable amounts of time.

Research by Fliege and Werner, Goldfarb and Iyengar, and Tütüncü and Koenig [26, 33, 34] in robust portfolio optimization indicates that most models focus on mean-variance optimization. We develop a robust formulation of the mean-semivariance portfolio problem based on moment constrained ambiguity sets. Using conic duality, we provide a tractable deterministic counterpart that is a semidefinite program. We provide conditions under which the optimal solution to this semidefinite program is equivalent to the mean-variance portfolio. Finally we extend the formulation to incorporate support constraints on the ambiguity set. Our work is related to a similar thread of a research focusing on the value-at-risk and log-optimal portfolio problems as studied by Rujeerapaiboon, Kuhn, and Weisemann [35] and Zymler, Kuhn, and Rustem [36]. To the best of our knowledge this is the first work applying robust optimization to the mean-semivariance portfolio problem. Our methodology allows us to analytically solve the robust mean-semivariance portfolio problem, which simplifies comparisons with the non-robust portfolios.

CHAPTER 2

2.1 Literature Review

Our approach to mean-semivariance portfolio optimization spans three topics: properties of the portfolio problem itself, extension to dynamic multi-period portfolios, and robust reformulations. First, we discuss the history of mean-semivariance portfolio optimization in the context of Markowitz's [2] seminal paper on the mean-variance portfolio. We show that our methodology fills a void in the sensitivity analysis and efficient frontier analysis of existing literature. Second, we discuss methods for the dynamic multi-period mean-variance portfolio model. Our work demonstrates that similar results are achievable by adapting existing literature on piecewise quadratic optimization. Third, we explore recent work in distributionally robust optimization. The yet untouched robust semivariance model is shown in our work to have an analytical solution which aligns with the mean-variance and second-moment portfolio problems.

The natural starting point for studying mean risk models for portfolio optimization begins with Markowitz's [2] seminal paper. He suggests both variance and semivariance as candidate risk measures. Variance is familiar to most investors, and he develops extensive theory around mean-variance portfolios in his later book [3]. Semivariance at the time was computationally difficult, so he did not pursue the mean-semivariance portfolio in the same depth. In his book he describes the Critical Line Algorithm for generating mean-variance efficient frontiers, which is applicable to mean-semivariance portfolio construction. Hogan and Warren [12] provide one of the earliest computational studies of the mean-semivariance portfolio, and they establish the basic properties of the mean-semivariance objective function. They base their numerical algorithm on the work of Frank and Wolfe [37].

The literature shows many different arguments for when semivariance is an appropriate risk-measure. In the case of symmetric return distributions, variance is equivalent to the mean-centered semivariance. In the asymmetric case, justification for semivariance was first cited by Markowitz [3] as intuitive for investors who do not wish to penalize exceptional returns. Using the concepts of admissibility and stochastic dominance, Quirk and Saposnik [38] show that efficiency in the mean-variance sense may contradict preferences defined in terms of stochastic dominance. Porter [10] expands on this idea by showing that (except in special cases) mean-semivariance efficiency implies second order stochastic-dominance efficiency. Mao [7] compares mean-variance and mean-semivariance in a capital budgeting context, arguing that semivariance is more consistent with utility functions elicited from executives. Levy and Markowitz [28] show that mean-variance performance is strongly correlated with different utility functions, which is a precursor to Markowitz's later work [29] on the same topic. Estrada [39] showed similar correlation results for mean-semivariance.

An economic approach to this discussion is generally founded on the capital asset pricing model (CAPM) developed by Sharpe [30]. His framework is based on Markowitz's portfolio theory, and the concept remains popular to this day. Hogan and Warren [6] develop a model analogous to Sharpe's [30] when the semivariance benchmark rate is equal to the risk-free rate. The analysis results in an asymmetric "cosemivariance" measure used to calculate the risk premium of an asset. Nantell and Price [8] show that under the assumption of a bivariate normal distribution between individual asset returns and the returns of the market portfolio that the beta calculated using CAPM is the same as the beta calculated using the semivariance CAPM due to Hogan and Warren [6]. Estrada [39] develops a behavioral model that assumes investors optimize mean-semivariance and justifies it using the same rational that Levy and Markowitz [28] used for mean-variance behavioral models. Motivated by these arguments, Estrada [5] develops a downside beta analogous to the CAPM beta. Estrada's [5] beta differs from the downside beta of Hogan and Warren [6] both in the benchmark rate and the construction of the semivariance analogue to covariance. Estrada [5] argues that downside beta consistently outperforms CAPM beta in emerging markets.

A recent generalization of semivariance was developed by Cumova and Nawrocki [31]. Their work studies the family of upper potential and downside risk measures that includes mean-semivariance as a special case. They suggest that the upper potential and downside measures can approximate many different practical utility functions. They provide a framework for generating approximate optimization problems using exogenous moment matrices developed in their previous work [11, 40, 41]. Their methods are typically generalizations of prior work based on semivariance, so we view extending our work to this family of risk measures as a possible avenue of future research.

Given an investor is interested in using semivariance as a risk measure, there is relatively limited literature on appropriate optimization methodology. Markowitz et al. [13] developed an alternative implementation of the Critical Line Algorithm for mean-semivariance portfolios by introducing an artificial variable for each sample path of the excess returns. To our knowledge, the approach developed by Markowitz et al. [13] is the only published approach for exact mean-semivariance portfolio optimization. Niedermayer and Niedermayer [20] modified the popular Critical Line Algorithm to increase its efficiency. Steuer et al. [21] develop a computationally efficient way to calculate the exact mean-variance efficient frontier by treating it as a multi-criteria optimization problem to which they apply a parametric quadratic programming algorithm, described in detail in their later paper [19]. Taking a similar approach, we treat mean-semivariance as a piecewise-quadratic function. We contribute a novel algorithm for the mean-semivariance efficient frontier based on a parametric piecewise-quadratic algorithm following the work of Sun and Patrinos and Sarimveis [14, 22]. Our approach provides straight forward tools for sensitivity analysis not readily apparent in existing mean-semivariance solution algorithms.

2.2 Efficient Frontiers and Multi-Period Portfolios

The other important use of the single-period portfolio optimization algorithms like the Critical Line Algorithm developed by Markowitz [3] is computing efficient frontiers. The goal of calculating efficient frontiers is to enumerate optimal portfolios and their trade-off between mean return and risk. More generally, the Critical Line Algorithm is a type of parametric optimization. Our work applies the general framework of piecewise quadratic parametric optimization developed by Patrinos and Sarimveis [22]. We establish efficient formulas for mean-semivariance efficient frontiers. More generally, we extend the parametric approach to an explicit multi-period mean-semivariance portfolio optimization problem. To our knowledge this is the first work giving an exact solution to the multi-period mean-semivariance formulation.

The simplest approach for enumerating the efficient frontier is to compute it at discrete points on a grid as done by Hogan and Warren [12]. Markowitz et al. [13] extend their Critical Line Algorithm to efficiently enumerate the mean-semivariance efficient frontier. Their formulation, however, introduces two artificial variables for each sample realization to reformulate the piecewise quadratic as a quadratic. The addition of artificial variables obscures the underlying analytical structure of the efficient frontier which we leverage. Our methodology for computing the efficient frontier is an extension of the same parametric approach used for our sensitivity analysis results.

The parametric approach we use also permits us to explore multi-period portfolio optimization problems. Multi-period models benefit investors with long term return goals by explicitly accounting for the ability to re-balance a portfolio dynamically. In the mean-variance case, an analytical solution is possible, which is provided by Li and Ng [42]. Similarly, in the absolute deviation case, Yu, Takahashi, Inoue, and Wang [43] provide an analytical solution. Li, Cui, Wang, and Zhu [44] extend the mean-variance work of Li and Ng [42] to account for time-inconsistency of mean-variance. Najafi and Mushakhian [45] adopt a heuristic approach for dynamic mean-semivariance-CVaR portfolios with transaction costs. In the continuous time Black-Scholes setting, Lari-Lavasann and Li [17] show that a nearly analytical solution is possible, reducing the problem to a one-dimensional optimization. We show that it is possible to analytically enumerate the value function of the multi-period semivariance portfolio at each period, leading to an efficient algorithm for solving the dynamic mean-semivariance problem. The convergence and tractability of such a formulation stems directly from the piecewise quadratic structure, with proofs given by Lau and Womersley [32].

2.3 Robust Portfolio Optimization

The reliance of variance optimization on the covariance matrix and the semivariance on second moment type matrices cause both methods to be sensitive to estimation errors. One approach is re-sampling as proposed by Michaud [4]. We instead look to robust optimization. In the context of mean-variance optimization, robust formulations typically involve uncertainty sets on the mean vector and covariance matrix. Halldórsson and Tütüncü [46] provide a classic example of this approach. Because the value of semivariance endogenously depends on the portfolio, we instead consider alternatives for modeling uncertainty over the full distribution of returns.

Several frameworks for distributionally robust optimization models are being developed in parallel. Jiang and Guan [47] study the case of ambiguity sets defined by ϕ-divergence measures. Kleywegt and Gao [48] study the Wasserstein distance. These ambiguity sets, however, do not immediately lend themselves to portfolio optimization. The classical moment problem as given by Shapiro [49] is often seen in applications to portfolio optimization. Jiang and Guan [50] in a working paper applied the moment based ambiguity set to a portfolio problem restricting the probability of not reaching a certain return threshold. Rujeerapaiboon, Kuhn, and Wiesemann [35] study worst-case growth optimal portfolios under moment-based ambiguity sets. Ling et al. [51] looked at moment-based ambiguity sets as a part of their overall robust model. Our robust formulation generalizes the approach of Jiang and Guan and Zymler, Kuhn, and Rustem [36, 50] to a robust piecewise quadratic objective under moment-based ambiguity sets. The formulation permits an analytical solution, which we use to show robust semivariance is equivalent to variance and second-moments. This result is analogous to that of Rujeerapaiboon et al. [35], and may also provide an alternative proof to their result.

A common argument for using mean-variance portfolios is based on the economic interpretation due to Markowitz [3]. Alternative interpretations are due to utility function approaches originating from Neumann and Morgenstern [52]. A recent paper by Markowitz [29] revisits the mean-variance portfolio as an approximation to various utility functions, including the log-wealth-utility. This idea is similar to active research by Gotoh, Kim, and Lim [53] showing ϕ-divergence based distributionally robust optimization problems are well-approximated by mean-variance optimization problems. Our results show that the mean-variance portfolio is often equivalent to the robust semivariance portfolio.

The standard semivariance formulation presupposes a static return benchmark. The natural generalization is to use another portfolio's return as the benchmark. Ling, Sun, and Yang [51] study tracking error portfolios designed to follow specific benchmark portfolios, such as indexes. They develop a robust optimization approach for two downside risk measures: probability of loss and expected loss. Their approach is based on a semidefinite program derived from a duality result due to Shapiro [54]. This duality approach parallels our work from the perspective of distributionally robust optimization. Our approach generalizes to piecewise quadratic risk measures, and for the case of semivariance, we showed an equivalence to variance or second-moment formulations. The equivalence in particular reduces the solution of the robust semivariance tracking error portfolio to a variance or second-moment tracking error portfolio.

CHAPTER 3

3.1 Preface on Notation

The elements of a vector $x \in \mathbb{R}^n$ are denoted by $x(i)$ for $i = 1, 2, \ldots, n$. We denote by $[p]$ the indicator function of the boolean predicate p, which is equal to 1 whenever the predicate holds and is equal to 0 otherwise. We denote the relative interior of a set X by $\operatorname{ri} X$. We denote the convex hull of a collection of elements $x_1, x_2, \ldots, x_m$ by $\operatorname{co}(x_1, x_2, \ldots, x_m)$.

3.2 Single-Period Mean-Semivariance Portfolio Problem

Let r denote an $n+1$ dimensional random vector of rates of return for $n+1$ risky assets. Let r_0 denote rate of return for asset $n+1$. For $i < n+1$, let the n dimensional vector p be defined by $p(i) = r(i) - r_0$, so that p denotes the excess rates of return for the first n risky assets relative to the reference rate r_0. In what follows we often assume that r_0 is a deterministic risk free rate. Suppose an investor has x units of wealth to invest among the $n+1$ assets. We denote by $u(i)$ the amount invested in asset $i = 1, 2, \ldots, n$. Any remaining wealth is invested in the reference asset, i.e. the amount invested in the reference asset is $x - \sum_{i=1}^{n} u(i)$. A negative investment denotes holding a short position in the asset. By setting $x = 1$ we arrive at a normalized portfolio denoting the fraction of wealth invested in each asset. Given a particular portfolio u the amount of wealth at the end of the period is denoted by x_T where the subscript T is to distinguish it from the initial wealth x. The two are connected by the following relationship: $x_T = r_0 x + p'u$.

To facilitate exact computation of optimal portfolios, we consider a finite discrete return distribution; i.e. the excess returns and reference return take a finite number of values $p(\omega)$ and $r_0(\omega)$ for $\omega \in \Omega$. The set Ω denotes the sample space of all the random variables in our model.

The investor would like to maximize the mean semivariance of the portfolio with respect to a target wealth of h and risk aversion coefficient c, and mean return weight b. The mean semivariance with respect to target h with risk aversion coefficient c and mean weight b is defined as

$$b\mathrm{E}(Z) - c\mathrm{E}(h - Z)_+^2,$$

where Z is a random variable. The investment problem can now be posed as the following convex optimization problem:

$$\begin{aligned} \max_u \quad & b\mathrm{E}x_T - c\mathrm{E}(h - x_T)_+^2, \\ \text{s.t.} \quad & x_T(\omega) = r_0 x + p(\omega)' u. \end{aligned}$$

The equality on x_T holds almost surely. Extensions where h itself is a random variable representing the returns of a benchmark portfolio are possible. It is often convenient to specify an excess return target that we denote by η. We define η as the wealth target above the wealth gained from the risk free return: $\eta = h - r_0 x$.

Define the random variable $\lambda(x_T) = [h - x_T \geq 0]$, where $[h - x_T \geq 0]$ denotes the indicator function of the statement $h - x_T \geq 0$, which is equal to 1 whenever the inequality holds and is equal to 0 otherwise. This indicator function allows the objective function to be rewritten in the following way:

$$\begin{aligned} & b\mathrm{E}x_T - c\mathrm{E}(h - x_T)_+^2 \\ = & \mathrm{E}(-ch^2\lambda(x_T)) + \mathrm{E}(b + 2ch\lambda(x))x_T - \mathrm{E}c\lambda(x)x_T^2 \\ = & \mathrm{E}\tilde{\gamma}(x_T) + \mathrm{E}\tilde{\alpha}(x_T)x_T - \mathrm{E}\tilde{\beta}(x_T)x_T^2. \end{aligned}$$

The random variables $\tilde{\gamma}$, $\tilde{\alpha}$, and $\tilde{\beta}$ have the following definitions:

$$\tilde{\gamma}(x_T) = -ch^2\lambda(x_T),$$

$$\tilde{\alpha}(x_T) = b + 2ch\lambda(x_T),$$

$$\tilde{\beta}(x_T) = c\lambda(x_T).$$

Because these random variables are piecewise constant random functions of x_T, the objective function is a piecewise quadratic function. Writing this function in terms of the decision variable u, we arrive at the following optimization problem:

$$\begin{aligned} \max_u \quad & \mathrm{E}(\gamma(x,u) + \alpha(x,u)r_0x - \beta(x,u)(r_0x)^2) \\ & + \mathrm{E}(\alpha(x,u) - 2\beta(x,u)r_0x)p'u - u'\mathrm{E}(b(x,u)pp')u. \end{aligned} \tag{3.1}$$

The following notation is used:

$$\gamma(x,u) = \tilde{\gamma}(r_0x + p'u),$$

$$\alpha(x,u) = \tilde{\alpha}(r_0x + p'u),$$

$$\beta(x,u) = \tilde{\beta}(r_0x + p'u).$$

Let u^* denote the optimal portfolio. We can evaluate the coefficients α, β, and γ at the optimal portfolio to yield the following shorthand notation:

$$\gamma^*(x) = \gamma(x,u^*),$$

$$\alpha^*(x) = \alpha(x,u^*),$$

$$\beta^*(x) = \beta(x,u^*),$$

$$B = \mathrm{E}(\beta^*(x)pp').$$

The optimal portfolio u^* must satisfy the following first order necessary conditions:

$$u^* = \frac{1}{2} B^{-1} \mathrm{E}(\alpha^*(x) - 2\beta^*(x) r_0 x) p. \tag{3.2}$$

First order necessary conditions for optimality are also sufficient because the mean-semi-variance objective is concave. We observe that the expression in (3.2) requires that B is invertible. This is true when the mean-semivariance objective is strictly convex. Strict convexity holds when there are no risk-free portfolios with respect semivariance under the given wealth target. This may not hold, for example, when the wealth target is smaller than the initial wealth. In these cases, an optimal mean-semivariance portfolio may still exist, but the following analysis generally focuses on the strictly convex case.

3.3 Mean-Semivariance as a Piecewise Quadratic Function

Piecewise quadratic functions that are continuously differentiable are also called polyhedral piecewise quadratic functions. This name derives from the following observation. The domain of the function can be partitioned into polyhedral regions such that on each region, the function f is equal to a quadratic function f_k uniquely associated with that region. On the interior of each region k we can uniquely define the Hessian as a matrix Q_k. This notation is formalized by the following definition given due to Patrinos and Sarimveis [22].

Definition 1. Consider a collection of nonempty sets $\mathcal{C} = \{C_k : k \in K\}$ for some finite index set K.

1. $\mathcal{C}$ is called a polyhedral decomposition of $D \subseteq \mathbb{R}^n$, if its members are polyhedral sets and (i) $\bigcup_{k \in K} C_k = D$, (ii) $\dim C_k = \dim D$ for all $k \in K$, (iii) $\operatorname{ri} C_k \cap \operatorname{ri} C_\ell = \emptyset$ for $k, \ell \in K$, $k \neq \ell$, where $\operatorname{ri} C_\ell$ denotes the relative interior of C_ℓ.

2. $\mathcal{C}$ is called a polyhedral subdivision if it is a polyhedral decomposition and furthermore the intersection of any two members of $\mathcal{C}$ is either empty or a common proper face of both.

For the mean-semivariance objective function, the polyhedral subdivision is given by the arrangement of hyperplanes of the form $\{u : \eta - p(\omega)'u = 0\}$. Each polyhedral region $C \in \mathcal{C}$ can therefore be associated with a partition of the sample space $\Omega^1 \cup \Omega^2 = \Omega$ such that

$$C = \{u : \eta - p(\omega)'u \leq 0,\, \omega \in \Omega^1;\, \eta - p(\omega)'u \geq 0,\, \omega \in \Omega^2\}.$$

On the boundary between regions, the function is not twice differentiable, but we can define a generalized Hessian, which is analogous to subdifferentials for functions that are not continuously differentiable. Let u be a point on the boundary between ℓ polyhedral regions with Hessians $Q_1, Q_2, \ldots, Q_\ell$. The generalized Hessian is defined by

$$\partial^2 f(u) = \mathrm{co}(Q_1, Q_2, \ldots, Q_k), \tag{3.3}$$

where $\mathrm{co}(Q_1, Q_2, \ldots, Q_k)$ denotes the convex hull of the Hessians $Q_1, Q_2, \ldots, Q_\ell$.

Let $\mathcal{Q}$ denote the set of all Hessians of the mean-semivariance objective (evaluated at points where the Hessian is defined and unique). Let L_i for $i \in \{2, 3, \ldots, N\}$ denote the collection of boundaries defined by i adjacent polyhedral regions, where N denotes the maximum number of adjacent regions. In every boundary $\partial \in L_i$, the collection of Hessians associated with the regions defining the boundary differ by at most $i - 1$ rank-1 factors. Let $\mathcal{L}$ denote the collection of L_i for all i.

Convex piecewise quadratic functions possess a variety of useful structures. First we introduce an additional piece of notation, for $u \in C_k$, we recall that f can be written as

$$f(u) = f_k(u) = a_k + d_k'u + u'Q_k u.$$

In the following proposition, we write $\tilde{f}_k(u) = f_k(u) + \delta_{C_k}(u)$ where δ_{C_k} is the convex indicator function equal to zero on C_k and $+\infty$ otherwise. The following proposition is adapted from Patrinos and Sarimveis [22].

Proposition 2. *The following properties hold for all choices of the parameters b, h, and x.*

1. *If* $\inf f$ *is finite, then* $\arg\min f$ *is polyhedral.*

2. $x \in \arg\min f$ *if and only if* $x \in \arg\min \tilde{f}_k$ *for all* $k \in K$.

3. *If* $\inf f = \alpha^* > -\infty$*, then* $\arg\min f = \bigcup_{k \in \mathcal{L}(\alpha)} \arg\min \tilde{f}_k$*, where* $\mathcal{L}(\alpha) = \{k \in \mathcal{K} : \{u : \tilde{f}_k(u) \leq \alpha\} \neq \emptyset\}$.

4. $\arg\min f$ *is nonempty if and only if there exists* $x_k \in \mathbb{R}^n$*,* $\beta_k \in \mathbb{R}_+^{m_k}$ *such that* $\nabla f_k(x_k) + (A^k)'\beta_k = 0$ *for every* $k \in \mathcal{K}$*, where* A^k *is the facet representation of* C_k.

3.4 Solution via Nonsmooth Newton's Method

We apply a nonsmooth version of Newton's method for the unconstrained portfolio problem, and we show how the method can be altered to accomodate equality and inequality constraints. The mean-semivariance objective function is a continuously differentiable piecewise quadratic function, for which Sun [14] showed the algorithm to be finitely convergent. The nonsmooth Newton's method generates a sequence of iterates of the form:

$$
\begin{aligned}
d_k &= \arg\max_d \nabla f(u_k)'d - \frac{1}{2}d'V_k d, \\
u_{k+1} &= u_k + d_k.
\end{aligned}
$$

Here V_k is an element of the generalized Hessian at u_k. By substitution, this sequence is equivalent to the following:

$$
u_{k+1} = u_k + V_k^{-1}\nabla f(u_k).
$$

Observing the structure of (3.1), we can identify an element of the generalized Hessian by selecting $V_k = \mathrm{E}(\beta(x, u_k)pp')$. This is an element of the generalized Hessian because it is trivially a convex combination of the Hessians defined by (3.3).

Applying Newton's method to the mean-semivariance objective gives the following sequence of iterates:

$$u_{k+1} = \frac{1}{2}\mathrm{E}(\beta(x, u_k)pp')^{-1}\mathrm{E}(\alpha(x, u_k) - 2\beta(x, u_k)r_0x)p.$$

Following the results of Sun [14], this sequence of portfolios converges in a finite number of iterations to the optimal mean-semivariance portfolio when the objective is strictly convex. In the non-strictly convex cases, the Newton method should be replaced by a proximal point methodology by adding a strictly convex quadratic term to the objective. Assuming the problem is not unbounded, the proximal point approach will also converge as discussed in Sun's [14] paper. Therefore, in the following we address only the strictly convex case.

3.4.1 Modifications for Equality Constraints

Suppose we add linear equality constraints to the mean-semivariance portfolio problem. We arrive at the following problem:

$$\begin{aligned} \max_u \quad & \mathrm{E}(\gamma(x, u) + \alpha(x, u)r_0x - \beta(x, u)(r_0x)^2) \\ & + \mathrm{E}(\alpha(x, u) - 2\beta(x, u)r_0x)p'u - u'\mathrm{E}(b(x, u)pp')u, \\ \text{s.t.} \quad & Au = a. \end{aligned}$$

To generate the iterates for Newton's method in the constrained case, we form the Lagrangian with Lagrange multipliers v,

$$\begin{aligned} L(u, v) = {} & \mathrm{E}(\gamma(x, u) + \alpha(x, u)r_0x - \beta(x, u)(r_0x)^2) \\ & + \mathrm{E}(\alpha(x, u) - 2\beta(x, u)r_0x)p'u - u'\mathrm{E}(b(x, u)pp')u \\ & - v'(Au - a). \end{aligned}$$

Forming the gradient with respect to u and setting it equal to zero gives the following equation:

$$\mathrm{E}(\alpha(x,u) - 2\beta(x,u)r_0 x)p - 2\mathrm{E}(b(x,u)pp')u - A'v = 0.$$

The solution u^* satisfies the following equation:

$$u^* = \frac{1}{2}\mathrm{E}(\beta(x,u^*)pp')^{-1}(\mathrm{E}(\alpha(x,u^*) - 2\beta(x,u^*)r_0 x)p - A'v).$$

To simplify the remaining steps, we use the following shorthand notation:

$$\begin{aligned}
\gamma^* &= \gamma(x,u^*), \\
\alpha^* &= \alpha(x,u^*), \\
\beta^* &= \beta(x,u^*), \\
B &= \mathrm{E}\beta^* pp'.
\end{aligned}$$

Using the requirement that $Au^* - a = 0$, we can solve for the value of v, which we denote by v^*, where

$$v^* = (AB^{-1}A')^{-1}(AB^{-1}\mathrm{E}(\alpha^* - 2\beta^* r_0 x)p - 2a).$$

Substituting this into the solution u^* we have

$$\begin{aligned}
\tilde{u} &= \frac{1}{2}B^{-1}\mathrm{E}(\alpha^* - 2\beta^* r_0 x)p, \\
u^* &= \tilde{u} - B^{-1}A'(AB^{-1}A')^{-1}(A\tilde{u} - a).
\end{aligned}$$

This leads to the following sequence of iterates for Newton's method:

$$\alpha_k = \alpha(x, u_k), \tag{3.4}$$

$$\beta_k = \beta(x, u_k), \tag{3.5}$$

$$B_k = \mathrm{E}(\beta_k pp'), \tag{3.6}$$

$$\tilde{u}_k = \frac{1}{2} B_k^{-1} \mathrm{E}(\alpha_k - 2\beta_k r_0 x)p, \tag{3.7}$$

$$u_{k+1} = \tilde{u}_k - B_k^{-1} A' (A B_k^{-1} A')^{-1} (A\tilde{u}_k - a). \tag{3.8}$$

3.4.2 Modifications for Inequality Constraints

To illustrate the general procedure for introducing inequality constraints, we describe modifications of Newton's method for nonnegativity constraints on the asset weights (which includes nonnegativity of cash). The nonnegativity constrained problem can be written in the following way:

$$\begin{aligned} \max_u \quad & b(r_0 x + \mu' u) - c\mathrm{E}(h - r_0 x - p'u)_+^2, \\ \text{s. t.} \quad & \sum_{i=1}^n u(i) \le 1, \\ & u \ge 0. \end{aligned}$$

The following lemma describes the necessary and sufficient conditions for optimality in the nonnegativity case, which are derived from the KKT conditions.

Lemma 3. *Let $g = \nabla f(u)$. The vector u is optimal if and only if one of the following two conditions is satisfied.*

1. $\sum_i u(i) < 1$, $g \le 0$ *and* $g_i u_i = 0$ *for all* i.

2. $\sum_i u(i) = 1$ *and there exists a scalar* t *such that* $g_i \le t$ *for all* i *such that* $u_i = 0$ *and* $g_i = t$ *for all* i *such that* $u_i > 0$.

Proof. The KKT conditions hold at u if there exists v and t such that

$$
\begin{aligned}
g + v - te &= 0, \\
v(i)u(i) &= 0, \quad i = 1, 2, \ldots, n, \\
\left(\sum_j u(j) - 1\right)t &= 0, \\
t, v &\geq 0,
\end{aligned}
$$

where e is a vector of ones. In case 1, complementary slackness implies $t = 0$. Therefore, the first equation gives the relation $g = -v$, so that $v \geq 0$ is equivalent to $g \leq 0$. Using this relation for the complementary slackness conditions from the second set of equations gives $v(i)u(i) = -g(i)u(i) = 0$, which is equivalent to $g(i)u(i) = 0$.

In case 2, setting $v_i = 0$ and $t = g_i$ for all i such that $u_i > 0$ satisfies the first equation and the associated complementary slackness equations. For i such that $u_i = 0$, we can choose $v_i = t - g_i$, so that $v \geq 0$ is equivalent to $g \leq t$. The convexity of f yields that these conditions are necessary and sufficient. □

Applying the conditions of Lemma 3 we construct algorithm 1. We introduce an additional notation to simplify the algorithm's presentation. Let I denote a subset of the indices $1, 2, \ldots, n$. Given a square matrix A, we denote by A_I the submatrix consisting of those rows and columns specified by the index set I. Given a vector u, a similar notation applies for a subvector u_I of u. We now prove the correctness of algorithm 1.

Proposition 4. *Algorithm 1 returns the nonnegative mean-semivariance optimal portfolio.*

Proof. The algorithms stops if one of two conditions are met, either $I = \emptyset$ or $u'e - 1 = 0$ and g_i is constant on J and equal to a value t with $g_i \leq t$ for $i \notin J$. The latter case is exactly the KKT conditions described in Lemma 3. Therefore the solution is optimal. In the former case, the KKT conditions are satisfied for all the nonnegativity constraints. Setting t equal 0 satisfies the KKT conditions from Lemma 3. □

We now give an intuition behind the construction of Algorithm 1. The outer loop identifies asset indices for which the nonnegativity constraint is tight and for which shorting the asset would improve the portfolio's mean-semivariance. The inner loop repeatedly tries to use Newton steps to improve the portfolio while maintaining feasibility. The two inner if-statements differentiate whether all wealth is currently invested, in which case the step must be adjusted to not short cash. The process repeats until either the optimal portfolio is found with no constraints being tight, or the only way to improve the portfolio is to violate one of the constraints. It is this basic structure that would be used to extend Algorithm 1 over general linear inequality constraints.

3.4.3 Detecting Unbounded Problems

Up to now we had assumed that the problem was well posed in the sense that the objective function is bounded. This condition can fail even if the objective is defined by n linearly independent return samples. We derive here the conditions that lead to an unbounded objective function. We introduce two notations that simplify the proposition.

Definition 5. Given a matrix A, the kernel of the matrix A is the following set:

$$\ker A = \{x : Ax = 0\}.$$

Definition 6. The horizon or recession cone of a set C is the following set:

$$C^\infty = \{d : x + d \in C, \forall x \in C\}.$$

Proposition 7. *The mean-semivariance objective is unbounded if and only if there exists* $C_k \in \mathcal{C}$ *such that* $\ker Q_k \cap \{d : (b\mu + 2\eta\mu_k)'d > 0\} \cap C_k^\infty \neq \emptyset$.

Proof. Let $d \in \ker Q_k \cap \{d : (b\mu + 2\eta\mu_k)'d > 0\} \cap C_k^\infty$ and $u \in C_k$. By construction $u + td \in C_k$ for all $t > 0$ and $f(u + td)$ is an increasing linear function of t for all $t > 0$.

Algorithm 1 Given historical rates of return r, risk free rate of return r_0, target wealth h, risk aversion coefficient c, and initial nonnegative portfolio u, the algorithm returns u^* the optimal mean-semivariance portfolio with nonnegative weights. The algorithm stops when no feasible improving direction exists.

Set $u \leftarrow u_0$.
$\lambda(\omega) \leftarrow [h - r_0 x - p(\omega)'u > 0]$ for $\omega \in \Omega$.
Calculate $g \leftarrow \nabla f_\lambda(u)$
Set $I \leftarrow \{i : g_i > 0\} \cup \{i : g_i < 0,\, u_i > 0\}$.
while $I \neq \emptyset$ **do**
    Set $S \leftarrow$ FALSE.
    while $S =$ FALSE **do**
        if $\sum_j u(j) < 1$ **then**
            Calculate $\tilde{u} \leftarrow Q_I(\lambda)^{-1}(\frac{b}{2c}\mu + (h - r_0 x)\mu_I(\lambda))$
            Let σ be the mapping from $\{1, 2, \ldots, |I|\}$ to I given by the indexing for μ_I.
            Set $d_i \leftarrow 0$ if $i \notin I$ and $d_i = \tilde{u}_{\sigma(i)} - u_i$ if $i \in I$.
            Set $\beta^* \leftarrow \max\{\beta : u + \beta d \geq 0, (u + \beta d)'e \leq 1\}$.
            if $\beta^* > 0$ **then**
                Set $S \leftarrow$ TRUE.
                Set $u \leftarrow u + \beta^* d$.
            else
                Set $I \leftarrow I \backslash \{i : d_i < 0,\, u_i = 0\}$.
        if $\sum_j u(j) = 1$ **then**
            Calculate $\bar{u} \leftarrow Q_I(\lambda)^{-1}(\frac{b}{2c}\mu + (h - r_0 x)\mu_I(\lambda))$
            if $(\bar{u} - u)'e_I \geq 0$ **then**
                Calculate $\tilde{u} \leftarrow \bar{u} - (e'Q_I(\lambda)^{-1}e)^{-1}(e'\bar{u} - 1)Q_I(\lambda)^{-1}e$
            else
                $\tilde{u} \leftarrow \bar{u}$
            Let σ be the mapping from $\{1, 2, \ldots, |I|\}$ to I given by the indexing for μ_I.
            Let $d_i \leftarrow 0$ if $i \notin I$ and $d_i \leftarrow \tilde{u}_{\sigma(i)} - u_i$ if $i \in I$.
            Set $\beta^* \leftarrow \max\{\beta : u + \beta d \geq 0, (u + \beta d)'e \leq 1\}$.
            if $\beta^* > 0$ **then**
                Set $S \leftarrow$ TRUE.
                Set $u \leftarrow u + \beta^* d$.
            else
                Set $I \leftarrow I \backslash \{i : d_i < 0,\, u_i = 0\}$.
        if $I = \emptyset$ **then**
            Exit Loop.
    $\lambda(\omega) \leftarrow [h - r_0 x - p(\omega)'u > 0]$ for $\omega \in \Omega$.
    Calculate $g \leftarrow \nabla f_\lambda(u)$
    Set $I \leftarrow \{i : g_i > 0\} \cup \{i : g_i < 0,\, u_i > 0\}$.
    if $\sum_j u(j) = 1$ **then**
        Set $J = \{i : u_i > 0\}$.
        if $g_J - g_J(1) = 0$ and $g_i \leq g_J(1)$ for all $i \notin J$ **then**
            Exit loop.
Return $u^* \leftarrow u$.

Hence f is unbounded.

Conversely suppose that f is unbounded. This means that for all $M > 0$ there exists u such that $f(u) > M$. Let u_i be a sequence such that $f(u_i)$ is strictly increasing in i and $f(u_i) \to \infty$ and $i \to \infty$. Because the functions f_k for all $k \in K$ are bounded on any compact set, there must exist I such that for all $i > I$ we have u_i only lie in a collection of C_k such that each C_k is an unbounded polyhedron. Let K_{ub} denote the finite collection such that $u_i \in \cup_{k \in K_{ub}} C_k$ for all $i > I$.

Because K_{ub} is finite, we can choose a subsequence $j > I$ such that $u_j \in C_{k^*}$ for all j and $f(u_j) \to \infty$ as $j \to \infty$. Because $f(u_j) = f_{k^*}(u_j)$ for all j, we have that f is a quadratic function when restricted to this subsequence. Furthermore, because C_{k^*} is polyhedral and u_j is an unbounded sequence, there must exist a further subsequence ℓ such that for some $d \in C_{k^*}^{\infty}$ we have $u_\ell / \|u_\ell\| \to d$.

Because Q_{k^*} is positive semidefinite, and contributes a nonpositive term to f_{k^*}, it must follow that $d' Q_{k^*} d = 0$ so that $d \in \ker Q_{k^*}$, and $(b\mu + 2\eta\mu_{k^*})' d > 0$. We have shown therefore that there exists C_k such that $d \in \ker Q_k \cap \{d : (b\mu + 2\eta\mu_k)' d > 0\} \cap C_k^{\infty}$. □

As a sufficient condition for boundedness of f, we can verify that the kernel of Q_k is empty at each iteration of the optimization algorithm. Although the kernel can be non-empty and the problem be bounded, this is still an indicator of the matrix being ill conditioned at that iterate which indicates that the resulting portfolio may be very unstable to changes in the problem data.

3.5 Sensitivity Analysis of Mean-Semivariance Portfolio

In this section, we investigate the behavior of the optimal mean-semivariance portfolio u^* as a function of the parameters b and $\eta = h - r_0 x$ where r_0 is deterministic and f has a unique solution. We develop these characterizations as a basis for understanding the behavior of optimal portfolios calculated for one set of parameters when evaluated against a fixed set of parameters. In other words, let f_0 denote the mean-semivariance objective function with parameters b_0, c and η_0. Let $U(b,\eta)$ denote the optimal mean-semivariance portfolio as a function of b and η. Our objective is to derive properties of $f_0(U(b,\eta))$ by understanding the behavior of $U(b,\eta)$.

We use the following notation:

$$U(b,\eta) = Q(\Lambda(b,\eta))^{-1}\left(\frac{b}{2c}\mu + \eta\mu(\Lambda(b,\eta))\right), \tag{3.9}$$

$$\Lambda(b,\eta) = [\eta - p'U(b,\eta) \geq 0]. \tag{3.10}$$

We observe that Λ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^{|\Omega|}$, i.e. for each (b,η), $\lambda = \Lambda(b,\eta)$ is a random vector taking values 0 or 1. We denote specific realizations by $\lambda(\omega)$. The convergence of the nonsmooth Newton's method ensures that $\Lambda(b,\eta)$ is uniquely defined for all choices of b and η. We will generally present our results for the typical cases where $b \geq 0$ and $\eta \geq 0$.

We first characterize the behavior of Λ. Our goal is to show that $U(b,\eta)$ is a continuous piecewise linear mapping from $\mathbb{R}^2$ to $\mathbb{R}^n$. We also give sufficient conditions for $\|U(b,\eta)\|^2$ to be nondecreasing in b.

To simplify the presentation in the following lemmas we define the following notation:

$$\mu_H = \mu(\lambda^H), \quad \tilde{\mu}_H = \mu(\tilde{\lambda}^H), \tag{3.11}$$

$$Q_H = Q(\lambda^H), \quad \tilde{Q} = Q(\tilde{\lambda}), \tag{3.12}$$

$$\tilde{Q}_H = Q(\tilde{\lambda}^H). \tag{3.13}$$

Lemma 8. *Let $\eta > 0$. The mapping Λ has the following properties.*

1. *The mapping $\Lambda(0,\eta)$ is constant with respect to η. Let the value of this constant be denoted by λ^H.*

2. *If $1 - p(\omega)'Q_H^{-1}\mu_H \neq 0$ for all $\omega \in \Omega$, then λ^H is the limit of the $\Lambda(b,\eta)$ as $\eta \to \infty$.*

3. *The mapping $\Lambda(b,0)$ is constant with respect to b. Let this constant value be denoted by λ^B.*

Proof. Part 1: The mapping $\Lambda(0,\eta)$ is the unique solution to the following system of binary equations:

$$\lambda(\omega) = [\eta\big(1 - p(\omega)'Q(\lambda)^{-1}\mu(\lambda)\big) \geq 0], \quad \forall \omega \in \Omega,$$

in terms of the binary vector λ. Because $\eta \geq 0$, the solution does not depend on η. The value of $\Lambda(0,\eta)$ does not depend on η and, therefore, is constant. Let λ^H denote this constant value.

Part 2: Now suppose that $b > 0$. To show the limiting result, it is sufficient to show that there exists M such that for all $\eta > M$, λ^H is a solution to the following system of binary equations:

$$\lambda(\omega) = \left[\eta\big(1 - p(\omega)'Q(\lambda)^{-1}\mu(\lambda)\big) \geq \frac{b}{2c}p(\omega)'Q(\lambda)^{-1}\mu\right], \quad \forall \omega \in \Omega,$$

in terms of the binary vector λ.

If $\lambda^H(\omega) = 1$, then $1 - p(\omega)'Q_H^{-1}(\mu_H) > 0$. If $\frac{b}{2c}p(\omega)'Q_H^{-1}\mu \leq 0$, then the inequality is satisfied for all $\eta \geq 0$. If $\frac{b}{2c}p(\omega)'Q_H^{-1}\mu > 0$, then the inequality is satisfied for all $\eta > m(\omega)$ where $m(\omega)$ is defined by

$$m(\omega) = \frac{\frac{b}{2c}p(\omega)'Q_H^{-1}\mu}{\big(1 - p(\omega)'Q_H^{-1}(\mu_H)\big)}.$$

If $\lambda^H(\omega) = 0$, then $1 - p(\omega)'Q_H^{-1}(\mu_H) < 0$. If $\frac{b}{2c}p(\omega)'Q_H^{-1}\mu \geq 0$, then the inequality is violated for all $\eta \geq 0$. If $\frac{b}{2c}p(\omega)'Q_H^{-1}\mu < 0$, then the inequality is violated for all $\eta > m(\omega)$ where $m(\omega)$ is defined by

$$m(\omega) = \frac{\frac{b}{2c}p(\omega)'Q_H^{-1}\mu}{\big(1 - p(\omega)'Q_H^{-1}(\mu_H)\big)}.$$

By choosing $M > \max_{\omega \in \Omega} m(\omega)$, we have that λ^H satisfies the binary equation for all $\eta > M$.

Part 3: Observe that $\Lambda(b, 0)$ solves the following equation:

$$\lambda(\omega) = \left[\frac{b}{2c} p(\omega)' Q(\lambda)^{-1} \mu \le 0\right],$$

in terms of the binary vector λ. The value of the indicator function depends only on the sign of $p(\omega)' Q(\lambda)\mu$, so the solution cannot depend on b. Hence, $\Lambda(b, 0)$ is constant with respect to b, which can be denoted by λ^B. □

Lemma 9. *Let $\eta_0 \ge 0$. The indicator function*

$$[\eta_0 - p(\omega)' U(b, \eta) \ge 0]$$

has a unique limit, denoted by $\tilde{\lambda}$, as $\eta \to \infty$. The limit has the following formula:

$$\tilde{\lambda}(\omega) = [p(\omega)' Q_H^{-1} \mu_H < 0] + [p(\omega)' Q_H^{-1} \mu_H = 0] \left[\frac{b}{2c} p(\omega)' Q_H^{-1} \mu \le \eta_0\right].$$

Furthermore, the first term in the definition, which we denote by $\tilde{\lambda}^H$ defined by

$$\tilde{\lambda}^H(\omega) = [p(\omega)' Q_H^{-1} \mu_H \le 0],$$

satisfies

$$\mu_H Q_H^{-1} Q(\tilde{\lambda}) Q_H^{-1} \mu_H = \mu_H Q_H^{-1} Q(\tilde{\lambda}^H) Q_H^{-1} \mu_H.$$

Proof. Let η_0 be given. By Lemma 8, we can select η such that $\Lambda(b, \eta) = \lambda^H$. Substitute the definition of $U(b, \eta)$ into the definition of the indicator function in the hypothesis and rearrange terms:

$$\left[\eta p(\omega)' Q_H^{-1} \mu_H \le \eta_0 - \frac{b}{2c} p(\omega)' Q_H^{-1} \mu\right].$$

We evaluate the limit of this indicator function by cases.

First, if $p(\omega)' Q_H \mu_H > 0$, then there exists $H_1 > 0$ such that the inequality is violated for all $\eta > H_1$ for any fixed η_0 and b. Second, if $p(\omega)'Q_H\mu_H < 0$, then there exists $H_2 > 0$ such that the inequality is satisfied for all $\eta > H_2$ for any fixed η_0 and b. In the third case, $p(\omega)'Q_H\mu_H = 0$, the left hand side of the inequality is independent of η, so this term is constant with respect to η. Hence, the limit exists. The formula for the limit follows by converting the cases into indicator functions.

The second part follows by direct computation. Let

$$\bar{\Omega} = \{\omega : p(\omega)'Q_H\mu_H = 0\}.$$

If $\bar{\Omega}$ is empty, then the result is trivial because $\tilde{\lambda} = \tilde{\lambda}^H$, so suppose otherwise. To simplify notation, let $v = Q_H^{-1}\mu_H$, then

$$\begin{aligned}
&v'Q(\tilde{\lambda})v \\
&\quad = \sum_{w\in\Omega} \mathbb{P}(\omega)\tilde{\lambda}(\omega)v'p(\omega)p(\omega)'v \\
&\quad = \sum_{w\in\bar{\Omega}^C} \mathbb{P}(\omega)\tilde{\lambda}(\omega)v'p(\omega)p(\omega)'v \\
&\qquad + \sum_{w\in\bar{\Omega}} \mathbb{P}(\omega)\tilde{\lambda}(\omega)v'p(\omega)p(\omega)'v \\
&\quad = \sum_{w\in\bar{\Omega}^C} \mathbb{P}(\omega)\tilde{\lambda}(\omega)v'p(\omega)p(\omega)'v \\
&\quad = \mu_H'Q_H^{-1}Q(\tilde{\lambda}^H)Q_H^{-1}\mu_H.
\end{aligned}$$

The second to last equality follows from the definition of $\bar{\Omega}$. The last equality follows because the summation holds for any choice of λ that differs from $\tilde{\lambda}$ only on $\omega \in \bar{\Omega}$. $\square$

Suppose we fix a portfolio $d \in \mathbb{R}^n$. We consider now the restricted optimization problem of how much wealth to invest in this portfolio to maximize mean-semivariance. In other words, we restrict our attention only to portfolios of the form τd for some scalar $\tau \in \mathbb{R}$ and fixed portfolio d, normalized to a unit vector. Under this restriction we arrive at a one-dimensional optimization problem,

$$\max_{\tau} br_0 x - c\eta^2 \mathrm{E}\lambda + ((b\mu + 2c\eta\mu(\lambda))'d)\tau - (cd'Q(\lambda)d)\tau^2,$$

which has a solution τ^* given by:

$$\tau^* = \frac{(b\mu + 2c\eta\mu(\lambda^*))'d}{2cd'Q(\lambda^*)d}.$$

Here, λ^* satisfies the relationship:

$$\lambda^*(\omega) = [\eta - p(\omega)'(\tau^* d) \geq 0].$$

We parameterize the optimal solution as a function of d, which we denote by $t(d)$, the optimal solution mapping:

$$t(d) = \frac{(b\mu + 2c\eta\mu(\lambda^*))'d}{2cd'Q(\lambda^*)d},$$

where λ^* is implicitly a function of d. The following proposition gives a particular characterization of $t(d)$. We first introduce lemmas for the properties of the solution mapping $t(d)$.

Lemma 10. *The derivative of $f'(\tau d)$ with respect to τ is convex for $\tau \geq 0$.*

Proof. We proceed by showing that the derivative of the semivariance function with respect to τd is concave in τ for $\tau \geq 0$:

$$\sum_{\omega \in \Omega} -2p(\omega)'d(\eta - p(\omega)'d\tau)[\eta - p(\omega)'d\tau \geq 0] \tag{3.14}$$

$$= \sum_{\omega \in \Omega} (-2\eta p(\omega)'d + 2(p(\omega)'d)^2\tau)[\eta - p(\omega)'d\tau \geq 0, p(\omega)'d < 0] \tag{3.15}$$

$$+ \sum_{\omega \in \Omega} (-2\eta p(\omega)'d + 2(p(\omega)'d)^2\tau)[\eta - p(\omega)'d\tau \geq 0, p(\omega)'d \geq 0] \tag{3.16}$$

$$= a + b\tau + \sum_{\omega \in \Omega'} (\alpha_\omega + \beta_\omega \tau)[\tau \leq \gamma_\omega]. \tag{3.17}$$

The first equality is a simple re-arrangement and grouping of terms. The second equality is more involved, so we first define the notation used:

$$\begin{aligned} \alpha_\omega &= -2\eta p(\omega)'d, & \beta_\omega &= 2(p(\omega)'d)^2, \\ a &= \sum_{\omega\in\Omega} \alpha_\omega[p(\omega)'d < 0], & b &= \sum_{\omega\in\Omega} \beta_\omega[p(\omega)'d < 0], \\ \gamma_\omega &= \frac{\eta}{p(\omega)'d}, & \Omega' &= \{\omega \in \Omega : p(\omega)'d \geq 0\}. \end{aligned}$$

Associating the first summation in (3.16) with a follows from the observation that $p(\omega)'d < 0$ implies $\eta - p(\omega)'d\tau \geq 0$ for all $\tau > 0$. The same applies to b. The mapping of terms in the second summation of (3.16) is a direct grouping of terms from the linear expression in τ and solving for τ in the break point of the indicator function. It is now sufficient to show that the second term in the piecewise linear expression in (3.17) is concave. We denote this piecewise linear function by g.

Let ω_i for $i = 1, 2, \ldots, |\Omega'|$ denote an ordering of Ω' such that γ_{ω_i} is an increasing sequence. Under this ordering it is clear that the slope of g is given by the following expression:

$$\sum_{i=1}^{k(\tau)} \beta_{\omega_i},$$

where $k(\tau) = \arg\max\{i : \tau \leq \gamma_{\omega_i}\}$. It is clear from the ordering of ω_i that $k(\tau)$ is non-increasing, and consequently so is the slope of g, which is equivalent to concavity for a piecewise linear function. □

Lemma 11. *If* $\mu'd > 0$*, then*

$$0 < t(d) = \max_\ell \left\{ \frac{(b\mu + 2c\eta\mu(\ell))'d}{2cd'Q(\ell)d} : \ell(\omega) = 1, \, \forall\omega \in \{p(\omega)'d < 0\} \right\}.$$

If $\mu'd < 0$*, then*

$$0 > t(d) = \min_\ell \left\{ \frac{(b\mu + 2c\eta\mu(\ell))'d}{2cd'Q(\ell)d} : \ell(\omega) = 1, \, \forall\omega \in \{p(\omega)'d > 0\} \right\}.$$

Proof. Suppose $\mu' d > 0$. We observe that at $\tau = 0$ we evaluate $f(\tau d)$ using the following vector:

$$\lambda_0 = [\eta - p'(\tau d)] = [\eta \geq 0] = 1,$$

so that $\mu(\lambda_0) = \mu$. For $\tau \neq 0$, the derivative of the objective can be written as

$$b\mu' d + 2c\eta\mu(\lambda)' d - 2cd' Q(\lambda) d\tau.$$

From concavity, $f'(\tau) > 0$ for all $\tau < t(d)$. So we see that

$$f'(\tau) = b\mu' d + 2c\eta\mu(\lambda^*)' d - 2cd' Q(\lambda^*) d\tau > 0 = f'(t(d)).$$

Hence, $t(d) > 0$. For $\mu' d < 0$, the analogous result immediately follows by the same argument.

Again assume $\mu' d > 0$. Given a fixed binary vector $\ell \in \mathbb{R}^{|\Omega|}$, let $t_\ell(d)$ be defined by the following expression:

$$t_\ell(d) = \frac{(\mu + 2c\eta\mu(\ell))' d}{2cd' Q(\ell) d}.$$

Let λ^* denote the binary vector at the optimal solution with respect to d. Because we know that $t(d) = t_{\lambda^*}(d) > 0$, we can restrict our consideration to binary vectors ℓ associated with $\tau > 0$. This restriction is equivalent to $\ell(\omega) = 1$ if $p(\omega)' d < 0$ because $\eta - p(\omega)'(\tau d) > 0$ for all $\tau > 0$ if $p(\omega)' d < 0$.

The last portion of the proof relies on the observation that $f'(\tau)$ is a convex function for $\tau > 0$, which is an immediate consequence of Lemma 10. Let τ be such that $0 < \tau < t(d)$. Let ℓ be generated by τ, and assume $\ell \neq \lambda^*$. We plan to show that $t_\ell(d) < t(d)$. By the convexity of f', we know that the following is true:

$$f'(\tau) + f''(\tau)(\tau' - \tau) \leq f'(\tau'),$$

for all τ'. The inequality is strict for all $\tau' \in (\tau_\ell, t(d))$, τ_ℓ corresponds to the boundary such that $f'(\tau')$ is linear for $\tau' \in [\tau, \tau_\ell]$. This implies that

$$f'(\tau) + f''(\tau)(t(d) - \tau) < f'(t(d)) = 0.$$

By construction $f'(\tau) + f''(\tau)(t_\ell(d) - \tau) = 0$, because $t_\ell(d)$ corresponds to the solution of a quadratic optimization problem. Because $f'(\tau) + f''(\tau)(\tau' - \tau)$ is linear as a function of τ' with $f'(\tau) > 0$ and $f''(\tau) < 0$, it must follow that $t_\ell(d) < t(d)$. The same argument can be repeated for $\tau > t(d)$. □

We now introduce two results that follow from Lemma 11. For this notation we consider the optimal solution mapping as a function of d, b, and η jointly. We denote this mapping by $T(d, b, \eta)$.

Lemma 12. *Given a unit vector d and η. Let the optimal mapping for the 1-dimensional solution be given by*

$$T(d, b, \eta) = \frac{(b\mu + 2c\eta\mu(\lambda^*))'d}{2cd'Q(\lambda^*)d}.$$

If $\mu'd > 0$, T is strictly increasing as a function of b. The function is strictly decreasing if $\mu'd < 0$. Given d and b, the function T is convex in η

Proof. The result follows by observing that, for $b_1 > b_0$ such that $b_1 - b_0$ is sufficiently small so that $\Lambda(b_0, \eta) = \Lambda(b_1, \eta)$, $T(d, b, \eta)$ is a linear function of b with positive slope. The result for $\mu'd < 0$ follows by the same argument except that the slope is negative.

Convexity of $T(d, b, \eta)$ is immediate when $b = 0$ because then it is a linear function of η. Suppose $b > 0$. Let $\omega_i \in \Omega$ such that $p(\omega_i)'d > 0$, and

$$p(\omega_i)'d > p(\omega_{i+1})'d$$

for $i \in \{1, 2, \ldots, B\}$. The gradient of the objective function is piecewise linear with change points located at

$$\beta_i = \frac{\eta}{p(\omega_i)'d}.$$

Let $f_d(b)$ denote the objective function for a fixed vector d evaluated as a function of b. We now denote the value of the gradient at each change point β_i as a function of η:

$$\alpha_i(\eta) = f_d'(\beta_i) = \beta_i \mu' d + 2c \frac{\mu_i' d - d' Q_i d}{p(\omega_i)'d} \eta.$$

We now argue that according to our choice of labels that the derivatives α_i' are decreasing in i. By the assumption that $\mu' d > 0$, the derivative f_d' is strictly decreasing as a function of b. We immediately conclude for $\eta > 0$ that $\alpha_i(\eta)$ is strictly decreasing in i.

We now distinguish two cases. First, suppose $\alpha_B' > 0$. Under the assumption that $\mu' d > 0$, the gradient is never zero at any of the breakpoints and so $T(d, b, \eta)$ is linear in η and therefore convex.

Second, suppose $\alpha_B' < 0$. As η increases, $T(d, b, \eta)$ moves from region B to region $B - 1$, and inductively from region $B - i + 1$ to region $B - i$ for $i = 1, 2, \ldots, K$, where K is the region with the least negative slope $\alpha_K' = \max_j \{\alpha_j : \alpha_j < 0\}$.

Following this argument, the derivative $D_\eta T(d, b, \eta)$ is equal to the following constant:

$$\frac{\mu_{B-i}' d}{d' Q_{B-i} d}, \tag{3.18}$$

for $i = 0, 1, \ldots, K$. It is now sufficient to show that the sequence in equation (3.18) is non-decreasing in i. The following calculations give the basis for this conclusion:

$$
\begin{aligned}
& d'Q'_{B-i}d \\
&= \sum_{\omega:p(\omega)'d<0} \mathbb{P}(\omega)(p(\omega)'d)^2 + \sum_{j=0}^{i-1} \mathbb{P}(\omega)(p(\omega_{B-j})'d)^2 \\
&\leq \sum_{\omega:p(\omega)'d<0} \mathbb{P}(\omega)(p(\omega)'d)^2 + \sum_{j=0}^{i} \mathbb{P}(\omega)(p(\omega_{B-j})'d)^2 \\
&= d'Q_{B-(i+1)}d
\end{aligned}
$$

and

$$
\begin{aligned}
& \mu'_{B-i}d \\
&= \sum_{\omega:p(\omega)'d<0} \mathbb{P}(\omega)p(\omega)'d + \sum_{j=0}^{i-1} \mathbb{P}(\omega)p(\omega_i)'d \\
&\leq \sum_{\omega:p(\omega)'d<0} \mathbb{P}(\omega)p(\omega)'d + \sum_{j=0}^{i} \mathbb{P}(\omega)p(\omega_i)'d \\
&= \mu'_{B-(i+1)}d.
\end{aligned}
$$

Now we make the following additional observation. The calculation $\mu'_B d$ is the sum of only negative terms, so it must be negative. In combining the previous results we have the following implications:

$$
\begin{aligned}
& \frac{1}{d'Q_{B-i}d} \geq \frac{1}{d'Q_{B-(i+1)}d} \\
\Rightarrow \quad & \frac{\mu'_{B-i}d}{d'Q_{B-i}d} \leq \frac{\mu'_{B-i}d}{d'Q_{B-(i+1)}d} \\
\Rightarrow \quad & \frac{\mu'_{B-i}d}{d'Q_{B-i}d} \leq \frac{\mu'_{B-(i+1)}d}{d'Q_{B-(i+1)}d}.
\end{aligned}
$$

The implication holds whenever $\mu'_{B-i}d < 0$.

To simplify exposition, we use the notation $m_i = \mu'_{B-i}d$, $q_i = d'Q_{B-i}$, $\pi_i = p(\omega_{B-i})'d$, and $\mathbb{P}_i = \mathbb{P}(\omega_{B-i})$. With this notation we utilize the following relationships:

$$m_{i+1} - m_i = \mathbb{P}_i \pi_i,$$
$$q_{i+1} - q_i = \mathbb{P}_i \pi_i^2.$$

We wish to show that m_i/q_i is an increasing sequence for i such that $m_i > 0$ and $i \leq K$. We now show how answering this question is equivalent to another:

$$\begin{aligned}
& \frac{m_{i+1}}{q_{i+1}} > \frac{m_i}{q_i} \\
\Leftrightarrow \quad & \frac{m_{i+1}}{m_i} > \frac{q_{i+1}}{q_i} \\
\Leftrightarrow \quad & \frac{m_{i+1} - m_i}{m_i} > \frac{q_{i+1} - q_i}{q_i} \\
\Leftrightarrow \quad & \frac{\mathbb{P}_i \pi_i}{m_i} > \frac{\mathbb{P}_i (\pi_i)^2}{q_i} \\
\Leftrightarrow \quad & \frac{q_i}{m_i} > \pi_i.
\end{aligned}$$

Because $m_i - q_i < 0$ for all $i \leq K$, we know that $q_i/m_i > 1$. From the definition of K, $m_K - q_K < 0$ and $m_{K+1} - q_{K+1} > 0$. This implies the following relationship:

$$\begin{aligned}
& (m_{K+1} - q_{K+1}) - (m_K - q_K) > 0 \\
\Leftrightarrow \quad & m_{K+1} - m_K - (q_{K+1} - q_K) > 0 \\
\Leftrightarrow \quad & \mathbb{P}_K \pi_K - \mathbb{P}_K (\pi_K)^2 > 0 \\
\Leftrightarrow \quad & \mathbb{P}_K \pi_K (1 - \pi_K) > 0.
\end{aligned}$$

Because $\mathbb{P}_K \pi_K > 0$, the last inequality implies $1 - \pi_K > 0$ or $\pi_K < 1 < q_K/m_K$. Because π_i is an increasing sequence, we have $\pi_i < 1$ for all $i < K$. This completes the proof that m_i/q_i is increasing, and so $T(d, b, \eta)$ is convex as a function of η. $\square$

Corollary 13. *Let E denote the following set:*

$$E = \{u \in \mathbb{R}^n : \|u\| \leq t(u/\|u\|)\}.$$

Let $\mathcal{E}$ denote the boundary of E. Let $\xi(\phi)$ denote a parameterization of $t(d)$ in terms of the spherical coordinates of the unit vectors. There exists a polyhedral partition in the space of spherical coordinates $\mathcal{S}$ such that on each partition $S \in \mathcal{S}$ there exists a binary vector λ^* that is optimal for all d associated with the coordinates $\phi \in S$. As represented in $\mathbb{R}^n$, the set $E \cap S$ is convex for all $S \in \mathcal{S}$.

Proof. The existence of the polyhedral partition of the set of spherical coordinates follows directly from the same argument for the existence of a polyhedral partition of $\mathbb{R}^n$ with similar properties. For a fixed binary vector ℓ, the function is $t_\ell(d)$ is an ellipsoid, so the set $E \cap S$ is the intersection of a cone S and an ellipsoid. Hence, the set is convex. □

Proposition 14. *The following properties hold for $U(b, \eta)$ as a function of (b, η).*

1. *The mapping $U(b, \eta) : \mathbb{R}^2 \mapsto \mathbb{R}^n$ is continuous piecewise linear.*
2. *$U(b, 0)$ and $U(0, \eta)$ are linear functions of b and η, respectively.*
3. *For (b, η) such that $U(b, \eta)'Q(b, \eta)U(b, \eta) \geq \frac{1}{4c}$, $\|U(b, \eta)\|$ is nondecreasing in b, (strictly increasing if the inequality is held strictly).*
4. *$U(0, 0) = 0$.*

Proof. Part 1: Because U is a piecewise linear function, we only need to check continuity at the points where $\Lambda(b, \eta)$ is not defined by strict inequalities in the indicator function. When (b, η) is such that $\eta - p(\omega)'U(b, \eta) = 0$, then there exist at least two distinct binary vectors ℓ_1 and ℓ_2 such that

$$\{\omega : \eta - p(\omega)'U(b, \eta) = 0\} = \{\omega : \eta - p(\omega)'\tilde{u}(\ell_1) = 0\} = \{\omega : \eta - p(\omega)'\tilde{u}(\ell_2) = 0\},$$

and the vectors ℓ_1 and ℓ_2 differ only on $\omega \in \Omega$ such that $p(\omega)'U(b, \eta) = \eta$. The portfolios $\tilde{u}(\ell_1)$ and $\tilde{u}(\ell_2)$ correspond to limits of the function $U(b', \eta')$ as (b', η') approaches (b, η) from regions neighboring the particular point (b, η). By showing that $\tilde{u}(\ell_1)$ is the same as $\tilde{u}(\ell_2)$, we ensure that U is a continuous function of b and η.

Let u_1 and u_2 denote $\tilde{u}(\ell_1)$ and $\tilde{u}(\ell_2)$, respectively. By definition u_1 is the solution to the following system of equations:

$$Q(\ell_1)u_1 = \frac{b}{2c}\mu + \eta\mu(\ell_1). \tag{3.19}$$

We now rearrange the terms of the equation:

$$\begin{aligned}
Q(\ell_1)u_1 &= \sum_{\omega\in\Omega} \mathbb{P}(\omega)\ell_1(\omega)p(\omega)p(\omega)'u_1 \\
&= \sum_{\omega\in\Omega:\ell_1=\ell_2} \mathbb{P}(\omega)\ell_1(\omega)p(\omega)p(\omega)'u_1 \\
&\quad + \sum_{\omega\in\Omega:\ell_1\neq\ell_2} \mathbb{P}(\omega)\ell_1(\omega)p(\omega)p(\omega)'u_1 \\
&= \sum_{\omega\in\Omega:\ell_1=\ell_2} \mathbb{P}(\omega)\ell_1(\omega)p(\omega)p(\omega)'u_1 \\
&\quad + \eta \sum_{\omega\in\Omega:\ell_1\neq\ell_2} \mathbb{P}(\omega)\ell_1(\omega)p(\omega).
\end{aligned}$$

Setting this result equal to the right hand side of the original equation gives

$$\begin{aligned}\sum_{\omega\in\Omega:\ell_1=\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega)p(\omega)'u_1 &= \frac{b}{2c}\mu+\eta\mu(\ell_1)\\ &\quad -\eta\sum_{\omega\in\Omega:\ell_1\neq\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega)\\ &= \frac{b}{2c}\mu+\eta\left(\sum_{\omega\in\Omega:\ell_1=\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega)\right.\\ &\quad +\sum_{\omega\in\Omega:\ell_1\neq\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega)\\ &\quad \left.-\sum_{\omega\in\Omega:\ell_1\neq\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega)\right)\\ &= \frac{b}{2c}\mu+\eta\sum_{\omega\in\Omega:\ell_1=\ell_2}\mathbb{P}(\omega)\ell_1(\omega)p(\omega).\end{aligned}$$

Hence, u_1 is a solution to the previous equation. By reversing the roles of 1 and 2 in the above calculations we arrive at the same set of equations for u_2:

$$\sum_{\omega\in\Omega:\ell_1=\ell_2}\mathbb{P}(\omega)\ell_2(\omega)p(\omega)p(\omega)'u_1=\frac{b}{2c}\mu+\eta\sum_{\omega\in\Omega:\ell_1=\ell_2}\mathbb{P}(\omega)\ell_2(\omega)p(\omega).$$

Because u_1 and u_2 solve the same system of equations, they are the same. The uniqueness follows from the assumption that the mean-semivariance problem has a unique solution.

Part 2: The linearity of $U(b,0)$ and $U(0,\eta)$ are an immediate consequence of $\Lambda(b,0)$ and $\Lambda(0,\eta)$ being constant functions of b and η, respectively, following from Lemma 8.

Part 3: We denote the inner product of two vectors u and v by $\langle u,v\rangle$. We next prove sufficient conditions such that $\|U(b,\eta)\|$ is strictly increasing in b. This result is equivalent to

$$\langle u_1-u_0,u_0\rangle>0,$$

where $b_1-b_0>0$, $u_1=U(b_1,\eta)$, and $u_0=U(b_0,\eta)$. The sufficiency of proving this inequality is because u_0 is collinear to the gradient of $\|\cdot\|$ evaluated at u_0. The gradient $\nabla t(u)$ provides the normal vector to the boundary of the locally ellipsoidal set E defined in

Corollary 13. The following expressions are of use:

$$
\begin{aligned}
\lambda^* &= \Lambda(b,\eta), \\
u_* &= U(b,\eta), \\
t(u) &= \frac{\|u\|(b\mu + 2c\eta\mu(\lambda^*))'u}{2cu'Q(\lambda^*)u}, \\
t(u_*) &= \|u_*\|, \\
\nabla t(u) &= \frac{(b\mu + 2c\eta\mu(\lambda^*))'u}{\|u\|2cu'Q(\lambda^*)u}u + \frac{\|u\|}{(2cu'Q(\lambda^*)u)^2}(b\mu + 2c\mu(\lambda^*)) \\
&\quad - \frac{4c\|u\|(b\mu + 2c\eta\mu(\lambda^*))'u}{(2cu'Q(\lambda^*)u)^2}Q(\lambda^*)u, \\
\nabla_u t(u_*) &= \frac{u_*}{\|u_*\|} + \frac{\|u_*\|}{2cu_*'Q(b,\eta)u_*}K^*(b\mu + 2c\eta\mu(\lambda^*)), \\
K^* &= \frac{1}{2cu_*'Q(b,\eta)u_*} - 2, \\
u_* &= U(b,\eta).
\end{aligned}
$$

The calculation of $\nabla t(u_*)$ uses two observations, first that

$$(b\mu + 2c\eta\mu(\lambda^*))'u_* = 2cu_*'Q(\lambda^*)u_*,$$

and $Q(\lambda^*)u_* = \frac{b}{2c}\mu + \eta\mu(\lambda^*)$.

We now show that $\langle u_1 - u_0, \nabla t(u_0)\rangle \geq 0$, which then implies the conditions of the proposition. Suppose $b_1 - b_0$ is sufficiently small such that $\Lambda(b_1,\eta) = \Lambda(b_0,\eta)$, then $u_1 - u_0 = \frac{b_1 - b_0}{2c}Q(\lambda^*)^{-1}\mu$.

Suppose, additionally, by contradiction, that $\langle u_1 - u_0, \nabla t(u_0)\rangle < 0$. In a neighborhood of u_0, the set E from Corollary 13 defined with respect to b_0 is an ellipse, which implies there exists $\bar{b}$ such that $\bar{b} \in (b_0, b_1)$ and $\bar{u} = U(\bar{b},\eta)$ is within the interior of E, which implies that

$$\|\bar{u}\| = T(\bar{u},\bar{b},\eta) < T(\bar{u},b_0,\eta).$$

However, by Lemma 12, $\bar{b} > b$ implies

$$T(\bar{u}, \bar{b}, \eta) > T(\bar{u}, b_0, \eta),$$

which is a contradiction.

Computing the inner product leads to

$$\left\langle u_1 - u_0, \frac{u_0}{\|u_0\|} \right\rangle + \frac{(b_1 - b_0)\|u_0\|}{(2c)^2 u_0' Q_0 u_0} K_0 (b_0 \mu + 2c\eta\mu^*)' Q_0^{-1} \mu. \tag{3.20}$$

By observing that

$$(b_0\mu + 2c\eta\mu^*)' Q_0^{-1} \mu = 2cu_0'\mu > 0,$$

the sign of the second term of equation (3.20) is then determined by the sign of K_0. If $K_0 \leq 0$, then necessarily $\langle u_1 - u_0, u_0 \rangle \geq 0$, with strict inequality holding for the latter if it holds for the former. The nonnegativity (positivity) of the inner product follows from the a priori knowledge that $\langle u_1 - u_0, \nabla t(u_0) \rangle \geq 0$. The condition that $K_0 \leq 0$ is the same as that of the proposition which completes the proof of this property. □

3.6 Ratio of Norm and Error

We investigate the behavior of the following ratio:

$$\frac{\|U(b_0, \eta_0) - U(b, \eta)\|^2}{\phi_0 - \phi(b, \eta)},$$

where $\phi(b, \eta)$ is the mean-semivariance objective with parameters b_0 and η_0 evaluated at the point $U(b, \eta)$, and ϕ_0 is the mean-semivariance evaluated at the point $u(b_0, \eta_0)$. Note that the denominator is strictly positive if $b \neq b_0$ or $\eta \neq \eta_0$. The limit as $b \to b_0$ and $\eta \to \eta_0$ does not is exist but is bounded.

Lemma 15. *The ratio,*

$$\frac{\|U(b_0,\eta_0)-U(b,\eta)\|^2}{\phi_0-\phi(b,\eta)},$$

has limits infimum and supremum as $(b,\eta)\to(b_0,\eta_0)$ *that are bounded by the minimum and maximum eigenvalues of the following matrix:*

$$(cv'Q(\lambda^*)v)^{-1}v'v,$$

where $\lambda*$ *is evaluated at* (b_0,η_0)*, and*

$$v=\begin{bmatrix}\frac{1}{2c}Q(\lambda^*)^{-1}\mu\\ Q(\lambda^*)^{-1}\mu(\lambda^*)\end{bmatrix}.$$

Proof. Let $\tilde{\Lambda}$ denote the following mapping from $\mathbb{R}^2$ to $\mathbb{R}^{|\Omega|}$:

$$\tilde{\Lambda}(b,\eta)=[\eta_0-p'U(b,\eta)].$$

In a neighborhood of (b_0,η_0), we have that

$$\Lambda(b_0,\eta_0)=\Lambda(b,\eta)=\tilde{\Lambda}(b,\eta).$$

We denote the value $\Lambda(b,\eta)$ by λ^* and $u(b_0,\eta_0)$ by u_0. Using this fact we can rewrite $u_0=v'\theta_0$ and $U(b,\eta)=v'\theta$, where

$$v=\begin{bmatrix}\frac{1}{2c}Q(\lambda^*)^{-1}\mu & Q(\lambda^*)^{-1}\mu(\lambda^*)\end{bmatrix},$$
$$\theta_0=\begin{bmatrix}b_0\\ \eta_0\end{bmatrix},$$
$$\theta=\begin{bmatrix}b\\ \eta\end{bmatrix}.$$

Using this notation in addition to the observation that

$$\phi_0 - \phi(b, \eta) = c(\theta - \theta_0)' v' Q(\lambda^*) v(\theta - \theta_0).$$

The limit is given as

$$\lim_{\theta \to \theta_0} \frac{(\theta - \theta_0)' v' v(\theta - \theta_0)}{c(\theta - \theta_0)' v' Q(\lambda^*) v(\theta - \theta_0)}.$$

By fixing $\eta = \eta_0$ we arrive at the following limit:

$$\frac{\mu' Q(\lambda^*)^{-1} Q(\lambda^*)^{-1} \mu}{c\mu' Q(\lambda^*)^{-1} \mu}.$$

By fixing $b = b_0$ we arrive at the following limit:

$$\frac{\mu(\lambda^*)' Q(\lambda^*)^{-1} Q(\lambda^*)^{-1} \mu(\lambda^*)}{c\mu(\lambda^*)' Q(\lambda^*)^{-1} \mu(\lambda^*)}.$$

Because we arrive at different limits by approaching (b_0, η_0) along different paths, the limit does not exist.

Because the ratio is that of two strictly convex quadratic functions, we may apply the well known results due to Gantmacher [55] to show the following. Let β_n and β_1 denote the maximum and minimum eigenvalues of the following matrix,

$$(cv' Q(\lambda^*) v)^{-1} v' v.$$

The following holds:

$$0 \leq \beta_1 \leq \frac{(\theta - \theta_0)' v' v(\theta - \theta_0)}{c(\theta - \theta_0)' v' Q(\lambda^*) v(\theta - \theta_0)} \leq \beta_n,$$

for all θ such that $\theta - \theta_0 \neq 0$. Hence the limit supremum and infimum must also satisfy the inequality. □

The ratio has a unique limit as $\eta \to \infty$ for any fixed b. The existence of such a limit is a direct consequence of Lemma 8. The following lemma gives the value of the limit.

Lemma 16. *For any b the following limit exists:*

$$\lim_{\eta\to\infty} \frac{\|U(b_0,\eta_0) - U(b,\eta)\|^2}{\phi_0 - \phi(b,\eta)} = \frac{\mu_H Q_H^{-1} Q_H^{-1} \mu_H}{c\mu_H Q_H^{-1} \tilde{Q}_H Q_H^{-1} \mu_H}.$$

The limit is approached monotonically from below if $\rho > 0$ and from above if $\rho < 0$, where ρ is given by the following expression:

$$\rho = a_2 b_1 - a_1 b_2,$$

where

$$\begin{aligned}
a_1 &= \left(\frac{b}{c}\mu' Q_H^{-1} - 2u(b_0,\eta_0)\right) Q_H^{-1}\mu_H,\\
a_2 &= \mu_H Q_H^{-1} Q_H^{-1} \mu_H,\\
b_1 &= \left(b\mu' Q_H^{-1}\tilde{Q}_H - b_0\mu' - 2c\eta_0\tilde{\mu}'_H\right) Q_H^{-1}\mu_H,\\
b_2 &= c\mu'_H Q_H^{-1} Q(\tilde{\lambda}^H) Q_H^{-1}\mu_H.
\end{aligned}$$

Proof. Following from Lemma 8, we know that λ^H is the value of $\Lambda(b,\eta)$ as $\eta \to \infty$. We compute the limit by determining the quadratic terms (as a function of η) of the numerator and denominator, and then by applying l'Hôpital's rule we can determine the limit by the ratio of the coefficients.

First we calculate the quadratic expression for $\|u(b_0,\eta_0) - U(b,\eta)\|^2$ as

$$\begin{aligned}
\|U(b_0,\eta_0) - U(b,\eta)\|^2 = U(b_0,\eta_0)'U(b_0,\eta_0) - 2U(b_0,\eta_0)'U(b,\eta)\\
+ U(b,\eta)'U(b,\eta).
\end{aligned}$$

The quadratic term can only come from $U(b,\eta)'U(b,\eta)$, so we focus on that term:

$$\begin{aligned} U(b,\eta)'U(b,\eta) = &\left(\frac{b}{2c}\right)^2 \mu' Q_H^{-1} Q_H^{-1} \mu \\ &+ \frac{b\eta}{c} \mu' Q_H^{-1} Q_H^{-1} \mu_H \\ &+ \eta^2 \mu_H Q_H^{-1} Q_H^{-1} \mu_H. \end{aligned}$$

The above calculation shows that the quadratic term for $\|U(b_0,\eta_0) - U(b,\eta)\|^2$ is

$$\eta^2 \mu_H Q_H^{-1} Q_H^{-1} \mu_H.$$

Second, we calculate the quadratic term for $\phi_0 - \phi(b,\eta)$. Necessarily, the quadratic term can only come from $\phi(b,\eta)$, so we focus on that term. We use the first part of Lemma 9 to observe that the semivariance is evaluated at $\tilde{\lambda}$ in the limit as $\eta \to \infty$. Therefore,

$$\begin{aligned} &-\phi(b,\eta) \\ &= -b_0 r_0 x + c\eta_0^2 E\tilde{\lambda} \\ &\quad - (b_0 \mu + 2c\eta_0 \mu(\tilde{\lambda}))' U(b,\eta) \\ &\quad + cU(b,\eta)'\tilde{Q}U(b,\eta). \end{aligned}$$

The function $U(b,\eta)$ is linear in η, so the term $cU(b,\eta)'\tilde{Q}U(b,\eta)$ is the only term quadratic in η. We focus on the following term:

$$\begin{aligned} &cU(b,\eta)'\tilde{Q}U(b,\eta) \\ &= c\left(\frac{b}{2c}\right)^2 \mu Q_H^{-1} q(\tilde{\lambda}) Q_H^{-1} \mu \\ &\quad + c\left(\frac{b\eta}{c}\right) \mu Q_H^{-1} q(\tilde{\lambda}) Q_H^{-1} \mu_H \\ &\quad + c\eta^2 \mu_H Q_H^{-1} q(\tilde{\lambda}) Q_H^{-1} \mu_H. \end{aligned}$$

The above calculation shows that the quadratic term for $\phi_0 - \phi(b, \eta)$ is

$$c\eta^2 \mu'_H Q_H^{-1} \tilde{Q} Q_H^{-1} \mu_H.$$

Applying l'Hôpital's rule, we calculate the limit as the ratio of the two quadratic coefficients. Using the second part of Lemma 9, we replace $\tilde{\lambda}$ by $\tilde{\lambda}^H$ in the resulting ratio.

From the analysis of the limit we have rewritten $\|U(b_0, \eta_0) - U(b, \eta)\|^2$ in the form $a_0 + a_1\eta + a_2\eta^2$ and $\phi_0 - \phi(b, \eta)$ in the form $b_0 + b_1 + b_2$. The values of a_1, a_2, b_1, and b_2 are as given in the statement of the lemma. The expressions for the constant terms a_0 and b_0 are unnecessary for what follows.

The monotonic approach of the limit comes directly from the following two observations. First, we rewrite the ratio in the following form:

$$\frac{a_0 + a_1\eta + a_2\eta^2}{b_0 + b_1\eta + b_2\eta^2}.$$

Second, this ratio has the following derivative:

$$\frac{(a_1b_0 - a_0b_1) + 2(a_2b_0 - a_0b_2)\eta + (a_2b_1 - a_1b_2)\eta^2}{(b_0b_1\eta + b_2\eta^2)^2}.$$

Because the denominator is strictly positive for $\eta \neq \eta_0$, the sign of the derivative as $\eta \to \infty$ is the same as the sign of $a_2b_1 - a_1b_2$ as $\eta \to \infty$. Continuity combined with the derivative being strictly positive or strictly negative guarantees the limit is approached monotonically. □

We now investigate the behavior of the ratio near $b = 0$ or $\eta = 0$. Using the property that $U(b, 0)$ and $U(0, \eta)$ are linear from Lemma 14 we see that $\|U(b_0, \eta_0) - U(b, 0)\|^2$ and $\|U(b_0, \eta_0) - U(0, \eta)\|^2$ are quadratic. Hence, the ratio becomes the ratio of a quadratic function divided by a piecewise quadratic function. In certain special cases, $\phi(b, 0)$ and $\phi(0, \eta)$ are also quadratic. The following lemma summarizes these properties.

Lemma 17. *Suppose that* $\eta_0 = 0$. *Then it follows that* $\phi_0 - \phi(b, 0)$ *and* $\phi_0 - \phi(0, \eta)$ *are quadratic functions of* b *and* η, *respectively.*

Proof. When $\eta_0 = 0$, we observe that the following indicator function:

$$\tilde{\Lambda}(b, 0) = [p(\omega)' U(b, 0) \geq 0],$$

satisfies $\tilde{\Lambda}(b, 0) = \lambda^B$ for all b. The indicator function,

$$\tilde{\Lambda}(0, \eta) = [p(\omega)' U(0, \eta) \geq 0],$$

is constant with respect to η. Let this constant value be denoted by $\tilde{\lambda}^\eta$. Because $U(b, 0)$ and $U(0, \eta)$ are linear functions and because $\tilde{\Lambda}(b, 0)$ and $\tilde{\Lambda}(0, \eta)$ are constant functions, it follows that $\phi(b, 0)$ and $\phi(0, \eta)$ are quadratic functions. □

Proposition 18. *The norm ratio,*

$$\frac{\|U(b_0, \eta_0) - U(b, \eta)\|^2}{\phi_0 - \phi(b, \eta)},$$

is bounded for all b *and* η *and continuous for* $(b, \eta) \neq (b_0, \eta_0)$.

Proof. Continuity and boundedness follows from the denominator being bounded away from zero for $(b, \eta) \neq (b_0, \eta_0)$, and by Lemma 15. □

3.6.1 Transaction Cost Interpretation of Norm Ratio

A practical application of the norm-error ratio is in terms of transaction costs. Suppose that changing from portfolio u to v causes a change in market price proportional to $\|u - v\|$. This implies that changing the portfolio from u to v incurs a cost to the investor proportional to $\|u - v\|^2$.

Suppose that we wish to change our portfolio parameters from (b, η) to (b_0, η_0). Further, assume that we are currently invested in portfolio $U(b, \eta)$. The improvement that is possible due to changing the portfolio to the new optimal portfolio is $\phi_0 - \phi(b, \eta)$. Suppose the cost of changing the portfolio is $c_t \|u_0 - U(b, \eta)\|^2$, then we would only benefit from changing portfolios if the transaction cost proportionality constant c_t satisfies

$$\frac{c_t \|u_0 - U(b, \eta)\|^2}{\phi_0 - \phi(b, \eta)} < 1.$$

We can therefore draw a few conclusions about requirements on transaction cost coefficients.

- Because the ratio of the norm and mean-semivariance error is bounded, there exists a threshold such that a transaction cost coefficient greater than this threshold implies the portfolio would never be rebalanced due to a change in parameters.
- Lemma 16 provides a bound on the transaction cost threshold above which the portfolio would not be rebalanced.
- These thresholds may have an impact on dynamic portfolio optimization in which the parameter η varies due to variations in total wealth.

Numerically, the norm of $U(b, \eta)$ behaves as a convex function monotone in b, which also has implications with respect to transaction costs. Convexity implies that if we choose to create a portfolio using the parameters of two other investors with parameters, (b_1, η_1) and (b_2, η_2) such that $(b, \eta) = \alpha(b_1, \eta_1) + (1 - \alpha)(b_2, \eta_2)$, then the cost of constructing the portfolio $U(b, \eta)$ is less than the convex combination of the costs of constructing the portfolios $U(b_1, \eta_1)$ and $U(b_2, \eta_2)$, i.e.

$$c_t \|U(b, \eta)\|^2 \leq \alpha c_t \|U(b_1, \eta_1)\|^2 + (1 - \alpha) c_t \|U(b_2, \eta_2)\|^2.$$

Table 3.1: Average value of limiting ratio divided by the observed maximum ratio.

No. Assets	No. Samples		
	200	600	1000
5	0.8181	0.8278	0.8495
10	0.6860	0.6949	0.7638
20	0.6054	0.5979	0.6101

3.6.2 Numerical Evaluation of Norm-Error Ratio

The importance of studying the ratio between the norm squared and the value function error is two fold. First the norm squared approximates the shape of the value function error for optimal solutions with differing parameters. Second, this characterization of error as being approximately proportional to the norm squared of the portfolio has implications toward normalizing mean-semivariance portfolios by using a norm penalty function.

The first relationship we investigate is the relationship between the maximum of the ratio and the limit from Lemma 16. The values in Table 3.1 show the limit from Lemma 16 divided by the observed maximum of the ratio. The values of the ratio decrease as the number of assets increases, indicating the limit becomes closer to lower bounding the ratio as the number of assets increases. The second relationship we study is that the norm, when rescaled by the limit from Lemma 16, is a good approximation for the change in objective function. The plot in Figure 3.1 shows the mean absolute deviation of the change in norm and the actual objective value change for different values of the mean coefficient b with the average taken over values of η. The observed error is approximately quadratic in a neighborhood of $b_0 = 1$.

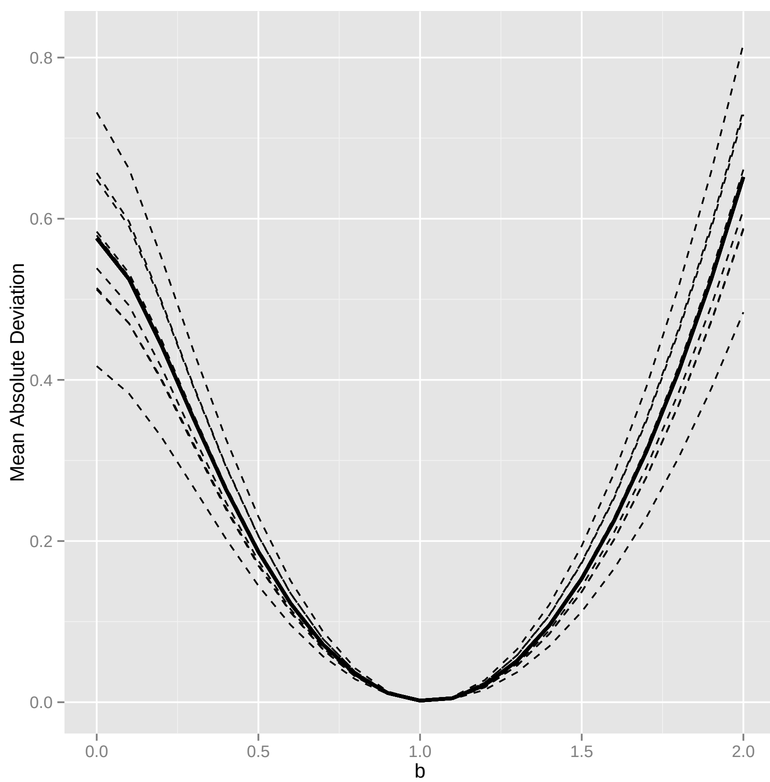

Figure 3.1: The dashed lines show the error, measured by mean absolute deviation with respect to different values of η, in approximating the change in objective by the norm, rescaled by the limit in Lemma 16.

CHAPTER 4

4.1 Efficient Frontier and Multi-Period Portfolio Problem

We establish in this chapter an application of parametric piecewise-quadratic methods to the mean-semivariance portfolio problem. We focus on applications relevant to calculating the efficient frontier of the single-period model and the value functions required to solve the multi-period mean-semivariance portfolio problem. We first describe a parametric expression for the mean-semivariance efficient frontier. The resulting algorithm is conceptually equivalent to the critical line algorithm of Markowitz [3]. However, the resulting explicit formula for the efficient frontier is novel. The formula provides the foundation for further analysis of optimal portfolio characteristics, and it provides practitioners with a means for efficient storage and calculation of these functions. We then develop a two-period stochastic optimization model to motivate the design of a parametric method for the multi-period mean-semivariance portfolio problem. We present the parametric formulas for the terminal period of the multi-period problem. We then generalize the procedure to accommodate the multi-period problem.

4.2 Parametric Solution to Mean-Semivariance Efficient Frontier

Efficient frontiers provide investors with a complete range of options. We propose a method to compute the mean-semivariance efficient frontier using parametric piecewise quadratic optimization based on ideas similar to those in the work of Patrinos and Sarimveis [22]. The result is essentially an alternative derivation of Markowitz's [3] critical line algorithm. The efficiency of this type of algorithm is made clear by Niedermayer and Niedermayer [20]. Our analysis and reformulation, however, provide explicit formulas for the parametric description of the efficient frontier, and we develop an algorithm on the basis of this description.

For the sake of simplicity we set the risk aversion parameter c to 1, and we paramterize the objective in terms of the mean weight b only. We first observe that the mean-semivariance objective function can be written in the following way:

$$f(u, b) = f_k(u, b) := \eta^2 \pi_k - b(r_0 x + \mu' u) - 2\eta \mu_k' u + u' Q_k u,$$

where the first equality is true whenever u lies in the polyhedral region C_k. The region C_k is defined by

$$C_k = \{u : p(\omega)' u \leq \eta,\, \omega \in \Omega_k^1,\, p(\omega)' u \geq \eta,\, \omega \in \Omega_k^2\},$$

where Ω_k^1 and Ω_k^2 partition Ω and $k \in K$ indexes all such partitions of two sets. We use the following notation,

$$\begin{aligned} \lambda_k(\omega) &= [\eta - p(\omega)' u \geq 0], \\ \pi_k &= \mathrm{E}(\lambda_k), \\ \mu_k &= \mathrm{E}(\lambda_k p), \\ Q_k &= \mathrm{E}(\lambda_k p p'). \end{aligned}$$

Values of u lying on the boundary of multiple regions $C_{k'}$ for k' in some subset K' of K the functions $f_{k'}$ evaluated at u are all equal, as are their gradients. This is a consequence of the mean-semivariance objective being continuously differentiable. Let the optimal solution mapping as a function of the mean weight b be denoted by

$$U(b) = \arg\min_u f(u, b).$$

For the following analysis we suppose that the mapping U is single valued for all b, which greatly simplifies the notation involved. The multi-valued cases are natural generalizations of the approaches discussed here. Let K^* denote the following set:

$$K^* = \{k \in K : \exists b, U(b) \in C_k\}.$$

Following a similar notation to Patrinos and Sarimveis [22] we define critical regions.

Definition 19. The critical region R_k associated with $k \in K$ is the following set:

$$R_k = \{b : U(b) \in C^k\}.$$

Observe that $R_k = \emptyset$, if $k \notin K^*$. We now provide formulas for the non-empty critical regions.

Lemma 20. *Let $k \in K^*$. The critical region associated with k is the following interval:*

$$R_k = \left[\max_{\omega \in \bar{\Omega}_k^2 \cup \bar{\Omega}_k^3} \frac{\psi_k(\omega)}{\phi_k(\omega)}, \min_{\omega \in \bar{\Omega}_k^1 \cup \bar{\Omega}_k^4} \frac{\psi_k(\omega)}{\phi_k(\omega)}\right], \tag{4.1}$$

where

$$\begin{aligned}
\phi^k(\omega) &= \frac{1}{2} p(\omega)' Q_k^{-1} \mu, \\
\psi^k(\omega) &= \eta(1 - p(\omega)' Q_k^{-1} \mu_k), \\
\tilde{\Omega}_k^1 &= \{\omega : \phi_k(\omega) > 0\}, \\
\tilde{\Omega}_k^2 &= \{\omega : \phi_k(\omega) < 0\}, \\
\bar{\Omega}_k^1 &= \Omega_k^1 \cap \tilde{\Omega}_k^1, \\
\bar{\Omega}_k^2 &= \Omega_k^1 \cap \tilde{\Omega}_k^2, \\
\bar{\Omega}_k^3 &= \Omega_k^2 \cap \tilde{\Omega}_k^1, \\
\bar{\Omega}_k^4 &= \Omega_k^2 \cap \tilde{\Omega}_k^2.
\end{aligned}$$

Proof. Equation (4.1) follows immediately by solving the inequalities resulting from $U(b) \in$

C_k for b. $\square$

Because we are only interested in $b \geq 0$, from this point forward we make the substitution $R_k \equiv R_k \cap [0, \infty)$.

The next lemma shows that given a critical region associated with index $k \in K^*$ how to determine adjacent critical regions associated with indices $k^- \in K^*$ and $k^+ \in K^*$. Two critical regions are adjacent if their intersection is a facet of both critical regions.

Lemma 21. *Let $k^* \in K^*$ be given and $U(b)$ be single valued. The associated critical region can be written $R_{k^*} = [b_\ell, b_u]$. If $b_u \neq \infty$, then*

$$\Omega^* = \arg\min_{\omega \in \bar{\Omega}_k^1 \cup \bar{\Omega}_k^4} \frac{\psi_{k^*}(\omega)}{\phi_{k^*}(\omega)}$$

is nonempty. If $\Omega^* = \{\omega^+\}$, then there exists k^+ such that $R_{k^+} = [b_u, b_{u^+}]$, where k^+ is associated with the following partition of Ω:

$$\omega^+ \in \Omega_{k^*}^1 \Rightarrow \begin{cases} \Omega_{k^+}^1 = \Omega_{k^*}^1 \backslash \{\omega^+\}, \\ \Omega_{k^+}^2 = \Omega_{k^*}^2 \cup \{\omega^+\}. \end{cases}$$

$$\omega^+ \in \Omega_{k^*}^2 \Rightarrow \begin{cases} \Omega_{k^+}^1 = \Omega_{k^*}^1 \cup \{\omega^+\}, \\ \Omega_{k^+}^2 = \Omega_{k^*}^2 \backslash \{\omega^+\}. \end{cases}$$

If Ω^* is not a singleton then k^+ is associated with the partition satisfying the following three requirements.

1. *$\Omega_{k^+}^1 \backslash \omega^* = \Omega_{k^*}^1 \backslash \omega^*$ and $\Omega_{k^+}^2 \backslash \omega^* = \Omega_{k^*}^2 \backslash \omega^*$.*

2. *For each $\omega \in \Omega^* \cap \Omega_{k^+}^1$ the following must hold:*

$$p(\omega)' Q_{k^+}^{-1} \mu \leq 0.$$

3. *For each* $\omega \in \Omega^* \cap \Omega^2_{k^+}$ *the following must hold:*

$$p(\omega)' Q_{k^+}^{-1} \mu \geq 0.$$

Similarly, if $b_\ell \neq 0$ *there exists* $k^- \in K^*$ *such that* $R_{k^-} = [b_{\ell^-}, b_\ell]$*, where* k^- *is associated the the same type of partition as* k^+ *but with*

$$\Omega^* = \arg\max_{\omega \in \bar{\Omega}^2_k \cup \bar{\Omega}^3_k} \frac{\psi_{k^*}(\omega)}{\phi_{k^*}(\omega)}.$$

Proof. We prove the case when k^+ exists. The result is analogous for k^-. When Ω^* is a singleton, the result follows immediately by recalling the set Ω^* is the set of inequalities that are tight at $U(b_u)$, so by the continuity of U, the region associated with the adjacent critical region must be the one resulting from swapping the violated inequality.

Now suppose that Ω^* is not a singleton. Condition 1 in the lemma indicates that k^+ must be one of the regions C_k, $k \in \tilde{K}$, adjacent to the point $U(b^*)$ where the $|\Omega^*|$ inequalities intersect. Let $U_k(b)$ denote the solution mapping of the quadratic functions f_k associated with the adjacent regions $\tilde{K}$. We now recall that $U_k(b) = U(b)$ if and only if $U_k(b) \in C_k$. This is equivalent to the following statement: there exists $\varepsilon > 0$ such that for all $\delta \in [0, \varepsilon]$, $U_k(b^* + \delta) \in C_k$. Furthermore, this expression is equivalent to $\nabla U_k(b^*) \in T_{C_k}(u^*)$, where $T_{C_k}(u^*)$ denotes the the tangent cone of C_k at u^*, which is defined by

$$T_{C_k}(u^*) = \{w : p(\omega)'w \leq 0,\, \omega \in \Omega^1_k \cap \Omega^*,\, p(\omega)'w \geq 0,\, \omega \in \Omega^2_k \cap \Omega^*\}.$$

The equivalency holds because U_k is linear and C_k is polyhedral. Observing that $\nabla U_k(b^*) = \frac{1}{2} Q_k^{-1} \mu$, we see that conditions 2 and 3 are equivalent to $\nabla U_k(b^*) \in T_{C_k}(u^*)$, which completes the proof. It is worth noting that, while condition 1 is actually redundant, it is of practical value for identifying the adjacent regions. □

Lemma 22. *The critical regions* R_k *for* $k \in K^*$ *form a polyhedral subdivision of* $[0, \infty)$.

Proof. Because $U(b)$ is defined for all b, the mapping from b to $k \in K^*$ such that $U(b) \in C_k$ must be surjective. Hence,

$$U^{-1}(K^*) = \bigcup_{k \in K^*} R_k = [0, \infty).$$

The proof of Lemma 21 actually shows that $U(b)$ enters and leaves each C_k for $k \in K^*$ exactly once. The relative interiors of the critical regions of R_k are equivalent $U(b)$ being in the relative interior of C_k. Because the $\mathcal{C} = \{C_k : k \in K\}$ is a polyhedral subdivision of $\mathbb{R}^n$, $U(b)$ being single valued does not permit it to take values simultaneously in distinct C_k. Hence the relative interiors of R_k are disjoint. Each interval is full dimensional. Hence $\mathcal{R} = \{R_k : k \in K^*\}$ is a polyhedral decomposition of $[0, \infty)$. Observing that the intersection of two critical regions is either empty or an end point of the two regions, we conclude that $\mathcal{R}$ is a polyhedral subdivision of $[0, \infty)$. □

Using these lemmas we can now prove a result that gives a procedure for inductively calculating the efficient frontier and the associated efficient portfolios.

Proposition 23. *Let $K^* = \{0, 1, \ldots, N\}$ be such that $R_0 = [0, b_1]$, $R_1 = [b_1, b_2]$, $\ldots$, $R_N = [b_N, \infty]$, with $b_1 < b_2 < \cdots < b_N$. The following recursive formula holds for $U(b)$:*

1. *If $b \in R_0$, $U(b) = b\frac{1}{2}Q_0^{-1}\mu + \eta Q_0^{-1}\mu_0$.*

2. *If $b \in R_k$, $U(b) = (b - b_k)\frac{1}{2}Q_k^{-1}\mu + U(b_k)$; or equivalently,*

3. *If $b \in R_k$, $U(b) = \frac{1}{2}\Big[(b - b_k)Q_k^{-1}\mu + \sum_{\ell=0}^{k-1} Q_\ell^{-1}\mu(b_{\ell+1} - b_\ell)\Big] + \eta Q_0^{-1}\mu_0$, where $b_0 = 0$.*

The value function $V(b) := \min_u f(u, b)$ is given by the following recursive formulas. Let $u_k := U(b_k)$.

1. *$V(0) = \eta^2(\pi_0 - \mu_0 Q_0^{-1}\mu_0)$.*

2. *$V(b_k) = f(u_k, b_k)$.*

3. If $b \in R_k$, $V(b) = V(b_k) - (r_0 x + \mu' u_k)(b - b_k) - \frac{1}{4}\mu' Q_k^{-1}\mu(b - b_k)^2$.

Finally, let $S(b)$ *and* $M(b)$ *denote the semivariance and mean of* $U(b)$*, respectively. We denote* $S_k := S(b_k)$ *and* $M_k := M(b_k)$*. The following formula suffices to compute the mean-semivariance efficient frontier* $\{(M(b), S(b)) : b \geq 0\}$*. For* $b \in R_k$*, the following formulas hold:*

$$M_k = r_0 x + \mu' u_k, \tag{4.2}$$

$$S_k = \sum_{\omega \in \Omega} (\eta - p(\omega)' u_k)_+^2, \tag{4.3}$$

$$M(b) = M_k + \frac{1}{2}\mu' Q_k^{-1}\mu(b - b_k), \tag{4.4}$$

$$S(b) = S_k + \mu'(u_k - \eta Q_k^{-1}\mu_k)(b - b_k) + \frac{1}{4}\mu_k' Q_k^{-1}\mu_k(b - b_k)^2. \tag{4.5}$$

Using equation (4.5)*, we rewrite* S *as a piecewise quadratic function of* M*:*

$$\begin{aligned} S(M) &= S_k + \frac{2\mu'(u_k - \eta Q_k^{-1}\mu_k)}{\mu' Q_k^{-1}\mu}(M - M_k) + \frac{\mu_k' Q_k^{-1}\mu_k}{(\mu' Q_k^{-1}\mu)^2}(M - M_k)^2, \\ k &= \sup\{k' : M_{k'} \leq M\}. \end{aligned}$$

Proof. The recursive formulas for $U(b)$ can be proven by the following computation. For $b \in R_k = [b_k, b_{k+1}]$,

$$U(b) = U(b_k) + \nabla U(b_k)(b - b_k),$$

which follows because U is linear on R_k. The proof of the third formula for $U(b)$ follows by repeatedly applying this expansion to yield

$$U(b) = U(0) + \nabla U(b_k)(b - b_k) + \sum_{i=0}^{k-1} \nabla U(b_k)(b_{i+1} - b_i).$$

We provide the formula for the value function from the following calculation. Let

$b \in R_k$. We rely on the fact that $\pi(U(b)) = \pi_k$, $\mu(U(b)) = \mu_k$, and $Q(U(b)) = Q_k$:

$$
\begin{aligned}
V(b) &= \eta^2 \pi_k - b r_0 x - b\mu' U(b) - 2\eta \mu_k' U(b) + U(b)' Q_k U(b) \\
&= \eta^2 \pi_k - b_k r_0 x - b_k \mu' (\Delta b \frac{1}{2} Q_k^{-1} \mu + u_k) - \Delta b r_0 x \\
&\quad - \Delta b \mu' (\Delta b \frac{1}{2} Q_k^{-1} \mu + u_k) - 2\eta \mu_k' (\Delta b \frac{1}{2} Q_k^{-1} \mu) - 2\eta \mu_k' u_k \\
&= V(b_k) - (r_0 + \mu' u_k) \Delta b - (\frac{1}{4} \mu' Q_k^{-1} \mu) \Delta b^2,
\end{aligned}
$$

which gives the desired formula for the value function. We note that all three of the formulas for $V(b)$ given in the statement of the proposition are a consequence of substituting different values into the given expression. We now simplify to give the formulas for the efficient frontier. Let M_k and S_k denote the mean and semivariance of u_k. Using the recursive formula for $U(b)$, we arrive at the given formula for $M(b)$. The formula for $S(b)$ comes from a direct calculation of the semivariance of $U(b_k + (b - b_k))$. The final formula for $S(M)$ is given by solving equation (4.4) for $(b - b_k)$ and substituting the result into equation (4.5), giving

$$
S(M^*) = S_k + \frac{2\mu'(u_k - \eta Q_k^{-1} \mu_k)}{\mu' Q_k^{-1} \mu} (M - M_k) + \frac{\mu_k' Q_k^{-1} \mu_k}{(\mu' Q_k^{-1} \mu)^2} (M - M_k)^2. \tag{4.6}
$$

We note that the k in equation (4.6) is the same as $\sup\{k' : M_{k'} \leq M^*\}$. □

The proposition implies an iterative procedure for exactly computing the efficient frontier. We present this iterative procedure for the case where Ω^* is a singleton at all iterations. The algorithm only needs to be modified to check all candidate regions C_k to handle this additional case. We now prove Algorithm 2 is correct and produces an exact representation of the efficient frontier.

Proposition 24. *The finite sequences u_k, M_k, S_k, and $\mu' Q_k^{-1} \mu$ resulting from Algorithm 2 completely determine the efficient frontier of the mean-semivariance portfolio problem.*

Proof. Observing that the calculations for u_k, M_k, S_k are correct, by Proposition 23 it is

Algorithm 2 Computes the mean-semivariance efficient frontier and the associated efficient portfolios.

$u_0 \leftarrow U(0) = \eta Q_0^{-1}\mu_0$
$\lambda(\omega) \leftarrow [\eta - p(\omega)'u_0 \geq 0]$ for all $\omega \in \Omega$.
$\phi_0(\omega) \leftarrow \frac{1}{2}p(\omega)'Q_0^{-1}\mu$ for all $\omega \in \Omega$.
$\psi_0(\omega) \leftarrow \eta(1 - p(\omega)'Q_0^{-1}\mu_0) = p(\omega)'u_0$ for all $\omega \in \Omega$.
$M_0 \leftarrow r_0 x + \mu' u_0 = r_0 x + \sum_{\omega\in\Omega} \mathbb{P}(\omega)\psi_0(\omega)$.
$S_0 \leftarrow \sum_{\omega\in\Omega} \mathbb{P}(\omega)(\eta - p(\omega)'u_0)^2\lambda(\omega)$.
Compute

$$b_1 \leftarrow \min_{\omega\in\bar{\Omega}_1^0\cap\bar{\Omega}_4^0} \left\{ \frac{\psi_0(\omega)}{\phi_0(\omega)} \right\}$$
$$\omega_1 \leftarrow \arg\min_{\bar{\Omega}_1^0\cap\bar{\Omega}_4^0} \left\{ \frac{\psi_0(\omega)}{\phi_0(\omega)} \right\}$$

Set $\lambda(\omega_1) \leftarrow 1 - \lambda(\omega_1)$.
Let $k \leftarrow 1$.
loop
    $Q_k^{-1} \leftarrow Q(\lambda)^{-1}$.
    $\mu_k \leftarrow \mu_{k-1} + (2\lambda(\omega) - 1)\mathbb{P}(()\omega_k)p(\omega_k)$
    Compute $Q_k^{-1}\mu$ and $Q_k^{-1}\mu_k$.
    $u_k \leftarrow \frac{b_k}{2}Q_k^{-1}\mu + \eta Q_k^{-1}\mu_k$.
    $\phi_k(\omega) \leftarrow \frac{1}{2}p(\omega)'Q_k^{-1}\mu$ for all $\omega \in \Omega$.
    $\psi_k(\omega) \leftarrow \eta(1 - p(\omega)'Q_k^{-1}\mu_k) = p(\omega)'u_k$ for all $\omega \in \Omega$.
    $M_k \leftarrow r_0 x + \mu' u_k$
    $S_k \leftarrow \sum_{\omega\in\Omega} \mathbb{P}(\omega)(\eta - p(\omega)'u_k)^2\lambda(\omega)$.
    Compute

$$b_{k+1} \leftarrow \min_{\omega\in\bar{\Omega}_1^k\cap\bar{\Omega}_4^k} \left\{ \frac{\psi_k(\omega)}{\phi_k(\omega)} \right\}$$
$$\omega_{k+1} \leftarrow \arg\min_{\bar{\Omega}_1^k\cap\bar{\Omega}_4^k} \left\{ \frac{\psi_k(\omega)}{\phi_k(\omega)} \right\}$$

    If $b_{k+1} = +\infty$, end algorithm.
    Set $\lambda(\omega_{k+1}) \leftarrow 1 - \lambda(\omega_{k+1})$ and $k \leftarrow k + 1$.

sufficient to demonstrate that the algorithm correctly enumerates all the critical regions. The calculation of λ is equivalent to selecting the next region $k \in K^*$ according to the procedure in Lemma 21. Let $R_0 = [0, b_0]$ denote the first critical region calculated. By Lemma 21, the next critical region calculated is given by $R_1 = [b_0, b_1]$. Again by Lemma 21, given $R_k = [b_k, b_{k+1}]$ the next critical region is given by $R_{k+1} = [b_{k+1,k+2}]$. Because K^* is finite, the terminating condition must be reached in a finite number of iterations. By the terminating condition of the algorithm the last critical region is given by $R_N = [b_N, \infty]$; therefore, $\bigcup_{i=0}^{N} R_i = [0, \infty)$.

In the case that Ω^* is not single valued then the algorithm continues to work by replacing the calculation of ω_1 by a loop that checks conditions 2 and 3 of Lemma 21 for all candidate regions $\tilde{K}$. □

A few remarks about Algorithm 2 are worthwhile toward its efficient implementation. In the case where Ω^* is a singleton for all iterations, the matrix Q_{k+1} is a rank-1 modification relative to Q_k, i.e. there exists $\omega^* \in \Omega$ such that $Q_{k+1} = Q_k \pm p(\omega^*)p(\omega^*)'$, with a plus sign corresponding to an update and a minus sign corresponding to a downdate. Hence, the efficiency of the algorithm depends on an efficient algorithm for performing rank-1 modifications of the inverses Q_k^{-1} that need to be computed throughout the algorithm. This an issue similar to the computational issue faced by the implementation of the critical line algorithm described by Markowitz et al. [13].

4.3 Two-Period Mean-Semivariance Portfolio Problem

The general form of the two-period mean-semivariance problem is the following:

$$\begin{aligned} \max_{u_0, u_1} \quad & \mathrm{E}(x_2) - c\mathrm{E}(h - x_2)_+^2, \\ \text{s. t.} \quad & x_2 = r_2 x_1 + p_2' u_1, \\ & x_1 = r_1 x_0 + p_1' u_0. \end{aligned}$$

This optimization problem has the following large scale formulation by substituting the values for x_2 and x_1:

$$\max_{u_0,u_1} \mathrm{E}(r_2 r_1 x_0 + r_2 p_1' u_0 + p_2' u_1) - c\mathrm{E}(h - r_2 r_1 x_0 - r_2 p_1' u_0 - p_2' u_1)_+^2.$$

However, such a formulation requires defining a set of second period decision variable for each first period sample path. By using dynamic programming we can significantly reduce the number of variables that must be treated explicitly.

First we define the optimization problems for the first and second periods, recursively:

$$\begin{aligned} J_2(x_1) &= \max_{u_1} \mathrm{E}(r_2 x_1 + p_2' u_1) - c\mathrm{E}(h - r_2 x_1 - p_2' u_1)_+^2, \\ J_1(x_0) &= \max_{u_0} \mathrm{E}(J_2(r_1 x_0 + p_1' u_0)). \end{aligned}$$

Here J_i is the value function for period i. Solving the dynamic programming formulation then reduces to developing an expression for the value function J_2.

To draw parallels later, it is useful to rewrite the second period problem in the following way:

$$\begin{aligned} \max_{u_1} \quad & \mathrm{E}(-ch^2\Lambda(x_2)) + \mathrm{E}((1+2ch)\Lambda(x_2)x_2) - \mathrm{E}(c\Lambda(x_2)x_2^2), \\ \text{s.t.} \quad & x_2 = r_2 x_1 + p_2' u_1. \end{aligned}$$

Here $\Lambda(x_2) = [h - x_2 \geq 0]$ is a mapping from $\mathbb{R}$ to $\mathbb{R}^{|\Omega|}$. This form is equivalent to the following problem:

$$\begin{aligned} \max_{u_1} \quad & \mathrm{E}(\tilde{\gamma}_2(x_2)) + \mathrm{E}(\tilde{\alpha}_2(x_2)x_2) - \mathrm{E}(\tilde{\beta}_2(x_2)x_2^2), \\ \text{s.t.} \quad & x_2 = r_2 x_1 + p_2' u_1. \end{aligned}$$

The three random variables $\tilde{\gamma}_2$, $\tilde{\alpha}_2$, and $\tilde{\beta}_2$ are piecewise constant according to the following definitions:

$$\tilde{\gamma}_2(x_2) = -ch^2\Lambda(x_2),$$
$$\tilde{\alpha}_2(x_2) = 1 + 2ch\Lambda(x_2),$$
$$\tilde{\beta}_2(x_2) = c\Lambda(x_2).$$

Using the same short hand notation as in the single period formulation, we have the following:

$$\gamma_2(x, u) = \tilde{\gamma}(r_2 x + p_2' u),$$
$$\alpha_2(x, u) = \tilde{\alpha}(r_2 x + p_2' u),$$
$$\beta_2(x, u) = \tilde{\beta}(r_2 x + p_2' u).$$

In terms of u_1, the second period optimization problem can be rewritten in the following way:

$$\max_{u_1} \mathrm{E}(\gamma_2(x_1, u_1) + \alpha_2(x_1, u_1) r_2 x_1 - \beta_2(x_1, u_1)(r_2 x_1)^2)$$
$$+ \mathrm{E}((\alpha_2(x_1, u_1) - 2\beta_2(x_1, u_1) r_2 x_1) p_2') u_1 - u_1' \mathrm{E}(\beta(x_2) p_2 p_2') u_1.$$

Let u_1^* denote the optimal mean-semivariance portfolio. Then it follows from the piecewise quadratic structure of mean-semivariance, u_1^* satisfies the following relationship:

$$u_1^* = \frac{1}{2} \mathrm{E}(\beta_2(x_1, u_1^*) p_2 p_2')^{-1} \mathrm{E}(\alpha_2(x_1, u_1^*) - 2\beta_2(x_1, u_1^*) r_2 x_1) p_2.$$

In the following, we use the notation, γ_2^*, α_2^*, and β_2^* to denote the following piecewise constant random variables according to the following definitions:

$$\gamma_2^*(x_1) = \gamma_2(x_1, u_1^*),$$
$$\alpha_2^*(x_1) = \alpha_2(x_1, u_1^*),$$
$$\beta_2^*(x_1) = \beta_2(x_1, u_1^*).$$

Given u_1^*, we can write a formula for $J_2(x_1)$:

$$J_2(x_1) = \mathrm{E}(r_2 x_1 + p_2' u_1^*) - c\mathrm{E}(h - r_2 x_1 - p_2' u_1^*)_+^2.$$

Simplification leads to the following function:

$$J_2(x_1) = \mathrm{E}(\tilde{\gamma}_1(x_1)) + \mathrm{E}(\tilde{\alpha}_1(x_1)x_1) - \mathrm{E}(\tilde{\beta}_1(x_1)x_1^2).$$

The three functions $\tilde{\gamma}_1$, $\tilde{\alpha}_1$, and $\tilde{\beta}_1$ are piecewise constant according to the following definitions:

$$\tilde{\gamma}_1(x_1) = \mathrm{E}\gamma_2^*(x_1) + \frac{1}{4}\mathrm{E}\alpha_2^*(x_1)p_2'\mathrm{E}(\beta_2^*(x_1)p_2p_2')^{-1}\mathrm{E}\alpha_2^*(x_1)p_2,$$
$$\tilde{\alpha}_1(x_1) = \mathrm{E}\alpha_2^*(x_1)r_2 - \mathrm{E}\alpha_2^*(x_1)p_2'\mathrm{E}(\beta_2^*(x_1)p_2p_2')^{-1}\mathrm{E}\beta_2^*(x_1)r_2p_2,$$
$$\tilde{\beta}_1(x_1) = \mathrm{E}\beta_2^*(x_1)r_2^2 - \mathrm{E}\beta_2^*(x_1)r_2p_2'\mathrm{E}(\beta_2^*(x_1)p_2p_2')^{-1}\mathrm{E}\beta_2^*(x_1)r_2p_2.$$

We note that this form is necessarily analogous to the derivation of Li and Ng [42]. In their paper, they embed the mean-variance problem in a particular second moment problem. The mean-semivariance problem, in a neighborhood of the optimal solution, is equivalent to a particular second moment problem, hence, the analogous form. We again use the shorthand notation:

$$\gamma_1(x,u) = \tilde{\gamma}(r_1 x + p_1' u),$$

$$\alpha_1(x,u) = \tilde{\alpha}(r_1 x + p_1' u),$$

$$\beta_1(x,u) = \tilde{\beta}(r_1 x + p_1' u).$$

Finding the optimal first period portfolio is now equivalent to solving the following optimization problem:

$$\max_{u_0} \mathrm{E}(\gamma_1(x_0,u_0) + \alpha_1(x_0,u_0) r_1 x_0 - \beta_1(x_0,u_0)(r_1 x_0)^2)$$
$$+ \mathrm{E}((\alpha_1(x_0,u_0) - 2\beta_1(x_0,u_0) r_1 x_0) p_1') u_0 - u_0' \mathrm{E}(\beta(x_1) p_1 p_1') u_0.$$

Let u_0^* denote the optimal solution to the first period problem. The optimal solution to this problem has the same form as the second period, except that the coefficients $\alpha_1^*(x_0)$ and $\beta_1^*(x_0)$ are determined by the second period optimal portfolio function:

$$u_0^* = \frac{1}{2} \mathrm{E}(\beta_1(x_0,u_0^*) p_1 p_1')^{-1} \mathrm{E}(\alpha_1(x_0,u_0^*) - 2\beta_1(x_0,u_0^*) r_1 x_0) p_1.$$

The basic idea behind solving the two period portfolio problem is illustrated in Algorithm 3. The correctness of the algorithm, given in the following proposition, is an immediate consequence of applying the results of Lau and Womersley [32], which describes the finite convergence of multi-period piecewise quadratic programming algorithm on which this algorithm is based.

Proposition 25. *Algorithm 3 converges to the two-period mean-semivariance optimal portfolio.*

Algorithm 3 The algorithm takes as input initial wealth x_0, risk free rates r_1 and r_2, excess returns p_1 and p_2, target wealth h, and risk aversion parameter c, and returns an optimal two period mean-semivariance portfolio u_0^*.

Let $u_0^{(0)}$ and $u_1^{(0)}$ be arbitrary portfolios.
Set $\leftarrow 1$
while $u_0^{(i)}$ is not optimal. **do**
    for $\omega_1 \in \Omega_1$ **do**
        $x_1^{(i)} \leftarrow r_1(\omega_1)x_0 + p_1(\omega_1)' u_0^{(i)}$
        $k \leftarrow 0$
        while $u_1^{(k)}$ is not optimal for P2$(x_1^{(i)}(\omega_1))$ **do**
            $\alpha_2^{(k)} \leftarrow \alpha_2(r_2 x_1^{(i)}(\omega_1) + p_2' u_1^{(k)}(\omega_1))$
            $\beta_2^{(k)} \leftarrow \beta_2(r_2 x_1^{(i)}(\omega_1) + p_2' u_1^{(k)}(\omega_1))$
            $u_1^{(k+1)} \leftarrow \frac{1}{2}\mathrm{E}_{\omega_2}^{-1}\beta_2^{(k)} p_2 p_2' \mathrm{E}_{\omega_2}(\alpha_2^{(k)} q - 2\beta_2^{(k)} r_2 x_1^{(i)}(\omega_1))p_2$
            $k \leftarrow k+1$
        $u_1^{*(i)}(\omega_1) \leftarrow u_1^{(k-1)}(\omega_1)$
        $\alpha_2^{*(i)}(\omega_1) \leftarrow \alpha_2(r_2 x_1^{(i)}(\omega_1) + p_2' u_1^{*(i)}(\omega_1))$
        $\beta_2^{*(i)}(\omega_1) \leftarrow \beta_2(r_2 x_1^{(i)}(\omega_1) + p_2' u_1^{*(i)}(\omega_1))$
        $\alpha_1^{(i)}(\omega_1) \leftarrow \mathrm{E}_{\omega_2}\alpha_2^{*(i)}(\omega_1) r_2$
            $-\mathrm{E}_{\omega_2}\alpha_2^{*(i)}(\omega_1) p_2 \mathrm{E}_{\omega_2}^{-1}\beta_2^{*(i)}(\omega_1) p_2 p_2' \mathrm{E}_{\omega_2}\beta_2^{*(i)}(\omega_1) r_2 p_2$
        $\beta_1^{(i)}(\omega_1) \leftarrow \mathrm{E}_{\omega_2}\beta_2^{*(i)}(\omega_1) r_2^2$
            $-\mathrm{E}_{\omega_2}\beta_2^{*(i)}(\omega_1) r_2 p_2 \mathrm{E}_{\omega_2}^{-1}\beta_2^{*(i)}(\omega_1) p_2 q p_2' \mathrm{E}_{\omega_2}\beta_2^{*(i)}(\omega_1) r_2 p_2$
    $u_0^{(k+1)} \leftarrow \frac{1}{2}\mathrm{E}_{\omega_1}^{-1}\beta_1^{(k)} p_1 p_1' \mathrm{E}_{\omega_1}(\alpha_1^{(k)} - 2\beta_1^{(k)} r_1 x_0^{(i)}(\omega_0))p_1$
    $i \leftarrow i+1$

4.3.1 Modifications for Equality Constraints

Adding equality constraints to the two-period formulation yields the following optimization problem. The general form of the two-period mean-semivariance problem is the following:

$$\begin{aligned} \max_{u_0, u_1} \quad & \mathrm{E}(x_2) - c\mathrm{E}(h - x_2)_+^2, \\ \text{s.t.} \quad & x_2 = r_2 x_1 + p_2' u_1, \\ & x_1 = r_1 x_0 + p_1' u_0, \\ & A_1 u_1 = b_1, \\ & A_0 u_0 = b_0. \end{aligned}$$

In a manner similar to the single-period problem, we modify the approach for generating the two-period mean-semivariance portfolio. From equations (3.4) to (3.8), we see that if u_1^* is the optimal second period portfolio, it satisfies the following equation:

$$u_1^* = \tilde{u}_1 - B_2^{-1}A_1'(A_1B_2^{-1}A_1')^{-1}(A_1\tilde{u}_1 - b_1).$$

The following notation is used:

$$\begin{aligned}
\tilde{u}_1 &= \frac{1}{2}B_2^{-1}\mathrm{E}(\alpha_2^*(x_1) - 2\beta_2^* r_2 x_1)p_2, \\
\alpha_2^*(x_1) &= \alpha_2(x_1, u_1^*), \\
\beta_2^*(x_1) &= \beta_2(x_1, u_1^*), \\
B_2 &= \mathrm{E}(\beta_2^*(x_1)p_2p_2').
\end{aligned}$$

Substituting this value for the optimal solution into the second period value functions gives the following first period optimization problem (after simplification):

$$J_2(x_1) = \mathrm{E}(\gamma_1(x_1)) + \mathrm{E}(\alpha_1(x_1)x_1) - \mathrm{E}(\beta_1(x_1)x_1^2).$$

The three functions γ_1, α_1, and β_1 are piecewise constant according to the following definitions:

$$\begin{aligned}
\gamma_1(x_1) &= \mathrm{E}\gamma_2^*(x_1) + \frac{1}{4}(\mathrm{E}\alpha_2^*(x_1)p_2'(B_2^{-1} - \tilde{H}_2)\mathrm{E}\alpha_2^*(x_1)p_2) \\
&\qquad + (\mathrm{E}\alpha_2^*(x_1)p_2 - 2\mathrm{E}\beta_2^*(x_1)r_2p_2)'H_2b - b'(A_1B_2^{-1}A_1')^{-1}b, \\
\alpha_1(x_1) &= \mathrm{E}\alpha_2^*(x_1)r_2 - \mathrm{E}\alpha_2^*(x_1)p_2'(B_2^{-1} - \tilde{H}_2)\mathrm{E}\beta_2^*(x_1)r_2p_2, \\
\beta_1(x_1) &= \mathrm{E}\beta_2^*(x_1)r_2^2 - \mathrm{E}\beta_2^*(x_1)r_2p_2'(B_2^{-1} - \tilde{H}_2)\mathrm{E}\beta_2^*(x_1)r_2p_2, \\
H_2 &= B_2^{-1}A_1'(A_1B_2^{-1}A_1'), \\
\tilde{H}_2 &= H_2A_1B_2^{-1}.
\end{aligned}$$

Let u_0^* denote the optimal solution to the first period problem,

$$u_0^* = \tilde{u}_0 - B_1^{-1}A_0'(A_0B_1^{-1}A_0')^{-1}(A_0\tilde{u}_0 - b_0),$$

where the following notation is used:

$$\begin{aligned}
\tilde{u}_0 &= \frac{1}{2}B_1^{-1}\mathrm{E}(\alpha_1^*(x_0) - 2\beta_1^* r_1 x_0)p_1,\\
\alpha_1^*(x_0) &= \alpha_1(x_0, u_0^*),\\
\beta_1^*(x_0) &= \beta_1(x_0, u_0^*),\\
B_1 &= \mathrm{E}(\beta_1^*(x_0)p_1p_1').
\end{aligned}$$

In this case, as with the unconstrained two-period problem, the first period coefficients are determined from the value function for the second period.

4.4 Multi-Period Portfolio Optimization

We now generalize the procedures discussed in the two-period formulation to a multi-period mean-semivariance portfolio optimization problem. The motivation for studying multi-period portfolio optimization is discussed by Li and Ng [42]. We now state the full form for the T-period multi-period portfolio optimization problem.

$$\begin{aligned}
\max_{u_0,u_1,\ldots,u_{T-1}} \quad & \mathrm{E}(x_T) - c\mathrm{E}(h - x_T)_+^2,\\
\text{s.t.} \quad & x_t = r_t x_{t-1} + p_t' u_{t-1}, \qquad \forall t \in \{1, 2, \ldots, T\}.
\end{aligned}$$

The high level approach we take is broken into the following steps. We first restate the problem as a dynamic programming problem in a manner similar to the two-period problem. As in the two-period case, we show that at each period the optimal reward-to-go function can be written in the following inductively defined form:

$$\mathrm{E}\gamma_t + \mathrm{E}\alpha x_t - \mathrm{E}\beta x_t^2.$$

We then use a parametric optimization approach to provide an algorithm for explicitly computing the entire value function, reducing each period of the problem to constructing an efficient frontier for the single period mean-semivariance portfolio problem.

We now present the lemma used for solving each period of the multi-period problem

Lemma 26. *Let f be a continuously differentiable piecewise-quadratic function, which we express in the following form:*

$$f(x) = \gamma + \alpha x - \beta x^2, \tag{4.7}$$

where α, β, and γ are piecewise constant functions of x defining the quadratic pieces of f. Let J be the optimal value function defined in the following way:

$$J(x) = \max_{u \in \mathbb{R}^n} Ef(rx + p'u), \tag{4.8}$$

where r and p are both finitely supported random variables in $\mathbb{R}$ and $\mathbb{R}^n$ respectively. Then, J can be written in the following way:

$$\begin{aligned}
J(x) &= \gamma' + \alpha' x - \beta' x^2, \\
\gamma' &= E\gamma^* + \frac{1}{4} E\alpha^* p' (E\beta^* pp')^{-1} E\alpha^* p, \\
\alpha' &= E\alpha^* r - E\beta^* rp' (E\beta^* pp')^{-1} E\alpha^* p, \\
\beta' &= E\beta^* r^2 - E\beta^* rp' (E\beta^* pp')^{-1} E\beta^* rp,
\end{aligned}$$

where the starred coefficients γ^*, α^*, and β^* are the coefficients of f evaluated at the maximizer with respect to u given x.

Proof. Grouping the terms of $\mathrm{E}f(rx + p'u)$ by u yields the followings expression:

$$\mathrm{E}\gamma + \mathrm{E}\alpha rx - \mathrm{E}\beta r^2x^2 + \mathrm{E}(\alpha p - 2\beta rxp)'u - u'\mathrm{E}\beta pp'u.$$

Applying the non-smooth Newton's method to this expression, maximizing with respect to u yields the following optimizer:

$$u^* = (\mathrm{E}\beta^* pp')^{-1}\mathrm{E}(\alpha^* p - 2\beta^* rxp),$$

where α^*, β^*, and γ^* are the functions α, β, and γ evaluated at x and the terminating fixed point of the non-smooth Newton method, which is identically u^*. Evaluating $\mathrm{E}f(rx + p'u^*)$ and grouping the resulting terms by x yields the result of the lemma. □

Using Lemma 26, we can concisely write the dynamic programming form of the problem. The following proposition both states the dynamic program and its tractability.

Proposition 27. *The multi-period mean-semivariance optimal portfolio can be calculated by solving the following dynamic program via standard backward-induction procedures:*

$$\begin{aligned} J_T(x_T) &= x_T - c(h - x_T)_+^2, \\ J_t(x_t) &= \max_{u_t} E_t J_{t+1}(r_{t+1}r_t + p'_{t+1}u_t), \quad \forall t \in \{0, 1, \ldots, T-1\}, \end{aligned}$$

where E_t denotes the expectation conditional on returns $r_{t'}$ and $p_{t'}$ for $t' \leq t$.

Proof. We observe that J_T can be written in the following form:

$$J_T(x_T) = (-ch^2\lambda) + (1 + 2ch\lambda)x_T - (c\lambda)x_T^2,$$

where $\lambda = [h - x_T \geq 0]$. Combining this reformulation with Lemma 26, we now can write each J_t in the following form:

$$J_t(x_t) = \gamma_t + \alpha_t x_t - \beta_t x_t^2.$$

Because each period is now a continuously differentiable piecewise-quadratic function, we can therefore apply the non-smooth Newton's method to each period to yield the following optimal portfolio functions, noting that for $t \geq 1$, the portfolios are random functions of the return realizations:

$$u_t^*(x_t) = \frac{1}{2}(\mathrm{E}_t \beta_{t+1}^* p_{t+1} p_{t+1}')^{-1}(\mathrm{E}_t \alpha_{t+1}^* p_{t+1} - 2\mathrm{E}_t \beta *_{t+1} r_{t+1} p_{t+1} x_t).$$

Evaluating the optimal portfolio function for $t = 0$ yields the desired result. □

We observe that directly evaluating the coefficients at period t requires applying the non-smooth Newton's method to numerous piecewise quadratic functions. We now describe parametric optimization procedures analogous to the efficient frontier calculations for the single-period problem that dramatically reduces the computational effort required by the multi-period portfolio problem.

We approach the parametric optimization in two parts, first for the terminal period, in which we have the explicit semivariance formulations, second for the intermediate periods, in which the piecewise quadratic function is given in terms of coefficient functions and breakpoints.

Lemma 28. *Let x_{T-1} be given. Let u_{T-1}^* be the optimal portfolio associated with x_{T-1}.*

Then λ^* associated with x_{T-1} is constant on the interval $[x^-, x^+]$, where

$$\begin{aligned}
x^- &= \max\big\{\delta(\omega) : \omega \in \Omega_0^- \cup \Omega_1^-\big\},\\
x^+ &= \min\big\{\delta(\omega) : \omega \in \Omega_0^+ \cup \Omega_1^+\big\},\\
\delta(\omega) &= \frac{\frac{1}{2}p_T(\omega)'Ba - h}{p_T(\omega)'Bb - r_T(\omega)},\\
B &= \boldsymbol{E}\beta_T^* p_T p_T',\\
a &= \boldsymbol{E}\alpha_T^* p_T,\\
b &= \boldsymbol{E}\beta_T^* r_T p_T,\\
\Omega_0^- &= \{\omega \in \Omega : p_T(\omega)'Bb - r_T(\omega) < 0, \lambda^*(\omega) = 0\},\\
\Omega_1^- &= \{\omega \in \Omega : p_T(\omega)'Bb - r_T(\omega) > 0, \lambda^*(\omega) = 1\},\\
\Omega_0^+ &= \{\omega \in \Omega : p_T(\omega)'Bb - r_T(\omega) < 0, \lambda^*(\omega) = 1\},\\
\Omega_1^+ &= \{\omega \in \Omega : p_T(\omega)'Bb - r_T(\omega) > 0, \lambda^*(\omega) = 0\}.
\end{aligned}$$

We take the maximum of an empty set to be $-\infty$ and the minimum of an empty set to be $+\infty$. Furthermore, either $x^+ = +\infty$ or there exists a neighboring interval associated with a vector λ_+^* that differs from λ^* in the following way:

$$\begin{aligned}
\lambda_+^*(\omega) &= \begin{cases} \lambda^*(\omega) & \omega \notin \Omega^*, \\ 1 - \lambda^*(\omega) & \omega \in \Omega^*, \end{cases}\\
\Omega^* &= \arg\min\left\{\frac{\frac{1}{2}p_T(\omega)'Ba - h}{p_T(\omega)'Bb - r_T(\omega)} : \omega \in \Omega_0^+ \cup \Omega_1^+\right\}.
\end{aligned}$$

Similarly, either $x^- = -\infty$ or there exists a neighboring interval associated with a vector

λ^*_- that differs from λ^* in the following way:

$$\lambda^*_-(\omega) = \begin{cases} \lambda^*(\omega) & \omega \notin \Omega^*, \\ 1 - \lambda^*(\omega) & \omega \in \Omega^*, \end{cases}$$
$$\Omega^* = \arg\max \left\{ \frac{\frac{1}{2} p_T(\omega)' Ba - h}{p_T(\omega)' Bb - r_T(\omega)} : \omega \in \Omega_0^- \cup \Omega_1^- \right\}.$$

Proof. The vector λ^* is a piecewise constant function, and at x_{T-1}, it satisfies the following equality:

$$\lambda^*(\omega) = [h - r_T(\omega) x_{T-1} - p_T(\omega)' u^*_{T-1} \geq 0]. \tag{4.9}$$

Substituting the value of u^*_{T-1} as a function of x_{T-1}, this equality in indicator function is equivalent to the following system of inequalities on x_{T-1}:

$$x_{T-1} \geq \delta(\omega) \quad \forall \omega \in \Omega_1^-, \tag{4.10}$$

$$x_{T-1} \leq \delta(\omega) \quad \forall \omega \in \Omega_1^+, \tag{4.11}$$

$$x_{T-1} \geq \delta(\omega) \quad \forall \omega \in \Omega_0^-, \tag{4.12}$$

$$x_{T-1} \leq \delta(\omega) \quad \forall \omega \in \Omega_0^+. \tag{4.13}$$

The bounds on increasing and decreasing x_{T-1} emerge by simplifying this system of inequalities, yielding x^+ and x^-. Identifying the ω associated with the tight inequalities as x_{T-1} increases or decreases yields the values of λ^* that change to yield λ^*_+ and λ^*_-. □

Lemma 29. *Let $t \in \{1, 2, \ldots, T-1\}$ and x_{t-1} be given. Let u^*_{t-1} be the optimal portfolio associated with x_{t-1}. Let the vector of optimal coefficients associated with x_{t-1} be denoted by $\xi^*_t = (\alpha^*_t, \beta^*_t, \gamma^*_t)$. The general coefficient vector $\xi = (\alpha_t, \beta_t, \gamma_t)$ have breakpoints X_k for $k = 1, 2, \ldots, K$ on which ξ is piecewise constant. From these breakpoints we construct intervals $I(\omega) = [L(\omega), U(\omega)]$ for each $\omega \in \Omega$ containing $x_t = r_t(\omega) x_{t-1} + p_t(\omega)' u^*_{t-1}$. We* *rewrite x_t for ease of computation by expanding u_t^*:*

$$x_t(\omega) = \phi(\omega)x_{t-1} + \psi(\omega), \tag{4.14}$$

$$\phi(\omega) = r_t(\omega) - p_t(\omega)'B^{-1}b, \tag{4.15}$$

$$\psi(\omega) = \frac{1}{2}p_t(\omega)'B^{-1}a, \tag{4.16}$$

$$B = E\beta_t^* p_t p_t', \tag{4.17}$$

$$a = E\alpha_t^* p_t, \tag{4.18}$$

$$b = E\beta_t^* r_t p_t. \tag{4.19}$$

The vector ξ_t^* as a function of x_{t-1} is constant on the interval $[x^-, x^+]$, where

$$x^- = \max\{\delta^-(\omega) : \phi(\omega) \neq 0\},$$

$$x^+ = \min\{\delta^+(\omega) : \phi(\omega) \neq 0\},$$

$$\delta^-(\omega) = \begin{cases} \frac{L(\omega)-\psi(\omega)}{\phi(\omega)} & \phi(\omega) > 0, \\ \frac{U(\omega)-\psi(\omega)}{\phi(\omega)} & \phi(\omega) < 0, \end{cases}$$

$$\delta^+(\omega) = \begin{cases} \frac{U(\omega)-\psi(\omega)}{\phi(\omega)} & \phi(\omega) > 0, \\ \frac{L(\omega)-\psi(\omega)}{\phi(\omega)} & \phi(\omega) < 0. \end{cases}$$

We take the maximum of an empty set to be $-\infty$ and the minimum of an empty set to be $+\infty$. Furthermore, either $x^+ = +\infty$ or there exists a neighboring interval associated with *a vector ξ^*_+ that differs from ξ^* in the following way:*

$$\xi^*_+(\omega) = \begin{cases} \xi^*(\omega) & \omega \notin \Omega^*_+, \\ \xi_+(\omega) & \omega \in \Omega^*_+, \end{cases}$$
$$\Omega^*_+ = \arg\min\{\delta^+(\omega) : \phi(\omega) \neq 0\},$$
$$\xi_+(\omega) = \begin{cases} \xi_{rhs}(\omega) & \phi(\omega) > 0, \\ \xi_{lhs}(\omega) & \phi(\omega) < 0. \end{cases}$$

where $\xi_{rhs}(\omega)$ is the value of $\xi(\omega)$ in the interval adjacent to $I(\omega)$ as x_t increases and $\xi_{lhs}(\omega)$ is the value of $\xi(\omega)$ in the interval adjacent to $I(\omega)$ as x_t decreases. Similarly, either $x^- = -\infty$ or there exists a neighboring interval associated with a vector ξ^*_- that differs from ξ^* in the following way:

$$\xi^*_-(\omega) = \begin{cases} \xi^*(\omega) & \omega \notin \Omega^*_-, \\ \xi_-(\omega) & \omega \in \Omega^*_-, \end{cases}$$
$$\Omega^*_- = \arg\max\{\delta^-(\omega) : \phi(\omega) \neq 0\},$$
$$\xi_-(\omega) = \begin{cases} \xi_{lhs}(\omega) & \phi(\omega) > 0, \\ \xi_{rhs}(\omega) & \phi(\omega) < 0. \end{cases}$$

Proof. The vector ξ^*_t is a piecewise constant function, and at x_{Tt-1}, it satisfies the following equality:

$$\xi^*_t(\omega) = \xi_t(\omega). \tag{4.20}$$

Substituting the value of u^*_{t-1} as a function of x_{t-1}, this equality holds for $x_t \in I(\omega)$, which is equivalent to the following system of inequalities on x_{t-1}:

$$x_{t-1} \geq \delta^{-}(\omega) \quad \forall \omega \in \{\phi(\omega) > 0\}, \tag{4.21}$$

$$x_{t-1} \leq \delta^{+}(\omega) \quad \forall \omega \in \{\phi(\omega) > 0\}, \tag{4.22}$$

$$x_{t-1} \geq \delta^{-}(\omega) \quad \forall \omega \in \{\phi(\omega) < 0\}, \tag{4.23}$$

$$x_{t-1} \leq \delta^{+}(\omega) \quad \forall \omega \in \{\phi(\omega) < 0\}. \tag{4.24}$$

The bounds on increasing and decreasing x_{t-1} emerge by simplifying this system of inequalities, yielding x^{+} and x^{-}. Identifying the ω associated with the tight inequalities as x_{T-1} increases or decreases yields the values of ξ^{*} that change to yield ξ^{*}_{+} and ξ^{*}_{-}. □

Combining Lemmas 28 and 29, we can solve the dynamic program defining the multi-period mean-semivariance problem in the following way. First, compute the value function of the terminal period by solving the single-period optimization problem for $x_{T-1} = 0$. Use Lemma 28 to extend the value function definition. Second, compute the value function the subsequent period using the pre-computed value function for the terminal period. Use Lemma 29 to extend the value function definition. Inductively, we can repeat this second step until we reach the first period where x_0 is given exogenously, or we compute an wealth-frontier of portfolios.

The computational improvement here is that the extensions provided by Lemmas 28 and 29 require only rank-1 updates to the inverse B^{-1} defining any given segment of the value function. This is in contrast to solving the single-period portfolio problem from scratch for each segment.

4.5 Numerical Experiment

We compare now the single-period portfolio to the multi-period portfolio. The relative performance of the two depends heavily on the investor's forecast of returns for the horizon being optimized. Because such forecasting is beyond the scope of this research, we focus instead on the differences between the objective functions in each optimization problem. The goal is to better understand the differences in investor utility implied by mean-semivariance and the multi-period mean-semivariance value function. The mean-semivariance investor optimizes a single-period mean-semivariance portfolio at each period. The multi-period investor optimizes using the dynamic programming value function calculated over the remaining number of periods.

The objective function of the mean-semivariance portfolio optimization is the expectation over a problem independent utility function: terminal wealth less semivariance. By contrast, the multi-period value function is problem dependent. First, we describe some qualitative differences:

- Mean-semivariance and the multi-period value function are both quadratic for wealth below a threshold and linear above a threshold.
- The multi-period value function does not have a unit slope in the linear component and is piecewise quadratic in the transition from quadratic to linear in an open interval around the wealth target h.

Figure 4.1 illustrates the differences between a 3-period value function and the mean-semivariance objective.

Second, we describe some common cases when the multi-period and single-period portfolios are most similar. Generally, as the wealth target h increases above the risk free rate of return, the multi-period portfolio becomes more similar to the single-period portfolio. Depending on the data set, one of the following will happen:

1. As the wealth target h crosses the risk free rate of return r_0, the multi-period portfolio becomes the same as the single-period portfolio.
2. As the wealth target h crosses the risk free rate of return r_0, the multi-period portfolio becomes co-linear to the single-period portfolio. However, a gap persists because the multi-period has a different degree of risk aversion.

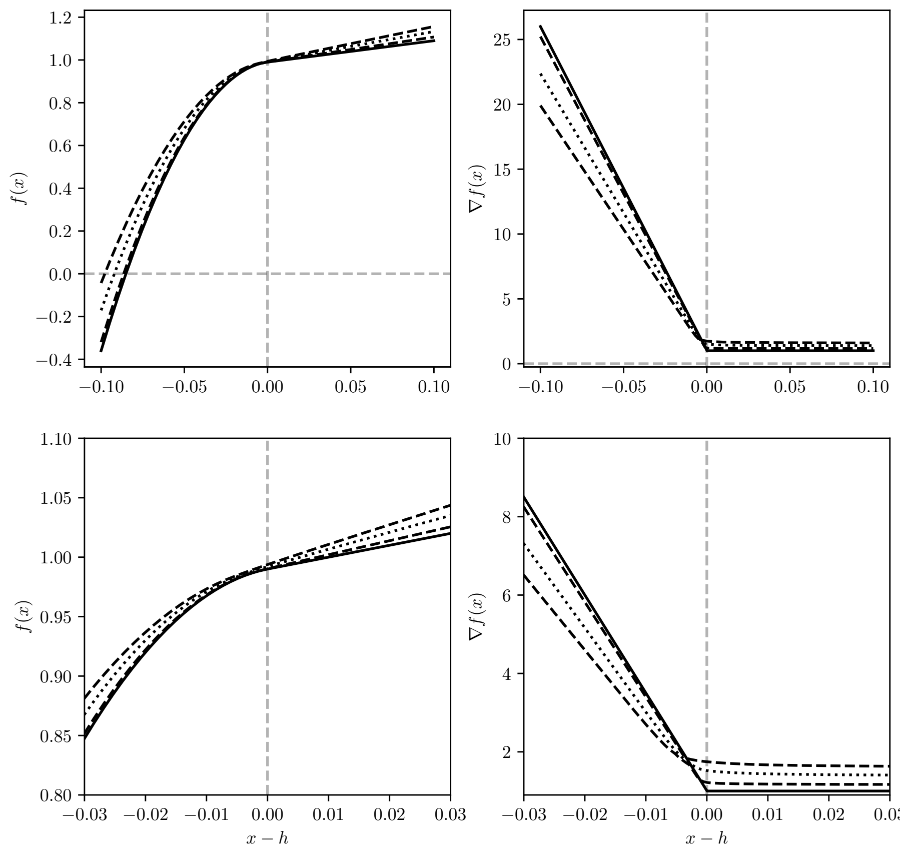

Figure 4.1: The left two figures compare the mean-semivariance objective (solid line) to a range of 3-period value functions. The dashed lines are the min and max over 5 data sets; the doted is the mean. The right two figures show the same comparison but in terms of gradients for the respective functions.

3. As the wealth target h crosses the a threshold materially larger than the risk free rate of return r_0, the multi-period portfolio becomes co-linear to the single-period portfolio, possibly with a difference in risk aversion.

Tables 4.1 and 4.2 provide comparisons between 3-period and 2-period portfolios for different values of the risk free return r_0 and the wealth target h. The differences are with respect to the L^1 norm. In Table 4.2 the portfolios are normalized to check co-linearity. We see the greatest similarity for $r_0 = 1$ because the effects of $r_0 \neq 0$ are multiplied in a multi-period portfolio. This is also illustrated by a 3-period portfolio generally being more different from the single-period portfolio than the 2-period portfolio.

Table 4.1: Each cell computes the mean, min, or max over 5 return scenarios of the L^1 difference between the multi-period and myopic portfolios. The multi-period portfolio is most similar to the single-period portfolio when the risk free return is near 1.0 and the wealth target is above the risk free return.

		T =3			T =2		
r_0	$h - r_0$	Mean	Min	Max	Mean	Min	Max
0.98	-0.01	1.18	0.25	2.72	2.97	2.02	4.50
	0.01	4.61	1.75	10.25	1.58	0.39	4.05
1.00	-0.01	3.07	0.39	9.65	1.21	0.26	3.40
	0.01	2.27	0.03	9.28	0.31	0.00	1.06
1.02	-0.01	9.19	3.09	24.20	6.56	4.88	11.17
	0.01	9.57	4.07	25.75	4.52	2.82	8.99

Table 4.2: Each cell computes the mean, min, or max over 5 return scenarios of the L1 difference between the normalized multi-period and myopic portfolios. The normalization highlights cases when portfolio weights have the same proportions and differ only in total investment, or similarly, risk aversion. The multi-period portfolio is most similar to the myopic portfolio when the risk free return less than or equal to 1.0 and the wealth target is above the risk free return.

		T =3			T =2		
r_0	$h - r_0$	Mean	Min	Max	Mean	Min	Max
0.98	-0.01	0.16	0.08	0.38	0.11	0.06	0.23
	0.01	0.01	0.00	0.03	0.00	0.00	0.01
1.00	-0.01	0.03	0.00	0.05	0.06	0.01	0.16
	0.01	0.03	0.01	0.08	0.01	0.00	0.02
1.02	-0.01	0.15	0.02	0.28	0.08	0.02	0.19
	0.01	0.12	0.02	0.26	0.08	0.05	0.11

4.5.1 Algorithmic Efficiency Gains

Our explicit formulation of the piecewise quadratic structure of the multi-period mean-semi-variance value functions increases the efficiency of the multi-period portfolio optimization. The problem is still very computationally intense, but the savings are significant compared to computing each segment of the piecewise quadratic value function by explicitly solving and caching the results of a single-period optimization. To illustrate, we compare run times for a 4 period portfolio with 2 assets and 30 samples per period. The 2, 3, and 4 period value functions have 29, 813, and 22,228 segments, respectively. On the system used to compute these numerical experiments, a single-period instance required 0.02 seconds to solve. Solving the 4 period model naively under these assumption would require 461.4 seconds or a 14 times increase compared to our approach which required 30.2 seconds. Tables 4.3 and 4.4 give detailed run times for a variety of configurations and illustrate the scalability limits of the current approach in terms of sample size and number of assets.

4.6 Conclusion and Future Research

For investors seeking to optimize performance over a fixed time horizon, multi-period portfolio optimization provides greater flexibility. Our framework provides the basis for further extensions to incorporate considerations such as transaction costs. The approach can in general be applied to piecewise linear and piecewise quadratic extensions to the objective function. The complexity gains of the parametric approach apply if these additional terms are also analytically tractable. Possible areas of future research:

- Replace the value function by a smooth approximation using the piecewise quadratic coefficients from the extreme minimum and maximum wealth as anchors. This approach could leverage the objective function being continuously differentiable to bound the error of such an approximation. This would significantly improve run time

[1] For $T = 5$ and $n \geq 2$, only a single instance was run, so there was no averaging of run times.

Table 4.3: Run times are in seconds. Each timing is the average of 3 replications. The variable T denotes the number of periods in the optimization. The variables n and m denote the number of assets and samples per period, respectively. Missing cells have run times in excess of 1 hour.

	$m = 30$					$m = 60$				
T	$n = 1$	2	3	4	5	$n = 1$	2	3	4	5
1	0.02	0.05	0.04	0.06	0.09	0.04	0.05	0.06	0.06	0.07
2	0.16	0.40	0.45	0.51	0.94	0.10	0.15	0.20	0.22	0.24
3	0.85	1.47	1.59	1.70	2.82	1.83	2.47	3.13	3.05	3.45
4	20.68	30.25	32.86	23.32	39.85	507.2	838.5	939.4	873.6	832.9
5[1]	3911	7021	5299	3344	2715	–	–	–	–	–
	$m = 90$					$m = 120$				
1	0.02	0.02	0.04	0.04	0.04	0.03	0.03	0.04	0.05	0.05
2	0.13	0.22	0.34	0.34	0.39	0.17	0.26	0.33	0.37	0.39
3	6.97	12.31	12.55	12.61	13.16	17.34	30.75	34.53	34.01	35.30
4	3699	–	–	–	–	–	–	–	–	–
	$m = 150$									
1	0.03	0.05	0.08	0.09	0.09					
2	0.19	0.42	0.44	0.60	0.60					
3	29.39	62.89	69.66	90.55	104.6					

Table 4.4: Run times are in seconds. Each timing is the average of 3 replications. The variable T denotes the number of periods in the optimization. The variables n and m denote the number of assets and samples per period, respectively.

T	n	$m = 90$	120	150	180	210	240	270	300
1	5	0.05	0.06	0.05	0.05	0.06	0.06	0.08	0.08
	10	0.08	0.09	0.07	0.07	0.10	0.11	0.12	0.14
	15	0.11	0.12	0.12	0.11	0.18	0.18	0.18	0.20
	20	0.14	0.16	0.15	0.16	0.25	0.25	0.26	0.27
	30	0.34	0.33	0.41	0.46	0.65	0.96	1.08	1.32
2	5	0.36	0.25	0.36	0.45	0.62	0.72	0.75	0.99
	10	0.79	0.41	0.59	0.70	0.85	1.03	1.23	1.36
	15	1.22	0.73	0.89	1.24	1.17	1.77	2.02	2.44
	20	1.99	1.10	1.63	1.84	1.53	2.46	2.54	3.40
	30	5.20	2.51	6.08	5.81	4.71	7.20	7.36	10.82

by reducing the number of pieces, which grows rapidly with the number of periods.

- Adding transaction costs would make for a more balanced comparison between a multi-period investor and a buy-and-hold investor. This could help understand the cases when active versus passive investing are favorable.

CHAPTER 5

5.1 Robust Mean-Semivariance Portfolio Problem

It is well known that mean-variance portfolios suffer significantly from sampling errors in the calculation of the mean and covariance used in the optimization. Although the academic literature is relatively limited with respect to the sensitivity of mean-semivariance portfolios, the piecewise quadratic structure of the portfolio results in a an optimal solution with similar structure. In particular the classical mean-semivariance formulation is computed using either an empirical distribution or an assumed distribution. Empirical distributions tend to underestimate rare events and increase the risk of the portfolio being over-fit against historical data. Assuming a distribution imposes a structural assumption which may not be reflective of future behavior.

Typically three approaches are taken to overcome the mean-variance portfolio's sensitivity. First, portfolios can be generated by resampling from a perturbed distribution, using the average portfolio that results from optimizing against each re-sampled outcome reduces the sensitivity of the resultng portfolio to any particular realization of the empirical distribution. Second, a robust optimization approach, usually based on theory in the vein of Ben-Tal, Ghaoui, and Nemirovski [56], can be used to create a robust counterpart where the mean and covariance range over a given uncertainty set. Third, the investor is assumed to impose a prior distribution over the mean and covariance. This prior distribution allows a Bayesian style estimate to be made of the portfolio returns. Bayesian estimates are especially useful for integrating multiple alternative forecasts. Our work explores a distributionally robust formulation for the mean-semivariance portfolio based on recent methods studied by Rujeerapaiboon et al. [35] and Jiang and Guan [50].

The resampling approach has been strongly advocated by Michaud [4], and the method has foundations in literature related to bootstrapping statistics. The method has the benefit of simplicity because it requires no additional optimization framework and is equivalent to a Monte Carlo method for estimating the unknown, true efficient frontier. The approach suffers from the slow convergence typical of Monte Carlo methods, and it requires repeating the same optimization procedure many times.

The robust optimization approach focuses on protecting the investor from the worst-case outcome among an uncertainty set constraining the conservativeness. Distributionally robust optimization considers the worst-case performance among a family of probability distributions. The most common robust formulation for the mean-variance portfolio is to assume the mean and covariance lie in an uncertainty set. The robust counterparts are deterministic, often second-order cone problems or semidefinite programs. These problems are computationally tractable with a difficulty similar to convex quadratic programming. A distributionally robust formulation for a growth-optimal portfolio based on optimizing the worst-case log-wealth over a range of distributions with fixed first and second moments was developed by Rujeerapaiboon et al. [35] and shown to be equivalent to a non-robust mean-variance portfolio.

5.2 Comparison of Distributionally Robust Optimization Methodologies

We compare three major branches of distributionally robust optimization. The key difference between these approaches is the choice for constructing the ambiguity set. First are moment-based methods that we study in our research. These methods constrain the ambiguity set to distributions having either fixed moments or moments constrained by inequalities. Second are the family of ϕ-divergence measures. Typically, these methods construct a reference distribution, empirical or fitted density, and constrain the ambiguity set based on a particular ϕ-divergence, e.g. Kullback-Leibler divergence. Third is the Wasserstein distance. Also called the earth-mover's distance, Wasserstein distance is determined both by an order parameter and a choice of distance metric based on the problem domain.

Our choice is to constrain the first and second moment, which Rujeerapaiboon et al. [35] call weak-sense white noise. This ambiguity set makes for a straight forward comparison with the non-robust mean-variance and mean-semivariance portfolios, as the non-robust portfolios are invariant under this family of distributions. The use of the first and second moment is intuitively clear, but as noted by Kleywegt and Gao [48], it does not fully utilize the information provided by the historical data used to generate the first and second moment estimates. However, we believe the moment-based approach is the natural first step for developing a theory of robust mean-semivariance portfolio optimization.

5.3 Robust Mean-Semivariance Optimization Problem

The mean-semivariance optimization problem has the following form:

$$\min_{u} f(u) = \mathrm{E}(\eta - p'u)_+^2 - b\mu' u.$$

Typical models of uncertainty for robust optimization introduce uncertainty on the moments of the random variables. However, because the calculation of semivariance is not independent of the decision variable u, this model of uncertainty is not straight forward to specify, compared to, for example, in variance models as presented by Halldórsson and Tütüncü [46]. We present an approach using moment-based distributionally robust optimization similar to the work of Rujeerapaiboon et al. and Jiang and Guan [35, 50]. The essential idea is to apply the classical problem of moments to construct a robust formulation of the mean-semivariance portfolio problem.

We consider the case where we wish to maximize the mean return subject to a constraint on the semivariance:

$$\begin{aligned} \max_{u \in U} \quad & \mathrm{E}_\pi(p'u), \\ \text{s. t.} \quad & \mathrm{E}_\pi(\eta - p'u)_+^2 \le s, \end{aligned}$$

where π is the probability measure associated with the distribution of p. We focus on a distributionally robust formulation of the semivariance constraint:

$$\max_{u \in U} \quad \mathrm{E}_\pi(p'u), \tag{5.1}$$

$$\text{s.t.} \quad \min_{\pi \in \Pi} -\mathrm{E}_\pi(\eta - p'u)_+^2 \geq -s. \tag{5.2}$$

We denote by Π the following set of probability measures:

$$\Pi = \{\pi' : \mathrm{E}_{\pi'} p = \mu, \mathrm{E}_{\pi'} pp' = P\},$$

where μ and P are the sample mean and sample second moments of p, respectively. We now provide a tractable expression for (5.2) using a generalization of the same procedure in the proof of Theorem 2 by Jiang and Guan [50] or the Appendix in Rujeerapaiboon et al. [35]. For this proof we utilize the S-lemma:

Lemma 30 (S-lemma)**.** *Let $f_i(\xi) = \xi' A_i \xi$ with $A_i \in \mathbb{S}^n$ be quadratic functions of $\xi \in \mathbb{R}^n$ for $i = 0, 1, \ldots, p$. Then, $f_0(\xi) \geq 0$ for all ξ with $f_i(\xi) \leq 0$, $i = 1, 2, \ldots, p$, if there exist constants $\tau_i \geq 0$ such that*

$$A_0 + \sum_{i=1}^{p} \tau_i A_i \succeq 0. \tag{5.3}$$

For $p = 1$, the converse implication holds if there exists a strictly feasible point $\tilde{\xi}$ with $f_1(\tilde{\xi}) < 0$.

Theorem 31. *Let E_1 and E_2 define the two following non-empty regions:*

$$E_1 = \{p \in \mathbb{R}^n : p \in \cup_{i=1}^m \{p' Q_i p + q_i' p + q_i^0 \leq 0\}\}, \tag{5.4}$$

$$E_2 = \{p \in \mathbb{R}^n : p \in \cap_{i=1}^m \{p' Q_i p + q_i' p + q_i^0 > 0\}\}. \tag{5.5}$$

Note, $E_1 = E_2^C$. *Let the* $f : \mathbb{R}^n \mapsto \mathbb{R}$ *be a piecewise quadratic function defined by*

$$f(p) = \begin{cases} p'A_1 p + b_1' p + c_1 & p \in E_1, \\ p'A_2 p + b_2' p + c_2 & p \in E_2. \end{cases} \tag{5.6}$$

We assume further that the following relationship holds between the two quadratic functions defining f*:*

$$\begin{bmatrix} A_2 - A_1 & \frac{1}{2}(b_2 - b_1) \\ \frac{1}{2}(b_2 - b_1) & c_2 - c_1 \end{bmatrix} \succeq 0. \tag{5.7}$$

In other words, the quadratic defined by A_2*,* b_2*, and* c_2 *is an upper bound on the quadratic defined by* A_1*,* b_1*, and* c_1*. Furthermore, we require that for each* $i = 1, 2, \ldots, m$*, there exist* $p_i \in \mathbb{R}^n$ *such that* $p_i' Q_i p_i + q_i' p_i + q_i^0 < 0$*. Then the value of* $z = \inf_{\mathbb{P} \in \mathcal{P}} E_{\mathbb{P}} f(p)$ *is equal to the optimal objective value of the following semidefinite program:*

$$\begin{aligned} z = \max_{H,h,h_0,y} \quad & P \cdot H + \mu' h + h_0, \\ \text{s. t.} \quad & \begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h' & -h_0 \end{bmatrix} + y_i \begin{bmatrix} Q_i & \frac{1}{2}q_i \\ \frac{1}{2}q_i' & q_i^0 \end{bmatrix} + \begin{bmatrix} A_1 & \frac{1}{2}b_1 \\ \frac{1}{2}b_1 & c_1 \end{bmatrix} \succeq 0, \quad \forall i = 1, 2, \ldots, m, \\ & \begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h' & -h_0 \end{bmatrix} + \begin{bmatrix} A_2 & \frac{1}{2}b_2 \\ \frac{1}{2}b_2 & c_2 \end{bmatrix} \succeq 0, \\ & y_i \geq 0, \qquad \forall i = 1, 2, \ldots, m, \\ & H \in \mathbb{S}. \end{aligned}$$

Proof. We denote by z the optimal value of the primal problem:

$$
\begin{aligned}
z = \inf_{\pi \in \mathcal{P}} & \int_{\mathbb{R}^n} f(p) d\pi, \\
\text{s. t.} & \int_{\mathbb{R}^n} p d\pi = \mu, \\
& \int_{\mathbb{R}^n} pp' d\pi = P, \\
& \int_{\mathbb{R}^n} d\pi = 1.
\end{aligned}
$$

Observing that the integrals behave as an inner product on the linear spaces of functions and probability measures, we can now apply the duality theory for conic linear programming to formulate the following dual:

$$
\begin{aligned}
z = \max_{H \in \mathbb{S}, h, h_0} \quad & P \cdot H + \mu' h + h_0, \\
\text{s. t.} \quad & p' H p + h' p + h_0 \leq f(p), \quad \forall p \in \mathbb{R}^n.
\end{aligned}
$$

Equality follows by strong duality for conic linear programs. We provide the details in Appendix A, Lemma 37, a direct application of Shapiro's work [54]. We separate the semi-infinite constraint according to the piecewise definition:

$$
\begin{cases}
p'(H - A_1)p + (h - b_1)'p + (h_0 - c_1) \leq 0, & \\
\qquad\qquad p \in \cup_{i=1}^m \{p' Q_i p + q_i' p + q_i^0 \leq 0\}, & (5.8) \\
p'(H - A_2)p + (h - b_2)'p + (h_0 - c_2) \leq 0, & \\
\qquad\qquad p \in \cap_{i=1}^m \{p' Q_i p + q_i' p + q_i^0 > 0\}. & (5.9)
\end{cases}
$$

We first note that condition (5.7) implies that the constraint on p is redundant in (5.9), as the quadratic with subscript 2 is greater than or equal to the quadratic with subscript 1. Second, we observe that constraint (5.8) is logically equivalent to $p \in \{p' Q_i p + q_i' p + q_i^0 \leq 0\}$ implies the constraint is satisfied. Combining these observations results in the following simplified form:

$$\begin{cases} p'Q_ip + q_i'p + q_i^0 \leq 0 \Rightarrow p'(H - A_1)p + (h - b_1)'p + (h_0 - c_1) \leq 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \forall p \in \mathbb{R}^n, i = 1, 2, \ldots, m, \quad (5.10) \\ p'(H - A_2)p + (h - b_2)'p + (h_0 - c_2) \leq 0, \quad \forall p \in \mathbb{R}^n. \quad (5.11) \end{cases}$$

Equation (5.11) is the statement that the quadratic function defined by H, h, and h_0 should be non-positive for all values of p, which is equivalent to the following matrix being positive semidefinite:

$$\begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h' & -h_0 \end{bmatrix} + \begin{bmatrix} A_2 & \frac{1}{2}b_2 \\ \frac{1}{2}b_2 & c_2 \end{bmatrix} \succeq 0.$$

Equation (5.10), requires additional analysis. For each $i = 1, 2, \ldots, m$, the implication can be rewritten in the following way. For all $p \in \mathbb{R}^n$ such that

$$p'Q_ip + q_i'p + q_i^0 \leq 0,$$

it holds that

$$-((H - A_1)p + (h - b_1)'p + (h_0 - c_1)) \geq 0.$$

This statement is equivalent to the conditions in Lemma 30 (the S-Lemma) for $p = 1$. Hence, (5.10) is equivalent to the following infinite system of quadratic inequalities having a non-negative solution in y:

$$-(p'(H - A_1)p + (h - b_1)'p + (h_0 - c_1)) + y(p'Q_ip + q_i'p + q_i^0) \geq 0, \forall p \in \mathbb{R}^n. \quad (5.12)$$

Finally, the quadratic system (5.12) is equivalent to following semidefinite constraint:

$$\begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h' & -h_0 \end{bmatrix} + y_i \begin{bmatrix} Q_i & \frac{1}{2}q_i \\ \frac{1}{2}q_i' & q_i^0 \end{bmatrix} + \begin{bmatrix} A_1 & \frac{1}{2}b_1 \\ \frac{1}{2}b_1 & c_1 \end{bmatrix} \succeq 0.$$

Combining these two equivalent expressions completes the reformulation. □

Corollary 32. *Given portfolio u, the worst-case semivariance $z = \min_{\pi \in \Pi} -E_\pi(\eta - p'u)_+^2$ is given by solving the following semidefinite program:*

$$
\begin{aligned}
z = \max_{H,h,h_0,y} \quad & P \cdot H + \mu' h + h_0, \\
\text{s.t.} \quad & \begin{bmatrix} -(H + uu') & -\frac{1}{2}(h - (2\eta + y)u) \\ -\frac{1}{2}(h - (2\eta + y)u)' & -(h_0 + y\eta + \eta^2) \end{bmatrix} \succeq 0, \\
& \begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h & -h_0 \end{bmatrix} \succeq 0, \\
& y \geq 0, \\
& H \in \mathbb{S}.
\end{aligned}
$$

Proof. This result is an immediate consequence of Theorem 31 with the following substitutions and $m = 1$:

$$A_1 = -uu', \quad b_1 = 2\eta u, \quad c_1 = -\eta^2, \tag{5.13}$$

$$A_2 = 0, \qquad b_2 = 0, \qquad c_2 = 0, \tag{5.14}$$

$$Q_1 = 0, \qquad q_1 = u, \qquad q_1^0 = -\eta. \tag{5.15}$$

With these substitutions, we observe that $f(p)$ is equal to the negative semivariance, and the resulting semidefinite program is as given. □

Corollary 33 addresses the simplification when the piecewise quadratic function is degenerate, meaning either $E_1 = \emptyset$ or $E_2 = \emptyset$.

Corollary 33. *Let f be a quadratic function denoted by the following expression:*

$$p'Ap + b'p + c.$$

Then the value of $z = \inf_{\mathbb{P}\in\mathcal{P}} E_{\mathbb{P}} f(p)$ is equal to the optimal objective value of the following semidefinite program:

$$
\begin{aligned}
z = \max_{H,h,h_0,y} \quad & P \cdot H + \mu' g + q, \\
\text{s. t.} \quad & \begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h' & -h_0 \end{bmatrix} + \begin{bmatrix} A & \frac{1}{2}b \\ \frac{1}{2}b & c \end{bmatrix} \succeq 0, \\
& H \in \mathbb{S}.
\end{aligned}
$$

Proof. This result follows from the same nonnegativity requirements of quadratic functions used to simplify (5.11). □

We note that the result in Theorem 31 generalizes Jiang and Guan's result [50] in that under the following substitution our result simplifies to theirs;

$$A_1 = 0, \quad b_1 = 0, \quad c_1 = 0, \tag{5.16}$$

$$A_2 = 0, \quad b_2 = 0, \quad c_2 = 1, \tag{5.17}$$

$$Q_i = 0, \quad q_i = -a_i, \quad q_i^0 = b_i. \tag{5.18}$$

The result also generalizes to include the Quadratic chance constraint used in the proof by Rujeerapaiboon et al. [35] with the following substitution with $m = 1$:

$$A_1 = 0, \quad b_1 = 0, \quad c_1 = 0, \tag{5.19}$$

$$A_2 = 0, \quad b_2 = 0, \quad c_2 = 1, \tag{5.20}$$

$$Q_1 = Q, \quad q_1 = q, \quad q_1^0 = q^0. \tag{5.21}$$

The resulting formulation is slightly different from the one used by Rujeerapaiboon et al. [35], so the analytical proof they provide would need to be reworked to show that an equivalent result can be derived.

Using corollary 32 we can now provide the complete worst-case semivariance portfolio problem statement:

$$\max_{H,h,h_0,y,u} \quad \mathrm{E}p'u, \tag{5.22}$$

$$\text{s. t.} \quad P \cdot H + \mu' h + h_0 \geq -s, \tag{5.23}$$

$$\begin{bmatrix} -(H + uu') & -\frac{1}{2}(h - (2\eta + y)u) \\ -\frac{1}{2}(h - (2\eta + y)u)' & -(h_0 + y\eta + \eta^2) \end{bmatrix} \succeq 0, \tag{5.24}$$

$$\begin{bmatrix} -H & -\frac{1}{2}h \\ -\frac{1}{2}h & -h_0 \end{bmatrix} \succeq 0, \tag{5.25}$$

$$y \geq 0, \tag{5.26}$$

$$H \in \mathbb{S}. \tag{5.27}$$

This problem is not convex jointly in all five variables, but it is convex for either y or u fixed. Because y is one-dimensional, we can perform a simple binary search on y to determine the value required to jointly optimize y and u. However, we show that such a search is unnecessary as the problem is solvable analytically.

5.4 Analytical Solution to Worst-Case Semivariance

We observe that semivariance is a function of the asset returns only through the portfolio's overall return. This observation allows us to project the ambiguity set from n dimensions to 1 dimension. Under this projection we reduce the semidefinite program to a nonlinear optimization problem in four variables.

Let $\mathcal{P}(\mu, P)$ denote the set of probability distributions with mean μ and second moment matrix P. We apply Corollary 1 of Yu, Li, Schuurmans, and Szepesvári [57] to conclude

$$\min_{p \sim \mathcal{P}(\mu,P)} -\mathrm{E}_\pi(\eta - p'u)_+^2 = \min_{\phi \sim \mathcal{P}(\mu'u, u'Pu)} -\mathrm{E}_\pi(\eta - \phi)_+^2.$$

We note that the semivariance in terms of ϕ is the expectation of a 1-dimensional distribution. Repeating the derivation of Theorem 31 for the 1-dimensional semivariance, yields the following SDP:

$$\max_{M\in\mathbb{S}^2, y\in\mathbb{R}} \quad M \cdot \begin{bmatrix} u'Pu & \mu'u \\ \mu'u & 1 \end{bmatrix}, \tag{5.28}$$

$$\text{s. t.} \quad -M - \begin{bmatrix} 1 & -\eta \\ -\eta & \eta^2 \end{bmatrix} - y \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & \eta \end{bmatrix} \succeq 0, \tag{5.29}$$

$$-M \succeq 0, \tag{5.30}$$

$$y \geq 0. \tag{5.31}$$

With this SDP in 4 decision variables, we provide an analytical solution which can be used to simplify the robust mean-semivariance portfolio problem.

Theorem 34. *If $\mu'u - \eta \geq 0$, the SDP in* (5.28)–(5.31) *has the following optimal value function:*

$$-u'Pu + (\mu'u)^2 = -V(p'u),$$

where $V(p'u)$ is the variance of $p'u$.

If $\eta - \mu'u > 0$, then the optimal value function is

$$-u'Pu + 2\eta\mu'u - \eta^2 = -E(\eta - p'u)^2.$$

Proof. We adopt the following notation:

$$M = \begin{bmatrix} m_1 & m_2 \\ m_2 & m_3 \end{bmatrix}, \tag{5.32}$$

$$\tilde{P} = u'Pu, \tag{5.33}$$

$$\tilde{\mu} = \mu'u. \tag{5.34}$$

The semidefinite constraints for a 2 x 2 matrix can be explicitly written as a system of nonlinear equations using the definition for semidefinite matrices to yield the following optimization problem:

$$\max_{m\in\mathbb{R}^3, y\in\mathbb{R}} \quad f(m,y) = \tilde{P}m_1 + 2\tilde{\mu}m_2 + m_3, \tag{5.35}$$

$$\text{s. t.} \quad g_1(m,y) = m_2^2 - m_1 m_3 \leq 0, \tag{5.36}$$

$$g_2(m,y) = (m_2 - \eta - \frac{1}{2}y)^2 - (m_1+1)(m_3 + \eta^2 + \eta y) \leq 0, \tag{5.37}$$

$$g_3(m,y) = m_3 + \eta^2 + \eta y \leq 0, \tag{5.38}$$

$$g_4(m,y) = m_1 + 1 \leq 0, \tag{5.39}$$

$$g_5(m,y) = m_3 \leq 0, \tag{5.40}$$

$$g_6(m,y) = -y \leq 0. \tag{5.41}$$

We divide the proof into two cases. The first case has $\tilde{\mu} > \eta$. The second case has $\eta > \tilde{\mu}$. The first case is verified explicitly by constructing a solution and Lagrange multipliers satisfying the KKT conditions. The solution is constructed by supposing certain constraints are tight thereby reducing the problem to an optimization in a single variable. Full details are provided in Appendix B. The second case is verified by showing the existence of a probability distribution in the set $\mathcal{P}(\tilde{\mu}, \tilde{P})$ achieving the given maximum.

We focus first on the case $\tilde{\mu} > \eta$. The solution is a result of supposing constraints (5.36) and (5.39) are tight. This supposition is based on the objective increasing with respect to m_3, and the intuition that we likely wish to maximize $|m_2|$ in the direction of the sign of $\tilde{\mu}$.

Together these hypotheses simplify the optimization problem to the following:

$$\begin{aligned}
\max_{m_2 \in \mathbb{R}, y \in \mathbb{R}} \quad & -m_2^2 + 2\tilde{\mu} m_2 - \tilde{P}, && (5.42)\\
\text{s.t.} \quad & m_2 = \eta + \frac{1}{2} y, && (5.43)\\
& -m_2^2 + \eta^2 + \eta y \leq 0, && (5.44)\\
& y \geq 0. && (5.45)
\end{aligned}$$

We further simplify the system by substituting (5.43) throughout to reduce the problem to a single variable optimization:

$$\begin{aligned}
\max_{y \in \mathbb{R}} \quad & -\frac{1}{4} y^2 + (\tilde{\mu} - \eta) y + 2\tilde{\mu}\eta - \eta^2 - \tilde{P},\\
\text{s.t.} \quad & y \geq 0.
\end{aligned}$$

In this case, the optimal solution is clearly either $y^* = 2(\tilde{\mu} - \eta)$ if $\tilde{\mu} - \eta > 0$ or $y^* = 0$, otherwise. Through the preceding process we arrive at a feasible solution. Substituting it into the Langragian system results in following optimal solutions, with Lagrangian multipliers denoted by λ_i:

$$\begin{aligned}
(m^*, y^*, \lambda^*) &= (m_1^*, m_2^*, m_3^*, y^*, \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*, \lambda_6^*)\\
&= (-1, \tilde{\mu}, -\tilde{\mu}^2, 2(\tilde{\mu} - \eta), 1, \frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2}, 0, 0, 0, 0).
\end{aligned}$$

Details of setting up the Lagrangian are given in Appendix B.

We prove the second case $\eta \geq \tilde{\mu}$ by explicitly constructing a distribution achieving the worst case bound. To motivate how we proceed, suppose we restricted ourselves to only consider distributions having a finite support ϕ_i for $i = 1, 2, \ldots, m$. Given a fixed support, we can formulate a linear program to find the worst case distribution satisfying the moment constraints over following support set:

$$
\begin{aligned}
\max_{\pi \in \mathbb{R}^m} \quad & \sum_{i=1}^{m} \pi_i (\eta - \phi_i)_+^2, \\
\text{s. t.} \quad & \sum_{i=1}^{m} \pi_i \phi_i^2 = \tilde{P}, \\
& \sum_{i=1}^{m} \pi_i \phi_i = \tilde{\mu}, \\
& \sum_{i=1}^{m} \pi_i = 1, \\
& \pi \geq 0.
\end{aligned}
$$

Naturally every basic feasible solution to this linear program has at most 3 nonzero probability weights π_i. In fact, this result is proved in generality by Lemma 3.1 in Shapiro's paper [54]. He shows that the worst-case distribution can be represented as finitely supported on a number of points equal to the number of constraints.

Proceeding with the finitely supported formulation, we state the problem of jointly optimizing π and ϕ for a 3 point distribution:

$$
\begin{aligned}
\max_{\pi \in \mathbb{R}^3, \phi \in \mathbb{R}^3} \quad & \pi_1 (\eta - \phi_1)_+^2 + \pi_2 (\eta - \phi_2)_+^2 + \pi_3 (\eta - \phi_3)_+^2, \\
\text{s. t.} \quad & \pi_1 \phi_1^2 + \pi_2 \phi_2^2 + \pi_3 \phi_3^2 = \tilde{P}, \\
& \pi_1 \phi_1 + \pi_2 \phi_2 + \pi_3 \phi_3 = \tilde{\mu}, \\
& \pi_1 + \pi_2 + \pi_3 = 1, \\
& \pi \geq 0.
\end{aligned}
$$

We first substitute the relation $\pi_3 = 1 - \pi_1 - \pi_2$:

$$
\begin{aligned}
\max_{\pi \in \mathbb{R}^2} \quad & \pi_1((\eta - \phi_1)_+^2 - (\eta - \phi_3)_+^2) + \pi_2((\eta - \phi_2)_+^2 - (\eta - \phi_3)_+^2) \\
& + (\eta - \phi_3)_+^2, \\
\text{s. t.} \quad & \pi_1(\phi_1^2 - \phi_3^2) + \pi_2(\phi_2^2 - \phi_3^2) = \tilde{P} - \phi_3^2, \\
& \pi_1(\phi_1 - \phi_3) + \pi_2(\phi_2 - \phi_3) = \tilde{\mu} - \phi_3, \\
& \pi \geq 0.
\end{aligned}
$$

To simplify further, we set $\phi_1 = \tilde{\mu}$, and we show later that this is the correct choice. With the optimization reduced to two variables we can explicitly solve for π_1 and π_2 as a function of the returns ϕ_2 and ϕ_3:

$$
\begin{aligned}
\pi_1 &= -\frac{\tilde{\mu}\phi_2 - \phi_2\phi_3 + \tilde{\mu}\phi_3 - \tilde{P}}{(\tilde{\mu} - \phi_2)(\tilde{\mu} - \phi_3)}, \\
\pi_2 &= -\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \phi_2)(\phi_2 - \phi_3)}.
\end{aligned}
$$

Now suppose that we can choose ϕ_2 and ϕ_3 so that $\eta \geq \phi_2 > \tilde{\mu}$ ands $\tilde{\mu} \geq \phi_3$, then we can simplify the objective in the following way:

$$
\begin{aligned}
& \pi_1(\eta - \phi_1)_+^2 + \pi_2(\eta - \phi_2)_+^2 + \pi_3(\eta - \phi_3)_+^2 \\
= & \pi_1(\eta - \phi_1)^2 + \pi_2(\eta - \phi_2)^2 + \pi_3(\eta - \phi_3)^2 \\
= & \pi_1(\eta^2 - 2\eta\phi_1 + \phi_1^2) + \pi_2(\eta^2 - 2\eta\phi_2 + \phi_2^2) + \pi_3(\eta^2 - 2\eta\phi_3 + \phi_3^2) \\
= & \eta^2 - 2\eta\tilde{\mu} + \tilde{P} \\
= & \mathrm{E}(\eta - \phi)^2 \\
\geq & \mathrm{E}(\eta - \phi)_+^2.
\end{aligned}
$$

Crucially, the last equality holds for all distributions in the set $\mathcal{P}(\tilde{\mu}, \tilde{P})$. The last inequality holds in general because the expectation is over a nonnegative function and the same nonnegative function constrained by an indicator function. This shows that any choice of ϕ_2 and ϕ_3 satisfying the prior conditions maximizes the semivariance over the family of distributions $\mathcal{P}(\tilde{\mu}, \tilde{P})$. We now proceed to show that such a choice is always possible when $\eta > \tilde{\mu}$.

Using the values of π_1 and π_2, the nonnegativity constraints imply the following constraints on the choice of ϕ_2 and ϕ_3:

$$\begin{aligned} 1 - \frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \phi_2)(\tilde{\mu} - \phi_3)} &\geq 0, \\ \frac{\tilde{\mu}^2 - \tilde{P}}{(\tilde{\mu} - \phi_2)(\tilde{\mu} - \phi_3)} &\geq 0, \\ \frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \phi_2)(\tilde{\mu} - \phi_3)} &\geq 0. \end{aligned}$$

Combined with the conditions from before, this system simplifies to a single inequality:

$$\phi_3 \leq \frac{\tilde{\mu}\phi_2 - \tilde{P}}{\phi_2 - \tilde{\mu}}. \tag{5.46}$$

Hence, for any ϕ_2 chosen in the interval $(\tilde{\mu}, \eta)$, by choosing ϕ_3 sufficiently negative, we can satisfy the given conditions. □

To help illustrate the results of Theorem 34, we provide some example worst case distributions in Table 5.1. We see that if ϕ_3 is chosen to equal the bound given in equation (5.46), then the distribution reduces to a two point distribution.

We now discuss the consequence of Theorem 34. Let u^* be the optimal robust mean-semivariance portfolio. If $\mu' u^* \geq \eta$, then the worst-case semivariance is equal to the variance, and so u^* is equal to the mean-variance portfolio. If $\mu' u^* < \eta$, then the worst-case semivariance is equal to $\eta^2 - 2\eta\mu' u^* + u^{*\prime} P u^*$, the second moment centered at η, which is fixed over this uncertainty set. This second moment optimal portfolio is not always equivalent to a mean-variance portfolio, but often similar. The following proposition makes this comparison rigorously.

Table 5.1: Example worst case distributions from Theorem 34.

η	$\tilde{\mu}$	$\tilde{P}$	π_1	π_2	π_3	ϕ_1	ϕ_2	ϕ_3	$\frac{\tilde{\mu}\phi_2-\tilde{P}}{\phi_2-\tilde{\mu}}$
1	0.1	1	0.0000	0.5500	0.4500	0.1	1	-1	-1
1	0.1	1	0.3125	0.4400	0.2475	0.1	1	-1.5	-1
1	0.1	1	0.4762	0.3667	0.1571	0.1	1	-2	-1
1	0.1	1	0.5769	0.3143	0.1088	0.1	1	-2.5	-1
1	0.1	1	0.6452	0.2750	0.0798	0.1	1	-3	-1
1	0.1	1	0.6944	0.2444	0.0611	0.1	1	-3.5	-1

Proposition 35. *Let u^* be the optimal solution to the robust mean-semivariance problem given in equations (5.22)–(5.27). There exists an $\alpha \geq 0$ such that the following holds:*

$$u^* = \begin{cases} \frac{1}{2\alpha}\Sigma^{-1}\mu & \mu' u^* \geq \eta, \\ \frac{1+2\eta\alpha}{2\alpha}P^{-1}\mu & \mu' u^* < \eta, \end{cases}$$

or equivalently

$$u^* = \begin{cases} \frac{1}{2\alpha}\Sigma^{-1}\mu & \frac{1}{2\alpha}\mu'\Sigma^{-1}\mu \geq \eta, \\ \frac{1+2\eta\alpha}{2\alpha}P^{-1}\mu & \frac{1+2\eta\alpha}{2\alpha}\mu' P^{-1}\mu < \eta. \end{cases}$$

Proof. This result is a direct consequence of Theorem 34. The key idea is that for a fixed portfolio, we can apply Theorem 34 to re-evaluate the worst-case semivariance in terms of either the covariance matrix or the second moment matrix. We then observe that the robust mean-semivariance problem (5.22)–(5.27) is equivalent to either a mean-variance problem or a mean-second-moment problem for all u within a neighborhood of the optimal solution u^*. The optimal solution for each of the variance or moment problems has the form given in the statement of the proposition. □

Proposition 35 shows that except for cases when the excess return target η is set aggressively large or the risk aversion parameter α is set to a large value, the robust mean-semivariance portfolio is equivalent to a mean-variance portfolio. This result mirrors a result by

Rujeerapaiboon et al. [35] in their study of robust growth-optimal (log-wealth) portfolios under the same uncertainty set that we consider (weak-sense white noise). They approximate the robust growth-optimal (log-wealth) portfolio by a worst-case value-at-risk of a quadratic approximation of the growth rate. They show that this worst-case value-at-risk problem can be solved explicitly and yields a mean-variance portfolio.

Our result, however, differs from the related work of Gotoh, Kim, and Lim [53] because they consider a different uncertainty set. They consider the family of probability measures constrained by ϕ-divergence, which is different from the weak-sense white noise ambiguity set. They show a general property that ϕ-divergence robust optimization problems can be approximated by mean-variance optimization problems. A generalization of the mean-variance similarity observed in this work and Rujeerapaiboon et al. [35] does not to our knowledge exist for the weak-sense white noise uncertainty paradigm.

Our proof methodology from proposition 32 extends the proof methodology of the working paper by Jiang and Guan [50] and the work of Zymler et al. [36]. Their proofs addressed worst-case chance constraints, which translate into indicator functions in the proof. We extend the methodology to piece-wise quadratic functions that can be expressed as the product of a quadratic term and an indicator function. The relationship to the mean-variance objective depended on the similarity of the semivariance to the second moment, which was held fixed in the robust formulation. The analytic result also relied on the ability to apply Corollary 1 of Yu et al. [57] to reduce the dimensionality of the optimization problem.

5.5 Robust Semivariance with Support Information

Following an approach similar to Rujeerapaiboon et al. [35], we add support constraints to the formulation for robust semivariance. The extension constrains the support of each distribution in the distribution family. The objective is to mitigate the risk of degenerate distributions, some consequences of which are studied by Kleywegt and Gao [48].

Let $\mathcal{P}$ be the set of probability measures with a given mean μ and second moment P, i.e.

$$\mathcal{P} = \left\{ \mathbb{P} : \int p d\mathbb{P} = \mu, \int pp' d\mathbb{P} = P \right\}.$$

Using the notion of support constraints defined by Zymler et al. [36], we define the following uncertainty set:

$$\mathcal{P}_{\Xi} = \left\{ \mathbb{P} : \int p d\mathbb{P} = \mu, \int pp' d\mathbb{P} = P, \mathbb{P}(\Xi) = 1 \right\}.$$

The set Ξ is defined by

$$\Xi = \left\{ p \in \mathbb{R}^n : \begin{bmatrix} p \\ 1 \end{bmatrix}' W_i \begin{bmatrix} p \\ 1 \end{bmatrix} \leq 0 \right\},$$

for some matrices W_i, $i = 1, 2, \ldots, m$. Applying a similar proof methodology as Zymler et al. [36], the worst case semivariance $\max_{\mathbb{P} \in \mathcal{P}\Xi} \mathrm{E}(\eta - u'p)_+^2$ can be conservatively approximated by an SDP.

Proposition 36. *The worst case semivariance with respect to the uncertainty set $\mathcal{P}_{\Xi}$,*

$$\max_{\mathbb{P} \in \mathcal{P}\Xi} \boldsymbol{E}(\eta - u'p)_+^2,$$

is conservatively approximated by the optimal value of the following SDP:

$$\begin{aligned}
\max_{\substack{M\in\mathbb{S}^{n+1},\\ y\in\mathbb{R},\\ \tau\in\mathbb{R}^m}} \quad & M \cdot \begin{bmatrix} P & \mu \\ \mu' & 1 \end{bmatrix}, \\
\text{s.t.} \quad & -M - \begin{bmatrix} uu' & -\eta u \\ -\eta u & \eta^2 \end{bmatrix} - y \begin{bmatrix} 0 & -\frac{1}{2}u \\ -\frac{1}{2}u & \eta \end{bmatrix} + \sum_{i=1}^{m} \tau_i W_i \succeq 0, \\
& -M + \sum_{i=1}^{m} \tau_i W_i \succeq 0, \\
& \tau \geq 0, \\
& y \geq 0.
\end{aligned}$$

Proof. By Lemma 38, the worst case semivariance is equivalent to the following semi-infinite program:

$$\begin{aligned}
\zeta = \max_{H\in\mathbb{S},h,h_0} \quad & P \cdot H + \mu' h + h_0, \\
\text{s.t.} \quad & p'Hp + h'p + h_0 \leq -(\eta - u'p)^2 [\eta - u'p > 0], \quad \forall p \in \Xi.
\end{aligned}$$

The indicator function allows the semi-infinite constraint to be separated into two semi-infinite constraints:

$$\begin{cases} \eta - u'p \geq 0 \Rightarrow p'Hp + h'p + h_0 \leq -(\eta - u'p)^2, & \forall p \in \Xi, \quad (5.47a) \\ p'Hp + h'p + h_0 \leq 0, & \forall p \in \Xi. \quad (5.47b) \end{cases}$$

We substitute the matrix variable M using the relation $M = \begin{bmatrix} H & \frac{1}{2}h \\ \frac{1}{2}h & h_0 \end{bmatrix}$. Equation (5.47b), by the S-lemma (Lemma 30), is implied by the following semidefinite constraint system:

$$\begin{cases} -M + \sum_{i=1}^{m} \tau_i W_i \succeq 0, \\ \tau \geq 0. \end{cases}$$

Similar to the proof without the support constraint on the uncertainty set, the implication constraint (5.47a) can be rewritten as the following:

$$\begin{bmatrix} p \\ 1 \end{bmatrix}' M \begin{bmatrix} p \\ 1 \end{bmatrix} + (\eta - u'p)^2 \leq 0, \tag{5.48}$$

for all $p \in \Xi$ such that

$$\begin{cases} -(\eta - u'p) \leq 0, & (5.49a) \\ \begin{bmatrix} p \\ 1 \end{bmatrix}' W_i \begin{bmatrix} p \\ 1 \end{bmatrix} \leq 0, \quad i = 1, 2, \ldots, m. & (5.49b) \end{cases}$$

This is the same as the conditions of the S-lemma (Lemma 30). Hence the following semidefinite constraint system:

$$-M - \begin{bmatrix} uu' & -\eta u \\ -\eta u & \eta^2 \end{bmatrix} - y \begin{bmatrix} 0 & -\frac{1}{2}u \\ -\frac{1}{2}u & \eta \end{bmatrix} + \sum_{i=1}^{m} \tau_i W_i \succeq 0, \tag{5.50}$$

$$\tau \geq 0, \tag{5.51}$$

$$y \geq 0, \tag{5.52}$$

implies (5.47a). Combining the two systems of semidefinite constraints, we conclude the semidefinite constraints imply the semi-infinite constraints, and the optimal value of the SDP is a conservative approximation of the original problem. □

5.6 Robust Error-Tracking Mean-Semivariance Problem

We illustrate a novel application of Theorem 31 to error-tracking portfolios. Our application can be viewed as a partial extension to the work of Ling, Sun, and Yang [51]. They consider the case of worst-case probability of shortfall and worst-case shortfall, in which shortfall is measured relative to a given benchmark portfolio. They consider of the problem of using this portfolio to track the SSE50 index of the Shanghai Stock Exchange. They consider both distribution uncertainty and parameter uncertainty. Our formulation extends to the case of worst-case semivariance under distribution uncertainty only.

Let p_0 denote the excess returns of the benchmark portfolio and $\bar{p}$ denote the excess returns of the assets which will compose our portfolio. Using the same basic notation as previously, we have the following worst case semivariance problem:

$$\max_{u \in U} \quad \mathrm{E}_\pi(\bar{p}'u), \tag{5.53}$$

$$\text{s.t.} \quad \min_{\pi \in \Pi} -\mathrm{E}_\pi(\eta - (p_0 - \bar{p}'u))_+^2 \geq -s. \tag{5.54}$$

To see how the semivariance can be expressed in the form necessary for Theorem 31, we rewrite it in the following way:

$$(\eta - (p_0 - \bar{p}'u))_+^2 = \left(\eta - \begin{bmatrix} p_0 \\ \bar{p} \end{bmatrix}' \begin{bmatrix} 1 \\ -u \end{bmatrix}\right)_+^2.$$

We can then apply Theorem 31 with the following substitutions:

$$\begin{aligned}
A_1 &= -\begin{bmatrix} 1 \\ -u \end{bmatrix} \begin{bmatrix} 1 & -u' \end{bmatrix}, \quad & b_1 &= 2\eta \begin{bmatrix} 1 \\ -u \end{bmatrix}, \quad & c_1 &= -\eta^2, \\
A_2 &= 0, & b_2 &= 0, & c_2 &= 0, \\
Q_1 &= 0, & q_1 &= \begin{bmatrix} 1 \\ -u \end{bmatrix}, & q_1^0 &= -\eta, \\
p &= \begin{bmatrix} p_0 \\ \bar{p} \end{bmatrix} \begin{bmatrix} p_0 & \bar{p} \end{bmatrix}, & P &= \mathrm{E}(pp'), & \mu &= \mathrm{E}(p), \\
\bar{\mu} &= \mathrm{E}(\bar{p}).
\end{aligned}$$

After expanding the the outer products and applying Theorem 31, we yield the following

SDP:

$$
\begin{aligned}
\max_{M,y,u} \quad & \bar{\mu}'u, \\
\text{s.t.} \quad & M \cdot \begin{bmatrix} P & \mu \\ \mu' & 1 \end{bmatrix} \geq -s, \\
& -M + y \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2}u \\ \frac{1}{2} & -\frac{1}{2}u' & -\eta \end{bmatrix} + \begin{bmatrix} -1 & u' & \eta \\ u & -uu' & -\eta u \\ \eta & -\eta u' & -\eta^2 \end{bmatrix} \succeq 0, \\
& -M \succeq 0, \\
& y_i \geq 0, \\
& M \in \mathbb{S}^{n+2}.
\end{aligned}
$$

Applying the results of Section 5.4, we can show that this formulation is actually equivalent to a variance or second-moment error-tracking portfolio.

5.7 Conclusions and Future Research

We have made two distinct contributions in this chapter. First, we showed that fixed moment-based robustness when applied to semivariance is equivalent to non-robust variance or the second-moment. This result suggests that semivariance is sensitive to the estimation procedures for covariance. We also observe that the worst case distribution is degenerate in the sense that a typical investor does not consider only 3 return outcomes when evaluating portfolio risk. Consequently, we should explore alternative forms of robustness, such as ϕ-divergence and Wasserstein distance as they may be more suitable for semivariance.

Second, we developed a novel SDP framework for modeling robust piecewise quadratic problems consisting of two-distinct pieces. This framework encompasses chance constraints, shortfall, and semivariance. We apply this formulation to semivariance to derive the analytical form for the worst case semivariance. We believe it may be possible to consolidate the work of Zymler et al. [36] and Rujeerapaiboon et al. [35] using our SDP formulation. Such a result would allows us to have a single proof for worst case chance constraints, value-at-risk, and semivariance with moment-based ambiguity sets.

After moment-base uncertatinty, the two most likely alternatives are ϕ-divergence and Wasserstein distance ambiguity sets. In Jiang and Guan's [47] work, they show that ϕ-divergence worst case chance constraints are equivalent to a re-scaled risk target. We posit the indicator function component of the semivariance function is structurally similar to a chance-constraint and may benefit from a similar analysis. An important drawback to ϕ-divergence is that for discrete distributions, such as the empirical distribution, the worst-case distribution may be unrealistically pathological. Kleywegt and Gao [48] provide an illustrative example based on image processing. Alternatively, Wasserstein distance shows particular promise because a distance metric can be chosen based on metrics relevance financial time series data. Kleywegt and Gao [48] provide an application to the Newsvendor problem, which suggests the formulation may be tractable for a piecewise function like the semivariance function.

Appendices

APPENDIX A

Conic Duality for Moment Problems

We apply the results of Shapiro [54] for conic linear programs to the moments problem from Chapter 5. An alternative approach to strong duality is given earlier by Kemperman [58] without reference to conic duality. Lemma 37 summarizes the result needed for Theorem 31. The extension to incorporate support constraints is given in Lemma 38.

Lemma 37. *Let f be given as in Theorem 31. The primal optimization problem,*

$$\begin{aligned} z = \inf_{\pi \in \mathcal{P}} \quad & \int_{\mathbb{R}^n} f(p) d\pi, \\ \text{s.t.} \quad & \int_{\mathbb{R}^n} p d\pi = \mu, \\ & \int_{\mathbb{R}^n} pp' d\pi = P, \\ & \int_{\mathbb{R}^n} d\pi = 1, \end{aligned}$$

has the same optimal value z as the following dual formulation:

$$\begin{aligned} z = \max_{H \in \mathbb{S}, h, h_0} \quad & P \cdot H + \mu' h + h_0, \\ \text{s.t.} \quad & p' H p + h' p + h_0 \leq f(p), \quad \forall p \in \mathbb{R}^n. \end{aligned}$$

Proof. Following the conventions in Shapiro's [54] work, we note that the integral is an inner product on the joint space of nonnegative measures and real-valued functions. Using this property he shows that these optimization problems form a primal-dual pair. The strong duality is then an immediate consequence of Proposition 3.4 in [54] by observing that the right hand side of the primal optimization satisfies condition (3.12) in [54]. □

Lemma 38. *Let f be given as in Theorem 31. The primal optimization problem,*

$$
\begin{aligned}
z = \inf_{\pi \in \mathcal{P}_\Xi} \quad & \int_{\mathbb{R}^n} f(p) d\pi, \\
\text{s.t.} \quad & \int_{\mathbb{R}^n} p d\pi = \mu, \\
& \int_{\mathbb{R}^n} pp' d\pi = P, \\
& \int_{\mathbb{R}^n} d\pi = 1,
\end{aligned}
$$

is equivalent to the following dual formulation:

$$
\begin{aligned}
z = \max_{H \in \mathbb{S}, h, h_0} \quad & P \cdot H + \mu' h + h_0, \\
\text{s.t.} \quad & p' H p + h' p + h_0 \leq f(p), \quad \forall p \in \Xi.
\end{aligned}
$$

Proof. The proof is essentially the same as Lemma 37 through the following observation. Shapiro's [54] argument requires only that $\mathcal{P}_\Xi$ be a cone generated by a convex set of nonnegative signed measures. It is therefore sufficient to show that $\mathcal{P}_\Xi$ is convex. Let π and ν be elements of $\mathcal{P}_\Xi$ and $t \in (0, 1)$. Observe:

$$
(t\pi + (1-t)\nu)(\Xi) = t\pi(\Xi) + (1-t)\nu(\Xi) = 1,
$$

because $\pi(\Xi) = \nu(\Xi) = 1$. Hence $t\pi + (1-t)\nu \in \mathcal{P}_\Xi$, and $\mathcal{P}_\Xi$ is convex. The transformation to the semi-infinite constraint can therefore be restricted to point masses in $\mathcal{P}_\Xi$, giving the desired result. □

APPENDIX B

Detailed Proof of Theorem 34

We present in the following a detailed proof of case 1 ($\mu' u - \eta \geq 0$) from Theorem 34 in section 5.4.

Proof: Case 1 ($\mu' u - \eta \geq 0$) from Theorem 34. The SDP (5.28)–(5.31) is equivalent to the following nonlinear program:

$$
\begin{aligned}
\max_{m \in \mathbb{R}^3, y \in \mathbb{R}} \quad & f(m, y) = \tilde{P} m_1 + 2\tilde{\mu} m_2 + m_3, \\
\text{s. t.} \quad & g_1(m, y) = m_2^2 - m_1 m_3 \leq 0, \\
& g_2(m, y) = (m_2 - \eta - \frac{1}{2} y)^2 - (m_1 + 1)(m_3 + \eta^2 + \eta y) \leq 0, \\
& g_3(m, y) = m_3 + \eta^2 + \eta y \leq 0, \\
& g_4(m, y) = m_1 + 1 \leq 0, \\
& g_5(m, y) = m_3 \leq 0, \\
& g_6(m, y) = -y \leq 0.
\end{aligned}
$$

The equivalence to the SDP is an immediate consequence of converting the semidefinite constraints into explicit constraints on the nonnegativity of all principal minors of the matrices.

The following Jacobian matrix is used in the Lagrangian:

$$
\begin{aligned}
J(m, y)' &= \begin{bmatrix} \nabla g_1 & \nabla g_2 & \nabla g_3 & \nabla g_4 & \nabla g_5 & \nabla g_6 \end{bmatrix} \\
&= \begin{bmatrix} -m_3 & -(m_3 + \eta^2 + \eta y) & 0 & 1 & 0 & 0 \\ 2m_2 & 2(m_2 - \eta - \frac{1}{2} y) & 0 & 0 & 0 & 0 \\ -m_1 & -(m_1 + 1) & 1 & 0 & 1 & 0 \\ 0 & -(m_2 - \eta - \frac{1}{2} y) - \eta(m_1 + 1) & \eta & 0 & 0 & -1 \end{bmatrix}.
\end{aligned}
$$

The gradient of the objective is given by the following:

$$\nabla f(m, y) = \begin{bmatrix} \tilde{P} \\ 2\tilde{\mu} \\ 1 \\ 0 \end{bmatrix}.$$

Let $G = (g_1, g_2, g_3, g_4, g_5, g_5)$ denote the vector of constraint functions. Then the KKT conditions are satisfied by a solution (m^*, y^*, λ^*) if

$$\nabla f(m^*, y^*) = J(m^*, y^*)'\lambda^*, \tag{B.1}$$

$$G(m^*, y^*) \leq 0, \tag{B.2}$$

$$\lambda^* \geq 0, \tag{B.3}$$

$$G(m^*, y^*)'\lambda^* = 0, \tag{B.4}$$

where λ_i is the dual variable associated with g_i.

Suppose $\tilde{\mu} - \eta > 0$. We now proceed to solve case 1 by showing that the following solution satisfies the KKT conditions:

$$\begin{aligned} (m^*, y^*, \lambda^*) &= (m_1^*, m_2^*, m_3^*, y^*, \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*, \lambda_6^*) \\ &= (-1, \tilde{\mu}, -\tilde{\mu}^2, 2(\tilde{\mu} - \eta), 1, \frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2}, 0, 0, 0, 0). \end{aligned}$$

First, we verify Lagrange stationarity (B.1).

$$
\begin{aligned}
\tilde{P} &= -m_3\lambda_1 - (m_3 + \eta^2 + \eta y)\lambda_2 + \lambda_4 \\
&= -(-\tilde{\mu}^2)(1) - (-\tilde{\mu}^2 + \eta^2 + 2\eta(\tilde{\mu} - \eta))\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} + (0) \\
&= \tilde{\mu}^2 - (-\tilde{\mu}^2 + 2\tilde{\mu}\eta - \eta^2)\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} \\
&= \tilde{P}. \\
2\tilde{\mu} &= 2m_2\lambda_1 + 2(m_2 - \eta - \frac{1}{2}y)\lambda_2 \\
&= 2(\tilde{\mu})(1) + 2(\tilde{\mu} - \eta - \frac{1}{2}(2)(\tilde{\mu} - \eta))\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} \\
&= 2\tilde{\mu} + (\tilde{\mu} - \tilde{\mu} - \eta + \eta)\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} \\
&= 2\tilde{\mu}. \\
1 &= -m_1\lambda_1 - (m_1 + 1)\lambda_2 + \lambda_3 + \lambda_5 \\
&= -(-1)(1) - (-1 + 1)\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} + (0) + (0) \\
&= 1. \\
0 &= -((m_2 - \eta - \frac{1}{2}y) + \eta(m_1 + 1))\lambda_2 + \eta\lambda_3 - \lambda_6 \\
&= -((\tilde{\mu} - \eta - \frac{1}{2}(2)(\tilde{\mu} - \eta)) - \eta(-1 + 1))\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} + \eta(0) - (0) \\
&= 0.
\end{aligned}
$$

Next, we verify primal feasibility (B.2).

$$
\begin{aligned}
& m_2^2 - m_1 m_3 \\
&= (\tilde{\mu})^2 - (-1)(-\tilde{\mu}^2) \\
&= 0. \\
& (m_2 - \eta - \frac{1}{2}y)^2 - (m_1 + 1)(m_3 + \eta^2 + \eta y) \\
&= (\tilde{\mu} - \eta - \frac{1}{2}(2)(\tilde{\mu} - \eta))^2 - (-1 + 1)(-\tilde{\mu}^2 + \eta^2 - \eta(2)(\tilde{\mu} - \eta)) \\
&= 0. \\
& m_3 + \eta^2 + \eta y \\
&= -\tilde{\mu}^2 + \eta^2 + \eta(2)(\tilde{\mu} - \eta) \\
&= -\tilde{\mu}^2 + 2\tilde{\mu}\eta - \eta^2 \\
&= -(\tilde{\mu} - \eta)^2 \\
&\leq 0. \\
& m_1 + 1 \\
&= -1 + 1 \\
&= 0. \\
& m_3 \\
&= -\tilde{\mu}^2 \\
&\leq 0. \\
& -y \\
&= -2(\tilde{\mu} - \eta) \\
&\leq 0.
\end{aligned}
$$

Next, we verify dual feasibility (B.3).

$$\begin{aligned}
\lambda_1 &= 1 \\
&\geq 0. \\
\lambda_2 &= \frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} \\
&= \frac{\mathbf{V}(p'u)}{(\tilde{\mu} - \eta)^2} \\
&\geq 0. \\
\lambda_3 &= 0. \\
\lambda_4 &= 0. \\
\lambda_5 &= 0. \\
\lambda_6 &= 0.
\end{aligned}$$

Next, we verify complementary slackness (B.4).

$$\begin{aligned}
g_1(m^*, y^*)\lambda_1 &= (0)(1) \\
&= 0. \\
g_2(m^*, y^*)\lambda_2 &= (0)\frac{\tilde{P} - \tilde{\mu}^2}{(\tilde{\mu} - \eta)^2} \\
&= 0. \\
g_3(m^*, y^*)\lambda_3 &= -(\tilde{\mu} - \eta)^2(0) \\
&= 0.
\end{aligned}$$

$$
\begin{aligned}
g_4(m^*, y^*)\lambda_4 &= (0)(0) \\
&= 0. \\
g_5(m^*, y^*)\lambda_5 &= -\tilde{\mu}^2(0) \\
&= 0. \\
g_6(m^*, y^*)\lambda_6 &= -2(\tilde{\mu} - \eta)(0) \\
&= 0.
\end{aligned}
$$

Thus, the proposed solution satisfies the KKT conditions. □

REFERENCES

[1] O. Goldsmith, *The citizen of the world: Or, letters from a chinese philosopher, residing in london, to his friends in the east*. Dublin, 1762, ch. Letter VII.

[2] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[3] ——, *Portfolio selection*. New York: John Wiley & Sons, 1959.

[4] R. Michaud, *Efficient asset management: A practical guide to stock portfolio management and asset allocation, survey and synthesis series*. HBS Press, Boston, MA, 1998.

[5] J. Estrada, "Mean-semivariance behavior: Downside risk and capital asset pricing," *International Review of Economics and Finance*, vol. 16, pp. 169–185, 2007.

[6] W. Hogan and J. Warren, "Toward the development of an equilibrium capital-market model based on semivariance," *The Journal of Financial and Quantitative Analysis*, vol. 9, no. 1, pp. 1–11, Jan. 1974.

[7] J. Mao, "Model of capital budgeting, e-v vs e-s," *The Journal of Financial and Quantitative Analysis*, vol. 4, no. 5, pp. 657–675, Jan. 1970.

[8] T. Nantell and B. Price, "An analytical comparison of variance and semivariance capital market theories," *Journal of Financial and Quantitative Analysis*, vol. 14, no. 2, pp. 221–242, Jun. 1979.

[9] D. Nawrocki, "A brief history of downside risk measures," *Journal of Investing*, vol. 8, no. 3, pp. 9–25, 1999.

[10] R. Porter, "Semivariance and stochastic dominance: A comparison," *The American Economic Review*, vol. 64, no. 1, pp. 200–204, Mar. 1974.

[11] D. Cumova and D. Nawrocki, "A symmetric lpm model for heuristic mean-semivariance analysis," *Journal of Economics and Business*, vol. 63, pp. 217–236, 2011.

[12] W. Hogan and J. Warren, "Computation of the efficient boundary in the e-s portfolio selection model," *Journal of Financial and Quantitative Analysis*, vol. 7, no. 4, pp. 1881–1896, Sep. 1972.

[13] H. Markowitz, P. todd, G. Xu, and Y. Yamane, "Computation of mean-semivariance efficient sets by the critical line algorithm," *Annals of Operations Research*, vol. 45, pp. 307–317, 1993.

[14] J. Sun, “On piecewise quadratic newton and trust region problems,” *Mathematical Programming*, vol. 76, pp. 451–467, 1997.

[15] H. Jin, H. Markowitz, and X. Zhou, “A note on semivariance,” *Mathematical Finance*, vol. 16, no. 1, pp. 53–61, Jan. 2006.

[16] H. Jin, J. Yan, and X. Zhou, “Continuous-time mean-risk portfolio selection,” *Annales de l’Institute Henri Poincaré*, vol. 41, pp. 559–580, 2005.

[17] A. Lari-Lavassani and X. Li, “Dynamic mean semi-variance portfolio selection,” in *Lecture Notes in Computer Science*, Computational Science — ICCS, vol. 2657, 2003, pp. 95–104.

[18] W. yan, R. Miao, and S. Li, “Multi-period semi-variance portfolio selection: Model and numerical solution,” *Applied Mathematics and Computation*, vol. 194, pp. 128–134, 2007.

[19] M. Hirschberger, Y. Qi, and R. Steuer, “Large-scale mv efficient frontier computation via a procedure of parametric quadratic programming,” *European Journal of Operational Research*, vol. 204, pp. 581–588, 2010.

[20] A. Niedermayer and D. Niedermayer, “Applying markowitz’s critical line algorithm,” University of Bern, Department of Economics, Tech. Rep., Jan. 2007.

[21] R. Steuer, Y. Qi, and M. Hirschberger, “Portfolio optimization: New capabilities and future methods,” *Zeitschrift für Betriebswirtschaft*, vol. 76, no. 2, pp. 199–219, 2006.

[22] P. Patrinos and H. Sarimveis, “Convex parametric piecewise quadratic optimization: Theory and algorithms,” *Automatica*, vol. 47, pp. 1770–1777, 2011.

[23] V. DeMiguel and F. Nogales, “Portfolio selection with robust estimation,” *Operations Research*, vol. 57, no. 3, pp. 560–577, 2009.

[24] L. Garlappi, R. Uppal, and T. Wang, “Portfolio selection with parameter and model uncertainty: A multi-prior approach,” *The Review of Financial Studies*, vol. 20, no. 1, pp. 41–81, Jan. 2007.

[25] L. Ghaoui, M. Oks, and F. Oustry, “Worst-case value-at-risk and robust portfolio optimization: A conic programming approach,” *Operations Research*, vol. 51, no. 4, pp. 543–556, 2003.

[26] D. Goldfarb and G. Iyengar, “Robust portfolio selection problems,” *Mathematics of Operations Research*, vol. 28, no. 1, pp. 1–38, Feb. 2003.

[27] C. Harvey, J. Liechty, and M. Liechty, "Bayes vs. resampling: A rematch," *Journal of Investment Management*, vol. 6, no. 1, pp. 1–17, 2008.

[28] H. Levy and H. Markowitz, "Approximating expected utility by a function of mean and variance," *The American Economic Review*, vol. 69, no. 3, pp. 308–317, Jun. 1979.

[29] H. Markowitz, "Mean-variance approximations to expected utility," *European Journal of Operational Research*, vol. 234, pp. 346–355, 2014.

[30] W. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, vol. 19, no. 3, pp. 425–442, Sep. 1964.

[31] D. Cumova and D. Nawrocki, "Portfolio optimization in an upside potential and downside risk framework," *Journal of Economic and Business*, vol. 71, pp. 68–89, 2014.

[32] K. Lau and R. Womersley, "Multistage quadratic stochastic programming," *Journal of Computational and Applied Mathematics*, vol. 129, pp. 105–138, 2001.

[33] J. Fliege and R. Werner, "Robust multiobjective optimization & applications in portfolio optimization," *European Journal of Operational Research*, vol. 234, pp. 422–433, 2014.

[34] R. Tütüncü and M. Koenig, "Robust asset allocation," *Annals of Operations Research*, vol. 132, pp. 157–187, 2004.

[35] N. Rujeerapaibon, D. Kuhn, and W. Wiesemann, "Robust growth-optimal portfolios," *Management Science*, 2015.

[36] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Math. Programming Ser. A*, vol. 137, no. 1–2, pp. 167–198, 2013a.

[37] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Naval Research Logistics*, vol. 3, no. 1–2, pp. 95–110, 1956.

[38] J. Quirk and R. Saposnik, "Admissibillity and measurable utility functions," *The Review of Economic Studies*, vol. 29, no. 2, pp. 140–146, Feb. 1962.

[39] J. Estrada, "Mean-semivariance behavior: An alternative behavioral model," Centro Internacional de Investigacion Financiera, Tech. Rep., 2003.

[40] D. Nawrocki, "Optimal algorithms and lower partial moment: Ex post results," *Applied Economics*, vol. 23, pp. 465–470, 1991.

[41] ——, "The characteristics of portfolios selected by n-degree lower partial moment," *International Review of Financial Analysis*, vol. 1, no. 3, pp. 195–209, 1992.

[42] D. Li and W. Ng, "Optimal dynamic portfolio selection: Multiperiod mean-variance formulation," *Mathematical Finance*, vol. 10, no. 3, pp. 387–406, Jul. 2000.

[43] M. Yu, S. Takahashi, H. Inoue, and S. Wang, "Dynamic portfolio optimization with risk control for absolute deviation model," *European Journal of Operational Research*, vol. 201, pp. 349–364, 2010.

[44] D. Li, X. Cui, S. Wang, and S. Zhu, "Better than dynamic mean-variance: Time inconsistency and free cash flow stream," *Mathematical Finance*, vol. 22, no. 2, pp. 346–378, Apr. 2012.

[45] A. A. Najafi and S. Mushakhian, "Multi-stage stochastic mean-semivariance-cvar portfolio optimization under transaction costs," *Applied Mathematics and Computation*, vol. 256, pp. 445–458, 2015.

[46] B. V. Halldórsson and R. H. Tütüncü, "An interior-point method for a class of saddle-point problems," *Journal of Optimization Theory and Applications*, vol. 116, no. 3, pp. 559–590, Mar. 2003.

[47] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Mathematical Programming*, pp. 1–37, 2015.

[48] A. J. Kleywegt and R. Gao, "Distributionally robust stochastic optimization with wasserstein distance," *arXiv:1604.02199*, 2016, visited on 2017-11-02.

[49] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming*. Philadelphia: MPS-SIAM, 2009.

[50] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Early Release Draft*, 2012.

[51] A. Ling, J. Sun, and X. Yang, "Robust tracking error portfolio selection with worst-case downside risk measures," *Journal of Economical Dynamics & Control*, 2014.

[52] J. Neumann and O. Morgenstern, *Theory of games and economic behavior*, Third. Princeton University Press, 1944/1953.

[53] J. ya Gotoh, M. J. Kim, and A. Lim. (Apr. 2015). Robust empirical optmization is almost the same as mean-variance optimization. `https://ssrn.com/abstract=2827400`, visited on 2017-11-02.

[54] A. Shapiro, “Semi–infinite programming: Recent advances,” in. Kluwer Acdemic Publishers, 2001, ch. On duality theory of conic linear problems.

[55] F. Gantmacher, *The theory of matrices*. New York: Chelsea Pub., 1959.

[56] A. Ben-Tal, L. Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009.

[57] Y. liang Yu, Y. Li, D. Schuurmans, and C. Szepesvári, “A general projection property for distribution families,” in *Advances in Neural Information Processing Systems 22*, Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, Eds., `http://papers.nips.cc/paper/3811-a-general-projection-property-for-distribution-families.pdf`, visited on 2017-11-02, Curran Associates, Inc., 2009, pp. 2232–2240.

[58] J. H. B. Kemperman, “The general moment problem, a geometric approach,” *The Annals of Mathematical Statistics*, vol. 39, no. 1, pp. 93–122, 1968.

VITA

Carl Morris was born in Lubbock, TX on November 9, 1986. He received his Bachelor's of Science in Industrial and Systems Engineering from Mississippi State University in May, 2010. He received his Ph.D. in Operations Research from the Georgia Institute of Technology in November, 2017. He began an applied science career in the retail and information technology industries in 2015.